Diabetic Retinopathy Grade Classification based on Fractal Analysis and Random Forest

Farrikh Alzami[1], Abdussalam, Rama Arya Megantara, Ahmad Zainul Fanani, Purwanto
Dian Nuswantoro University
alzami@dsn.dinus.ac.id[1*]

Abstract—**Glaucoma which part of Diabetic Retinopathy is the disease which distorted the optical nerve system and resulted loss in vision. Fractal dimension is one of feature extraction that can been used in retinopathy fields due to the fractal dimension able to characterize the retinal vasculature. In this paper, we presented a research based on fractal dimension which not only distinguish the healthy subjects and diabetic retinopathy patients, but also severe level of diabetic retinopathy patients. By using MESSIDOR dataset and Random Forest as Classifier, we obtained the results that fractal dimensions are able to distinguish the healthy subjects and diabetic retinopathy patients, but it did not obtain satisfactory results for classifying the severity of diabetic retinopathy patients (grade level). Thus, future directions which need to be explore is the other features such as univariate, multivariate and other statistical features. It also needs to pay attention to red lesion detection to gain more information about diabetic retinopathy grade level.**

Keywords—*glaucoma, diabetic retinopathy, random forest, fractal dimension*

I. INTRODUCTION

Glaucoma is the disease which distorted the optical nerve system and resulted impairment in vision[1]. The glaucoma itself is estimated affected at least 2% percent in world population at 2020[2]. The bright region in fundus eye image which composed from optic cup and optic disk is called optic nerve head region. Glaucoma indeed is part of diabetic retinopathy symptoms among cataract at the retina and diabetic macular edema. The problem of glaucoma is the progression itself is asymptomatic in early stages but resulted in gradually loss vision. Hence, it is crucial to have the detection and classification system for diabetic retinopathy disease especially glaucoma.

Diabetic Retinopathy can be detected by several characteristic, such as[3]: 1) Microaneurysms (MA), which deformations in walls of blood vessels and can be seen as balloon shaped; 2) Hemorrhages (HMA) which resulted from leakage of blood from the damaged capillaries; 3) Exudates, which resulted from blood leakage from veins.

Fractal objects is commonly used in biomedical field[4][5][6]. Fractal objects are the structures where the pattern persists at multiple scale. Moreover, the retinal vasculature could be characterized using fractal dimension[7]. Thus, Fractal dimensions is common features which used for pathological in retinal vasculature. Such as: Aliahmad *et.al*[8] and Mudigonda *et.al* [9] reported that there is a distinction between healthy subject and diabetic patients which can be seen from the fractal dimensions. Even though fractal dimensions reported able to distinct the healthy subjects and diabetic patients, it has lack report about fractal dimension is able to distinct the diabetic retinopathy grades, especially from grade 1 to grade 3 which is mild to severe.

In this paper, we presented diabetic retinopathy classification methods which classifying the healthy eyes and diseased pathologies eyes using fractal analysis as feature extraction and random forest is used as classifier. We initially segmented the images, then we computed the fractal dimension as features. After we obtained the features, we fed those features into random forest classifier. The experiment used MESSIDOR dataset which contains normal, mild (only Microaneurysms is found), moderate (beside Microaneurysms MA, the Hemorrhages HMA is also found even those HMA and MA is less than 20 in each quadrant), and severe (which more than 20 for Hemorrhages HMA in each quadrant).

II. RELATED WORKS

A. Diabetic Retinopathy

Diabetic patients are more likely suffer of neuropathy, ventricular arrhythmia sudden cardiac death and stroke, and vision loss compared with other people. Diabetic retinopathy refers to collection of disorder which caused by the damage the blood vessel of retina. Early stages of diabetic retinopathy are called Non proliferative DR (NPDR) which affect the nutrition supply for the sensitive-light layer to retina. NPDR itself did not cause blindness, but if not cared enough and got quick treatment, the NPDR comes to the next stages, called Proliferative DR (PDR) which caused blindness and the effect is not reversible[10]. NPDR itself contains several stages, such as mild, moderate and severe. When NPDR becomes severe, the retinal sensory layers will ask to the brain to create new blood vessels which called neovascularization. Most of the patients develop changes of blood vessel in optic disk.

