

---

# Retinal Vessel Image Segmentation

---

Utsav Garg (2021108) Parthiv Dholaria (2021078)

## Dataset

The **CHASEDB1** dataset is widely used for retinal vessel segmentation tasks. Below are the key details of the dataset:

- The dataset comprises **28 retinal images**, representing both eyes (**left (L) and right (R)**) of **14 participants**.
- Each image has a resolution of **2,688,000 pixels** and is accompanied by **two ground truth annotations** provided by different human observers, labeled as "**1stHO**" and "**2ndHO**".
- This results in a total of **84 images**, including the original images and their corresponding annotations.
- The file naming convention includes:
  - Participant numbers: **01–14**.
  - Eye identifiers: **L (Left)** or **R (Right)**.
  - Ground truth sources: **1stHO** or **2ndHO**.

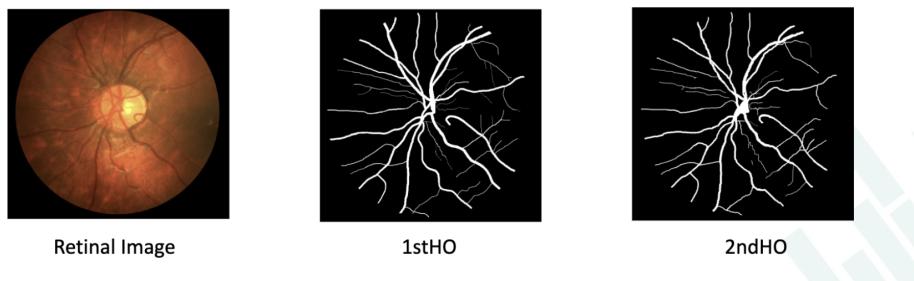


Figure 1: Overview of the CHASEDB1 Dataset: (Left) Original retinal image, (Center) Ground truth by 1st observer (1stHO), and (Right) Ground truth by 2nd observer (2ndHO).

## 1 Problem Statement

The primary objective of this project is to develop an accurate and robust system for the segmentation of blood vessels in retinal images. Retinal vessel segmentation plays a vital role in medical image analysis, as it helps in diagnosing and monitoring various ophthalmological and cardiovascular diseases.

This segmentation task is crucial because:

- Analyzing retinal blood vessels enables the diagnosis and tracking of diseases such as diabetic retinopathy and glaucoma.
- Accurate segmentation allows clinicians to detect issues early, facilitating better treatment planning.

## 2 Challenges Faced

Despite its importance, retinal vessel segmentation poses significant challenges due to:

- Variability in the appearance of vessels across images.
- Imbalanced data, where thinner vessels are underrepresented compared to larger ones.

## 3 Methodology

### 3.1 Image Acquisition:

- **Objective:** Retrieve high-resolution retinal images captured in the RGB color space through medical imaging devices.
- Image acquisition involves obtaining pixel values from the captured retinal images in the RGB color space.

### 3.2 Green Channel Isolation:

- **Objective:** Separate individual color channels—Red, Green, and Blue—and prioritize the Green channel for its efficacy in highlighting retinal vessel structures.
- Utilizes RGB decomposition to extract the Green channel:  $G(x, y) = \text{Green channel intensity at } (x, y)$ .

### 3.3 Contrast Enhancement: CLAHE, Histogram Equalization, Gamma Correction:

- **CLAHE (Contrast-Limited Adaptive Histogram Equalization):**
  - **Objective:** Enhance local contrast while preventing over-amplification in specific regions.
  - For each pixel  $(x, y)$  in the image, CLAHE computes a transformation function based on the histogram of pixel intensities in the local neighborhood.
- **Histogram Equalization:**
  - **Objective:** Adjust overall image contrast by redistributing pixel intensities.
  - The transformation function is given by

$$T(I) = \text{round} \left( \frac{\text{CDF}(I) - \text{min\_CDF}}{\text{max\_CDF} - \text{min\_CDF}} \right) \times (\text{num\_intensities} - 1) \quad (1)$$

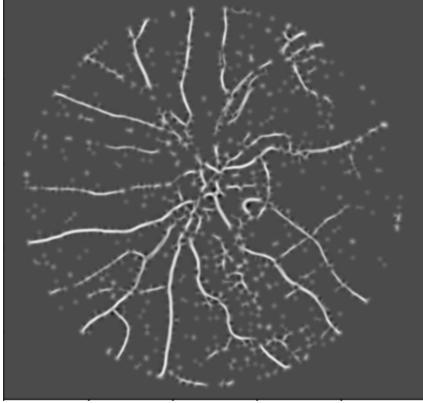
where num\_intensities is the total number of possible intensity levels in the image and the Cumulative Distribution Function (CDF) is calculated from the image histogram. It represents the cumulative probability of pixel intensities up to a certain level. The transformation function uses the CDF to normalize pixel intensities, ensuring that the output values cover the entire intensity range of the image.

