Artificial intelligence on diabetic retinopathy diagnosis: an automatic classification method based on grey level co-occurrence matrix and naive Bayesian model

Kai Cao, Jie Xu, Wei-Qi Zhao

Beijing Institute of Ophthalmology, Beijing Tongren Hospital of Capital Medical University, Beijing 100005, China
Correspondence to: Kai Cao. Beijing Institute of Ophthalmology, Beijing Tongren Hospital of Capital Medical University, Beijing 100005, China. anzhen602@163.com

Abstract

● **AIM:** To develop an automatic tool on screening diabetic retinopathy (DR) from diabetic patients.

● **METHODS:** We extracted textures from eye fundus images of each diabetes subject using grey level co-occurrence matrix method and trained a Bayesian model based on these textures. The receiver operating characteristic (ROC) curve was used to estimate the sensitivity and specificity of the Bayesian model.

● **RESULTS:** A total of 1000 eyes fundus images from diabetic patients in which 298 eyes were diagnosed as DR by two ophthalmologists. The Bayesian model was trained using four extracted textures including contrast, entropy, angular second moment and correlation using a training dataset. The Bayesian model achieved a sensitivity of 0.949 and a specificity of 0.928 in the validation dataset. The area under the ROC curve was 0.938, and the 10-fold cross validation method showed that the average accuracy rate is 93.5%.

● **CONCLUSION:** Textures extracted by grey level co-occurrence can be useful information for DR diagnosis, and a trained Bayesian model based on these textures can be an effective tool for DR screening among diabetic patients.

● **KEYWORDS:** grey level co-occurrence matrix; Bayesian; textures; artificial intelligence; receiver operating characteristic curve; diabetic retinopathy

INTRODUCTION

Diabetic retinopathy (DR) is a common microvascular complication of diabetes, and is one of the main causes of preventable blindness in population within working age[1-2]. According to an analysis on global causes of blindness, DR ranks the 7th leading reason and takes responsibility for 0.85% of all blindness worldwide[3]. As the life expectancy getting extended, more and more diabetic patients are diagnosed. China alone has 92.4 million diabetic patients[4]. Meanwhile, DR population is up to 15.2 million[2].

Such large crowd of DR patients poses a challenge to ophthalmologists, because reading eye fundus takes lots of time. Doctors are in sore need of an automatic tool to help them do preliminary screening, and they can save time for the intractable cases.

Recently, many researchers cast their eyes over recognition filed, in this field many machine learning or deep learning methods are widely used to make classification of diseases based on images, and neural network algorithm is one the most commonly used method[5-7]. Other algorithms like support vector machine[8], random forest method *etc.* are also have been applied and showed good performance[9].

Naive bayes is also one kind of machine learning algorithm. Unlike other classifiers, naive bayes is a classification algorithm based on probability theory. Naive bayes is simple but efficient, and has been widely used in many fields like cardiovascular disease risk's level detection[10], biomarker selection and classification from genome-wide SNP data[11], diagnosis of Parkinson's disease[12], glaucoma management[13] and so on. What is more, at the early days of DR, symptoms like retinal hard exudates, microaneurysm and cotton version are very few and very hard to identify macroscopically. Some patients may be mis-diagnosed and thus miss the chance to stop the progress. However, we all know that DR could be stopped if treated by laser photocoagulation at the early stage[14]. Under this circumstance, early detection and early diagnosis is quite important. Some researchers even tried to find out some biomarkers of DR to help with early detection[15-17], but compared with auto-detection through eye fundus using artificial intelligence, biomarkers cost much more money and is of low efficiency[18].

In our study, we used grey-level co-occurrence matrix firstly to help extract textures from eye fundus images. Grey-level co-occurrence matrix method is a useful method on analyzing textural features and was originally proposed by Haralick[19], grey-level cooccurrence matrix method has been widely used in many fields[20-22]. Specifically, for DR, textures of symptoms like microaneurysm, exudation *etc.* are quite different from textures of normal images, so we can use the texture information to help make diagnosis. In our study, we applied naive Bayesian method to train a prediction model based on these textures. At last we used an external group of diabetic subjects' eye fundus images to make validation on the trained model.

