Hybrid Learning of Vessel Segmentation in Retinal Images

Worapan Kusakunniran [1], **Peeraphat Charoenpanich** [2], **Perapat Samunyanoraset** [3], **Sarocha Suksai** [4], **Sarattha Kanchanapreechakorn** [5], **Qiang Wu** [6], and **Jian Zhang** [7], Non-members

ABSTRACT

In this paper, a novel technique of vessel segmentation in retinal images is proposed, using a hybrid learning based approach. Unlike most of other existing methods, a double-layers segmentation technique combining supervised and instance learning steps is introduced to enhance a sensitivity score of segmenting retinal blood vessels. The supervised learning based approach alone may not cope with unseen patterns caused by intrinsic variations in shapes, sizes, and color intensities of blood vessels across different retinal images. Thus, in the proposed hybrid learning solution, the supervised learning part is adopted to compute initial seeds of segmented vessels. They are then fed into the instance learning part as an initial foreground to further learn specific characteristics of vessels in each individual image. In the supervised learning step, the support vector machine (SVM) is applied on three types of features including green intensity, line operators, and Gabor filters. While, the iterative graph cut is adopted in the instance learning step, with the pre-processing of morphological operations and watershed algorithm. The proposed method is evaluated using two well-known datasets including DRIVE and STARE. It shows the promising sensitivity scores of 82.6% and 82.0% on DRIVE and STARE datasets respectively, which outperforms other existing methods in the literature.

Keywords: Vessel segmentation, Retinal image, Hybrid learning, Supervised learning, Instance learning.

1. INTRODUCTION

This paper focuses on developing a technique of blood vessels segmentation in retinal images. Segmentation results of blood vessels can be further used in many analysis tasks for detecting abnormalities of vessels including a severe stage of diabetic retinopathy. Thus, it is important to have the technique for segmenting retinal vessels in an automatic way with a promising performance. Several related methods have been proposed to address this challenge [1]. The existing techniques of vessel segmentation can be classified into two groups including instance-learning and supervised-learning based approaches.

A key challenge of the retinal blood vessels segmentation is about variations in shapes, scales, sizes, and colors of the vessels themselves. Across different retinal images, pixels' intensities of blood vessels may be varied significantly. Even within an individual image, shapes and sizes of vessels can be varied on different locations and types (e.g arteries and veins). Especially on retinal images with signals of a disease, for example, severe non-proliferative and proliferative stages of diabetic retinopathy contain abnormalities of blood vessels. Thus, variations of retinal blood vessels could be very significant.

Starting with instance learning based approaches, the main difficulty is where preset parameters could not cope well with different characteristics in different retinal images. While, for supervised learning based approaches, the learned segmentation model could not work efficiently with unseen vessels' characteristics caused by variations mentioned above.

Lately, the CNN-based approaches, which also belong to the category of supervised learning, have been proposed. However, they could not breakthrough the limited performance of sensitivity scores, since they need a sufficiently more training data to address variations of blood vessels. This paper, therefore, proposes a hybrid learning solution to overcome limita-

[1] The author is with Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom, Thailand 73170, Phone(+66-2)441-0909, E-mail: worapan.kun@mahidol.edu (Corresponding author)
[2] The author is with Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom, Thailand 73170, Phone(+66-2)441-0909, E-mail: p.peeraphat@outlook.com
[3] The author is with Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom, Thailand 73170, Phone(+66-2)441-0909, E-mail: perapat1972@gmail.com
[4] The author is with Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom, Thailand 73170, Phone(+66-2)441-0909, E-mail: sarocha.suks@gmail.com
[5] The author is with Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom, Thailand 73170, Phone(+66-2)441-0909, E-mail: j.sarattha@gmail.com
[6] The author is with School of Electrical and Data Engineering, University of Technology Sydney, Sydney, Australia 2007, Phone(+61-2)9514-2000, E-mail: qiang.wu@uts.edu.au
[7] The author is with School of Electrical and Data Engineering, University of Technology Sydney, Sydney, Australia 2007, Phone(+61-2)9514-2000, E-mail: jian.zhang@uts.edu.au

tions of each learning category.

Examples of instance learning and supervised learning (including CNN) based approaches of the retinal blood vessels segmentation are reviewed below, to demonstrate how they are developed and work in different perspectives.

Using the instance-learning, for example, the paper [2] began with extracting image ridges. They were later grouped into sets of primitives to approximately form up straight-line elements which were parts of vessels. The paper [3] used pixels' intensities with vessel enhancement, as features. Then, the vessel segmentation was performed using the Self-Organizing Map (SOM) with the Otsu's method to estimate the neuron class in neural network as an unsupervised clustering method. The proposed method in [4] used Heidelberg Retina Tomograph (HRT) images. They contained typical double-edge shapes of retinal vessels with different qualities. The reconstruction of vascular structures in retinal images was done on the green channel with the two-stage approach. The dedicated detector was applied for certain widths of the vessels. In addition, the Gaussian function was used for helping to detect the center line of vessels. Then, the unbiased detector of curvilinear structures was applied to detect vessels.

In [5], it used the histogram equalization to enhance the contrast in the retinal image. The distance transform algorithm was then applied to create a distance map image to calculate the directions and the magnitudes of the vessels' gradients. These pieces of information were used to construct the graph. Then, the graph cut algorithm was applied to segment the graph into vessel and non-vessel parts.

The techniques in this first category have a key advantage of not requiring a training process and a large set of labeled data for training. However, they could be fairly sensitive to the intra-variations among the retinal images, which may include variations in color tones and intensities of vessels.

In the group of supervised-learning based approaches, for example, in [6], it used basic and orthogonal line operators and the green channel as features, without any pre-processing or enhancement applied on the input image. This was to preserve the vessel structure before being further analyzed. The line operators were calculated at twelve different angles. Also, vessels and backgrounds were more contrast in the green channel, when compared with the red and blue channels. Then, the SVM was used as the final classification tool. In [7] used seven features extracted based on the inverted green channel, Gabor Wavelet, and line operator. Similarly, the Bayesian and SVM were attempted as the classification tools.

