Analysis of Texture-based Features for Image Classification of Retinal Exudates

Hanung Adi Nugroho
Department of Electrical Engineering and Information Technology, Faculty of Engineering, Universitas Gadjah Mada, Jalan Grafika 2, Kampus UGM, Yogyakarta 55281, Indonesia
adinugroho@ugm.ac.id

Widhia KZ Oktoeberza
Department of Electrical Engineering and Information Technology, Faculty of Engineering, Universitas Gadjah Mada, Jalan Grafika 2, Kampus UGM, Yogyakarta 55281, Indonesia
widhia.oktoeberzakz.mti13@mail.ugm.ac.id

Ratna Lestari Budiani
Department of Electrical Engineering and Information Technology, Faculty of Engineering, Universitas Gadjah Mada, Jalan Grafika 2, Kampus UGM, Yogyakarta 55281, Indonesia
buana.ratna@gmail.com

Teguh Bharata Adji
Department of Electrical Engineering and Information Technology, Faculty of Engineering, Universitas Gadjah Mada, Jalan Grafika 2, Kampus UGM, Yogyakarta 55281, Indonesia
adji@ugm.ac.id

ABSTRACT

Diabetic retinopathy is one of the primary causes of blindness as complication of long term diabetes. The permanent vision loss can be avoided by conducting early detection of retinopathy symptoms such as retinal exudates. This paper proposes a scheme to classify fundus images whether containing exudates based on analysis of extracted texture features. Removal of optic disc and detection of exudate candidate area were firstly conducted. Afterwards, some texture features consisting of five features of grey level co-occurrence matrices (GLCM), eleven features of grey level of run-length matrices (GLRLM) and six histogram-based features were extracted from candidate exudates detected. These extracted features subsequently underwent classification process by using multi-layer perceptron (MLP) classifier. The performance of proposed scheme was evaluated on 80 fundus images taken from DIARETDB1 comprising of 38 images with exudates and 42 images without exudates. The best evaluation result of classification was achieved by using five GLCM features with 95% of accuracy, 97.37% of sensitivity and 92.86% of specificity. These results indicate that the proposed scheme successfully detects exudates and also classifies fundus images either containing exudates or no exudates. In addition, the implementation of the proposed scheme is expected to assist the ophthalmologists in monitoring and diagnosing diabetic retinopathy especially on the presence of retinal exudates.

CCS Concepts

• **Applied computing~Imaging**

ICISPC 2017, July 26–28, 2017, Penang, Malaysia

ACM ISBN 978-1-4503-5289-5/17/07…$15.00
DOI: https://doi.org/10.1145/3132300.3132333

Keywords

Diabetic retinopathy; retinal exudates; fundus images; texture analysis; grey level co-occurrence matrices (GLCM); grey level of run-length matrices (GLRLM); histogram features.

1. INTRODUCTION

As reported by the International Diabetes Federation (IDF), around 415 million adults worldwide in 2015 suffered from diabetes and it is predicted to increase up to 642 million by 2040. A person dies from diabetes in every six seconds. Even in high-income countries, diabetes is one of the leading causes of blindness due to its complication on the retina (retinopathy) caused by damaging of retinal vessels. Regular eye screening is an essential step for early retinopathy detection to prevent permanent loss of vision [1].

REF

One of the symptoms in retinopathy is the presence of retinal exudates [2-5]. Exudate is the bright lesion in the retina generated by discharging of plasma and white blood cells from damaged retinal vessels. Exudates have some characteristics similar to the optic disc (OD) especially in terms of colour, luminance, contrast and texture [6] as depicted in Figure 1. The characteristic of retinal colour also depends on the properties of optical pigments and the layers structure of retina [7].

