

Advancing retinal image analysis: from preprocessing to lesion identification in diabetic retinopathy

Sowmyashree B.^{1*}, Mahesh K. Rao¹ and Chethan H.K.²

Department of Electronics and Communications Engineering, Maharaja Institute of Technology Mysore-571477 Karnataka, India¹

Department of Computer Science, Maharaja Institute of Technology Mysore-571477 Karnataka, India²

Received: 24-September-2023; Revised: 24-October-2024; Accepted: 28-October-2024

©2024 Sowmyashree B. et al. This is an open access article distributed under the Creative Commons Attribution (CC BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Diabetic retinopathy (DR) is a leading cause of vision loss, making early detection and classification vital for effective management. Despite advancements in machine learning and deep learning techniques, these methods often require substantial computational resources and specialized domain knowledge, which are not always readily available. To address these challenges, a robust and computationally efficient methodology is proposed for early-stage DR detection, leveraging classical image enhancement and multi-stage segmentation techniques. Utilizing publicly available datasets—DRIVE, CHASEDB1, and DIARETDB1, our study applied data augmentation techniques to bolster the training set and mitigate overfitting. The framework employs a range of image enhancement methods, including Median Filter, Weiner Filter, and contrast limited adaptive histogram equalization (CLAHE), followed by lesion segmentation techniques. Accurate isolation of critical features like blood vessels (BVs) and the optic disc is achieved through principal component analysis (PCA) and iterative self-organizing data analysis technique algorithm (ISODATA). The experimental analysis demonstrates that the proposed methodology achieves high performance, with sensitivity, specificity, and accuracy rates of 91.50%, 88.20%, and 90.40% for microaneurysms (MA); 92.20%, 89.10%, and 91.30% for hemorrhages (HE); and 93.10%, 90.20%, and 92.60% for exudates, respectively. The image enhancement techniques improved the peak signal-to-noise ratio (PSNR) to 21.80 and the normalized cross-correlation (NCC) to 0.812. These results indicate the effectiveness of the proposed methods in accurately detecting and classifying lesions in retinal images.

Keywords

Diabetic retinopathy, Lesions detection, Retinal image segmentation, Contrast limited adaptive histogram equalization (CLAHE), Data augmentation.

1.Introduction

Deep learning and neural networks have advanced artificial intelligence (AI) by simulating interconnected neurons to process inputs and generate outputs, significantly enhancing AI's capabilities [1]. With AI achieving diagnostic accuracy comparable to healthcare experts, automated disease screening offers increased efficiency and risk stratification [2]. Diabetes mellitus, affecting 422 million people globally [3], has seen its incidence double over the past 30 years, particularly in Asia [4]. Diabetic retinopathy (DR), a complication of diabetes, is a leading cause of blindness in working-age adults [5], affecting one-third of diabetic patients.

DR diagnosis is labour-intensive, especially in rural areas lacking specialized medical professionals [6]. As diabetes incidence rises, manual screening methods may struggle to meet the growing demand [7]. Globally amongst the working-age adults, DR is the leading preventable cause of blindness, accounting for more than 24,000 [8] incidents of blindness annually and being the primary emphasis in ophthalmological AI screening algorithms. Patients with persistent diabetes mellitus, especially those with poor clinical management, are at higher risk of developing blindness. Automated methods for diagnosing DR are crucial to finding answers to these questions. DR manifests in two forms: non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR) [9] as illustrated in *Figure 1*. NPDR, the early stage of the disease, is characterized by the swelling or bumping

*Author for correspondence

of blood vessels (BVs), known as aneurysms. Lesions, which are abnormal changes or damages in the retina's structure due to injury or disease, play a crucial role in the early detection of DR. These lesions are categorized as dark or bright based on their intensity value and include micro aneurysms, hemorrhages, exudates, and cotton wool spot [10].

PDR, the more severe stage of DR, occurs when irregular new BV and damaged tissue appear on the retina's surface. Micro aneurysms, tiny dark spots in the retinal fundus image, are often the initial symptoms of DR in diabetic patients. As the disease progresses, hemorrhages occur, which can cause severe vision impairment [11].

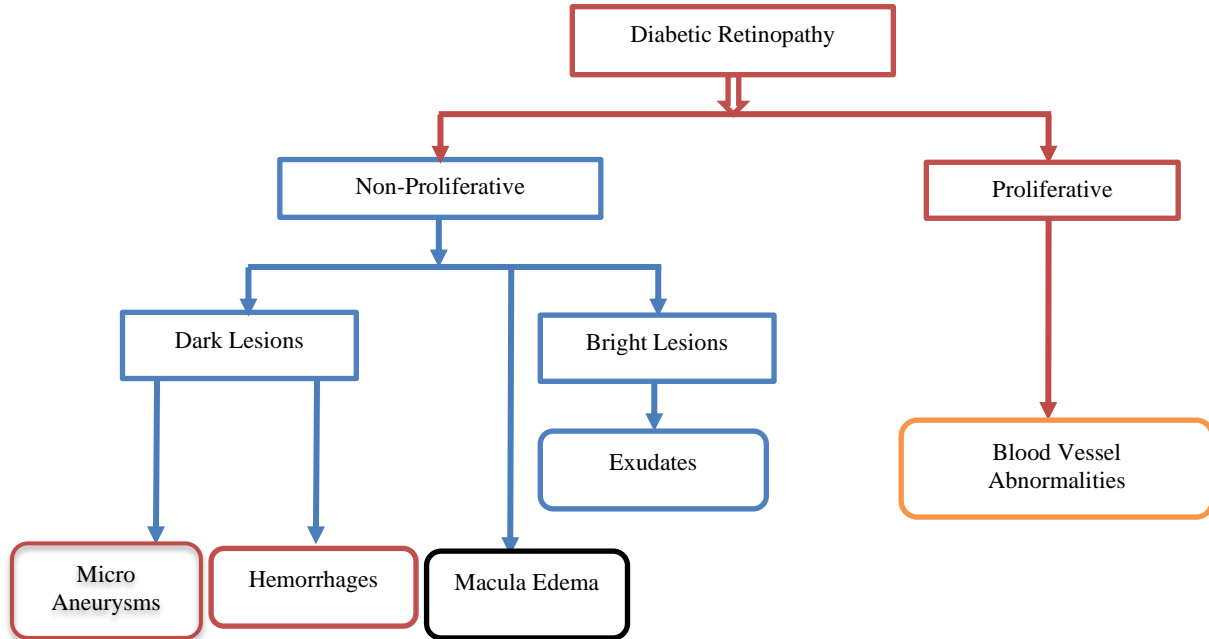


Figure 1 DR Stages based on lesions

Manual DR screening is prone to variability in interpretation and may not meet the increasing demand for screening services as diabetes prevalence escalates [12, 13]. The scarcity of skilled readers in remote regions further complicates early diagnosis [14], which is crucial for preventing vision loss. Additionally, data labelling accuracy remains a challenge, necessitating effective supervised learning models for DR detection [15]. Automated AI-driven methods are essential to address these challenges, improving screening accuracy and efficiency [16]. The experimental studies relied on two main datasets as shown in *Figure 2*: the CHASEDB1 dataset for

retinal vascular segmentation, which contains 28 colour retina images, and the DRIVE dataset, a collection of 400 diabetic people aged 25 to 90 years and is publicly available. Throughout this work, a more efficient method for segmenting retinal lesions is clarified, with the aim of potentially identifying DR in its early stages to improve clinical outcomes. The representative images are depicted in *Figure 2*. From left (L) to right (R): A is the fundus image of the left eye and B is the corresponding right eye of CHASEDB1 dataset. Similarly, C is the left fundus image and D, the right eye image of the DRIVE dataset.

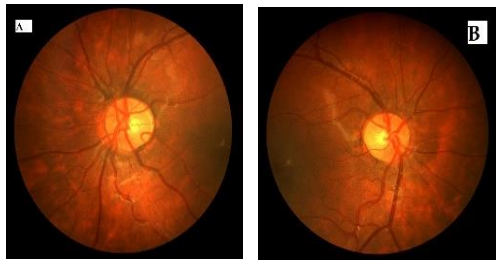


Figure 2(a) Sample images of the CHASEDB1 and 1450

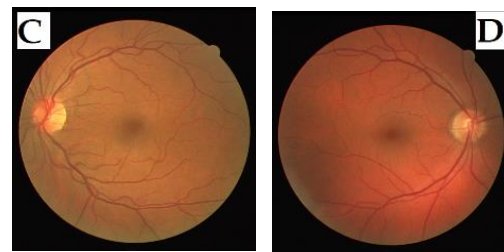


Figure 2(b) DRIVE datasets

Ophthalmologists diagnose DR by looking for and classifying anomalies in retinal images. The four main types of lesions are microaneurysms (MA), hemorrhages (HEM), soft exudates (SE), and hard exudates (HE). Small, red spots appear on the retina due to macular atrophy (MAT), a hallmark of early-onset DR caused by vessel wall weakening. They have a diameter of 125 micrometres or less with sharp edges [17].

This paper aims to develop a more efficient method for segmenting retinal lesions to facilitate early DR detection. The study seeks to optimize clinical outcomes through automated image processing techniques and AI algorithms tailored for DR detection.

This work introduces a robust preprocessing pipeline incorporating image enhancement techniques like luminosity adjustment, Weiner filtering, and contrast-limited adaptive histogram equalization (CLAHE). It proposes a novel luminance gain matrix for enhancing retinal image visibility, and presents a BV segmentation method using PCA and iterative self-organizing data analysis technique algorithm (ISODATA) thresholding. A detailed data augmentation strategy is employed to improve model generalization, and lesion detection is refined with morphological methods and the extended-minima transform. These computationally efficient methods are designed for real-time applications and large datasets.

The paper is structured as follows: Section 2 reviews deep learning approaches for DR detection. Section 3 discusses materials and methods. Experimental findings are presented in Section 4, followed by a discussion. Section 5 concludes with future directions.

2.Literature review

In medical image processing, segmentation plays a vital role in the early detection of chronic eye diseases. The recent improvement is essentially the result of the timely convergence of two factors:

- The amount of training data is now significantly higher than it was five decades ago; and
- Its coincidence with the advent of super-powerful graphical processing units (GPU) [18].

Most research on DR detection are usually studied under two main categories, and highlighted in the following section.

