Exudate Detection for Diabetic Retinopathy With Convolutional Neural Networks

Shuang Yu, Di Xiao and Yogesan Kanagasingam

***Abstract*— Exudate detection is an essential task for computer-aid diagnosis of diabetic retinopathy (DR), so as to monitor the progress of DR. In this paper, deep convolutional neural network (CNN) is adopted to achieve pixel-wise exudate identification. The CNN model is first trained with expert labeled exudates image patches and then saved as off-line classifier. In order to achieve pixel-level accuracy meanwhile reduce computational time, potential exudate candidate points are first extracted with morphological ultimate opening algorithm. Then the local region (64×64) surrounding the candidate points are forwarded to the trained CNN model for classification / identification. A pixel-wise accuracy of 91.92%, sensitivity of 88.85% and specificity of 96% is achieved with the proposed CNN architecture on the test database.**

***Index Terms*— Deep Learning, Convolutional Neural Networks, Exudate Detection, Retinal Imaging, Diabetic Retinopathy**

I. Introduction

Diabetic retinopathy (DR) is the leading cause of blindness and vision loss for working age population in developed countries [1], [2]. Early detection and yearly screening is necessary to prevent further vision loss for people with diabetes [1]. Considering the rapidly increasing number of diabetic patients, automatic DR grading has the potential to alleviate the workload of ophthalmologist, improve the efficiency and reduce the cost for DR screening [3]–[5]. Exudate is resulted by the breakdown of the blood-retinal barrier, allowing leakage of serum proteins, lipids, and protein from the vessels [2]. It is one of the preliminary clinical signs of DR. Therefore, the accurate and automatic detection of exudate is crucial for the diagnosis of DR. Exudates are presented as white or yellow bright structures in color fundus images, with variable shapes and contrast. A typical DR image with exudate is shown on Fig 1.

The automatic detection of exudate has been extensively investigated with different techniques proposed. Typically, the detection of exudates can be broadly divided into three steps: getting exudate candidates, extracting features and machine learning. Various algorithms have been developed for extracting the exudate candidates, including morphological operation based approaches, e.g. [6], [7], clustering based approaches e.g., [8], [9] and pixel-level feature based machine learning, e.g. [10], [11]. After the exudate candidates are obtained, generally feature extraction and machine learning is used to further classify the candidate points. Zhang *et al.* extracted 28 features including intensity, geometrical and texture related features for all the exudates candidate, then random forest was utilized for the classification [7]. Garcia *et al.* extracted 18 features including color and shape features for the exudate candidates and tested different machine learning methods, including multilayer perceptron, radial basis function and support vector machine for the classification [12].

Shuang Yu, Di Xiao and Yogesan Kanagasingam are with Australian e-Health Research Centre, Commonwealth Scientific and Industrial Research Organization (CSIRO). (email: Shuang.Yu@csiro.au, Di.Xiao@csiro.au, Yogi.Kanagasingam@csiro.au)

Fig. 1. DR color fundus image with exudates presented.

In this paper, in order to improve the accuracy of exudates detection in the pixel level, we combined the image processing steps with deep neural networks. Exudates candidates are obtained with ultimate opening algorithm. Then the local region surrounding the seed points are extracted and forwarded to trained deep convolutional neural networks for classification. Therefore, pixel-level accurate exudate detection is realized.

II. Method and Methodology

The overall framework for exudates detection is illustrated in Fig. 2. The CNN network is trained on the $64 * 64$ patches extracted from the illumination corrected green channel image and then the trained model is stored in computer. For the image processing procedure, the retinal images are first processed to remove the optic disc and blood vessels. Then ultimate opening algorithm is utilized to obtain the potential exudate candidates, i.e. the seed points. At last, the local patches that surrounds the seed points are passed to the trained deep learning model to identify whether they are exudates or not.