REF

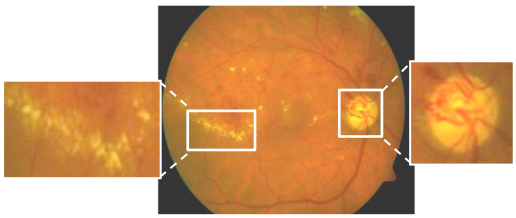

Figure 1. The characteristics of exudates and optic disc [8]

Various research works on the detection of exudates from colour fundus images have been proposed. Eadgahi and Pourreza [9] conducted a segmentation of hard exudates based on morphological operation, i.e. combining the process of top and

bottom hats and reconstruction operation. The bottom hat was used to remove retinal vessels from fundus images while the top one was applied to obtain hard exudates areas. Optic disc area was eliminated using Sobel operator. Aftab and Akram [10] used a closing as a morphological operation to eliminate dark lesion, i.e. haemorrhages and retinal vessels. Having filtered bright lesions by using Gabor filter, the image was subsequently used as a candidate image for exudates segmentation based on adaptive thresholding. Finally, a feature extraction was obtained from the basic properties of exudates classified using Gaussian mixture model (GMM).

Garcí a *et al*. [11] focused on the detection of hard exudates based on statistical and shape features extraction. In their works, the candidate areas of exudates were obtained from a normalised image segmented by combining the global and adaptive thresholding methods. Finally, the best features were classified using radial basis function (RBF) classifier. This research was continued in [12] by adding multi-layer perceptron (MLP) classifier method. It was found that the classification result of MLP was better than that of RBF. Nugroho *et al*. [13] used the combination of grey level co-occurrence matrices (GLCM) and lacunarity features for detecting exudates. Clustering-based method was applied to yield the candidates of exudate. The extracted features were later classified by using Naïve Bayes classifier. On the other side, Nugroho *et al*. also applied texture-based histogram and GLCM features to identify the masses of malignant on digital mammogram images [14].

To complete the study of retinal exudates detection, this paper proposes a scheme to detect exudates based on the analysis of texture features for feature extraction. The extracted texture features, which comprise of GLCM, grey level of run-length matrices (GLRLM) and histogram features, subsequently underwent classification by using MLP classifier. The structure of this paper is organised as follows. Section II explains the experimental approach. Section III presents the result and discussion of this study followed by conclusion in Section IV.

2. APPROACH

This study consists of four main steps, i.e. removal of optic disc, candidate exudate detection, features extraction and classification as shown in Figure 2.

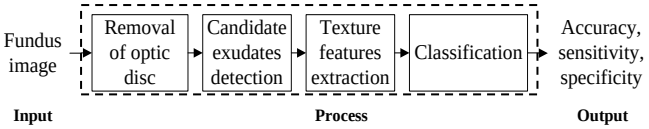

Figure 2. Block diagram of the proposed scheme

For the first two steps, i.e. removal of optic disc and candidate exudates detection, we referred to our proposed method in [15]. Removal of optic disc was processed based on radius enlargement of initial OD obtained. The L-band of HSL colour model was used to facilitate the process accompanied by contrast enhancement and filtering stages. Detection of exudate candidates was conducted by applying Butterworth high pass filtering (BHPF). Since exudates were spotted as the bright lesion, BHPF could preserve the important information of exudates by passing the high frequency and suppressing the low frequency. Afterward, multiplying the resulted images from the step 1 and step 2 eliminated the optic disc that was remaining in the detected candidates of exudates. Thus, the final candidates of exudates were obtained.

Furthermore, some texture features were extracted from detected candidates of exudates based on GLCM, GLRM and histogram. GLCM calculated textures in the second order by considering the relation of pixel adjacent. The four directions were used for GLCM, including 0, 45, 90 and 135 degrees. Five extracted GLCM features were contrast, correlation, inverse different moment (IDM), angular second moment (ASM) and entropy [13].

