Research Article

A Hybrid Unsupervised Approach for Retinal Vessel Segmentation

Khan Bahadar Khan, Muhammad Shahbaz Siddique, Muhammad Ahmad, and Manuel Mazzara

1Department of Telecommunication Engineering, Faculty of Engineering, The Islamia University of Bahawalpur, Bahawalpur, Pakistan
2Department of Electronic Engineering, Faculty of Engineering, The Islamia University of Bahawalpur, Bahawalpur, Pakistan
3Department of Computer Science, National University of Computer and Emerging Sciences, Islamabad, Chiniot-Faisalabad Campus, Chiniot 35400, Pakistan
4Institute of Software Development and Engineering, Innopolis University, Innopolis, Russia

Correspondence should be addressed to Khan Bahadar Khan; kb.khattak@gmail.com

Received 22 January 2020; Accepted 26 November 2020; Published 12 December 2020

Copyright © 2020 Khan Bahadar Khan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Retinal vessel segmentation (RVS) is a significant source of useful information for monitoring, identification, initial medication, and surgical development of ophthalmic disorders. Most common disorders, i.e., stroke, diabetic retinopathy (DR), and cardiac diseases, often change the normal structure of the retinal vascular network. A lot of research has been committed to building an automatic RVS system. But, it is still an open issue. In this article, a framework is recommended for RVS with fast execution and competing outcomes. An initial binary image is obtained by the application of the MISODATA on the preprocessed image. For vessel structure enhancement, B-COSFIRE filters are utilized along with thresholding to obtain another binary image. These two binary images are combined by logical AND-type operation. Then, it is fused with the enhanced image of B-COSFIRE filters followed by thresholding to obtain the vessel location map (VLM). The methodology is verified on four different datasets: DRIVE, STARE, HRF, and CHASE_DB1, which are publicly accessible for benchmarking and validation. The obtained results are compared with the existing competing methods.

1. Introduction

The most essential sensory system for gathering information, navigation, and learning is the human visual system [1]. The retina is the sensitive part of the eye that contains fovea, light receptors, Optic disk, and macula. The retina is a layered tissue, coating the interior of the eye, which is an initial sensor of the communication system and gives a sense of sight. Moreover, it allows understanding the colors, dimensions, and shape of objects by processing the amount of light it reflects or emits. Retina image of an eye is captured with a fundus camera [2]. RGB photographs of the fundus are the protrusion of the internal surface of an eye. Imaging of the retina has emerged swiftly and now one of the most common practices in healthcare and for screening the patients suffering from ophthalmologic or systemic diseases. For identifying numerous ophthalmologic diseases, the ophthalmologist uses vessel condition as an indicator which is a vital component in retinal fundus images.

Critical diagnostic to eye diseases in human retinal images can be indicated by its shape analysis, its appearance, blood vessels, morphological features, and tortuosity [3]. Structure of RVS is also used for screening of brain and heart stock diseases [4, 5]. Retinal vessel structures play a significant role among other structures in fundus images. RVS is the elementary phase utilized for the examination of retina images [6]. Vascular-related diseases are diagnosed with the help of vessel delineation which is an important component of medical image processing. Additionally, ongoing research in the area of deep learning suggested multiple approaches with emphasis on the separation and the delineation of the vasculature.
The inadequate number of images and having low-contrast in publicly available retina datasets is challenging for deep learning-based research. A dataset having a large number of retina images captured with a different imaging system and under diverse environmental conditions is required to train the supervised network. Deep learning-based methods will aid to control blindness, timely and precise identification of diseases for successful remedy, and thus vividly increase the life quality of patients with eye ailments [7]. RVS is a very difficult task due to many reasons:

(1) The structure and formation of retinal vessels are very complex and there is a prominent dissimilarity in various local parts regarding the shape, size, and intensity in vessels.

(2) Some structures have the same intensity and shape as vessels, e.g., hemorrhage. Moreover, there are also thin microvessels, whose width is normally between ranges from one to a few pixels and which can be easily mixed with the background. There are irregular illumination in the images and having low-varying contrast [7, 8]. Typically, noise in fundus images is added by the image-capturing procedure such as artifact on the lens or movement of the patient [9]. It is hard to differentiate vessels from other structures that are similar or noises in the retina image. In other words, thicker vessels are more prominent in comparison to the thinner ones as shown in Figure 1.

(3) Different manual graders have different segmentation results. Manual RVS is also a very hard and tedious task. Over the recent two decades, automatic RVS has caught noteworthy attention and numerous such techniques are developed but they have performance degradation with the change of datasets. Some of the techniques are not fully automatic while others are incapable to handle pathological images. Some of these methods are evaluated on the datasets having a limited number of images while others have problems of oversegmentation or undersegmentation with abnormal images [10]. Hence, the dilemma of perfect RVS is still not answered.

Automated RVS techniques provide incredible support to the ophthalmologist in terms of identification and medication of numerous ophthalmological abnormalities. In this article, an automatic unsupervised approach is developed for RVS that consists a combination of the preprocessing steps, segmentation, vessel structure-based enhancement, and postprocessing steps. The preprocessing steps aim at exterminating noise and improving the contrast of the fundus image. Segmentation is performed by using the Modified Iterative Self Organizing Data Analysis Technique (MISODATA) to acquire a binary image that is fused with the segmented image of the Combination Of Shifted Filter Responses (B-COSFIRE). Then, the fused image is multiplied with the enhanced image of the B-COSFIRE to obtain the initial vessel location map (VLM). Lastly, the VLM and the fused image are combined by logical OR-type operators to obtain final results. In a nutshell, the main contributions of this research are the following:

(1) A mask image is not provided with all retina datasets. Automatic masking creation is proposed for each image to extract ROI which suppresses the false positive rate (FPR).