In addition, the cellular neural networks with the line detection on the inverted green channel were used for the segmentation [8]. It could avoid the problems of noises and low contrast between vessels and backgrounds. The proposed method in [9] extracted features based on local area shape-based features combined with multi-scale local statistical features from the green channel. The three types of morphological operators were applied. First, the opening operation was to remove smaller bright details such as noises' points or pathological areas. Second, the closing operation was to enhance dark details which mainly contained vessels. Third, the difference between erosion and dilation was to enhance the edge information. The SVM was used to find the optimal hyperplane for segmenting vessels from backgrounds.

In [10], it applied the adaptive histogram equalization and the morphological operations in the pre-processing stage. Then, the local binary pattern and the gray level co-occurrence matrix were then applied on the pre-processed image. For the feature extraction, the energy, contrast, correlation, and homogeneity were used. The SVM was also used in this paper as the vessel classification. In [11], the 7D feature vector from the gray-level-based features and the moment invariants-based features were computed. Then, the neural network was used as the pixel-based classification tool. It also included the two post-processing steps which were filling of pixel-gaps and removing of false detections.

More recently, several methods using convolutional neural network (CNN) are proposed for solving the segmentation problem of vessels in retinal images. Their techniques are very similar relying on CNN, but with varieties of CNN architectures. In [12], the deformable U-Net is adopted, where deformable convolutional layers are added into the conventional neural networks. This could address the challenges of high-variations on shapes and scales of blood vessels. While, in [13], U-Net and Dense-Net structures are combined and used to construct the segmentation model. The rotation and image mirroring are adopted in the data augmentation process, where CLAHE algorithm is applied for the pre-processing. In [14], dense U-Net is used under an assumption that retinal vessels are tiny which could be learned effectively using patch-based learning strategy.

In [15], an encoder-decoder architecture is adopted into a fully convolutional deep neural network, using a dilated spatial pyramid pooling with multiple dilation rates. In [16], an encoder-decoder based octave convolution network is introduced for addressing the vessels segmentation with the challenges of variations in sizes and shapes. Also, in [17], a convolutional neural network with a multi-scale residual is adopted to solve the scale-variation of retinal vessels. In [18], five different models of deep convolutional neural networks are attempted using high-resolution patches. M-AlexNet provides the best segmentation result. The main limitation of CNN based methods is about requiring high computational resources and a sufficiently large set of labeled data for training the

segmentation model. Also, as mentioned in the experimental comparisons of this paper, they could achieve high specificity values, but not for sensitivity scores.

The challenge of techniques in this second category could be the overfitting problem, especially when there are not enough numbers of training data samples. This occurs often for the case of medical image analysis, where it is difficult to obtain a large number of medical images with related groundtruth. As a result, many small vessels could be missed from the segmentation. Moreover, the main advantage of this category is a well-trained model can be directly used without any additional learning steps or calculations.

Therefore, in this paper, the hybrid solution combining good aspects of both instance-learning and supervised-learning concepts is proposed to solve the segmentation problem of vessels in retinal images. The supervised-learning using SVM learns pixels of blood vessels based on features of green intensity, line operators, and Gabor filters. The color transfer technique is also adopted to normalize color tones of retinal images before using in the SVM learning process. Then, the detected vessel-pixels with high confident values are used as the initial seeds to be learned in the instance learning using graph cut. This could extend the segmentation results to cover remaining blood vessels in the retinal image, especially small branches of the vessels. The proposed method is validated on the two well-known datasets, namely DRIVE and STARE. It achieves the outstanding sensitivity performance, when compared with the other methods in the literature.

Motivations on adopting each technique in the proposed method are explained in this paragraph. The SVM is applied in this paper as the classification tool because it focuses on producing a single binary output of vessel and non-vessel. It also contains a small number of hyper-parameters and guarantee a global optimum [19]. In addition, the color transfer technique is adopted in this paper to normalize color tones of retinal images. This is important to reduce variations of retinal blood vessels for a learning process in the supervised learning step. Also, this makes it easier for the instance learning step to extend coverage areas of blood vessels. While, the Gabor filter is used as one of vessel features because it can describe vessel textures/patterns in various scales and orientations. Moreover, the iterative graph cut is used as a main tool in the instance learning step, due to its flexibility and ability to capture multiple characteristics of vessels within individual retinal images.

The rest of this paper is organized as follow. The details of the proposed method is explained in the section 2. The experimental results and discussions are summarized in the section 3. Then, the conclusion is drawn in the section 4.

2. PROPOSED METHOD

The figure 1 shows the overview framework of the proposed method hybridizing supervised and instance learning approaches for the coarse-level and fine-level segmentation of vessels in retinal images. In the preprocessing stage, the green channel is split and used in the further processes because it gives the best contrast between vessels and background, when compared with other two channels of red and blue. Then, the histogram equalization is applied to enhance the contrast. It begins with the construction of a color histogram of the green channel ($hist$), with 256 bins.

$$hist = \{h_i\}_{i=0}^{255} \tag{1}$$

Then, the accumulative color histogram ($C = \{c_j\}_{j=0}^{255}$) is computed as below.

$$c_j = \sum h_i{}_{i=0}^{j} \tag{2}$$

The pixel value p_j is then updated using the equation below.

$$round[\frac{c_j - c_0}{W \times H - c_0}(255)] \tag{3}$$

where W and H are width and height of the input image respectively, and 255 is the maximum value of the pixel intensity. The sample output of this enhancement process is shown in the figure 2.

Next, in the coarse-level segmentation stage, the supervised learning approach is applied to train the extracted high-level features including green intensities, basic-line operators, orthogonal-line operators and Gabor features, for recognizing vessels. These segmented vessels are used as pre-seeds in the fine-level segmentation stage based on the instance learning approach. The pre-seeds are refined and used as the initial foreground (i.e. vessels) in the iterative graph cut for finalizing the vessel areas. The detailed processes are explained in the following sub-sections.

2.1 Segmentation Level 1 using Supervised Learning (Coarse-Level Segmentation)

In this level, four types of features are extracted for each pixel in the retinal image. The first feature is the intensity of the green channel since it provides the best contrast information between vessels and background, when compared with the red and blue channels.

