Detecting diabetic retinopathy exudates in fundus images using fuzzy c-means (FCM)†

كشف افرازات اعتلال شبكية العين بمرض السكري باستخدام خوارزمية متوسطات المراكز الضبابية

Tahreer Dwickat* & Hadi Hamad**

تحرير دويكات، وهادي حمد

*MA student :Department of Mathematics, Faculty of Science, An-Najah National University, Nablus, Palestine. **Department of Mathematics, Faculty of Science, An-Najah National University, Nablus, Palestine.

**Corresponding Author: hadihamad@najah.edu

Received: (26/11/2019), Accepted: (28/10/2020)

Abstract

Diabetic Retinopathy (DR) is the main cause of blindness for diabetic patients. As the exudates are the primary sign of DR, therefore early detection and timely treatment can prevent and delay the risk of vision loss. Automatic screening could facilitate the screening process, reduce inspection time, and increase accuracy, which is vital in ophthalmic treatment, this development of exudates detection will help doctors in detecting symptoms faster. In this research, we use an automatic method to detect exudates from retinal digital images with non-dilated pupils of retinopathy patients; starting by detecting both the optic disc (OD) and retinal vessels, then probable exudates are defined through morphological techniques, in the last main phase, four features are implemented as input data for the fuzzy C-means (FCM) clustering to define the existing exudates in the fundus images. The overall detection performance is evaluated through measuring sensitivity, specificity,

† This research was submitted as part of Master thesis in Mathematics by Tahreer Dwickat at An-Najah National University dated.2018

positive predictive value (PPV), positive likelihood ratio (PLR) and accuracy. These measures are done by comparing results to hand-drawn ground truth (GT) done by an expert; which are comparatively analyzed. It is found that the proposed method detects exudates successfully with average values of sensitivity, specificity, PPV, PLR and accuracy of 86.3%, 98.4%, 20.8%, 86.2 and 98.4% respectively on the testing studied database.

Keywords: Diabetic Retinopathy, Retina, Exudates, FCM.

ملخص

اعتلال الشبكية بمرض السكري يعتبر من أكثر الأمراض خطورة على العين والذي قد يؤدي الى فقدان البصر وخاصة عند مرضى السكري. تعتبر الافرازات من أولى العلامات الدالة على الاصابة بالمرض، لذلك من المهم اكتشافها واحصائها من أجل العلاج والوقاية من تفاقم المرض وفقدان البصر. في هذا البحث قمنا باقتراح طريقة تلقائية (اوتوماتيكية) للكشف عن الافرازات من صور شبكية العين لمرضى السكري، والتي بدورها تساعد الأطباء في عملية مراقبة التغير في وضعية العين واكتشاف اعراض امراض العين في مراحل مبكرة. وقد تم التطبيق على صور شبكية العين لمرضى السكري، وكانت الخوارزمية تعتمد اولا على استخراج قرص العين والاوعية الدموية وذلك لتحسين عملية اكتشاف الافرازات، ثم قمنا بتعيين وتحديد الافرازات وذلك بالاعتماد على خوارزمية التجميع الضبابي بالإضافة الى مجموعة من تقنيات معالجة الصور باستخدام برنامج الماتلاب. في النهاية تم تقييم النتائج عن طريق حساب مجموعة معايير منها: الحساسية، النوعية، والدقة وذلك بمقارنتها بالصور التي تم رسمها عن طريق الاخصائيين والأطباء، وكانت متوسطات النتائج كما يلي :86.3%، 98.4% و98.4% على التوالي.

الكلمات المفتاحية: مرض السكري، شبكية العين، افرازات، متوسطات المراكز الضبابية.

Introduction

Eyes are the visualizing organs, where retina is considered as one of the main parts of the eye, it can be seriously damaged by diabetes. This damage affects the tiny blood vessels in it, this disease is known as diabetic retinopathy (DR), by blurring or distorting patients' vision. To overcome this problem an early detection of this disease will help in treating and protecting patients from blindness (Sopharak, Uyyanonvara, & Barman, 2009, p.2194).

Diabetes is the fourth leading cause of death in Palestine after cardiovascular, cancer and cerebrovascular diseases. It is considered as the main cause of blindness associated with DR, where more than 80% of people with diabetes are at risk of developing DR (Ministry of Health, 2017).

A primary sign of DR is exudates, which can be described as a common retinal complication associated with diabetes (Habashy, 2013, p.9), it appears as random whitish or yellowish colored areas with varying shapes, sizes and locations. In general, they exist near the leaking capillaries within the retina (Jelinek, & Cree, 2010).

In machine learning, many clustering methods are present where each one has its merits and limitations. They are mainly accomplished either with crisp (hard) or fuzzy (soft) clustering; the first defines specific boundaries between clusters; hence each pixel belongs to a particular cluster. However, fuzzy clustering can tolerate to define a data to more than one cluster with different memberships; as the membership got larger value, this indicates a higher confidence in assigning the object to the cluster. It is convenient to use fuzzy clustering in non-deterministic problems such as: overlapping intensities, poor contrast, and noise homogeneities variation. This motivated us to use the FCM clustering technique to locate the exudates in fundus images as they are of non-deterministic properties.

Detecting and evaluating DR methods are expensive, manual, not accurate and time consuming. This motivated us (Dwickat, 2018) to contribute for an alternative method towards automatic detection; that has attracted large popularity in medical society in recent years, where computer vision is employed in addition to image processing techniques to detect different features associated with retinopathy (Wellmer, 1998). Automated methods of DR screening help to save time, cost and vision of patients compared to the manual methods of diagnosing (Babu, Vijayan, Susan, et al., 2015).

Related Work

Automatic exudates detection is a desire as it is an important step in improving human health. Hence many scientists worked to implement it in real life; in the following we would like to mention some of the related work in this aspect.

Sopharak et al. (2009) have proposed an automatic method to detect the exudates using FCM clustering algorithm. Firstly, applying contrast enhancement preprocessing, then extract the features from the modified image. These image features consist of intensity, standard deviation of the intensity, and hue. Numbers of edge pixels are used as input parameters to construct a coarse segmentation by using FCM clustering method. After that using some subsequent morphological reconstruction are implemented to obtain better segmentation results. Finally, validated there results by comparing them with ground truth (GT). They have used sensitivity, specificity, positive predictive value (PPV), positive likelihood ratio (PLR) and accuracy to evaluate the overall performance which came to be 87.28%, 99.24%, 42.77%, 224.26 and 99.11%, respectively.

Gandhi, & Raghavan (2015), proposed an automatic method for the detection of exudates using the FCM clustering technique and reconstruction through a superimposition process in the absence of dilating patient's eye. After that, they compared the results of the FCM clustering to the results of another clustering technique called Fuzzy K-Means segmentation. The sensitivity and specificity values for the exudate's detection using the FCM algorithm were 87.38% and 96.94%, respectively. On the other hand, sensitivity and specificity values for exudates detection using the K-Means algorithm were 75.04% and 93.73%, respectively.

