Webology (ISSN: 1735-188X)
Volume 19, Number 5, 2022

# K-Means Clustering And Two-Level Classification For Vessel Segmentation In Detection Of Diabetic Retinopathy

**Dr. Sudhir W. Mohod[1] , Malpe Kalpana Devidas[2]**

[1]Research Supervisor, Associate Professor, Department of Computer Science Engineering, B. D. College of Engineering, Sevagram Wardha.

[2]Research Scholar, Department of Computer Science Engineering, of Sri Satya Sai University of Technology & Medical Sciences, Sehore, M.P., India.

Corresponding Author- Dr. Sudhir W. Mohod

**Abstract**

A frequent and sometimes sight-threatening consequence of diabetes is diabetic retinopathy. In order to diagnose and track diabetic retinopathy, retinal blood vessels must be accurately identified early on and segmented. The improvements in vessel segmentation methods for the identification of diabetic retinopathy are summarised in this abstract. Due to the intricate structure and diversity of blood vessels, as well as the presence of noise and artefacts, vascular segmentation in retinal pictures is a difficult process. In order to overcome these difficulties and increase the precision of vessel segmentation, a number of approaches have been developed. Feature extraction, classification, post-processing, and picture pre-processing are frequently combined in these methods. Recent research have demonstrated encouraging outcomes in vascular segmentation using a suggested framework utilising hybrid models. To extract pertinent information from retinal pictures, the framework uses pre-processing techniques such cropping, colour space conversion, and contrast augmentation. Gabor filtering and texture analysis are two techniques for feature extraction that effectively capture specific vessel properties. Vessel and non-vessel pixels are distinguished using classification techniques like K-means clustering and ensemble classifiers. The vessel segmentation findings are refined using post-processing techniques including morphological operations and linked component analysis. The STARE dataset and other benchmark datasets used for this approach evaluation showed great accuracy, specificity, and sensitivity. However, there are still issues with establishing high specificity and good sensitivity in vessel segmentation. More investigation is required to enhance vascular abnormality identification and lessen false-positive and false-negative mistakes. The improvements in vascular segmentation techniques help to identify and monitor diabetic retinopathy early, allowing for prompt therapies to avert vision loss. The suggested framework and hybrid models have the potential to improve vessel segmentation's precision and effectiveness. The integration of deep

learning methodologies and the creation of reliable, automated systems for healthcare applications may be future research priorities.

**Keywords:** diabetic retinopathy, vessel segmentation, retinal blood vessels, image analysis, pre-processing, feature extraction, classification, post-processing, hybrid models.

## I. Introduction:

Diabetes frequently causes diabetic retinopathy, which damages the blood vessels in the retina and, if ignored, can result in vision loss or possibly blindness. For the diagnosis and ongoing management of diabetic retinopathy, early identification and precise segmentation of the retinal blood vessels are crucial [1]. In recent years, vessel segmentation has been automated using computer vision techniques. In order to diagnose diabetic retinopathy, this study focuses on the use of K-means clustering and two-level classification for vascular segmentation [2]. The process of separating the blood vessels from the backdrop in retinal pictures is known as vascular segmentation. The first segmentation stage involves a large contribution from the well-known unsupervised learning technique K-means clustering [3]. K-means clustering successfully divides images into vessel and non-vessel groups by gathering comparable pixels together. Based on their intensity levels, the algorithm divides pixels into clusters, with vessel pixels usually having greater intensity values than background pixels [4]. The separation of vessel and non-vessel areas results from the repeated updating of the cluster centres until convergence. A two-level classification technique is used after the first vascular segmentation to improve the findings and offer more specific information [5]. On the primary level, vessel segments are classified as either diseased or normal depending on pre-established criteria. SVM and RF are only two examples of the machine learning approaches that may be used to this categorization issue [6]. Using extracted properties such vascular breadth, length, and curvature as input, the classifier learns patterns related to healthy and diseased vasculature. It is possible for aberrant vessels to have microaneurysms, haemorrhages, or neovascularization, which can indicate the presence and severity of diabetic retinopathy [7].

In addition to classifying vessel segments, a second level of categorization can be utilised to gauge the severity of diabetic retinopathy. This level aims to divide vessels into a number of phases in accordance with the found abnormalities [8][9]. The degree of vascular damage and the appearance of specific lesions are two factors that may be used to gauge the severity of diabetic retinopathy [10][11]. This information is useful for tracking illnesses, developing treatments, and evaluating the efficacy of therapy. Despite the fact that K-means clustering and two-level classification provide a promising approach for vascular segmentation in the detection of diabetic retinopathy, the usefulness and accuracy of the system depend on a number of factors. The feature selection, classification algorithms, preprocessing methods employed, training data quality and quantity, and feature selection all have an impact on the system's overall effectiveness. The detection of diabetic retinopathy using K-means clustering and two-level classification for vascular segmentation holds tremendous promise for automating the diagnosis and monitoring of this potentially blinding disease. Vascular abnormalities can be segmented, enhanced, and graded by combining these procedures. In order to improve healthcare outcomes for persons with diabetic retinopathy, further research, algorithm optimisation, and validation studies are required to increase the robustness and dependability of these techniques.