Angular second moment (ASM) was used to calculate the homogeneity of image using Equation (1) with the number of levels for computation expressed as L.

$$ASM = \sum_{i=1}^{L}\sum_{j=1}^{L} GLCM(i,j)^2 \quad (1)$$

The measurement of variation of grey level pixels image known as contrast is formulated in Equation (2). Whilst, inverse different moment (IDM) was used to measure homogeneity as formulated in Equation (3).

$$contrast = \sum_{n=1}^{L} n^2 \left\{ \sum_{|i-j|=n} GLCM(i,j) \right\} \quad (2)$$

$$IDM = \sum_{i=1}^{L}\sum_{j=1}^{L} \frac{GLCM(i,j)}{1+(i-j)^2} \quad (3)$$

Entropy describes the irregularity of grey level image. If elements of GLCM are relative the same, high entropy value would be obtained. Low entropy value is achieved if the elements of GLCM near 0 or 1. Correlation was used to measure the linear dependence of grey level value of image. Entropy and correlation are denoted in Equation (4) and Equation (5).

$$entropy = -\sum_{i=1}^{L}\sum_{j=1}^{L} GLCM(i,j)\log(GLCM(i,j)) \quad (4)$$

$$correlation = \frac{\sum_{i=1}^{L}\sum_{j=1}^{L}(ij)(GLCM(i,j)) - \mu_i'\mu_j'}{\sigma_i'\sigma_2'} \quad (5)$$

GLRLM is a set of pixels sequence on the same grey level value. Run represents how many times the value of grey level occurs sequentially in a particular direction. Run-length technique is an extraction feature technique by using a statistical approach. This method used pixel distribution with the same intensity in sequence within a particular direction as its primitive. Each of primitive was defined as direction, length and grey level. The length of texture primitive in different direction was used to describe a texture. Similar to GLCM, GLRM also uses 4 directions in its process [16].

A total of 11 GLRLM features were extracted comprising of short run emphasis (SRE), long run emphasis (LRE), grey-level non-uniformity (GLN), run length non-uniformity (RLN), run percentage (RP), low grey-level run emphasis (LGRE), high grey-level run emphasis (HGRE), short run low grey-level emphasis (SRLGE), short run high grey-level emphasis (SRHGE), long run low grey-level emphasis (LRLGE) and long run high grey-level emphasis (LRHGE).

Short run emphasis (SRE) divides each run-length value by the square of its length with the total number of runs in an image as a denominator. It is calculated using Equation (6). Each value of LRE is multiplied to the square of its length to allow a higher

weight to the long runs that are calculated using Equation (7). For run percentage (RP), a ratio is between the total numbers of observed runs in image and the total number of possible runs if all runs have a length of one with A as an image area. Equation (8) was used to calculate run percentage. For GLN, grey level non-uniformity feature has its lowest value if the runs are evenly distributed over all grey levels for being affected by high run length values. RLN, on the other hand, will be the lowest if the runs are evenly distributed over all runs length. GLN and RLN are formulated mathematically in Equation (9) and Equation (10), respectively.

$$SRE = \sum_i \sum_j \frac{P_\theta(i,j)}{j^2} / \sum_i \sum_j P_\theta(i,j) \quad (6)$$

$$LRE = \sum_i \sum_j j^2 P_\theta(i,j) / \sum_i \sum_j P_\theta(i,j) \quad (7)$$

$$RP = \sum_i \sum_j P_\theta(i,j)/A \quad (8)$$

$$GLN = \sum_i \left\{ \sum_j P_\theta(i,j) \right\}^2 / \sum_i \sum_j P_\theta(i,j) \quad (9)$$

$$RLN = \sum_j \left\{ \sum_i P_\theta(i,j) \right\}^2 / \sum_i \sum_j P_\theta(i,j) \quad (10)$$

Other obtained GLRLM features include low grey-level run emphasis (LGRE), high grey-level run emphasis (HGRE), short run low grey-level emphasis (SRLGE), short run high grey-level emphasis (SRHGE), long run low grey-level emphasis (LRLGE) and long run high grey-level emphasis (LRHGE). They are calculated using formula from Equation (11) to Equation (16) as follows.