2.1Statistical and machine learning-based approaches

Numerous studies have focused on the early detection of DR and the classification of its various stages using handcrafted feature extraction. A crucial aspect of early DR diagnosis is the detection of MA. In a study by [19], the k-nearest neighbor classifier (KNN) was used for discrimination of outliers. Another research introduced a multi-scale approach [20] for microaneurysm detection in fundus images, achieving promising results on the state of the arts (SoTA) datasets. For the exudate's detection, Bhimavarapu et al. [21] used a matching filter to exclude the vessel and optic disc before detecting exudate, and an ensemble of classifiers was employed to locate the exudate area using the saliency map. Another study used mathematical morphology and support vector machines (SVM) classifiers for exudate identification [22], achieving an area under the curve (AUC) of 99.7% using the MESSIDOR dataset [23]. Several machine learning-based algorithms have shown significant results for lesion detection [24].

Nahiduzzaman et al. [25] proposed a method for enhancing dark lesions using curvelet-based edge detection and morphological closing with an optimum Bandpass filter for bright lesion enhancement. Phridviraj et al. [26] used the CLAHE method for retinal image enhancement for super-pixel classification of DR symptoms. They also detected exudates in color fundus images using contrast enhancement and global and local luminosity enhancement. Hu et al. [27] proposed a technique for hemorrhage detection using histogram equalization (HE) with a matched filter for retinal image enhancement. Mousavi et al. [28] used a multiple-scale hidden Markov enhancement approach for exudate detection. It was Akram et al. [29], who presented a method for image enhancement, background partitioning using an averaging and variance-based method, and noise elimination by applying an adaptive contrast enhancement technique in the hue, saturation, and value (HSV) channel.

2.2Deep Learning-based approaches

Many objects identification and biological image segmentation challenges [30, 31] have previously been won by efficient convolutional neural networks (CNNs), in particular the works of Krizhevsky et al. [32], and Dieleman et al. [33]. Recent idea involving the use of semi-supervised approach [34], was adapted by Melinscak et al. [35] to handle vascular segmentation. In the study by Liskowski and Krawiec

[36], a structured prediction technique was employed to emphasize context information while testing a wide range of architectures, with a 7-layer no-pooling CNN emerging as the most effective.

U-Net is a common and well-known backbone network in medical image segmentation [37–40]. A U-Net primarily comprises a “skip connection” between a standard down-sampling encoder and up-sampling decoder structure. It mixes local and global context information through the encoding and decoding procedures. Ikechukwu and Murali [41] proposed a neural network that used less computational resources efficiently. To accomplish retinal vascular segmentation, Wu et al. presented vessel-Net [42], which for the first time employs a method that combines the benefits of the initial method with the residual method. To help people better retain structural knowledge, Zhang et al. presented the attention network (AG-Net) [43], which created the “Attention Guide Filter” as an attention mechanism. These U-Net variations work well but always make the network more complicated and harder to understand. By applying context-aware patch-based loss, Xu and Fan [44], relied on the morphological properties of the retinal BV. They proposed MainSegment-Net fuses the convolutional layer with skip connection to achieve an accurate thin vessel region. Kumar and Singh [45] presented a comprehensive review of various techniques for BV segmentation. Emphasis was placed on research involving deep learning techniques benchmarked on freely available datasets such as DRIVE, STARE, CHASE and MESSIDOR. It further explains the advantages and limitations of existing SoTA model while highlighting the future direction. When tested on the DRIVE and STARE datasets, the technique shows an average sensitivity of 82.61% and 81.67%, respectively. It effectively increases the sensitivity of retinal vascular segmentation and outperforms existing state-of-the-art methods in this regard. Remeseiro et al. [46] used only the local contrast between BV and their background for the automatic arteries BV classification. The method computes a graph representing the vascular structure and then applies a multilevel thresholding to obtain a preliminary classification. A novel graph propagation approach was created. Evaluation on two publicly available datasets (INSPIRE and DRIVE), gave promising results: especially in the main vessels, that are strikingly close to those supplied by human experts.

Traditional methods like k-means, fuzzy c-means clustering, and SVM are more accurate at classifying the outcomes. Deep learning outperforms standard machine learning due to more hidden layers and hyperparameters. However, training deep-learning models generally requires lots of computational resources and adequate domain knowledge about automatic feature extraction. To address this issue, the research seeks to propose a robust approach that will be computationally effective using classical image enhancement and multi-stage segmentation of retinal lesions, especially for the early-stage DR detection.

3. Materials and method

3.1 Dataset description

The DRIVE, CHASEDB1, and DIARETDB1 were the publicly available datasets used in this study comprising of images with various abnormalities such as MA, HEM, and EX. The datasets are relatively balanced, with diverse patient demographics and varying disease severity levels.

The DRIVE dataset [47] was created to facilitate quantitative research on segmentation of BV in retinal pictures. The first phase of the collection contains fundus images of 400 diabetic patients aged 25 and 90 in the joint photographic experts group (JPEG) format. It contains a description of seven abnormalities that was categorized by experienced ophthalmologist and subdivided into training and validation sets. The dimensions of the images are 768 by 584 pixels with their corresponding masked images. On the other hand, the CHASEDB1 dataset was created for retinal vascular segmentation. It contains 28 colour retina images measuring 999 by 960 pixels, which were taken from the left and right eyes of 14 school children. Two separate human annotators create the annotations for each image. The DIARETDB1 dataset was used to validate our findings. This dataset consists of 89 retinal fundus images of dimension 1500 by 1152.

The proportions of images with various abnormalities, patient demographics, disease severity, and image quality are as follows:

Proportion of images with various abnormalities

They include images that exhibit various retinal abnormalities such as MA, and EX. Here is the distribution:

MA: Often the most prevalent abnormality in DR datasets, MA are small, red lesions that are an early indicator of DR. They are present in about 60-70% of the images in these datasets.

Hemorrhages: These are more severe lesions and are found in approximately 30-40% of images. Hemorrhages vary in size and intensity, with flame-shaped hemorrhages indicating more severe disease progression.

Exudates: These bright lesions are indicative of advanced DR stages. They are present in around 20-30% of images in the datasets, often appearing alongside other abnormalities like hemorrhages.

Dataset balance

Overrepresentation of MA: MA tend to be overrepresented due to their prevalence in early-stage DR.

Underrepresentation of severe lesions: Hemorrhages and exudates might be underrepresented, leading to a class imbalance where the model might perform better on detecting MA but struggle with less frequent lesions.

To counteract this, data augmentation techniques have been used to balance the dataset by artificially increasing the number of images showing severe lesions.

Patient demographics

- **Age distribution:** The majority of the patients in these datasets are likely adults over the age of 40, as DR prevalence increases with age.
- **Gender distribution:** There is usually a roughly equal distribution between male and female patients, though some datasets might slightly favor one gender over the other.
- **Ethnic representation:** In many datasets, there might be a bias towards certain ethnic groups depending on the dataset's origin (e.g., a dataset collected in Europe might predominantly feature Caucasian patients).

Disease severity

- **Severity distribution:** The datasets typically contain images representing various stages of DR:
- **Mild NPDR:** A significant portion of the images (40-50%) may belong to this category, showing early signs like MA.
- **Moderate to severe NPDR:** This category might account for 30-40% of the images, featuring more pronounced abnormalities like hemorrhages and cotton wool spots.
- **Proliferative DR (PDR):** The most severe cases, where abnormal BV proliferate, might represent around 10-20% of the dataset.

Image quality

- **Quality analysis:** Image quality in these datasets can vary significantly:
- **High-quality images:** High-resolution, well-focused images make up the majority (60-70%),

crucial for accurately identifying fine details like MA and small hemorrhages.

- **Medium-quality images:** Around 20-30% of images may suffer from minor issues like slight blurriness or uneven illumination, which can still be useful with proper preprocessing.
- **Low-quality images:** A small portion (5-10%) might be low-quality due to factors like poor focus, significant noise, or artifacts. These images are challenging for both manual and automated analysis but provide an opportunity to test the robustness of image enhancement techniques.

3.2 Data augmentation

Lack of data is a common concern when adopting machine learning in real-time. Obtaining such data can be costly and time-consuming. To mitigate this issue, a group of techniques known as “data augmentation” was employed to increase the dataset size by creating additional data points from existing images. These techniques enhance the diversity of the training set, thereby improving the model’s generalizability and robustness with the chosen parameters indicated in *Table 1*.

Table 1 Data augmentation approaches

Method	Default	Augmented
Horizontal flip	None	True
Rescale	-	1./255
Zoom range	-	0.25
Rotation (°)	-	60, 90 & 120
Fundus image sizes	999x960; 565x584; 1500x1152	224 x 224

The initial step involved flipping the images along the x-axis, followed by applying a rescale factor. This was part of the data augmentation technique, which also included rotating each unique patch by 60 degrees, 90 degrees, and 120 degrees, as depicted in *Figure 3*. Subsequently, the training vector was constructed using a random arrangement of these patches. Most pre-trained networks favour a standard input size of 224 by 224, so the image dimensions were rescaled accordingly.

Rationale for choosing specific augmentation Parameters:

- **Rotation angles (60, 90, and 120 degrees):** The rotation angles of 60, 90, and 120 degrees were selected to introduce sufficient variability in the training images while maintaining the anatomical integrity of retinal structures. Rotations at these angles simulate common variations in the orientation of retinal images due to patient

movement or imaging conditions, helping the model learn to recognize lesions from multiple perspectives. Using these specific angles ensures balanced coverage of possible orientations without introducing extreme rotations that could distort the retinal features.

- **Zoom range (0.25):** The zoom range of 0.25 was chosen to mimic the slight variations in image scale that may occur during retinal imaging. This level of zoom allows the model to become robust to changes in the apparent size of retinal features, ensuring that the model remains effective even when the images are captured at slightly different distances or resolutions. The chosen zoom range preserves the critical details necessary for accurate lesion detection, thus maintaining the effectiveness of the augmentation.

Only multiples of 30 degrees were used to avoid the necessity for any interpolation, which could introduce artifacts into the images. Each augmentation operation was applied consistently to both the

original image patches and their corresponding annotations, ensuring that the integrity of the dataset was maintained. These augmentation techniques were carefully selected to enhance the model's generalizability, particularly within the framework of fully convolutional networks (FCNs). The literature provides valuable suggestions for implementing data augmentation with rotations in fully connected CNNs (FC-CNNs), supporting the effectiveness of these specific parameters. By incorporating this rationale, the data augmentation strategy becomes a more integral and justified part of the overall methodology, ensuring that the choices made are transparent and well-founded.