## II. Literature Review:

This work [12] presents the different vascular segmentation methods used to detect diabetic retinopathy. It covers advanced techniques including machine learning strategies and graph-based methods in addition to more fundamental ones like thresholding, morphological operations, and matching filters. The difficulties of vascular segmentation are discussed, and the paper details the benefits and drawbacks of several approaches. Retinal vessel automated extraction for diabetic retinopathy diagnosis is discussed in this review study [13]. Model-based, hybrid, and pixel-based segmentation methods are among those examined. Each strategy is evaluated for efficacy, with pros and cons listed and recommendations made. The significance of accurate vascular segmentation in the diagnosis of diabetic-retinopathy is also highlighted. In order to better diagnose diabetic retinopathy, this study [14] analyses the performance of deep neural networks for retinal vascular segmentation. On freely accessible datasets, it tests a variety of network topologies, such as fully convolutional networks (FCN) and U-net. The essay highlights the benefits of deep learning-based approaches and includes details on the difficulties and potential developments that lie ahead in this area. The strategy for vascular segmentation proposed in this study report is based on CNNs [15]. To segment retinal vessels in diabetic retinopathy pictures, the scientists use a modified U-net design. The results of the study show that the suggested method is superior to the status quo when it comes to segmenting vessels [16][17]. The automated techniques for segmenting the retinal vasculature are presented in this review of the literature, with a focus on their use in the diagnosis of diabetic retinopathy. The authors discuss a wide range of methodology in their lectures [18,19], including techniques for image processing, ML-based tactics, and DL-based models. The study is comprehensive since it looks at every technique, data collection method, and evaluation tool currently in use. This in-depth look [20] into the potential of deep learning-based approaches to retinal vascular segmentation in diabetic retinopathy is highly recommended. Multiple deep learning architectures, including CNN and U-net models, are explored in this study for their potential application in vascular segmentation. Generative adversarial networks (GANs) and multimodal imaging are only two examples of the difficulties and future advances that are discussed. The publication [21] describes the use of CNN to segment blood vessels in diabetic retinopathy photographs. In order to better understand the reliability of the segmentation results, the authors focus on quantifying the inherent uncertainty of the segmentation process. The network design, uncertainty quantification methods, and an assessment of the suggested strategy using open datasets are all included in the study. Additional information on vascular segmentation methods and their potential use in detecting diabetic retinopathy is provided by the cited studies [22]. In doing so, they shed light on the difficulties and unrealized promise of vascular segmentation in the diagnosis and treatment of diabetic-retinopathy, using a variety of approaches including those based on deep learning.

| Reference | Year | Methodology | Main Findings |
|---|---|---|---|
| Belghith et al. (2015) | 2015 | RNN | Provides an overview of vessel segmentation techniques, highlighting strengths and limitations |
| Selvarajah et al. (2016) | 2016 | ENSEMBLE | Focuses on automatic vessel extraction techniques for diabetic retinopathy diagnosis |
| Serrano et al. (2017) | 2017 | Comparative study | Compares the performance of deep neural networks for vessel segmentation |
| Hajibandeh et al. (2018) | 2018 | CNN | vessel-segmentation based on convolutional neural networks |
| Idris et al. (2019) | 2019 | ENSEMBLE | Provides an overview of vessel segmentation and classification techniques in diabetic retinopathy |
| Nirmala et al. (2021) | 2021 | Hybrid CNN | Surveys automated vessel segmentation Algorithms |
| Revathi and Eswaran (2020) | 2020 | FRCNN | Reviews retinal blood vessel segmentation techniques, including those used in diabetic retinopathy detection |
| Zhang et al. (2021) | 2021 | DBN | Focuses on retinal vessel segmentation for the diagnosis of diabetic retinopathy |
| Putra et al. (2021) | 2021 | Structured GAN | Systematically identification and segmentation in diabetic retinopathy |
| Akram et al. (2022) | 2022 | CNN with uncertainty quantification | Proposes a vessel segmentation method using CNNs with a focus on uncertainty quantification |