(2) The proposed efficient denoising process (preprocessing steps) improves the selection of a suitable threshold.

(3) The basic ISODATA algorithm only one-time process the retina image locally and then globally, which sometimes makes it unable to find an optimal threshold. The modified ISODATA technique is introduced to find the global threshold of the entire image which is compared and equated with the individual local threshold of each segment in order to find the optimal threshold for more precise detection of vessels.

(4) The vessel location map (VLM) is a new scheme to achieve better performance. In this scheme, the background noise eradication and vessel enhancement are accomplished independently.

(5) A distinctive postprocessing steps (AND-type and OR-type operations) to reject misclassified foreground pixels.

2. Related Works

Numerous methodologies for RVS have been developed in literature [4, 10]. These methodologies are arranged into two sets: supervised and unsupervised procedures. Supervised techniques utilizing a trained classifier for pixel classification into the foreground or background. Supervised techniques utilized various classifiers, for instance, adaptive boosting (AdaBoost), support vector machines (SVM), neural networks (NN), Gaussian mixture models (GMM) and k-nearest neighbors (k-NN).

A RVS method utilizing a supervised k-NN classifier for isolation of foreground and background pixels was recommended by Niemeijer et al. [11], with a feature vector (FV) formation based on a multiscale (MS) Gaussian filter. Staal et al. [12] projected an equivalent RVS methodology using an FV generated based on a ridge detector. A feed-forward NN built classifier was applied by Marin et al. [13], using 7-D FV generated based on moment-invariant.

An SVM-based approach was presented by Ricci et al. [14], utilizing FV constructed through a rotation-invariant linear operator and pixel intensity. An AdaBoost classifier was suggested by Lupascu et al. [15], utilizing a 41 − D feature set. An ensemble-based RVS system applying a simple linear iterative clustering (SLIC) algorithm was presented by Wang et al. [16]. A GMM classifier-based scheme was recommended by Roychowdhury et al. [17], utilizing 8 − D FV extracted from the pixel neighborhood on first and second-order gradient images.
Figure 1: Graphical representation of thinner and thicker vessels.

Figure 2: Sequential sketch of the proposed framework.
Figure 3: (a) DRIVE dataset color image. (b) Green plane. (c) CLAHE output.

Figure 4: Histogram pictorial effects. (a) Green plane. (b) CLAHE.
and the multiorientation Gabor function (ELM) based RVS scheme utilizing a random forest classification, and vessel fusion-based) deep learning approach were proposed by Aslani et al. [20], utilizing a 17-D FV created based on MS vessel filtering and the Gabor wavelet features. A random forest classifier-based RVS system was proposed by Thangaraj et al. [22] constructed a directionally sensitive vessel enhancement-based scheme combined with NN derived from the U-Net model was presented in [21]. Thangaraj et al. [22] constructed a 13-D FV from the Gabor filter responses, Frangi’s vesselness measure (1D), local binary pattern feature (1D), Hu moment invariants (7D), and grey-level cooccurrence matrix features (3D) for RVS utilizing NN-based approach. Memari et al. [23] recommended an arrangement of various enhancement techniques with the AdaBoost classifier to segregate foreground and background pixels.

A three-stage (thick vessel extraction, thin vessel extraction, and vessel fusion-based) deep learning approach were proposed in [24]. Guo et al. [25] suggested an MS deeply supervised network with short connections (TS-DSN) for RVS. Local intensities, local binary patterns, a histogram of gradients, DVF, higher-order local autocorrelations, and morphological transformation features were used for RVS in [26]. Random forests were used for the selection of feature

Algorithm 1: MISODATA algorithm.
Figure 5: (a) Average filter image. (b) Difference image. (c) Segmented initial vessel map.

Figure 6: (a) B-COSFIRE binary segmented image (b) Enhanced image based on symmetric-asymmetric filter responses. (c) AND-type operation output.

Figure 7: Inspecting the results of AND-type operator. (a) Manual image. (b) AND-type. (c) OR-type.

Figure 8: Inspecting the results of AND-type operator. (a) Manual image. (b) AND-type. (c) without AND-type.
sets which were utilized in combination with the hierarchical classification methodology to extract the vessels.

Alternatively, unsupervised systems are categorized based on matched filtering (MF), mathematical morphology (MM), and multiscale-based approaches. In matched filtering approaches, thick and thin vessels are extracted by the selection of large and small filter kernels, respectively. However, the application of large kernels can accurately detect major vessels with the misclassification of thin vessels by increasing its width. Similarly, smaller kernels can accurately extract thin vessels along with the extraction of thick vessels in reduced widths. To obtain a complete vascular network, a conventional MF technique can be applied with a large number of diverse filter masks in various directions.
Similar methods were employed using MF [27–32], combined filters [33], COSFIRE filters [3, 5, 34–36], Gaussian filters [37], wavelet filters [38], and Frangi’s filter [39]. The MM-based approaches are utilized for isolating retinal image segments such as optic disk, macula, fovea, and vasculature. Morphological operators utilized the application of structuring elements (SE) to images for extraction and representation of region contours. A morphological operation for detecting particular structures has the benefit of speed and noise elimination. But they are unable to achieve the known vessel cross-sectional shape. Moreover, there is an issue to extract extremely tortuous vessels in case of superimposing large SE. Morphological operations were utilized for both enhancement and RVS [2, 40–44]. On the other hand, retinal blood vessels of variable thickness at various scales were obtained by multiscale approaches [45–50].

3. Proposed Model

The complete structure of the proposed RVS framework is introduced in this section. The information and description of every stage are also presented in subsections.