While *Figure 3(a)* depicts the original image, *Figures 3(b), (c), and (d)* demonstrate how the retinal fundus images were rotated by 60, 90, and 120 degrees to ensure model recognition. These approaches helped handle over-fitting, allowing the model to generalize beyond the training set.

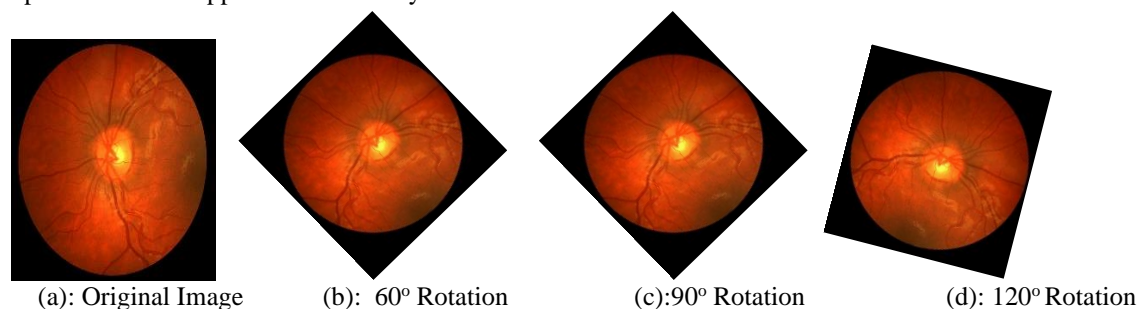


Figure 3 Data augmentation highlighting rotations effects at (b)60o, (c)90o, and (d)120o respectively

3.3 Image enhancement approaches on retinal fundus images

The image enhancement techniques employed were based on the literary findings [5, 10, 20]:

Median filter: This non-linear digital filtering technique effectively removes salt-and-pepper noise from an image while preserving edges. It is particularly useful for retaining the structural integrity of retinal features, such as BV, by eliminating isolated pixel artifacts without blurring important details.

Weiner filter: This statistical signal processing filter minimizes the mean square error in the presence of noise, enhancing image details while suppressing residual noise. Weiner filtering is applied after median filtering to refine the image further, ensuring that subtle features like MA and small vessels are highlighted.

Alternating sequential filter (ASF): This morphological filter smooths objects within an image

while preserving their shapes. It effectively removes noise and small structures, though its application is context-dependent and often reserved for specific cases where shape preservation is critical.

Gaussian blur: Gaussian blur reduces image noise and detail by applying a low-pass filter. However, it may cause a slight blur, reducing the sharpness of the image. It is used sparingly in cases where noise reduction is prioritized over detail preservation.

Luminosity: This measure of brightness in an image is calculated using a weighted sum of each pixel's RGB values. Luminosity adjustment is useful for standardizing the brightness across the dataset, which can improve the effectiveness of subsequent enhancement techniques. The most common formula to calculate the luminosity is presented in Equation 1:

$$L = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad (1)$$

CLAHE improves the local contrast of an image and reduces noise amplification introduced by traditional

HE. This technique is particularly beneficial in enhancing fundus images, making it easier to detect fine details like lesions.

3.4 Preprocessing and lesion segmentation

The quality of an image depends heavily on the sequence and effectiveness of the image enhancement techniques used during preprocessing. The chosen sequence: Median Filtering, Weiner Filtering, and CLAHE was selected based on their complementary strengths, which together prepare the retinal images for accurate lesion detection.

Justification for the preprocessing sequence:

- **Median filtering (First step):** Median filtering is applied first to remove salt-and-pepper noise while preserving the edges of critical retinal features. This step ensures that the image is free from isolated pixel artifacts, which could otherwise interfere with the effectiveness of subsequent enhancement techniques.
- **Weiner filtering (second step):** After median filtering, Weiner filtering is applied to further reduce any remaining noise and enhance image details. This step refines the image by emphasizing fine details, making it easier to detect small and

subtle lesions. By applying Weiner filtering after the initial noise removal, the image is progressively enhanced, ensuring that it is clean and detailed before contrast enhancement.

- **CLAHE (final step):** This is applied last to enhance the local contrast of the retinal images. With the noise reduced and details enhanced by the previous steps, CLAHE effectively highlights the critical features needed for accurate lesion detection. Applying this at the end of the pipeline maximizes its ability to improve the visibility of subtle features, ensuring that the images are well-prepared for segmentation and analysis.

The sequence: Median Filtering → Weiner Filtering → CLAHE, ensures that each preprocessing step builds on the improvements made by the previous one, resulting in a final image that is clean, detailed, and has enhanced contrast. This systematic approach is critical for maximizing the effectiveness of the lesion segmentation and detection process.

Therefore, the CLAHE method, a well-researched approach in image denoising, is incorporated to further improve image clarity, as shown in *Figure 4*.

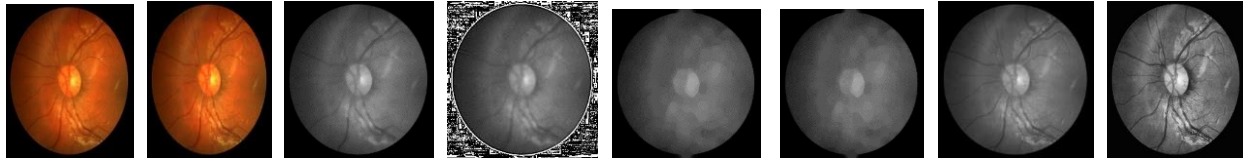


Figure 4 (i) Original Input Image (ii): After Rescaling, (iii) Median Filter (iv) Weiner Filter, (v) Luminosity (vi) ASF (vii) Gaussian Blur (viii) CLAHE

The following is the mathematical justification why CLAHE is generally considered superior to other methods:

Local vs global enhancement: Traditional HE applies a global transformation function that can wash out details. CLAHE, on the other hand, divides the image into small grids and applies HE to each one, which enables it to adapt to local changes in contrast as resented in Equation 2. This makes it more suited for medical images with important local features like BV and lesions.

$$CLAHE = \frac{Local\ Maximum + Local\ Minimum}{Local\ Maximum - Local\ Minimum} \quad (2)$$

Contrast limiting: In HE, noise can be amplified, making the image grainy. CLAHE has a “clip limit”, which helps to reduce the noise by clipping the histogram at a predefined value and redistributing it, which is a process often lacking in other methods. Thus,

$$Noise\ (CLAHE) < Noise\ (Traditional\ HE)$$

Preservation of mean brightness: CLAHE tends to preserve the mean brightness of an image, which is important in medical imaging to ensure that the characteristics of the tissue (like the color of BV and retina) are preserved.

$$\begin{aligned} Mean\ Brightness\ (CLAHE) &\approx \\ Mean\ Brightness\ (Original\ Image) \end{aligned}$$

The HE aims to transform an image $f(x, y)$ to an image $g(x, y)$ such that $g(x, y)$ has a uniform histogram. The transformation (T) is given by Equation 3:

$$T(v) = \int_0^v Pr(w) dw \quad (3)$$

Where:

$Pr(w)$ is the probability density function (pdf) of the image's intensity values.

While there is no definitive mathematical equation that proves CLAHE is superior, its advantages come from its adaptivity and the ability to limit contrast enhancement, making it more suitable for images with varying contrast and brightness.

CLAHE's apparent benefits in making finding and identifying vessels simpler are displayed in *Figure 4* (viii) which shows how the approach has improved the region of interest (ROI). More accurately distinguishing retinal images with neovascularization from those without is made possible by the proposed preprocessing approaches.

3.5 Model training and validation

The dataset for DR detection was divided into training (70%), validation (15%), and test (15%) sets. To prepare a representative validation set, the images were shuffled, then sampled to include a balanced distribution of abnormalities (MA, HEM, and EX) and disease severity levels. This approach ensured that the validation set closely reflected the diversity within the overall dataset, aiding in effective hyperparameter tuning and minimizing overfitting. The test set was reserved exclusively for final performance evaluation.

Use of the validation set during training:

During the model training process, the validation set was used to evaluate the model's performance at the end of each training epoch. The following practices were employed:

- **Hyperparameter Tuning:** The validation set was used to fine-tune hyperparameters such as learning rate, batch size, and regularization techniques (e.g., dropout rates). By observing the model's performance on the validation set, adjustments were made to the hyperparameters to optimize the model's ability to generalize to new data.
- **Early Stopping:** An early stopping criterion was implemented based on the validation loss. If the validation loss did not improve over a specified number of epochs (patience), the training was halted to prevent overfitting. This approach ensured that the model did not become overly specialized to the training data, maintaining its ability to perform well on unseen data.
- **Model selection:** Throughout the training process, the model with the lowest validation loss and highest validation accuracy was saved as the best-performing model. This model was then used for the final evaluation on the test set.

Importance of the validation set:

The use of the validation set was crucial for achieving a balanced model that is both accurate and generalizable. By continuously monitoring the model's performance on the validation set, overfitting was minimized, and the model was fine-tuned to perform optimally on a diverse range of retinal images.

3.6 Software and hardware requirements

The proposed methodology for DR detection was implemented using Python 3.8 due to its robust support for image processing and data analysis. The experiments were conducted on systems running Ubuntu 20.04 LTS and Windows 10 for versatility. The hardware setup comprised an Intel Core i7 processor with 16 GB RAM and an NVIDIA GeForce GTX 1080 GPU, which facilitated efficient model training and high-resolution image analysis. Additionally, a 1 TB SSD supported rapid data retrieval and storage, ensuring smooth handling of large datasets and intensive computations.

3.7 Proposed methodology

3.7.1 BV and optic disc segmentation

a) Enhancing luminance in retinal images

One of the challenges in retinal imaging is the inconsistent lighting conditions, which can obscure essential details such as lesions. Standard retinal images are captured in Red, Green and Blue (RGB) colour format. However, the RGB channels not only carry color information but are also sensitive to changes in luminosity. To account for that, a specialized luminance gain matrix $G(p, q)$ is introduced to uniformly adjust the luminosity across the Red, Green, and Blue (RGB) channels at each pixel location (p, q) . The enhanced R, G, and B channels $r'(p, q)$, $g'(p, q)$, $b'(p, q)$ are calculated from the original channels $r(p, q)$, $g(p, q)$, $b(p, q)$ through this matrix. To maintain the integrity of the original colors while focusing solely on luminance, the RGB image is converted into the Hue-Saturation-Value (HSV) color space. Here, the Value (V) channel can be modified independently, without affecting the Hue (H) and Saturation (S). In this framework, the gain matrix $G(p, q)$ is computed as a function of the luminance channel $V(p, q)$, denoted as $V'(p, q)$. This function aims to amplify or reduce the original luminance levels in a uniform manner across the image. Gamma correction is a commonly used technique for adjusting luminance levels. The transformation is given by Equation 4:

$$w = \mu^Y \quad (4)$$

Where μ = the normalized value of the luminance channel

w = the normalized output, and

γ = a constant that determines the degree of correction.

