Connected Component Clustering Based Hemorrhage Detection in Color Fundus Images

Arati Manjaramkar[1*]   **Manesh Kokare[2]**

[1]*Department of Information Technology,*
Shri Guru Gobind Singhji Institute of Engineering & Technology, Nanded, Maharashtra, India
[2]*Department of Electronics and Telecommunication,*
Shri Guru Gobind Singhji Institute of Engineering & Technology, Nanded, Maharashtra, India
* Corresponding author's Email: akmanjaramkar@sggs.ac.in

Abstract: Damage of retina due to diabetes is termed diabetic retinopathy. Hemorrhages and Microaneurysms are the first clinically visible symptoms of diabetic retinopathy. Detecting and treating diabetic retinopathy early can prevent vision loss. Accurate segmentation of retinal hemorrhage in color fundus image (CFI) has become a challenging task today; as retinal hemorrhages have varied size, shape and texture. We propose a connected component clustering method based on maximally stable extremal regions (MSER) for detecting many occurrences of hemorrhages with different shape and size in a fundus image. Proposed method has three main steps: firstly hemorrhage candidate generation, second is feature extraction and finally third step is hemorrhage detection. We have is evaluated our method on the DIARETDB1 and MESSIDOR dataset and experimental results show that the proposed system outperforms other state-of-the-art methods in detecting large and vessel connected hemorrhages. The proposed method achieves image level sensitivity, specificity of 96.45, 97.64 and lesion level sensitivity, specificity of 94.89, 98.9 respectively.

Keywords: Diabetic retinopathy Hemorrhages, Connected component clustering, MSER, Color fundus images.

1. Introduction

The last decades have witnessed a global hike in diabetes. Eye disorders in which bleeding occur in the light sensitive tissue of the retina is termed as retinal hemorrhage. One of the commonest causes for hemorrhage is diabetic retinopathy (DR), in which easily damageable weak, fragile vessels are formed, and ruptured due to insufficient oxygen supply. DR keeps on increasing with time and cause complete vision loss if left untreated. World health organization survey statistics reveal exponential rise in diabetes [1].

In the early stage of DR, the patient is usually not aware of its existence; however, in progressive stages, patients might experience blurred or distorted vision, floaters, etc. DR can be of broadly categorized into Non-proliferative DR and Proliferative DR.
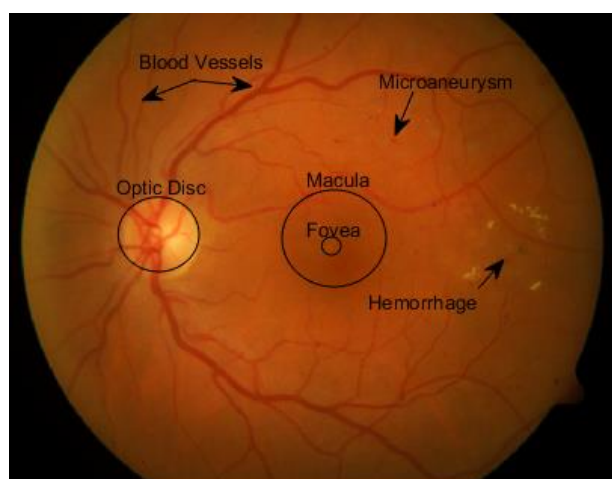

Figure. 1 Fundus image with retinal landmarks

Symptoms of NPDR include red lesions (RL) like Microaneurysms (MA): The earliest clinical sign of DR; appear like small, red dots in retina's

superficial layer and hemorrhages (HM). Dot and blot HM: look alike MAs when small; caused by rupturing of MAs in the deep layers of the retina as seen in Fig.1. Dot and blot HMs are usually found in the inmost nuclear or superficial plexiform layers. Their organization is because of intraregional compression, limiting the HMs in the certain position. Intraretinal HMs need more time to settle than superficial HMs as they are deeper inside than flame-shaped HMs.

Flame-shaped HMs arises in external nerve fiber layer (NFL). These HMs are placed inside the NFL. The property of being flame shape is due to the nerve cell of the ganglion cells passing blood in it, indicating a form of NFL. Flame-shaped HMs is detected at back of pole and has a tendency of settling in short duration of about six weeks. Subhyaloid and preretinal HMs are often so difficult to distinguish these terms are used interchangeably. They are located on the surface of the retina. A subhyaloid HM is located between the posterior vitreous base and the internal limiting membrane (ILM). The preretinal hemorrhage is found at the back of ILM and the front of the NFL. These HMs appear as D or boat-shaped.

One of the crucial areas of research in CFI analysis on which researcher across globe are working actively is detection of irregular shaped HMs [2]. Red lesions are the first clinically observable symptoms of NPDR so their early detection plays vital role in assuring successful diagnosis and effective treatment. The development of computer aided diagnosis (CAD) system for DR detection is of paramount significance for effective and timely treatment. In the past few years, so many algorithms are presented for detecting HMs. HMs have irregular shape and varying sizes; a few are also connected with retinal vasculature, its challenging task to detect all of them by a CAD system.