Osareh et al. (2002) published their work using FCM clustering to classify the segmented regions into exudates and non-exudates. An artificial neural network classifier was investigated. Their study indicates that automated evaluation of digital retinal images could be used to screen exudative diabetic retinopathy. The proposed system can achieve

a diagnostic performance with 95.0% sensitivity and 88.9% specificity for the identification of images containing any evidence of retinopathy.

Habashy (2013), detects the abnormalities in the retina such as the structure of blood vessels, micro-aneurysms, and exudates using image processing techniques. Identification of DR stages using Fuzzy C-means Classifier. This system intends to help ophthalmologists in the process of screening DR to detect symptoms faster and more easily. The sensitivity, precision and accuracy were 98.01%, 99%, and 97% respectively.

In this research, we organized the work as follow: an introduction was introduced in addition to the related work, while section two introduces the proposed technique which is divided into five parts; preprocessing, extracting probable exudate areas, features employed for abnormality detection, explaining the FCM clustering algorithm and finally describing the system performance by using accuracy measurements. In section three we discuss some numerical results, after that in section four reviews of conclusions and future work.

Proposed Technique

The proposed method is used to automatically discriminate the exudates from normal background in retinal images, based on FCM technique. This method is divided into five parts: first, detect and eliminate the OD and blood vessels from retinal image, after that apply morphological treatment to detect candidate exudates, then implement the FCM clustering algorithm to the candidate results of the previous stage, finally evaluate the result by using accuracy measurements. A flow chart of our proposed methodology is shown in Figure (1).

Preprocessing

The aim of the preprocessing step is to improve the image quality prior to the detection step, this preprocessing is important for many reasons, for instance: retinal images are often poorly and noisy illuminated, camera motion, incorrect settings and eye structure. Another main obstacle for identification of retinal exudates are the wide variety in color of fundus images for different patients, as variations are strongly

correlated to skin pigmentation and iris color (Osareh, 2002; Jelinek et al., 2010, p.504).

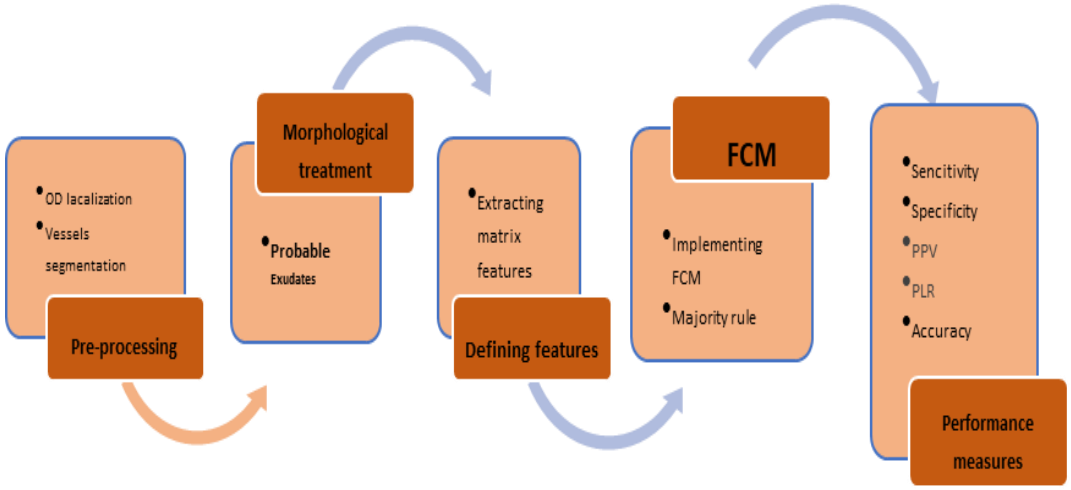


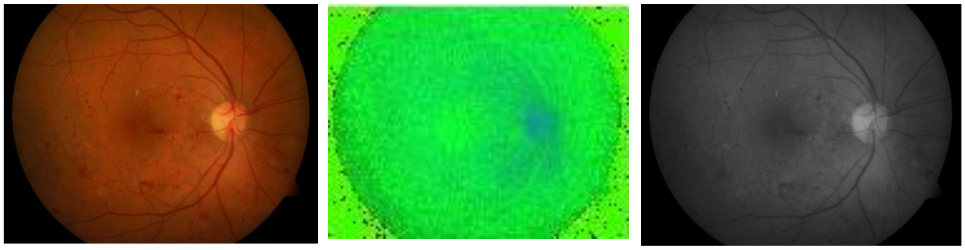
Figure (1): A block diagram for the exudate's detection implemented method.

To reduce the effect of these difficulties we start our preprocessing by eliminating the OD and the vessels as they have similar color as exudates.

Eliminating the OD

The OD is one of the major parts of the retinal image (Wisaeng, Hiransakolwong, & Pothiruk, 2014). Determining the location of the OD is important because of its similarity in brightness, color and contrast to the exudates (Mirajkar, & Patil, 2013).

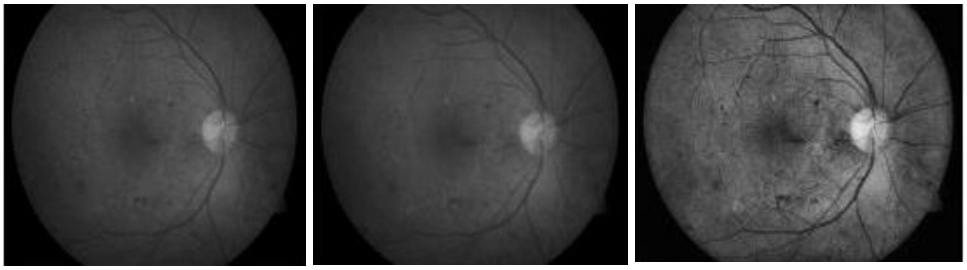
To detect the OD, the intensity component of the HSI (hue-saturation-intensity) color space of the fundus image is used; returning to the idea that the HSI model is suitable for image enhancement (Sopharak et al., 2009, p.2153) (see Figure (2): (b-c)).



(a) (b) (c)

Figure (2): (a) RGB (red-green-blue) (b) HSI, (c) Intensity Channel.