By adopting these methods, the luminance of the retinal image can be uniformly enhanced, making it easier to identify lesions and other features. This approach aligns with existing research, such as the

work by Yang et al. [48], which also advocates for luminance adjustments in imaging.

Sequel to the preprocessing and image enhancement, an approach to segment BV and optic disc are proposed and highlighted in *Figure 5*. These two are segmented and removed from the fundus images, and what is left are the lesions, viz: Exudates, Hemorrhages, and MA.

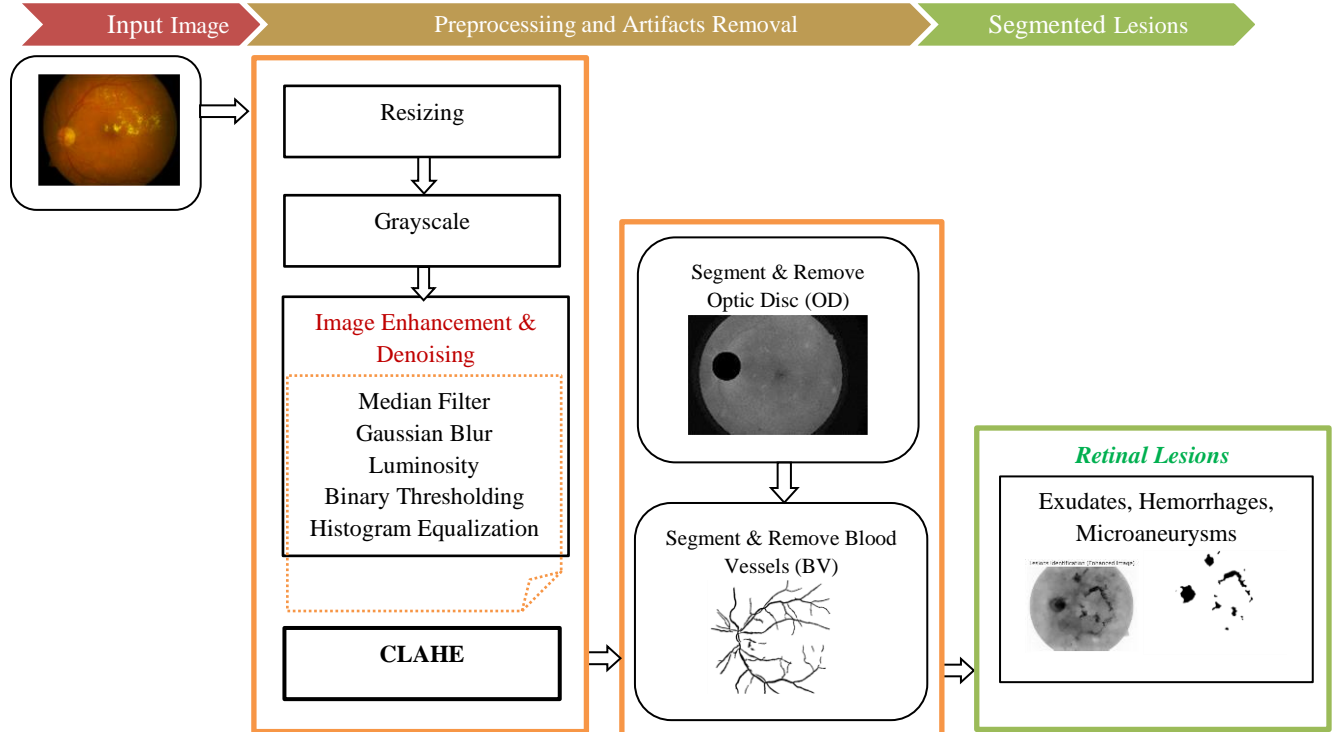


Figure 5 Block diagram of the proposed methodology

b) BV segmentation in retinal imaging

Isolating the BV in retinal images is crucial for diagnosing and analyzing DR. To isolate BV in retinal images for DR diagnosis, a multi-stage pipeline was developed. Principal Component Analysis (PCA) was initially applied to extract uncorrelated principal components that capture BV patterns, reducing data dimensionality and enhancing processing efficiency. CLAHE then enhanced contrast to make the BV more distinguishable. This was followed by average filtering to smooth the image and reduce noise, preparing it for thresholding. ISODATA thresholding was used for BV segmentation by dynamically adjusting clusters based on pixel intensities. ASF and background removal further isolated the BV, while binary thresholding and noise removal techniques finalized the vessel

isolation by separating BV structures from background noise.

The maximum intensity value is 255 (white) and the minimum 0 (for black). Given an input image $Img(x,y)$, threshold value was computed using Equation 5.

$$Img(x,y) = \begin{cases} \maxVal: f(src)[x,y] > threshold \\ 0: Otherwise \end{cases} \quad (5)$$

Where:

$Img(x,y)$ = Input Image with intensity values, and
 $f(src)$ = Source Image.

The proposed pipeline leverages PCA, a statistical method that transforms the correlated variables into a new set of linearly independent variables known as principal components. At first, the images undergo a

binary thresholding. A recommended threshold value of 11 was used for optimal results in filtering out unwanted outliers, especially bright pixels, from an image. Thresholding is a type of binarizing that eliminates exceptionally high or low pixel values that can lead to significant differences in extracted features. Because they are darker than the surrounding retinal tissue, BV benefit from thresholding to stand out clearly. To successfully isolated them, the images were binarized. There was noise and other artefacts in the output image, and erosion operation was used to get rid of those. As only blood arteries are linear in shape, this helps eliminate unwanted parts as illustrated in *Figure 6*.

c) Optic Disc identification in retinal images

The precise localization of the optic disc is vital for accurate retinal image analysis. *Figure 7* presents a detailed flowchart illustrating the methodology behind this essential task. The technique relies on a

series of image processing steps, beginning with morphological closing and concluding with area-based filtering for background exclusion.

i- Initial Enhancement and morphological closing

The first step involves applying a morphological closing operation to the pre-filtered image. This is followed by adaptive histogram equalization (AHE) to enhance the image quality. The morphological closing helps highlight areas with higher intensity, which typically correspond to the optic disc.

ii- Transition to convex hull representation

Once the image is prepared, it is converted into its convex hull form. In computer vision, a convex hull refers to the smallest convex polygon that can enclose all the white pixels in a binary image. This convex polygon has the property that any line segment drawn between two points inside it will remain inside the polygon. This representation aids in focusing on significant areas while ignoring irrelevant details.