3.1. Overview. The proposed framework consists of two major blocks to obtain a final binary image: (1) retina image denoising and segmentation and (2) vessel structure-based enhancement and segmentation. The key objective of this framework is to extract vasculature excellently along with the elimination of noise and supplementary disease falsifications. The complete structure of the proposed framework is labeled in Figure 2. In which Block-I consists of the selection of suitable retina channel, contrast enhancement, noise filtering, region of interest (ROI) extraction, thresholding, and post processing steps. Block-II includes the application of B-COSFIRE filter, logical operations, and postprocessing steps. The initial binary vessel map of Block-I is fused with the B-COSFIRE filter segmented image in Block-II. Then, it is multiplied with the B-COSFIRE filter-enhanced image which is further thresholded. This output image is combined with the initial postprocessed image by the logical OR-type operation to obtain the final binary.

3.2. Block-I: Retina Image Denoising and Segmentation. In the first block, the retina image is passed through selected techniques to extract the initial denoised vessel map. The green band of the RGB retina image is extracted and nominated for subsequent operation due to its noticeable contrast difference between the vessel and other retina structures. The RGB retina images generally have contrast variations, low resolution and noise. To avoid such variations and produce more appropriate image for further processing, the vessel light reflex elimination and background uniformity operations are performed. Retinal vessel structures have poor reflectance when equated to other retinal planes. Some vessels contain a bright stripe (light reflex) which runs down the central length of the vessel. To overcome this problem, a disc-shape opening operator with a 3-pixel width SE is used on the green plane. A minimal value of disc width is selected to avoid the absorption of close vessels. The background uniformity and smoothness of random salt-and-pepper noise are obtained by the application of a $3 \times 3$ mean filter. Additional noise flattening is achieved with the application of a Gaussian kernel of size $9 \times 9$, mean $= 0$, and variance 1.8.

CLAHE [51, 52] is applied on the preprocessed green channel to make vessel structures prominent. The CLAHE operation divides the input image into blocks (size $8 \times 8$ in our case) with the constraint of contrast improvement which is set to 0.01. The clip limit suppresses the noise level and escalates the contrast. The effect of the CLAHE process ($I_{\text{clahe}}$) along with the green plane is displayed in Figure 3. Histogram-based graphical demonstration of the contrast improvement operations is displayed in Figure 4. An averaging filter of size $49 \times 49$ is applied for smoothness and elimination of anatomical regions (e.g., optic disk, macula, and fovea). $I_{\text{avg}}$ symbolizes the output image of the averaging filter. The difference image ($I_{d}$) is computed for all pixels as follows.

$$I_{d}(m, n) = I_{\text{avg}}(m, n) - I_{\text{clahe}}(m, n).$$  \hspace{1cm} (1)

The extra regions of the retina image are cropped by the utilization of the masking method to extract ROI which reduced the computational complexity. An automatic mask...
is created from the red band of the retinal image. The reason behind using the red channel for mask construction is that it has a good vessel-background dissimilarity. The automatic mask is created for all datasets because the mask image is not available in some datasets. \( I_4 \) is thresholded by the MISO-DATA algorithm. The subsequent procedure is used to compute the threshold level, and the application of MISO-DATA is shown in Algorithm 1.
Table 5: Average efficiency scores on the HRF dataset (Glaucomatous images).

<table>
<thead>
<tr>
<th>Img</th>
<th>Acc</th>
<th>Sn</th>
<th>Sep</th>
<th>AUC</th>
<th>MCC</th>
<th>CAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.952</td>
<td>0.813</td>
<td>0.962</td>
<td>0.887</td>
<td>0.688</td>
<td>0.743</td>
</tr>
<tr>
<td>2</td>
<td>0.954</td>
<td>0.805</td>
<td>0.966</td>
<td>0.886</td>
<td>0.704</td>
<td>0.679</td>
</tr>
<tr>
<td>3</td>
<td>0.962</td>
<td>0.830</td>
<td>0.970</td>
<td>0.900</td>
<td>0.688</td>
<td>0.610</td>
</tr>
<tr>
<td>4</td>
<td>0.962</td>
<td>0.788</td>
<td>0.974</td>
<td>0.881</td>
<td>0.711</td>
<td>0.695</td>
</tr>
<tr>
<td>5</td>
<td>0.963</td>
<td>0.833</td>
<td>0.972</td>
<td>0.902</td>
<td>0.721</td>
<td>0.663</td>
</tr>
<tr>
<td>6</td>
<td>0.962</td>
<td>0.811</td>
<td>0.973</td>
<td>0.892</td>
<td>0.708</td>
<td>0.713</td>
</tr>
<tr>
<td>7</td>
<td>0.969</td>
<td>0.773</td>
<td>0.977</td>
<td>0.875</td>
<td>0.721</td>
<td>0.609</td>
</tr>
<tr>
<td>8</td>
<td>0.957</td>
<td>0.835</td>
<td>0.964</td>
<td>0.899</td>
<td>0.703</td>
<td>0.720</td>
</tr>
<tr>
<td>9</td>
<td>0.962</td>
<td>0.774</td>
<td>0.975</td>
<td>0.875</td>
<td>0.710</td>
<td>0.646</td>
</tr>
<tr>
<td>10</td>
<td>0.963</td>
<td>0.780</td>
<td>0.976</td>
<td>0.878</td>
<td>0.718</td>
<td>0.665</td>
</tr>
<tr>
<td>11</td>
<td>0.962</td>
<td>0.749</td>
<td>0.976</td>
<td>0.863</td>
<td>0.717</td>
<td>0.641</td>
</tr>
<tr>
<td>12</td>
<td>0.958</td>
<td>0.799</td>
<td>0.967</td>
<td>0.883</td>
<td>0.717</td>
<td>0.688</td>
</tr>
<tr>
<td>13</td>
<td>0.958</td>
<td>0.753</td>
<td>0.974</td>
<td>0.863</td>
<td>0.692</td>
<td>0.664</td>
</tr>
<tr>
<td>14</td>
<td>0.951</td>
<td>0.764</td>
<td>0.965</td>
<td>0.865</td>
<td>0.663</td>
<td>0.646</td>
</tr>
<tr>
<td>15</td>
<td>0.956</td>
<td>0.748</td>
<td>0.973</td>
<td>0.860</td>
<td>0.691</td>
<td>0.620</td>
</tr>
<tr>
<td>Avg</td>
<td>0.959</td>
<td>0.790</td>
<td>0.979</td>
<td>0.881</td>
<td>0.703</td>
<td>0.667</td>
</tr>
</tbody>
</table>