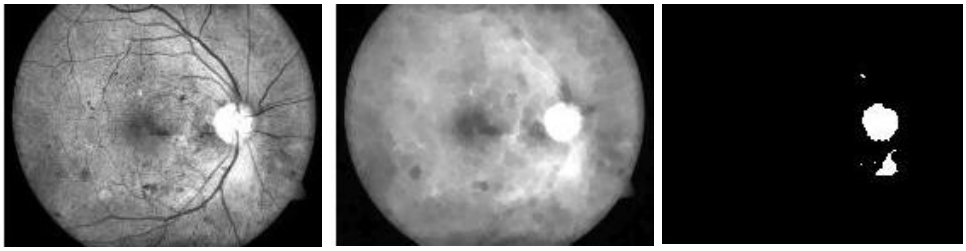
To reduce the noise and preserve edges we apply a median filter on the I (intensity) channel (Figure (3): (b)). Followed by applying contrast limited adaptive histogram equalization (CLAHE) for contrast enhancement. This will improve the contrast of the image and enhance the reparability between the OD and the background (Sopharak, Uyyanonvara, & Barman, 2009). An example is given in Figure (3): (c), the dark and bright areas are clearer hence the image shows more detail (Sinthanayothin, Boyce, Cook, & Williamson, 1999, p.903).



(a) (b) (c)

Figure (3): (a) Intensity image, (b) Median filter, (c) CLAHE.

These steps are followed by morphological closing for the CLAHE image to remove and close blood vessels, consequently creating a fairly constant OD region (Wisaeng et al., 2014). The effect of closing on the CLAHE intensity band images is shown in Figure (4).



(a) (b) (c)

Figure (4): Locating OD (a) An adjusted image (b) Closing, (c) Thresholding.

We note that some areas are not properly representing the OD, see Figure (4): (c), so there is a need to discard these areas. To solve this problem, the label filtering is used to remove isolated pixels by using the concept of labeling connected pixels. The result after applying label filtering is shown in Figure (5), small areas about the OD are deleted and hence the OD is detected a lone.

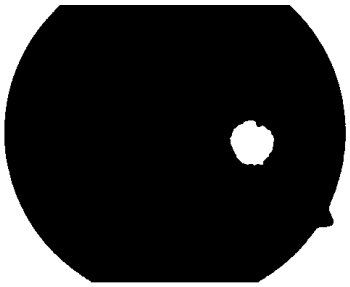
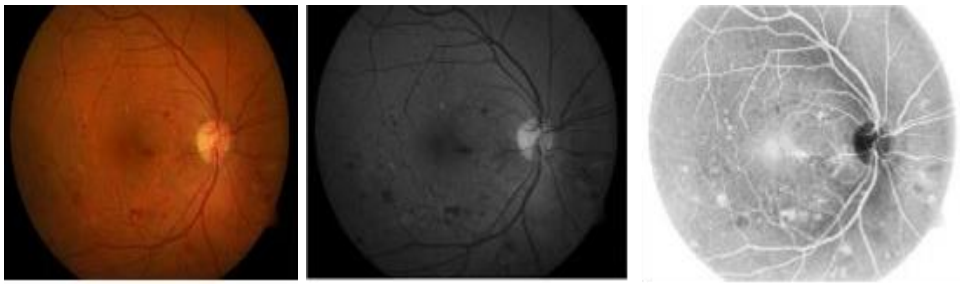


Figure (5): OD detection after applying label filtering.

Eliminating Blood Vessels

Although fundus images are generally saved in RGB format (red-green-blue), the green channel of a retinal image represents the best one to localize blood vessels; as darker blood vessels on a brighter background (Akhavan, & Faez, 2014, p. 562), see Figure (6): (b).

The "CLAHE" technique is applied to the green channel for contrast enhancement (Babu et al., 2015), this makes the image has well contrast inversion as shown in Figure (6)**:** (c).



(a) (b) (c)

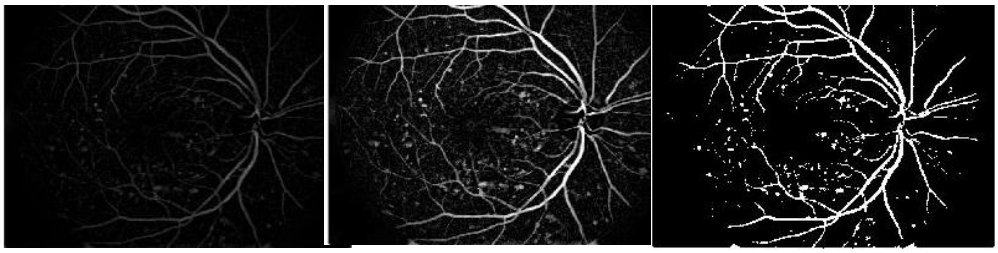
Figure (6): (a) RGB image, (b) Green channel, (c) CLAHE on green channal.

After that, the Gaussian filter is applied to the CLAHE image, this step is done to remove noise and unnecessary details that are not belonging to vessels.

1. And now, applying the top-hat transformation on the filtered image with a disk structure element (SE). In this work, a disk SE of radius 3 and 15 are used to fill all holes in blood vessels. The top-hat returns an image, containing the parts that are smaller than the SE and are darker than their surroundings as shown in Figure (7): (a). Actually, blood vessels appear as clear elongated objects while the background spreads to be black (Akhavan et al., 2014, p.565).

To increase the contrast within the image, an adjustment is performed on it. This modification gives a modified image as shown in

Figure (7): (b). in removing unnecessary details thresholding is applied, the resulted image is a binary one shown in Figure (7): (c).



(a) (b) (c)

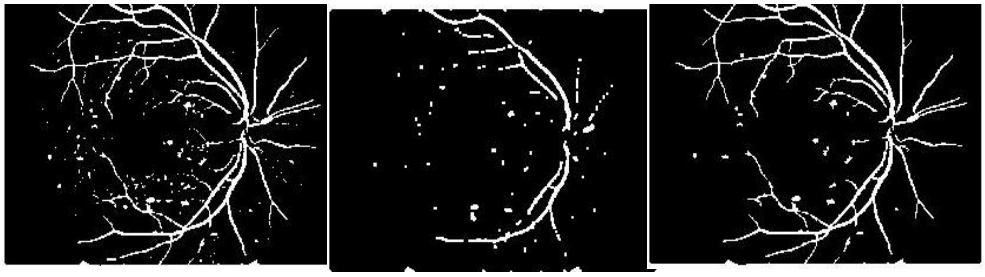
Figure (7): Thresholding vessels: (a) Top-hat transformation
(b) Adjusted intensity (c) Thresholding.

2. Mathematical morphology is mostly used to analyze the shape of the image and to remove imperfections by accounting for the form and structure of the image (Akhavan et al., 2014, p.562).

In this part we used opening and opening-closing by reconstruction to clean up the image, but before these steps, a median filter is applied to reduce the noise; this is illustrated in Figure (8): (a). Reconstruction is a morphological transformation involving two images and a structuring element; one of them is marker and the other is the mask.

3. The final step is to complement the output of the morphological reconstruction as shown in Figure (8). We notice that the reconstruction based on opening and closing are more effective than standard opening and closing for removing small blemishes without affecting the overall shapes of the objects (El-Abbadi et al., 2013; Kumar, & Yanan, 2010).

