

Integrated Optic Disc and Cup Segmentation with Deep Learning

Gilbert Lim, Yuan Cheng, Wynne Hsu, Mong Li Lee

School of Computing

National University of Singapore

{limyongs,cyuan,whsu,leeml}@comp.nus.edu.sg

Abstract—Glaucoma is a widespread ocular disorder leading to irreversible loss of vision. Therefore, there is a pressing need for cost-effective screening, such that preventive measures can be taken. This can be achieved with an accurate segmentation of the optic disc and cup from retinal images to obtain the cup-to-disc ratio. We describe a comprehensive solution based on applying convolutional neural networks to feature-exaggerated inputs emphasizing disc pallor without blood vessel obstruction, as well as the degree of vessel kinking. The produced raw probability maps then undergo a robust refinement procedure that takes into account prior knowledge about retinal structures. Analysis of these probability maps further allows us to obtain a confidence estimate on the correctness of the segmentation, which can be used to direct the most challenging cases for manual inspection. Tests on two large real-world databases, including the publicly-available MESSIDOR collection, demonstrate the effectiveness of our proposed system.

Keywords-Glaucoma screening, optic disc segmentation, optic cup segmentation

I. INTRODUCTION

Glaucoma is an ocular disorder characterized by the progressive degeneration of optic nerve fibers. Due to its lack of symptoms in its developing stages, and concurrent irreversibility once significant vision loss is detected, early detection of its onset is highly desirable if eventual blindness is to be avoided. While there has been work done on predicting glaucoma from other retinal image characteristics such as vessel distribution [1], FFT/B-spline coefficients [2] and various texture descriptors [3], most of the known literature has endeavoured to directly determine the CDR, generally by segmenting the disc and cup pixels.

Optic disc and cup segmentation has traditionally involved the judgment of skilled human graders, and can potentially be time-consuming, requiring about eight minutes per eye under the Klein protocol [4]. In current clinical practice, pallor contrast and vessel kinks are among the key features utilized by graders to estimate the cup boundary [5]. However, it remains an open problem as to how best to translate this knowledge into an automated procedure. We propose a novel data-driven segmentation on convolutional neural network outputs, instead of the conventional approach of working directly on the raw or filtered retinal images.



(a) Original disc image

(b) With overlays

Figure 1: An example of ground truth disc boundary (in blue) and cup boundary (in green), with vertical heights

The major contribution of this work is in leveraging the ability of convolutional neural networks to accurately segment the optic disc and cup, especially after relevant input features are appropriately exaggerated. The segmentations on two databases of retinal images are demonstrated to be superior or comparable to those from other known state-of-the-art methods. Additionally, we describe a metric for quantifying segmentation noise, which allows us to flag especially challenging images for manual inspection. This is necessary if our system is to eventually be deployed in a screening environment.

II. RELATED WORK

There have been a number of attempts at automatically segmenting the optic disc and cup. To begin with, the general location of the optic disc needs to be isolated as a precursor to further classification. This has usually been achieved by exploiting the properties that the optic disc tends to be much brighter than its surroundings, and is the ultimate source of retinal blood vessels [6] [7].

After localization, more exact determinations of the optic disc boundary have tended to involve adaptations of deformable energy-based models. Examples include gradient vector flow snakes on various colour spaces [8], circular deformable models [9], iterative knowledge-based clustering [10], circular Hough transforms [11] [12], ellipse refitting [13] and the multiresolution sliding band filter [14].

Cup segmentation has as mentioned been recognized to be a challenging task even for human graders. The generally-brighter pallor of the optic cup is an obvious characteristic, and is the primary consideration of many methods. Similar to the optic disc, deformable models can be applied directly, as with level-set thresholding following histogram analysis [15], hull fitting [16], active shape models [10], active contour models [17] and graph cuts [18].

However, a traditional limitation of such approaches is their reliance on pallor, which may not be definitive on images with weak or non-standard contrast. Therefore, the use of prior information to mitigate such deficiencies is common. As examples, we compare our method against pixel-level class estimates obtained from graph cuts on Gaussian Mixture Models (GMM) [18], as well as from physiologically-plausible features [19].

Some authors have attempted to extract and operate on secondary features, such as with a watershed morphological transform [20], grouped sparsity constraints on sliding windows [21], optimal image selection [12] and superpixel propagation [22]. Other than these intensity-based features, progress has also been made in spatial heuristic ensembling of multiple candidate segmentations [23], the incorporation of vessel kink data [17] and sector analysis using structural information [24].

Finally, optimal linear reconstruction with codebooks [25] has resulted in the best currently-known optic cup segmentation results. This involves first segmenting and normalizing the disc to a fixed circular template, with illumination equalization. While the reconstruction-based method has been shown to perform well in cup segmentation, it depends on a separate optic disc segmentation algorithm, and cannot adequately describe irregular cups since its outputs are limited to ellipse parameters. We address all the limitations raised above with our proposed method.

III. PROPOSED METHOD

Our method consists of four main stages, applied in order – the coarse localization of a square region around the optic disc, the conversion of this region to exaggerate relevant visual features, the classification of the converted image with a convolutional neural network to produce a pixel-level probability map on the retinal background, optic rim and optic cup classes, and finally the segmentation of that map to produce the predicted disc and cup boundaries.