B. Fractal Dimension as features for Diabetic Retinopathy

Many researchers using fractal dimension as their main experiment for diabetic retinopathy detection and classification. Thus, many of them giving remarkable conclusion, which will describe as follows.

For diabetic type 1, fractal dimensions are successful to detect early diabetic microvascular. Cheung *et.al*[11] using retinal fractal dimension to uncover geometric complexity of the retinal vasculature. Cheung stated that Euclidean geometric

alone is not sufficient to describe the retinal vascular tree. Thus, by using fractal dimension, concluded that increased the retinal fractal dimension have strong correlation with early retinopathy signs for diabetic type 1. The experiment itself used for early diabetic microvascular damage, and did not assessed the mild, moderate and severe diabetic retinopathy. Grauslund *et.al*[12] using fractal dimension to uncover the geometric pattern and complexity of the retinal vessels for diabetic type 1. Grauslund concluded that fractal dimension can be used as retinal vasculature measurement for diabetic type 1, thus diabetic type 2 did not explored.

Nagaoka and Yoshida[13] investigated the relationship for early-staged diabetic type 2 with fractal dimension. Resulted there is no correlation between fractal dimension and any retinal circulatory parameters from the retinal arteries. But Nagaoka and Yoshida found correlation between fractal dimension with vessel in retinal venas. Nagaoka and Yoshida did not mention about diabetic retinopathy grades detection and classification. Huang *et.al*[14] using messidor, drive and 5 camera dataset to find the correlation between fractal dimension with diabetic retinopathy. By using segmentation method such as: Frangi's Vesselness, Zhang's Method, Soares' Segmentation and BIMSO Method, concluded that fractal dimension is not robust enough to be used.

From literature reviews above, our research direction that we presented in this paper is: we are not only using vessels segmentation, but also using skeletonization segmentations. Then we also used the dataset to utilized the robustness of fractal dimension to characterized the diabetic retinopathy grades. Finally, the random forest classifier is used to classify the diabetic retinopathy grades. The summarized research can be seen as follows:
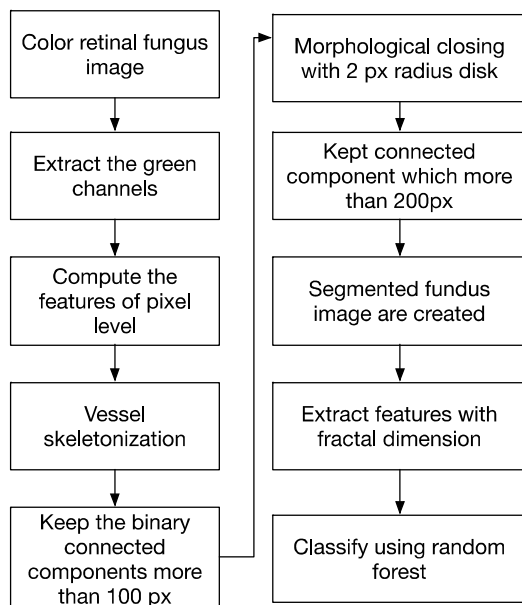
Figure 1: schematic research representation