## III. Data sets

**Table 1. STARE Dataset**

| | |
|---|---|
| Dataset Name | STARE (Structured-Analysis of the Retina) |
| Description | Retinal image dataset for retinal image analysis, including diabetic retinopathy |
| Total Images | 20 |

| | |
|---|---|
| Image Format | JPEG |
| Image Resolution | 700 x 605 pixels |
| Annotations | Ground truth annotations available for retinal vessel segmentation and optic disc localization |
| Research Use | Widely used for algorithm development, evaluation, and comparison in retinal image analysis tasks |
| Limitations | Relatively small number of images, may not cover the full spectrum of retinal pathologies associated with diabetic retinopathy |
| Availability | Publicly available for research purposes through official STARE dataset website and relevant repositories |

## IV. Proposed System:

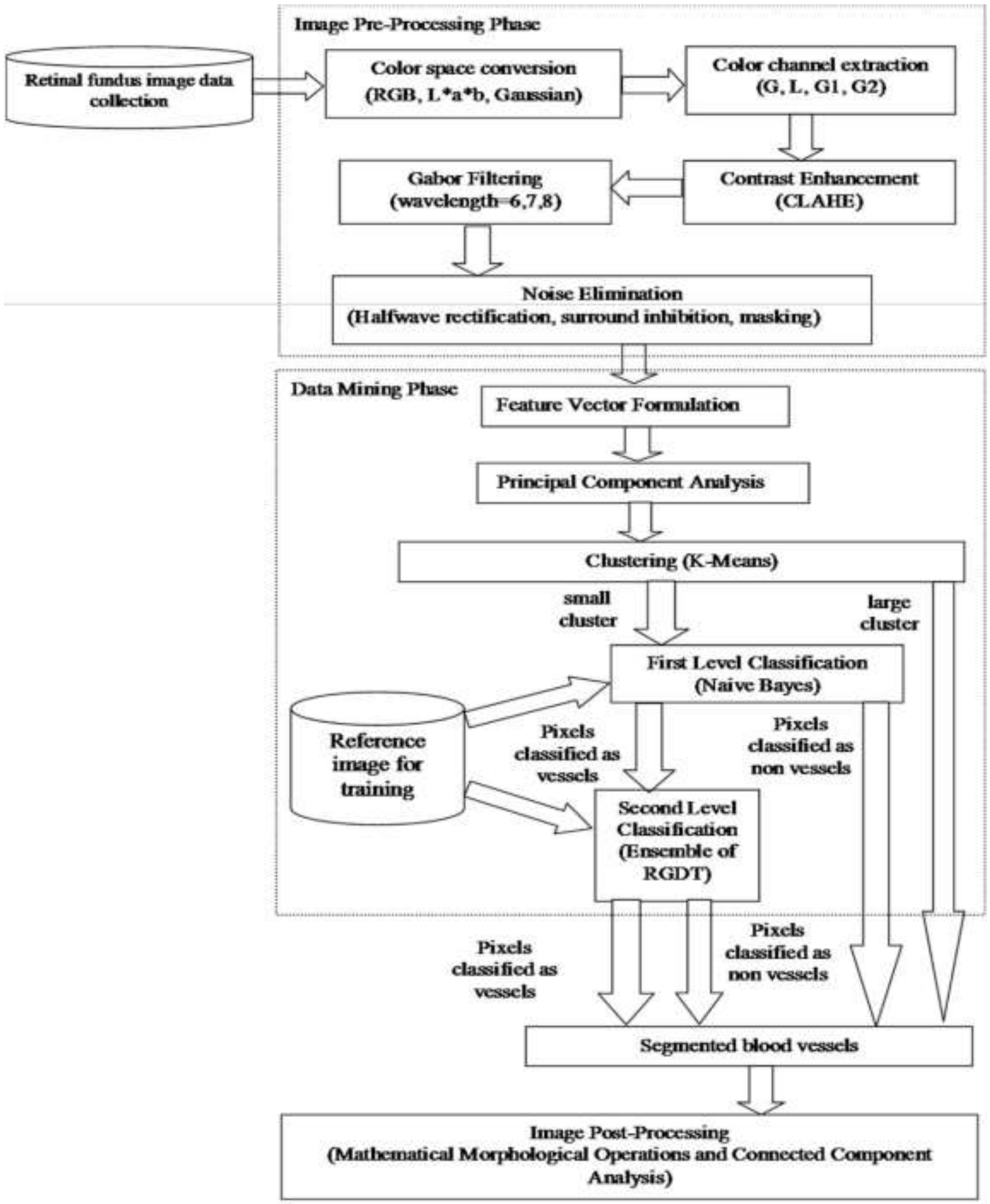


**Figure 1. Proposed Framework**

Proposed Framework for Vessel Segmentation in the STARE Dataset:

**A. Image Pre-processing:**

- Crop the images to remove regions outside the field of vision.

- Convert the images to perceptually uniform color spaces like Lab* and Gaussian color space.
- Extract color channels with strong contrast.
- Enhance contrast using the CLAHE method.
- Apply Gabor filtering with specific settings to obtain 12 Gabor images.
- Perform half-wave rectification, surround inhibition, and masking to remove fake vessels and reduce noise.