Due to the intra-variation of color statistics across different images, the green channels (I_g) of all images are normalized to match the statistic of the reference image (I_R). This process is to reduce such variation, using the color transfer based-approach [20]. Since this color-statistic normalization process is performed on the single color channel, it could be done on the channel directly without converting into the $l\alpha\beta$ color space. The process is defined as below.

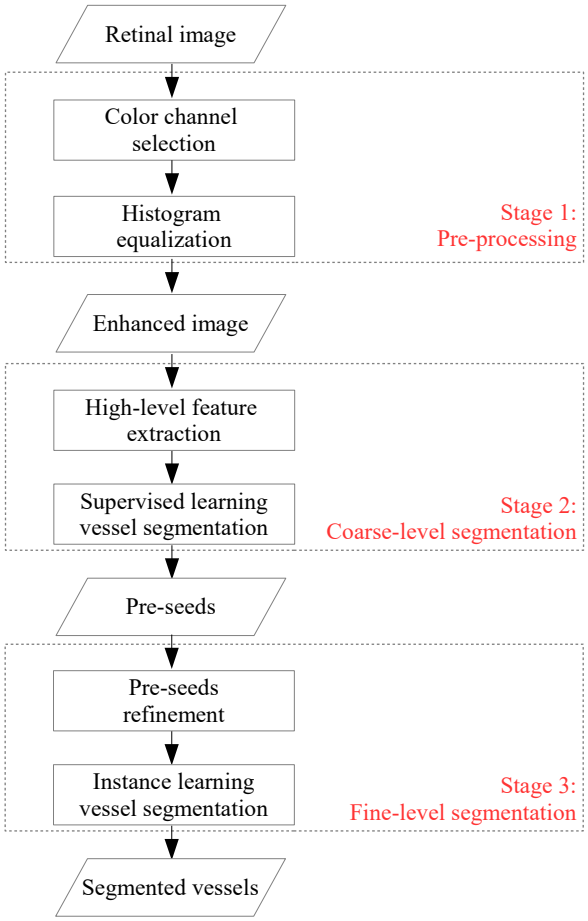

Fig.1: *The overview of the proposed method.*

$$\widetilde{I_g(p)} = (I_g(p) - \mu_g)\frac{\sigma_R}{\sigma_g} + \mu_R \tag{4}$$

where $I_g(p)$ is the pixel p of the image I_g, $\widetilde{I_g(p)}$ is the transferred value of $I_g(p)$, μ_g and μ_R are the means of I_g and I_R respectively, and σ_g and σ_R are the standard deviations of I_g and I_R respectively. Also, the transferred values must be trimmed up or down into the range of the green channel.

The two more features are extracted using the line operators [6, 7]. The second feature is computed based on the basic line operator. It is a line with fixed length centered at the considered pixel with 12 orientations spanning 360 degrees, including $0^o, 30^o, 60^o, 90^o, 120^o, 150^o, 180^o, 210^o, 240^o, 270^o, 300^o$, and 330^o. The average gray level of image is evaluated along lines. This is done, in order to find the line direction among the 12 directions noted above that provides the maximum gray level, denoted as $\widetilde{\theta}$. The $\widetilde{\theta}$ at the pixel of the coordinate (i, j) is calculated as below.

$$\widetilde{\theta} = argmax_\theta \sum_{x=i-k}^{i+k} \widetilde{I_g(x,q)} \tag{5}$$

where $\theta \in \{0^o, 30^o, 60^o, 90^o, 120^o, 150^o, 180^o, 210^o, 240^o, 270^o, 300^o, 330^o\}$, $\widetilde{I_g(x,q)}$ is the pixel at the coordinate (x, q) of $\widetilde{I_g}$, $q = x\tan(\theta) - i\tan(\theta) + j$, and k is a parameter defining the length of the considered line as $2k + 1$.

Then, the next step is to compute the difference by subtracting the mean value of all pixels on the line of the rotation $\widetilde{\theta}$ centered at the coordinate (i, j) with the length of $2k+1$ from the average gray level within the square window with the size of $(2n + 1) \times (2n + 1)$. This difference (S) is called the line strength, calculated as below.

$$S = \frac{\sum_{x=i-k}^{i+k} \widetilde{I_g(x,r)}}{2k+1} - \frac{\sum_{a=-n}^{n}\sum_{b=-n}^{n} \widetilde{I_g(i+a, j+b)}}{(2n+1)\times(2n+1)} \tag{6}$$

where $r = x\tan(\widetilde{\theta}) - i\tan(\widetilde{\theta}) + j$.

The value S is then used to determine whether the considered pixel is a vessel or not. The high value is supposed to be a part of vessels. It will be used as the feature in the supervised learning process. The figure 3 shows the sample output of the basic line operator.

In addition, to improve the segmentation of a pixel that is located on the thin vessels, the third feature is extracted from the orthogonal line operators by evaluating the average gray level of the neighbor pixels along the line that is perpendicular to the second feature. The rotation of such perpendicular line is denoted as $\widetilde{\theta_o}$. Its strength is denoted as S_o. It is again that S_o can be obtained by subtracting the mean value of all pixels on the line of the rotation $\widetilde{\theta_o}$ centered at the coordinate (i, j) with the length of $2k + 1$ from the average gray level within the same square window used in the second feature. The calculation uses the same equation (6) above, where $y = x\tan(\widetilde{\theta_o}) - i\tan(\widetilde{\theta_o}) + j$.

The fourth feature is based on the Gabor filter [21–23]. It is a linear filter that has been broadly used for multi-scale and multi-directional edge detection. The Gabor filter can be used to detect oriented features of vessels in the retinal image. It can fine-tune to particular scales and directions. The response of a Gabor filter kernel is defined by the product of a function f and a complex sinusoid, as below.

$$f(x, y, \theta, \lambda) = e^{-\frac{\acute{x}^2+\gamma^2\acute{y}^2}{2\sigma^2}} e^{i(2\pi\frac{\acute{x}}{\lambda}+\psi)} \tag{7}$$

where $\acute{x} = xcos(\theta) + ysin(\theta)$, $\acute{y} = -xsin(\theta) + ycos(\theta)$, θ is the orientation, σ is the scale, λ is the wavelength of the sinusoidal factor, γ is the spatial aspect ratio, and ψ is the phase offset of the sinusoidal factor.