III. METHODS

A. Segmentation

For segmentation, we follow the description from Orlando, *et.al* [15] to segment the images. In here we only summarized the approach, for the detailed explanation, please refer to the respected paper. In the preprocessing step, the green channel is extracted from RGB images due to the channel is able to provide reliable contrast between structures on the fundus images. Then features of pixel level is computed. In the postprocessing step, we only keep the binary connected components more than 100 pixels. To remove holes in the main arteries due to the central reflex and connect the isolated elements with the vascular structure, a morphological closing with a 2 pixels radius disk structuring element is applied. Then, all connected components which more than 200 pixels is kept for preserve the pathological areas. It is worth mentions, the segmentation used here is used for vessel segmentation and skeletonization segmentations. From figure 2, it can be seen that figure (a) and (c) is the fundus images, then by using segmentation methods, we can obtain the detailed vessel and skeletonization of fundus images as figure (b) and (d).
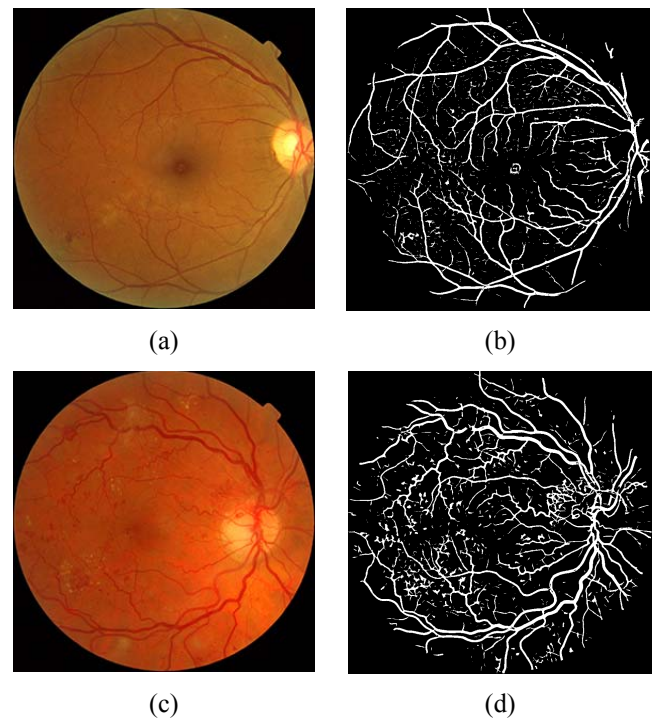
(a) (b)

(c) (d)

Figure 2: (a) the healthy subjects fundus image (b) healthy subjects segmented images (c) diabetic retinopathy patient fundus image (d) diabetic retinopathy patient segmented images

B. Fractal Characteristics

Fractal dimension are the metrics that count the similarity of fractal objects, which can be seen as:

$$D = -\lim_{r \to 0} \frac{\log N(r)}{\log r} \qquad (1)$$

Where D is fractal dimension, N(r) is the count for the pattern at given scale r. the parameter r can be seen as side length of the objects. in the literature, there exists several fractal dimension measurements, such as: 1) box-fractal dimension; 2) information-fractal dimension; 3) correlation-fractal dimension.

For the box-fractal dimension, the equation can be seen as:

$$D_{box} = -\lim_{r \to 0} \frac{\log N(r)}{\log \left(\frac{1}{r}\right)} \tag{2}$$

Then, information-fractal dimension can be seen as:

$$D_{inf} = -\lim_{r \to 0} \frac{\log Sh(r)}{\log \left(\frac{1}{r}\right)} \tag{3}$$

$$Sh(r) = -\sum_{i=1}^{n_r} p_i(r) \log p_i(r) \tag{4}$$

$$p_i(r) = \frac{q_i(r)}{M} \tag{5}$$

Where n_r is the total of cell of side-length r in the objects (image), $p_i(r)$ represents the probabilities obtaining part of binary objects in the i-th cell of side-length r, $q_i(r)$ is total pixel of objects which contained in the i-th cell, them M is the total pixel in the objects (images).

Finally, correlation dimension can be seen as:

$$D_{cor} = -\lim_{r \to 0} \frac{\log Corr(r)}{\log \left(\frac{1}{r}\right)} \tag{6}$$

$$Corr(r) = \frac{i}{n_r^2} \sum_{i=1, i=1, i \neq i}^{n_r} \theta(r - \|p_i - p_j\|) \approx \sum_{i=1}^{n_r} p_i^2(r) \tag{7}$$

Where θ is the heaviside step function, p_i is the position of the i-th pixel which belong to fractal objects.