We obtain the most probable coarse location of the optic disc using the Daubechies wavelet transform method described in [7]. Since we do not depend on any specific parameters (e.g. the disc centre) from this stage, it is sufficient that the entirety of the disc is contained within the square region returned by the localization procedure. To provide sufficient context about the surrounding retinal background, we set the width of the square region to about twice the width of the typical optic disc within.

A. Vessel-free Red-green-kink Feature Exaggeration

We next convert the image to a representation suited for optic disc and cup segmentation. The blood vessels within the optic disc tend to obscure the more important pallor and contrast features, and provide little useful information for optic cup segmentation with the exception of the location of vessel kinks. Therefore, the basic idea behind this preprocessing is to replace vessel pixels with the extrapolated disc intensity as far as is possible, while separately adding back kink indicators at the relevant locations.

Firstly, we obtain a mask of the vessels (Figure 2b) by trench detection [26]. Ideally, the width of the mask should completely cover the vessels, with an additional slight margin. A certain amount of vessel over-segmentation is acceptable, since the intent is to completely remove the influence of vessels on the optic disc image.

The masked vessel pixels are then filled in with plausible optic disc background pixels from outside the vessel mask, through nonparametric patch sampling (Figure 2c). This process attempts to automatically reconstruct the appearance of the optic disc as if the vessels were not present [2] [14] [24]. From this inpainted image, we extract the red and green channels, since it is popularly recognized that the blue channel contains little information of value in retinal image analysis, and equalize the channel histograms (Figure 2d).

We then skeletonize the vessel mask (Figure 2e). Unlike previous methods [5] [17], we need not try to qualify the “correct” kinks explicitly, as they will be inferred during CNN classification. We therefore represent all vessel kinks in the input. The raw kink strength is defined at a skeletonized vessel pixel p , by considering the $p - k$ and $p + k$ pixels before and after it along the same vessel segment, if and only if $\{p - k, p + k\}$ exist unambiguously (i.e. there are no junctions along the vessel segment within the range). A circular arc is fitted to the three points $\{p, p - k, p + k\}$, by estimating the chord w and height h of the arc, and thereby the circle radius r . The raw kink strength \mathcal{K} is then derived from r :

$$\mathcal{K} = \begin{cases} \sqrt{1/r}, & \text{if } r > 0 \\ 0.0, & \text{otherwise} \end{cases} \quad (1)$$

The values of the raw kink strength map (Figure 2f) are first linearly normalized to lie between 0 and 1, and then expanded by first dilating nonzero vessel strength pixels by d pixels, before applying a Gaussian blur with kernel σ (Figure 2g). For our experiments, we use $k = 9, d = 7, \sigma = 10$.

The finished converted representation (Figure 2h) then comprises the histogram-equalized vessel-free red-green image (Figure 2d), combined with the blue channel from the expanded kink strength map (Figure 2g), and is designed to exaggerate the most visually-relevant features used for optic disc and cup segmentation.