In this paper, the Gabor filter is applied to the inverted green channel of the retinal image, using a 2D convolution operator. The maximum Gabor response across the orientation (θ) spanning from 0 to π degrees in steps of $\pi/18$, is calculated for each pixel at four different scales (σ) of 2, 3, 4 and 5. Then, the

Fig.2: *The enhancement on the green channel. (a) An original image. (b) A split green channel image. (c) An applied histogram equalization of (b).*

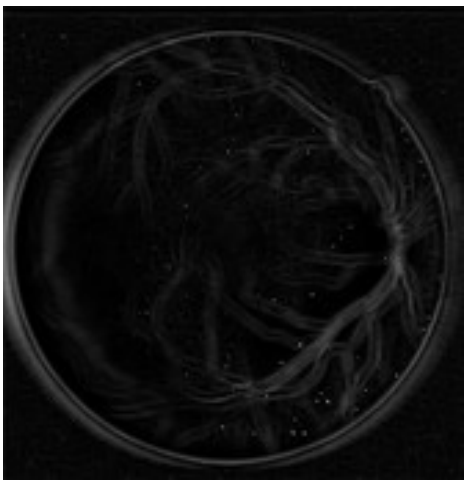

Fig.3: *The sample output of the basic line operator.*

maximum response $f_s(x, y, \theta, \lambda_s)$ across the orientation at each scale λ_s is taken as the pixel feature, as below.

$$f_s(x, y, \theta, \lambda_s) = max_{\theta \in \{0, \frac{\pi}{18}, \frac{2\pi}{18}, \ldots, \pi\}} f_s(x, y, \theta, \lambda_s) \tag{8}$$

In this stage of the coarse-level segmentation of vessels in the retinal image, Support Vector Machine (SVM) [24–28] is used as the supervised classifier for the pixel classification. In the classification process, each pixel of the retinal image takes into one of two label classes: vessel and non-vessel, which can be separated into the two classes of samples with the best generalization. In the training phase of using SVM to build up the vessel classification model, sample positive (i.e. vessel) and negative (i.e. non-vessel) pixels are taken from the groundtruth images with the 4 types of extracted feature vectors as explained above.

With using SVM [24–28], in the training phase of this binary-classes problem, the labeled training data $\{(d_m, c_m)\}_{m=1}^{M}$ is required, where m is the total number of training samples. The training dataset needs both positive samples where d_m is the extracted feature vector of the vessel pixel and the class label c_m is represented by $+1$, and negative samples where d_m is the extracted feature vector of the non-vessel pixel and the class label c_m is represented by -1.

In this paper, the SVM solves the hyperplane for the two-classes ($c_m = +1, c_m = -1$) classification problem by using the optimization below.

$$\begin{aligned} min_{w,b,\xi} \quad & \frac{1}{2} w^T w + C \sum_{m=1}^{M} \xi_m \\ subject \;\; to \quad & c_m (w^T \phi(d_m) + b) \geq 1 - \xi_m \end{aligned} \tag{9}$$

where $\xi_m \geq 0$ is the margin width, w is a weight vector, b is a bias parameter, $\phi(d_m)$ is the kernel mapping for a higher-dimensional space, and $C \geq 0$ is the regularization parameter. In this paper, the histogram intersection is used as the kernel. It is then figured out using the dual optimization, as below.

$$\begin{aligned} min_{\alpha} \quad & \frac{1}{2} \alpha^T Q \alpha - I\alpha \\ subject \;\; to \quad & c^T \alpha = 0 \end{aligned} \tag{10}$$

where $0 \leq \alpha \leq C$, I is the all-one vector of size M, Q is the $M \times M$ positive-value matrix, $Q_{i,j} = c_i c_j K(d_i, d_j)$, $K(d_i, d_j) = \phi(d_i)^T \phi(d_j)$, and $c = [c_m]_{m=1}^{M}$. Then, the optimal w is solved as below.

$$w = \sum_{m=1}^{M} c_m \alpha_m \phi(d_m) \tag{11}$$

Then, the final decision function is defined as below for the given input feature vector d_{in}.

$$f(w^T \phi(d_{in}) + b) = f(\sum_{m=1}^{M} c_m \alpha_m K(d_m, d_{in}) + b) \tag{12}$$

where $f(x)$ returns $+1$ or the vessel class if $f(x) \geq 0$,

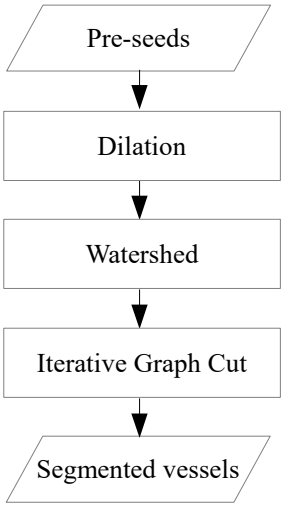

Fig.4: *The overall framework of the proposed fine-level segmentation.*

otherwise $f(x)$ returns -1 or the non-vessel class.

The trained model is then used to classify pixels in the retinal image into vessel and non-vessel pixels, as stated in the equation (12). The vessel pixels with the returned high confidential values in the top 25% percentiles are used further as pre-seeds for the next stage of instance-based learning.

2.2 Segmentation Level 2 using Instance Learning (Fine-Level Segmentation)

The figure 4 shows the overall framework of the proposed fine-level segmentation. This level of segmentation is to use the pre-seeds of vessels computed in the previous level of segmentation, as the initial input. First, the pre-seeds are refined using the morphological operations [29, 30]. The dilation is applied on the pre-seeds to enlarge candidate areas of vessels, which will be refined at the end of this process using the iterative graph cut. The dilation factor is set to be 3, with a disk-shaped structuring element of the radius = 3. The sample result is shown in the figure 5.

