Automatic Retinal Vessel Segmentation through Gabor Filtering, Principal Component Analysis and Ensemble of Classifiers (C4.5 with Bagging)

1R. Geetha Ramani and 2Lakshmi Balasubramanian

1Associate Professor, Department of Information Science and Technology, College of Engineering, Guindy, Anna University, Chennai – 600025, India.
2Research Scholar, Department of Information Science and Technology, College of Engineering, Guindy, Anna University, Chennai – 600025, India.

ABSTRACT

Background: Computational techniques have been exploited in the domain of ophthalmology for better retinal fundus image analysis. Blood vessels, one of the most important retinal anatomical structures are analyzed for diagnosis of many diseases like retinopathy, occlusion and much other vision threatening diseases. Objective: In this paper, retinal blood vessel segmentation is attempted through sequential application of image pre-processing data mining and image post-processing techniques. The image pre-processing phase comprises of image cropping, color space conversion and color channel extraction, contrast enhancement, Gabor filtering and half wave rectification. Image pre-processing is followed by the data mining phase which includes feature vector construction through principal component analysis and classification through C4.5 with bagging. The output of the data mining phase is post processed through connected component analysis. Results: The proposed methodology is evaluated on a publicly available DRIVE database. The results reported 95.08% accuracy on the test set of the dataset. A comparison of these results with the existing techniques is also reported. Conclusion: This methodology can help ophthalmologists in better and faster retinal analysis and hence earlier treatment to the patients.

INTRODUCTION

Retinal image analysis has been an interesting area of research in the recent times. Fundus imaging of the retina reveals its anatomical structures viz., blood vessels, optic disc, macula, fovea and any other abnormal structures such as exudates and microneurysms if present (Niall Patton et al., 2006). Changes in the properties associated with the blood vasculature aids in identification of various retinal diseases such as Diabetic Retinopathy, Retinal artery and vein occlusion, Hypertensive Retinopathy etc. Disorders such as Hypertension, Diabetes, Cardiovascular abnormalities and Stroke also cause alterations in the blood vessels of the retina. Hence, analysis of the retinal blood vessel positions itself as a core area in retinal image analysis. Manual segmentation of blood vessels requires high expertise. It is also a time consuming task. Hence, employing computational techniques for automatic retinal blood vessel segmentation would highly simplify the task of the ophthalmologists.

In this work, automatic retinal blood vessel segmentation is attempted through image pre-processing followed by data mining techniques and subsequently image post processing.

Subsequent sections present the literature work on automatic blood vessel segmentation, the proposed methodology and the experimental results.

1. Related Work:

Modifications in retinal blood vessels reveal many retinal and other disorders. Hence automatic blood vessel segmentation helps in efficient disease diagnosis. A few works have been previously attempted in this area. The works which report results on DRIVE database are concisely presented below.
In 2004, supervised classification through KNN classifier on 31 features including the green intensity plane, Gaussian and its derivative at various scales resulted in an accuracy of 94.16% (Niemeijer et al., 2004). In 2006, retinal blood vessel segmentation was attempted through extraction of vessel centerlines followed by vessel filling through region growing achieving an accuracy of 94.63% (Mendonca and Campilho, 2006). In 2006, yet another attempt to classify vessel pixels through Bayes classifier was employed on feature vector formed with pixel intensities and Gabor wavelet responses at various scales, yielding an accuracy of 94.66% (Soares et al., 2006). In 2007, a multi-pass region growing algorithm using local maxima over scales of magnitude of gradient and maximum principal curvature of hessian tensor obtained an accuracy of 93.44% (Martinez-Perez et al., 2007). In 2010, a local adaptive thresholding was applied to form a binary image, in which the largest component was considered as large vessels and SVM was applied to the residual segments followed by application of track growing to grow the small vessels (Xu and Luo, 2010). This methodology yielded 93.28% accuracy. In 2011, neural network based supervised classification was applied on a 7-D feature vector comprising of gray level and moment invariant features providing an accuracy of 94.52% (Diego Marin et al., 2011). In 2012, Isotropic Undecimated Wavelet Transform was applied for vessel segmentation resulting in an accuracy of 93.71% (Bankheard et al., 2012). In 2013, another methodology with Gabor wavelet for vessel enhancement and multi-layered thresholding for vessel segmentation attained an accuracy of 94.69% (Usman Akram and Shoab, 2013). In 2014, an approach to segment vessels using combination of bee colony optimization for identifying cluster centers followed by fuzzy c-means clustering achieved an accuracy of 93.9% (Eid et al., 2014). Yet another framework which comprised of bias correction and adaptive histogram equalization for vessel enhancement and probabilistic modeling optimized by expectation maximization algorithm for vessel segmentation obtained an accuracy of 94.10% (Djibril Kaba et al., 2014). In the same year, an enhanced multi-scale operator was proposed, which combines the responses of line operator at various scales by assigning appropriate weights to each scale, accomplishing an accuracy of 94.15% (Yanli Hou, 2014). One more attempt in 2014, employed multilayer perceptron on RGB color channels to segment the blood vessels and this technique achieved an accuracy of 95.03% (Wilfred and Edward, 2014).

