An Effective Automated System for Grading Severity of Retinal Arteriovenous Nicking in Colour Retinal Images

Pallab Kanti Roy, Uyen T. V. Nguyen, Alauddin Bhuiyan and Kotagiri Ramamohanarao

***Abstract*— Retinal arteriovenous (AV) nicking is a precursor for hypertension, stroke and other cardiovascular diseases. In this paper, an effective method is proposed for the analysis of retinal venular widths to automatically classify the severity level of AV nicking. We use combination of intensity and edge information of the vein to compute its widths. The widths at various sections of the vein near the crossover point are then utilized to train a random forest classifier to classify the severity of AV nicking. We analyzed 47 color retinal images obtained from two population based studies for quantitative evaluation of the proposed method. We compare the detection accuracy of our method with a recently published four class AV nicking classification method. Our proposed method shows 64.51% classification accuracy in-contrast to the reported classification accuracy of 49.46% by the state of the art method.**

I. INTRODUCTION

Retinal arterivenous nicking (AV nicking or AVN) can be defined as the narrowing of venular calibre by a stiff artery at their crossing point in response to a rise in the blood pressure (i.e., hypertension) [1]. In colour retinal images, AV nicking appears as a decrease in venular calibre at both side of an artery vein cross-over point (Fig. 1). Different research studies show that, AV nicking is strongly associated with hypertension, systemic diseases and stroke [2], [3]. This implies the importance of the quantification of AV nicking to identify people who are at high risk of cardiovascular heart disease.

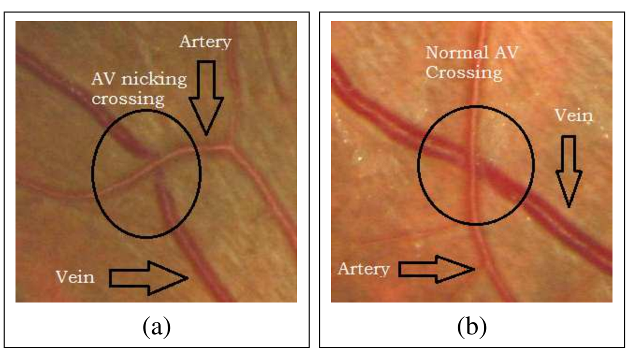


Fig. 1: Examples of AV crossing (a) AV nicking crossing; (b) normal AV crossing.

Pallab Kanti Roy is with the Department of Computing and Information Systems, The University of Melbourne, VIC 3010, Australia, e-mail: royp@student.unimelb.edu.au.

Uyen T. V. Nguyen is with the Department of Computing and Information Systems, The University of Melbourne, VIC 3010, Australia.

Alauddin Bhuiyan is with the Computational Informatics, Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia.

Kotagiri Ramamohanarao is with the Department of Computing and Information Systems, the University of Melbourne, Australia.

At present, AV nicking is manually graded in a subjective and qualitative manner, which is highly time consuming and depends on the grader's expertise. Another problem of manual grading is that more often the results are not reproducible. Therefore, an automated system for the quantification of AV nicking grading is highly required for large scale longitudinal studies.

A few works have been done on the quantification of AV nicking. Nguyen et al. [4] have proposed an automated method for the measurement of AV nicking, where a continuous value is produced to represent the severity of AV nicking. The continuous value provided by Nguyen's method is difficult to interpret. In addition, the widths are not normalized in this work. Therefore, the AV nicking severity score may vary depending on the variability of the vessel width's of different patient. In this method [4], vein widths are directly computed from the segmented image which may be influenced by the threshold value used in the binarization stages. More precise width computation which is not affected by the threshold value can produce more accurate AV nicking measurement. In this paper, we propose a fully automated method which can classify AV nicking from colour retinal imaging. The main contributions of the proposed method include:

- An fully automated AV nicking classification method, which will produce better information to predict cardiovascular diseases by providing discrete classification of AV nicking .
- A new width computation method is used for the precise estimation of venular widths.
- A novel feature vector is proposed for enhanced feature analysis and classification of AV nicking based on venular widths and Random Forest (RF) classifier.