Table 6: Average performance achieved on the CHASE_DB1 dataset.

<table>
<thead>
<tr>
<th>Images</th>
<th>Acc</th>
<th>Sn</th>
<th>Sep</th>
<th>AUC</th>
<th>MCC</th>
<th>CAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>01L</td>
<td>0.995</td>
<td>0.753</td>
<td>0.996</td>
<td>0.874</td>
<td>0.582</td>
<td>0.552</td>
</tr>
<tr>
<td>01R</td>
<td>0.997</td>
<td>0.705</td>
<td>0.999</td>
<td>0.852</td>
<td>0.707</td>
<td>0.485</td>
</tr>
<tr>
<td>02L</td>
<td>0.995</td>
<td>0.647</td>
<td>0.997</td>
<td>0.822</td>
<td>0.623</td>
<td>0.516</td>
</tr>
<tr>
<td>02R</td>
<td>0.995</td>
<td>0.744</td>
<td>0.996</td>
<td>0.870</td>
<td>0.620</td>
<td>0.488</td>
</tr>
<tr>
<td>03L</td>
<td>0.996</td>
<td>0.783</td>
<td>0.997</td>
<td>0.890</td>
<td>0.646</td>
<td>0.571</td>
</tr>
<tr>
<td>03R</td>
<td>0.998</td>
<td>0.848</td>
<td>0.999</td>
<td>0.923</td>
<td>0.793</td>
<td>0.582</td>
</tr>
<tr>
<td>04L</td>
<td>0.996</td>
<td>0.588</td>
<td>0.998</td>
<td>0.793</td>
<td>0.648</td>
<td>0.616</td>
</tr>
<tr>
<td>04R</td>
<td>0.997</td>
<td>0.651</td>
<td>0.998</td>
<td>0.825</td>
<td>0.692</td>
<td>0.535</td>
</tr>
<tr>
<td>05R</td>
<td>0.997</td>
<td>0.758</td>
<td>0.998</td>
<td>0.878</td>
<td>0.746</td>
<td>0.634</td>
</tr>
<tr>
<td>06L</td>
<td>0.998</td>
<td>0.620</td>
<td>0.999</td>
<td>0.809</td>
<td>0.694</td>
<td>0.548</td>
</tr>
<tr>
<td>06R</td>
<td>0.998</td>
<td>0.705</td>
<td>0.999</td>
<td>0.852</td>
<td>0.726</td>
<td>0.503</td>
</tr>
<tr>
<td>07L</td>
<td>0.997</td>
<td>0.722</td>
<td>0.998</td>
<td>0.860</td>
<td>0.678</td>
<td>0.574</td>
</tr>
<tr>
<td>07R</td>
<td>0.997</td>
<td>0.630</td>
<td>0.999</td>
<td>0.815</td>
<td>0.675</td>
<td>0.553</td>
</tr>
<tr>
<td>08L</td>
<td>0.997</td>
<td>0.850</td>
<td>0.998</td>
<td>0.924</td>
<td>0.650</td>
<td>0.507</td>
</tr>
<tr>
<td>08R</td>
<td>0.997</td>
<td>0.867</td>
<td>0.997</td>
<td>0.932</td>
<td>0.641</td>
<td>0.566</td>
</tr>
<tr>
<td>09L</td>
<td>0.991</td>
<td>0.816</td>
<td>0.991</td>
<td>0.903</td>
<td>0.315</td>
<td>0.509</td>
</tr>
<tr>
<td>09R</td>
<td>0.994</td>
<td>0.915</td>
<td>0.994</td>
<td>0.955</td>
<td>0.368</td>
<td>0.493</td>
</tr>
<tr>
<td>10L</td>
<td>0.994</td>
<td>0.672</td>
<td>0.995</td>
<td>0.834</td>
<td>0.408</td>
<td>0.483</td>
</tr>
<tr>
<td>10R</td>
<td>0.996</td>
<td>0.900</td>
<td>0.996</td>
<td>0.948</td>
<td>0.590</td>
<td>0.458</td>
</tr>
<tr>
<td>11L</td>
<td>0.997</td>
<td>0.933</td>
<td>0.997</td>
<td>0.965</td>
<td>0.617</td>
<td>0.622</td>
</tr>
<tr>
<td>11R</td>
<td>0.998</td>
<td>0.878</td>
<td>0.998</td>
<td>0.938</td>
<td>0.656</td>
<td>0.570</td>
</tr>
<tr>
<td>12L</td>
<td>0.997</td>
<td>0.699</td>
<td>0.998</td>
<td>0.848</td>
<td>0.566</td>
<td>0.609</td>
</tr>
<tr>
<td>12R</td>
<td>0.998</td>
<td>0.770</td>
<td>0.999</td>
<td>0.885</td>
<td>0.693</td>
<td>0.513</td>
</tr>
<tr>
<td>13L</td>
<td>0.998</td>
<td>0.833</td>
<td>0.998</td>
<td>0.916</td>
<td>0.667</td>
<td>0.503</td>
</tr>
<tr>
<td>13R</td>
<td>0.998</td>
<td>0.650</td>
<td>0.998</td>
<td>0.824</td>
<td>0.548</td>
<td>0.445</td>
</tr>
<tr>
<td>14L</td>
<td>0.997</td>
<td>0.813</td>
<td>0.998</td>
<td>0.905</td>
<td>0.652</td>
<td>0.639</td>
</tr>
<tr>
<td>14R</td>
<td>0.997</td>
<td>0.692</td>
<td>0.999</td>
<td>0.845</td>
<td>0.628</td>
<td>0.554</td>
</tr>
<tr>
<td>Avg</td>
<td>0.997</td>
<td>0.757</td>
<td>0.997</td>
<td>0.877</td>
<td>0.629</td>
<td>0.547</td>
</tr>
</tbody>
</table>
The isolated pixels with an area less than 25 pixels in the image \(I_{s}\) are trimmed and fused with the B-COSFIRE filter segmented image of Block-II by AND-type operation. The physical stats (eccentricity and area) are utilized for the rejection of nonvessel structures. The vessel structures have a higher area and eccentricity as their pixels are linked and having an elongated structure. Figure 5 indicates the graphical results of the \(I_{avg}\), \(I_{d}\), and \(I_{s}\).