$$LGRE = \sum_i \sum_j \frac{P_\theta(i,j)}{i^2} / \sum_i \sum_j P_\theta(i,j) \quad (11)$$

$$HGRE = \sum_i \sum_j i^2 P_\theta(i,j) / \sum_i \sum_j P_\theta(i,j) \quad (12)$$

$$SRLGE = \sum_i \sum_j \frac{P_\theta(i,j)}{i^2 . j^2} / \sum_i \sum_j P_\theta(i,j) \quad (13)$$

$$SRHGE = \sum_i \sum_j \frac{i^2 . P_\theta(i,j)}{j^2} / \sum_i \sum_j P_\theta(i,j) \quad (14)$$

$$LRLGE = \sum_i \sum_j \frac{j^2 . P_\theta(i,j)}{i^2} / \sum_i \sum_j P_\theta(i,j) \quad (15)$$

$$LRHGE = \sum_i \sum_j i^2 . j^2 . P_\theta(i,j) / \sum_i \sum_j P_\theta(i,j) \quad (16)$$

For histogram-based texture, there are six features extracted such as mean, standard deviation, skewness, energy, entropy and smoothness as formulated in Equation (17) to Equation (22).

$$m = \sum_{i=0}^{L-1} i \, . p(i) \quad (17)$$

$$\sigma = \sqrt{\sum_{i=1}^{L-1} (i-m)^2 p(i)} \quad (18)$$

$$skewness = \sum_{i=1}^{L-1} (i-m)^3 p(i) \quad (19)$$

$$energy = \sum_{i=0}^{L-1} [p(i)]^2 \quad (20)$$

$$entropy = -\sum_{i=0}^{L-1} p(i) \log_2(p(i)) \quad (21)$$

$$R = 1 - \frac{1}{1+\sigma^2} \quad (22)$$

A primary step after feature extraction is classification. These extracted features were classified using MLP to categorise fundus image whether containing exudates or not. Generally, data are presented at the input layer and then are processed in the network by multiplying it to the weight layer. Afterwards, the multiplication result was processed by output layer of nodes using a function to determine whether the output node fires. The error values were back propagated from the output nodes to the input nodes via the hidden nodes [17].

Some statistical parameters involved in evaluating the proposed scheme were accuracy, sensitivity and specificity which were mathematically formulated from Equation (23) to Equation (25).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (23)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (24)$$

$$Specificity = \frac{TN}{TN + FP} \quad (25)$$

3. RESULT AND DISCUSSION

A total of 80 colour fundus images collected from DIARETDB1 were used in this study. The dataset consists of 38 exudates images and 42 non-exudates images in PNG (RGB) format with the resolution of 1500 x 1152 pixels. DIARETDB1 provides the ground truth image marked by four experts.

3.1 Candidate Exudates Detection

The characteristics of OD and exudates tend to be similar. Hence, removal of optic disc from the fundus image was the best way to reduce false positive area of detected exudate candidate. L-band was selected as the input image for removal of optic disc since it contained sufficient information to detect optic disc. Some processes were employed to obtain initial OD area. Then, optic disc was detected based on radius enlargement of initial OD obtained [15] as described in Figure 3.

(a) (b)

Figure 3. (a) Initial OD (b) The result of radius enlargement [15]

The green channel contained useful information for exudates detection. BHPF was applied to preserve the high frequency in the extracted green channel as shown in Figure 4 (a). The results of the detected OD and detected exudate candidate areas were

then multiplied to remove OD. Subsequently, the final candidates of exudates [13] were obtained as illustrated in Figure 4 (b).

(a) (b)

Figure 4. (a) The result of BHPF (b) Final candidate of exudates [15]

3.2 Features Extraction and Classification

Feature extraction process obtained five features based on GLCM texture feature (i.e. ASM, contrast, IDM, entropy and correlation), eleven features based on GLRLM texture feature (i.e. SRE, LRE, LP, GLN, RLN, LGRE, HGRE, SRLGE, SRHGE, LRLGE, LRHGE) and six histogram-based features (i.e. mean, standard deviation, skewness, energy, entropy and smoothness). These extracted features were then used in the classification process using MLP classifier. The five-folds cross validation was used to optimise the performance of MLP classifier.