Generally, for a pure deep learning procedure, in order to

978-1-5090-2809-2/17/$31.00 ©2017 IEEE

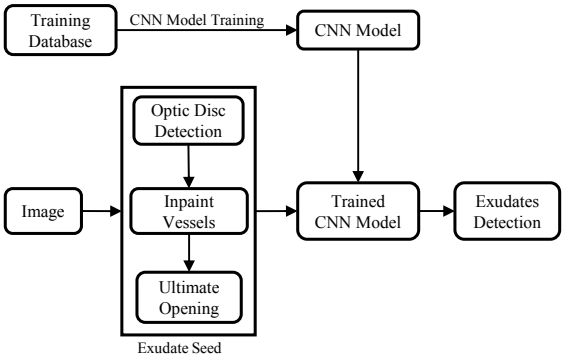


Fig. 2. Framework for exudate detection with deep learning.

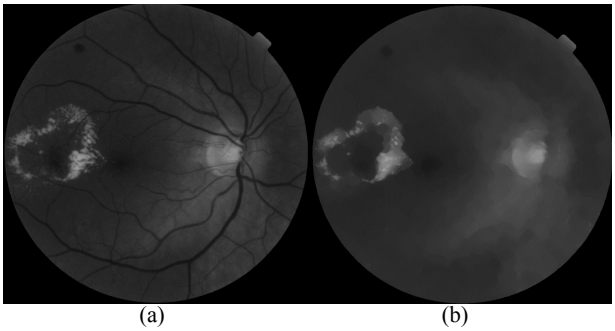


Fig. 3. Removal for retinal vessels and small dark lesions. (a), the green channel color fundus image; (b), the vessel and small dark lesion removed result image.

achieve pixel-level identification, the CNN model needs to be applied on each single pixel, which is a time consuming task if all the pixels inside the fundus image are calculated. Therefore, the advantage of using ultimate opening algorithm is that it will effectively select the potential exudate candidate and remove the points that are background. Depending on the exudates symptoms in the fundus image, ultimate opening will reduce the candidate numbers to 1-20% of the total pixel number in the image, therefore, speeding up the process.

A. Removal of Optic Disc Detection

Since optic disc and exudates are both bright objects and share similar color in color fundus image, as a result, before effective exudates detection, the optic disc shall be detected and masked out. For the detection of optic disc, local phase symmetry algorithm is adopted. Local phase symmetry algorithm is a dimensionless measure of the degree of feature symmetry in the local region [13]. Since optic disc appears as a bright circular or elliptic region on the retinal image, it gives strong response at the center of the optic disc. Then region growing is utilized to obtain the full optic disc region from the maximum response of local phase symmetry. The smallest circle radius that encloses the region growing result is detected as optic disc radius.

B. Removal of Retinal Vessels

Generally, the removal of retinal vessels are performed by vessel segmentation and then inpainting the vessel based on the segmentation result. However, this widely used method suffers from the limitation that the fine vessels can hardly be segmented and removed. Zhang *et al.* proposed to use morphological closing and opening operation, following by the supremum operation to inpaint the dark structures (vessels, microaneurysms, haemorrhages etc) [7]. In this research, we adopted this approach because of its effectiveness in removing the fine vessels and other small dark lesions, meanwhile preserving the bright lesions. Fig. 3(b) shows the inpainting result for the green channel retinal imaging shown on Fig 3(a).

C. Ultimate Opening

After the retinal vessel and small dark lesions are removed, ultimate opening is employed to obtain the exudates candidate points. Ultimate opening, first introduced by Beucher, is a residual operator that highlights patterns with the highest contrast [14], [15]. The opening operations with increasing structure element size are successively applied on the image and the maximum residue is selected as the temporary result. This step is repeated until the maximum given structure element size is reached. The ultimate opening result of the image Fig 3(a) is shown on Fig. 4(a). Region based Otsu algorithm is adopted to calculate the suitable threshold within a given structuring element surrounding the local region, the result of which is shown on Fig 4(b).

However, compared with the original image in Fig. 3(a), the binary image in Fig. 4(b) is not properly reflecting all the potential exudate candidate points. Some of the regions are smaller than the actual exudates and others larger than that. In order to refine the potential candidate result, we first inpaint the slightly dilated binary image (Fig. 4(b)) on the original green channel image, and then get the difference between the inpainted image with the original image, the result of which is shown on Fig. 4(c) [16]. Afterwards, a more accurate exudate candidate image, shown on Fig. 4(d), is obtained by segmenting the difference image.

