31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Intelligent Fusion of Cup-to-Disc Ratio Determination Methods for Glaucoma Detection in ARGALI

D.W.K. Wong, J. Liu, J. H. Lim, N.M. Tan, Z. Zhang, S. Lu, H. Li, M.H. Teo, K.L. Chan, T.Y. Wong

Abstract—**Glaucoma is a leading cause of permanent blindness. ARGALI, an automated system for glaucoma detection, employs several methods for segmenting the optic cup and disc from retinal images, combined using a fusion network, to determine the cup to disc ratio (CDR), an important clinical indicator of glaucoma. This paper discusses the use of SVM as an alternative fusion strategy in ARGALI, and evaluates its performance against the component methods and neural network (NN) fusion in the CDR calculation. The results show SVM and NN provide similar improvements over the component methods, but with SVM having a greater consistency over the NN, suggesting potential for SVM as a viable option in ARGALI.**

I. Introduction

GLAUCOMA is the second leading cause of permanent, irrecoverable blindness, accounting for 13.4% of visual impairment globally[1]. It has been estimated that glaucoma will affect an estimated 60 million individuals by 2010, and up to 80 million by 2020[2], with over half in Asia[3]. In Singapore, a recent study of the Chinese [4] and Malay [5] populations has reported an age-standardized prevalence rate of more than 3% for the population aged 40 and above. The report also corroborated well with other population-based studies showing a high correlation of the disease with age, with a ten-fold increase in prevalence for a person in his 70s compared to a person in his 40s. The high correlation with age is a serious social implication in many societies where the population is aging and living longer.

In glaucoma, early progression of the disease often goes unnoticed by the individual due to the pattern of glaucomatous visual loss which proceeds inwards from the peripheral visual field of view. As such, when visual loss is noticeable by the individual, severe progression of the disease has already occurred, with irrecoverable sight loss. This emphasizes the need for a screening system for the early detection of glaucoma to allow timely intervention in order to save sight. However, current glaucoma detection methods are manual and impractical for screening.

An essential pathological characteristic in the progression of glaucoma is the increase in the size of the optic cup with respect to the optic disc. The optic disc, or optic nerve head, is the region in the retina where the optic nerve fibers, or ganglion cells, aggregate to form the optic nerve, which is connected to the brain. Within this optic disc is a depression known as the optic cup. During glaucomatous progression, the death of the ganglion nerve cells leads to increased excavation of the optic cup, and a corresponding increase in the optic cup to disc ratio (CDR). The CDR is thus a vital indicator of glaucomatous neuropathy as it is indicative of actual pathological changes in the retina during glaucoma.

We are currently developing ARGALI, an Automatic cup-to disc Ratio measurement system for Glaucoma detection and AnaLysIs, which automatically utilizes the CDR to assess the risk of glaucoma using retinal fundus images without manual grading. Unlike other reported work [6,7], ARGALI is not reliant on stereoscopic images, facilitating its use in widely available and low-cost retinal fundus cameras for mass screening.

In this paper, we report on the work done on ARGALI, in particular the comparing the performance of support vector machines (SVM) to a neural network for intelligent fusion of the component methods in ARGALI. In Section I, we have given an overview of the background and motivation for ARGALI. Subsequently, in Section II, we briefly describe the ARGALI framework and how intelligent fusion is to be utilized within the system. Section III reviews SVM, in particular its application in regression. The experiments carried out to evaluate the performance of SVM and a neural network, including comparisons to the individual methods are described in Section IV, and lastly, Section V presents the conclusions drawn from this paper.

II. ARGALI

The ARGALI system [8] is motivated by the need to develop a feasible, automated system to facilitate screening for the early detection of glaucoma. In ARGALI, various methods are utilized for the determination of the optic cup and disc from retinal images in order to calculate the CDR. In this section, the basic ARGALI framework, as shown in Figure 1, will be briefly described. A fuller discussion of the component modules can be found in [9, 10].

The initial step involves locating a region of interest (ROI) in the retinal fundus image containing the optic disc. By limiting succeeding analysis to this ROI, a considerable reduction in the computational resources utilized is achieved, as the ROI is typically less than 15% of the total

J.Liu, D.W.K. Wong, J.H. Lim, N.M. Tan, Z. Zhang, S. Lu and H. Li are with the Institute for Infocomm Research, A*STAR, Singapore (phone: (065)6408-2526; fax: (065)6776-1378; e-mail:wkwong@ i2r.a-star.edu.sg).