The confusion matrix is a useful tool to analyse how well a classifier can recognise the instance of different classes. The number of TP, TN, FP and FN of confusion matrix affects the evaluation results. Accuracy expresses the successful rate of classification process. Sensitivity is a capability of the classifier to predict positive class as positive while specificity is a capability of the classifier to predict negative class as negative.

In this work, seven types of classification based on extracted features were conducted. They were GLCM, GLRLM, histogram and the combination of these three kinds of features. Comparison of classification results of these features is presented in Table1.

Table 1. Comparison evaluation results of classification

Features	Accuracy (%)	Sensitivity (%)	Specificity (%)
GLCM (5)	**95**	**97.37**	**92.86**
GLRLM (11)	91.25	86.84	95.23
Histogram (6)	90	86.84	92.86
GLCM+GLRLM (16)	92.50	92.10	92.86
GLCM+Histogram (11)	95	97.37	92.86
GLRLM+Histogram (17)	90	86.84	92.86
GLCM+GLRLM +Histogram (22)	93.75	94.73	92.86

As depicted in Table 1, five GLCM features yield the best classification rate with accuracy, sensitivity and specificity of 95%, 97.37% and 92.86%, respectively. While 11 GLRM features obtain the low sensitivity rate at 86.84%. The six features of histogram-based do not have significant effect of classification either in self or combination with other features. The combination of GLCM features and histogram features obtains the same of evaluation rate with GLCM features in self. For the combination of GLRLM features and histogram features, the resulted evaluation rate is similar to histogram features in self. Moreover, the combination of histogram with GLCM and GLRM features even reduces the rate of evaluation result compared that of five GLCM features.

The high sensitivity rate obtained by five GLCM features indicates that the proposed scheme is able to classify fundus image containing exudates by recognising the texture characteristics of the detected candidates of exudates. A total of 37 images that contain exudates were classified as exudates image (TP) and merely one image was classified as non-exudates (FN). Forty-two images did not contain exudates, 39 images were classified as non-exudates image (TN) while 3 images were classified as exudates image (FP).

4. CONCLUSION

This study proposed a scheme to classify fundus images whether containing exudates or not based on the analysis of extracted texture features. The three kinds of texture feature comprise of 5 GLCM features, 11 GLRLM features and 6 histogram-based features. In addition, there are three types of classification conducted by using these extracted features either in self or the combination of each other.

The best evaluation rate is achieved by using five features of GLCM with the accuracy of 95%, sensitivity of 97.37% and specificity of 92.86%. This result indicates that the proposed scheme is not only successful to recognise and to detect exudates but also is able to classify fundus images either containing exudates or containing no exudates. Moreover, this achievement denotes that the performance of proposed scheme has a potential to be implemented in developing a computerised aided retinopathy diagnosis system for assisting the ophthalmologists.

5. ACKNOWLEDGMENTS

The authors would like to thank the Rapid Assessment Diabetic Retinopathy research team and the Intelligence Systems research group for sharing meaningful knowledge and inspiring discussion.. This study is financially supported by the Directorate General of Higher Education, Ministry of Research, Technology and Higher Education, Republic of Indonesia through the Research Grant "Penelitian Unggulan Perguruan Tinggi" Universitas Gadjah Mada, No. 2365/UN1.P.III/DIT-LIT/LT/2017.

6. REFERENCES

[1] Federation, I. D. 2015. The IDF diabetes atlas. 7th ed. International Diabetes Federation. Brussels.

[2] SC, P. K. and Devaraj, M. D. 2013. Automatic Exudate Detection For The Diagnosis Of Diabetic Retinopathy. *Int. J. of Innovative Research & Studies*. 2, 5, (May, 2013), 658-669.

