IMPROVING THE ACCURACY OF OPTIC DISC DETECTION BY FINDING MAXIMAL WEIGHTED CLIQUE OF MULTIPLE CANDIDATES OF INDIVIDUAL DETECTORS

Balazs Harangi, Andras Hajdu

harangi.balazs@inf.unideb.hu, hajdu.andras@inf.unideb.hu
Faculty of Informatics, University of Debrecen, POB 12, 4010 Debrecen, Hungary

ABSTRACT

In this paper, we propose a method to improve the automatic detection of the optic disc on fundus images. We have studied and implemented some of the optic disc detectors from concerning literature to organize them into an ensemble system. As a former work, we proposed an ensemble-based optic disc detection system, based on simple majority voting which already outperformed the individual detectors. To improve further the performance of the ensemble-based system, now we examine how we can extract more candidates from the individual algorithms to have the appropriate location of the optic disc among them. We also assign weights to each candidate based on the priority suggested by the algorithms. We consider these weighted candidates as vertices of a graph and look for a subgraph with a maximal sum of weights constrained by the geometry of the optic disc. Experimental results are also presented to see the improvement.

Index Terms— optic disc, fundus image, graph

1. INTRODUCTION

Nowadays, diabetes is among the rapidly growing diseases affecting a large portion of the population. Screening examinations have great importance as the earlier the disease is recognized, the less serious the symptoms become. Since 6.4% of the Earth's population suffer from diabetes (more than 44 million people), developing automatic screening systems is strongly needed.

Diabetic retinopathy is a specific field, when diabetes threatens the vision through ruining the retina. For this disease, the main task of an automatic screening system is to detect the typical symptoms, like exudates, and cotton-wool spots. These lesions appear on retinal images as bright patches which are similar to the optic disc (OD can be seen in Figure 1) in color and shape characteristics. Thus, OD localization with its consequent removal is highly recommended before detecting these lesions.

In the corresponding literature, several optic disc detection algorithms are published [1-9]. Most of them try to extract the optic disc based on color, size, shape, entropy, direction of vessels, etc. The majority of these algorithms are able to locate the optic disc on healthy images. However, their accuracies decrease if image quality is not sufficiently high or severe lesions appear on the image. To overcome the imperfectness of the individual algorithms, we have suggested an ensemble of them [4] which approach outperformed the individual methods. In this ensemble, we considered simple majority voting, where every individual algorithm [5-8] had a single candidate. Now, we apply further improvements to this model. Namely, on the one hand, we extract more than one candidate for each algorithm to increase the chance of getting one of the candidates as a true one for the OD. Furthermore, we assign weight to each candidate to replace simple majority voting by a weighted one and treat them as weighted nodes of a complete graph. For the location of the OD we select that subgraph of the candidates which meets the OD geometry constraint and has a maximal total weight. For this selection we borrow a graph theoretical approach supplying the optimal solution.

The rest of paper is organized as follows: in section 2, we list the individual OD detector algorithms that we used to build up the ensemble. Section 3 deals with the extraction of multiple candidates of these algorithms. In section 4, we explain how a clique of candidates with maximal total weight can be found by a graph theoretical approach. Experimental results showing the improvement over simple majority voting is given in section 5. Finally, we draw some conclusions in section 6.

2. INDIVIDUAL OD DETECTOR ALGORITHMS

In this section, we give a short description for each OD detector we included in the ensemble. We collected seven methods based on different principles.

2.1. Based on pyramidal decomposition

Lalonde *et al.* [5] proposed an algorithm which generates a pyramid with simple Haar-based discrete wavelet transform. The image has low resolution on the four- and five-level decomposition provided that the input image size was approximately 600×500 pixel. The pixel which has the highest intensity value in the low-resolution image is considered as the center of the OD.

2.2. Based on edge detection

This method [5] uses edge detection algorithm which is based on Rayleight-based CFAR threshold. This solution guarantees the strong edges which are associated to the border of anatomical structures. In the next step the Hausdorff distance is calculated between the set of edge points and a circular template like the average OD. The lowest distance value means the center of the OD.

2.3. Based on entropy measurement

Sopharak *et al.* [6] proposed this method which applies a median and a CLAHE filter on the intensity image. The entropy of intensity in the local region around each pixel is calculated and the location of the highest entropy values is considered as belonging to the OD.

2.4. Based on kNN classification

Niemeijer *et al.* [7] proposed feature extraction and classification steps to determine the area of optic disc. They suggest the next features: number, width, orientation and density of vessels and combination of these. In the next step a kNN classifier made a decision about each pixel whether it belongs to the OD or not.

2.5. Based on fuzzy convergence of blood vessels

In this method [8] the vessel system is thinned and each line-shape segment is modeled by a fuzzy segment. This model creates a voting map and the pixel which obtained the most votes is considered as the center of the OD.