SUBJECTS AND METHODS

Ethical Approval The study followed the tenets of the Declaration of Helsinki. Oral consent was obtained from each patient, without any financial compensation.

All eye fundus images are from diabetic patients at Outpatient in Beijing Tongren Hospital during 2017 January to 2017 August. Diabetic patients were diagnosed by doctors or they reported a diabetic history and were using oral hypoglycemic drugs. All eye fundus pictures were taken with a Topcon TRC-NW6S/7S (Topcon; www.global.topcon.com) camera, and dilated 45° digital stereoscopic color retinal photographs of Early Treatment DR standard field 1 (centred on the optic disc, stereoscopic) and standard field 2 (centred on the macula, non-stereoscopic) were taken for each eye by trained and certified photographers. The eye fundus images were in a size of 1536 pixel × 1248 pixel. The Bayesian model was trained based on readable fundus images, blurry images due to cataract or other diseases were excluded.

Gray Level Co-occurrence Matrix The gray level co-occurrence matrix (GLCM) is defined as the probability that the gray value at a point away from a certain fixed position. We use P_δ to represent gray level co-occurrence, $P_\delta = L \times L$ (L represents the number of different grey levels in a picture). Thus actually, P_δ is the frequency of occurrences of two pixels (i and j) with spatial positional relationship and gray levels.

For intuitive interpretation, we draw Figure 1. There are 3 grey levels in Figure 1.

Thus we get a 3×3 matrix

$$P_{\delta=} \begin{bmatrix} 0 & 10 & 10 \\ 10 & 0 & 10 \\ 10 & 10 & 0 \end{bmatrix}$$

After normalization, the matrix will be

$$P_{\delta=} \begin{bmatrix} 0 & 1/6 & 1/6 \\ 1/6 & 0 & 1/6 \\ 1/6 & 1/6 & 0 \end{bmatrix}$$

Characteristic Quantities of Matrix The GLCM describes textural features with the spatial relationships among the gray
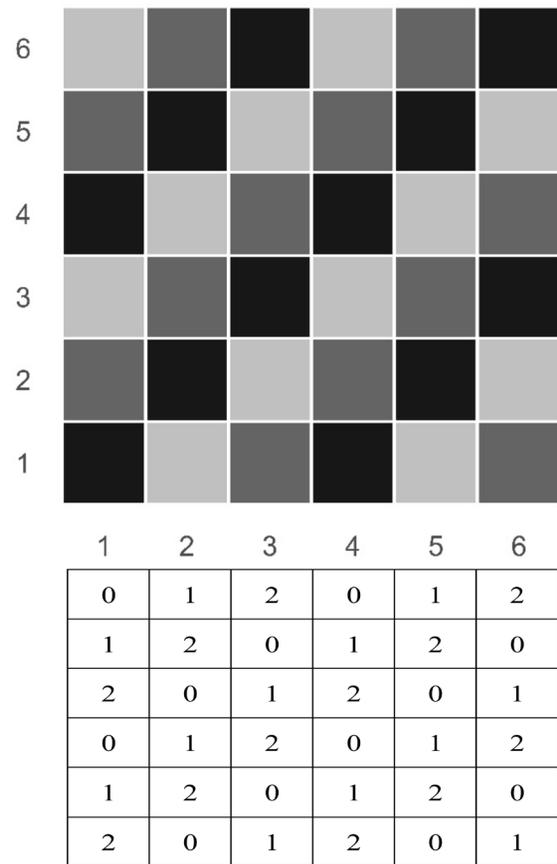
Figure 1 Grey level co-occurrence matrix The number shows the grey level.

levels through examining all pixels within the image space and counting the gray-level dependencies between two pixels separated by a certain interval distance along a particular direction. Haralick *et al*[23] described 14 statistics that can be calculated from the co-occurrence matrix. However, Ulaby *et al*[24] pointed out that only 4 statistics (contrast, angular second moment, entropy and correlation) are uncorrelated. In our study, we used four features to make diagnosis of DR.