Then, the watershed algorithm [31] is applied to reduce the noises from the pre-seeds. The output from the watershed is then used as the initial foreground (i.e. vessels) in the iterative graph cut algorithm [32]. The existing usages of graph cut based approaches [33–36] applied on the initialization of bounding box or region covering the target object. However, in this paper, it is applied with the pre-seed pixels obtained from the coarse-level segmentation based on the supervised learning as the sure foreground pixels for building up the initial foreground model using Gaussian Mixture Models (GMMs).

The graph is first built up, in which each node represents each pixel in the image. There are two additional special nodes, including the source node representing foreground or vessels and the sink node

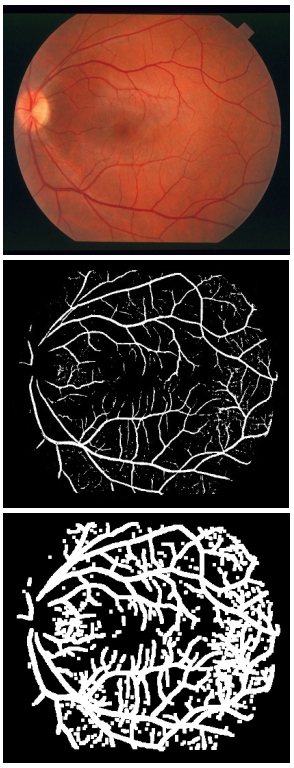

Fig.5: *The dilation applied on the pre-seed image.*

representing background or non-vessels. Each node in the graph is then connected to neighboring nodes and the two special nodes.

The weight of the link connecting each node to the source node is computed using the foreground Gaussian mixture model. While, the weight of the link connecting each node to the sink node is computed using the background Gaussian mixture model. Then, the weight w_{p_1,p_2} of the link connecting between two nodes of two pixels (p_1, p_2) is computed using their pixels' similarity, using the equation below [37, 38].

$$w_{p_1,p_2} = \frac{50}{dist(p_1,p_2)} \exp^{-\beta||pv_1-pv_2||^2} \qquad (13)$$

where $dist(p_1, p_2)$ is the distance between the pixels p_1 and p_2 coordinates, pv_1 and pv_2 are the pixel intensities of the pixels p_1 and p_2 respectively, $\beta = \frac{1}{2<(pv_1-pv_2)^2>}$, and $< . >$ denotes the expectation over the image sample. The constant value of 50 is used in the equation since it is the value suggested for the segmentation task in [39] for model calculations.

The models are initially created based on the refined pre-seeds. Then, the updated foreground and background pixels are used for constructing the models in the next iteration. In each iteration of the iterative graph cut, the mincut algorithm [40, 41] is used to cut the graph into two segments with the minimum cutting cost. The cost function is the summation of weights of the links being cut.

Then, the pixels on the nodes connecting to the

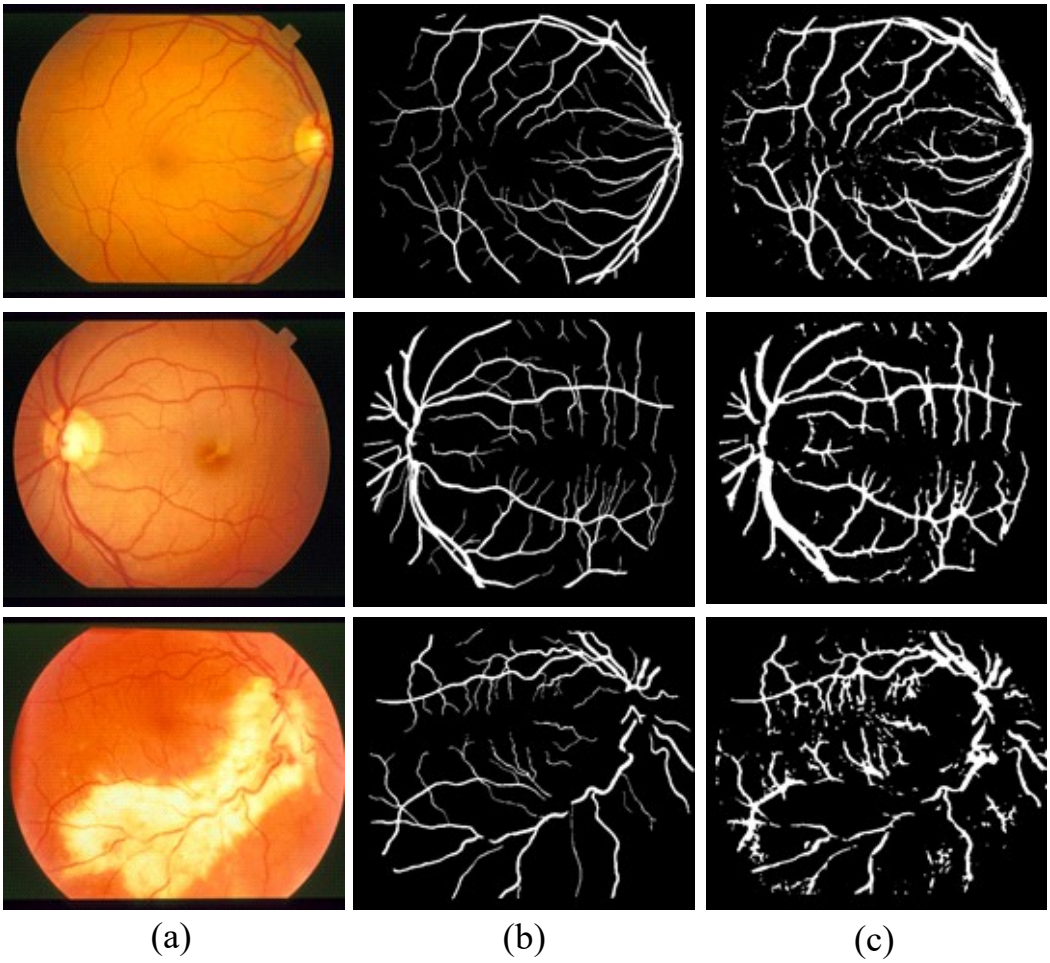

Fig.6: *Sample outputs of the vessel segmentation using the proposed method. (a) Original images. (b) Ground truth images. (c) Segmented images.*