M.H. Teo and K.L. Chan are with the School of Electronic and Electric Engineering, Nanyang Technological University, Singapore.

T.Y. Wong is with the Singapore Eye Research Institute and the National University of Singapore.

retinal fundus image size. The identification of the ROI involves patch analysis of the fundus image based on pixel intensities, followed by an expansion of the identified candidate patch into the ROI.

Subsequently, a variational level set approach is then applied to this ROI to determine the optic disc, typically in the red channel where the visibility of the optic disc is greater. However, the detected contour is potentially affected by the presence of blood vessels traversing across the optic disc edges, resulting in an uneven contour. This unevenness can be mitigated through the application of direct ellipse fitting on the detected contour.

Next, level set is applied to the region within the optic disc for the detection of the optic cup, where accurate segmentation is more difficult due to the denser vascular architecture within the optic disc, and decreased visibility of the optic cup boundary. An alternate method via histogram-based analysis of the color pixel intensity is also employed for optic cup segmentation. Smoothing of the cup contours

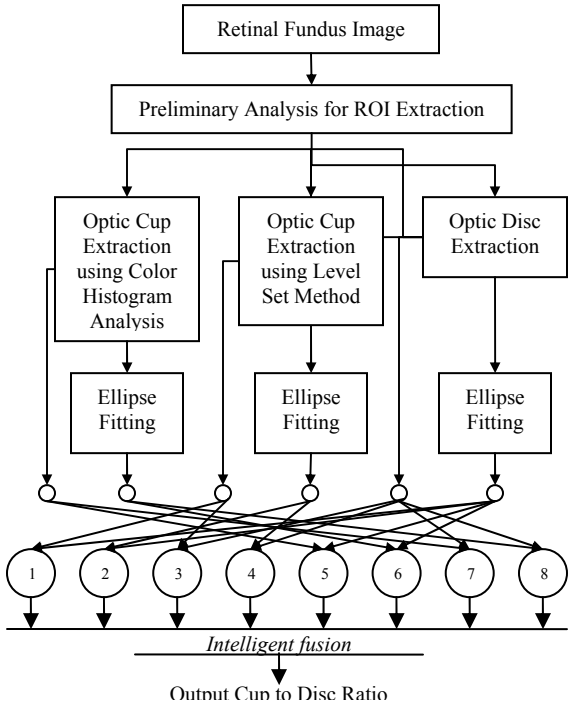

Figure 1. Overall framework presenting the different methods for cup to disc ratio determination and intelligent fusion in ARGALI. Component methods of the numbered circles (*G*) are described in Table 1 below.

TABLE I
COMBINATION OF METHODS FOR CDR DETECTION

G	Disc Detection	Disc Ellipse Fitting	Cup Detection	Cup Ellipse Fitting
1	Variational level set	Yes	Threshold level set	Yes
2	Variational level set	Yes	Threshold level set	None
3	Variational level set	None	Threshold level set	Yes
4	Variational level set	None	Threshold level set	None
5	Variational level set	Yes	Color intensity	Yes
6	Variational level set	Yes	Color intensity	None
7	Variational level set	None	Color intensity	Yes
8	Variational level set	None	Color intensity	None

detected through these methods is similarly performed using direct ellipse fitting.

Finally, CDR values calculated using different combinations of the component methods, as shown in Table I, are determined and fused. Due to the variability of image quality and the presence of ocular disease pathologies in retinal fundus images, the performance of each combination was found to vary, with no single combination being most favourable for every image. To make the best use of all available methods and to optimally merge the results, an intelligent adaptive network based on a neural network architecture was adopted. Training of the neural network would allow the network to learn the optimal fusion of inputs, and incorporate clinical knowledge. However, although extremely powerful, neural networks can be a challenge to use due to potential over-fitting issues and the possibility of being trapped in local minima.

III. SVM

Support vector machines (SVM) are an increasingly popular tool for classification and regression problems with widespread data mining and analysis applications such as in computer vision and bioinformatics. SVM is a supervised learning method that aims to determine the boundaries or hyperplanes of input data to distinguish between two or more classes. In a two-class classification scenario, two parallel hyperplanes are constructed on each side of the hyperplane which separates the data. The goal of SVM then is a quadratic problem to determine the optimal generalizing hyperplane which maximizes the margin between the two classes, and the identification of support vectors which constrain this margin. An important feature of SVM is the ability to select a kernel to transform non-linear data into a higher dimensional feature space for improved data separation. Some of the key features of SVM over neural networks are improved generalization due to maximization of the separating margin, avoidance local minima through global solutions to the quadratic problem, and improved tolerance or rejection outliners through the cost function *C*.