In Figure (8): (c) after applying opening-closing by reconstruction, we note that some areas are not blood vessels, we used label filter to overcome this problem.



(a) (b) (c)

Figure (8): (a) Median filter, (b) Opening, (c) Opening-closing reconstruction.

Label filtering removes isolated pixels by using the concept of connected pixels labeling. The final result for blood vessels network after implementing the label filtering is shown in Figure (9): (a).

After applying the algorithm to extract both the OD and blood vessels, the final result is superimposed on the original RGB image (see Figure (9): b). Consequently, the image is ready to be used to identify predicted exudates.

(a) (b)

Figure (9): (a) Blood vessels network after implementing label filter, (b)Masking OD and blood vessels network in a retinal image.

Predicted Exudates

The green channel from the modified RGB image (excluding the OD and vessels) is used, as it gives the best contrast between background and

exudates (the histogram shows the intensity distribution where most dark pixels appear in a small range representing the background and the lighter components such as exudates represent the rest of the histogram, see Figure (10): (f)), note that exudates appear brighter than the background in the green channel, but in red and blue channels exudates look less clear.

In this part the proposed method to predict probable exudate areas from retinal images is introduced, segmented areas will be considered as an input for FCM clustering. However, this reduction of data (exclude the non-probable exudate pixels) will reduce the miss-classification and time of processing.

The process of extracting probable exudates is illustrated in Algorithm (1) below.

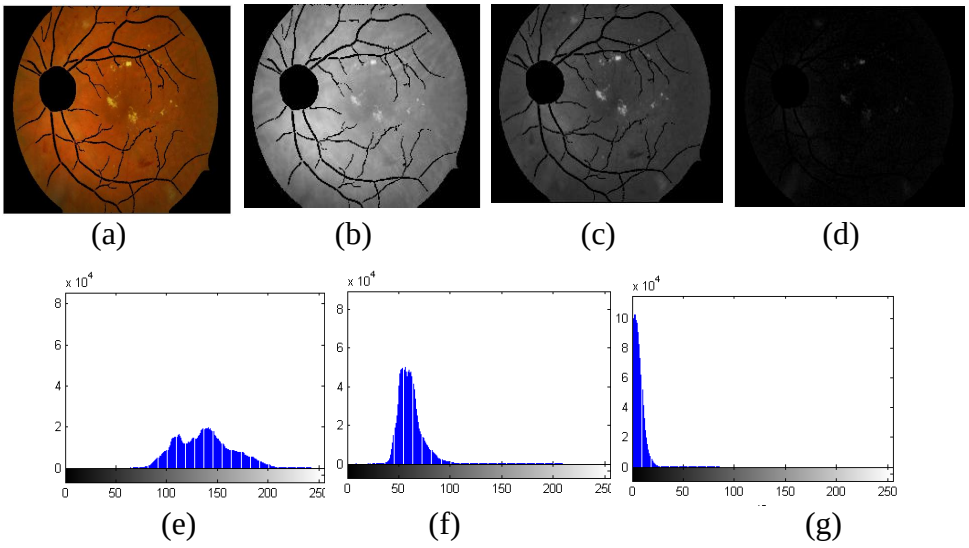


Figure (10): Illustrating the expected exudates in the pre-detection step (a) Eliminating OD and blood vessels, (b) Red; (c) Green; (d) Blue; (e-g) Corresponding intensity distribution red, green and blue respectively.

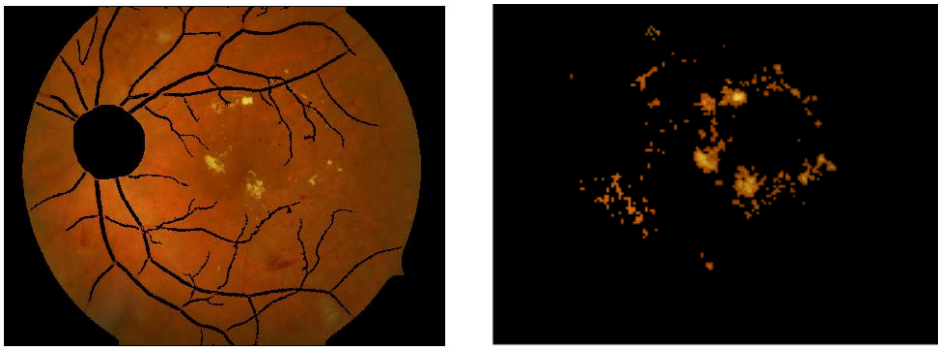
ALGORITHM (1)

Input: RGB Retinal image I.
Output: Retinal image contains the candidate exudate areas that can be considered as probable exudates.
Step (1): Extract Green channel image I_G. **Step (2):** Pixel's values are squared in the green channel (I_{sq}), then change the resulted matrix to the type double (I_d). **Step (3):** Apply the sobel edge filter on the (I_d) image giving image (I_s). **Step (4):** Take the complement of the (I_d) image, name it as (I_c). **Step (5):** Apply the sobel edge filter on the (I_c) image, the result is called (I_{sc}). **Step (6):** Apply the dilation and complement on the (I_{sc}) image, then the result is (I_{di}). **Step (7):** Take the "AND" between (I_s) and (I_{di}), name it as (I_{AND}). **Step (8):** Perform the close operation on the (I_{AND}) image, the result is (I_{close}). **Step (9):** Apply the dilation operation on the (I_{close}) image, the result is (I_{out}).

After applying the mentioned preprocessing stages, the output result contains the candidate exudate areas (see Figure (11): (b)).

We notice in Figure (11): (b) that some areas are included in the extracted exudates which are not really exudates. Therefore, to solve this

problem a method to detect the actual exudates has been added, in this work we are using FCM clustering method.



(a) (b)

Figure (11): Expected exudates after preprocessing stage (a) Input modified retinal RGB image, (b) Expected exudates.

The probable exudate output will be used in the next phase to extract the features that will be implemented as an input for the FCM clustering.

Used Features

Proper features where chosen depending on properties of exudates. Traditionally, exudates are defined by ophthalmologists as bright yellow spots with variable shape and size, but they have indistinctive edges. These different characteristics and others are implemented to extract proper data (from the probable exudates) to represent the input for the FCM clustering algorithm.

About seventy features were implemented and tested, finally four most influential features were experimentally selected to be used in FCM, and these are: hue, entropy and standard deviation for intensity and Y-channel. Selected features and their details are explained in **(Appendix B).**

Segmentation Using FCM Clustering

FCM is an overlapping clustering algorithm which can be used for segmentation, where each pixel may belong to one or more clusters with

different degrees of membership. It assigns any unclassified pixel to the closest cluster based on similarity measures; which means the distance of the features vector to the cluster centers ($\boldsymbol{V}$) (e.g., exudates and non-exudates). The Euclidean distance is used to measure the similarity, consequently the data will be associated to an appropriate membership value (Balasko, Abonyi, & Feil, 2002; Mohammed, Alnoamani, & AbdulZahraa, 2016).