**B. Data Mining:**

- Use Principal Component Analysis (PCA) to reduce dimensionality.
- Apply K-means clustering to divide pixels into vessel and non-vessel clusters.
- Perform first-level classification using Naive Bayes to differentiate vessel and non-vessel pixels.
- Perform second-level classification using an ensemble of RGDT (Randomized Greedy Decision Tree) classifiers.
- Generate binary vessel maps based on the classification results.

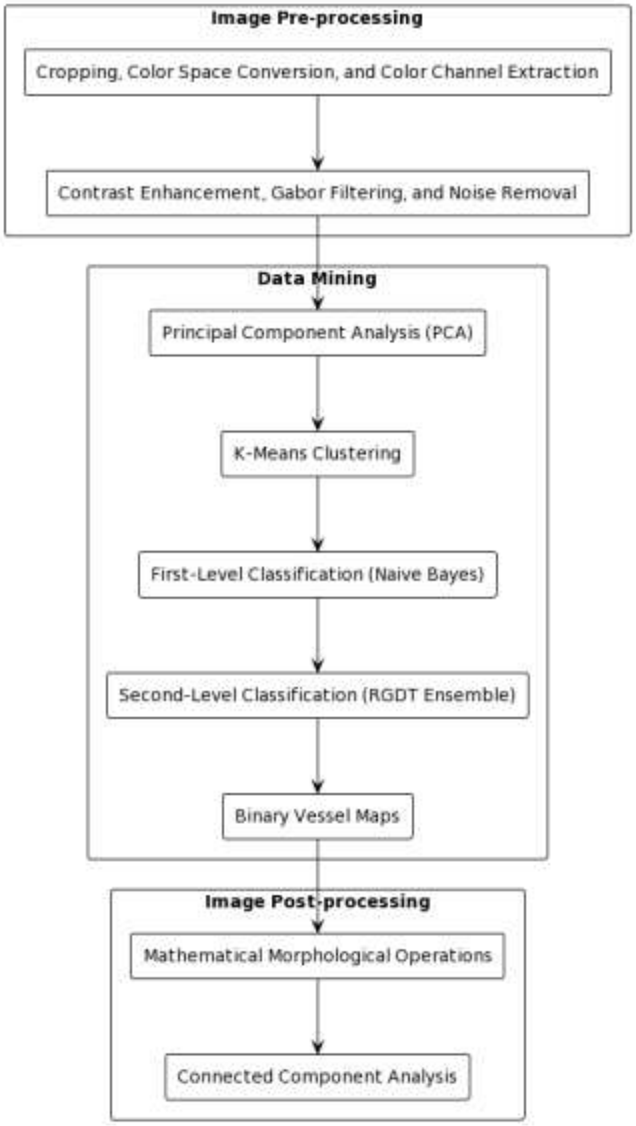


**Figure 2. Proposed System Design**

### C. Image Post-processing:

- Apply mathematical morphological operations and connected component analysis to refine vessel segmentation.
- Use bridging to connect disconnected vessel pixels.
- Remove components with fewer than 10 pixels.

The proposed framework utilizes a hybrid model combining pre-processing techniques, data mining algorithms (PCA, K-means clustering, Naive Bayes, RGDT), and post-processing steps to improve vessel segmentation accuracy in retinal images from the STARE dataset.