- **Gamma Correction:**
  - **Objective:** Adjust image contrast by modifying the gamma value.
  - The gamma correction operation is defined as  $I_{\text{out}} = I_{\text{in}}^{\gamma}$ , where  $I_{\text{out}}$  is the output intensity,  $I_{\text{in}}$  is the input intensity, and  $\gamma$  is the gamma value.

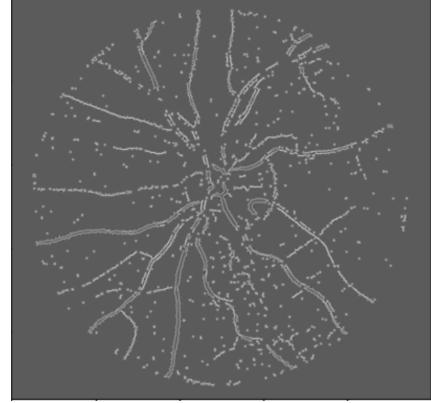
### 3.4 Hessian Matrix Computation:

- **Objective:** Analyze local geometric properties to identify vessel structures.
- Given an image  $I(x, y)$ , the Hessian matrix  $H$  is calculated using the second-order partial derivatives of intensity:  $H = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix}$ .
- We have used an innovative method to enhance images that highlight wide and thin vessels. By applying a morphological filter, the Hessian matrix and eigenvalue transformations, we were able to successfully separate varying-width vessels.

- We computed the second derivative of the image at different scales to focus on either wide or thin vessels. This method utilizes a Hessian matrix and eigenvalue-based approach to isolate the vessels with diverse widths. By leveraging the eigenvalues obtained from the Hessian matrix and their differences, we further improved contrast and reduced non-vasculature structures in the images.



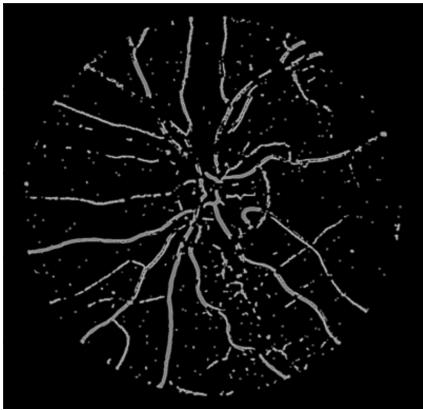
(a) Wide Vessels Enhanced Image



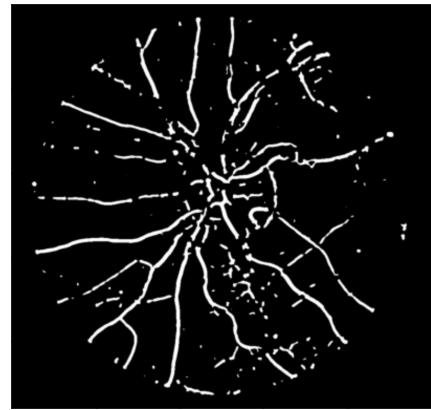
(b) Thin Vessels Enhanced Image

### 3.5 Otsu's Thresholding:

- **Objective:** Automatically determine an optimal threshold for image segmentation.
- Otsu's method maximizes the variance between two classes (foreground and background) by iteratively selecting a threshold value.
- We have modified Otsu's method to specifically target noise and geometric objects based on vessel structure.
- Traditionally, Otsu's approach determines pixel classification thresholds globally or locally across images. However, applying it directly to the whole image can be ineffective. To address this, we applied it separately to wide and thin vessel images. By applying a global threshold on enhanced wide vessels and merging the result with thin vessel enhancements, we were able to amplify both thin and wide vessels, increasing their visibility. This unified image was then subjected to localized thresholding using vessel-based thresholds tied to vessel location, which further refined the thresholds. To better handle noise near wide vessels, we adjusted the global threshold with an offset. Areas away from wide vessels used a lower threshold, which was achieved by subtracting an offset. This fine-tuning aided in extracting smaller vessels from backgrounds with lower intensity, thereby improving the accuracy of the results.



(a) Local Otsu Threshold Image



(b) Global Otsu Threshold Image

## 4 Hypothesis testing and Results

### 4.1 Null Hypotheses

- **Null Hypothesis 1:** Choosing the green channel for the image over the grayscale image does not improve contrast and visibility of retinal vessels, and segmentation accuracy remains unchanged.
- **Null Hypothesis 2:** Applying CLAHE does not improve contrast and visibility of retinal vessels, and segmentation accuracy remains unchanged.