Contrast Contrast reflects[23] how the matrix is distributed and to what degree the local change is. It shows the sharpness of the image and the depth of the texture. The deeper the groove from the gray texture is, the greater contrast and the clearer visual effect will be. Conversely, a small contrast comes with a shallow groove and a blurred vision effect.

$$Con = \sum_i^{J-1} \sum_j^{J-1} (i-j)^2 P(i,j) \qquad (1)$$

Angular Second Moment Angular Second Moment (ASM) is also called energy. ASM reflects the uniformity of image gray distribution and the thickness of the texture. If the element values of the GLCM were similar, the energy would be small; if some of the values were large, and other values were small, then the energy value would be large.

$$ASM = \sum_i^{J-1} \sum_j^{J-1} P(i,j)^2 \qquad (2)$$

Entropy The image contains a measure of the randomness of the amount of information. When all the values in the co-occurrence matrix are equal or the pixel values exhibit the
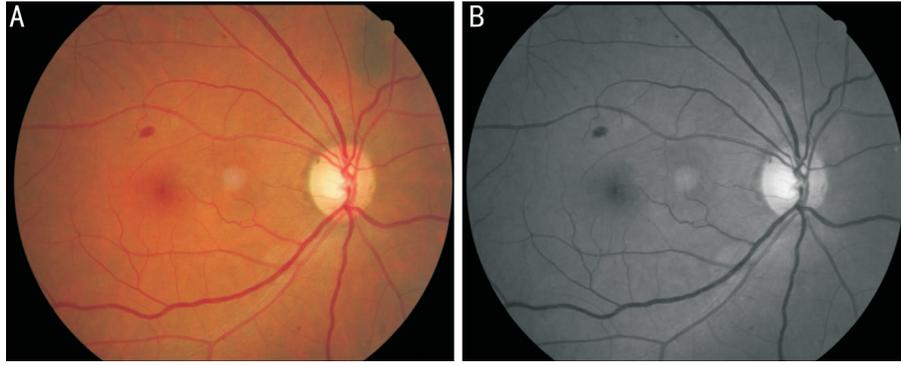
Figure 2 Figure preparation To save time on computation, we compressed the grey levels of images from 256 levels to 16 levels which helped make the matrix smaller, with no influence on the texture extracting. Then we cut each eye fundus image into small pieces by a unit of 128×128 pixels. For each small piece, we extracted four textures from four directions: 0°, 45°, 90°, 135°, and then calculated the mean value of each texture. A: Original eye fundus; B: G band of eye fundus.

greatest randomness, the entropy is the largest. The entropy value indicates the complexity of the gray distribution of the image. The larger the entropy value is, the more complex the image will be[25].

$$Ent = -\sum_i^{l-1}\sum_i^{l-1} P(i,j)lgP(i,j) \qquad (3)$$

Correlation Correlation is used to measure the similarity of the gray level of the image in the row or column direction. So it is worth to reflect the local gray correlation. The larger the value, the greater the correlation.

$$Cor = \frac{\sum_i^{l-1}\sum_i^{l-1}(i-\mu_i)(j-\mu_j)P(_{i,j})}{\sigma_i \cdot \sigma_j} \qquad (4)$$

μ_i and μ_j are respectively the means for the rows and columns of the normalized matrix, and σ_i and σ_j are the corresponding standard deviations.

$$\mu_i = \sum_i^{l-1} i \sum_j^{l-1} p(_{i,j}) \qquad (5)$$

$$\mu_j = \sum_j^{l-1} j \sum_i^{l-1} p(_{i,j}) \qquad (6)$$

$$\sigma_i = \sum_i^{l-1}(i-\mu_i)^2 \sum_j^{l-1} p(_{i,j}) \qquad (7)$$

$$\sigma_j = \sum_j^{l-1}(j-\mu_j)^2 \sum_i^{l-1} p(_{i,j}) \qquad (8)$$

Image Preparation Before extracting the texture feature, we need to convert the multi-band image (RGB image) into a grayscale image. Because the texture feature is a structural feature, thus the texture features obtained from each band is the same. We arbitrarily chose the G band in our study. The image preparation process is shown in Figure 2.