In Support vector regression (SVR), the input data is first mapped into a higher dimension feature space using non-linear kernel, and subsequently regression is performed in this space. Given a set of inputs $\{(\mathbf{x}_1, y_1),\ldots(\mathbf{x}_n, y_n),\}$, where $\mathbf{x_n}$ is the input vector and y_n is the target output, the goal then is determine the regression function f(x) which can predict the outputs corresponding to new input vectors

$$f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\phi(x) + b \tag{1}$$

where **w** and *b* are coefficients. In particular, εSVR [10] specifies an upper limit *L* to the difference ε between the mapped values *f(***x***)* and the target values *y* via an ε-insensitive loss functional, described as

$$L = \begin{cases} |f(\mathbf{x}) - y| - \varepsilon & |f(\mathbf{x}) - y| \geq \varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

This leads to the standard SVR formulation by Vapnik [11]

$$\begin{aligned} \min_{\mathbf{w},b,\xi,\xi^*} \quad & \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \\ \text{subject to} \quad & \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}_i) + b - y_i \le \varepsilon + \xi_i \\ & y_i - \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}_i) - b \le \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \ge 0, i = 1,\ldots, n. \end{aligned} \tag{3}$$

where C is the empirical cost function, ξ,ξ^* are the slack variables to tolerate some errors, analogous to the soft margin approach. The mapping or kernel function is represented by $\phi(\cdot)$. The solution of the optimization problem in (3) for parameters w and b will allow future outputs to be predicted using (1).

IV. Experiments

To compare the performance in the use of SVR against that of neural network for the fusion of the outputs obtained in the ARGALI system, an experiment was conducted using a set of retinal fundus images from the Singapore Eye Research Institute (SERI), collected as part of the recently concluded Singapore Indian Malay Eye Study (SIMES) [5]. The fundus images were from healthy individuals from the normative population without glaucoma. Each retinal fundus image has an image size of 3072x2048, and had the cup-to-disc ratio (CDR) individually measured by a senior ophthalmologist as part of the SIMES protocol. The values of the measured CDR ranged from 0.2CDR units to 0.65CDR units. This CDR value was set as ground truth, CDR_{GT}, and was used as the clinical standard against which the performances of the various methods are compared to.

The ARGALI system was then used to segment the optic cup and disc from the sample set of retinal fundus images using a variety of approaches, including level-set based methods, histogram-based methods, and smoothing of the contours via a direct ellipse-fitting method. An example of the segmented cup and disc using the various approaches is shown in Figure 2. The cup and disc outputs from each method were then combined in a certain fashion, as described in Table I, to generate the eight inputs into the fusion networks for each retinal image, and are labeled CDR_1 to CDR_8, corresponding to the combinations in Table 1. Subsequently, the retinal images were divided into a set of 100 images for training the SVR and neural network models, with the remaining 40 images used as a test set for the overall evaluation of the performance of the system.

For implementation of SVR, we adopted libsvm [12], a popular toolbox for support vector classification and regression. In the use of support vectors for the classification and regression, an important parameter is the selection of an appropriate kernel for mapping of inputs into another space for separation of the data. The radial basis function was selected for our use due to its flexibility in allowing non-linear mapping into a higher-dimensional space. Subsequently, we optimized the model parameters through a grid search method, and selected the optimally performing parameters through 5-fold cross validation of the model, using the training set of images. Finally, the SVR model was applied onto the test set, and the outputs of the model were termed as CDR_{SVR}.