2.6. Based on Hough transformation of vessels

Ravishankar *et al.* [9] proposed this algorithm which uses Hough transformation on the thinned vessel system. Lines which have slope smaller than 45° are eliminated. The intersection points of the rest of the lines result a voting map. Number of votes is weighted with intensity values which are under the intersection points on the image. Finally the highest value provides the center of the OD.

2.7. Based on circle detection

This algorithm locates the border of the OD in terms of a circle with a given diameter using the Hough transformation. First, we apply edge detection and find that circle which contains the most edge points.

3. EXTRACTION OF MULTIPLE CANDIDATES

In our former work [4], we have already examined how the results of individual algorithms can be combined, with selecting that OD-sized region, where the most algorithms voted to. In [4] all the individual detectors gave their single vote for the OD center. Now, we let them to suggest multiple candidates, and also assign weights to them.

3.1. Multiple candidates and OD-geometry constraint

All the individual algorithms are capable to assign a value $P(x) \in \Re^+$ to each image pixel x to show the probability of being the OD center. These values together compose the probability map of an algorithm for the image; see Figure 1 for the probability maps of the algorithms. To have multiple candidates from an algorithm, we can simply select not only one, but more pixels as OD center candidates having large probability value (see Figure 2 for a specific algorithm).

As a geometric constraint, the candidates of the same algorithm are not allowed to be closer than the average OD diameter (Figure 2). In this way we can assure that a possible OD region is suggested by only one candidate of each algorithm.

Figure 1. Probability maps of individual algorithms and the optic disc. Narrows show the maximal values.

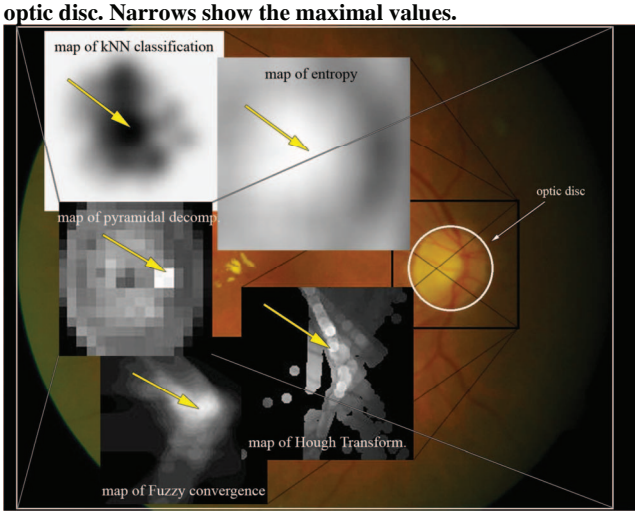
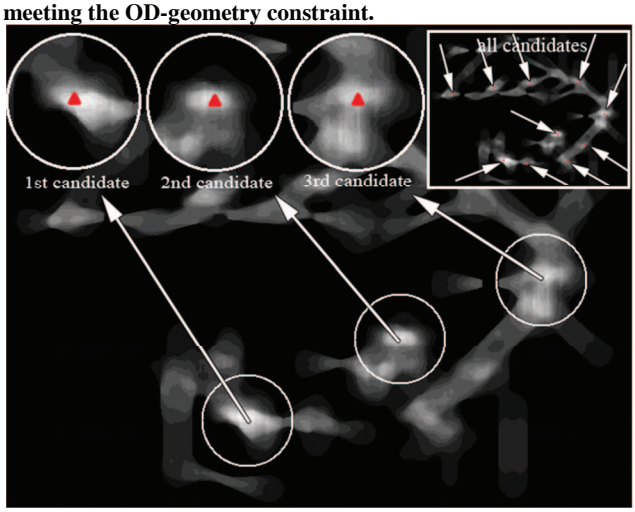


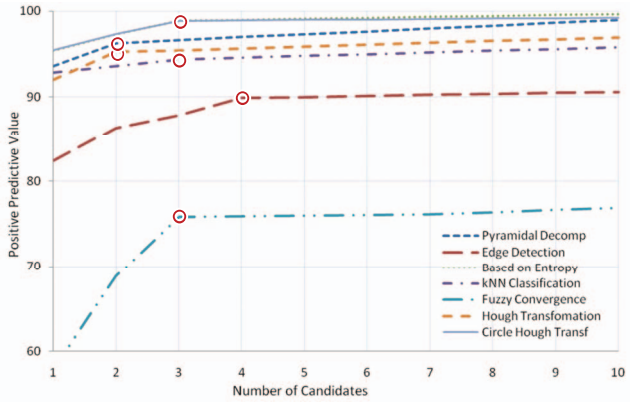
Figure 2. The probability map of algorithm [8] with candidates meeting the OD-geometry constraint.