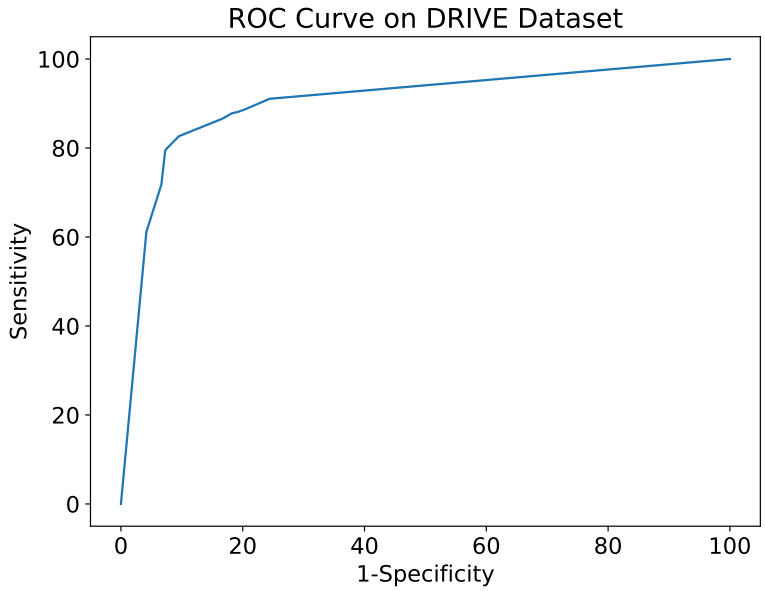

Fig.7: *ROC curve on the DRIVE dataset.*

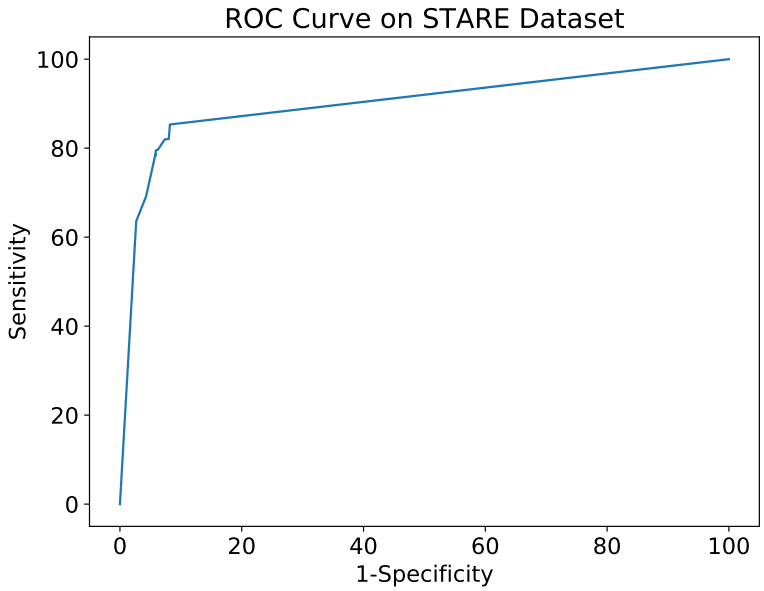

Fig.8: *ROC curve on the STARE dataset.*

source node are said to be foreground pixels or vessel pixels. While, the pixels on the nodes connecting to the sink node are said to be background pixels or non-vessel pixels. All of these processes is iteratively repeated until the segmentation converges or the maximum number of iterations is reached.

In the final step, the post-processing is applied on the segmented image. The morphological operations are then applied again to fill holes in vessels and remove noises. Also, the outer boundary of the retina contour is detected and removed from the final output image.

3. EXPERIMENTS AND DISCUSSIONS

The two public retinal image datasets are used to evaluate the proposed method, including the structured analysis of the retina (STARE) dataset [42] and the digital retinal images for vessel extraction (DRIVE) dataset [2]. The STARE dataset consists of 20 digitized images captured by the TopCon TRV-50 fundus camera at 35° filed-of-view (FOV), where ten of them present pathology. They are available in the ppm format. The image' size is 700 × 605 pixels with eight bits per color channel.

The DRIVE dataset is mainly to enable comparative studies on blood vessel segmentation in retinal images. It consists of 40 eye-fundus color images acquired by the Canon CR5 nonmydriatic 3CCD camera at 45° field-of-view (FOV), where seven of them present pathology. They are available in the TIF format. The image's size is 768 × 584 pixels with eight bits per color channel. For both datasets, 20% of images are used in the training process, where the rest of 80% of images are used in the testing phase.

In our experiments, the proposed method is implemented on the computer with Core(TM) i7 CPU@2.70 GHz and 16 GB RAM. It is developed using the library functions of OpenCV in Microsoft Visual C/C++ environment.

The figure 6 illustrates the sample segmented vessel images using the proposed method. The performances are evaluated based on the sensitivity and the specificity. The sensitivity reflects the ability of the method to correctly detect vessel pixels, while the specificity presents the ability of the method to correctly detect non-vessel pixels.

The sensitivity is calculated by diving the true positive with the summation of true positive and false negative. The specificity is calculated by dividing the true negative with the summation of true negative and false positive. The true positive is the number of vessel-pixels that are correctly classified as vessels. The true negative is the number of non-vessel pixels that are correctly classified as non-vessels. The false positive is the number of non-vessel pixels are incorrectly classified as vessels. The false negative is the number of vessel pixels are incorrectly classified as non-vessels.

The ROC curves based on the DRIVE and STARE datasets are shown in the figure 7 and figure 8 respectively. They are computed by varying the confidential score of SVM for classifying vessels and the dilation factor used in the fine-level segmentation step. The tables 1 and 2 demonstrates the performance comparisons between the proposed method and the other existing methods in the literature.