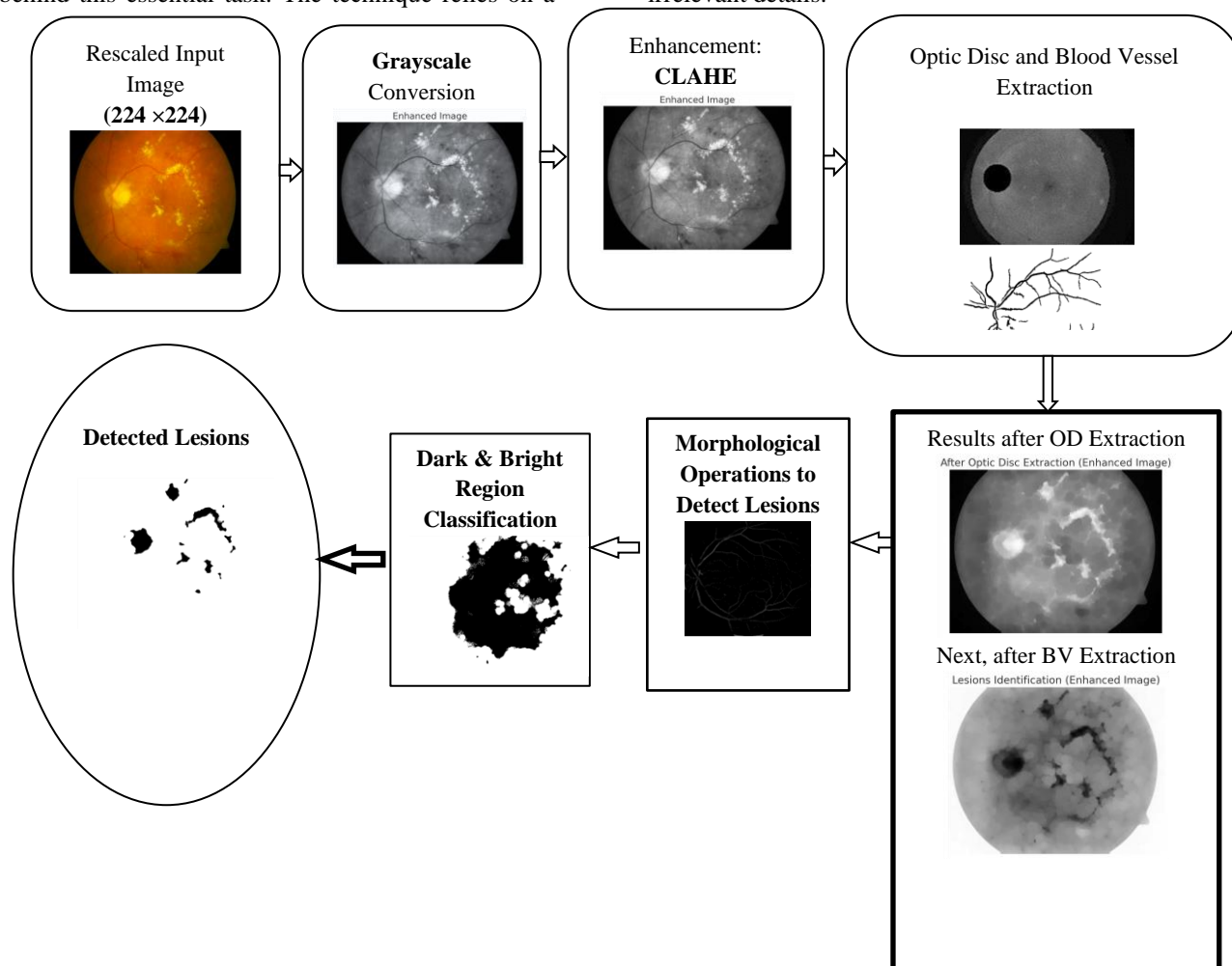


Figure 6 Block diagram of the image enhancement, and lesion segmentation

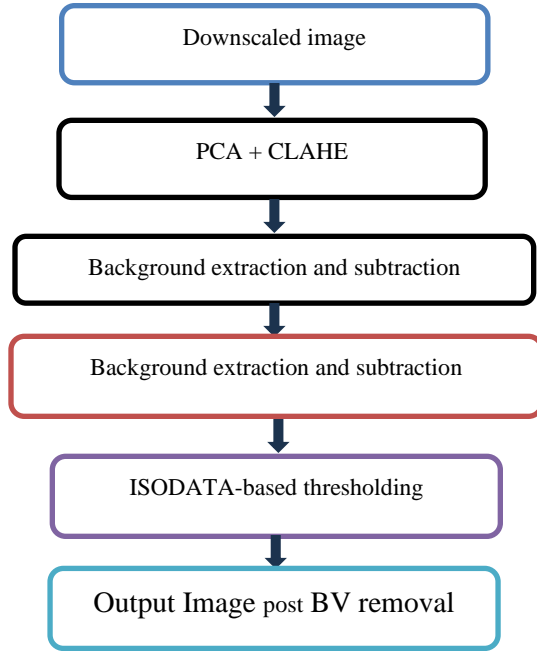


Figure 7 The process involves in the BV detection

iii- Area-based optic disc detection

After obtaining the convex hull image, the optic disc is identified based on its area. The smaller regions, which are unlikely to be the optic disc, are then eliminated using a background removal algorithm.

optic disc Segmentation:

The optic disc is a critical anatomical feature in retinal images, and accurate segmentation is essential to avoid confusion with lesions such as exudates. The algorithm 2 was proposed for optic disc extraction:

Algorithm 2: OpticDiscExtraction(ODE)

Begin ODE(Img):

Do:

Create a window size $w \times w$: A sliding window of size $w \times w$ is created to scan the retinal image.

Slide the window with stride s : From top-left corner of the image, slide the window with stride s in both directions.

FOR each window, compute its mean intensity and store it in a list.

Identify the window with the highest mean intensity corresponding to the optic disc.

Create a new image optic disc (od) with the exact dimensions as e .

FOR each pixel in e , check if the corresponding pixel falls within the boundaries of the od window.

IF True, set its value to 0 in the od image,

ELSE copy the value of the pixel from e to the od image.

ENDFOR

ENDFOR

Output the od image.

End Do

End Procedure

An appropriately sized window scans the image with a stride s , calculating the mean intensity of each window for each scan. Given that the optic disc is larger and has a high intensity, it typically exhibits the most prominent brightness. To eliminate its influence, a dark window is utilized. The intensity compensation technique, inspired by the work of Sanjani et al. [49], begins by identifying pixels likely to represent the background. The average and standard deviation (STD) of intensities from this region are then used to calibrate the image's brightness.

iv-Fine-tuning with pixel indexing and area properties

Smaller regions within the image are further analyzed by examining the index values of the pixels. The optic disc is then precisely extracted based on the property of area, ensuring a highly accurate identification.

v- Background exclusion for final extraction

The final step involves the use of a background exclusion method to isolate the optic disc from the remaining structures in the retinal image. Background subtraction is a pivotal technique in image processing that isolates the foreground, which usually contains the regions or objects of interest, for subsequent analyses. *Figure 8* provides a visual illustration of this technique, which utilizes basic arithmetic operations to accomplish complex tasks like object segmentation in computer vision. In computer vision, image subtraction is a straightforward yet effective method for segmenting out important objects or features. For each pixel in the image $I_{(x, y)}$, the corresponding pixel value, denoted as $P[I(x, y)]$, is subtracted from the same pixel in a background image $B_{(x, y)}$.

The background subtraction equation is represented in Equation 6:

$$P[F(x, y)] = P[I(x, y)] - P[B(x, y)] \quad (6)$$

Where $P[F(x, y)]$ = the pixel value in the foreground image after subtraction.

$P[B(x, y)]$ = the background, which in this context is composed of the optic disc and BV that have been previously extracted. By applying Equation 6, the

optic disc and BV were effectively removed, allowing for more focused and accurate lesion detection in retinal images, as illustrated in *Figure 8*.

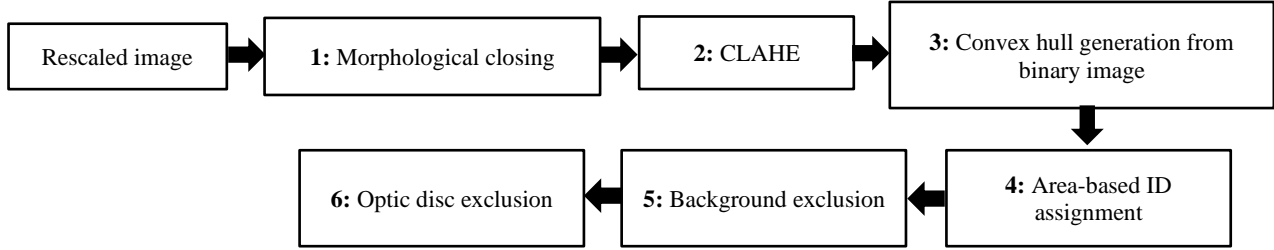


Figure 8 The flow diagram for OD detection

d) Lesion identification in retinal images

Once the optic disc and BV are efficiently isolated and removed, the next crucial step is the identification of lesions. Morphological operations are employed to facilitate this task, primarily leveraging intensity values and advanced transforms like the extended-minima method [50]. The extended-minima transform, an evolved form of the H-minima transform, identifies regional minima in an image. Each regional minimum comprises pixels with consistent intensity values, bordered by pixels with higher intensity. This technique is particularly useful for identifying the boundaries of lesions [51]. In computational frameworks like Python, the extended-minima transform can be achieved using the scikit-image library, specifically the *extrema.h_minima* function. This function suppresses all minima in the image that are shallower than a given depth h . For two-dimensional (2D) images, 8-associated neighborhoods are usually considered, while 26-associated neighborhoods can be employed for three-dimensional (3D) images. *Figure 9* illustrates the systematic approach adopted for lesion detection and classification. Post-detection, lesions are categorized into three main types: MA, HEM, and EX, based on their area. As outlined in recent studies[52]:

- MA: Lesions with an area smaller than 100 pixels.
- Hemorrhages: Lesions having an area ranging from 100 to 500 pixels.
- Exudates: Lesions with an area exceeding 500 pixels.

The severity of the condition is gauged based on the total count of lesions present in the retinal image

e) Performance evaluation

The metrics presented in equations 7-16, were used to evaluate the model's performance in two phases viz:

- Image Enhancement and
- Lesion Detection

i) Evaluation metrics for enhanced images

Mean Squared Error (MSE)

This evaluates the average squared differences between the pixels of the original image (A) and the enhanced image (B). A lower MSE indicates a closer match to the original image as provided in Equation 7.

$$MSE = \frac{1}{pq} \sum_{k=0}^{p-1} \sum_{l=0}^{q-1} (A(k, l) - B(k, l))^2 \quad (7)$$

Peak signal-to-noise ratio (PSNR)

PSNR is another quality metric, which uses MSE for its calculation. A higher PSNR indicates a higher-quality image as per Equation 8.

$$PSNR = 10 \times \log_{10} \left(\frac{(255)^2}{MSE} \right) \quad (8)$$

Normalized cross-correlation (NCC)

NCC is used to measure the similarity between the original (A) and enhanced (B) images. Higher values indicate greater similarity (Equation 9).

$$NCC = \frac{\sum_{k=0}^{p-1} \sum_{l=0}^{q-1} (A(k, l) - \bar{A})(B(k, l) - \bar{B})}{\sqrt{\sum_{k=0}^{p-1} \sum_{l=0}^{q-1} (A(k, l) - \bar{A})^2 \sum_{k=0}^{p-1} \sum_{l=0}^{q-1} (B(k, l) - \bar{B})^2}} \quad (9)$$

Mean absolute difference (MAD)

This metric calculates the average absolute differences between corresponding pixels in the original and enhanced images. Lower values indicate higher similarity presented in Equation 10.

$$MAD = \frac{1}{pq} \sum_{k=0}^{p-1} \sum_{l=0}^{q-1} |A(k, l) - B(k, l)| \quad (10)$$

ii) Evaluation metrics for lesion detection

The performance for lesion detection is calculated based on accuracy, precision, recall, sensitivity, specificity and F1-Score as illustrated in Equations 11 to 16.

Accuracy indicators are calculated based on the number of true and false positives (TP, FP) and true and false negatives (TN, FN) that occur on a pixel-by-pixel basis during the evaluation process. At the same time, recall is the fraction of correct classification in the dataset. While sensitivity measures the tendency to correctly detect DR lesions, specificity measures the tendency to correctly reject DR lesions.