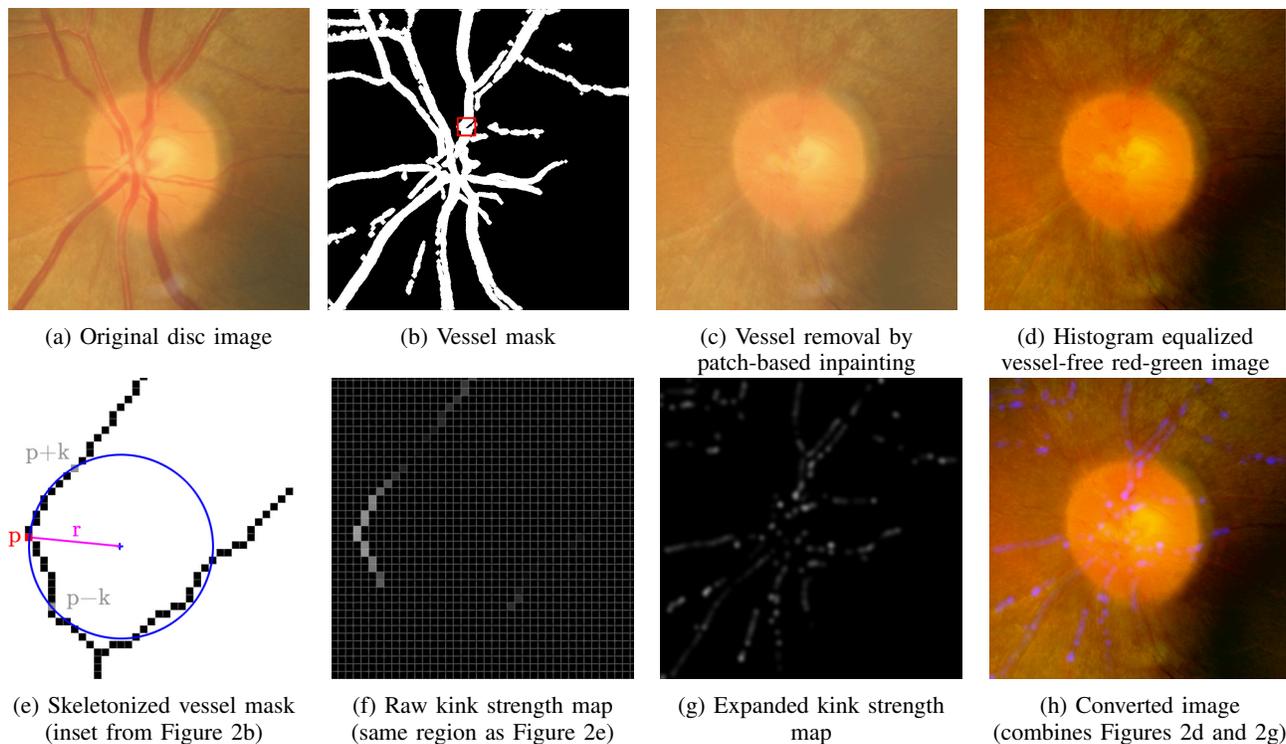


Figure 2: Vessel-free red-green-kink conversion

B. Three-Class CNN Pixel-Level Classification

Each pixel within the converted representation is then individually classified with a convolutional neural network (CNN). CNNs have recently shown themselves to be among the most effective techniques for image classification tasks. In the biomedical field, they have been applied to the identification of mitotic cells [27] and segmentation of neuronal membranes [28], but not yet to the optic disc, as far as is known.

The theoretical justification for utilizing CNNs is their capability to consider the detailed neighbourhood context of each and every pixel, in computing its class probability. The variability between different optic disc images has been a major challenge faced by other recent approaches. In some [18], the probability estimate for a pixel is fixed depending solely on the input colour. In others [22], the potential accuracy is restricted by the pre-grouping of pixels into larger superpixel units. In contrast, CNNs suffer neither of these limitations.

CNNs are specialised feedforward multi-layer perceptrons that consist of alternating convolutional and subsampling layers [29]. Each layer consists of a number of nodes, organized into maps, that are connected to nodes in the next layer. Layers are composed of three main types: convolutional, max-pooling and fully-connected.

1) *Convolutional layer*: A convolutional (C) layer emphasizes spatially-local correlations by accepting connec-

tions only from the immediate neighbourhood of a pixel. It is composed of multiple layer maps of identical dimensions. Each node in a convolutional layer map is connected to nodes of a map in the previous layer by a kernel, as illustrated in Figure 3b. Each unique combination of maps between layers has its own kernel, and thus shares a set of weights.

2) *Max-pooling layer*: A max-pooling (MP) layer map downsamples a convolutional layer map by first dividing it into disjoint uniform-sized pooling windows, and then assigning to each node in the max-pooling layer map the maximum individual value from the corresponding window. This provides a degree of spatial invariance, and reduces the amount of computation required by succeeding layers due to the reduction in the number of nodes. It has been found to perform better compared to conventional subsampling by averaging.

3) *Fully-connected layer*: A fully-connected (FC) layer is defined by having each of its nodes connected to every node in the previous layer, with each connection having an independent weight. This is necessary at the conclusion of the CNN to combine the final max-pooled results into a single output vector. By using a softmax function, this vector can be interpreted as the probabilities $\{P_b, P_r, P_c\}$ s.t. $P_b + P_r + P_c = 1$, which are the predicted likelihoods that the input pixel belongs to the background, rim and cup classes respectively.

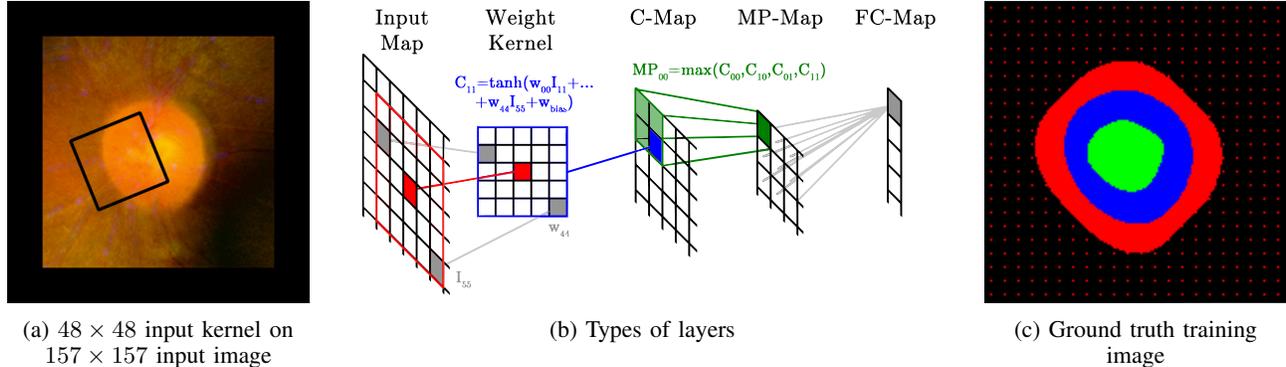


Figure 3: Convolutional neural network operation

For all convolutional and fully-connected layers, we use a scaled hyperbolic tangent activation function \tanh .

Training of a CNN consists of first randomly initializing all weights with values uniformly distributed in the range $[-0.05, 0.05]$. An equal number of source pixels from each class are then sampled from the training set, and randomly ordered. During each training iteration, an image patch of random orientation centered on the current source pixel, and of dimension equal to the input layer map, is used as the input to the CNN (Figure 3a). These input values are then forward propagated through the CNN layers, with the output of each layer feeding into the next layer as input, eventually producing an output vector from the last layer.