### 4.2 Experiments

To validate the above hypotheses, we will conduct the following experiments:

- **Experiment 1:** Perform segmentation on retinal images using the green channel and grayscale images, measuring segmentation accuracy, structural similarity index (SSIM), F1 score, and peak signal-to-noise ratio (PSNR) in both cases.
- **Experiment 2:** Perform segmentation on retinal images with and without CLAHE applied, measuring segmentation accuracy, structural similarity index (SSIM), F1 score, and peak signal-to-noise ratio (PSNR) in both cases.

### 4.3 Experiment Results

#### Experiment 1



We have taken a subset of the dataset to experiment for our hypothesis. The flow of the operations are - Grayscale/Green Channel selection, CLAHE, Morphological Operations, Hessian Widening and then Thresholding.

Evaluation Metrics	Grayscale#1	Green Channel#1	Grayscale#2	Green Channel#2
Accuracy	0.898	0.9139	0.9139	0.925
SSIM	0.6945	0.7314	0.7384	0.7639
PSNR	9.6262	10.298	10.346	10.8388
F1	0.5164	0.5782	0.6084	0.6632

(a) Experiment 1 Results

#### Experiment 2



Segmentation results with and without applying CLAHE on the retinal images.

Evaluation Metrics	Without CLAHE#1	With CLAHE#1	Without CLAHE#2	With CLAHE#2
Accuracy	0.9288	0.9139	0.9346	0.925
SSIM	0.7901	0.7314	0.8096	0.7639
PSNR	11.3494	10.298	11.7395	10.8388
F1	0.6498	0.5782	0.7217	0.6632

(b) Experiment 2 Results

### Overall Insights

- The green channel is more effective than grayscale for retinal vessel segmentation, improving both visual and structural metrics.
- CLAHE, while useful for visual enhancement, does not consistently improve segmentation performance and may need further fine-tuning for specific datasets or tasks.

## 5 Results

### Comparing Results



Dataset	Accuracy	Sensitivity	Specificity	F1	SSIM	PSNR
Original Dataset	0.9634	0.6191	<b>0.9903</b>	0.7543	0.8721	62.5135
Scaled Dataset (50%)	0.9678	0.6643	0.9901	0.8354	0.8743	62.8732
Scaled Dataset	<b>0.9703</b>	<b>0.7647</b>	0.9869	<b>0.8614</b>	<b>0.8884</b>	<b>63.4093</b>

Figure 5: Results

Scaling techniques improve the segmentation model's performance across most metrics. The scaled dataset achieves **higher accuracy (0.9703)**, significantly **better sensitivity (0.7647)**, and **improved F1-score (0.8614)**, indicating better vessel detection and overall performance. While **specificity slightly decreases (0.9903 → 0.9869)**, the **gains in sensitivity and structural metrics like SSIM (0.8884) and PSNR (63.4093)** make scaling a valuable preprocessing step for enhancing segmentation quality.

## Conclusion

- Among the pre-processing techniques evaluated, random brightness scaling combined with Green Channel isolation and CLAHE filtering demonstrated notable improvements in key metrics, confirming its effectiveness in handling intensity variations in the dataset.
- The scaled dataset achieved superior performance, particularly in F1 score and sensitivity, indicating the value of effective pre-processing in semantic segmentation tasks.
- U-Net demonstrates better performance after scaling the dataset.

## 6 Future Work

Future work could focus on exploring additional augmentation techniques and advanced architectures to further refine segmentation accuracy and robustness. We can explore other randomized scaling techniques to make our dataset even more scale invariant.

### Note

We have listed other relevant details that will help you understand and get a better insights of this project in the Annexure section at the end.

## Code Availability

### 6.1 Code

The code used for this project is publicly available on GitHub from the following link:

[https://github.com/parthivdholaria/Retinal-Image-Segmentation/  
tree/main](https://github.com/parthivdholaria/Retinal-Image-Segmentation/tree/main)

### 6.2 Dataset Source

The dataset can be downloaded from:

<https://www.kaggle.com/datasets/khoongweihao/chasedb1>

## References

1. Almotiri, J., Elleithy, K., & Elleithy, A. (2018). Retinal Vessels Segmentation Techniques and Algorithms: A survey. *Applied Sciences*, 8(2), 155. <https://doi.org/10.3390/app8020155>
2. GeethaRamani, R., & Balasubramanian, L. (2016). Retinal blood vessel segmentation employing image processing and data mining techniques for computerized retinal image analysis. *Biocybernetics and Biomedical Engineering*, 36(1), 102-118. <https://doi.org/10.1016/j.bbe.2015.06.004>
3. Khan, K. B., Khaliq, A. A., & Shahid, M. (2016). Correction: A morphological Hessian based approach for retinal blood vessels segmentation and denoising using region based OTSU thresholding. *PLOS ONE*, 11(9), e0162581. <https://doi.org/10.1371/journal.pone.0162581>
4. Maison, T., Lestari, T., & Luthfi, A. (2019). Retinal Blood Vessel Segmentation using Gaussian Filter. In *Journal of Physics: Conference Series*, **1376**, 012023. IOP Publishing Ltd. <https://doi.org/10.1088/1742-6596/1376/1/012023>
5. Shabani, M. (2017). Master thesis: Retinal Blood Vessel Extraction Based on a Combination of Matched Filter and Level Set Algorithm. <https://doi.org/10.13140/RG.2.2.12055.85929/1>
6. Soomro, T. A., Ali, A., Jandan, N. A., Afifi, A. J., Irfan, M., Alqhtani, S. M., Głowacz, A., Alqahtani, A., Tadeusiewicz, R., Kańtoch, E., & Zheng, L. (2021). Impact of novel image preprocessing techniques on retinal vessel segmentation. *Electronics*, 10(18), 2297. <https://doi.org/10.3390/electronics10182297>
7. Srinidhi, C. L., Aparna, P., & Rajan, J. (2017). Recent advancements in retinal vessel segmentation. *Journal of Medical Systems*, 41(4). <https://doi.org/10.1007/s10916-017-0719-2>

## Annexure

### 6.3 Exploring and Analyzing Different Color Channels

In the initial stages of our retinal vessel segmentation process, we begin by dividing the original image into its constituent channels: Red, Green, and Blue (RGB). This is a standard procedure in image processing, as it allows us to isolate and manipulate individual color components separately, which can be crucial in tasks such as segmentation where color information can be a key differentiator. However, for our specific task of retinal vessel segmentation, we focus on the Green channel. This is due to the fact that the Green channel has been found to exhibit superior contrast between the vessels and the background compared to the other channels. This contrast is beneficial as it enhances the visibility of the vessels, making them easier to detect and segment.

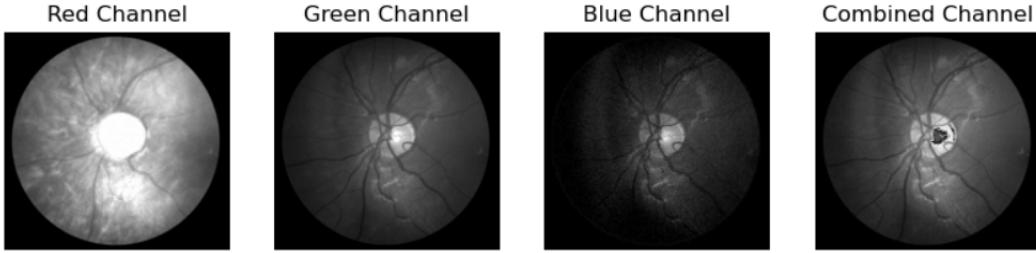


Figure 6: The different channel confirms our hypothesis that the green channel exhibits better contrast between the vessels and the background.

### 6.4 Contrast Enhancement Techniques and Analysis

The next step in our preprocessing pipeline is contrast enhancement. This is a crucial step as it significantly impacts the visibility of the vessels in the image, making the subsequent segmentation process more effective. There are several techniques used for contrast enhancement, including Gamma Correction, Histogram Equalization, and Contrast Limited Adaptive Histogram Equalization (CLAHE).

Gamma Correction is a technique used to enhance the contrast of an image by modifying its gamma value, which is a measure of the nonlinear chromaticity of a display device. Histogram Equalization is a method in image processing of contrast adjustment using the image's histogram. It transforms the intensity values so that the histogram of the output image approximately matches a specified histogram. CLAHE, on the other hand, is an advanced normalization technique that can handle a wide range of contrast in an image.

We first apply Gamma Correction to our Green channel image to generate a gamma corrected image. We then compare the Green channel image and the gamma corrected image by plotting their histograms and making relevant observations.

Next, we attempt to equalize the histogram of both images. However, we observe that this process adds a significant amount of noise to the image, which we found to be detrimental to our segmentation process.

Finally, we apply CLAHE to both the Green channel image and the gamma corrected image. To our surprise, we found that applying CLAHE to the Green channel image produced better results than applying it to the gamma corrected image. This suggests that the Green channel, even after gamma correction, might not be the optimal image for CLAHE.