Quality Control All DR patients are diagnosed by two ophthalmologists after training. Although GLCM method doesn't require the ophthalmologists to mark DR-related symptoms on the eye fundus images, the ophthalmologists were still asked to offer a brief description of diagnostic basis. When there is divergence, the eye fundus images will be sent to a third senior ophthalmologist to make final decision. We calculated the kappa coefficient to evaluate the consistency between two ophthalmologists.

Statistical Analysis We use mean value and standard deviation to make statistical description on continuous variables, and we use frequency and percentage to make statistical description on categorical variables. A naive Bayesian model was applied to train prediction model and a 10-fold cross validation method was applied to evaluate the accuracy of the model. We used receiver operating characteristic (ROC) curve to make evaluation on the classification ability of our built model. All the analysis was done with open source R-program (version 3.5.1).

RESULTS

Among all the included 1000 diabetic subjects, there are 298 DR patients and 702 control subjects. Specifically, among DR patients, there were 253 non-proliferative DR cases, 35 proliferative DR cases, and 10 diabetic macular edema cases. Two ophthalmologists misdiagnosed 49 and 25 cases respectively, and the kappa value (Table 1) between two doctors is 0.817 (95%CI: 0.776, 0.857). We think the diagnosis is trustable.

We firstly split all cases into two datasets by a probability of 7:3. The dataset with larger sample size was used to train Bayesian model and the smaller dataset was used to test the Bayesian model. As shown in Table 2, the Bayesian model performed quite good. We got a sensitivity of 0.949 and a specificity of 0.928 in the validation dataset. The area under the ROC curve is 0.938 (Figure 3).

At last, we further used a 10-fold cross validation method to evaluate the accuracy of the Bayesian model, and we got an average accuracy rate of 93.5%.

DISCUSSION

DR has been rising globally. It is estimated that in 1990s DR contributed to 0.85% of all blindness worldwide while in 2015 this percentage reached 1.05%[3].

DR in the early days is reversible while at the late stage, it is irreversible[26-27]. Thus the screening of diabetic population is of great public health meaning.
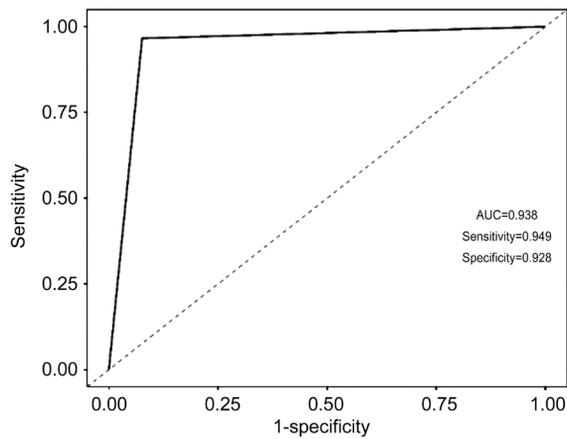
Figure 3 Receiver operating characteristic curve.

Table 1 Agreement on diabetic retinopathy diagnose between two doctors

Doctor 2	Doctor 1		Kappa (95%CI)
	No	Yes	
No	683	42	0.817 (0.776, 0.857)
Yes	32	243	

Table 2 Cross table of predicted diagnosis and real diagnosis

Predicted	Real		Sensitivity	Specificity
	No	Yes		
No	192	15	0.949	0.928
Yes	4	74		

Nowadays, researchers indeed have developed some auxiliary diagnosis system using artificial neural network[28], support vector machine and Gaussian mixture model[29]. But all these methods underutilize the grey level and texture information contained in eye fundus images. In 2012, Acharya et al[30] did apply grey level co-occurrence matrix to extract texture information from fundus photograph and they indeed achieve a good sensitivity of 0.989. However, the specificity is only 0.895. What is more, Acharya et al[30] tried to detect diabetic features namely vessels, exudates, and hemorrhages in fundus images first, and then use these features to diagnose DR. During this process, there must be some information loss, and this process ask ophthalmologists mark each DR feature on the images before training model, which adds too much work to them. It is not good for clinical practice, and I think that explains why only 450 fundus images were used in this study.