C. Reliability measurements

In here, we are using accuracy as measurements, which can be obtained by using confusion matrix as follows:

Table 1: confusion matrix

	Total population	True Condition	
		Positive condition	Negative condition
Prediction condition	Positive prediction condition	True positive (TP)	False positive (FP)
	Negative prediction condition	False negative (FN)	True negative (TN)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

IV. RESULTS

We conduct our research in images collected from MESSIDOR[16]. In the MESSIDOR dataset, the Diabetic retinopathy is divided into 5 grades, but in here we only used 4 grades, as follows:

Table 2: diabetic retinopathy grades in MESSIDOR

Grade	Description	Number of images
G0	MA=0 & HMA=0	546
G1	(0<MA≤5) & HMA = 0	153
G2	(5<MA<15) & (0<HMA<5)	247
G3	(MA≥15) or (HMA≥5)	254

Where G0 can be seen as healthy subjects, then G1 to G3 can be seen as diabetic retinopathy subjects with illness degrees.

Those images we are segmented according to previous section, then the segmented images are processed using fractal dimension to obtain features. Those features then fed into random forest. We applied k-fold validation to generalize the findings. Thus, the detailed parameter that we used as follows:

Table 3: list of parameters used

Parameters	Value
k-fold validation	5
Number repetition	3
Random forest: number-estimator	200 | 500
Random forest: max depth	5 | 8
Random forest: criterion	Entropy| Gini
Random forest: max features	auto

Then, we reported the features which we are using, such as: 1) using box fractal dimension only; 2) using information fractal dimension only; 3) using correlation fractal dimension only; 4) and combined box, information and correlation together.

Table 4: experiment results using random forest (accuracy)

datasets	Box	Info	Corr	combined
G0-G1	80.37±0.2	80.37±0.2	80.37±0.2	80.37±0.2
G0-G1-G2	55.8±0.17	55.8±0.17	55.8±0.17	55.8±0.17
G0-G1-G2-G3	48.27±0.07	48.27±0.07	48.27±0.07	48.27±0.07

Table 5: experiment to obtain grade level using random forest (accuracy)

Datasets	Box	Info	Corr	combined
G1-G2	62.37±0.13	62.37±0.13	62.37±0.13	62.37±0.13
G1-G2-G3	40.77±0.02	41.45±0.10	41.45±0.10	41.45±0.10

A. Effect of Fractal Dimension for DR classification

From the table 4, we can see, there is no difference using either box fractal dimension, information fractal dimension, correlation fractal dimension, or combination of those three. The possible explanation is those three features exhibits same characteristics which even using one of those features alone is enough.

From table 4, we also can conclude, that the fractal dimension is able to distinguish between healthy subjects and diabetic retinopathy patients, but in table 5, it failed when it used to distinguish the diabetic retinopathy grades. It means, fractal dimension alone is not sufficient to distinguish those diabetic retinopathy grade level from G1 to G3.

B. Discussion

From effects section, we obtained several findings, that fractal dimension which have several measurements, such as box fractal dimension, information fractal dimension, correlation fractal dimension. Those fractal dimension exhibits same characteristics, thus if we want use the fractal dimension, only used one is sufficient.

Due to fractal dimension alone is not able to distinguish the diabetic retinopathy grade level from G1 to G3, it means we need to face another direction to other features, such as statistical features, univariate and multivariate features.

Then, in side of classification, we are using optimized random forest as seen from the previous section. Thus, we need to explore which classifier is more robust to handle diabetic retinopathy grade level classification.