Several methods for red lesion detection have been reported in the literature. Red lesion detection methods are often grouped into three main types: Pixel-Based, Lesion-based and Image-based detection. Pixel-based method only locates the HM in retinal images. Lesion-based method identifies as well as counts them using morphology [3-5]. Image-based method only tells whether a retinal fundus image contains HM or not [6].

Niemeijer et al. [7] presented a hybrid red lesion detection system based on pixel classification which separated foreground objects like a blood vessel (BV), red lesions from the background image. Later connected vasculature was removed, leaving out candidate red lesions. Extracted features were classified into true candidates by K-nearest Neighbor. Sensitivity, specificity of 100%, 87%, respectively is reported by this method which is found close to that of a human expert.

Bae et al. [8] proposed hybrid method for HM detection. HM candidates were extracted by template matching with cross correlation. Tang et al. [9] presented splat feature classifier for HM detection in fundus images. A supervised method was used for partitioning image into non-overlapping segments called splats, comprising of similar color pixels. Features were extracted from these splats which describes its characteristics. Optimal feature subset was selected by wrapper approach. Messidor dataset was used for evaluation.

Kande et al. [10] presented the pixel classification approach for red lesion detection. The intensity details of red and green components are used along with matched filter and entropy thresholding. Both blood vasculature and lesion candidates are extracted at once. Later vasculature was segmented out from lesion candidates.

Zhang and Chutatape [11] used two dimensional principal component analyses for extracting image patch features in order to locate HMs. But all these methods did not consider irregular shaped HMs.

Fleming et al. [12] applied maximum of multiple linear top-hats (MMLTH) to an inverted image for blot HM candidate's extraction. MMLTH detected objects like HM at different scales. BV candidates were segmented using thresholding and skeletonizing method. SVM classifier was used to classify blot HM. Sensitivity and specificity of 98.60%, 95.50% was achieved by this method.

Marino et al. [13] presented a different resolution correlation filter set for extracting candidate red lesions. Then region growing was used for eliminating false positives.

Kose et al. [14] developed a novel method to detect DR called inverse segmentation. It segments the healthy and unhealthy region from the retinal image. Healthy region's texture does not vary much as compared to the unhealthy regions, using this fact healthy regions can be extracted accurately. Segmentation is done based on intensity. Low intensities and high intensities are segmented by considering background image as dynamic threshold. Hatanaka et al. [15] proposed a method for HM detection in fundus image. Emphasis was on new preprocessing and FP elimination. The histogram of each RGB plane image were extended. HM candidates were detected using density analysis. And the FPs was eliminated by the rule-based method and three Mahalanobis distance classifiers.
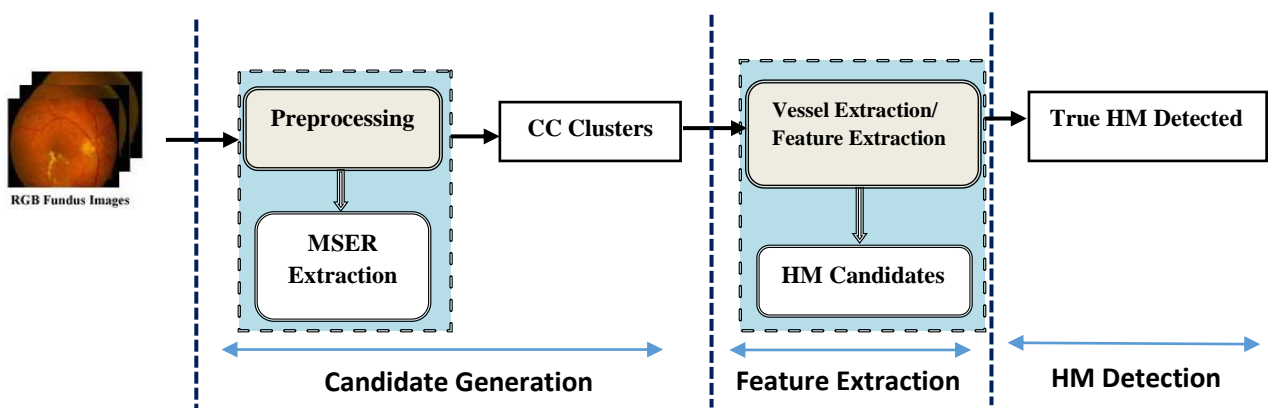
Figure. 2 Flow of proposed system

125 images were examined and sensitivity and specificity of 80% and 88% were reported. Limitation of the proposed system is the ability to detect only small HMs.

Srivastava et al. [16] proposed novel filters for discriminating blood vessels and red lesions. For dealing with varied size of lesions the proposed novel filters were applied on patches formed by using grid instead of complete image. Use of Multiple Kernel Learning reported area under receiver operating characteristic curve of 0.97 and 0.92 for MA and HM respectively.