3.3. **Block-II: Vessel Structure-Based Enhancement and Segmentation.** In Block-II, the masked image of the Block-I is used as an input for vessel structure-based enhancement and RVS. B-COSFIRE filter [5] is applied for contrast improvement of vessel structures that will enhance noise also along with the enhancement of vessel structures if the image is not preprocessed. Therefore, the masked image is used for further processing. B-COSFIRE filter produced two results: binary segmented image \(I_{s_c}\) and vessel structure-based enhanced image \(I_{s_e}\). The outputs of B-COSFIRE filter are displayed in Figure 6. The AND-type operation is used to combine \(I_{s}\) with \(I_{s_c}\) that produced output image denoted by \(I_{And}\). The effect of AND-type operation is shown in Figure 7, which demonstrates that if an alternative operator like OR-type is utilized, it will introduce noise and misclassification. The advantage of using an AND-type operator is exposed in Figure 8 by displaying the visual results with and without using the AND-type operator. The \(I_{And}\) is postprocessed \(I_{p}\) and multiplied with \(I_{E}\), which is further thresholded to obtain a segmented image \(I_{s}\). Pixel-by-pixel multiplication aims at ensuring the detection of vessels at their correct position. The logical OR-type operation is used to produce the final result by coupling of \(I_{p}\) and \(I_{s}\). The visual effects of the OR-type operator are presented in Figure 9.

The B-COSFIRE filter application includes convolution with difference of Gaussian (DoG) filters, its blurring effects, shifting the blurred responses, and an approximate point-
wise weighted geometric mean (GM). A DoG function $\text{DoG}_\sigma(x, y)$ is given by [5]

$$\text{DoG}_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-x^2 + y^2}{2\sigma^2}\right) - \frac{1}{2\pi(0.5\sigma)^2} \exp\left(\frac{-x^2 + y^2}{2(0.5\sigma)^2}\right),$$

where $\sigma$ is the standard deviation (SD) of the Gaussian function (GF) that decides the range of the boundary. $0.5\sigma$ is manually set SD value of the internal GF, and $(x, y)$ symbolizes the pixel position of the image. Response of DoG filter $C_\sigma(x, y)$ with kernel function of $\text{DoG}_\sigma(x - x', y - y')$ has been estimated by convolution, where $(x', y')$ denotes pixels intensity distribution.

$C_\sigma(x, y) = \left| I \ast \text{DoG}_\sigma \right|^+$, \hspace{1cm} (3)

where $\left| \cdot \right|^+$ represents the half-wave rectification process to reject negative values.

In the B-COSFIRE filter, three factors $(\sigma_i, \rho_i, \theta_i)$ are used to represent each point $i$, where $\sigma_i$ = SD of the DoG filter, while $\rho_i$ and $\theta_i$ denote the polar coordinates. This set of parameters is indicated by $S = \{(\sigma_i, \rho_i, \theta_i) | i = 1, \ldots, n\}$, where $n$ represents the figure of measured DoG responses. The blurring process indicates the calculation of the extreme limit of the weighted thresholded responses of a DoG filter. The blurring operation is shown as follows.

$$\sigma' = \sigma_0' + \alpha \rho_i,$$

where $\sigma_0'$ and $\alpha$ are constants. Each DoG-blurred outcome is moved in the reverse direction to $\theta_i$ by a gap $\rho_i$, and as a result, they can merge at the support center of the B-COSFIRE filter. Blurred and shifted responses of the DoG filter is indicated by $S_{\sigma, \rho, \theta}(x, y)$ for every tuple $(\sigma_i, \rho_i, \theta_i)$ in set $S$. The $i_{th}$ blurred and shifted response of the DoG filter is defined as

$$S_{\sigma, \rho, \theta}(x, y) = \max_{x', y'} \left\{ c_\sigma \left( x - \delta x_i - x', y - \delta y_i - y' \right) G_\sigma \left( x', y' \right) \right\}.$$

\hspace{1cm} (5)
where $-3\sigma' \leq x', y' \leq 3\sigma'$. The output of the filter is shown as GM of all the blurred and shifted DoG responses.

$$r_s(x, y) = \left| \prod_{i=1}^{S} \left( S_{\sigma_i, \rho_i, \phi_i}(x, y)^{\omega_i} \right) \right|_t,$$

(6)

where $\omega_i = \exp^{-\rho_i^2/2\sigma^2}$ and $|.|_t$ symbolizes the thresholding response at $t, (0 \leq t \leq 1)$. Equation (6) represents the AND-type operation that is attained by the B-COSFIRE filter only when all DoG filter responses $S_{\sigma_i, \rho_i, \phi_i}$ are larger than zero.