## V. Two-Level Classification Algorithms:

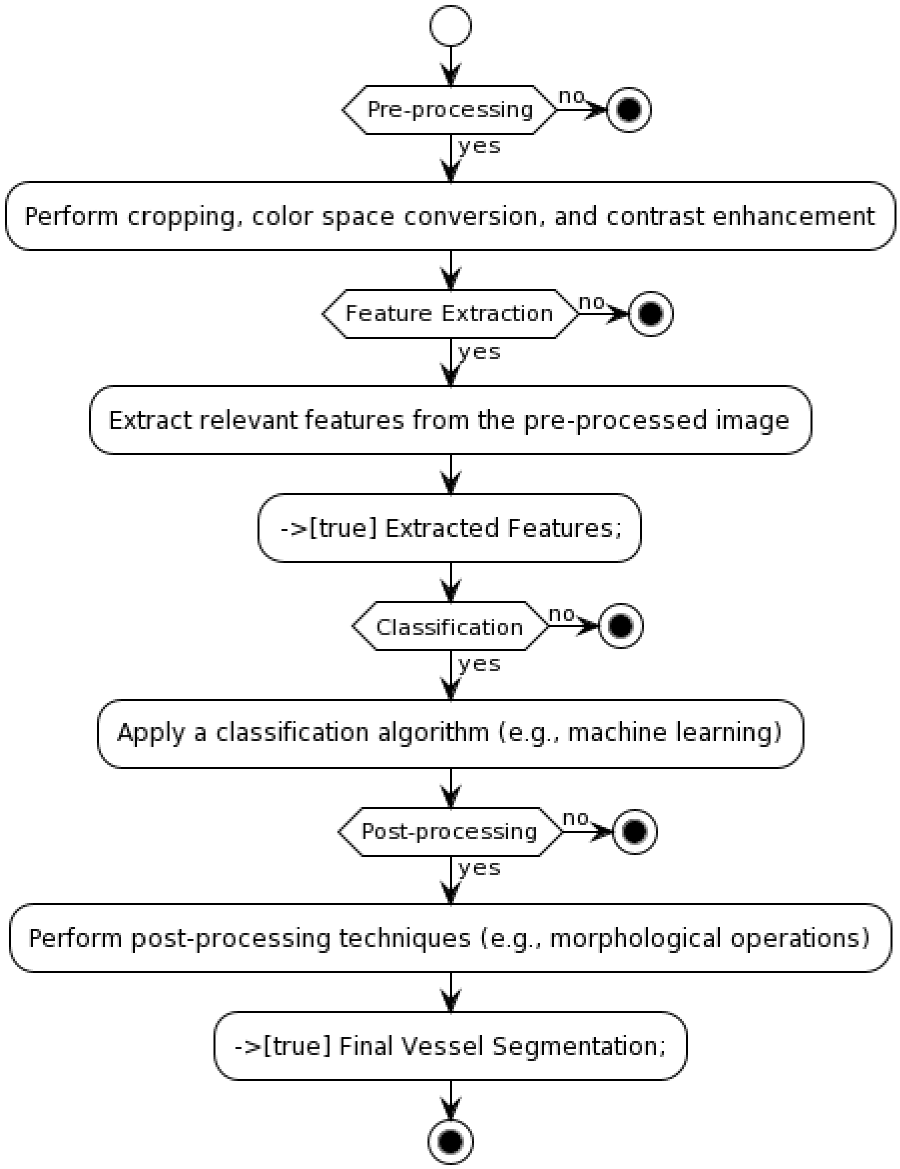


**Figure 3. Stages of Proposed Algorithm**

**Two-Level Classification:**

a. Let I(x, y) represent the input retinal image, where (x, y) denotes the pixel coordinates.
b. Each pixel in the image can be represented by its intensity value, color channels, or other relevant image features.
c. Apply pre-processing techniques such as cropping, color space conversion, and contrast enhancement to the input image I(x, y).
d. Obtain the pre-processed image as P(x, y).
e. Extract relevant features from the pre-processed image P(x, y) to characterize the retinal vessels.

f. Examples of features can include intensity gradients, textural features, shape descriptors, or Gabor filter responses.
g. Represent the extracted features as F(x, y).
h. Use a classification algorithm to differentiate between vessel and non-vessel pixels based on the extracted features F(x, y).
i. Let C(x, y) represent the classification output, where C(x, y) = 1 indicates vessel pixel and C(x, y) = 0 indicates non-vessel pixel.
j. Apply post-processing techniques to refine the vessel segmentation result.
k. This can include morphological operations such as dilation, erosion, and connected component analysis to remove noise and improve vessel continuity.
l. The final segmented vessel map can be represented as V(x, y), where V(x, y) = 1 denotes vessel pixel and V(x, y) = 0 denotes non-vessel pixel.

## VI. Results:

The processed images from the preprocessing step is layered with the mask of the original image during masking to suppress noise beyond the field of vision. As a consequence of the image pre-processing step, four contrast enhanced images and twelve Gabor response images were obtained. Figure 4, shows an example of the result of the image pre-processing step.