In our study, because the textures of DR patients are different from diabetes patients, we use textures of each fundus photograph directly to train our model. There is no information loss and could save lots of ophthalmologists' time. Usually, when we use artificial intelligence to diagnosis DR, a model training process is necessary, and the training of the diagnose model requires the ophthalmologists to mark the symptoms of DR on the eye fundus images firstly. Then the model treat these marks as a reference and learn by itself. For a new eye fundus image, artificial intelligence will try to judge whether there is a similar symptom what it has learned, and then make

diagnosis of DR. The image marking procedure takes a lot of ophthalmologists' time. However, GLCM takes use of disease-related information in another way. GLCM doesn't require any so-call gold-standard mark because GLCM extracts and transforms images' textures into quantitative values. None diagnosis on DR-related symptoms needed to be done. The ophthalmologists only need to spend time on diagnosis of DR. Using GLCM is a much more cost-effective way to diagnose DR. The most important thing is that we got both high sensitivity and specificity.

The limitation of our study is that eye fundus images of low quality were excluded. A small part of patients, like cataract patients, were unable to be screening by GLCM and Bayesian model. Actually, low quality of image is a common barrier for automatic diagnosis. Image enhancement techniques might be a promising method to figure out this problem[31].

Textures extracted by grey level co-occurrence can be useful information for DR diagnosis, and a trained Bayesian model based on these textures could be an effective tool for DR screening among diabetic patients.

REFERENCES

1 Zatic T, Bendelic E, Paduca A, Rabiu M, Corduneanu A, Garaba A, Novac V, Curca C, Sorbala I, Chiaburu A, Verega F, Andronic V, Guzun I, Căpăţină O, Zamă-Mardari I. Rapid assessment of avoidable blindness and diabetic retinopathy in Republic of Moldova. *Br J Ophthalmol* 2015;99(6):832-836.

2 Song PG, Yu JY, Chan KY, Theodoratou E, Rudan I. Prevalence, risk factors and burden of diabetic retinopathy in China: a systematic review and meta-analysis. *J Glob Health* 2018;8(1):010803.

3 Flaxman SR, Bourne RRA, Resnikoff S, Ackland P, Braithwaite T, Cicinelli MV, Das A, Jonas JB, Keeffe J, Kempen JH, Leasher J, Limburg H, Naidoo K, Pesudovs K, Silvester A, Stevens GA, Tahhan N, Wong TY, Taylor HR; Vision Loss Expert Group of the Global Burden of Disease Study. Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta-analysis. *Lancet Glob Health* 2017;5(12):e1221-e1234.

4 Zuo H, Shi ZM, Hussain A. Prevalence, trends and risk factors for the diabetes epidemic in China: a systematic review and meta-analysis. *Diabetes Res Clin Pract* 2014;104(1):63-72.

5 Franklin SW, Rajan SE. An automated retinal imaging method for the early diagnosis of diabetic retinopathy. *Technol Health Care* 2013;21(6):557-569.

6 Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016;316(22):2402-2410.

7 Lam C, Yi D, Guo M, Lindsey T. Automated detection of diabetic retinopathy using deep learning. *AMIA Jt Summits Transl Sci Proc* 2018;2017:147-155.

8 Ganesan K, Martis RJ, Acharya UR, Chua CK, Min LC, Ng EY, Laude A. Computer-aided diabetic retinopathy detection using trace transforms on digital fundus images. *Med Biol Eng Comput* 2014;52(8):663-672.

9 Casanova R, Saldana S, Chew EY, Danis RP, Greven CM, Ambrosius WT. Application of random forests methods to diabetic retinopathy classification analyses. *PLoS One* 2014;9(6):e98587.

10 Miranda E, Irwansyah E, Amelga AY, Maribondang MM, Salim M. Detection of cardiovascular disease risk's level for adults using naive Bayes classifier. *Healthc Inform Res* 2016;22(3):196-205.

11 Sambo F, Trifoglio E, Di Camillo B, Toffolo GM, Cobelli C. Bag of Naïve Bayes: biomarker selection and classification from genome-wide SNP data. *BMC Bioinformatics* 2012;13(Suppl 14):S2.