In this article, a methodology to segment retinal blood vessels through sequential application of image pre-processing, data mining and image post processing techniques to the fundus image is proposed. The following section describes the proposed methodology.

2. Proposed Methodology:

Retinal vessel segmentation through application of computational intelligence would be highly appreciated by the ophthalmologists. The proposed framework is depicted in fig. 1.

![Proposed Framework for Automatic Retinal Vessel Segmentation](image-url)
The proposed framework to segment the retinal blood vessels includes sequential application of image pre-processing, data mining techniques and post processing of the classified pixels. Image pre-processing comprises of image cropping, color space conversion and color channel extraction, contrast enhancement, application of Gabor filtering and half wave rectification done in the specified order. Then, a feature vector representation including the pixel intensities and Gabor responses for every pixel is formed which serves as input to the data mining phase. The data mining phase includes application of principal component analysis to the feature vector followed by classification through the ensemble classification model. Each pixel is classified either as a vessel or a non-vessel pixel. The classified image hence represents a binary vessel segmented image. Now, this image is post processed using connected component analysis, yielding to final segmented vessel image.

The entire framework is described in detail in the following sub-sections.

**Image Pre-Processing:**

The fundus image is pre-processed to make it suitable for further processing. The image pre-processing phase includes image cropping, color space conversion and color channel extraction, contrast enhancement, Gabor filtering and half wave rectification.

**Image Cropping:**

The image is cropped such that border of the image touches the field of view in all the sides. This reduces the processing area of the image thus reducing the computational complexity. The cropped area will not contain any vessel pixels as it is the area out of the region of field of view. Hence the pixels discarded in the step (approximately 40,000 pixels; varies with every image) would be included during performance evaluation as correctly identified non-vessel pixels.

**Color Space Conversion and Color Channel Extraction:**

Followed by image cropping, the RGB color model is converted into YCbCr, L*a*b and Gaussian models (Guesebroek et al., 2001). Since the blood vessels are well depicted at high contrast, the color channel displaying the highest contrast is chosen from every color model and considered further. Hence Green channel from RGB; Y Channel from YCbCr; L channel from L*a*b model and G1 channel from Gaussian color model is considered for further processing.

**Contrast Enhancement:**

The four channel images thus obtained are contrast enhanced using Contrast Limited Adaptive Histogram Equalization (CLAHE) (Pizer et al., 1987). The CLAHE technique is adopted so that the thin vessels are also enhanced.

**Gabor Filtering:**

2-D Gabor filters (Fogel and Sagi, 1989), which are sinusoidally modulated Gaussian functions, have been used to further enhance the blood vessels in the contrast enhanced images. The Gabor responses obtained at wavelength 9, 10 and 11 are investigated further. Since this step is applied to four images, twelve resultant images are obtained.

**Half wave rectification:**

The process of half wave Rectification operates on the Gabor response images based on a percentage value (10 in this case) of the maximum intensity of the image. Pixels with Gabor response value lesser than the specified percentage will be set to zero.

After the image pre-processing phase, a feature vector representing each pixel is formed with the twelve half wave rectified Gabor responses and the pixel intensities of the Green channel image. The feature vector is given as input to the data mining phase.