The FCM clustering algorithm is briefly illustrated below:

The input data matrix contains the features values for candidate exudates:

$$\boldsymbol{X} = [x_1, x_2, \dots, x_M], \qquad x_i \in \mathbb{R}^N \qquad \dots\dots\dots\dots\dots (1)$$

The objective function, the member ship, and center of the clusters are given by:

$$\boldsymbol{J}(\boldsymbol{X}; \boldsymbol{U}, \boldsymbol{V}) = \sum_{i=1}^{M} \sum_{k=1}^{c} (\mu_{ik})^2 \|x_i - v_k\|^2 \ , \qquad \dots\dots\dots (2)$$

$$\mu_{ik} = \left[\sum_{j=1}^{c} \left(\frac{\|x_i - v_k\|}{\|x_i - v_j\|} \right)^2 \right]^{-1} , \qquad \dots\dots\dots\dots\dots\dots\dots\dots (3)$$

$$v_k = \frac{\sum_{i=1}^{M} (\mu_{ik})^2 \, x_i}{\sum_{i=1}^{M} (\mu_{ik})^2} , \qquad \dots\dots\dots\dots\dots\dots\dots\dots (4)$$

where M is the number of features used as input for FCM to identify the exudates (four features in our case), N is the number of pixels in an image that contains the candidate exudates, $\boldsymbol{c}$ is the number of clusters (in our case, experimentally we tried two), x_i is the i^{th} item of the d -dimensional measured data, μ_{ik} is the degree of membership of x_i in the cluster k, v_k is the center of the cluster, and the symbol $\|*\|$ is used to

represent any norm expressing the similarity between any measured feature and the corresponding centers.

Fuzzy partitioning is carried out through an iterative optimization of the objective function by updating μ_{ik} and v_k. The iteration stops when:

$$\max_{ik}\{|\mu_{ik}{}^{(j+1)} - \mu_{ik}{}^{(j)}|\} < \varepsilon \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

where ε was set to 0.00001, and j is the iteration number which is set to a maximum of 100.

1.1.1. Executing the FCM Clustering Algorithm

The FCM algorithm is executed for a selected set of input features as described in the following steps:

Step 1: Initialize the fuzzy membership (partition) matrix

$$\boldsymbol{U} = [\mu_{ik}]\left(\boldsymbol{U}^{(0)}\right)$$

by generating random numbers in the range 0 to 1 subject to eq.(6):

$$\sum_{i=1}^{M}\sum_{k=1}^{c}\mu_{ik} = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

where c is the number of clusters and *M* is the number of features.

Step 2: Compute the centers vectors at j-iteration

$$V^{(j)} = [v_k]$$

according to eq. (4).

Step 3: Update the fuzzy partition (membership) matrix $\boldsymbol{U}^{(j)}$ to obtain $\boldsymbol{U}^{(j+1)}$ according to the new computed μ_{ik}, refer to eq.(3).

Step 4: Compute the objective function (J), refer to eq.(2). If the difference between $U^{(j+1)}$ and $U^{(j)}$ is less than or equal (ε) or the number of iterations is equal to L (maximum number of iterations), then stop, otherwise return to step 2.

The FCM output defines cluster centers and membership grade for each pixel.

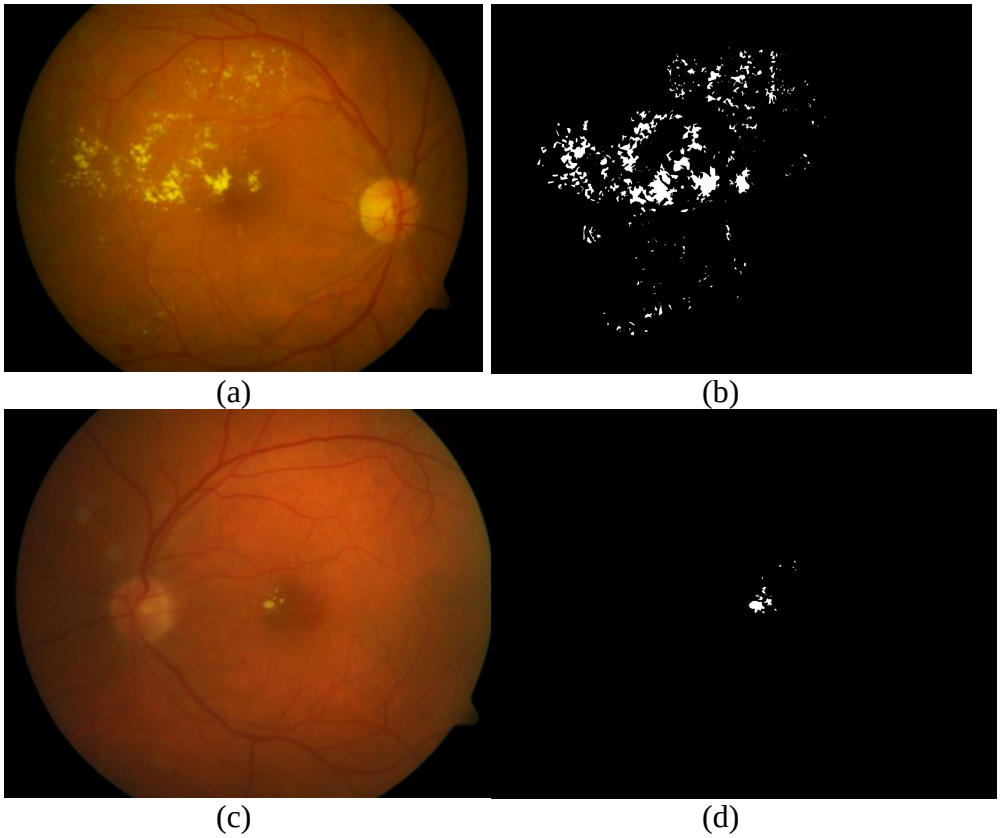
Evaluating the FCM Clustering

Several standard metrics are used to evaluate the segmentation performance, it is measured by comparing the automated detected results with hand-drawn ground truth (GT) provided by expert ophthalmologists for retinal images (Sopharak et al., 2009, p.2156), see Figure (12).

True positive (TP), false positive (FP), false negative (FN), true negative (TN), sensitivity, specificity, accuracy, positive predictive value (PPV) and positive likelihood ratio (PLR) are used here to analyze the performance of the FCM clustering. Their definitions accommodated to this work are given in the following:

- TP: is the number of pixels detected as exudates in both GT and result image.
- FP: is the number of pixels detected as exudates in result image, but are non-exudates in GT image.
- FN: is the number of non-exudate pixels detected in result image, but are identified as exudates in GT image.
- TN: is the number of non-exudate pixels which are correctly identified as non-exudate pixels.
- To be able to measure and compare results and performance with other algorithms, a set of popular metrics are used. Their definition regarding this work is introduced here:
- Sensitivity (or true positive rate): measures the potential of FCM clustering to positively detect the exudates, on other words, the ability to identify exudate pixels correctly, and is hence given by:

$$\text{Sensitivety} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad \ldots\ldots\ldots\ldots\ldots (7)$$



(a) (b)

(c) (d)

Figure (12): (a) & (c) RGB retinal images, (b) & (d) GT exudates, respectively.