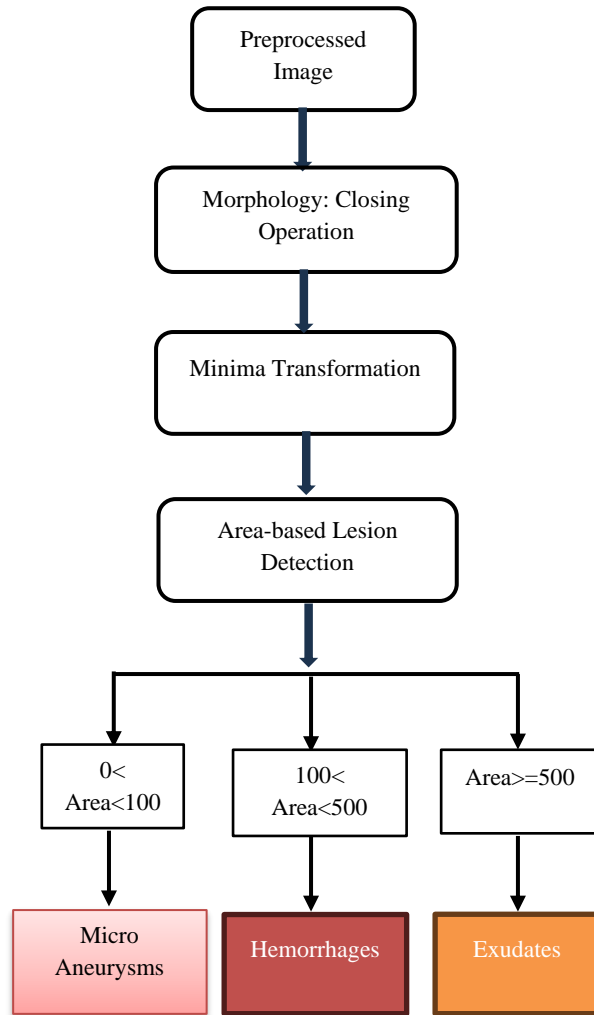


Figure 9 The systematic approach adopted for lesion detection and classification

F1-Score highlights the system that more accurately depicts the future with a lower value. Specificity measures the ration of correctly classified negative instances to the positive. If the results are closer to 1, they are favourable, whereas values closer to 0 indicate uncertainty. The area under the receiver operating characteristic curve (AUROC) is a

performance measurement for classification problems, summarizing the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity) across various thresholds. In the context of lesion detection for DR, a high AUROC value would indicate that the system is capable of distinguishing between pixels that correctly represent lesions and those that do not, regardless of the threshold set for classification, shown in Equations 11 to 16.

$$Accuracy\% = \left(\frac{TP + TN}{TP + FP + TN + FN} \right) * 100 \quad (11)$$

$$Precision\% = \left(\frac{TP}{TP + FP} \right) * 100 \quad (12)$$

$$Sensitivity(Recall)\% = \left(\frac{TP}{TP + FN} \right) * 100 \quad (13)$$

$$Specificity\% = \left(\frac{TN}{TN + FP} \right) * 100 \quad (14)$$

$$F1 - Score\% = 2 \left(\frac{Precision * Sensitivity}{Precision + Sensitivity} \right) \quad (15)$$

$$AUROC = \sum_{i=0}^n ((x_{i+1} - x_i) * (y_{i+1} - y_i)) / 2 \quad (16)$$

Where:

(x_i, y_i) and (x_{i+1}, y_{i+1}) = consecutive points on the ROC Curve, sorted by their x-values.

n = the number of points on the curve

TP = True Positives

TN = True Negatives

FP = False Positives

FN = False Negatives

4.Results and discussion


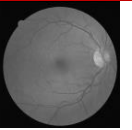
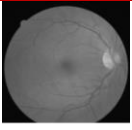
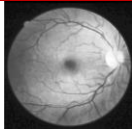




The section highlights evaluation results obtained from our proposed segmentation methodology. Initially focused on two datasets namely the CHASEDB1, and DRIVE dataset, but extended the analysis to DIARETDB1[53] dataset.

The study utilized the DIARETDB1 dataset, enhancing images through gamma correction for better analysis of retinal conditions. Subsequent to this enhancement, vital structures like BV and the optic disc were isolated and removed via background subtraction methods. Morphological operations were employed on these pre-processed images to identify lesions indicative of DR. Special attention was given to the green channel of the RGB images, which provided clearer views of vessels and other background elements, making it easier to differentiate lesions as illustrated in Table 2. Post gamma correction and matched filtering, lesions were identified based on their intensity levels. These were then categorized into hemorrhages, MA, and EX based on their area. The final images were color-

coded—blue for MA, red for HEM, and green for EX—to aid in the diagnostic process. This color-coding, combined with overlapping the identified

lesions on the original retinal images, offers a comprehensive view for efficient and accurate diagnosis of DR stages.

Table 2 Effects of channel extraction and contrast enhanced image on CHSEDB1 and DRIVE datasets

Dataset	Original Image	Grayscale (224×224)	Green Channel Image	Channel	Luminosity and CLAHE enhanced image
CHASEDB1					
DRIVE					

The proposed approach was tested on both the DRIVE and CHASEDB1 datasets. It was observed that our proposed approach was on par with the ground truth provided by seasoned ophthalmologists.

4.1 Detection of exudates (HE and SE)

For the exudate’s segmentation, the CHASEDB1 and DIAREDB1 were used since both contains validated

ground-truths. However, the latter contains hard and SE unlike the former without sub-classes. *Table 3* showcases the exudates as detected by the applied algorithm. These are then compared with the ground truth. This is further validated using the DIARETDB1 datasets as presented in *Table 4*.

Table 3 Segmentation results on the CHASEDB1 and DRIVE datasets

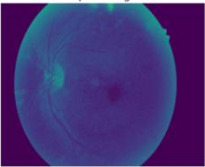


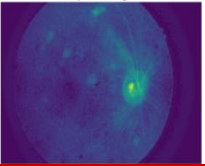






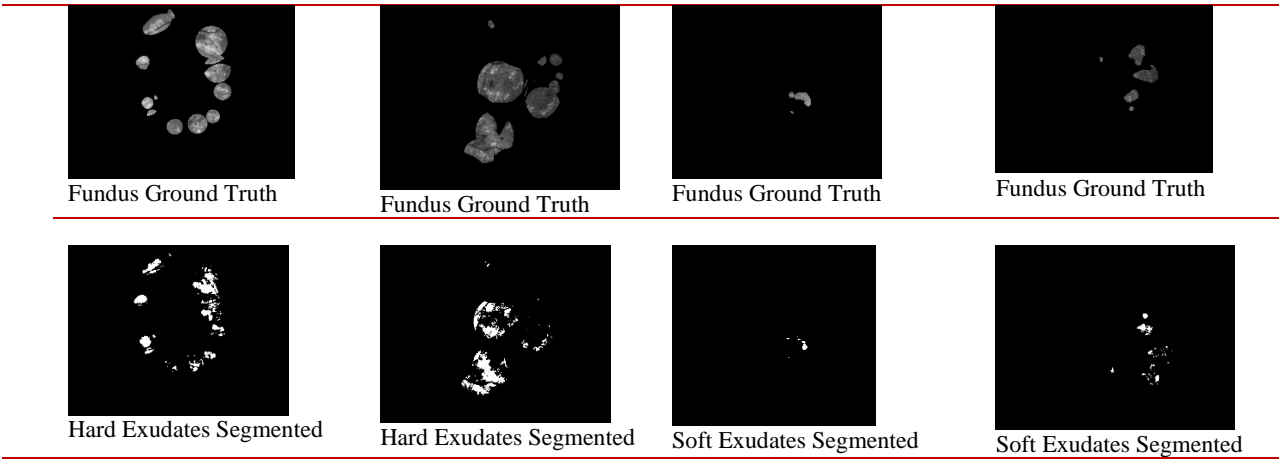
Datasets	Input image	Ground truth	Segmented exudates
CHASEDB1			
DRIVE			

Table 4 Segmentation results on the DIARETDB1 dataset that contains hard and SE

D I A R E T D B 1				
	Fundus Image1	Fundus Image2	Fundus Image3	Fundus Image4



4.2 Detection of MA

Results for identifying MA are demonstrated using “image013”, a select image from the DIARETDB1

dataset. The MA identified is illustrated in Table 5, and further compared with the ground truth for validation.

Table 5 Segmentation results of MA using the DIARETDB1 dataset

Datasets	Input image	Ground truth	Segmented exudates
DIARETDB1			

4.3 Detection of hemorrhages (HM)

Further, another sample of an image (image016) from the DIABETDB1 dataset serves as the test input for detecting hemorrhages. The identified hemorrhages

are displayed in Table 6 highlighted in red bounding box in comparison with the ground truth. Tables 7 Shows the performance metrics for image enhancement, and lesion detection.

Table 6 Segmentation results of HM

Datasets	Input image	Ground truth	Segmented exudates
DIARET1			

Table 7 Performance metric for lesion detection

S. No.	Lesions	Sensitivity (%)	Specificity (%)	Accuracy (%)
1	MA	91.50	88.20	90.40
2	Hemorrhages	92.20	89.10	91.30
3	Exudates	93.10	90.20	92.60

Table 8 Performance metric for image enhancement

S. No.	Performance metrics	Values
1	PSNR	21.80
2	NCC	0.812

S. No.	Performance metrics	Values
3	Average difference (AD)	15.20
4	Maximum difference (MD)	40.00
5	Normalized absolute error (NAE)	0.315

The high-performance metrics in *Table 7* indicate that the system is reliable in detecting various types of retinal lesions, crucial for early and accurate diagnosis. Meanwhile, the strong metrics in *Table 8* suggests that the image enhancement techniques employed are effective in preserving image quality and structural integrity which in turn aids in more accurate lesion detection. *Figure 10* shows the ROC curves comparing the performance of the proposed and existing models in detecting various lesion types, with the area under the curve (AUC) values indicating strong classification accuracy, particularly for bright lesions (AUC=0.97) and exudates (AUC=0.94). The proposed method's performance metrics, as shown in *Table 9*, highlight its superiority over several state-of-the-art (SoTA) approaches,

including Hough Transform, SVM, morphological models, Gabor filters, and ridge-based filters, particularly when applied to the CHASEDB1 dataset. The method achieved an accuracy of 98.51%, exceeding other models, with sensitivity values of 80% for hemorrhages and 99.50% for exudates. Its AUROC of 99.51% further emphasizes its exceptional capability in distinguishing DR lesions. These metrics, alongside high F1-scores and specificity, underscore the proposed method's reliability and robustness for accurately detecting various types of DR lesions. This strong performance across CHASEDB1 and additional datasets such as DRIVE and DIARETDB1 suggests the method's potential for clinical applications, offering high diagnostic precision and computational efficiency.