3.2. Number of candidates

It is a key issue, how many candidates of specific algorithm we should consider. Let N_i denote the number of the candidates considered for the i^{th} algorithm ($i=1,2,\ldots,7$). We have to keep the N_i values relatively small to avoid too many false positive candidates confusing the final combination. On the other hand, we try to have the true OD location among the N_i candidates of each individual OD detector. Thus, for each individual algorithm we investigated how probable that the true location of the OD is among the candidates as we increase N_i. This probability value was calculated using a training dataset and checking the ratio of the images, where the true OD was among the N_i candidates. We found that for each individual detector we can fix a small N_i above which this probability value will increase only slowly (see Figure 3 for the proper behavior). Note that, this selection is rather close to the elbow method in kNN clustering.

Figure 3. Number of candidates of algorithms with a suggestion to true OD location. Red circles show the chosen N_i values.



3.3. Weights of the candidates

Let $C_i = (c_{i,1}, c_{i,2}, \ldots, c_{i,N_i})$ denote the candidates of the i^{th} algorithm in a priority order so that $P(c_{i,k}) \geq P(c_{i,l})$ for any $1 \leq k < l \leq N_i$. Then, we assign weights to the candidates, as:

$$\omega_{i,j} = \frac{P(c_{i,j})}{\max_{k=1}^{N_i} P(c_{i,k})}$$

for $i=1,\ldots,7$, $j=1,\ldots,N_i$. Note that, this normalization will give $\omega_{i,1}=1$ (weight 1 for every first candidate) for all $i=1,2,\ldots,7$. Moreover, the weights also reflect the priority order of the algorithms regarding their own candidates.

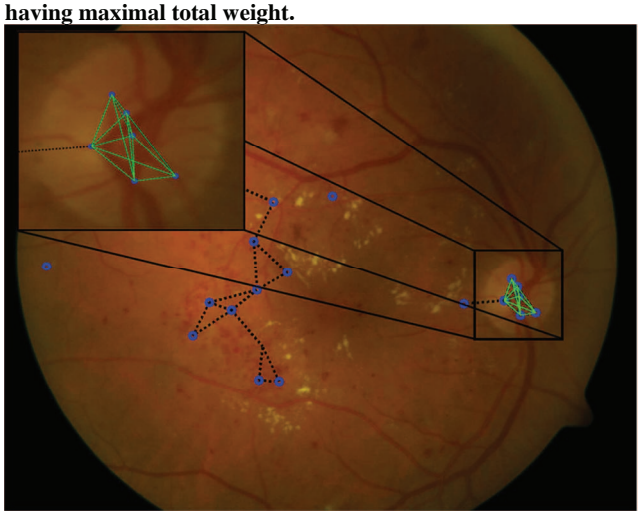
4. FINDING A CLIQUE WITH MAXIMAL TOTAL WEIGHTS

As a generalization of simple majority voting, we locate the OD as the region containing candidates having largest weight in total. Note that the OD-geometry constraint provides that within an OD-sized region only candidates belonging to different algorithms may fall. To solve this problem we build up a graph, whose vertices are the $i*N_i$ candidates weighted as given in section 3.3. Moreover, we connect those vertices with edges that fall inside a possible OD region, that is, their distance is smaller than the OD diameter. For such a graph see Figure 4.

In this graph we have to look for the maximal clique with maximal sum of weights of nodes. The corresponding graph theoretical problem in this representation is the maximum-weight clique problem. This problem is known to be NP-hard. For the solution of this problem we borrow the algorithm [10] which is based on a heuristic vertex-coloring and a backtrack search. This algorithm provides the maximal-weighted cliques and works quickly on a dense graph, too.

The obtained maximal-weighted clique is considered as the solution of the proposed ensemble-based OD detection system with deriving a single centroid from them.

Figure 4. The clique obeying the geometric constraint and having maximal total weight.



5. RESULTS

We have tested the proposed method on the DiaretDB0 (containing 130 images) and DiaretDB1 (containing 89 images) databases. The resolution of the images in these databases is 1500x1152.

For the measurability of the algorithm accuracies, a manually selected OD is also supplied for the images. We consider two ways for accuracy measurement. First, as a primary error, we expected that a good candidate falls inside the manually selected OD region. Then, as a secondary measurement, we check the distance of a candidate from the centroid of the manually selected OD.

As for the primary error function, the final OD center candidate is considered as a true positive if it falls in the manually selected OD, while it is a false positive when it

does not. Moreover, let *TP* (resp. *FP*) denote the number of total number of true positive (resp. false positive) candidates within the given test database. Then, we can use the Positive Predicted Value to measure the primary error as:

$$PPV = \frac{TP}{TP + FP}.$$

As for the secondary error function, for a fair comparison we exclude those images where the result of the ensemble system falls outside the OD region:

$$Err_2 = \frac{1}{|TP|} \sum_{C \, is \, true \; poitive} d(C, C_{OD}),$$

where C is the proposed candidate, while and C_{OD} is that center of the manually given OD, respectively.