D. Convolutional Neural Networks

CNN is different from traditional machine learning method in that the features are automatically learned from the training images, instead of expert designed features, as for support vector machine and random forest. It has been proven very effective in image recognition and classification [17], [18]. Typically, a CNN architecture is consisted of three main types of layers: convolutional layer, pooling layer and fully-connected layer. The three types of layers are stacked together to form a full CNN architecture.

Convolutional layer contains a set of learnable filters, the size and number of which are tunable by user, while the weights are automatically learned and optimized from the training procedure. Generally, the first convolutional layers will learn simple features, such as edges, while the deeper

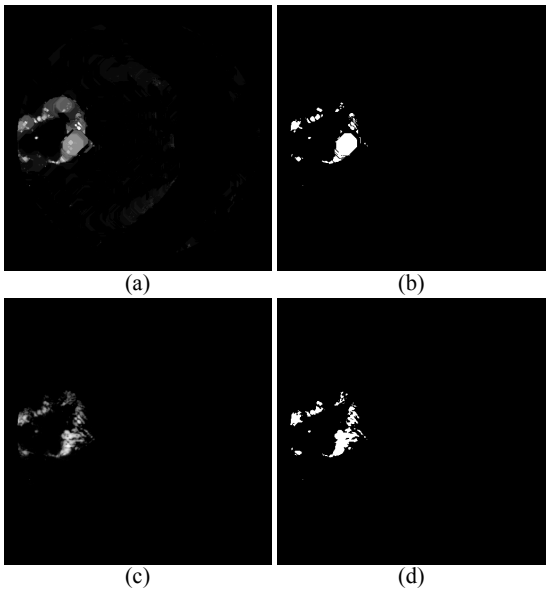


Fig. 4. Obtain exudate candidate points with ultimate opening algorithm. (a), ultimate opening result; (b), segment ultimate opening result with local Otsu thresholding; (c), difference image obtained by subtracting the inpainted image from the green channel image; (d), potential exudate candidates image

layers can learn more abstract, i.e., high-level features. The output of convolutional layer is calculated by summing up all the convolution responses between the input channels with the filters, and then applying an element-wise non-linear activation function, a rectified linear unit in this research, on the sum.

It is necessary to add a pooling layer, usually a max-pooling layer, periodically between successive convolutional layers, so as to progressively reduce the spatial dimension and further reduce the computation time of the network. The max-pooling will select the most prominent response from the convolutional layer in the local region. The tunable parameters in the max-pooling parameter is the stride, i.e., the downsample ratio, usually 2 is used.

After a few stacks of convolutional layers and max-pooling layers, high-level abstract features are extracted with proper dimension. The output of the last max-pooling layer is then passed to fully connected layer for final classification. Fully connected layer, in essence, is a traditional multi-layer perceptron that uses a softmax activation function in the output layer. The purpose of fully connected layer is using the generated high-level features from previous layers to classify the input image into various classes based on the training dataset. In order to regularize the network and avoid over-fitting, a dropout layer is added between the fully-connected layers, which will randomly drop neurons along with their connections during the training and thus preventing neurons from co-adapting too much [19].

The architecture of the CNN network is listed on Table I. The input to the network is 64×64 patches extracted from the illumination corrected green channel image . The training

TABLE I
CNN ARCHITECTURE.

No.	LayerType	Maps	Size	KerSize
0	Input layer	1	64×64	–
1	Convolutional	32	62×62	3×3
2	Convolutional	32	60×60	3×3
3	MaxPool	32	30×30	2×2
4	Convolutional	64	28×28	3×3
5	Convolutional	64	26×26	3×3
6	MaxPool	64	13×13	2×2
7	Convolutional	96	11×11	3×3
8	Convolutional	96	9×9	3×3
9	MaxPool	96	4×4	2×2
10	Convolutional	128	2×2	3×3
11	MaxPool	128	1×1	2×2
12	Fully-Connected	64 Neurons		
13	DropOut	64 Neurons		
14	Fully-Connected	64 Neurons		
	Output Layer	2 Neurons		

of CNN model is an independent procedure using expert annotations and illumination corrected green channel image. Limited image processing algorithms are involved in the CNN training procedure, except for illumination correction to remove the brightness variation. After the CNN model is properly trained, it will be saved as a classifier for later use.

