International Journal of Advanced Scientific Technologies ,Engineering and Management Sciences (IJASTEMS-ISSN: 2454-356X) Volume.3,Special Issue.1,April.2017

RED LESION DETECTION USING ADAPTIVE GLOBAL MAXIMUM CLUSTERING ALGORITHM FOR DIABETIC RETINOPATHY SCREENING

[1]Elizabeth J S, [2]Keerthi Krishnan, [3] Dr.Sheeja Agustin

[1]PG Scholar, [2, 3] Associate Professor

Department of Computer Science and Engineering,

Marian Engineering College, Kazhakuttom, Trivandrum

[1]js.elizabeth023@gmail.com

Abstract.- Now a day's computer aided design and diagnosis is very popular. Diabetic retinopathy (DR) is one of the diabetic eye diseases found in the patients who have diabetic in last 20-30 years. DR detection and screening done via Optic Disk(OD) detection, segmentation and classification. The main objective of this work is to effectively segment the DR detected area in the retinal image by using Adaptive Global Maximum Clustering(AGMC). Segmenting an image into set of pixels of similar intensities in AGMC is called regions. First, we develop the adaptive global maximum clustering. In the procedure of AGMC process image histogram and automatically obtain the number of significant local maxima of the histogram. This number indicates the number of different regions in the image. Then derive a simple and fast calculation to segment an image composed of distinct multiple regions. Then, we split an image into multiple regions according to the previous procedure. So effectively found the DR in the retinal image.

Index Terms— Computer Aided Diagnosis(CAD), Diabetic Retinopathy(DR),Adaptive Global Maximum Clustering(AGMC), Image Histogram, Means Clustering, Multiple-Region Segmentation.

I. INTRODUCTION

A computer-aided screening and grading system deals the automatic detection of lesions in the retinal image. Fundus images with red lesions shows microaneurysms (MA) and hemorrhages (HE), and bright lesions shows exudates and cottonwool spots[4].Diabetic retinopathy (DR) is a diabetic eye disease present the patient those who have high sugar in the retinal blood vessels. DR usually affects blood vessels in the light sensitive tissue are called retina. Retina are always located at the back side of the eye .They detect light and convert it to signals and sent through the optic nerve to the brain. Diabetic retinopathy results retina to leak the fluid or haemorrhage (bleed) [13].So it is the most common cause of vision loss among people those who have diabetes and lead to the major cause of vision impairment mainly in the working age of adults.

Diabetic Retinopathy having four stages [1]. They are Mild Non Proliferative Diabetic Retinopathy (Mild NPDR), Moderate Non Proliferative Diabetic Retinopathy (Moderate NPDR), Severe Non Proliferative Diabetic Retinopathy (Severe NPDR), and Proliferative Diabetic Retinopathy (PDR).Mild NPDR is a small areas of balloon like structure swelling in the retina's tiny blood vessels called micro aneurysms [13]. It comes at earliest stage of diabetic retinopathy. These micro aneurysms may leak fluid into the retina. Leakage of fluid may or may not be present in hard/soft exudates. Moderate NPDR second stage of diabetic retinopathy. Here blood vessels that nourish the retina may swell and damage[14]. So they loss their ability to transport blood. Both of these conditions cause large changes to the appearance and shape of the retina. So it may contribute DME. Also cause the presence of cotton wool spots. Severe NPDR is the third stage of diabetic retinopathy. In this case many more blood vessels are blocked, damage blood supply to surface areas of the retina. These surface areas has secrete growth factors that signal the retina to grow new blood vessels. Finally Proliferative Diabetic Retinopathy (PDR) is the advanced stage of diabetic retinopathy. In this stage the growth factors of retina are secreted. It trigger the proliferation of new blood vessel which grow along the inside surface of the retina and to the vitreous gel. So that fluid is fill in the eye[9]. Hence the new blood vessels are fragile which makes more to leak and bleed.