– Specificity (or true negative rate): measures the potential of FCM clustering to correctly detect the non-exudate pixels, or the ability to identify non-exudate pixels correctly, and is given by:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad \ldots\ldots\ldots\ldots\ldots (8)$$

- Accuracy: measures the potential of FCM clustering to get accurate results while detecting exudates and non-exudates, it is given by:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad \ldots\ldots\ldots\ldots\ldots(9)$$

- Positive Predictive Value (PPV): is the ratio of exudate and non-exudate pixels which are positively detected, or is the probability to detect correct exudate pixels out of all detected exudate pixels by the method. Mathematically can be expressed as:

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \ldots\ldots\ldots\ldots\ldots(10)$$

- Positive Likelihood Ratio (PLR): is defined as follows:

$$\text{PLR} = \frac{\text{Sensitivety}}{1 - \text{specificity}} = \frac{\text{TP} * (\text{TN} + \text{FP})}{\text{FP} * (\text{TP} + \text{FN})} \quad \ldots\ldots(11)$$

We notice from eqs. (7), (10) and (11) that the higher TP values are, it will increase the sensitivity, PPV, and PLR values; this means; the identification of exudates is happening more correctly (Sopharak, 2009, p.2156; Jelinek, & Cree, 2010; Kansal, & Nishi, 2016).

Numerical Results

In this research, we have used five images for training and another five for testing to evaluate the performance of the proposed method; we used DIARETDB1 dataset (Kauppi, et al. 2007). The used images are a subset of the DIARETDB1 Project's dataset which consists of 89 color fundus images. The wide variability in color of fundus images for different patients and containing normal and abnormal with symptoms of retinopathy made the database a pinch mark one. Ten abnormal images out of them where selected. The images have a size of 1500×1152 pixels stored in PNG format.

The proposed algorithm is implemented using MATLAB version R2012b and run on a Laptop with 4GB Ram and 2.2GHz Intel Core i3.

After features selection, we applied the FCM clustering algorithm on images for detecting exudates, this technique can be used to identify how much exudate pixels are present in a retinal fundus image, results are shown in Figure (13) and Tables (1-2).

As the candidate exudates contain both exudates and non-exudates, we separate the data using the FCM clustering. In Figure (13). the first cluster contains non exudate areas in the probable exudates, while the second cluster contains FCM classified exudate areas (Figure (13): (c), (d)). From the experimental results, we notice that the FCM clustering technique gives a high true positive value, while pixels located around the OD have higher false positive values, this is shown in Figure (13): (e). To overcome this difficulty; we used the erosion mathematical morphology process to reduce the false positive pixels, see Figure (13): (f).

To evaluate the performance of our technique, we compare the results of the extracted exudates with ophthalmologist's hand-drawn GT images (Figure (13): (b)). This approach aims to measure the correctness of the algorithm at the pixel level by using eqs. (8), (9), (10) and (11). Using the FCM clustering algorithm on training and testing images, the maximum value of sensitivity, specificity, accuracy, PPV, and PLR are 97.6%, 99.8%, 99.9%, 41.7, and 622.6 respectively. Detailed results of performance measurement are presented in Tables (1-2).

Sensitivity of exudates detection (true positive rate) is an important metric as it measures the potential of FCM to do correct clustering. We note in eq. (7); if TP is increasing the sensitivity will increase, and if FN is decreasing the sensitivity will increase and gives the best result. If TP is high and FP is low the PPV will be high, but in this work the FP is high so the value of PPV is low. To increase the PPV value we used erosion morphological process, results are presented in Tables (3-4).

Another important metric which measures the potential of FCM clustering to get accurate results is accuracy, which increases as TP increases.

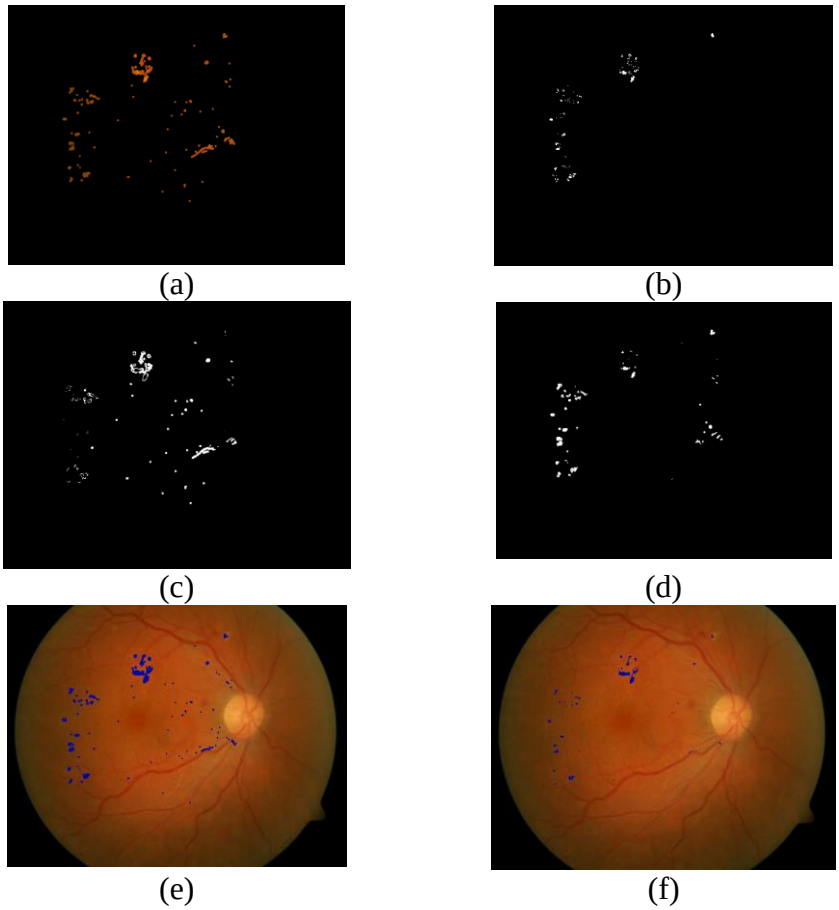


Figure (13): FCM clustering results: (a) RGB retinal expected exudates, (b) GT, (c) cluster 1 (non-exudates) (d) cluster 2 (exudates), (e-f) FCM clustering results before and after applying mathematical morphology.