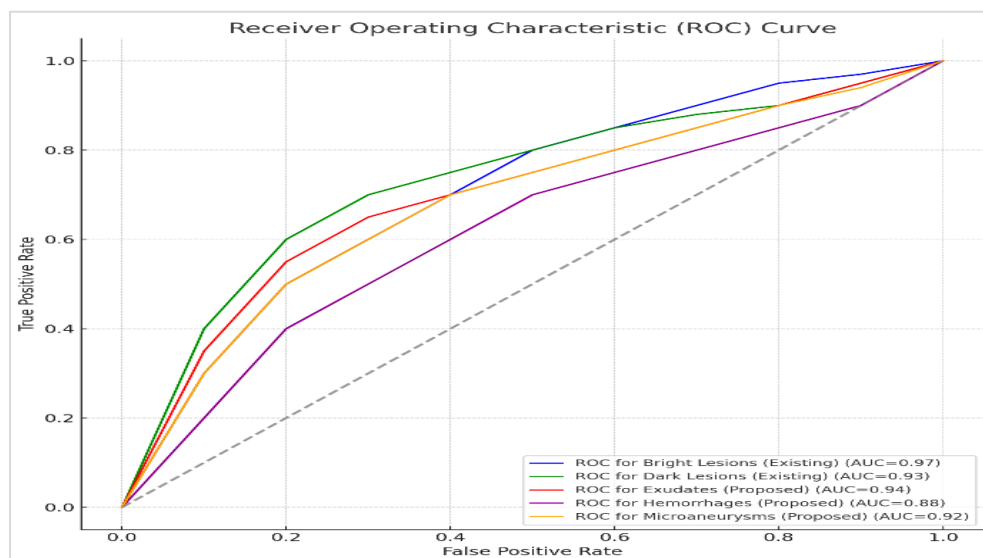


Figure 10 The AUROC curve on the CHASEDB1 dataset

Table 9 Comparison with SoTA

Source	Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUROC (%)
[49]	Hough Transform	80	72	85	74	83
[50]	SVM	97	85	96	90	95
[51]	Morphological	99	78	99	88	98
[52]	Gabor Filter	85	75	89	80	87
[53]	Ridge-based filter	94	79	93	85	92
[49]	Hough Transform	80	72	85	74	83
[50]	SVM	97	85	96	90	95

Source	Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUROC (%)
[51]	Morphological	99	78	99	88	98
[52]	Gabor Filter	85	75	89	80	87
[53]	Ridge-based filter	94	79	93	85	92
Proposed	*Proposed	98.51	80	99.50	90.18	99.51

NB: Best performance metrics are highlighted in bold figures

4.1 Limitations of the study

Despite the promising results, the study has certain limitations:

1. The datasets used in this study, while standard in the field, may not fully represent the diversity of retinal images encountered in real-world clinical settings. The model's performance on more diverse datasets remains to be validated.
2. Although the proposed method is more efficient than some SoTA models, it still requires substantial computational power, particularly during the training phase. This could limit its applicability in low-resource environments.
3. The model was specifically designed to detect three types of lesions—MA, HE, and EX. Other retinal abnormalities were not considered, which limits the comprehensiveness of the tool.
4. As discussed earlier, the model did produce instances of false positives and false negatives. While these were minimized through careful tuning and preprocessing, they still represent a limitation in terms of the model's reliability. False positives could lead to unnecessary follow-up procedures, while false negatives could result in missed diagnoses, both of which have significant clinical implications.
5. The methodology was specifically designed and tested for retinal images. Its applicability to other imaging modalities (e.g., MRI, CT scans) has not been tested and would require further research to adapt and validate the approach for different types of medical images.

A complete list of abbreviations is listed in *Appendix I*.

5. Conclusion and future work

The research presented a robust approach to retinal image analysis for the detection and classification of MA, HEM, and EX, which are essential markers for DR. Employing advanced image processing and computational techniques, the proposed methodology achieved notable performance across various metrics such as accuracy, sensitivity, and specificity. Comparative analysis with SoTA methods further validated the model's robustness, as indicated by the high AUROC value.

However, the study is constrained by limitations like data specificity, computational intensity, and a restricted focus on three lesion types. These limitations point to avenues for future research. First, the algorithm could be tested on different and larger datasets to validate its generalizability. Second, optimizing the computational aspects of the model can make it more suitable for real-time medical applications. Third, extending the model to identify other retinal abnormalities will make it more comprehensive.

Further, the high AUROC score suggests that the automated system can effectively distinguish between normal and DR images, a critical aspect of screening for this condition. In addition, this indicates that the system can identify positive cases of DR, which can help improve patient outcomes by facilitating early detection and treatment. Lastly, a clinical validation involving expert ophthalmologists could provide insights into the model's practical applicability and potential for aiding in early DR detection.

The results suggest that the automated system is a promising tool for DR screening, with high accuracy, specificity, and F1-Score and a solid ability to distinguish between normal and DR images. Future work will involve using real-time datasets such as the Asia Pacific Tele-Ophthalmology Society (APTOS) dataset that focused explicitly on DR grading. Expert ophthalmologists validate the dataset to create a reliable computer-assisted diagnostic tool to generalize to new datasets. Overall, this study contributes to developing a valuable tool for the early detection and management of DR.

Acknowledgment

None.

Conflicts of interest

The authors have no conflicts of interest to declare.

Data availability

The datasets used in this research were obtained from publicly available sources: <https://drive.google.com/drive/dataset/chase-db1>, <https://paperswithcode.com/dataset/chase-db1>, and <http://www2.it.lut.fi/project/imageret/diaretdb1>.

Author's contribution statement

Sowmyashree B.: Conceptualization, investigation, data collection, design, writing - original draft, writing - review and editing, analysis and interpretation of results. **Maresh K. Rao:** Study Conception, Supervision, Investigation on challenges and Draft manuscript verification. **Chethan H.K.:** Study conception, supervision, investigation on challenges and draft manuscript verification.

References

- [1] Forghani R. Machine learning and other artificial intelligence applications, an issue of neuroimaging clinics of North America, E-book: machine learning and other artificial intelligence applications. Elsevier Health Sciences; 2020.
- [2] Changhez J, James S, Jamala F, Khan S, Khan MZ, Gul S, et al. Evaluating the efficacy and accuracy of AI-assisted diagnostic techniques in endometrial carcinoma: a systematic review. *Cureus*. 2024; 16(5):1-11.
- [3] Wilder B, Pinedo A, Abusin S, Ansell D, Bacong AM, Calvin J, et al. A global perspective on socioeconomic determinants of cardiovascular health. *Canadian Journal of Cardiology*. 2024.
- [4] Ismail R, Ismail NH, Tamil AM, Ja'afar MH, Isa ZM, Nasir NM, et al. Prevalence and factors associated with prediabetes and diabetes mellitus among adults: baseline findings of pure Malaysia cohort study. *Clinical Epidemiology and Global Health*. 2023; 21:101279.
- [5] Kropp M, Golubnitschaja O, Mazurakova A, Koklesova L, Sargheini N, Vo TT, et al. Diabetic retinopathy as the leading cause of blindness and early predictor of cascading complications—risks and mitigation. *Epm Journal*. 2023; 14(1):21-42.
- [6] Hou X, Wang L, Zhu D, Guo L, Weng J, Zhang M, et al. Prevalence of diabetic retinopathy and vision-threatening diabetic retinopathy in adults with diabetes in China. *Nature Communications*. 2023; 14(1):4296.
- [7] Chaurasia S, Thool AR, Ansari KK, Saifi AI. Advancement in understanding diabetic retinopathy: a comprehensive review. *Cureus*. 2023; 15(11):1-11.
- [8] Li Z, Han Y, Yang X. Multi-fundus diseases classification using retinal optical coherence tomography images with swin transformer V2. *Journal of Imaging*. 2023; 9(10):1-18.
- [9] Zhang J, Ma K, Luo Z, Wang G, Feng Z, Huang Y, et al. Combining functional and morphological retinal vascular characteristics achieves high-precision diagnosis of mild non-proliferative diabetic retinopathy. *Journal of Translational Medicine*. 2024; 22(1):798.
- [10] Rajput Y, Gaikwad S, Dhupal R, Gaikwad J. The development of a tool for the detection of cotton wool spots, haemorrhage, and exudates using multi-resolution analysis. In *international conference on network security and blockchain technology 2023* (pp. 299-310). Singapore: Springer Nature Singapore.
- [11] Sandhya SG, Suhasini A. An efficient detection of micro aneurysms from fundus images with CDLNN algorithm. *Materials Today: Proceedings*. 2023; 81:553-62.
- [12] Care D. Standards of care in diabetes-2023. *Diabetes care*. 2023; 46:S1-267.
- [13] Sugandh FN, Chandio M, Raveena FN, Kumar L, Karishma FN, Khuwaja S, et al. Advances in the management of diabetes mellitus: a focus on personalized medicine. *Cureus*. 2023; 15(8):1-13.
- [14] Wolf RM, Channa R, Liu TA, Zehra A, Bromberger L, Patel D, et al. Autonomous artificial intelligence increases screening and follow-up for diabetic retinopathy in youth: the access randomized control trial. *Nature communications*. 2024; 15(1):421.
- [15] Naz H, Ahuja NJ, Nijhawan R. Diabetic retinopathy detection using supervised and unsupervised deep learning: a review study. *Artificial Intelligence Review*. 2024; 57(5):1-66.
- [16] Abushawish IY, Modak S, Abdel-raheem E, Mahmoud SA, Hussain AJ. Deep learning in automatic diabetic retinopathy detection and grading systems: a comprehensive survey and comparison of methods. *IEEE Access*. 2024; 12:84785-802.
- [17] Hemalakshmi GR, Santhi D, Mani VR, Geetha A, Prakash NB. Classification of retinal fundus image using MS-DRLBP features and CNN-RBF classifier. *Journal of Ambient Intelligence and Humanized Computing*. 2021; 12:8747-62.
- [18] Agughasi VI. Leveraging transfer learning for efficient diagnosis of COPD using CXR images and explainable AI techniques. *Inteligencia artificial*. 2024; 27(74):133-51.
- [19] Agughasi VI. The superiority of fine-tuning over full-training for the efficient diagnosis of COPD from CXR images. *Inteligencia Artificial*. 2024; 27(74):62-79.
- [20] Soares I, Castelo-branco M, Pinheiro A. Microaneurysms detection in retinal images using a multi-scale approach. *Biomedical Signal Processing and Control*. 2023; 79:104184.
- [21] Bhimavarapu U, Chintalapudi N, Battineni G. Automatic detection and classification of hypertensive retinopathy with improved convolution neural network and improved SVM. *Bioengineering*. 2024; 11(1):56.
- [22] Medhi JP, Sandeep R, Datta P, Nizami TK. Intelligent identification and classification of diabetic retinopathy using fuzzy inference system. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*. 2023; 11(6):2386-99.
- [23] Luo X, Wang W, Xu Y, Lai Z, Jin X, Zhang B, et al. A deep convolutional neural network for diabetic retinopathy detection via mining local and long-range dependence. *CAAI Transactions on Intelligence Technology*. 2024; 9(1):153-66.
- [24] Özbay E. An active deep learning method for diabetic retinopathy detection in segmented fundus images using artificial bee colony algorithm. *Artificial Intelligence Review*. 2023; 56(4):3291-318.
- [25] Nahiduzzaman M, Islam MR, Goni MO, Anower MS, Ahsan M, Haider J, et al. Diabetic retinopathy identification using parallel convolutional neural