For the fundus image there are three features for pixel classification work [6]:which include the grey level of the

inverted green channel image called green channel extraction. Response of two line detectors to the neighbourhood pixel, so that one perpendicular to another. Then basic line detector has length 15 pixels which rotate at 12 different orientations in between 0 to 360 degrees. So the response of line operator at each pixel along a specific angle is obtained by taking average of the grey level of pixels along the line operator. Then to find the largest response in one of two line features. The average grey level of line with equal length to three orthogonal pixels to the basic line detector and is also used as another line feature. A linear support vector machine (SVM)[19] ,Random Forest(RF)[4] is act as a classifier for segmenting vessels in retinal image.

II. LITERATURE SURVEY

Diabetic Retinopathy Detection (DRD) mainly include image acquisition, image pre-processing, optic disc elimination, feature extraction and classification [4]. Image acquisition usually done by using camera these images are named as fundus images. Image pre processing step is used to the contrast enhancement of the fundus image. Optic disc elimination is the major part of Diabetic Retinopathy Detection(DRD) because proper removal of optic disc make the detection process is more easier. Feature extraction of fundus image based on the shape or texture feature. Finally Classification is done by using any of the classifier such as Bayesian classifier, k- nearest classifier, Support Vector Machine(SVM) classifier. segment blood vessels automatically have concentrated primarily on their local attributes.

Mainly blood vessels may be characterized by the expected color (reddish), shape (curvilinear), gradient (strength of boundary), and contrast (with background)[3].These attributes, include boundaries of the optic nerve and few amount of hemorrhages and lesions, can show the same local attributes as vessels. Segmentation of blood vessels are of three types such as window-based segmentation [16,17,18], classifier-based segmentation [19,20], and tracking-based segmentation [21,22]. Window-based method of segmentation include edge detection, it estimate a match for each pixel for a given model against the pixel's surrounding window. Classifier-based method of segmentation proceed in two steps. First step, a low-level algorithm produces a segmentation of spatially-connected regions. Secondly identify the candidate regions are vessel or not vessel. In [19], region based segmentation, segmented by user-assisted thresholding. Here classification done as a blood vessel or not according to their leakage length to width ratio. The [20], regions segmented by the method in [16] were classified as vessel or not vessel according to their properties, including their response to a classic operator designed to detect roads in aerial imagery [23]. The drawback of the methods is the large-scale properties of vessels can't be applied to the problem unless to finish low-level segmentation .Therefore, these properties can't be used to drive the segmentation, and evaluation purpose. Tracking-based methods use a profile model to segment a vessel in the increasing manner. Here[21], Hough transform is used to locate the papilla in the retinal image. Vessel tracing proceeds iteratively from the papilla, it will halt the response of dimensional (cross-section) matched filter that falls below a given threshold value. In [22], a similar method was implemented to detect vessels in coronary arteriograms, from user-given starting points. One drawback of this approach is to calculate branch points in their proclivity for termination which are not well-modeled in one dimensional filters. Another drawback is that how to locate starting points. Segmentation part of DR part includes thresholding [3], adaptive histogram equalization and clustering. In thresholding method image open with different orientations. The total 12 patches with resolution of 15^0. Apply thresholding as length of structuring element larger than the biggest red lesion present.

Then THRESHOLD $t = i_{seed} - x(i_{seed} - i_{bg})$

Where i_{seed} - Intensity at the starting pixel position.

i_{bg} - Intensity of the same pixel in the background image.

x = value of 0/1.

Adaptive histogram equalization method is a type of histogram equalisation[5]. In histogram method enhance image contrast by adjusting image intensity value. Also computes several histograms each represent distinct section of the image and use them to redistribute the values o f the image. A variant of adaptive histogram equalisation is called "Limited Adaptive Histogram Equalisation". In clustering method includes k means, improved k mean, fuzzy c mean (FCM) and improved fuzzy c mean algorithm (IFCM).