Table (1): Exudates detection results using FCM on training images.

24-bit images	TP	FP	FN	TN	Sensitivity (%)	Specificity (%)	PPV (%)	PLR	Accuracy (%)
Image 1	12244	25973	714	1689069	94.5	98.5	32	62.6	98.5
Image 2	1950	25927	48	1700075	97.6	98.5	7	65.1	98.5
Image 3	1826	3116	260	1722798	87.5	99.8	37	486.3	99.8
Image4	845	2204	199	1724752	80.9	99.9	27.7	622.6	99.9
Image 5	3930	12597	843	1710630	82.3	99.3	23.8	112.8	99.2

Table (2): Exudates detection results using FCM on testing images.

24-bit Images	TP	FP	FN	TN	Sensitivity (%)	Specificity (%)	PPV (%)	PLR	Accuracy (%)
Image 1	19095	26752	2763	167939	87.4	98.4	41.7	55.6	98.3
Image 2	3465	21551	553	1702431	86.2	98.8	13.9	69	98. 7
Image 3	11166	63772	3547	1649515	75.9	96.3	14.9	20.4	96.1
Image4	3411	7790	351	1716448	90.7	99.6	30.5	201.5	99.5
Image 5	452	15434	43	1712071	91.3	99.1	2.8	102.6	99.1

Table (3): Exudates detection results using FCM, after applying morphological processing on training images.

24-bit images	TP	FP	FN	TN	Sensitivity (%)	Specificity (%)	PPV (%)	PLR	Accuracy (%)
Image 1	11152	7917	1806	1707125	86.06	99.5	58.5	187.1	99.4
Image 2	1827	5261	171	1720741	91.44	99.7	25.8	304.8	99.7
Image 3	1713	934	373	1724980	82.12	99.9	64.7	1642.4	99.9
Image4	783	1597	261	1725359	75.00	99.9	32.9	750	99.9
Image 5	3345	5127	1428	1718100	70.08	99.7	39.5	233.6	99.6

Table (4): Exudates detection results using FCM, after applying morphological processing on testing images.

24-bit images	TP	FP	FN	TN	Sensitivity (%)	Specificity (%)	PPV (%)	PLR	Accuracy (%)
Image 1	16908	10484	4950	1695658	77.4	99.4	61.7	126.8	99.1
Image 2	2875	5766	1143	1718216	71.6	99.7	33.3	216.8	99.6
Image 3	8694	18449	6019	1694838	59.1	98.9	32.0	54.7	98.6
Image4	3031	3223	731	1721015	80.6	99.8	48.5	424.1	99.8
Image 5	364	1478	131	1726027	73.5	99.9	19.8	817.1	99.9

Conclusions and Future Work

Scientific research is oriented to improve quality of life; this work is a small contribution devoted to medical image processing, particularly to detect the DR disease which is considered as the main cause of blindness if not treated in its early stages. Current detection methods of DR are expensive, manual, time consuming, and require trained ophthalmologists.

Our proposed method detects exudate areas automatically, it is based on the FCM clustering; starting by detecting OD and blood vessels to prevent misclassification and facilitate exudate detection. Four features are used as input data for FCM: hue, entropy and standard deviation for both intensity and Y-channel.

Processing retinal images have many obstacles due to the existence of artifacts, variability of color in fundus images due to the human pigment, being non sharp, different angles of views, in addition to the spherical shape of the retina. For these reasons and more, current works are still in developing stages to overcome these difficulties.

Finally, to evaluate the performance on selected DIARETDB1 database images, it is found that the proposed method detects exudate areas successfully with average sensitivity, specificity, and accuracy of 86.3%, 98.4%, and 98.4% respectively on the testing used images.

Suggestions to be implemented to improve the performance of the algorithm and avoid the difficulties are introduced in the following:

- Increasing the number of retinal images for both training and testing.
- Modifying the methods to localize and segment the OD and blood vessels precisely.
- Using other clustering algorithms and comparing results.
- Extending the proposed approach to detect spot lesions, namely micro-aneurysms and hemorrhages.
- Improving the features to define exudates characteristics more effectively.

References

- Akhavan, R., & Faez, K. (August 2014). A Novel Retinal Blood Vessel Segmentation Algorithm using Fuzzy Segmentation, *vol. 4, No. 4, pp. 561-572,* ISSN: 2088-8708. 561.
- Babu, S., Vijayan, S., Susan, A., & et al. (2015), Automatic Segmentation of Fovea and Classification of Different Stages of Diabetic Retinopathy, *Hindusthan College of Engineering and Technology, Coimbatore, Surya Engineering College*, India, vol. 10, No. 7.
- Balasko, B., Abonyi, J., & Feil, B. (2002). Fuzzy Clustering and Data Analysis Toolbox for Use with Matlab, *Department of Process Engineering University of Veszprem*.
- Dwickat, T. N. (2018*). Automatic Detection of Diabetic Retinopathy in Fundus Images by Using Fuzzy C-means (FCM) Clustering*. (Unpublished Master thesis). An-Najah National University, Nablus, Palestine.
- El-Abbadi, N., & Al-Saadi, E. (2013), Blood Vessel Extraction Using Mathematical Morphology, *Journal of Computer Science, vol. 9, No. 10, pp. 1389-1395*, ISSN: 1549-3636.
- Gandhi, M., & Raghavan, D. (April 2015), Detection and Localization of Retinal Exudates for Diabetic Retinopathy, *Journal of Biological Systems, vol. 23, No. 2*, pp.195-212.
- Habashy, S. (2013). Identification of Diabetic Retinopathy Stages using Fuzzy C-means Classifier, Electronic and Communication Department, Engineering Collage, *Helwan University, Egypt, International Journal of Computer Applications, vol. 77, No. 9*, pp. 7-13.
- Jelinek, H., & Cree, M. (2010). Automated Image Detection of retinal pathology, ISBN: 978-0-8493-7556-9.