Table 1 compares the respective performances of the individual algorithms, simple majority voting and the proposed weighted majority voting method on the tested databases. As we can see, the proposed solution achieved highest accuracy.

6. CONCLUSION

Table 1. Accuracies of individual detectors for the primary and secondary error functions.

	DiaretDB0		DiaretDB1	
	primary	secondary	primary	secondary
OD_{decomp}	96.15%	37.11	91.01%	40.26
OD_{edge}	83.84%	38.67	80.89%	26.50
$OD_{entropy}$	95.38%	29.40	95.50%	25.43
OD_{kNN}	94.61%	18.40	91.01%	26.50
OD_{fuzzy}	56.15%	48.51	59.55%	43.45
OD_{hough}	90.76%	39.11	93.25%	39.69
OD_{circle}	88.46%	18.13	87.64%	20.00
majority voting	97.69% FP = 3	16.36	94.38% FP = 5	16.93
proposed model	**98.46% FP = 2**	**10.38**	**98.87% FP=1**	**10.60**

In this paper, we have proposed a robust solution to the automatic optic disc detection on fundus images. For this aim, we have used several OD detectors to build up an ensemble. We let these individual algorithms to suggest multiple candidates, and assigned weights to the candidates based on the priority of the algorithm. The final OD location is found as a group of candidates fulfilling the OD-geometry constraint and having largest total weight. To find the solution we applied a graph theoretical maximum-weight clique extraction method. Experimental results were also presented to see the improvement.

7. ACKNOWLEDGEMENT

This work was supported in part by the Janos Bolyai grant of the Hungarian Academy of Sciences, by the TÁMOP-4.2.2/B-10/1-2010-0024 project which is co-financed by the European Union and the European Social Fund, and by the TECH08-2 project DRSCREEN – Developing a computer based image processing system for diabetic retinopathy screening of the National Office for Research and Technology of Hungary (contract no.: OM-00194/2008, OM-00195/2008, OM-00196/2008).

8. REFERENCES

[1] Abramoff, M. D., Niemeijer, M., "The Automatic Detection of the Optic Disc Location in Retinal Images using Optic Disc Location Regression", In 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), pp. 4432–4435, 2006.

[2] C. Sinthanayothin, J. F. Boyce, T. H. Williamson, H. L. Cook, E. Mensah, S. Lal, and D. Usher, "Automated Detection of Diabetic Retinopathy on Digital Fundus Images", Diabetic Medicine, Vol. 19, pp. 105–112, 2002.

[3] M. Niemeijer, M. D. Abràmoff, and B. van Ginneken, "Segmentation of the Optic Disc, Macula and Vascular Arch in Fundus Photographs", IEEE trans. on medical imaging, Vol. 26, pp. 116–127, 2007.

[4] B. Harangi, R. J. Qureshi, A. Csutak, T. Peto, A. Hajdu, "Automatic Detection Of The Optic Disc Using Majority Voting In A Collection Of Optic Disc Detectors", IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1329-1332, 2010.

[5] Lalonde, M., Beaulieu, M. and Gagnon, L., "Fast and Robust Optic Disk Detection using Pyramidal Decomposition and Hausdorff-Based Template Matching", IEEE Trans. Medical Imaging, Vol. 20, pp. 1193-1200, Nov. 2001.

[6] Sopharak, A., Thet New, K., Aye Moe, Y., N. Dailey, M., Uyyanonvara, B., "Automatic Exudate Detection with a Naive Bayes Classifier", International Conference on Embedded Systems and Intelligent Technology, Grand Mercure Fortune Hotel, Bangkok, Thailand, pp. 139–142, 2008.

[7] M. Niemeijer, M. D. Abràmoff , B. van Ginneken, "Fast detection of the optic disc and fovea in color fundus photographs", Medical Image Analysis, vol. 13, pp. 859–870, 2009

[8] A. Hoover and M. Goldbaum, "Locating the Optic Nerve in a Retinal Image using the Fuzzy Convergence of the Blood Vessels", IEEE Trans. Med. Imag., Vol. 22, pp 951-958, 2003.

[9] S. Ravishankar, A. Jain, A. Mittal, "Automated Feature Extraction for Early Detection of Diabetic Retinopathy in Fundus Images", CVPR - IEEE Conference on Computer Vision and Pattern Recognition, pp. 210-217, 2009.

[10] D. Kumlander, "A new exact algorithm for the maximum-weight clique problem based on a heuristic vertex-coloring and a backtrack search", Proceedings of the Fourth International Conference on Engineering Computational Technology, 2004.