III. RESULTS AND DISCUSSIONS

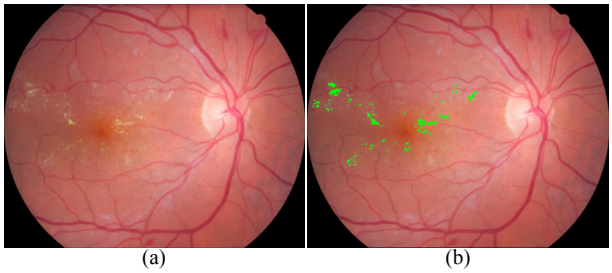
The database used in this research is E-Ophtha EX, a newly published open-source database for exudates segmentation with pixel level annotations by two experts [7]. E-Optha Ex contains in total 82 images, including 47 exudates images and 35 normal images. All the images are acquired under field of view of 45°, with image size ranging from 1440×960 pixels to 2544×1696 pixels. In this research, in order to unify the processing parameters, the image sizes are normalized to have a disc radius of 70 pixels, which results in the image dimension around 1200×800.

For the training of CNN network, we extracted all the positive expert-labeled annotations from the 32 randomly selected training images, which takes around 70% of all the 47 exudates images in the database. 30% of the training data are randomly selected and reserved for performance test. The negative patches are randomly selected from the pixels where no exudates present. In order to keep the balance of two classes, the number of negative samples is selected comparable to that of positive samples. The CNN network is trained and validated on 249,448 patches, including 122,036 positive patches and 127,412 negative patches. The test set contains 65,510 positive patches and 49,412 negative patches, in total 114,922 patches from 15 exudate images. We utilize the deep learning library of Theano, Lasagne and Nolearn for the training of the CNN model. The network has been trained for 150 epochs.

TABLE II

PERFORMANCE OF THE PROPOSED CNN ARCHITECTURE.

	Training Set (%)	Test Set (%)
Accuracy	96.21	91.92
Sensitivity	94.28	88.85
Specificity	98.06	96.00
F-Score	96.05	92.61



(a) (b)

Fig. 5. Exudate identification result with the proposed procedure. (a), the original image with exudate; (b), exudate identified by the algorithm is marked in green.

Table II lists the performance of the proposed CNN architecture on training and test database. An accuracy of 91.92% is achieved on the test set. In comparison, Zhang *et al.* achieved a sensitivity of 74% and positive predicative value of 72% using random forest method on the same dataset at pixel level. Giving the fact that the training and test dataset are randomly divided in the image level, the accuracy on the test set indicates that the trained CNN model has good generalization ability, even though the contrast and image quality may vary.

Fig. 5(b) shows an example of the identification result with the proposed procedure for the image on Fig. 5(a). It is able to identify the exudate in the pixel level.

IV. CONCLUSION

In this paper, we proposed a deep learning based method for exudate detection in the pixel level. A set of exudate candidates are first extracted with morphological ultimate opening techniques and then the candidate points are passed to the trained CNN deep networks for classification. The method achieved a high pixel-wise accuracy for the training and test set.

Future work of this research will involve testing the method on more publicly available databases, e.g. Messidor and DIARETDB databases. Moreover, due to the nature of machine learning and deep learning, the increasing in training data quantity and variation will generally lead to the improvement of model performance. Further work can be done to include more exudate images in the training set, such as the manual labeling provided by DIARETDB1.

V. ACKNOWLEDGMENT

This research is supported by Australian National Health and Medical Research Council (NHMRC Grant no. APP1093682).

REFERENCES

[1] M. D. Abràmoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 169–208, 2010.