The error between the actual and expected output vector values at the last layer, as defined by the ground truth training output (Figure 3c), is then used to update the connection weights by gradient descent backpropagation [29]. The objective of this process is to minimize the mean output error, over many iterations. Note that since we wish to focus on learning the difference between the background and rim, we undersample background pixels that are more than 15 pixels distant from the rim boundary.

Training of the CNNs was performed using a custom C++/OpenCL implementation on four AMD Radeon HD 7970 GPUs. We use an epoch size of 1.5 million and a learning rate of 0.001, with a decay of 0.9 per epoch. Training is stopped upon convergence on the softmax loss of the output layer, which was generally achieved after about twenty epoches. The architecture of the CNN used is described in Table I.

C. Probability Map Segmentation

The CNN probability map predictions may be noisy, which is accounted for by the following segmentation procedure. The upscaled probability map outputs from the CNN are first averaged over four input orientations ($\theta = 0^\circ, 90^\circ, 180^\circ, 270^\circ$), before the application of a small Gaussian blur ($\sigma = 21$) to smooth out noise (Figure 4a). To enforce the constraint that the optic cup must be wholly

Layer	Type	Maps \times (nodes)	Kernel	Weights	Connections
0	I	$3 \times (48 \times 48)$	–	–	–
1	C	$20 \times (44 \times 44)$	5×5	1520	2942720
2	MP	$20 \times (22 \times 22)$	2×2	–	–
3	C	$20 \times (18 \times 18)$	4×4	10020	3246480
4	MP	$20 \times (9 \times 9)$	2×2	–	–
5	C	$40 \times (6 \times 6)$	4×4	12840	462240
6	MP	$40 \times (3 \times 3)$	2×2	–	–
7	FC	$100 \times (1 \times 1)$	–	36100	36100
8	FC	$3 \times (1 \times 1)$	–	303	303

Type: I=Input, C=Convolutional, MP=Max-pooling, FC=Fully-connected

Table I: 9-layer CNN architecture used

contained within the optic disc, we start by determining the disc boundary. First, a *disc boundary probability* is defined at each pixel as:

$$P_D = \begin{cases} r/(r+b), & \text{if } r/(r+b) < T_D \\ 1.0, & \text{otherwise} \end{cases} \quad (2)$$

where r is the original rim class probability, b is the original background class probability, and T_D is a parameter thresholding the boundary strength. The optic disc centerpoint and approximating ellipse are estimated using an elliptical Hough transform. We then obtain the final disc segmentation by transforming the boundary strength map into polar coordinates ($\theta = 2^\circ, d = 2$) on the estimated centerpoint, and then apply dynamic programming to find the best path. This has previously been used only with features directly extracted from the image, such as edge strength, texture and pixel variation [30]. The cup segmentation is likewise obtained on $P_C \rightarrow B_C$ by the same procedure, with all cup pixels falling outside the rim area removed.

Finally, we quantify the reliability of the final disc and cup segmentation. Let the segmentation *noise* be defined as:

$$\mathcal{N} = (1 - (C_m \cap C)/C_m) + (1 - (R_m \cap R)/R_m) \quad (3)$$

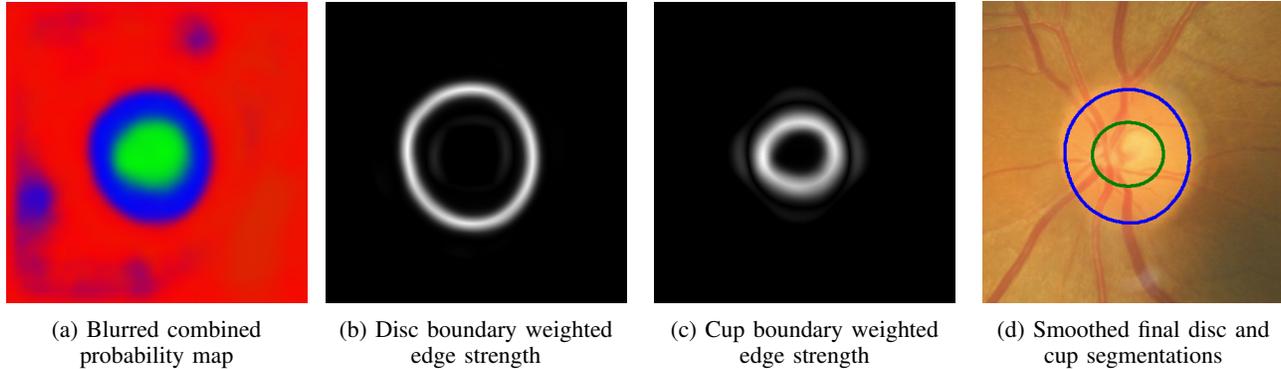


Figure 4: Probability map segmentation procedure

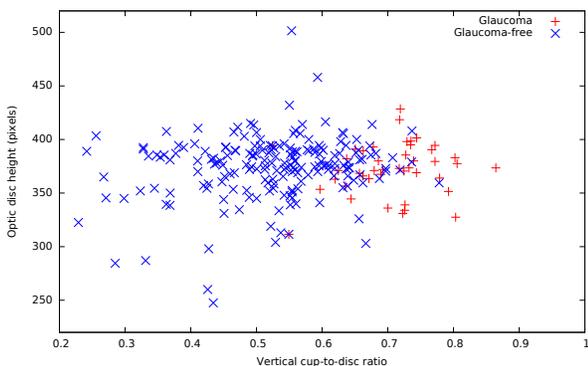


Figure 5: Distribution of optic disc height vs. CDR, SEED-DB database

where R_m and C_m are the set of pixels classified as rim and cup respectively in the blurred probability map, and R and C are the set of pixels classified as rim and cup respectively in the final segmentation. The more closely the pixel-level classifications agree with the final segmentation, the more reliable the segmentation is considered to be.

IV. EMPIRICAL EVALUATION

To demonstrate the effectiveness of our proposed CNN-FE method, we conduct comprehensive experiments on two large databases of real-world retinal images.

The MESSIDOR database is a public fundus image database of 1200 images of dimensions ranging from 1440×960 to 2304×1536 , with corresponding optic disc ground truth segmentations available separately.

The SEED-DB database consists of 235 images of dimension 3504×2336 from the Singapore Epidemiology of Eye Disease (SEED) program, 43 with glaucoma, and 192 without (Figure 5). Each image has a ground truth segmentation GT marked by a trained grader. For each image, we localize a region of dimension 791×791 containing the optic disc and its immediate neighbourhood.

We first evaluate the performance of CNN-FE on optic disc segmentation against a collection of recent state-of-the-art methods on both MESSIDOR and SEED-DB. We then similarly evaluate the performance of CNN-FE on optic cup segmentation, vertical cup-to-disc ratio determination and as a glaucoma screening measure on SEED-DB.

A. Methodology

For all of the following experiments, we divide the set of images into four approximately equal-sized folds. Then, to segment images in any particular fold, all images in the other three folds are used in cross-validation training of the required CNN models. To evaluate whether the feature exaggeration is in fact essential, we train our CNNs with both the feature-exaggerated (CNN-FE) and original (CNN-OR) retinal images as input, with performance maximized on a grid search on $\{T_D, T_C\}$ as described in Section III-C. We apply CNN-FE and CNN-OR on the regions after they are rescaled to a size of 157×157 pixels.

We focus on two commonly-used metrics to quantify our segmentation accuracy. Firstly, the non-overlap ratio (NOR) m_1 measures the extent to which the segmented area overlaps with the ground truth, and is defined as:

$$m_1 = 1 - \frac{E_{dt} \cap E_{gt}}{E_{dt} \cup E_{gt}} \quad (4)$$

where E_{dt} is the predicted optic cup (or disc) region, and E_{gt} is the corresponding ground truth cup (or disc) region.

Secondly, the absolute vertical cup-to-disc (CDR) error δ is the difference between the predicted and ground truth vertical cup-to-disc ratios, and is defined as:

$$\delta = |CDR_{dt} - CDR_{gt}| \quad (5)$$

where CDR_{dt} is the predicted cup-to-disc ratio, and CDR_{gt} is the ground truth cup-to-disc ratio. The cup-to-disc ratio is defined as the maximum vertical height of the cup region, divided by the maximum vertical height of the disc region. An example of these heights can be found in Figure 1b.

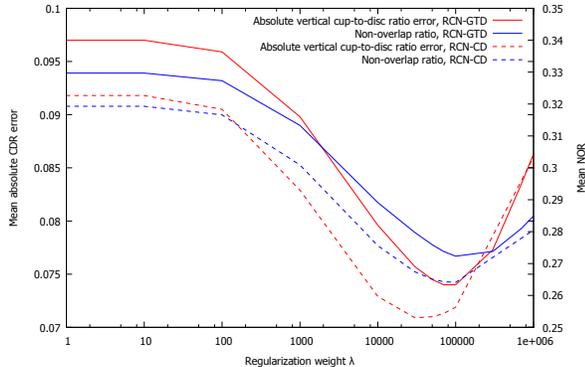


Figure 6: Mean absolute CDR error and NOR vs. regularization weight λ , RCN method

B. Optic disc segmentation

For optic disc segmentation, ground truth data is available on both the MESSIDOR and SEED-DB databases. In the case of MESSIDOR, our training input consists only of the background and optic rim classes. We compare the obtained non-overlap ratio statistics with Hausdorff-based template matching [31], circular Hough transform [11], superpixel [22], ellipse refitting [13] and sliding band filter [14] methods. Their results listed in Table II are as obtained from the relevant papers.

We have reimplemented the prior-based graph cut (PBGC) [18] and physiologically-plausible features (PPF) [19] methods. For PBGC, we explored using the original, greyscale-converted and histogram-equalized images, and report the best results, which were obtained using histogram-equalized inputs. For PPF, we have omitted the sole stereo-dependant disparity feature among the twelve used.