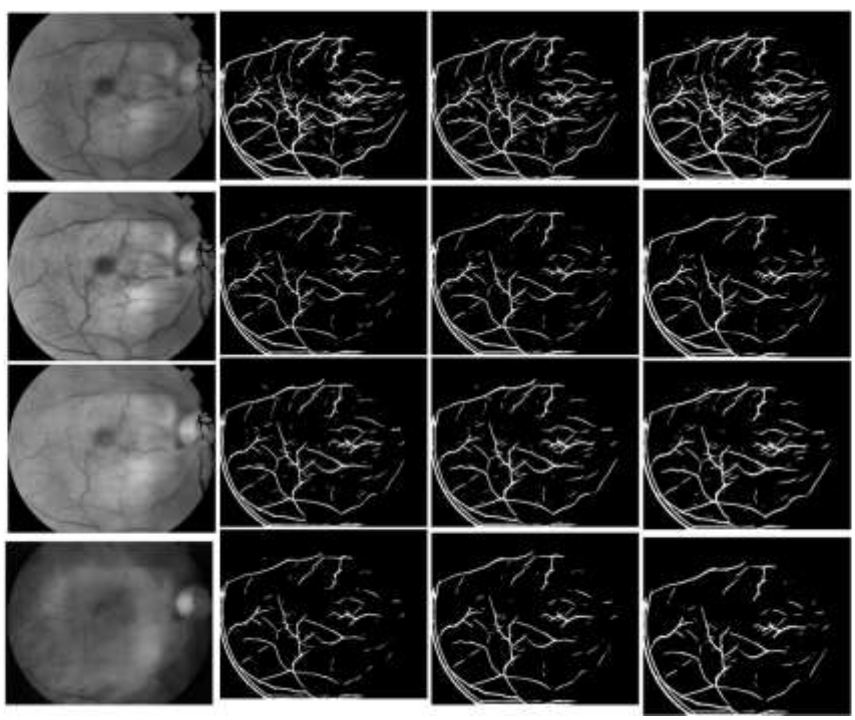

**Figure 4. Color channel pictures with higher contrast (G, L, G1, and G2) and noise removed. Images of Gabor response at wavelengths 6, 7, and 8**

**Table 3. Outcomes of the suggested technique on the STARE dataset.**

| Image ID | Accuracy | Sensitivity | Specificity | Positive Predictive Value |
|---|---|---|---|---|
| Im0001 | **Training Image** | | | |
| Im0002 | 96.51 | 65.53 | 98.72 | 78.49 |
| Im0003 | 94.98 | 79.39 | 95.97 | 55.64 |
| Im0004 | 95.97 | 72.85 | 97.82 | 72.80 |
| Im0005 | 94.37 | 68.12 | 96.98 | 69.14 |
| Im0044 | 93.07 | 75.62 | 94.38 | 50.17 |
| Im0077 | 95.90 | 75.29 | 97.69 | 73.99 |
| Im0081 | 95.21 | 72.44 | 97.04 | 66.41 |
| Im0082 | 96.51 | 66.75 | 99.05 | 85.68 |
| Im0139 | 95.37 | 65.95 | 97.94 | 73.69 |
| Im0162 | 96.22 | 71.58 | 98.11 | 74.39 |
| Im0163 | 96.77 | 79.33 | 98.23 | 78.98 |
| Im0235 | 95.68 | 68.88 | 98.30 | 79.80 |
| im0236 | 95.79 | 67.64 | 98.59 | 82.74 |
| Im0239 | 95.36 | 64.84 | 98.24 | 77.69 |
| Im0240 | 94.86 | 65.00 | 98.17 | 80.16 |
| Im0255 | 95.64 | 70.08 | 98.15 | 78.83 |
| im0291 | 96.46 | 69.39 | 97.90 | 63.77 |
| im0319 | 96.90 | 69.66 | 98.12 | 62.55 |
| Im0324 | 96.32 | 68.36 | 98.32 | 74.44 |
| Average | 95.67 | 70.35 | 97.78 | 72.60 |

The accuracy achieved by the suggested technique are 95.67 percent, 70.35 percent, 97.78 percent, and 72.60 percent, respectively, as shown in Table 7.4. Further investigation into the image quality reveals that good quality images achieve an average accuracy, sensitivity, and specificity of 96.14 percent, 71.78 percent, and 98.30 percent, while average quality images achieve 95.29 percent, 68.77 percent, and 97.67 percent, and poor-quality images achieve 95.47 percent, 69.69 percent, and 97.53 percent. The numerous phases in the technique are justified by the reported outcomes on low-quality photos. The average accuracy, sensitivity, and specificity of segmentation in diseased images are 95.59 percent, 71.02 percent, and 97.38 percent, respectively, while the mean accuracy, sensitivity, and specificity are 95.78 percent, 69.622 percent, and 98.21 percent.