12 de Souza JWM, Alves SSA, Rebouças ES, Almeida JS, Rebouças Filho PP. A new approach to diagnose parkinson's disease using a structural cooccurrence matrix for a similarity analysis. *Comput Intell Neurosci* 2018;2018:7613282.

13 An GZ, Omodaka K, Tsuda S, Shiga Y, Takada N, Kikawa T, Nakazawa T, Yokota H, Akiba M. Comparison of machine-learning classification models for glaucoma management. *J Healthc Eng* 2018;2018:6874765.

14 Mistry H, Auguste P, Lois N, Waugh N. Diabetic retinopathy and the use of laser photocoagulation: is it cost-effective to treat early? *BMJ Open Ophthalmol* 2017;2(1):e000021.

15 Jenkins AJ, Joglekar MV, Hardikar AA, Keech AC, O'Neal DN, Januszewski AS. Biomarkers in diabetic retinopathy. *Rev Diabet Stud* 2015;12(1-2):159-195.

16 Maghbooli Z, Hossein-nezhad A, Larijani B, Amini M, Keshtkar A. Global DNA methylation as a possible biomarker for diabetic retinopathy. *Diabetes Metab Res Rev* 2015;31(2):183-189.

17 Cunha-Vaz J, Ribeiro L, Lobo C. Phenotypes and biomarkers of diabetic retinopathy. *Prog Retin Eye Res* 2014;41:90-111.

18 Khan AA, Rahmani AH, Aldebasi YH. Diabetic retinopathy: recent updates on different biomarkers and some therapeutic agents. *Curr Diabetes Rev* 2018;14(6):523-533.

19 Haralick RM. Statistical and structural approaches to texture. *Proc IEEE* 1979;67(5):786-804.

20 Song CI, Ryu CH, Choi SH, Roh JL, Nam SY, Kim SY. Quantitative evaluation of vocal-fold mucosal irregularities using GLCM-based texture analysis. *Laryngoscope* 2013;123(11):E45-E50.

21 Chen X, Wei XH, Zhang ZP, Yang RM, Zhu YJ, Jiang XQ. Differentiation of true-progression from pseudoprogression in glioblastoma treated with radiation therapy and concomitant temozolomide by GLCM texture analysis of conventional MRI. *Clin Imaging* 2015;39(5):775-780.

22 Lian MJ, Huang CL. Texture feature extraction of gray-level co-occurrence matrix for metastatic cancer cells using scanned laser pico-projection images. *Lasers Med Sci* 2018.

23 Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst, Man, Cybern* 1973;SMC-3(6):610-621.

24 Ulaby F, Kouyate F, Brisco B, Williams TH. Textural information in SAR images. *IEEE Trans Geosci Remote Sensing* 1986;GE-24(2):235-245.

25 Baraldi A, Parmiggiani F. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. *IEEE Trans Geosci Remote Sensing* 1995;33(2):293-304.

26 Al-Shabrawey M, Zhang WB, McDonald D. Diabetic retinopathy: mechanism, diagnosis, prevention, and treatment. *Biomed Res Int* 2015;2015:854593.

27 Zhang XY, Zhao L, Hambly B, Bao SS, Wang KY. Diabetic retinopathy: reversibility of epigenetic modifications and new therapeutic targets. *Cell Biosci* 2017;7:42.

28 ShuangYu, DiXiao, Kanagasingam Y. Exudate detection for diabetic retinopathy with convolutional neural networks. *Conf Proc IEEE Eng Med Biol Soc* 2017;2017:1744-1747.

29 Roychowdhury S, Koozekanani DD, Parhi KK. DREAM: diabetic retinopathy analysis using machine learning. *IEEE J Biomed Health Inform* 2014;18(5):1717-1728.

30 Acharya UR, Ng EY, Tan JH, Sree SV, Ng KH. An integrated index for the identification of diabetic retinopathy stages using texture parameters. *J Med Syst* 2012;36(3):2011-2020.

31 Khojasteh P, Aliahmad B, Arjunan SP, Kumar DK. Introducing a novel layer in convolutional neural network for automatic identification of diabetic retinopathy. *Conf Proc IEEE Eng Med Biol Soc* 2018;2018: 5938-5941.