The next major step is to classify the extracted candidate regions. In DRD mainly classification is done by any of the classifier. Bayesian classifier is one of the probabilistic classifier. It is based on the Bayes' theorem [6]. Let $x=(x1,x2,....,x_n)$ be the vector representing some 'n' features such as instance probability for independent variables as $p(C_k |x1,x2,....,x_n)$ for K outcomes or classes using Bayes theorem [6]$p(C_kx)=\{p(C_k)p(x|C_k)\}/p(x)$.

III. PROPOSED METHOD

A new, simple, fast, and unsupervised multiple-region segmentation method. By developing the AGMC method to find the number of different regions automatically. From the AGMC method, we have not only the number of distinct regions but also subintervals of associated with gray levels of distinct regions.

In AGMC We use an image histogram to get the number of different regions as in [23] and [24]. In most images, there are too many local maxima of image histograms. However, we need to find only significant local maxima since those maxima are necessary to discriminate regions. To extract the significant local maxima, first search for an interval, including the global maximum of an original histogram, and then fix the interval.

Figure 1 depicts the proposed method for Red Lesion Detection.

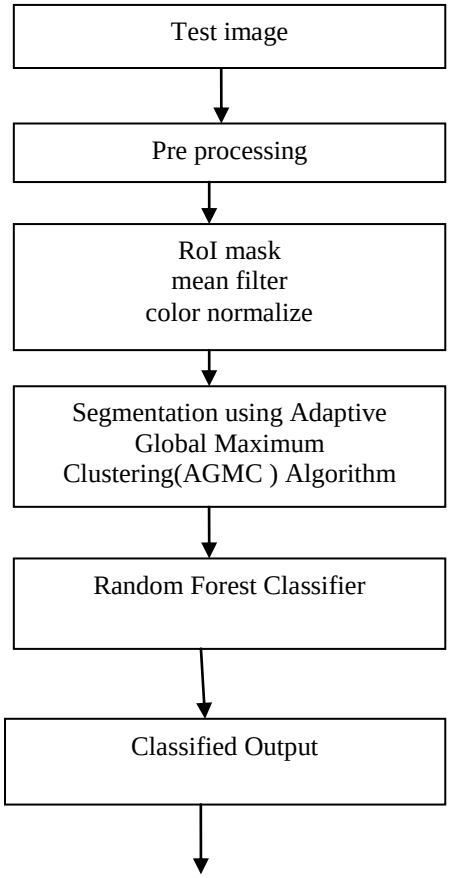

Fig.1 Proposed method

We call such a fixed interval a cluster. Next, we remove the cluster, gained from the previous searching process, from the original histogram to find another new cluster. We then search for the new cluster of the reduced histogram and repeat this process until we have the desired result. Since the histogram changes every iteration, which is precisely the reduced version of the original histogram, the global maximum also adaptively changes every iteration. We therefore call such a maximum an adaptive global maximum that corresponds to one of the significant local maxima of the original histogram. This whole process is a series of clustering . Divid gray level interval into several subintervals so that the original histogram has the adaptive global maximum over each subinterval. Thus, we call this process the AGMC process. The goal of this process is

$$[1,L] = (U_{i=1}{}^{n} I_i)U^R$$

where each subinterval $I_{i=}$ $[a_i, b_i]$ is a cluster containing the i^{th} adaptive global maximum, which is the global maximum of the i^{th} histogram. R is the domain of the $(n+1)^{th}$ histogram or R is empty. Such very small histogram values are usually useless in the detection of regions. To find a cluster that is a subinterval with an adaptive global maximum at each iteration, we fix and repetitively implement the standard -means clustering under the rules described below. The standard -means clustering method is a process to solve the following minimization problem:

$$\arg\min_{I1,I2,...Ik} \Sigma^{k}_{j=1} \Sigma_{x\varepsilon Ij} \mid x-d_j|^2$$

Where $|x-d_j|^2$ is a distance between a data point 'x' and the j^{th} cluster center d_j . The aim of the -means clustering is to congregate a set of distributed data into clusters. For an image histogram, the k-means clustering problem is described as

$$\arg\min_{I1,I2,...Ik} \Sigma^{k}_{j=1} \Sigma_{l\varepsilon Ij}\, h(l)(l-d_j)^2 .$$

where h is an image histogram, and I_j is subinterval of gray levels such that $U_{j=1}{}^{k} I_j = [1,L]$. Each center d_j is computed by

$$d_j = \frac{\Sigma_{l\varepsilon Ij}\, l*h(l)}{\Sigma_{l\varepsilon I}\, h(l)} \quad \text{for } j=i...k.$$

The k value ,the total number of clusters, has to be given a priori, and the result depends on the initial centers.