Seoud et al. [17] proposed red lesion detection system using dynamic shape features (DSF) without prior vessel segmentation. These DSF state evolved shape while flooding image, which allows discrimination between lesion and blood vessel segments. Six datasets were used for evaluation: per-lesion level and per-image level. Performance was found comparable with human experts.

In all these methods, vessel-connected HMs and flame shaped HMs detection is still missing. As HMs is of blood, their appearance is similar to BVs due to which it's difficult to distinguish them from retinal BVs with pixel features. In this paper, we propose an effective method to detect HMs which addresses these problems. We have tested our method on standard retinal datasets with diverse characteristics.

We have been attempting to develop an automated system. We reported methods of detecting MA to help in the diagnosis of diabetic retinopathy [3-5]. In this study, we aim to develop a method which segments hemorrhages from fundus image irrespective of its shape, size and appearance, which includes large, isolated and vessel connected HMs as well.

The remainder of the paper is organized as follows. Section 2 elaborates HM segmentation.

Section 3 presents feature extraction and classification. Section 4 is about HM segmentation. Results and discussions of the proposed system are given in Section 5. We end with conclusion in Section 6.

2. HM segmentation scheme

The proposed system uses some image processing and computer vision techniques for the HM detection purpose. The flow of proposed system is shown in Fig. 2. It has three main steps: HM candidate generation, feature extraction and finally HM detection. Each block of the flowchart is explained briefly in the following subsections.

2.1 Preprocessing

CFI often has poor quality due to patient movement, light-reflections and variation in illumination. Pre-processing of CFI is thus required for shade correction and noise removal before subsequent image analysis. The green plane (I_g) of the original RGB input image is selected for further processing; since the green plane of retinal image clearly exhibits HMs and BVs well compared to other planes. Noise in I_g image is removed by applying a median filter of M × M kernel (M= 56) size due to its properties of degrading the edges minimally while removing noise effectively. The median filter is given by:

$$y[p,q] = median\{ x[i,j], (i,j) \in w \} \qquad (1)$$

Here w- represents a neighbourhood, centred at [p, q] in the image. The median filtered image is shade corrected by equalizing histograms. Taking difference of shade corrected and median filtered image results into new image I_{sc} shown in Fig. 3b. Since proposed method is based on connected

components(CC), I_{sc} image needs to be binarized for extracting CC.

The multilevel Otsu thresholding method is used for binarizing I_{sc}. Assuming that image is a representation in gray levels *(0, 1, ..., L-1)*. Pixel count at level *i* is given as f_i; so all pixels count N= $f_0 + f_{1 +......+} f_{L-1}$. For segmenting gray image I_{sc} into *K* clusters *(C₀+ C₁... C_{K-1})*, *K-1* thresholds needs to be selected[18]. Optimal thresholds *(t₀*, t₁*,..., t_{K-2}*)* are decided by maximizing between-class variance as:

$$\{t_0^*, t_1^*,..., t_{K-2}^*\} = arg\ _{0 \le t_0,\ t_1,...,t_{K-2} < L-1} max$$
$$\{\sigma^2_B (t_0,\ t_1,,,, t_{K-2})\} \quad (2)$$

Where σ^2_B is between class-variance which we get by subtracting the within-class variance from the total variance: $\sigma^2_B(T) = \sigma^2 - \sigma^2 W (T)$, where σ^2 is combined variance $\sigma^2 W (T)$ is within class variance defined as the weighted sum of the variances of each cluster.

$$\sigma^2_W(T) = n_B(t)\sigma^2_B(T) + n_0(T)\sigma^2_0(T) \quad (3)$$

Here $\sigma^2_B(T)$ is variance of pixels below threshold, σ^2_0 is variance of pixels above threshold, $n_B(t) = \sum_{i=0}^{T-1} p(i)$; $n_0(t) = \sum_{i=T}^{N-1} p(i)$ where [0, N-1] is the range of intensity levels.

2.1 MSER connected component clustering

After image binarization, maximally stable extremal regions (MSER) algorithm is used for extracting co-variant regions. Regions are also known as connected components, objects, or blobs. MSER is an intensity based detector that identifies regions satisfying certain uniqueness and stability criteria. The set of all extremal regions in the sequence is the set of all CC are partitioned into clusters[19].

Let $Q1...,Q_{i-1},\ Q_i...$ be nested extremal regions, $Q_1 \subset Q_{i+1}$. As seen in Fig. 5(a), MSERs are attached elliptical frames by fitting ellipses extremal regions.