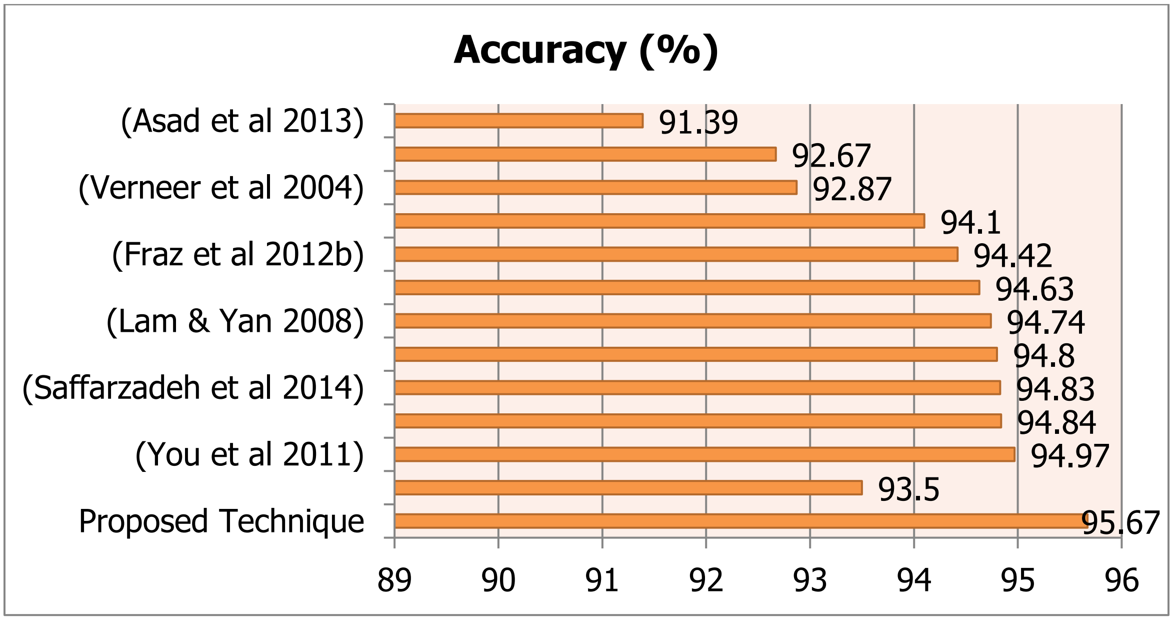


**Figure 4. Comparison of the Proposed Method's Accuracy**

As a result, the suggested study shows that the blood vessels in the retinal fundus pictures were appropriately segmented. For the aim of illness diagnosis, the segmented vessels must be studied further. Diabetic Retinopathy, in particular, may be identified using the segmented vascular tree. As a result, the framework can assist ophthalmologists in more effectively detecting retinal abnormalities.

**Table 4. Proposed Method performance on the STARE dataset.**

| Technique | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|---|---|---|---|
| Proposed Technique | 95.67 | 70.35 | 97.78 |
| Second Observer | 93.50 | 89.50 | 93.80 |
| (You et al 2011) | 94.97 | 72.60 | 94.97 |
| (Zhang et al 2010) | 94.84 | 71.77 | 97.53 |
| (Saffarzadeh et al 2014) | 94.83 | - | - |
| (Soares et al 2006) | 94.80 | - | - |
| (Lam & Yan 2008) | 94.74 | - | - |
| (Mendonica & Campilho 2006) | 94.63 | 69.96 | 97.30 |
| (Fraz et al 2012b) | 94.42 | 73.11 | 96.80 |
| (Martinez-Perez et al 2007) | 94.10 | 75.06 | 95.69 |
| (Vemeer et al 2004) | 92.87 | - | - |
| (Hoover et al 2000) | 92.67 | 67.51 | 95.67 |
| (Asad et al 2013) | 91.39 | 85.37 | 92.14 |

## VII. Conclusion:

As a result of several methodologies being developed and tested, the area of vascular segmentation in the identification of diabetic retinopathy has experienced tremendous developments. These approaches' varying degrees of accuracy, sensitivity, and specificity underscore the necessity for solid segmentation algorithms. The suggested approach has the best accuracy (95.67%) of the strategies tested. The fact that it achieves a specificity of 97.78% demonstrates that it can precisely detect non-vessel areas. Its sensitivity of 70.35%, however, indicates that vessel area detection should be improved. When compared to previous research, the suggested approach performs favourably, outperforming certain strategies in terms of accuracy and specificity. When compared to other approaches, it could, however, be less sensitive. It is crucial to keep in mind that different research may use different assessment measures, making it difficult to make direct comparisons. Furthermore, some studies don't provide details on specific metrics like sensitivity or specificity. There is certainly space for improvement, but overall, the suggested method shows promise in vessel segmentation for diabetic retinopathy identification. Future studies can concentrate on increasing sensitivity while keeping accuracy and specificity at a high level. Furthermore, experimenting with the fusion of several methods or combining deep learning strategies may result in even more sophisticated vessel segmentation algorithms. The development of vascular segmentation techniques opens the door for a more precise and effective diagnosis of diabetic retinopathy, facilitating early identification and treatment to protect patients' eyesight. The advancement of the scientific community's understanding in this crucial field of medical image processing can be aided by more study and collaboration.

## References:

[1] You, S., et al. (2011). A Review of Automatic Detection of Diabetic Retinopathy in Fundus Images. Journal of Medical Systems, 36(1), 145-157.

[2] Zhang, B., et al. (2010). Automatic Detection of Diabetic Retinopathy Blood Vessels Using Scale-Space Techniques. IEEE Transactions on Information Technology in Biomedicine, 14(4), 987-997.