We compare the classification accuracy of our method with the existing state of the art method where the classification is performed on a four class problem (0=normal to 3=most severe).

The rest of the paper is organized as follows. In Section II, the technical details of the proposed method are described. The dataset and experimental results are presented in Section III and Section IV to demonstrate the performance of the proposed method. Finally, the paper is concluded with Section V.

II. METHODOLOGY

The overall framework of the proposed method is illustrated in Fig. 2. The system takes retinal image as input and classifies the AV nicking severity level of crossover

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

point. The system is divided into two main modules: the Image Processing module and the Machine Learning module. Image Processing module includes the task of retinal vessel segmentation, AV crossover point detection, ROI selection, artery vein classification, separation of venular segment, vein edge and centreline refinement and width computation. On the otherhand, Machine Learning module includes the feature extraction from computed widhts, test and training data creation and classification. Detail of each module of the proposed method are presented in the following sub-sections of the paper.

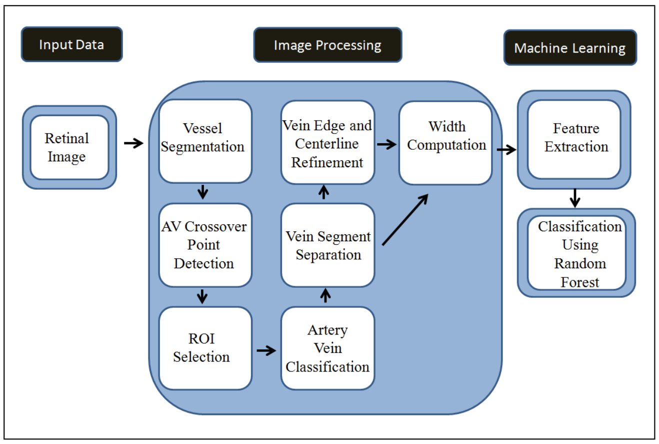


Fig. 2: The proposed AV nicking severity level classification method.

A. IMAGE PROCESSING

1) Vessel Segmentation , Artery Vein Classification, Venular Segment Separation: We have applied multi-scale line detection to segment the blood vessels and following this the vessel crossover points are detected and arteries and veins are classified [4] as shown in Fig. 3.

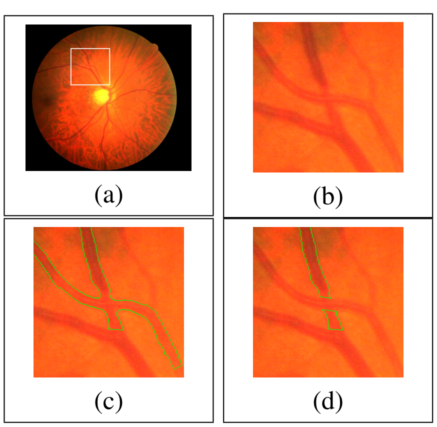


Fig. 3: Output images of the different steps of image processing module (a) original Image; (b) region of interest; (c) simplified segmentation of the AV crossover point; (d) detection of venular segments.