The color codes in MSER correspond to threshold *t* and frame *t*. Extremal property of MSER indicates pixels inside extremal regions, *Q* are either high intensity that is bright or low intensity that is dark intensity than all outside it and is given as:
$Q \subset D\ such\ that\ for\ all$

$$p \in Q, q \in \partial Q: I(p) > I(q)\ or\ I(p) < I(q) \quad (4)$$

For candidate normalization, we have computed convex hull to find the largest area bounded between them and binary region. Normalization is to portray a *N: $\mathbb{R}^2 \to \mathbb{R}^2$* getting region *Q* and reciprocate another region *N (Q)* such that *N (TQ) = N (Q)* for any affine transformation T. The change of area, normalized by the area of the CC, is used as the stability criterion (*φ*) given as:

$$\varphi(R_i) = \frac{\frac{d}{dt}A(Q)}{A(Q)} \quad (5)$$

Such a function prefers shapes with irregular boundaries, while being affine-invariant. Affine transformation from ellipses to circular regions plus intensity normalization.

3. Feature extraction

To segment out HMs we need to extract blood vasculature which is done by region filtering. MSER region descriptors are used as features for segmenting BV. After sorting extremal regions top six regions were selected by experimentation. Fig. 3(c) shows results of BV segmentation. After eliminating BV from CC, leftovers are considered as HM candidates. But not all leftover components are true HMs.

Based on the characteristics extracted from this CC, rules were applied for removal of small noisy components and left out vessel components, which usually have long shapes. Many features can be extracted from CC. However, if a selected feature is not good, it does not benefit classification. As shape and context are very important features with which retinal we extract features from size, shape information, and position of CC.

We have selected set of following features which are listed here:

Feature	Description
1	Area: $Area(\theta k)$ is number of pixels included in MSER region. Area of HM is not important for HM detection, but it depicts disease severity.
2	Min_Int: The minimum gray value of the MSER.
3	Max_Int:The maximum gray value of the MSER.
4	C_M: Center of mass.
5	Stability: Evaluating criterion for regions relative change in size.
6	Bounding box: smallest surrounding rectangle position.

7	Covariance: statistical distribution of shape.
8	Texture: Information of underlying image.
9	Shape: defined by included pixels.
10	Perimeter: total number of objects pixels having one or more background pixels. $P8 = \{\{r,c\} \in R \mid N4(r,c) - R \neq \emptyset$
11	Elongated components, $I_{elong} = \frac{Area(\theta k}{Len(\theta k)}$
12	Compactness, $I_{comp} = \frac{-2.\sqrt{\pi.Area(\theta k)}}{Peri(\theta k)}$
13	Aspect ratio: ratio of major - minor axis length, $\frac{l}{h}$.
14-18	Response of the five Difference of Gaussian (DoG) filter. Take difference of one version of the image from another version. Convolution of six different Gaussian kernels of standard deviation *(σ) 0.5, 1, 2, 4, 8.*
19	Standard Deviation $SD = \sum_{i,j} Q_{i,j} (i - \mu i)$

Feature selection if done well can reduce the number of false positives thereby reducing the computational cost. Fig. 4 shows examples of some feature extracted from the image and the value of feature extracted from candidate MA pixel which are used as an input to the classifier.

4. Author name(s) and affiliation(s)

HMs are either isolated or they are connected to vessel. The vessel connected HMs will have totally different size and shape compared to isolated ones making it more crucial for extraction. On the basis of observation so far BV normally has certain ranges of diameter.

At places where HM is connected with BV it becomes larger than normal diameter range. So an iterative CC extraction step is applied to deal with this problem of HM segmentation. The proposed method extracts vessel connected HM as well (see Fig. 3(d) encircled in red).

A structure based compactness test of CC is performed for HM candidate selection.

5. Results and discussion

Retinal images from publicly available standard datasets were used for evaluation of the proposed method. DIARETDB1 dataset is provided by Finnish research group with an intention to provide common public database for researchers in this area.

It has 89 images of size 1500 x 1152 and 50^0 field of view. Out of 89, 84 are abnormal and 5 images are normal. In DIARETDB1 lesions are manually highlighted by four experts forming ground truth [20].

MESSIDOR database is available publicly since 2008. It contains 1200 images captured by Topcon TRC NW6 with field of view 45^0. Every image is provided with DR grade and edema risk. In our study, we have used DR grades [21]. A set of 24 DIARETDB1 and 10 MESSIDOR images were used for training. Set of 65 DIARETDB1 and some randomly selected MESSIDOR images were used for testing the classifier. Features are extracted from the MSER components (Fig. 5 (a)).

Earlier state-of-the art methods [14, 15, 23] were having limitation in extracting large and vessel connected HMs it is overcome in proposed method based on connected component clustering. Proposed method results show that it is able to extract out vessel-connected, flame-shaped, irregular-shaped and isolated HMs from the retinal fundus image (as seen in Figs. 3 and 6).

We have used a Random Forest (RF) classifier to classify true HM candidates. RF is more widely used in computer vision in past few years, because of its convenience for non-linear classification. It has been found robust for outliers. RF is a combination of decision trees trained independently from the training set [22].