We resolve those problems by fixing k=2,setting the initial two centers as starting index of I and ending index of I , and repeating the k-means clustering in the following way. Here I , is the domain of histogram. Note that, in our method, the two means clustering is applied to the reduced histograms and the original histogram. Thus, the starting and the ending indexes are changeable, not fixed as 1 and 256. Once we implement the two-means clustering, the histogram is divided into two clusters I_1 and I_2 since k=2 is chosen. Then, we have two maxima, i.e., one is obtained in I_1 and the other is obtained in I_2 .It is clear that if the maximum value of the histogram in one I_1 cluster is larger than the maximum value of the histogram in the other cluster I_2 , then cluster I_1 contains the global maximum of the histogram. The three rules for AGMC as Rule 1 to choose a cluster

Rule 1

$$I^* = \begin{cases} I_1 \text{ ; if } \max_{l\,\varepsilon\, I1} h(l) > \max_{l\,\varepsilon\, I2} h(l) \\ I_2 \text{ ; otherwise.} \end{cases}$$

Then, we again perform the two-means clustering over the chosen cluster and repeat this process until Rule 2 holds.

Rule 2

$$\left| \arg\max_{l\,\varepsilon\, I_1} h(l) - \arg\max_{l\,\varepsilon\, I_2} h(l) \right| < \alpha.$$

Where α least difference in intensities of distinct regions. If α is very small then separate regions with similar intensities. So get the larger number of regions. Rule2 is satisfied then get a cluster, which is a subinterval including one of the significant local maxima of the original histogram. To stop the AGMC procedure if Rule 3 holds.

Rule 3 : Max $h^i < \omega$ mean(h^0)

Where h^i – i^{th} histogram which is reduced after $(i-1)^{th}$ histogram. h^0 – original histogram and ω - constant.

Rule 3 signifies that h^i is too small compared with original histogram. Then h^i is useless in the detection of regions. Also h^i prevents from finding small local maxima points. Larger ω value result in find the larger local maxima of the histogram.

IV. PROPOSED ALGORITHM

The algorithm of the AGMC procedure is as follows:

Require : let h = $\{h(l)\}_{l=1}^{L}$ be an original image histogram .

L=256, h^0 = h , k = 2, i = 0 and I = [1,L].

repeat :

i <− i+1 and I^{old} <− I.

repeat :

i. Implement two means clustering by minimizing $h(I)=\{h^0(l) : l \varepsilon I\}$ where h(I) – histogram whose domain is restricted to cluster I.

ii. Choose cluster I^* between two clusters.

iii. I <− I^*

until Rule 2 holds.

set a cluster D_i <− I^*

I <− I^{old} / D_i = I^{old} / I^* .

until Rule 3 holds for h(I) or h (I) = 0.

Finally To distinguish between lesions and non-lesions, we use a Random Forest(RF) classifier [25]. This powerful approach has been widely used in computer vision over the last few years, due to its numerous advantages. It is convenient for non-linear classification with high-dimensional and noisy data. It is robust against outliers and over-fitting. Moreover, it incorporates an implicit features selection step.

A RF is a combination T of decision trees trained independently using T bootstrap samples drawn with replacement from the training set. Each node is split using the best of a randomly selected subset of features chosen, according to the decrease in the Gini index [25]. The RF returns, for each candidate, a probability of being a lesion $P(S_j)$, equal to the proportion of trees returning a positive response.

Using MATLAB interface [26] to the RF implementation in [27]. The classifier relies on two user-defined parameters T, and m , but usually the outcome is not very sensitive to their values [27]. We thus follow the recommendations provided in[25]and set T=200 trees $m=\sqrt{M}$ and ,where M =4+6x(K+4) is the total number of features.