As can be seen in the tables 1 and 2, the proposed method outperforms the other methods, based on the sensitivity values. This is because of the two-levels

Table 1: *The performance comparisons based on the DRIVE dataset.*

Method	Sensitivity	Specificity
[5]	75.1	96.8
[6]	77.5	97.3
[9]	67.7	98.7
[10]	78.0	98.0
[11]	70.7	98.0
[14]	74.5	97.8
[15]	81.9	98.2
[16]	79.6	98.3
[18]	79.9	98.0
[43]	79.9	97.4
The proposed method	82.6	90.5

Table 2: *The performance comparisons based on the STARE dataset.*

Method	Sensitivity	Specificity
[5]	78.9	96.3
[6]	90.3	93.9
[9]	70.4	98.7
[10]	78.0	98.0
[11]	69.4	98.2
[14]	71.9	98.2
[15]	-	-
[16]	81.6	98.7
[18]	80.9	98.4
[43]	79.1	97.2
The proposed method	82.0	94.7

segmentation deployed in the proposed method. The first level using the supervised learning can detect initial seeds of vessel pixels. The performance on different input images could be varied because they contain different details of illuminations and other signals of disease. The same trained model could not cover all input images with the same performance. In the second level using the instance learning, it could begin with the segmentation result from the first level and extend to cover more foreground pixels based on internal contexts of individual input images.

In addition, the CNN-based approaches [14][15][16] [18][43] could achieve high specificity scores, but with significantly lower sensitivity scores. This is because, in each image, the negative class or non-vessel pixel has a significantly higher proportion, when compared to the positive class or vessel pixel. More importantly, the CNN-based approaches belong to the supervised-based learning category which could work well on seen data/vessel patterns. However, in facts, retinal vessels contain high variations of shapes, sizes, and colors. So, the trained CNN model may not be able to segment vessels that are variate from the seen data in the training process. In addition, the complexity of CNN architecture must be trained using a sufficiently large set of labeled data.

When compared to the proposed method, it could address these intrinsic variations in some extent, since it uses the instance-based learning technique on pre-seeds of vessels segmented from the supervised-based learning technique. So, it could learn and detect vessels instantly and specifically for individual retinal images. The proposed method attempts to achieve a high sensitivity because the medical diagnosis commonly needs a high sensitivity [44]. It is important to not miss any parts of positive class or signals of disease, for a further medical analysis.

The further enhancement of the segmentation performance could be done in the future work by filtering and/or clustering the pre-seeds obtained from the first level segmentation. The pre-seeds of vessels could be clustered into multiple groups based on their intensity's levels. Then, the more reliable pre-seeds could be the cluster(s) containing pixels of approximate connected line(s). The Hough transform could be used to deal with the approximate line detection. This process should improve the pureness of pre-seeds for being sure vessel pixels, which will lead to the better initialization of the graph cut and better segmentation results.

4. CONCLUSION

This paper proposes the two-levels segmentation of vessels in the retinal image. The main contribution of this paper is the development of this hybrid learning solution combining both aspects of instance learning and supervised learning processes. They could minimize limitations of each other as mentioned in the main content of this paper. The segmented vessels from the supervised learning step is further used as the initial seeds in the instance learning step, in order to enhance the sensitivity score. The proposed method starts with the pre-processing stage using the histogram equalization and color transfer. Then, the four types of features are extracted and used in the first level of the segmentation, including the green intensities, basic-line operators, orthogonal-line operators, and Gabor features. The SVM is used as the vessel classifier. The segmentation output from this level are refined using the second level based on the instance learning approach. It includes the morphological operators, watershed, and iterative graph cut. The proposed method is validated using the two well-known datasets including DRIVE and STARE. It achieves the outstanding sensitivity values, when compared with the other methods in the literature.

ACKNOWLEDGMENT

This work was supported by the Office of Higher Education Commission (OHEC) Thailand and the Thailand Research Fund (TRF) [grant number MRG6080267].

References

[1] A. Imran, J. Li, Y. Pei, J.-J. Yang, and Q. Wang, "Comparative analysis of vessel segmentation techniques in retinal images," *IEEE Access*, vol. 7, pp. 114 862–114 887, 2019.

[2] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE transactions on medical imaging*, vol. 23, no. 4, pp. 501–509, 2004.

[3] J. Zhang, Y. Cui, W. Jiang, and L. Wang, "Blood vessel segmentation of retinal images based on neural network," in *International Conference on Image and Graphics*. Springer, 2015, pp. 11–17.

[4] D. Paulus, S. Chastel, and T. Feldmann, "Vessel segmentation in retinal images," in *Medical Imaging 2005: Physiology, Function, and Structure from Medical Images*, vol. 5746. International Society for Optics and Photonics, 2005, pp. 696–706.

[5] A. G. Salazar-Gonzalez, D. Kaba, Y. Li, and X. Liu, "Segmentation of the blood vessels and optic disk in retinal images." *IEEE J. Biomedical and Health Informatics*, vol. 18, no. 6, pp. 1874–1886, 2014.

[6] E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE transactions on medical imaging*, vol. 26, no. 10, pp. 1357–1365, 2007.

[7] R. Kharghanian and A. Ahmadyfard, "Retinal blood vessel segmentation using gabor wavelet and line operator," *International Journal of Machine Learning and Computing*, vol. 2, no. 5, p. 593, 2012.

[8] R. Perfetti, E. Ricci, D. Casali, and G. Costantini, "Cellular neural networks with virtual template expansion for retinal vessel segmentation," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 54, no. 2, pp. 141–145, 2007.

[9] Z. Han, Y. Yin, X. Meng, G. Yang, and X. Yan, "Blood vessel segmentation in pathological retinal image," in *Data Mining Workshop (ICDMW), 2014 IEEE International Conference on*. IEEE, 2014, pp. 960–967.

[10] D. S. S. Raja, S. Vasuki, and D. R. Kumar, "Performance analysis of retinal image blood vessel segmentation," *Advanced Computing*, vol. 5, no. 2/3, p. 17, 2014.

[11] D. Marín, A. Aquino, M. E. Gegúndez-Arias, and J. M. Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," *IEEE transactions on medical imaging*, vol. 30, no. 1, p. 146, 2011.

[12] Q. Jin, Z. Meng, T. D. Pham, Q. Chen, L. Wei, and R. Su, "Dunet: A deformable network for retinal vessel segmentation," *Knowledge-Based Systems*, vol. 178, pp. 149–162, 2019.