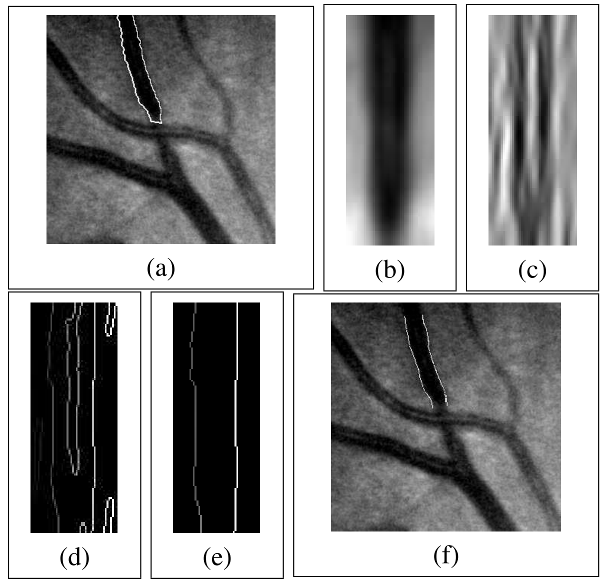


Fig. 4: Output images of the different steps of edge refinement process (a) initial segmentation of vein segment; (b) straightened vessel image after applying anisotropic Gaussian smoothing, that image is created by stacking many image profiles alongside one another; (c) subsequently another filter is applied to approximate the second derivative computed perpendicular to the vessel (i.e. horizontally); (d) here gray pixels represent positive-to-negative and white pixels represent negative-to-positive transitions; (e) the length of each connected line is computed and only the longest lines that fall close to the estimated vessel boundaries are preserved; (f) superimposition of the new edge on the original green channel image.

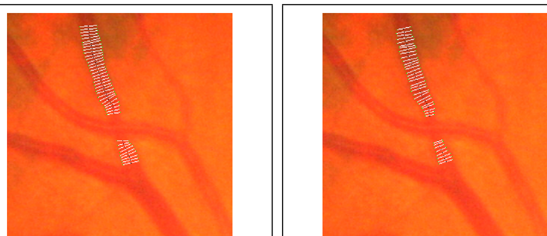
2) Refinement of the Edge and Centreline of Venular Segment: The width of the venular segment plays an important role in AV nicking measurement. To obtain more accurate measurements, the venular widths are computed based on its edge rather than the threshold given initial segmentation. The steps of the edge detection and final selection are as follows:

- The average width of a venular segment is computed from its intitial segmenation.
- A spline is fitted to smooth the centreline of the venular segment. The centreline is computed from the initial segmentation of venular segment.
- Using the Green channel image, the pixel intensity profile is computed by applying linear interpolation to find the normal points of spline fitted centreline points. The length of the intensity profile is set as double of the estimated average width so that it can cover the whole vessel cross-section. The intensity profiles are then stacked into one after another for the computational simplicity. The intensity values of these profiles are then smoothed by using a anisotropic Gaussian filter (Fig. 4b)
.
- To identify the edge points, we apply a filter proposed in [7] to compute the second derivative perpendicular

to the vessel (Fig. 4c). At the edge points there should be change in the sign of second derivative. It can be bright to dark (e.g. background to vessel) and dark to bright (vessel to background). These zero crossing lines are identified by using connected component labeling (Fig. 4d).

- Finally, the most longest lines which are close to the boundary of the intial segmentation is identified as the vein edge (Fig. 4e). Detected edges are superimposed on the green channel image in (Fig. 4f).

3) Width Computation: For width computation we draw a line perpendicular to each of the centreline pixel of the vein segment. The Euclidean distance of the intersection points of each perpendicular line and vein edge are estimated as the vein widths. Fig. 5 shows the vein widths computed from the initial vein segmentation and after the final detection of the edge points.



(a) (b)

Fig. 5: (a) Widths of the vein segment based on initial segmentation; (b) widths of the vein segment after the refinement of the edge.

B. MACHINE LEARNING

1) Feature Extraction: In case of AV nicking, there is a decrease in the venular widths near the crossover point compared to the mean width of the full vein segment. Based on this property, we create our feature set to classify AV nicking. Let W_i represent the vein widths which are sorted based on the Euclidean distance from the crossover point and Cr_n (Fig. 6) represent the mean width of first n centre points starting from the crossover point as $Cr_n = \frac{1}{n}\Sigma_{i=1}^{n} W_i$. If c (Fig. 6) is the middle point of the vessel centreline then width distribution of the overall vessel is computed as follows: $MW = \frac{1}{s}\Sigma_{i=c-s/2}^{c+s/2} W_i$, ($s$ is empirically set to 20). The difference between Cr_n and MW at different scale is used to classify AV nicking.

$$F_n = \frac{MW - Cr_n}{MW}, where, n = [10, 20, 30]$$

In this equation, the difference between MW and Cr_n is normalized by MW, which makes our features less sensitive to the variation of vein widths. We extract these features from the vein widths given by initial segmentation and edge refined segmentation. Along with the width difference at three different scales, our feature set also includes mean and standard deviation of the widths of the vein segment on both sides of the crossover point.