Performance measures such as sensitivity and specificity were used for system evaluation. Sensitivity (SE) is measure of true lesions identified as lesions and is given as:

$$SE = \frac{Number\ of\ TP}{Number\ of\ TP + Number\ of\ FN} \qquad (6)$$

Specificity (SP) is a measure of false identified as false.

$$SP = \frac{Number\ of\ TN}{Number\ of\ TN + Number\ of\ FP} \qquad (7)$$

Where TP, TN, FP, FN are true positive/negative and false positive/negative.

ROC plot of proposed method on DIARETDB1 and MESSIDOR dataset is shown in Fig.7.

Our proposed system achieves image level sensitivity of 96.45% and specificity of 97.64 respectively. We have also compared our proposed system with earlier existing state-of-the-art methods (Table 1). Proposed method outperforms existing methods.

<div align="center">(a)   (b)   (c)   (d)   (e)</div>
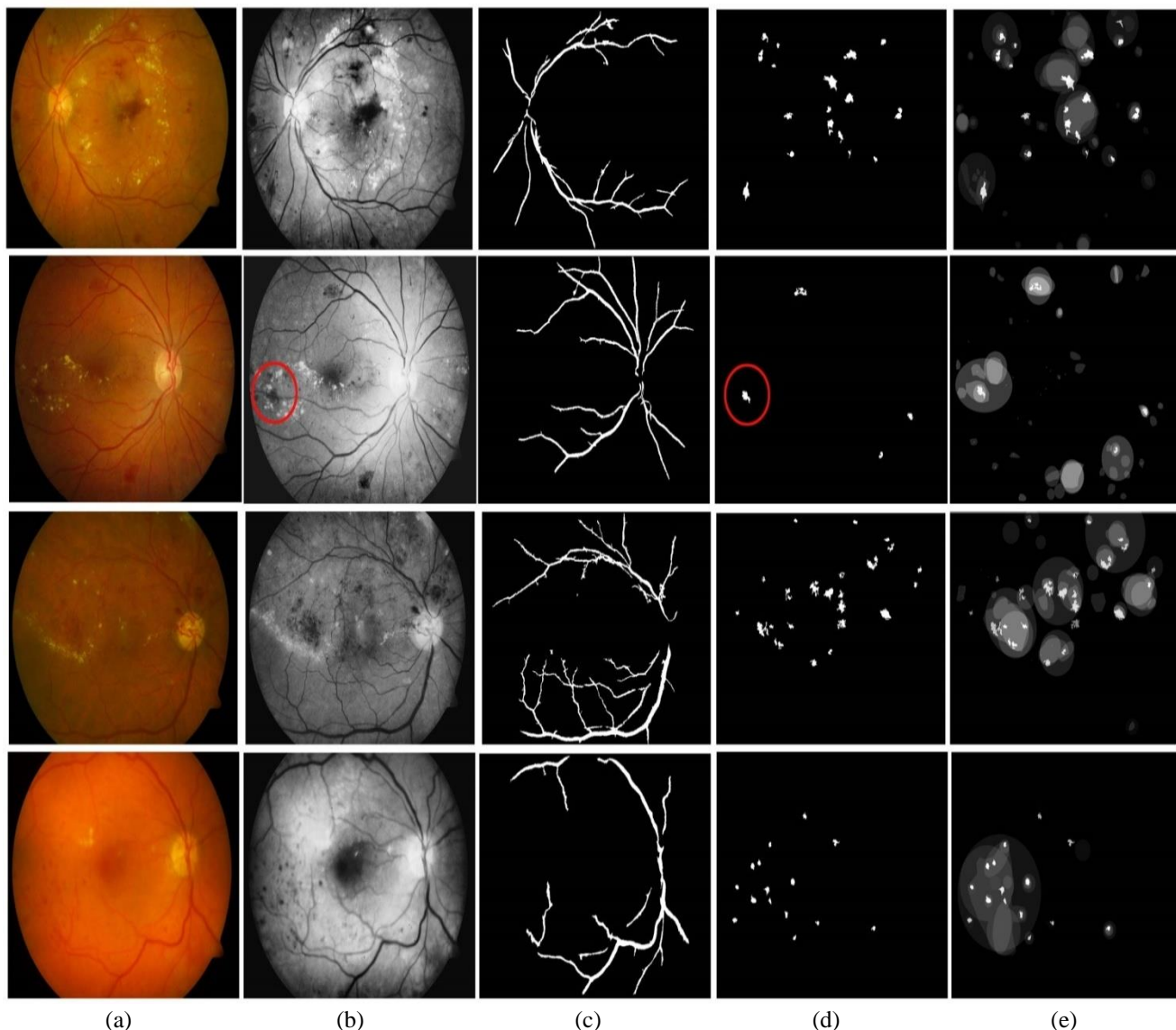
Figure. 3 Steps of the proposed system on four DIARETDB1 dataset images: (a) RGB fundus images, (b) Preprocessed image, (c) Blood vessel segmentation results, (d) Hemorrhages detected by the proposed system, and (e) Detected hemorrhages superimposed on ground truth to depict confidence intervals