	CNN-FE	[14]	[13]	[22]	[11]	[31]
Mean m_1	0.112	0.114	0.120	0.125	0.140	0.190
$m_1 \leq 0.30$	0.964	0.94	0.94	–	0.93	0.85
$m_1 \leq 0.25$	0.948	0.92	0.92	0.92	0.90	0.77
$m_1 \leq 0.20$	0.922	0.87	0.88	0.86	0.84	0.70
$m_1 \leq 0.15$	0.864	0.81	0.82	0.76	0.73	0.45
$m_1 \leq 0.10$	0.614	0.66	0.59	0.51	0.46	0.25
$m_1 \leq 0.05$	0.073	0.23	0.13	0.08	0.07	0.02

Table II: Optic disc NOR results, MESSIDOR

For SEED-DB, we compare the CNN-FE results only against those obtained by PBGC and PPF (Table III), since reconstruction-based localization is purely an optic cup segmentation method.

The first row in Tables II and III give the mean NOR obtained by each of the methods. The remaining rows in the tables describe the *distribution* of the NOR values.

We observe that the optic disc segmentation performance of CNN-FE is superior to that of known state-of-the-art

	Method/Ground truth			
	CNN-FE	CNN-OR	PBGC	PPF
Mean m_1	0.0843	0.0961	0.1050	0.1255
$m_1 \leq 0.30$	0.996	0.974	0.871	0.953
$m_1 \leq 0.25$	0.991	0.962	0.849	0.909
$m_1 \leq 0.20$	0.987	0.949	0.806	0.858
$m_1 \leq 0.15$	0.970	0.923	0.754	0.767
$m_1 \leq 0.10$	0.759	0.774	0.703	0.552
$m_1 \leq 0.05$	0.108	0.119	0.058	0.047

Table III: Optic disc NOR results, SEED-DB

methods. While CNN-FE segmentation tends not to be as exactly precise on the best-segmented images, relative to results from the sliding band filter and ellipse refitting, this is made up with there being fewer severe errors as compared to other methods, which may be due to discs with less straightforward boundaries being better interpreted by the CNN pixel-level classifier.

C. Optic cup segmentation

For optic cup segmentation, ground truth data is available only on the SEED-DB database. As a further comparison in addition to PBGC and PPF, we have reimplemented the reconstruction-based optic cup localization method with codebooks (RCN), as described in [25], which currently produces the best optic cup segmentations based on both absolute CDR error and NOR metrics, to the best of our knowledge. The reconstruction-based with codebooks method has been empirically demonstrated to strongly dominate the recent pixel [15], sliding window [21], threshold [32], r-bend [17] and superpixel [22] based approaches, as reported in [22] and [25].

Since reconstruction-based localization requires a prior optic disc segmentation, we provided it with both the ground truth disc segmentation (RCN-GTD), as well as the CNN-FE disc segmentation (RCN-CD), to check for consistency and to simulate real-world operation. We have also taken care to search for an optimal regularization weight, and found that performance is maximized at $\lambda = 70000$ (Figure 6) on both the absolute CDR error and NOR metrics.

The results for optic cup NOR, and for the absolute CDR error (including weighted kappa κ_W as defined in [33]) are given in Table IV. For all methods, the three images (1.3% of the total) rejected due to having segmentation noise $\mathcal{N} \geq 1$ were ignored in calculating the results.

We observe that CNN-FE offers a considerable improvement over the reconstruction-based, physiologically-plausible features and prior-based graph-cut methods, on both the optic cup NOR and absolute CDR error metrics (Table IV). The relative improvement of CNN-FE against RCN-CD, PPF and PBGC are 14.2%, 40.9% and 65.5% on the absolute CDR error metric, and 13.7%, 28.7% and 65.2% on the NOR metric respectively.

	CNN-FE	CNN-OR	RCN-GTD	RCN-CD	PPF	PBGC
Mean δ	0.0608	0.0754	0.0740	0.0709	0.1028	0.1764
Std dev δ	(± 0.0503)	(± 0.0851)	(± 0.0578)	(± 0.0582)	(± 0.0853)	(± 0.1120)
$\kappa_W \delta$	0.707	0.590	0.586	0.609	0.155	0.130
Mean m_1	0.2302	0.2400	0.2729	0.2667	0.3230	0.6617
Std dev m_1	(± 0.1326)	(± 0.1459)	(± 0.1240)	(± 0.1220)	(± 0.1731)	(± 0.1613)
$m_1 \leq 0.30$	0.793	0.759	0.671	0.655	0.547	-
$m_1 \leq 0.25$	0.692	0.664	0.495	0.464	0.427	-
$m_1 \leq 0.20$	0.524	0.496	0.292	0.270	0.237	-
$m_1 \leq 0.15$	0.304	0.269	0.116	0.107	0.082	-
$m_1 \leq 0.10$	0.052	0.050	0.022	0.021	0.013	-
$m_1 \leq 0.05$	0.002	0.000	0.002	0.004	0.000	-

Table IV: Optic cup NOR & absolute CDR error results, SEED-DB

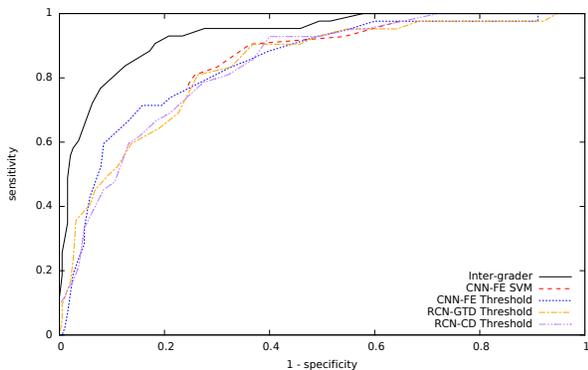


Figure 7: Glaucoma screening ROC curve

The reconstruction-based localization method indeed performs well in general. However, it does not accommodate adequately the challenging cases where the optic cup pallor is not well-defined.