The overall step-by-step visual results according to the block diagram (Figure 2) are portrayed in Figure 10.

4. Experimental Outcomes and Deliberation

This section will provide the information about datasets, performance metrics, analysis of experimental results, and time complexity of the proposed method.

4.1. Datasets. The proposed system obtained remarkable results on the freely online available datasets: DRIVE [11, 12], STARE [53], HRF [54], and CHASE_DB1 [55]. The magnificence of the framework is justified in terms of assessment with state-of-the-art systems. The datasets used for endorsement of the suggested framework are encapsulated in Table 1. The manually labeled results in all datasets are utilized as a gold standard for performance assessment of the proposed framework.

4.2. Performance Judgment Parameters. The quantitative results are obtained by equating the proposed segmentation’s
There are numerous performance standards mentioned in the literature. The performance metrics used for evaluation of the proposed framework are visible in Table 2. Six performance standards (Acc, Sn, Sp, AUC, MCC, and CAL) are selected for the justification of the proposed methodology. The Acc metric tells about the overall valuation of the proposed method. Sn is a measure of the quantity of correctly identified positive instances. Sp is the proportion of negative instances that are correctly identified. AUC is the area under the receiver operating characteristic curve, which provides a measure of the model’s ability to discriminate between classes. MCC is the Matthews correlation coefficient, which is a robust measure of the agreement of a binary classification. CAL is the Cramer’s V, a measure of association between two nominal variables.

Table 7: Mean results comparison of the proposed method on different datasets.

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Images</th>
<th>Acc</th>
<th>Sn</th>
<th>Sp</th>
<th>AUC</th>
<th>MCC</th>
<th>CAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRIVE (1st observer)</td>
<td>20</td>
<td>0.954</td>
<td>0.766</td>
<td>0.972</td>
<td>0.869</td>
<td>0.721</td>
<td>0.690</td>
</tr>
<tr>
<td>DRIVE (2nd observer)</td>
<td>0.958</td>
<td>0.797</td>
<td>0.973</td>
<td>0.885</td>
<td>0.739</td>
<td>0.696</td>
<td></td>
</tr>
<tr>
<td>HRF_Nonormal</td>
<td>15</td>
<td>0.963</td>
<td>0.801</td>
<td>0.982</td>
<td>0.881</td>
<td>0.761</td>
<td>0.699</td>
</tr>
<tr>
<td>HRF_DR</td>
<td>0.954</td>
<td>0.726</td>
<td>0.974</td>
<td>0.826</td>
<td>0.666</td>
<td>0.603</td>
<td></td>
</tr>
<tr>
<td>HRF_Glaucoma</td>
<td>0.959</td>
<td>0.790</td>
<td>0.979</td>
<td>0.881</td>
<td>0.703</td>
<td>0.667</td>
<td></td>
</tr>
<tr>
<td>HRF_Average</td>
<td>45</td>
<td>0.959</td>
<td>0.772</td>
<td>0.978</td>
<td>0.863</td>
<td>0.710</td>
<td>0.656</td>
</tr>
<tr>
<td>CHASE_DB1 (1st observer)</td>
<td>28</td>
<td>0.997</td>
<td>0.757</td>
<td>0.97</td>
<td>0.877</td>
<td>0.629</td>
<td>0.547</td>
</tr>
<tr>
<td>CHASE_DB1 (2nd observer)</td>
<td>0.996</td>
<td>0.814</td>
<td>0.996</td>
<td>0.905</td>
<td>0.569</td>
<td>0.547</td>
<td></td>
</tr>
<tr>
<td>STARE</td>
<td>20</td>
<td>0.996</td>
<td>0.792</td>
<td>0.997</td>
<td>0.895</td>
<td>0.707</td>
<td>0.566</td>
</tr>
</tbody>
</table>

Figure 14: Best and worst cases visual results from the CHASE_DB1 dataset.
classified vessel pixels, while Sp is an assessment of the competency of differentiating nonvessel pixels. The AUC is the ratio of Sn and Sp. The MCC [5, 56] is a more appropriate indicator of the accuracy of binary categorization in the case of unbalanced structures. For a comprehensive judgment of the superiority of segmentation, the CAL metric [57, 58] is computed. This metric provides justification based on the properties (connectivity-area-length) of the segmented structures beyond the correctly classified image pixels.

In Table 2, \( N = TN + TP + FN + FP \), \( S = (TP + FN)/N \) and \( P = (TP + FP)/N \) [58]. The terms TP, TN, FP, and FN denote the true positive (exactly matched vessel pixels), true...
negative (exactly matched nonvessel pixels), false positive (invalidly predicated vessel pixels), and false negative (invalidly predicated nonvessel pixels), correspondingly.