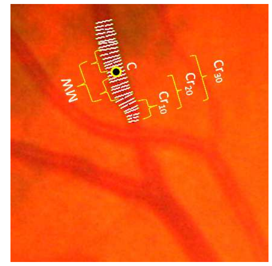


Fig. 6: Illustration of the feature extraction process. Here Cr_{10}, Cr_{20} and Cr_{30} represent mean width of the first 10, 20 and 30 center points from the crossover point. MW represent the mean width computed in 20 pixels window from the middle (represented by C) of vein centerline.

2) Classification: We use Random Forest (RF) [5] to classify the severity of each AV crossover point. Given a set of training data consisting of N samples each of which is a $D-dimensional$ feature vector labelled as belonging to one of C classes, a random forest contains a set of tree predictors created from the training data. Each tree in the forest is built from a bootstrap sample of the training data (that is, a set of N samples chosen randomly, with replacement, from the original data). The trees are built using a standard classication and regression tree (CART) algorithm [6]. However, rather than assessing all $D-dimensions$ for the optimal split at each tree node, only a random subset of $d < D-dimensions$ is considered. The trees are built to full size (i.e. until a leaf is reached containing samples from only one class) and are not pruned. During classification, unseen feature vectors are classified independently by each tree in the forest; each tree casts a unit class vote, and the most popular class can be assigned to the input vector. We empirically set the number of trees as 500 in our method.

III. DATASET

We use 47 high-resolution retinal images for the evaluation of our study. The images are obtained from Blue Mountain Eye Study (BMES) [8] and Singapore Malay Eye Study (SiMES) [9]. The resolution of the retinal images of BMES and SiMES study is 3888×2592 and 3072×2592. Images in both studies are converted to 24-bit (8 bit for each color space respectively red, green and blue without any enhancement).

From these images, 93 detected crossover points are selected to evaluate the performance of the proposed method. Each crossover point was manually graded by two experts at the Centre for Eye Research Australia (Melbourne, Australia) using 4 scale grading system (from 0=normal to 3=most severe). Disagreement between graders are then reassessed in a joint discussion, which resulted in a single grading for each crossover point.

IV. EXPERIMENTAL RESULTS

For each AV crossover point we extract 8 features to train the Random Forest (RF) classifier as discussed in section II-B.1. Since, the classes are not evenly distributed, we use

TABLE I: Confusion matrix for the predicted classification of Nguyen et al. [4]. The rows represent manual classification while columns represent automatic classification.

Actual/Predicted	0	1	2	3	Total
0	36	12	2	1	51
1	9	3	5	0	17
2	4	1	2	7	14
3	1	0	5	5	11
# Correct Pre.	36	3	2	5	46
Accuracy	70.58%	17.64%	14.28%	45.45%	49.46%

TABLE II: Confusion matrix for the predicted classification of proposed method. The rows represent manual classification while columns represent automatic classification.

Actual/Predicted	0	1	2	3	Total
0	40	10	1	0	51
1	8	5	2	2	17
2	4	1	6	3	14
3	1	0	1	9	11
# Correct Pre.	40	5	6	9	60
Accuracy	78.43%	29.41%	42.85%	81.81%	64.51%

leave-one-out cross-validation method for the evaluation of the prediction accuracy. For one test case, we consider all remaining cases as our training dataset. We follow similar strategy for Nguyen et al. [4]. The AV nicking severity score given by this method is used as a feature for the classification of AV nicking level. Table I and II shows the confusion matrix of the predicted classification of Nguyen et al. [4] and proposed method. In each AV nicking severity level, our method shows better classification accuracy compared to the current method [4]. Among the 93 crossover points our method correctly classifies 60 crossover points (accuracy rate is 64.51%) which is significantly higher compared to the classification accuracy of 49.46% shown by [4].

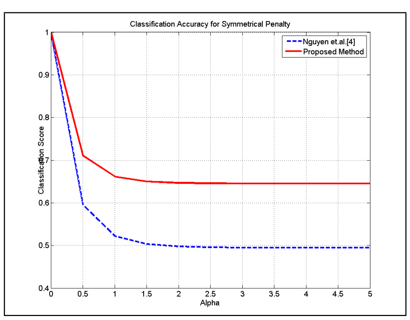


Fig. 7: Classification accuracy for symmetrical penalty.