We observe that the prior-based graph-cut method is strongly affected by vessel boundaries. This could be expected as an intensity-based GMM cannot mitigate for the neighbourhood context (i.e. a pixel of a certain pallor may be found within the cup in a dark image, but within the rim in a bright image). While this can be mitigated to a certain extent by histogram equalization, the pixel classification remains less reliable, especially given colour variance in real-life datasets. This is compensated somewhat with the inclusion of cup and rim priors for the physiologically-plausible features method, which however causes the segmentation result to be sensitive to the input being accurately centered.

Since the desired end-product is a screening model that should remove glaucoma-free images from further consideration, while retaining as many possible glaucoma cases as possible, we perform four-fold cross-validation using radial basis function support vector machines (SVM) on two features: the predicted vertical cup-to-disc ratios, and the

predicted disc height. The receiver operating characteristic (ROC) curves obtained from the SVM classifier on features from CNN-FE, from simple thresholds on CDR from CNN-FE, RCN-GTD and RCN-CD, as well as from the provided inter-grader agreement rate, are shown in Figure 7.

Overall CNN-FE screening performance ($AUC = 0.847$) is superior to the reconstruction-based method ($AUC = 0.838$), but remains somewhat below the inter-grader agreement rate ($AUC = 0.936$), due to several of the most challenging glaucoma images. The improved accuracy on CDR has translated to improved performance at sensitivity levels between about 0.4 and 0.75, as compared to reconstruction-based localization. The lower NOR error rate does not have a direct impact on screening as of yet, but could be helpful in any follow-up inspections by human graders.

V. CONCLUSION

We have presented an accurate and reliable method for segmenting the optic disc and optic cup regions from retinal images, that outperforms all other known methods on the commonly-used absolute vertical cup-to-disc error and non-overlap ratio metrics, as demonstrated on two large databases of real-world images. This is essential both for the proposed system to serve as a screening tool, and as an aid that provides an initial suggested boundary for human graders. Further work is planned in investigating the incorporation of additional features, such as parapapillary atrophy.

ACKNOWLEDGMENTS

This research was supported by SERI grant R-252-000-396-490. The authors thank Deng Fanbo for his early implementation of the vessel kink algorithm, the graders at the Singapore Eye Research Institute for annotating the SEED-DB database, the Messidor program partners for their provision of the MESSIDOR database (<http://messidor.crihan.fr/>), as well as the Control and Robotics Research Group of the University of Huelva for their optic disc ground truth annotations for the MESSIDOR database (<http://www.uhu.es/retinopathy/>).