<div align="center">Table 1: Comparison of proposed method with existing methods</div>

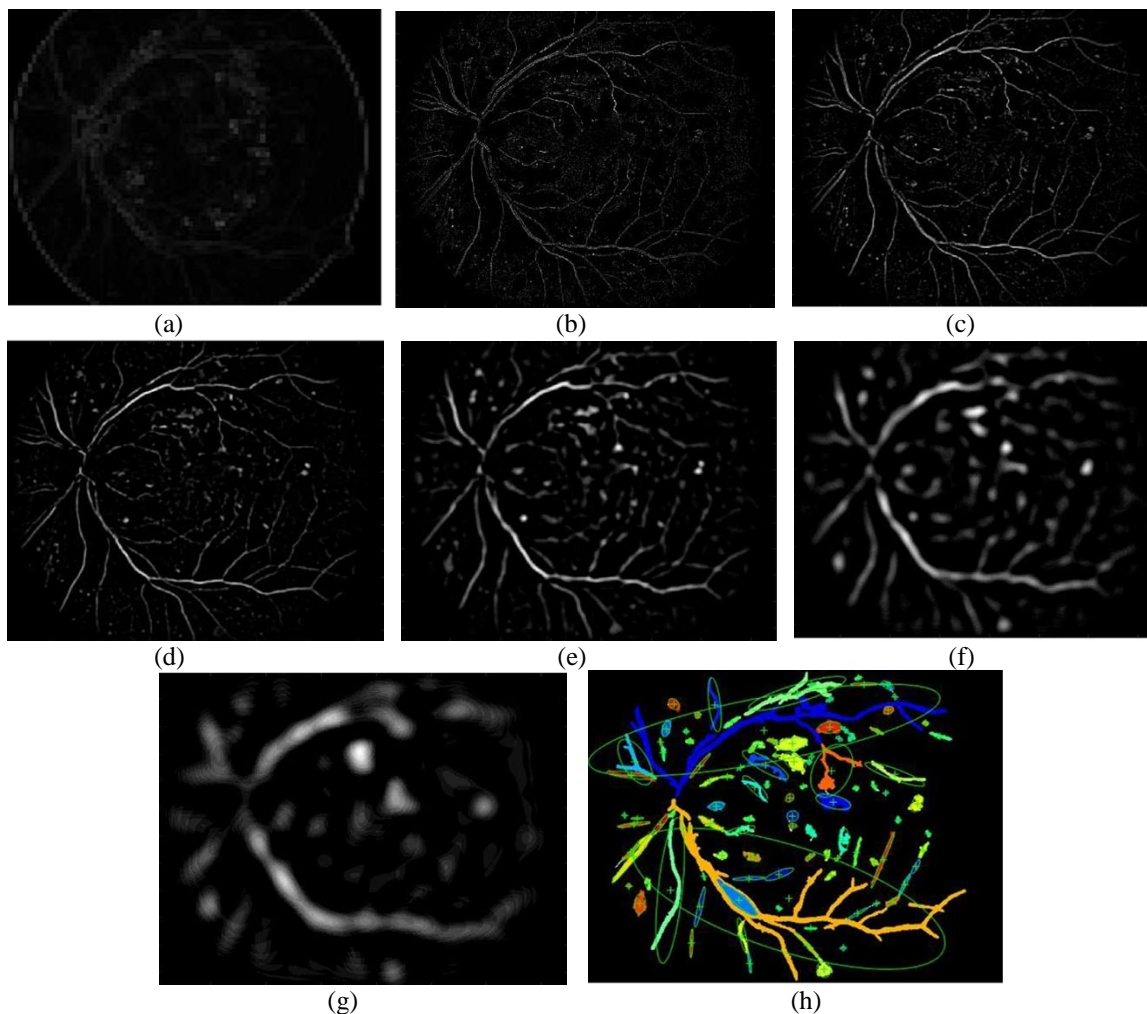| Method	Dataset Used	Image level		Lesion level	
		SE	SP	SE	SP
Hatanaka et al.[15]	Local dataset	80	80	--	--
Junior et al. [23]	DIARETDB1	87.69	92.44	--	--
Tang et al. [9]	Messidor	93	66	--	--
Proposed method	DIARETDB1, Messidor	96.45	97.64	94.89	98.9
* SE-Sensitivity, SP- Specificity					