Let $I_S$ be the extracted final binary image and $I_G$ the corresponding manual segmented image. The considered metric evaluates the following [57, 58]:

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(i) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

<table>
<thead>
<tr>
<th>Technique</th>
<th>Year</th>
<th>HRF Acc</th>
<th>HRF Sn</th>
<th>HRF Sp</th>
<th>HRF AUC</th>
<th>CHASE_DB1 Acc</th>
<th>CHASE_DB1 Sn</th>
<th>CHASE_DB1 Sp</th>
<th>CHASE_DB1 AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odstrcilik [54]</td>
<td>2013</td>
<td>0.949</td>
<td>0.774</td>
<td>0.967</td>
<td>0.871</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Azzopardi [5]</td>
<td>2015</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.939</td>
<td>0.759</td>
<td>0.959</td>
<td>0.859</td>
</tr>
<tr>
<td>Zhang [66]</td>
<td>2016</td>
<td>0.957</td>
<td>0.798</td>
<td>0.974</td>
<td>0.886</td>
<td>0.946</td>
<td>0.763</td>
<td>0.968</td>
<td>0.866</td>
</tr>
<tr>
<td>Biswal [29]</td>
<td>2017</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.940</td>
<td>0.760</td>
<td>0.970</td>
<td>0.865</td>
</tr>
<tr>
<td>Rodrigues [67]</td>
<td>2017</td>
<td>0.948</td>
<td>0.722</td>
<td>0.964</td>
<td>0.843</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Badawi [3]</td>
<td>2018</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.953</td>
<td>0.800</td>
<td>0.964</td>
<td>0.882</td>
</tr>
</tbody>
</table>

Let $I_G$ be the extracted final binary image and $I_G$ the corresponding manual segmented image. The considered metric evaluates the following [57, 58]:

(i) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(i) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(iii) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(i) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(iii) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(i) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(iii) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(i) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(iii) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(i) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(iii) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(i) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]

(iii) Connectivity (C): it calculates the fragmentation grade of $I_S$ with respect to the manual segmentation $I_G$ and penalizes fragmented segmentation. It is computed as

\[
C(I_S, I_G) = 1 - \min \left( 1, \frac{\#_C(I_G) - \#_C(I_S)}{\#(I_G)} \right),
\]
We set \( \epsilon \) where \( \delta \) logical dilation with a disc SE of \( \phi \) pixel radius. The value of \( \epsilon \) is a skeletonization process and \( \delta \) is a morphological dilation with a disc SE of \( \beta \) pixel radius. The value of \( \beta \) controls the tolerance to lines of various sizes. We set \( \epsilon = 2 \) and \( \beta = 2 \). The final assessment parameter, named CAL, is demarcated as \( \phi \) and \( \mu \) respectively. The average statistical results of the proposed framework for all selected datasets are displayed in Table 7, which reflects that the highest mean score of Acc 0.997, Sn 0.814, Sp 0.997, and AUC 0.905 is achieved on the DRIVE_DB1 dataset. The lowest FPR is also observed using the same dataset. The highest value of MCC 0.761 and CAL 0.699 is recorded on the HRF dataset. The highest value of each parameter is italicized in the respective column of the Table 7.

The average performance parameter scores of the proposed framework on the DRIVE and STARE datasets are compared with the existing literature in Table 8, while Table 9 shows the result comparison of the HRF and CHASE_DB1 datasets. The Acc, Sn, and Sp results of all techniques in Tables 8 and 9 are acquired from their respective published articles while the AUC result is calculated by using the formula in Table 2.

In Table 8, the obtained results of the framework are compared with 19 unsupervised and 18 supervised existing techniques. The proposed framework achieved the highest Acc result than all unsupervised methods on the DRIVE dataset except Khan et al. [40], Memari et al. [59] which is 0.003%, and Fan et al. [60] which is 0.002% better than ours. The supervised methods Ricci and Perfetti [14], Lupascu et al. [15], Wang et al. [16], Zhu et al. [18], Thangaraj et al. [22], Memari et al. [23], Khowaja et al. [26], and Fan et al. [61] show 0.001%, 0.001%, 0.019%, 0.003%, 0.003%, 0.014%, 0.017%, and 0.008% better results than the proposed method, respectively. But some of these methods are only validated on one dataset, which reflects that they are tuned for a single dataset. Some of these methods produce a very low AUC score, which is a trade-off between Sn and Sp. Moreover, supervised methods are computationally very

The framework performs well on both healthy and pathological images of all selected datasets. The statistical results in Tables 3–6 validates that the suggested system is robust and has the capability to handle the bright lesions images of the STARE dataset, higher resolution images of the HRF dataset, low resolution images of the DRIVE dataset, and left/right eyes images of the CHASE_DB1 dataset. The anatomical structures are also efficiently omitted to avoid any misclassification.

### Table 11: Processing time evaluation of the systems.

<table>
<thead>
<tr>
<th>Method</th>
<th>Time</th>
<th>Hardware particulars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roychowdhury [17]</td>
<td>3.11 sec</td>
<td>Intel Core i3 CPU 2.6 GHz, 2 GB RAM</td>
</tr>
<tr>
<td>Zhu [18]</td>
<td>12.160 sec</td>
<td>4.0 GHz Intel i7-4790K CPU and 32 GB RAM</td>
</tr>
<tr>
<td>Memari [23]</td>
<td>8.2 mins</td>
<td>Intel i5-M480 CPU, 2.67 GHz, 4 GB RAM</td>
</tr>
<tr>
<td>Biswal [29]</td>
<td>3.3 sec</td>
<td>Intel i3 (4010U CPU) 1.7 GHz, 4 GB RAM</td>
</tr>
<tr>
<td>Badawi [3]</td>
<td>8 sec</td>
<td>CPU 2.7 GHz, 16 GB RAM</td>
</tr>
<tr>
<td>Yue [50]</td>
<td>4.6 sec</td>
<td>Intel i5-6200U CPU 2.3 GHz, 8 GB RAM</td>
</tr>
<tr>
<td>Khan [39]</td>
<td>6.1 sec</td>
<td>5•Intel Core i3 CPU, 2.53 GHz, 4 GB RAM</td>
</tr>
<tr>
<td>Khan [40]</td>
<td>1.56 sec</td>
<td></td>
</tr>
<tr>
<td>Azzopardi [5]</td>
<td>11.83 sec</td>
<td></td>
</tr>
<tr>
<td>Vlachos [47]</td>
<td>9.3 sec</td>
<td></td>
</tr>
<tr>
<td>Bankhead [38]</td>
<td>22.45 sec</td>
<td></td>
</tr>
<tr>
<td>Proposed</td>
<td>5.5 sec</td>
<td></td>
</tr>
</tbody>
</table>