- network based feature extractor and ELM classifier. *Expert Systems with Applications*. 2023; 217:119557.
- [26] Phridviraj MS, Bhukya R, Madugula S, Manjula A, Vodithala S, Waseem MS. A bi-directional long short-term memory-based diabetic retinopathy detection model using retinal fundus images. *Healthcare Analytics*. 2023; 3:100174.
- [27] Hu D, Pan L, Chen X, Xiao S, Wu Q. A novel vessel segmentation algorithm for pathological en-face images based on matched filter. *Physics in Medicine & Biology*. 2023; 68(5):055014.
- [28] Mousavi N, Monemian M, Ghaderi DP, Mirmohammadsadeghi M, Zekri M, Rabbani H. Cyst identification in retinal optical coherence tomography images using hidden markov model. *Scientific Reports*. 2023; 13(1):1-15.
- [29] Akram MU, Khalid S, Tariq A, Khan SA, Azam F. Detection and classification of retinal lesions for grading of diabetic retinopathy. *Computers in Biology and Medicine*. 2014; 45:161-71.
- [30] Ikechukwu AV, Murali S, Deepu R, Shivamurthy RC. ResNet-50 vs VGG-19 vs training from scratch: a comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images. *Global Transitions Proceedings*. 2021; 2(2):375-81.
- [31] Ikechukwu AV, Murali S. i-Net: a deep CNN model for white blood cancer segmentation and classification. *International Journal of Advanced Technology and Engineering Exploration*. 2022; 9(95):1448-64.
- [32] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*. 2012.
- [33] Dieleman S, De FJ, Kavukcuoglu K. Exploiting cyclic symmetry in convolutional neural networks. In *international conference on machine learning 2016* (pp. 1889-98). PMLR.
- [34] Agughasi VI, Srinivasiah M. Semi-supervised labelling of chest x-ray images using unsupervised clustering for ground-truth generation. *Applied Engineering and Technology*. 2023; 2(3):188-202.
- [35] Melinscak M, Prentasac P, Loncaric S. Retinal vessel segmentation using deep neural networks. In *VISAPP 2015* (pp. 577-582). Scite Press.
- [36] Liskowski P, Krawiec K. Segmenting retinal blood vessels with deep neural networks. *IEEE Transactions on Medical Imaging*. 2016; 35(11):2369-80.
- [37] Zhang Y, He M, Chen Z, Hu K, Li X, Gao X. Bridge-net: context-involved U-net with patch-based loss weight mapping for retinal blood vessel segmentation. *Expert Systems with Applications*. 2022; 195:116526.
- [38] Wang B, Qiu S, He H. Dual encoding u-net for retinal vessel segmentation. In *medical image computing and computer assisted intervention: 22nd international conference, Shenzhen, China, 2019* (pp. 84-92). Springer International Publishing.
- [39] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In *medical image computing and computer-assisted intervention: 18th international conference, Munich, Germany, 2015* (pp. 234-41). Springer International Publishing.
- [40] Guo C, Szemenyei M, Yi Y, Wang W, Chen B, Fan C. Sa-unet: spatial attention u-net for retinal vessel segmentation. In *25th international conference on pattern recognition 2021* (pp. 1236-42). IEEE.
- [41] Ikechukwu AV, Murali S. CX-Net: an efficient ensemble semantic deep neural network for ROI identification from chest-x-ray images for COPD diagnosis. *Machine Learning: Science and Technology*. 2023; 4(2):025021.
- [42] Wu Y, Xia Y, Song Y, Zhang D, Liu D, Zhang C, et al. Vessel-net: retinal vessel segmentation under multi-path supervision. In *medical image computing and computer assisted intervention: 22nd international conference, Shenzhen, China, 2019* (pp. 264-72). Springer International Publishing.
- [43] Zhang S, Fu H, Yan Y, Zhang Y, Wu Q, Yang M, et al. Attention guided network for retinal image segmentation. In *medical image computing and computer assisted intervention: 22nd international conference, Shenzhen, China, 2019* (pp. 797-805). Springer International Publishing.
- [44] Xu Y, Fan Y. Dual-channel asymmetric convolutional neural network for an efficient retinal blood vessel segmentation in eye fundus images. *Biocybernetics and Biomedical Engineering*. 2022; 42(2):695-706.
- [45] Kumar KS, Singh NP. Analysis of retinal blood vessel segmentation techniques: a systematic survey. *Multimedia Tools and Applications*. 2023; 82(5):7679-733.
- [46] Remeseiro B, Mendonça AM, Campilho A. Automatic classification of retinal blood vessels based on multilevel thresholding and graph propagation. *The Visual Computer*. 2021; 37(6):1247-61.
- [47] <https://drive.grand-challenge.org/DRIVE/>. Accessed 05 September 2024.
- [48] Yang B, Zhao H, Cao L, Liu H, Wang N, Li H. Retinal image enhancement with artifact reduction and structure retention. *Pattern Recognition*. 2023; 133:108968.
- [49] Sanjani SS, Boin JB, Bergen K. Blood vessel segmentation in retinal fundus images. Stanford University. 2013.
- [50] Tamim N, Elshrkawey M, Abdel AG, Nassar H. Retinal blood vessel segmentation using hybrid features and multi-layer perceptron neural networks. *Symmetry*. 2020; 12(6):1-27.
- [51] Cervantes J, Cervantes J, García-lamont F, Yee-rendon A, Cabrera JE, Jalili LD. A comprehensive survey on segmentation techniques for retinal vessel segmentation. *Neurocomputing*. 2023; 556:126626.
- [52] Jena PK, Khuntia B, Palai C, Nayak M, Mishra TK, Mohanty SN. A novel approach for diabetic retinopathy screening using asymmetric deep learning features. *Big Data and Cognitive Computing*. 2023; 7(1):1-16.
- [53] Kauppi T, Kalesnykiene V, Kamarainen JK, Lensu L, Sorri I, Raninen A, et al. The diaretdb1 diabetic

retinopathy database and evaluation protocol. In BMVC 2007 (pp. 1-18).



Sowmyashree B. holds a Bachelor of Engineering in Electronics and Communication Engineering and a Master of Technology in VLSI and Embedded Systems. She is currently pursuing a Ph.D. in Electronics and Communication Engineering at Visvesvaraya Technological University (VTU), Belagavi, India. In addition to her academic qualifications, she has earned several professional certifications and has developed expertise in various technical skills. Presently, she serves as an Assistant Professor in the Department of Electronics and Communication Engineering at Maharaja Institute of Technology, Mysore. Sowmyashree is an active member of professional organizations, including the Indian Society for Technical Education (ISTE), the International Association of Engineers (IAENG), and IEEE. Her research interests lie in the areas of Image Processing and the applications of Artificial Intelligence and Machine Learning. Email: sowmyashreeb.kp@gmail.com



Dr. Mahesh K. Rao is a distinguished expert in Electrical Engineering, with a Ph.D. from the University of Wyoming USA, earned in 1988, where his research focused on 3D actuator design for robotics. He began his academic journey with a Bachelor of Engineering in Electronics and Communication from SJCE, University of Mysore, in 1981, followed by a Master's degree from the University of Windsor in 1984. Dr. Rao worked as an Assistant Professor at the University of Nevada, Reno, before transitioning to the industry, where he joined Intel in 1992. At Intel, he progressed to the role of Validation Manager for Chipsets and later served as Program Manager for Integrated Microprocessors and returned to Intel India as General Manager. In 2002, he founded Aspire Communication, a successful company specializing in embedded product solutions which was sold to Alten in 2011. After a successful stint in the industry, Dr. Rao returned to academia as Professor and Head of the Department of Electronics and Communication Engineering at Maharaja Institute of Technology, Mysore. He has authored over 50 research papers and holds numerous patents, with his research focusing on Image Processing, Artificial Intelligence, and Machine Learning. Email: raomahesh08@gmail.com



Dr. Chethan H.K. holds a Ph.D. in Computer Science from the University of Mysore, India, and currently serves as a Professor in the Department of Computer Science & Engineering at Maharaja Institute of Technology Mysore, India. His expertise covers a

diverse range of disciplines, including Algorithms, Artificial Intelligence, and Software Engineering. Dr. Chethan has developed deep proficiency in specialized areas such as Pattern Recognition, Computer Vision, Document Analysis, Image Processing, Feature Extraction, Image Segmentation, and Digital Image Processing. Email: hkchethan@gmail.com

Appendix I

S. No.	Abbreviation	Description
1	2D	Two-Dimensional
2	3D	Three-Dimensional
3	AD	Average Difference
4	AHE	Adaptive Histogram Equalization
5	AG-Net	Attention Network
6	APTOS	Asia Pacific Tele-Ophthalmology Society
7	ASF	Alternating Sequential Filter
8	AUROC	Area Under the Receiver Operating Characteristic Curve
9	BV	Blood Vessel
10	CLAHE	Contrast-Limited Adaptive Histogram Equalization
11	CNN	Convolutional Neural Network
12	DEV-Net	Dual Encoding U-Net
13	DR	Diabetic Retinopathy
14	FC-CNNs	Fully Connected CNNs
15	FCN	Fully Convolutional Network
16	FCNs	Fully Convolutional Networks
17	FN	False Negative
18	FP	False Positive
19	GPU	Graphical Processing Units
20	HE	Histogram Equalization
21	HSV	Hue, Saturation, and Value
22	ISODATA	Iterative Self-Organizing Data Analysis Technique Algorithm
23	JPEG	Joint Photographic Experts Group
24	KNN	K-Nearest Neighbor
25	MA	Microaneurysms
26	MAD	Mean Absolute Difference
27	MAT	Macular Atrophy
28	MD	Maximum Difference
29	MSE	Mean Squared Error
30	NCC	Normalized Cross-Correlation
31	NAE	Normalized Absolute Error
32	NPDR	Non-Proliferative Diabetic Retinopathy
33	ODE	Optic Disk Extraction
34	PCA	Principal Component Analysis
35	PDR	Proliferative Diabetic Retinopathy
36	PSNR	Peak Signal-to-Noise Ratio
37	ResNet50	Residual Networks with 50 Layers
38	RGB	Red, Green, and Blue (Color Model)
39	ROI	Region of Interest
40	ROC	Receiver Operating Characteristic
41	SE	Soft Exudates
42	SoTA	State-of-the-Art
43	STD	Standard Deviation
44	SVM	Support Vector Machine
45	TN	True Negative
46	TP	True Positive
47	VGG	Visual Geometry Group with 19 Layers