As the classification of AV nicking level is a multi-class classification problem with some similarity between the classes (for example, severity level 0 is close to severity level 1). Because of this reason, we use a new measurement to compute how close our classification is to the manual classification. We give symmetric penalty for misclassification. The cost function is computed by Eq. 1 and the average classification score using this cost function is given by Eq. 2

$$Cost(\alpha, A_i, P_i) = e^{-\alpha \times (A_i+1) \times |A_i - P_i|} \quad (1)$$

$$S_1 = \frac{1}{n} \Sigma_{i=1}^{n} Cost(\alpha, A_i, P_i) \quad (2)$$

Here, we are assigning more penalty for the misclassification as the severity of AV nicking increase and penalties are given in a symmetrical way. Fig. 7 shows the classification accuracy S_1 score of two methods at different values of α. At $\alpha = 0$ we are giving no penalty for misclassification and hence, the accuracy is 1. As the value of α increases, the accuracy starts to drop. At $\alpha = 2$ we are getting the standard classification accuracy. For all the non-zero values of α, our method is showing better accuracy compared to Nguyen et al. [4].

V. CONCLUSION

In this paper, we proposed an effective method for the classification of AV nicking. The accuracy of the method was evaluated on 93 AV crossover points which were graded in 4 classes by two experienced ophthalmologists. In our proposed method, the vein widths are precisely computed using its edge and intensity information. The proposed method showed 64.51% detection accuracy in classifying AV nicking severity level (0=normal to 3=most severe) compared to 49.46% showed by the current method [4]. The proposed method provides a discrete quantification of AV nicking which can help to study the relationship between AV nicking and different diseases such as hypertension and stroke.

REFERENCES

[1] Tien Y. Wong and Paul Mitchell, Hypertensive retinopathy, New England Journal of Medicine, vol. 351, number. 22, pp. 2310-2317, 2004.

[2] Tien Wong and Paul Mitchell, The eye in hypertension, The Lancet, vol. 369, number. 9559, pp. 425-435, 2007.

[3] P. Mitchell, J J Wang, TY Wong, W Smith, R Klein and SR Leeder, Retinal microvascular signs and risk of stroke and stroke mortality, Neurology, vol. 65, number. 7, pp. 1005-1009, 2005.

[4] Uyen T. V. Nguyen, Alauddin Bhuiyan, Laurence A.F. Park, Ryo Kawasaki, Tien Y Wong, Jie Jin Wang, Paul Mitchell and Kotagiri Ramamohanarao, An Automated Method for Retinal Arteriovenous Nicking Quantification From Color Fundus Images, IEEE Transactions on Biomedical Engineering, vol. 60, number.11, pp. 3194–3203, 2013.

[5] Leo Breiman, Random forests, Machine Learning:ECML, pp. 359–370, 2004.

[6] Dan Steinberg and Phillip Colla, CART: classification and regression trees,The Top Ten Algorithms in Data Mining, pp. 179–201, 2009.

[7] Peter Bankhead, C Norman Scholfield, J Graham McGeown and Tim M Curtis, Fast retinal vessel detection and measurement using wavelets and edge location refinement, PloS one, vol.7, number. 3, pp. 324–325, 2012.

[8] S. Cugati, R.G. Cumming, W. Smith, G.Burlutsky, P. Mitchell and J. J. Wang, Visual impairment, Age-related macular degeneration, cataract and long-term mortality: The Blue Mountains Eye Study, Arch. Ophthalmol., Epidemiol., vol.125, no.7, pp. 917-924, 2007.

[9] A.W.P. Foong, S.M. Saw, J. L. Loo, S.Shen, S.C. Loon, M. Rosman, T. Aung, D. T. H. Tan, E. S. Tai and T. Y. Yong, Rantionale and methodology for population-based study of eye diseases in Malay people: The Singapore Malay study (SiMES), Ophthalm. Epidemiol., vol.14, pp. 25–35, 2007.