- Kansal, K., & Nishi, E. (2016). Automated Detection of Exudates for Diabetic Retinopathy Screening in Fundus Images Using CS-ACO Optimization Approach, *International Journal of Bio-Science and Bio-Technology, vol. 8, No.3, pp. 323-340*, ISSN: 2233-7849.
- Kauppi, T., Kalesnykiene, V., Kamarainen, J., Lensu, L., Sorri, I., Raninen, A., Voutilainen, R., Uusitalo, H., Kalviainen, H., Pietila, J., 2007. Diaretdb1 Diabetic Retinopathy Database and Evaluation Protocol. Proc. Medical Image Understanding and Analysis (MIUA), pp. 61–65.
- Kumar, V., & Yanan, N. (2010). Blood Vessel Extraction Using Wiener Filter and Morphological Operation, *International Journal of Computer Science & Emerging Technologies, India, (E-ISSN: 2044-6004), vol. 1*, Issue 4.
- Ministry of Health, Health Annual Report – Palestine 2016, (July 2017). General Directorate of Health Policies and Planning, *Palestinian Health Information Center*, Palestine.
- Mirajkar, S., & Patil, M. (2013). Feature Extraction of Diabetic Retinopathy Images, *International Journal of Computer Applications (IJCA), Proceedings on Emerging Trends in Electronics and Telecommunication Engineering (NCET).*
- Mohammed, H., Alnoamani, H., & AbdulZahraa, A. (2016). Improved Fuzzy C-Mean Algorithm for Image Segmentation, *International Journal of Advanced Research in Artificial Intelligence (IJARAI), vol. 5*, No. 6.
- Osareh, A., Mirmehdi, M., & Thomas, B. (2002). Classification and Localizations of Diabetic-Related Eye Disease, *European Conf. on Computer Vision (ECCV)*, pp. 502-516.
- Rashid, S. (2013). Computerized Exudate Detection in Fundus Images Using Statistical Feature based Fuzzy C-mean Clustering,

International Journal of Computing and Digital Systems, vol. 2, No. 3, pp.135-145.

- Sinthanayothin, C., Boyce, J., Cook, H., & Williamson, T. (1999). Automated Localization of the Optic Disc, Fovea and Retinal Blood Vessels from Digital Color Fundus Images, *Br. J. Opthalmol. vol. 83,* pp. 902-910.
- Sopharak, A., Uyyanonvara, B., & Barman, S. (2009). Automatic Exudate Detection from Non-Dilated Diabetic Retinopathy Retinal Images Using Fuzzy C-means Clustering, *Sensors (Basel, Switzerland), vol. 9,* pp. 2148-2161.
- Sopharak, A., Uyyanonvara, B., & Barman, S. (2009). Automatic Exudates Detection for Diabetic Retinopathy Screening, *doi: 10.2306/scienceasia1513-1874.2009.35.080*.
- Wang, X., Hansch, R., Ma, L., & Hellwich, O. (2014). Comparison of Different Color Spaces for Image Segmentation using Graph-cut, *International Conference on Computer Vision Theory and Applications (VISAPP), IEEE, Lisbon*, pp. 301-308.
- Wellmer, F. (1998). Statistical Evaluations in Exploration for Mineral Deposits, *ISBN: 978-3-642-64325-5, Springer-Verlag, Berlin, Heidelberg*.
- Wisaeng, K., Hiransakolwong, N., & Pothiruk, E. (2014). Automatic Detection of Optic Disc in Digital Retinal Images, *International Journal of Computer Applications, Thailand,* vol. 90, No 5.

Appendices

A. Color Spaces

Different color spaces are used to describe color images in different modes with specific properties in image processing. The most important types of color spaces are: RGB (Red-Green-Blue), HSI (Hue-Saturation-Intensity), XYZ, Lab, and luv which can be used as features for FCM clustering (Wang, Hansch, Ma, & Hellwich, 2014)

B. Variance, Standard Deviation and Entropy

The **variance** (which we denote by σ^2) is a measure of how spread out a data set is; in other words, it is the average squared deviation of values from the mean (Wellmer, 1998). The variance of a set of values is defined as:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots (B1)$$

where x_i is the i^{th} data value, N is the number of data values, μ represents the data mean which is calculated by:

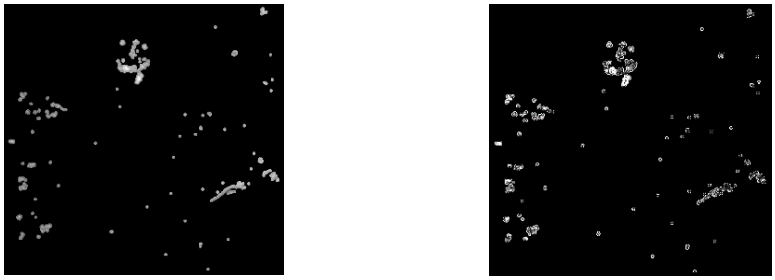
$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad \ldots\ldots\ldots\ldots\ldots\ldots (B2)$$

In this research, the variance is calculated for all pixels in an image of size ($n \times n$) by partitioning it into sub matrices of smaller size ($m \times m$). Then, calculating the mean for these sub matrices using eq. (B2). Finally, computing the variance of the center pixel by using eq. (B1). The result is a matrix of size ($n \times n$) that contains the variance value for each pixel compared with the surrounding pixels, an example is shown in Figure (B1).

The variance value of color/intensity ranges between 0 and 255, from black to white, respectively. The white refers to the high variance, while the black pixels refer to zero variance which means the data in the sub-matrix are of the same color, in between values represent lower contrast between the calculated pixels. Another important relevant feature is the **standard deviation** (σ), it is a useful measure of data spreading, commonly used as a statistical measure. This feature is used in this work to compare between exudate and non-exudate pixels. Standard deviation is calculated by taking the square root of the variance of a sub matrix, (Wellmer, 1998) as shown in eq. B3.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (B3)$$

Standard deviation shows the amount of existing variation from the mean. When the standard deviation value is small, it indicates that the pixel values tend to be very close to the mean. Whereas, high value indicates that the pixel values are spread out over a large range of values, which means pixels are different, see Figure (B2): (a-b).



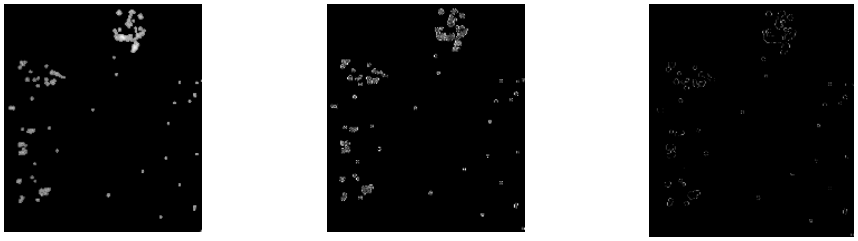
(a) (b)

Figure (B1):(a) Green image, (b) Variance.

Entropy is a statistical measure of randomness; it can be used to characterize the texture of an input image. It tries to describe how much randomness is there in an image. When the value is high, the pixel is different from the pixels around it, accordingly, low entropy images represent very little contrast and large runs of pixels with the same intensity. An image that is perfectly flat will have zero entropy (Rashid, 2013, p.142). Entropy is defined as:

$$H(x) = -\sum_{i \in W(x)} p_i \log_2 p_i \quad \ldots\ldots\ldots\ldots\ldots\ldots (B4)$$

where x is a set of all pixels in a sub-window $W(x)$, p_i is the histogram count in sub-window i. A window of size 3×3 pixels is implemented as shown in the next figure; its result is shown in Figure (B2): (c):



(a) (b) (c)

Figure (B2): (a) Green image, (b) Standard deviation, (c) Entropy.