V.SIMULATION RESULTS

To evaluate the performance of proposed method on a large variety of images use six independent databases. This method was evaluated at two levels:

- When a delineation of the lesions was provided, method was evaluated on a per-lesion basis, meaning we analyzed its performance in detecting every single lesion.
- When only a diagnosis was provided for each image, our method was evaluated on a per-image basis.

VI.CONCLUSION

In this paper proposes a novel approach to find diabetic retinopathy detection by using using adaptive global maximum clustering algorithm Retinal images are acquired and pre processed via ROI mask, mean filter and color normalize method. The new unsupervised multiple-region segmentation method by using the AGMC procedure. Proposed method is composed of two procedures, namely, the segmentation and the AGMC. In the AGMC procedure, we automatically obtain the number of distinct regions and clusters, which are subintervals of gray levels containing adaptive global maxima. We fix and repeat the k-means clustering under a few rules in this procedure. In the segmentation procedure, we decompose the original image into the multiple regions according to the number and clusters obtained from the AGMC procedure. Finally a Random Forest (RF) classifier is used to classify the lesion and non lesion in a retinal image. The method employed in this study will help in improving diagnostic accuracy as well as in improving the workflow efficiency of the DR screening at peripheral healthcare centres and diabetic clinics.Very promising outcome for diabetic retinopathy detection was provided.

ACKNOWLEDGEMENT

This work was supported by Libray of Marian Engineering College, Kazhakuttom, Trivandrum.

REFERENCES

[1] Meindert Niemeijer, Bram van Ginneken, and Michael D. Abràmoff “Automatic Detection Of Red Lesions In Digital Color Fundus Photographs” IEEE Transactions on medical imaging,Vol:24,pp 584-592,2005.

[2] Reza Kharghanian and Alireza Ahmadyfard “Retinal Blood Vessel Segmentation Using Gabor Wavelet And Line Operator” International Journal of Machine Learning and Computing, Vol. 2, PP 593-597,2012.

[3] Adam Hoover, Valentina Kouznetsova, and Michael Goldbau “Locating Blood Vessels In Retinal Images By Piece-Wise Threshold Probing Of A Matched Filter Response” Department of Ophthalmology, University of California and San Diego La Jolla,2000.

[4] Lama Seoud , Thomas Hurtut, and Jihed Chelbi “Red Lesion Detection Using Dynamic Shape Features for Diabetic Retinopathy Screening”, IEEE Transactions on Medical Imaging ,vol no:35,pp 1116-1126,2016.

[5] Saumitra Kumar Kuri “Automatic Diabetic Retinopathy Detection Using Gabor Filter With Local Entropy Thresholding” IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS),2012.

[6] Reza Kharghanian and Alireza Ahmadyfard” Retinal Blood Vessel Segmentation Using Gabor Wavelet and Line Operator ” International Journal of Machine Learning and Computing, Vol. 2, No. 5, October 2012.

[7] Xiaoyi Jiang, and Daniel Mojon” Adaptive Local Thresholding by Verification-Based Multi threshold Probing with Application to Vessel Detection in Retinal

Images" IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 25, No. 1, January 2003.

[8] Carla Agurto*," Multiscale Am-Fm Methods For Diabetic Retinopathy Lesion Detection" IEEE Transactions On Medical Imaging, Vol. 29, No. 2, February 2010.

[9]] N. Cheung, P. Mitchell, and T. Y. Wong, "Diabetic Retinopathy," Lancet, vol. 376, no. 9735, pp. 124–36, 2010.

[10] J.Dingand, T.Y.Wong," Current Epidemiology Of Diabetic Retinopathy And Diabetic Macular Edema," Curr. Diabetes Rep., vol. 12, no. 4, pp. 346–54, 2012.

[11] C.Sinthanayothinetal.," Automated Detection Of Diabetic Retinopathy On Digital Fundus Images," Diabetic Med. A J. Brit. Diabetic Association , vol. 19, no. 2, pp. 105–12, 2002.