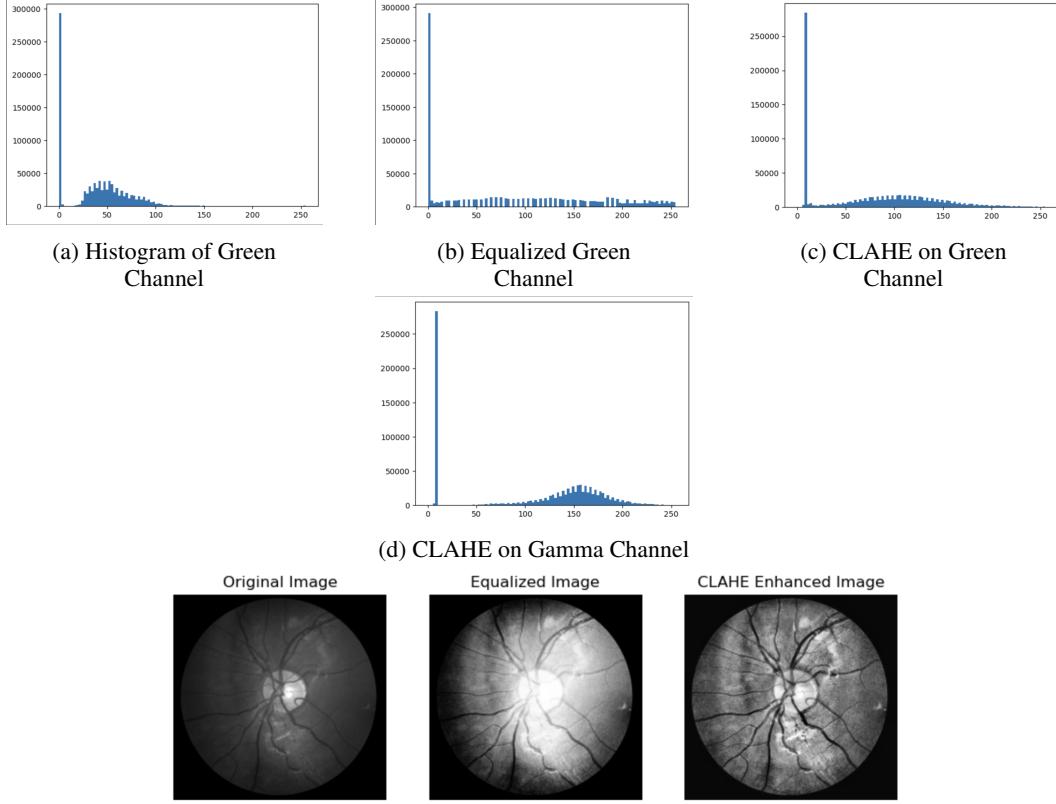


Figure 7: Comparison between different Contrast Enhancement Techniques

Gamma correction is often used to correct for variations in image illumination. However, in this case, the green channel alone already provides sufficient and better contrast for vessel segmentation, thereby eliminating the need for using gamma correction (See Fig. 3).

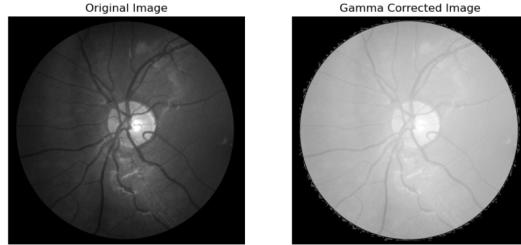


Figure 8: Original image vs Gamma corrected image

## 6.5 Background Homogenization

The next step in our pipeline involves the process of background homogenization. Background homogenization, also known as uneven illumination removal, is a critical step in image processing that aims to improve the visibility of objects within an image by reducing variations in lighting. This is particularly important for our task of retinal vessel segmentation, as uneven illumination can obscure the details of the vessels. We will explore different morphological operations in this part.

### 6.5.1 Morphological Operations

Morphological operations are mathematical operations based on the shape of objects within an image. In our pipeline, we apply the top-hat, bottom-hat, and their subtraction to the preprocessed Green channel image.

The top-hat transform is an operation that extracts small elements and details from given images. It is defined as the difference between the input image and its opening by some structuring element. Similarly, a bottom-hat transform is defined as the difference between the closing and the input image.

In mathematical terms, the top-hat transform can be represented as:

$$\text{Tophat}(f) = f - (f \circ b)^{\text{open}} \quad (2)$$

where  $(I)$  is the input image and  $(\text{Open}(I))$  is the opening of the input image.