**Data Mining:**

The image pre-processing phase is followed by the application of data mining techniques (GeethaRamani, Lakshmi and Shomona, 2012a) to the feature vector formed from the previous phase. The data mining phase comprises of feature vector formulation through principal component analysis and classification through ensemble of classifiers.

**Feature Contraction through Principal Component Analysis:**

Principal Component Analysis (Jollife, 1986) has always been applied to analyse patterns in high dimensional data. A new set of more representative features is constructed from the feature vector through application of Principal Component Analysis. This new set of features is considered for further processing.
Classification through Ensemble of Classifiers:
Supervised Classification techniques (Shomona Gracia Jacob and Geetha Ramani, 2012) can be applied to the feature vector formed from the preceding step. The segmentation of vessels is viewed here as a binary classification problem i.e., the pixel belongs to either a vessel class or non-vessel class. The supervised classification technique demands training data to form the decision rules (Shomona Gracia Jacob, Geetha Ramani and Nancy, 2012). In this work, one of the images is taken as reference and used for training. The performance of the classifier is always increased through a meta-learning procedure. In this study, bagging technique (Breiman 1996) is applied to the classifier model to achieve increased performance.

C4.5 (Steven, 1994) with Bagging is applied on the data to build the classification model to classify the vessel and non-vessel pixels. The output of this stage would be a binary image consisting of vessel pixels represented as white and non-vessel pixels represented as black. This image is further post processed to eliminate false vessels and false non-vessels.

Image Post-Processing:
The result from the data mining phase is a classified binary image. This image contained a few white pixels hanging all over the image. Since, even small vessels would occupy greater than 10 pixels, all components lesser than 10 pixels were eliminated. The entire process comprising of image pre-processing, data mining and image post-processing is illustrated on a sample fundus image in Fig. 2.

![Fig. 2: Sample illustration of the proposed framework on image 20 of the DRIVE dataset.](image-url)

Experimental results of the proposed methodology are discussed in the next section.

RESULTS AND DISCUSSION
The proposed methodology is validated on the fundus images from the publicly available database DRIVE (Niemeijer et al., 2004). DRIVE database is the widely used database for assessing the performance of the vessel segmentation algorithms. The database consists of forty images, 20 of which are packed as training and the remaining 20 as test set. The results reported here are the average performance measures for the twenty
images in the test set of the database. A manual graded ground truth is available for these images and a second human observer prediction is also provided.

The proposed methodology is implanted using Matlab r2008a and weka 3.6.11. Matlab r2008a is utilised for implementing the image pre-processing and image post processing phases. Weka, an open source data mining tool is used for implementing the data mining phases. Performance comparison is facilitated through average accuracy of the segmentation algorithm on all the twenty images.

Various classification algorithms (GeethaRamani and Jacob, 2013) were attempted for the purpose of classification of vessels and non-vessels. C4.5 with bagging was observed to yield the best results. C4.5 algorithm utilises gain ratio for the purpose of evaluation of the attributes to form the decision tree. As supervised classification algorithms demands training data, one of the images from the training set is taken as reference and considered for training (first image of the training set in this case). The classification prediction from the C4.5 algorithm yields a binary segmented image, which is post processed for enhanced performance.

Accuracy (GeethaRamani, Lakshmi and Shomona, 2012b, 2013)(GeethaRamani and Lakshmi, 2013) is calculated as the ratio of sum of correctly classified non-vessel pixels and correctly classified vessel pixels to the total number of image pixels. Sensitivity is defined as the ratio of correctly identified vessel pixels to the total number of actual vessel pixels. Specificity is defined as the ratio of correctly identified non-vessel pixels to the total number of non vessel pixels. In all these calculations, the total numbers of non-vessel pixels indicate the sum of pixels discarded during image cropping and the correctly identified non-vessel pixels in the current step.

Table 1 presents the accuracy of the proposed algorithm in every image.