V. CONCLUSION AND FUTURE WORKS

In this paper, we presented the fractal analysis for diabetic retinopathy grade classification. The experiment concludes that fractal analysis able to distinguish healthy subjects and diabetic retinopathy patients, but it failed to distinguish mild diabetic retinopathy to severe diabetic retinopathy. Thus, in next research, we will analyze another feature, such as univariate and multivariate or statistical features. Then, we will move the direction to red lesion detection to extract more features from it. It worth to mention, we believed the segmentation is done and served the purpose, but we will seek another direction to segment the retinal fundus image which could give more information about grade level.

REFERENCES

[1] J. Zilly, J. M. Buhmann, and D. Mahapatra, "Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation," *Comput. Med. Imaging Graph.*, vol. 55, pp. 28–41, 2017.

[2] A. P. H. Organization, "VISION 2020: Global Campaign to Eliminate Avoidable Blindness," 2011. .

[3] M. R. K. Mookiah, U. R. Acharya, C. K. Chua, C. M. Lim, E. Y. K. Ng, and A. Laude, "Computer-aided diagnosis of diabetic retinopathy: A review," *Comput. Biol. Med.*, vol. 43, no. 12, pp. 2136–2155, 2013.

[4] F. Alzami *et al.*, "Adaptive Hybrid Feature Selection-Based Classifier Ensemble for Epileptic Seizure Classification," *IEEE Access*, vol. 6, pp. 29132–29145, 2018.

[5] E. Abdulhay, E. V., C. M., B. V.S., and N. K., "Automated diagnosis of epilepsy from EEG signals using ensemble learning approach," *Pattern Recognit. Lett.*, May 2017.

[6] A. Chan and J. A. Tuszynski, "Automatic prediction of tumour malignancy in breast cancer with fractal dimension," *R. Soc. Open Sci.*, vol. 3, no. 12, p. 160558, Dec. 2016.

[7] B. R. Masters, "Fractal Analysis of the Vascular Tree in the Human Retina," *Annu. Rev. Biomed. Eng.*, vol. 6, no. 1, pp. 427–452, Aug. 2004.

[8] A. B., K. D.K., S. M.G., and J. R., "Relationship between diabetes and grayscale fractal dimensions of retinal vasculature in the Indian population," *BMC Ophthalmol.*, vol. 14, p. 152, 2014.

[9] S. Mudigonda, F. Oloumi, K. M. Katta, and R. M. Rangayyan, "Fractal analysis of neovascularization due to diabetic retinopathy in retinal fundus images," *2015 E-Health Bioeng. Conf. EHB 2015*, pp. 1–4, 2016.

[10] K. M. F. M. Stephen J.H., *Parsons' Diseases of the Eye*. Harcourt Publishers Ltd, 1990.

[11] N. Cheung *et al.*, "Quantitative Assessment of Early Diabetic Retinopathy Using Fractal Analysis," *Diabetes Care*, vol. 32, no. 1, pp. 106–110, Jan. 2009.

[12] J. Grauslund, A. Green, R. Kawasaki, L. Hodgson, A. K. Sjølie, and T. Y. Wong, "Retinal Vascular Fractals and Microvascular and Macrovascular Complications in Type 1 Diabetes," *Ophthalmology*, vol. 117, no. 7, pp. 1400–1405, Jul. 2010.

[13] T. Nagaoka and A. Yoshida, "Relationship Between Retinal Fractal Dimensions and Retinal Circulation in Patients with Type 2 Diabetes Mellitus," *Curr. Eye Res.*, vol. 38, no. 11, pp. 1148–1152, Nov. 2013.

[14] B. M. ter Haar Romeny *et al.*, "Reliability of Using Retinal Vascular Fractal Dimension as a Biomarker in the Diabetic Retinopathy Detection," *J. Ophthalmol.*, vol. 2016, pp. 1–13, 2016.

[15] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "Convolutional neural network transfer for automated glaucoma identification," 2017, p.

101600U.

[16] E. Decencière *et al.*, "Feedback on a publicly distributed image database: The Messidor database," *Image Anal. Stereol.*, vol. 33, no. 3, pp. 231–234, 2014.