Figure. 4 Examples of some feature extracted: (a) Standard deviation of green band image, (b) DoG1, (c) DoG2, (d) DoG3, (e) DoG4, (f) DoG5, (g) DoG6, and (h) Color coded MSER
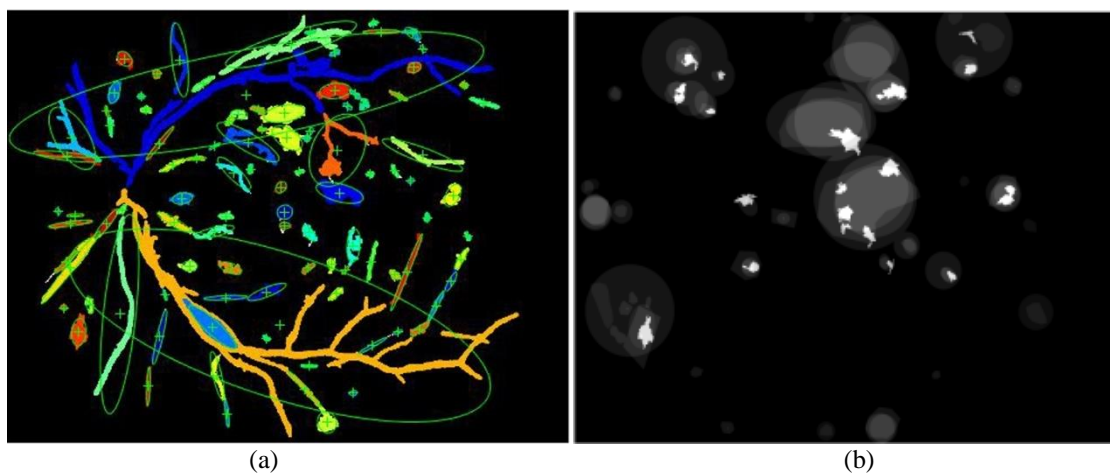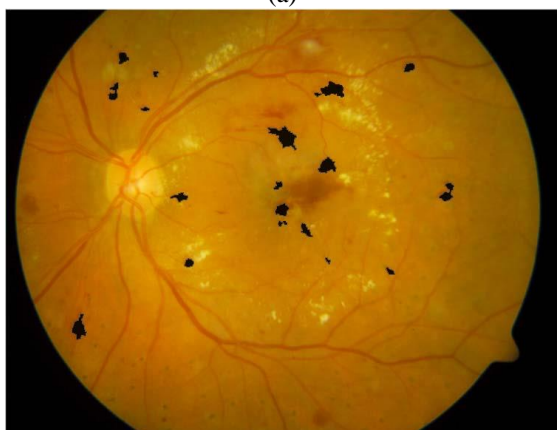

Figure. 5 Illustration of the connected components: (a) Color coded MSER components for a sample image from DIARETDB1 dataset and (b) HMs extracted by the proposed method shown in white color is superimposed on ground truth image for showing confidence interval

(a)

(b)

Figure. 6 Proposed method HM detection results: (a). Original color fundus image and (b). True HMs detected by proposed method
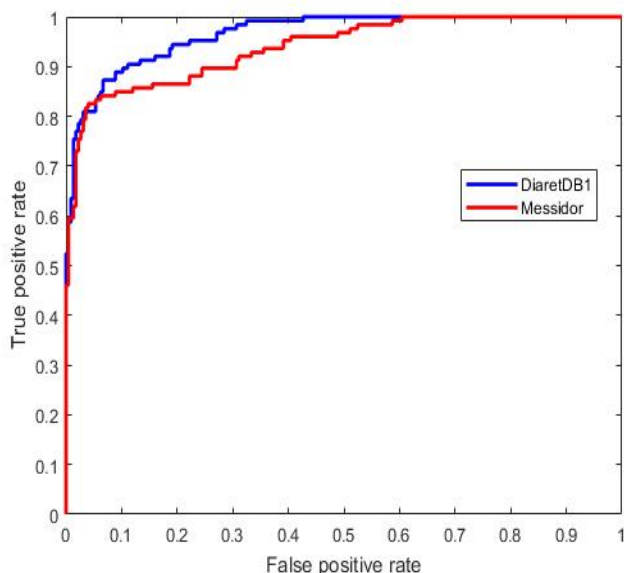

Fig.7. ROC plots of proposed method on DIARETDB1 and MESSIDOR dataset.

6. Conclusions

HM detection is a crucial task in DR. We propose a method based on MSER for extraction of CC which is a dominant technique in blob detection.

The proposed method results demonstrate its satisfactory performance in detecting isolated, large and vessel connected HMs. The method outperforms several other state-of-the-art techniques by achieving lesion level sensitivity of 94.89% and specificity of 98.90% and image level sensitivity of 96.45% and specificity of 97.64% which is quiet encouraging to be used in practice. Hence, proposed method will help experts by acting as a supporting tool in screening programs thereby minimizing their load and filtering patients with DR. As future work, same method can be applied for detection of other lesions and retinal diseases.

References

[1]   D. Klonoff and D. Schwartz, "An economic analysis of interventions for diabetes", *Diabetes Care*, Vol. 23, No. 3, pp. 390-404, 2000.

[2]   M. Abramoff, M. Garvin, and M. Sonka, "Retinal Imaging and Image Analysis", *IEEE Reviews in Biomedical Engineering*, Vol. 3, pp. 169-208, 2010.