[3] Mookiah, M. R. K., Acharya, U. R., Chua, C. K., Lim, C. M., Ng, E. and Laude, A. 2013. Computer-aided diagnosis of diabetic retinopathy: A review. *Computers in biology and medicine*. 43, 12, 2136-2155.

[4] Kande, G. B., Subbaiah, P. V. and Savithri, T. S. 2009. Feature extraction in digital fundus images. *J. Med. Biol. Eng*. 29, 3, (Jan, 2009), 122-130.

[5] Sathiyasekar, K., Karthick, S. and Priyadharsini, A. 2014. A Survey on Hard Exudates Detection and Segmentation. *International Journal of Scientific Engineering and Technology*. 3, 2, 154-158.

[6] Benalcázar, M., Brun, M. and Ballarin, V. 2013. Automatic Segmentation of Exudates in Ocular Images using Ensembles

of Aperture Filters and Logistic Regression. *Journal of Physics: Conference Series*. 477, 1, 012021.

[7] Fadzil, M. A., Nugroho, H., Venkatachalam, P., Nugroho, H. and Izhar, L. 2008. Determination of retinal pigments from fundus images using independent component analysis. In *Proceedings of the 4th Kuala Lumpur International Conference on Biomedical Engineering*. Springer, 555-558.

[8] Tomi Kauppi, V. K., Joni-Kristian Kamarainen, Lasse Lensu, Iiris Sorri, Asta Raninen, Raija Voutilainen, Juhani Pietilä, Heikki Kälviäinen, and Hannu Uusitalo. 2007. DIARETDB1 - Standard Diabetic Retinopathy Database Calibration level 1.

[9] Eadgahi, M. G. F. and Pourreza, H. 2012. Localization of hard exudates in retinal fundus image by mathematical morphology operations. In *Proceedings of the Computer and Knowledge Engineering (ICCKE), 2012 2nd International eConference on*. IEEE, 185-189.

[10] Aftab, U. and Akram, M. U. 2012. Automated identification of exudates for detection of macular edema. In *Proceedings of the Biomedical Engineering Conference (CIBEC), 2012 Cairo International*. IEEE, 27-30.

[11] García, M., Sánchez, C. I., Poza, J., López, M. I. and Hornero, R. 2009. Detection of hard exudates in retinal images using a radial basis function classifier. *Annals of biomedical engineering*. 37, 7, 1448-1463.

[12] García, M., Valverde, C., López, M. I., Poza, J. and Hornero, R. 2013. Comparison of logistic regression and neural network classifiers in the detection of hard exudates in retinal images. In *Proceedings of the Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE, 5891-5894.

[13] Nugroho, H. A., Oktoeberza, K., Adji, T. B. and Najamuddin, F. 2015. Detection of Exudates on Color Fundus Images using Texture Based Feature Extraction. *International Journal of Technology*. 6, 2, 121-129.

[14] Nugroho, H. A., Faisal, N., Soesanti, I. and Choridah, L. 2014. Identification of malignant masses on digital mammogram images based on texture feature and correlation based feature selection. In *Proceedings of the Information Technology and Electrical Engineering (ICITEE), 2014 6th International Conference on*. IEEE, Yogyakarta, Indonesia, 1-6.

[15] Nugroho, H. A., Oktoeberza, K. W., Adji, T. B. and Sasongko, M. B. 2015. Segmentation of exudates based on high pass filtering in retinal fundus images. In *Proceedings of the 7th International Conference on Information Technology and Electrical Engineering (ICITEE)* (Chiang Mai, Thailand, October 29-30, 2015). IEEE, Thailand, 436-441.

[16] Wibawanto, H., Susanto, A., Widodo, T. S. and Tjokronegoro, S. M. 2010. Discriminating cystic and non cystic mass using GLCM and GLRLM-based texture features. *Int J Electron Eng Res*. 2, 569-580.

[17] Akrimi, J. A., RahimAhmad, A., George, L. E. and Aziz, S. 2013. Review of Artificial Intelligence. *Int. J. Sci. Res*. 2, 2, 487-505.