Similarly, the bottom-hat transform can be represented as:

$$\text{Bottomhat}(f) = (f \circ b)^{\text{close}} - f \quad (3)$$

where  $(\text{Close}(I))$  is the closing of the input image.

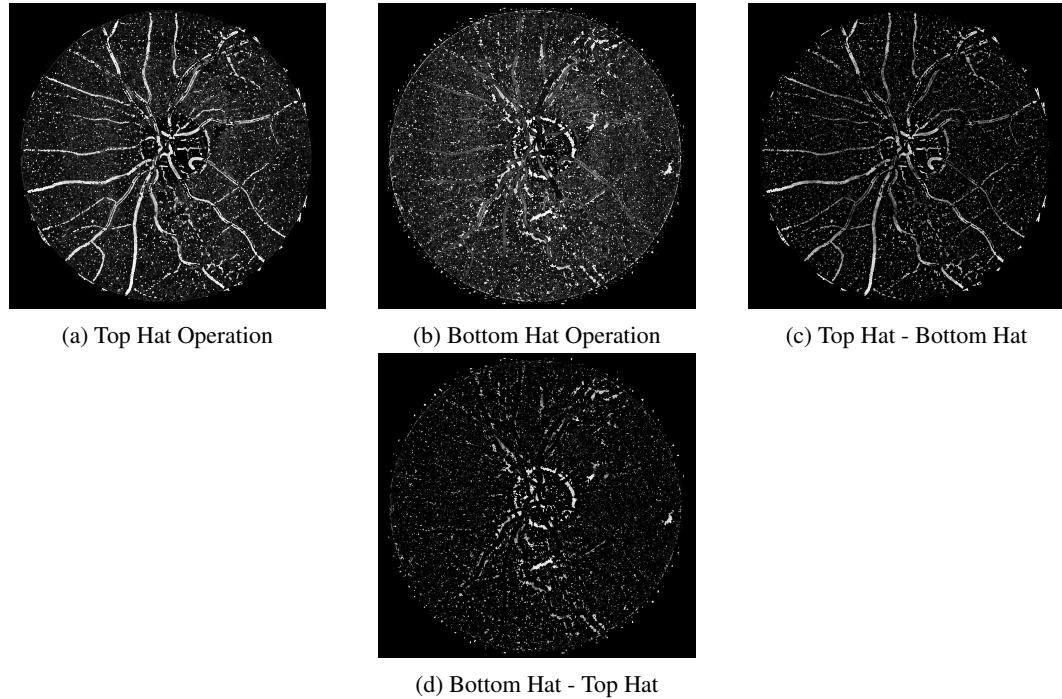


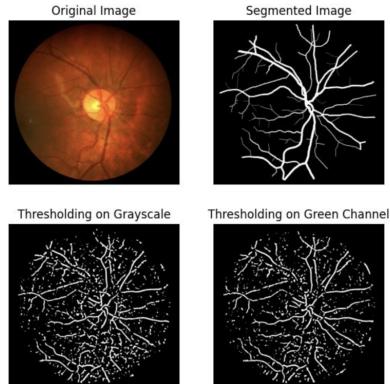
Figure 9: Different Morphological Operations

## 6.6 Hypothesis Qualitative Analysis

### Experiment 1

One can clearly see choosing the green channel over the grayscale image reduces noise leading to better segmentation results and accuracy. This is also quite evident from slide 4 where we can visually also verify that green channel shows less noise in comparison to the grayscale.

Image: Image\_01R.jpg

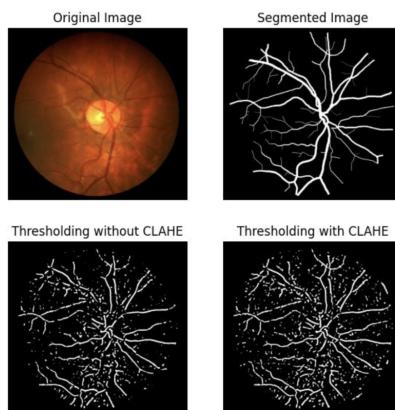


(a) Experiment 1 Results

### Experiment 2

Here we have kept the consistent green channel for both cases; with or without CLAHE. One can see without CLAHE we have less noise leading to better segmentation results and accuracy. This is actually a little contradictory to the visual inspection of the both images, and to the general convention followed for retinal vessel segmentation. One reason for this could be the preprocessing steps for supportive towards without CLAHE ones.