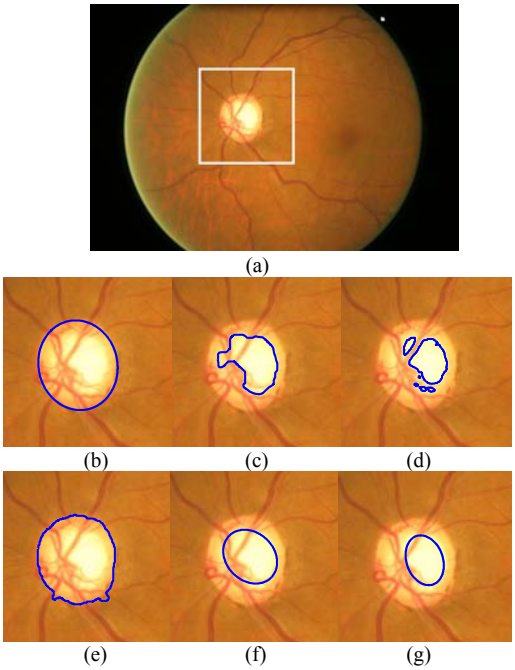

Figure 2. (a) Input retinal fundus image with region of interest indicated by a white border; (b) Disc segmentation using variational level set; (c) Cup segmentation using threshold-initialized level set; (d) Cup segmentation using color histogram analysis; (e), (f) and (g) are the ellipse-fitted refinements of the contours determined in (b), (c) and (d) respectively.

In a similar fashion, multiple experiments using the same training set were also run to optimize the neural network. The optimal network configuration was found to be a single hidden layer architecture with 6 neurons, 8 input neurons and 1 output neuron. Gradient descent method with momentum backpropagation was specified as the learning method, and the performance of the network was determined using a mean squared error function. The individual numerical parameters were also optimized in the process. Using this trained neural network, the test set was processed and the corresponding results were labeled as CDR_{NN}.

After obtaining the CDR values from the fusion networks, the results were compared through calculating the error ε with respect to the clinical ground truth CDR_{GT} i.e. $E = \left|CDR_m - CDR_{GT}\right|, m \in \{1,\ldots 8, NN, SVM\}$. The errors were then classified into three different categories E <0.1, E <0.15 and E <0.2, where the maximum acceptable error is 0.2 corresponding to reported inter-observer variability for CDR determination by human observers [13]. Table II presents the results of this classification. It can be seen from the table that the errors calculated from CDR obtained from the individual methods are higher than those obtained using the fusion networks, with the percentage of acceptable CDR performance ranging from 40% to 80%, and the mean error from ranging from 0.1143 to 0.2362 CDR units. In contrast, the results obtained using the neural

TABLE II
PERFORMANCE OF INDIVIDUAL METHODS AND FUSION NETWORKS FOR CUP-TO-DISC RATIO DETERMINATION

G	Error *E* (CDR units)			Mean Error (CDR units)	Mean Error Range (CDR units)
	<0.1[a]	<0.15[b]	<0.2[c]		
1	10 (25%)	19 (47.5%)	23 (57.5%)	0.1674	0.0576
2	8 (20%)	9 (22.5%)	16 (40%)	0.2362	0.0705
3	17 (42.5%)	25 (62.5%)	31 (77.5%)	0.1349	0.0655
4	11 (27.5%)	14 (35%)	21 (52.5%)	0.1947	0.0808
5	20 (50%)	26 (65%)	31 (77.5%)	0.1143	0.0824
6	15 (37.5%)	25 (62.5%)	28 (70%)	0.1682	0.1488
7	18 (45%)	27 (67.5%)	32 (80%)	0.1165	0.0821
8	17 (42.5%)	25 (62.5%)	29 (72.5%)	0.1517	0.1389
NN	22 (55%)	31 (77.5%)	35 (87.5%)	0.1053	0.0482
SVM	20 (50%)	31 (77.5%)	37 (92.5%)	0.1076	0.0219

The groupings *G* are as in Table I. [a,b,c]Figures indicate number of images within that error category; figures in parentheses indicate the percentage against the test set size of 40 retinal images.

network and SVM report substantially better performance, with 87.5% and 92.5% of results within the inter-observer variability, and with similar mean error rates of 0.1053 and 0.1076. Furthermore, the fusion networks were also found to be able to better detect the CDR with a greater accuracy, which can be seen by the higher proportion of CDR results in the E<0.1 and E <0.15 classes.

Another important means of evaluating the methods and systems is the range of the calculated error, which can give a measure of the consistency and reliability of the calculated CDR. The average error range, calculated by taking the mean of the range of errors across the clinical CDR, was determined and is also presented in Table II. It can be seen that the fusion systems have a smaller error range then the individual methods. Furthermore, the error range for SVM is less than that for the neural network, suggesting that SVM has a potentially more consistent performance as compared to the neural network for data fusion in ARGALI.