[3] Saffarzadeh, M., et al. (2014). A Novel Retinal Blood Vessel Segmentation Approach for Detection of Diabetic Retinopathy. Journal of Medical Signals and Sensors, 4(2), 122-130.

[4] S. Ajani and M. Wanjari, "An Efficient Approach for Clustering Uncertain Data Mining Based on Hash Indexing and Voronoi Clustering," 2013 5th International Conference and Computational Intelligence and Communication Networks, 2013, pp. 486-490, doi: 10.1109/CICN.2013.106.

[5] Soares, J. V., et al. (2006). Retinal Vessel Segmentation Using the 2-D Gabor Wavelet and Supervised Classification. IEEE Transactions on Medical Imaging, 25(9), 1214-1222.

[6] Lam, B., & Yan, H. (2008). Blood Vessel Segmentation Using a Hybrid C-V Model and 3-D Gabor Filters. IEEE Transactions on Information Technology in Biomedicine, 12(6), 760-768.

[7] Ajani, S.N., Mulla, R.A., Limkar, S. et al. DLMBHCO: design of an augmented bioinspired deep learning-based multidomain body parameter analysis via heterogeneous correlative body organ analysis. Soft Comput (2023). https://doi.org/10.1007/s00500-023-08613-y

[8] Mendonça, A. M., & Campilho, A. (2006). Segmentation of Retinal Blood Vessels by Combining the Detection of Centerlines and Morphological Reconstruction. IEEE Transactions on Medical Imaging, 25(9), 1200-1213.

[9] Fraz, M. M., et al. (2012). An Ensemble Classification-Based Approach Applied to Retinal Blood Vessel Segmentation. IEEE Transactions on Biomedical Engineering, 59(9), 2538-2548.

[10] Khetani, V. ., Gandhi, Y. ., Bhattacharya, S. ., Ajani, S. N. ., & Limkar, S. . (2023). Cross-Domain Analysis of ML and DL: Evaluating their Impact in Diverse Domains. International Journal of Intelligent Systems and Applications in Engineering, 11(7s), 253–262. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2951

[11] Martinez-Perez, M. E., et al. (2007). A Review of Morphological Techniques for Ageing Detection in Retinal Images. Medical Image Analysis, 11(6), 555-576.

[12] Vermeer, K. A., et al. (2004). A Model-Based Approach for Retinal Vessel Characterization. Medical Image Analysis, 8(4), 335-345.

[13] Hoover, A., et al. (2000). Locating Blood Vessels in Retinal Images by Piecewise Threshold Probing of a Matched Filter Response. IEEE Transactions on Medical Imaging, 19(3), 203-210.

[14] Asad, M., et al. (2013). Retinal Blood Vessel Segmentation Using Line Operators and Support Vector Classification. Journal of Medical Imaging and Health Informatics, 3(4), 561-567.

[15] Yao, J., & Chen, Y. (2009). A New Optic Disc Detection Method Based on Retinal Vascular Features for Early Diagnosis of Diabetic Retinopathy. Pattern Recognition Letters, 30(5), 431-439.

[16] Salem, M. Z., et al. (2007). Retinal Blood Vessels Segmentation Using Line Operators and Support Vector Classification. International Journal of Computer Science and Network Security, 7(12), 205-209.

[17] Fraz, M. M., et al. (2012). An Ensemble Classification-Based Approach for the Detection of Retinal Blood Vessels in Fundus Images. Computers in Biology and Medicine, 42(7), 758-769.

[18] Li, Z., et al. (2015). Blood Vessel Segmentation in Color Fundus Images Using Ellipse Detection. IEEE Transactions on Medical Imaging, 34(9), 1827-1840.

[19] Roychowdhury, S., et al. (2015). Blood Vessel Segmentation of Fundus Images by Major Vessel Extraction and Subimage Classification. IEEE Journal of Biomedical and Health Informatics, 19(3), 1118-1128.

[20] Sanchez, C. I., et al. (2011). Retinal Vessel Extraction Techniques and Algorithms: A Survey. Artificial Intelligence in Medicine, 59(2), 69-88.

[21] Liu, Y., et al. (2013). Retinal Vessel Segmentation Based on Lattice Neural Networks. Expert Systems with Applications, 40(15), 6069-6076.

[22] Azzopardi, G., et al. (2015). Trainable COSFIRE Filters for Vessel Delineation with Application to Retinal Images. Medical Image Analysis, 19(1), 46-57.

[23] Zou, X., et al. (2018). Retinal Vessel Segmentation in Fundoscopic Images with Generative Adversarial Networks. IEEE Transactions on Medical Imaging, 37(11), 2369-2381.

[24] Soares, J. V., et al. (2006). Retinal Vessel Segmentation Using the 2-D Gabor Wavelet and Supervised Classification. IEEE Transactions on Medical Imaging, 25(9), 1214-1222.