REFERENCES

- [1] J. Nayak, U. R. Acharya, P. S. Bhat, N. Shetty, and T.-C. Lim, "Automated diagnosis of glaucoma using digital fundus images," *J. Med. Syst.*, vol. 33, no. 5, pp. 337–346, 2009.
- [2] R. Bock, J. Meier, L. G. Nyúl, J. Hornegger, and G. Michelson, "Glaucoma risk index: Automated glaucoma detection from color fundus images," *Med. Imag. Anal.*, vol. 14, no. 3, pp. 471–481, 2010.
- [3] M. R. K. Mookiah, U. R. Acharya, C. Chakraborty, L. C. Min, E. Y. Ng, and J. S. Suri, "Automated glaucoma identification using retinal fundus images: A hybrid texture feature extraction paradigm," *Image Analysis and Modeling in Ophthalmology*, pp. 9–22, 2014.
- [4] B. E. Klein, Y. L. Magli, K. A. Richie, S. E. Moss, S. M. Meuer, and R. Klein, "Quantitation of optic disc cupping," *Ophthalmol.*, vol. 92, no. 12, pp. 1654–1656, 1985.
- [5] D. W. K. Wong, J. Liu, N. M. Tan, F. Yin, and T. Y. Wong, "Automatic detection of the optic cup using vessel kinking in digital retinal fundus images," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2012, pp. 1647–1650.
- [6] K. G. Goh, W. Hsu, M. L. Lee, and H. Wang, "Adris: an automatic diabetic retinal image screening system," *Studies in Fuzziness and Soft Computing*, vol. 60, pp. 181–210, 2001.
- [7] P. Pallawala, W. Hsu, M. L. Lee, and K.-G. A. Eong, "Automated optic disc localization and contour detection using ellipse fitting and wavelet transform," in *ECCV 2004*, 2004, pp. 139–151.
- [8] A. Osareh, M. Mirmehdi, B. Thomas, and R. Markham, "Comparison of colour spaces for optic disc localisation in retinal images," in *ICPR 2002*, vol. 1, 2002, pp. 743–746.
- [9] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher *et al.*, "Optic nerve head segmentation," *IEEE Trans. Med. Imag.*, vol. 23, no. 2, pp. 256–264, 2004.
- [10] J. Xu, O. Chutatape, E. Sung, C. Zheng, and P. Chew, "Optic disk feature extraction via modified deformable model technique for glaucoma analysis," *Pattern Recognit.*, vol. 40, no. 7, pp. 2063–2076, 2007.
- [11] A. Aquino, M. E. Gegúndez-Arias, and D. Marín, "Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques," *IEEE Trans. Med. Imag.*, vol. 29, no. 11, pp. 1860–1869, 2010.
- [12] F. Yin, J. Liu, D. W. K. Wong, N. M. Tan, C. Y. Cheung, M. Baskaran *et al.*, "Automated segmentation of optic disc and optic cup in fundus images for glaucoma diagnosis," in *CBMS 2012*, 2012, pp. 1–6.
- [13] A. Giachetti, L. Ballerini, and E. Trucco, "Accurate and reliable segmentation of the optic disc in digital fundus images," *J. Med. Imag.*, vol. 1, no. 2, pp. 024001–024001, 2014.
- [14] B. Dashtbozorg, A. M. Mendonça, and A. Campilho, "Optic disc segmentation using the sliding band filter," *Comp. Biol. Med.*, vol. 56, pp. 1–12, 2015.
- [15] J. Liu, D. W. K. Wong, J. H. Lim, H. Li, N. M. Tan, Z. Zhang *et al.*, "Argali: an automatic cup-to-disc ratio measurement system for glaucoma analysis using level-set image processing," in *13th Int. Conf. Biomed. Eng.*, 2009, pp. 559–562.
- [16] Z. Zhang, J. Liu, N. S. Cherian, Y. Sun, J. H. Lim, D. W. K. Wong *et al.*, "Convex hull based neuro-retinal optic cup ellipse optimization in glaucoma diagnosis," in *Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2009, pp. 1441–1444.
- [17] G. D. Joshi, J. Sivaswamy, and S. Krishnadas, "Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment," *IEEE Trans. Med. Imag.*, vol. 30, no. 6, pp. 1192–1205, 2011.
- [18] Y. Zheng, D. Stambolian, J. OBrien, and J. C. Gee, "Optic disc and cup segmentation from color fundus photograph using graph cut with priors," in *MICCAI 2013*, 2013, pp. 75–82.
- [19] M. D. Abràmoff, W. L. Alward, E. C. Greenlee, L. Shuba, C. Y. Kim, J. H. Fingert *et al.*, "Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features," *Invest. Ophth. Vis. Sci.*, vol. 48, no. 4, pp. 1665–1673, 2007.
- [20] K. Stapor, A. Brueckner, and A. Świtoński, "Mathematical morphology and support vector machines for diagnosis of glaucoma on fundus eye images," in *Comp. Vis. Graph.*, 2006, pp. 888–893.
- [21] Y. Xu, D. Xu, S. Lin, J. Liu, J. Cheng, C. Y. Cheung *et al.*, "Sliding window and regression based cup detection in digital fundus images for glaucoma diagnosis," in *MICCAI 2011*, 2011, pp. 1–8.
- [22] J. Cheng, J. Liu, Y. Xu, F. Yin, D. W. K. Wong, N. M. Tan *et al.*, "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1019–1032, 2013.
- [23] D. W. K. Wong, J. Liu, N. M. Tan, F. Yin, C. Y. Cheung, M. Baskaran *et al.*, "An ensembling approach for optic cup detection based on spatial heuristic analysis in retinal fundus images," in *Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2012, pp. 1426–1429.
- [24] D. Xiao, J. Lock, J. M. Manresa, J. Vignarajan, M.-L. Tay-Kearney, and Y. Kanagasingam, "Region-based multi-step optic disk and cup segmentation from color fundus image," in *SPIE Med. Imag.*, 2013, pp. 86702H–86702H.
- [25] Y. Xu, S. Lin, D. W. K. Wong, J. Liu, and D. Xu, "Efficient reconstruction-based optic cup localization for glaucoma screening," in *MICCAI 2013*, 2013, pp. 445–452.
- [26] Q. P. Lau, M. L. Lee, W. Hsu, and T. Y. Wong, "The singapore eye vessel assessment system," *Image Analysis and Modeling in Ophthalmology*, pp. 143–160, 2014.
- [27] D. C. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," in *MICCAI 2013*, 2013, pp. 411–418.
- [28] D. C. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," in *NIPS 2012*, 2012, pp. 2843–2851.
- [29] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [30] M. B. Merickel Jr, X. Wu, M. Sonka, and M. D. Abràmoff, "Optimal segmentation of the optic nerve head from stereo retinal images," in *Med. Imag.*, 2006, pp. 61433B–61433B.
- [31] M. Lalonde, M. Beaulieu, and L. Gagnon, "Fast and robust optic disc detection using pyramidal decomposition and hausdorff-based template matching," *IEEE Trans. Med. Imag.*, vol. 20, no. 11, pp. 1193–1200, 2001.
- [32] G. D. Joshi, J. Sivaswamy, K. Karan, and S. Krishnadas, "Optic disk and cup boundary detection using regional information," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2010, pp. 948–951.
- [33] S. Hanson, S. K. Krishnan, and J. Phillips, "Observer experience and cup: disc ratio assessment," *Optom. Vis. Sci.*, vol. 78, no. 10, pp. 701–705, 2001.