V. CONCLUSIONS

Glaucoma is a leading cause of blindness, and it is vital for early detection in order to save sight. ARGALI is being developed as a leading system employing multiple techniques to extract the optic cup and disc to determine the cup to disc ratio, an important parameter for assessing the risk of glaucoma. In this paper, we compared the performance of SVR against a neural network in the fusion of the results from the CDRs obtained using the individual methods. It was found that the use of fusion strategies is key to improving CDR accuracy. Both SVR and the neural network were shown to perform comparably in determining the CDR, with similar performance in overall accuracy and quality of results, suggesting that SVR is a viable alternative to the use of a neural network in fusion of the individual methods. Furthermore, the experiments also indicate a lower error range for the SVR method, implying a better consistency of CDR determination. The results motivate further exploration of SVR for data fusion in ARGALI, perhaps resulting to an ensemble of SVR and a neural network for improved performance. Through this and other concurrent research, we hope to develop ARGALI as a clinical system which can be fielded in primary healthcare providers for the early detection of glaucoma.

REFERENCES

[1] D. Pascolini, S. P. Mariotti, G. P. Pokharel, R. Pararajasegaram, D. Etya'ale, A. D. Negrel, and S. Resnikoff, "2002 global update of available data on visual impairment: a compilation of population-based prevalence studies," *Ophthalmic Epidemiol*, vol. 11, 67-115, Apr 2004.

[2] H.A. Quigley and A.T. Broman, "The number of people with glaucoma worldwide in 2010 and 2020," *Br J Ophthalmol*., vol 90, 262-267, 2006.

[3] T. Y. Wong, S. C. Loon, and S. M. Saw, "The epidemiology of age related eye diseases in Asia," *Br J Ophthalmol*, vol. 90, 506-511, 2006.

[4] P. J. Foster, F. T. S. Oen, D. Machin, T.-P. Ng, J. G. Devereux, G. J. Johnson, P. T. Khaw, and S. K. L. Seah, "The Prevalence of Glaucoma in Chinese Residents of Singapore: A Cross-Sectional Population Survey of the Tanjong Pagar District," *Arch Ophthalmol*, vol. 118, 1105-1111, August 2000.

[5] S.Y. Shen, T.Y. Wong, P.J. Foster, J.L. Loo, M. Rosman, S.C. Loon, W.L. Wong, S.M. Saw, T. Aung, "The prevalence and types of glaucoma in malay people: the Singapore Malay eye study," *Invest. Ophth. Vis. Sci*., vol. 49, 3846-51, 2008.

[6] M.D. Abramoff, W.L.M. Alward, E.C.Greenlee, L. Shuba, C.Y. Kim, J.H. Fingert, Y.H. Kwon, "Automated Segmentation of the Optic Disc from Stereo Color Photographs Using Physiologically Plausible Features," *Invest. Ophth. Vis. Sci.* vol. 48, 1665-1673, April 2007.

[7] J. Xu, O. Chutatape, E. Sung, C. Zheng, P.C.T. Kuan, "Optic disk feature extraction via modified deformable model technique for glaucoma analysis," *Pattern Recognition*, vol. 40, 2063-2076, July 2007.

[8] J. Liu, D.W.K. Wong, J.H. Lim, H. Li, N.M. Tan, Z. Zhang, T. Y Wong, R. Lavanya, "ARGALI : An Automatic Cup-To-Disc Ratio Measurement System For Glaucoma Analysis Using Level-Set Image Processing", 13th International Conference on Biomedical Engineering (ICBME2008), 2008.

[9] J. Liu, D. W. K. Wong, J.H. Lim, X. Jia, F. Yin, H. Li, W. Xiong, T.Y. Wong, "Optic cup and disk extraction from retinal fundus images for determination of cup-to-disc ratio," *Conf Proc IEEE ICIEA*, 1828-1832, 2008.

[10] D. W. K. Wong, J. Liu, J.H. Lim, X. Jia, F. Yin, H. Li, T.Y. Wong, Level-set based automatic cup-to-disc ratio determination using retinal fundus images in ARGALI," *Conf Proc IEEE Eng Med Biol Soc.,* 2266-2269, 2008

[11] V. Vapnik, *The Nature of Statistical Learning Theory*, Second Edition, Springer, New York, 2001.

[12] C.-C. Chang and C.-J. Lin, LIBSVM : a library for support vector machines, 2001.

[13] R. Varma, G.L. Spaeth, W.C. Steinmann and L.J. Katz, "Agreement between clinicians and an image analyzer in estimating cup-to-disc ratios," *Arch Ophthalmol*, vol. 107, 526-529, 1989.