[13] D. Yang, M. Ren, and B. Xu, "Retinal blood vessel segmentation with improved convolutional neural networks," *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 6, pp. 1112–1118, 2019.

[14] S. Guo, K. Wang, H. Kang, Y. Zhang, Y. Gao, and T. Li, "Bts-dsn: Deeply supervised neural network with short connections for retinal vessel segmentation," *International journal of medical informatics*, vol. 126, pp. 105–113, 2019.

[15] A. Hatamizadeh, H. Hosseini, Z. Liu, S. D. Schwartz, and D. Terzopoulos, "Deep dilated convolutional nets for the automatic segmentation of retinal vessels," *arXiv preprint arXiv:1905.12120*, 2019.

[16] Z. Fan, J. Mo, and B. Qiu, "Accurate retinal vessel segmentation via octave convolution neural network," *arXiv preprint arXiv:1906.12193*, 2019.

[17] K. J. Noh, S. J. Park, and S. Lee, "Scale-space approximated convolutional neural networks for retinal vessel segmentation," *Computer methods and programs in biomedicine*, vol. 178, pp. 237–246, 2019.

[18] Q. Jin, Q. Chen, Z. Meng, B. Wang, and R. Su, "Construction of retinal vessel segmentation models based on convolutional neural network," *Neural Processing Letters*, pp. 1–18, 2019.

[19] J. Li, "An empirical comparison between svms and anns for speech recognition," in *The First Instructional Conf. on Machine Learning*, vol. 951, 2003, p. 2003.

[20] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Computer graphics and applications*, vol. 21, no. 5, pp. 34–41, 2001.

[21] W. Kusakunniran, A. Wiratsudakul, U. Chuachan, S. Kanchanapreechakorn, and T. Imaromkul, "Automatic cattle identification based on fusion of texture features extracted from muzzle images," in *2018 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2018, pp. 1484–1489.

[22] D. Barina, "Gabor wavelets in image processing," *arXiv preprint arXiv:1602.03308*, 2016.

[23] M. M. Fraz, P. Remagnino, A. Hoppe, S. Velastin, B. Uyyanonvara, and S. Barman, "A supervised method for retinal blood vessel segmentation using line strength, multiscale gabor and morphological features," in *Signal and Image Processing Applications (ICSIPA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 410–415.

[24] I. Steinwart and A. Christmann, *Support vector machines*. Springer Science & Business Media, 2008.

[25] A. Iranmehr, H. Masnadi-Shirazi, and N. Vasconcelos, "Cost-sensitive support vector machines," *Neurocomputing*, 2019.

[26] W. Kusakunniran, K. Ngamaschariyakul, C. Chantaraviwat, K. Janvittayanuchit, and

K. Thongkanchorn, "A thai license plate localization using svm," in *2014 International Computer Science and Engineering Conference (ICSEC)*. IEEE, 2014, pp. 163–167.

[27] P. Druzhkov, V. Erukhimov, N. Y. Zolotykh, E. Kozinov, V. Kustikova, I. Meerov, and A. Polovinkin, "New object detection features in the opencv library," *Pattern Recognition and Image Analysis*, vol. 21, no. 3, p. 384, 2011.

[28] C.-C. Chang and C.-J. Lin, "Libsvm: A library for support vector machines," *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, no. 3, p. 27, 2011.

[29] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. "O'Reilly Media, Inc.", 2008.

[30] P. K. Ghosh and R. M. Haralick, "Mathematical morphological operations of boundary-represented geometric objects," *Journal of Mathematical Imaging and Vision*, vol. 6, no. 2-3, pp. 199–222, 1996.

[31] J. B. Roerdink and A. Meijster, "The watershed transform: Definitions, algorithms and parallelization strategies," *Fundamenta informaticae*, vol. 41, no. 1, 2, pp. 187–228, 2000.

[32] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: Interactive foreground extraction using iterated graph cuts," in *ACM transactions on graphics (TOG)*, vol. 23, no. 3. ACM, 2004, pp. 309–314.

[33] M. Eapen, R. Korah, and G. Geetha, "Swarm intelligence integrated graph-cut for liver segmentation from 3d-ct volumes," *The Scientific World Journal*, vol. 2015, 2015.

[34] L. Massoptier and S. Casciaro, "Fully automatic liver segmentation through graph-cut technique," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2007, pp. 5243–5246.

[35] J. Cha, M. M. Farhangi, N. Dunlap, and A. A. Amini, "Segmentation and tracking of lung nodules via graph-cuts incorporating shape prior and motion from 4d ct," *Medical physics*, vol. 45, no. 1, pp. 297–306, 2018.

[36] J. won Cha, N. Dunlap, B. Wang, and A. Amini, "3d segmentation of lung ct data with graph-cuts: analysis of parameter sensitivities," in *Medical Imaging 2016: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 9788. International Society for Optics and Photonics, 2016, p. 97882O.

[37] J. F. Talbot and X. Xu, "Implementing grabcut," *Brigham Young University*, vol. 3, 2006.

[38] T. Malmer, "Image segmentation using grabcut," *IEEE Transactions on Signal Processing*, vol. 5, no. 1, pp. 1–7, 2010.

[39] A. Blake, C. Rother, M. Brown, P. Perez, and P. Torr, "Interactive image segmentation using an adaptive gmmrf model," in *European conference on computer vision*. Springer, 2004, pp. 428–441.

[40] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 9, pp. 1124–1137, 2004.

[41] M. Stoer and F. Wagner, "A simple min-cut algorithm," *Journal of the ACM (JACM)*, vol. 44, no. 4, pp. 585–591, 1997.

[42] A. Hoover and M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels," *IEEE transactions on medical imaging*, vol. 22, no. 8, pp. 951–958, 2003.

[43] C. Wang, Z. Zhao, Q. Ren, Y. Xu, and Y. Yu, "Dense u-net based on patch-based learning for retinal vessel segmentation," *Entropy*, vol. 21, no. 2, p. 168, 2019.

[44] B. Zupan, E. T. Keravnou, and N. Lavrac, *Intelligent Data Analysis in Medicine and Pharmacology*. Kluwer Academic Publishers, 1997.