(ii) Area (\( A \)): it estimates the intersecting area between \( I_S \) and \( I_G \), based on the jaccard coefficient. Let \( \delta_\epsilon(\cdot) \) be a morphological dilation that utilizes a disc structuring element (SE) with a radius of \( \epsilon \) pixels. The magnitude \( A \) is calculated as follows:

\[
A(I_S, I_G) = \frac{\#((\delta_\epsilon(I_S) \cap I_G) \cup (I_S \cap \delta_\epsilon(I_G)))}{\#(I_G \cup I_S)}.
\]

The value of \( \epsilon \) controls the tolerance to lines of various sizes. We set \( \epsilon = 2 \).

(iii) Length (\( L \)): it determines the equivalent degree between \( I_S \) and \( I_G \) by computing the length of the two line networks:

\[
L(I_S, I_G) = \frac{\#((\phi(I_S) \cap \delta_\beta(I_G)) \cup (\delta_\beta(I_S) \cap \phi(I_G)))}{\#(\phi(I_S) \cup \phi(I_G))},
\]

where \( \phi(\cdot) \) is a skeletonization process and \( \delta_\beta(\cdot) \) is a morphological dilation with a disc SE of \( \beta \) pixel radius. The value of \( \beta \) controls the tolerance to dissimilarity of the line tracing output. We set \( \beta = 2 \). The final assessment parameter, named CAL, is demarcated as \( f(C, A, L) = C \cdot A \cdot AL \).

#### 4.3. Experimental Results and Inspection

The success of the proposed framework is established by utilizing four freely obtainable datasets: DRIVE, STARE, HRF, and CHASE_DB1 for testing and evaluation. The average performance parameters results in Table 3 are computed by processing 20 test images of the DRIVE and STARE datasets. The performance scores of HRF dataset (15 normal images, 15 DR, and 15 glaucomatous) and CHASE_DB1 are presented in Tables 4 and 5 and Table 6, respectively. The best and worst results within Tables 3–6 are highlighted in italic font. The best and worst image results from each dataset are selected based on their accuracy’s scores. Their pictorial results are shown in Figures 11–14.
expensive. In the case of the STARE dataset, the framework produced highest Acc scores than all other methods. Table 9 reflects that there are very few techniques that used both HRF and CHASE_DB1 datasets for validation. The Acc score of the framework is higher than both supervised and unsupervised approaches on the HRF and CHASE_DB1 datasets except Soomro et al. [62] and Fan et al. [61] which is slightly higher than ours on HRF dataset only. Fan et al. [61] showed higher Sp value than all other methods on the HRF dataset. The highest Sp value on CHASE_DB1 dataset is obtained by the proposed method. All the other supervised and unsupervised methods acquired a bit greater or equivalent values of Sn and AUC metric on the HRF and CHASE_DB1 datasets as compared to ours.

In Table 10, the MCC and CAL values are recorded by the proposed method and other existing supervised and unsupervised methods. The MCC and CAL values of Chauduri et al. [31], Niemeijer et al. [11], Hoover et al. [53], and BCOSFIRE [5] are calculated by utilizing their publicly accessible segmented images. The results of Fraz et al. [68, 69], RUSTICO [58], Yang et al. [70, 71], Vega et al. [72], FC-CRF [73], and UP-CRF [73] are extracted from their published articles.

The average value of MCC attained by the proposed method is higher than all compared unsupervised approaches on the DRIVE, STARE, and HRF datasets, while it is statistically lower than the supervised methods (i.e., FC-CRF [73] and UP-CRF [73]) on the DRIVE, STARE, and CHASE_DB1 datasets. The CAL value of the proposed method is observed higher than all supervised and unsupervised methods on the HRF dataset, while it is statistically lower than or equivalent to CAL values of other methods on the DRIVE, STARE, and CHASE_DB1 datasets.

4.3.1. Processing Time. The proposed framework processes a single image in a very short time as equated to other approaches in Table 11. The time values are computed on the single image taken from the DRIVE and STARE datasets.

5. Conclusion

Vessel extraction is momentous for inspecting abnormalities inside and around the retinal periphery. The retinal vessel segmentation is a challenging task due to the existence of pathologies, unpredictable dimensions and contour of the vessels, nonuniform clarification, and structural inconsistency between subjects. The proposed methodology is consistent, faster, and completely automated for isolation of retinal vascular network. The success of the proposed framework is evidently revealed by the RVS statistics on the DRIVE, STARE, HRF, and CHASE_DB1 datasets. The eradication of anomalous structures prior to enhancement boosted the efficiency of the proposed method. The application of logical operators avoids misclassification of foreground pixels which enhances the accuracy and makes the method robust. Pictorial representation validates that the framework is able to segment both healthy and unhealthy images. Furthermore, the method does not include any hand-marked data by experts for training, which makes it computationally fast.

Data Availability

All the data are fully available within the manuscript without any restriction.

Conflicts of Interest

The authors declare no conflict of interest.

References