Image: Image\_01R.jpg



(b) Experiment 2 Results

## 7 Model Used

### Model Description:

U-Net is a convolutional neural network primarily designed for image segmentation. Its hallmark “U” shape is formed by a symmetric encoder and decoder path connected via skip connections. On the encoder (left) side, the architecture applies repeated  $3 \times 3$  convolutions followed by ReLU activations, then performs  $2 \times 2$  max pooling to halve the spatial resolution at each step. Concurrently, the network doubles the number of feature channels as it progresses downward, allowing it to capture increasingly complex features at reduced resolution.

At the bottom (bottleneck), U-Net has the fewest spatial dimensions and the largest number of channels, extracting abstract, high-level feature representations. On the decoder (right) side, each step begins with a  $2 \times 2$  transposed convolution (up-convolution) that doubles the spatial dimensions while halving the number of channels. Skip connections then merge the upsampled features with the correspondingly cropped feature maps from the encoder. This “copy and crop” step preserves the fine-grained spatial information lost during max pooling, ensuring that boundary details are regained in the segmentation.

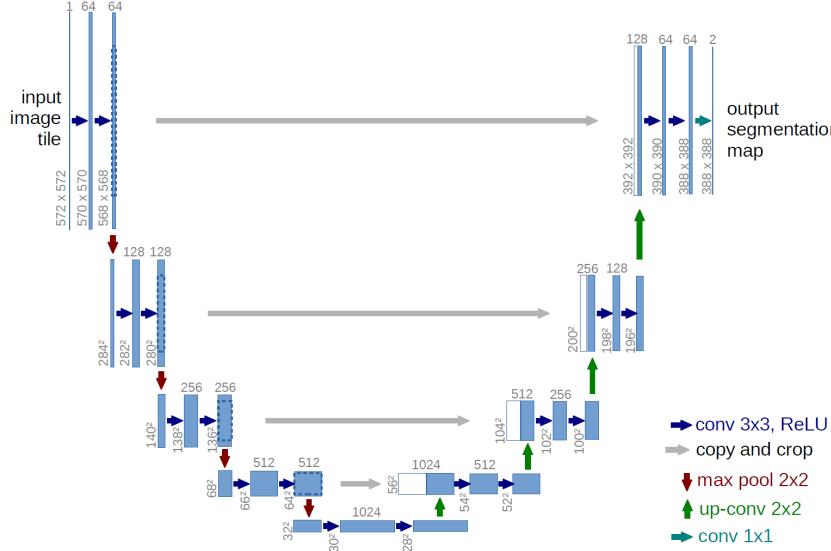


Figure 11: U-Net Architecture

Each merging is followed by two additional  $3 \times 3$  convolutions and ReLU activations, refining the fused features at the higher resolution. This process continues until the feature maps return to the input scale. Finally, a  $1 \times 1$  convolution converts the feature maps to a segmentation map with one output channel per class (commonly two or more channels).

## Randomized Scaling Techniques

### Random Brightness Scaling

- **Description:**
  - This technique adjusted the brightness of the image by multiplying every pixel value in the green channel of the retinal image with a randomly sampled scaling factor.
  - The scaling factor was sampled uniformly from the range  $[0.80, 1.6]$ , ensuring that the brightness could both increase and decrease, simulating various lighting conditions.
  - After applying the scaling factor, the pixel values were clipped to lie within the valid range  $[0, 255]$  to avoid overflows or negative values.
- **Rationale:**
  - This technique ensures that the model is exposed to images with varying brightness levels, improving its generalization capabilities.
  - It helps simulate real-world scenarios where the retinal images might be captured under different lighting conditions.

### Random Contrasting using CLAHE

- **Description:**
  - CLAHE (Contrast Limited Adaptive Histogram Equalization) was applied to the green channel of the retinal image.
  - The clip limit, which controls the contrast amplification, was chosen randomly for each image from a uniform range of  $[2, 5]$ . This variation introduced randomness in how much contrast enhancement was applied to each image.
  - A grid size of  $8 \times 8$  was used to divide the image into smaller tiles, allowing local contrast adjustments to enhance vessel details.
  - This process enhanced the visibility of fine details, particularly the blood vessels, while preventing over-saturation by limiting the contrast amplification.

- **Rationale:**

- By applying a random clip limit, the technique ensured the model would not rely on a fixed contrast level, making it more robust to variations in image contrast due to differing capture conditions.
- It also enhanced the visibility of retinal vessels, which are critical for accurate segmentation.