[12] S. Ravishankar, A.Jain, and A.Mittal, "Automated Feature Extraction For Early Detection Of Diabetic Retinopathy In Fundus Images,"inProc. IEEE Conference Computation Vis .Pattern Recognition ,2009,pp.210–7.

[13] L.Seoudetal, "Automatic Detection Of Microaneurysms And Haemorrhages In Fundus Images Using Dynamic Shape Features," in Proc. IEEE 11th Int. Symp .Bio Medical Imaging Beijing,pp.101–104,2014.

[14] X. Zhang et al., "Exudate Detection In Color Retinal Images For Mass Screening Of Diabetic Retinopathy," Medical Image Analysis, vol. 18, no. 7, pp. 1026–1043, 2014.

[15] S. Chaudhuri, N. Katz, M. Nelson and M. Goldbaum, ``Detection Of Blood Vessels In Retinal Images Using Two-Dimensional Matched Filters", in IEEE Transaction on Medical Imaging, vol. 8, no. 3, Sep. 1989, pp. 263-269.

[16] T. Pappas and J. Lim, ``A New Method for Estimation of Coronary Artery Dimensions in Angiograms", in IEEE Transaction on Acoustics, Speech, and Signal Processing, vol. 36, no. 9, pp. 1501-1513, Sep. 1988.

[17] R. Nekovei and Y. Sun, ``Back-Propagation Network and its Configuration for Blood Vessel Detection in Angiograms", in IEEE Transaction on Neural Networks, vol. 6, no. 1,pp. 64-72, Jan. 1995.

[18] S. Tamura, A. Okada and M. Hoshi, ``Semiautomatic Leakage Analyzing System for Time Series Fluorescein Ocular Fundus Angiography", in Pattern Recognition, vol. 16, no. 2,pp. 149-162, 1983.

[19] J.-P. Antoine, P. Vandergheynst and S. Twareque Ali,"Two-Dimensional Wavelets and their Relatives" Cambridge University Press, 2004.

[20] S. Tamura, Y. Okamoto and K. Yanashima, ``Zero-Crossing Interval Correction in Tracing Eye-Fundus Blood Vessels", in Pattern Recognition, vol. 21, no. 3, 1988, pp. 227-233.

[21] Y. Sun, ``Automated Identification of Vessel Contours in Coronary Arteriograms by an Adaptive Tracking Algorithm", in IEEE Transaction on Medical Imaging, vol. 8, no. 1, pp. 78-88, March 1989.

[22] M. Fischer, J. Tenenbaum and H. Wolf, ``Detection Of Roads And Linear Structures In Low Resolution Aerial Imagery Using A Multisource Knowledge Integration Technique", in Computer Graphics and Image Processing, vol. 15, no. 3, pp. 201-223, 1981.

[23] M. Luessi, M. Eichmann, G. Schuster, and A. Katsaggelos, "Framework For Efficient Optimal Multilevel Image Thresholding," J. Electronics Imaging, vol. 18, p. 013004, Feb. 2009.

[24] M.Eichmannand, and M.Lussi,"Efficient multilevel image thresholding," M.S. thesis, Science in Engineering, Hochschule fur Technik Rapperswil, Rapperswil, Switzerland, Dec. 2005.

[25] Textbook of Rafael C. Gonzalez and Richard E. Woods "Digital Image Processing Third Edition"

[26] L.Breiman,"Randomforests,"Mach.Learn.,vol.45,pp.5–32,2001.

[27] Random Forest Implementation for Matlab [Online].Available :https:/ /code.google.com/p/random forest- matlab/

[28] A. Liaw and M. Wiener, "Classification And Regression By Random Forest," R News, vol. 2, no. 3, pp. 18–22, 2002.

[29] S.N. Sivanandam, S.N. Deepa, "Principles of Soft Computing", 2/e, John Wiley India, 2012

[30] https://en.wikipedia.org/diabetic retinopathy.