[2] T. A. Ciulla, A. G. Amador, and B. Zinman, "Diabetic retinopathy and diabetic macular edema: Pathophysiology, screening, and novel therapies," *Diabetes Care*, vol. 26, no. 9, pp. 2653–2664, 2003.

[3] Wild Sarah, Roglic Gojka, Green Anders, Sicree Richard, and K. Hilary, "Global Prevalence of Diabetes: Estimates for the year 2000 and projection for 2030," *Diabetes Care*, vol. 27, no. 5, pp. 1047–1053, 2004.

[4] T. Teng, M. Lefley, and D. Claremont, "Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy.," *Medical & biological engineering & computing*, vol. 40, no. 1, pp. 2–13, 2002.

[5] N. Patton, T. M. Aslam, T. MacGillivray, I. J. Deary, B. Dhillon, R. H. Eikelboom, K. Yogesan, and I. J. Constable, "Retinal image analysis: Concepts, applications and potential," *Progress in Retinal and Eye Research*, vol. 25, no. 1, pp. 99–127, 2006.

[6] A. Sopharak, B. Uyyanonvara, S. Barman, and T. H. Williamson, "Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods," *Computerized Medical Imaging and Graphics*, vol. 32, no. 8, pp. 720–727, 2008.

[7] X. Zhang, G. Thibault, E. Decencière, B. Marcotegui, B. Laÿ, R. Danno, G. Cazuguel, G. Quellec, M. Lamard, P. Massin, A. Chabouis, Z. Victor, and A. Erginay, "Exudate detection in color retinal images for mass screening of diabetic retinopathy," *Medical Image Analysis*, vol. 18, no. 7, pp. 1026–1043, 2014.

[8] A. Osareh, M. Mirmehdi, B. Thomas, and R. Markham, "Automated identification of diabetic retinal exudates in digital colour images.," *The British journal of ophthalmology*, vol. 87, no. 10, pp. 1220–3, 2003.

[9] Feroui Amel and Messadi Mohammed and Bessaid Abdelhafid, "Improvement of the Hard Exudates Detection Method Used For Computer- Aided Diagnosis of Diabetic Retinopathy," *I.J. Image, Graphics and Signal Processing*, no. May, pp. 19–27, 2012.

[10] M. Niemeijer, B. Van Ginneken, S. R. Russell, M. S. A. Suttorp-Schulten, and M. D. Abràmoff, "Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis," *Investigative Ophthalmology and Visual Science*, vol. 48, no. 5, pp. 2260–2267, 2007.

[11] A. Sopharak, K. T. Nwe, Y. A. Moe, M. N. Dailey, and B. Uyyanonvara, "Automatic Exudate Detection with a Naive Bayes Classifier," *Proceedings of the International Conference on Embedded Systems and Intelligent Technology*, vol. 1, pp. 139–142, 2008.

[12] M. García, C. I. Sánchez, M. I. López, D. Abásolo, and R. Hornero, "Neural network based detection of hard exudates in retinal images," *Computer Methods and Programs in Biomedicine*, vol. 93, no. 1, pp. 9–19, 2009.

[13] P. Kovesi, "Symmetry and asymmetry from local phase," *Tenth Australian Joint Converence on Artificial Intelligence*, pp. 2–4, 1997.

[14] J. Fabrizio and B. Marcotegui, "Fast implementation of the ultimate opening," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5720 LNCS, pp. 272–281, 2009.

[15] S. Beucher, "Numerical residues," *Image and Vision Computing*, vol. 25, no. 4 SPEC. ISS., pp. 405–415, 2007.

[16] T. Walter, J.-C. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy–detection of exudates in color fundus images of the human retina.," *IEEE transactions on medical imaging*, vol. 21, no. 10, pp. 1236–1243, 2002.

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances In Neural Information Processing Systems*, pp. 1–9, 2012.

[18] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *International Conference on Learning Representations*, pp. 1–14, 2015.

[19] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout : A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research (JMLR)*, vol. 15, pp. 1929–1958, 2014.