### Random Gaussian Noise Addition

- **Description:**

- Gaussian noise was added to the green channel of the retinal image to simulate sensor noise or real-world imperfections.
- The noise was generated using a normal distribution  $\mathcal{N}(\mu, \sigma^2)$ , where the mean  $\mu$  was fixed at 0, and the standard deviation  $\sigma$  was randomly sampled from the range [10, 30].
- The generated noise was added to the pixel intensities of the image, and the resulting pixel values were clipped to the valid range of [0, 255].
- This process introduced a controlled level of randomness, mimicking noisy conditions in imaging systems.

- **Rationale:**

- This technique was applied to evaluate whether CLAHE and the overall segmentation pipeline were resistant to noise.
- It also aimed to improve the model's robustness to noisy input images, ensuring reliable segmentation even in the presence of noise.

## 7.1 Results from random\_brightness\_scaling

### Randomized scaling Techniques : Results



- We randomly sampled 10 images and analysed Green Channel isolation + CLAHE filtering on grey scale images with and without random brightness scaling.

Setup\Metrics	accuracy	Sensitivity	Specificity	F1	SSIM	PSNR
Without scaling	0.9114	0.3777	0.9698	0.5420	0.7239	10.1236
With scaling	0.9163	0.3983	0.9707	0.5626	0.7355	10.2677

Figure 12: random\_brightness\_scaling

## 7.2 Results from random\_CLAHE\_Clipping\_scaling

### Randomized scaling Techniques : Results



- We randomly sampled 10 images and analysed Green Channel isolation + CLAHE filtering on grey scale images with and without random CLAHE clipping.

Setup\Metrics	accuracy	Sensitivity	Specificity	F1	SSIM	PSNR
Without clipping	0.9280	0.5071	0.9614	0.6607	0.7971	11.4131
With clipping	0.9145	0.4352	0.9641	0.5958	0.7404	10.3739

Figure 13: random\_CLAHE\_Clipping\_scaling

## 7.3 Inference from Scaling

- We ran the inference test for 10 randomly selected images and averaged out the result. For Random Brightness scaling : F1 improved by 2% , rest all metrics (sensitivity,specificity,PSNR,SSIM) showed a slight improvement.
- CLAHE filtering by Randomly clipping shows a degradation in F1 as well as all the other metrics.
- Hence , Green Channel isolated followed by random factor multiplication and CLAHE is resistant to intensity scaling and is the best method to move forward with.
- CLAHE filtering with random clipping , Addition of Random Gaussian Noise , Intensity normalisation do not perform well for segmentation.

## 8 Evaluation Strategy

Evaluating the performance of retinal vessel image segmentation is a crucial step in assessing the accuracy and effectiveness of the segmentation algorithms. Two commonly used metrics for this purpose are the Structural Similarity Index (SSIM) measure and the Peak Signal-to-Noise Ratio (PSNR).

### 1. Structural Similarity Index (SSI):

- Objective:** Quantitatively assess the similarity between segmented and ground truth images.
- $$\text{SSI} = \frac{2 \cdot \mu_x \cdot \mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \cdot \frac{2 \cdot \sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}$$
, where  $\mu_x$ ,  $\mu_y$  are means,  $\sigma_x$ ,  $\sigma_y$  are variances,  $\sigma_{xy}$  is the covariance, and  $c_1$ ,  $c_2$  are constants.
- This metric provides a comprehensive assessment of the similarity between the segmented vessels and the ground truth. SSIM considers luminance, contrast, and structure in its evaluation, making it suitable for assessing image quality. It produces a value between -1 and 1, with 1 indicating perfect similarity. In the context of retinal vessel segmentation, a higher SSIM score, such as the achieved value of **0.87**, signifies better accuracy and fidelity of the segmented vessels.

### 2. Peak Signal-to-Noise Ratio (PSNR):

- Objective:** Measure the quality of the segmented image by evaluating the ratio of signal power to noise.
- $$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{Max intensity value}^2}{\text{Mean squared error}} \right)$$
.
- PSNR is a widely used metric to quantify the quality of an image by measuring the ratio between the maximum possible power of a signal and the power of corrupting

noise. In retinal vessel segmentation, a higher PSNR indicates a clearer and more accurate segmentation result, as it reflects the quality of the segmented vessels concerning the original image. The obtained PSNR value of **55** further affirms the quality and accuracy of the segmented vessel structures, providing insights into the preservation of vessel details during the segmentation process.

### Note

We also measured additional metrics such as:

- **Accuracy** – proportion of correct predictions across all classes.
- **Sensitivity (Recall)** – ability to correctly identify positive instances.
- **Specificity** – ability to correctly identify negative instances.
- **F1 Score** – harmonic mean of precision and recall, balancing both.
- Since we used only a portion of the images in 50% of the dataset. The running time for each epoch was reduced with the same proportion. The preprocessing scaling was separately done before training the model.