<table>
<thead>
<tr>
<th>Image ID</th>
<th>Accuracy (%)</th>
<th>Image ID</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95.58</td>
<td>11</td>
<td>94.94</td>
</tr>
<tr>
<td>2</td>
<td>95.84</td>
<td>12</td>
<td>95.46</td>
</tr>
<tr>
<td>3</td>
<td>94.19</td>
<td>13</td>
<td>94.97</td>
</tr>
<tr>
<td>4</td>
<td>95.31</td>
<td>14</td>
<td>95.13</td>
</tr>
<tr>
<td>5</td>
<td>94.85</td>
<td>15</td>
<td>91.74</td>
</tr>
<tr>
<td>6</td>
<td>95.00</td>
<td>16</td>
<td>95.63</td>
</tr>
<tr>
<td>7</td>
<td>94.78</td>
<td>17</td>
<td>95.55</td>
</tr>
<tr>
<td>8</td>
<td>94.78</td>
<td>18</td>
<td>95.91</td>
</tr>
<tr>
<td>9</td>
<td>95.56</td>
<td>19</td>
<td>95.16</td>
</tr>
<tr>
<td>10</td>
<td>95.21</td>
<td>20</td>
<td>95.96</td>
</tr>
</tbody>
</table>

The proposed technique thus achieved an average accuracy, sensitivity and specificity of 95.08%, 68.20% and 97.70% respectively. Performance comparison of the presented technique with the existing techniques is presented in Table 2.

```markdown
<table>
<thead>
<tr>
<th>Technique</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed technique</td>
<td>95.08</td>
</tr>
<tr>
<td>Wilfred and Edward, 2014</td>
<td>95.03</td>
</tr>
<tr>
<td>Second Observer</td>
<td>94.73</td>
</tr>
<tr>
<td>Usman Akram and Shoaib, 2013</td>
<td>94.69</td>
</tr>
<tr>
<td>Soares et al., 2006</td>
<td>94.66</td>
</tr>
<tr>
<td>Mendonca and Campilho, 2006</td>
<td>94.63</td>
</tr>
<tr>
<td>Diego Marin et al., 2011</td>
<td>94.52</td>
</tr>
<tr>
<td>Niemwijer et al., 2004</td>
<td>94.16</td>
</tr>
<tr>
<td>Yanli Hou, 2014</td>
<td>94.15</td>
</tr>
<tr>
<td>Djibril Kaba et al., 2014</td>
<td>94.10</td>
</tr>
<tr>
<td>Eid et al., 2014</td>
<td>93.90</td>
</tr>
<tr>
<td>Bankhead et al., 2012</td>
<td>93.71</td>
</tr>
<tr>
<td>Martinez Perez et al., 2007</td>
<td>93.44</td>
</tr>
<tr>
<td>Xu and Luo, 2010</td>
<td>93.28</td>
</tr>
</tbody>
</table>
```

The proposed methodology is thus efficient in segmenting blood vessels from the retinal fundus images. It can be of great use to the ophthalmologist society in analyzing the patient’s retina.

4. Conclusion:

Computational techniques viz. image processing and data mining are exploited for the purpose of retinal image analysis. Changes in the properties associated with the retinal blood vessels are the main indicator to diagnose many sight threatening diseases. In this paper, segmentation of retinal blood vessel is attempted through sequential application of image cropping, color space conversion and color channel extraction, Contrast Enhancement, Gabor Filtering, Half wave rectification, Principal Component Analysis, Classification through C4.5 with Bagging and image post-processing through connected component analysis.. Results reported outperform the existing retinal segmentation algorithms. An average accuracy of 95.08% is accomplished on the
test set of the DRIVE dataset. The insight for future directions includes building new classification models, and enhanced post processing techniques for better segmentation. The proposed framework can assist the medical practitioners in better retinal analysis, earlier diagnosis and hence better treatment to patients.

REFERENCES


Geetha Ramani, R., B. Lakshmi and Shomona Gracia Jacob, 2012b. Data Mining Method of Evaluating Classifier Prediction Accuracy in Retinal Data, Proc. of IEEE Int. Conf. on Computational Intelligence and Computing Research, 426-429.


Martinez-Perez, M.E., A.D. Hughes, S.A. Thom, et al., 2007. Segmentation of blood vessels from red-free and fluorescein retinal images, Medical Image Analysis, 11: 47-61.