[3]   A. Manjaramkar and M. Kokare, "A rule based expert system for microaneurysm detection in digital fundus images", In: *Proc. of International Conf. on Computational Techniques in Information and Communication Technologies*, New Delhi, pp. 137-140, 2016.

[4]   A. Manjaramkar and M. Kokare, "Statistical Geometrical Features for Microaneurysm Detection", *Journal of Digital Imaging*, pp. 1-11, 2017.

[5]   A. Manjaramkar and M. Kokare, "Decision trees for microaneurysms detection in color fundus images", In: *Proc. of International Conf. on Innovations in Green Energy and Healthcare Technologies*, Coimbatore, pp. 1-4, 2017.

[6]   P. Jitpakdee, P. Aimmanee, and B. Uyyanonvara, "A survey on hemorrhage detection in diabetic retinopathy retinal images", In: *Proc. of International Conf. on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, Phetchaburi, pp. 1-4, 2012.

[7]   M. Niemeijer, B. Ginneken, J. Staal, M. Schulten, and M. Abramoff, "Automatic detection of red lesions in digital color fundus photographs", *IEEE Transactions on Medical Imaging*, Vol. 24, No. 5, pp. 584-592, 2005.

[8]   J. Bae, K. Kim, H. Kang, C. Jeong, K. Park, and J. Hwang, "A study on hemorrhage detection using hybrid method in fundus

images", *Journal of digital imaging*, Vol. 24, No. 3, pp.394-404, 2011.

[9] L. Tang, M. Niemeijer, J. Reinhardt, M. Garvin, and M. Abràmoff, "Splat feature classification with application to retinal hemorrhage detection in fundus images", *IEEE Transactions on Medical Imaging*, Vol. 32, No. 2, pp. 364-375, 2013.

[10] G. Kande, T. Savithri, and P. Subbaiah, "Automatic detection of microaneurysms and hemorrhages in digital fundus images", *Journal of digital imaging,* Vol. 23, No.4, pp.430-437, 2010.

[11] X. Zhang and O. Chutatape, "A SVM approach for detection of hemorrhages in background diabetic retinopathy", In: *Proc. of International joint Conf. on Neural Networks,* pp. 2435-2440, 2005.

[12] A. Fleming, K. Goatman, G Williams, S. Philip, P Sharp, and J Olson, "Automated detection of blot haemorrhages as a sign of referable diabetic retinopathy", In: *Proc. of Medical Image Understanding and Analysis,* pp. 235-239, 2008.

[13] C. Marino, E. Ares, M. Penedo, M. Ortega, N. Barreira, and F. Gomez-Ulla, "Automated three stage red lesions detection in digital color fundus images", *WSEAS Transactions on Computers*, Vol. 7, No. 4, pp.207-215, 2008.

[14] C. Kose, U. ŞEvik, C. İKibaş, and H. Erdöl, "Simple methods for segmentation and measurement of diabetic retinopathy lesions in retinal fundus images", *Computer methods and programs in biomedicine,* Vol. 107, No. 2, pp. 274-93, 2012.

[15] Y. Hatanaka, T. Nakagawa, Y. Hayashi, T. Hara, and H. Fujita, "Improvement of automated detection method of hemorrhages in fundus images", In: *Proc. of International Conf. on Medicine and Biology Society*, Vancouver, pp. 5429-5432, 2008.

[16] R. Srivastava, L. Duan, D. Wong, J. Liu, and T. Wong, "Detecting retinal microaneurysms and hemorrhages with robustness to the presence of blood vessels", *Computer methods and programs in biomedicine*, Vol. 138, pp. 83-91, 2017.

[17] L. Seoud, T. Hurtut, J. Chelbi, F. Cheriet, and J. Langlois, "Red lesion detection using dynamic shape features for diabetic retinopathy screening", *IEEE transactions on medical imaging*, Vol. 35, No. 4, pp.1116-1126, 2016.

[18] D. Huang, and C. Wang, "Optimal multi-level thresholding using a two-stage Otsu optimization approach", *Pattern Recognition Letters*, Vol. 30, No. 3, pp.275-284, 2009.

[19] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions", *Image and Vision Computing*, Vol. 22, No.10, pp.761-767, 2004.

[20] T. Kauppi, V. Kalesnykiene, J. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kälviäinen, and J. Pietilä, "The DIARETDB1 Diabetic Retinopathy Database and Evaluation Protocol", In: *Proc. of British Machine Vision Conference* , pp. 1-10, 2007.

[21] E. Decencière, X.Zhang, G. Cazuguel, B. Laÿ, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, and B. Charton, "Feedback on a publicly distributed image database: The Messidor database", *Image Anal. Stereology*, Vol. 33, No. 3, pp.231–234, 2014.

[22] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, No. 1, pp.5-32, 2001.

[23] S. Junior and W. Daniel, "Automatic detection of microaneurysms and hemorrhages in color eye fundus images", *International Journal of Computer Science & Information Technology*, Vol. 5, No.5, pp. 21-37, 2013.