Imperial College London

SURG70004

IGI COURSEWORK 1

IMAGE GUIDED INTERVENTION

Module Leader: Stamatia (Matina) Giannarou

Student Name: Zhiling (Gavin) Yue
CID Number: 02469416

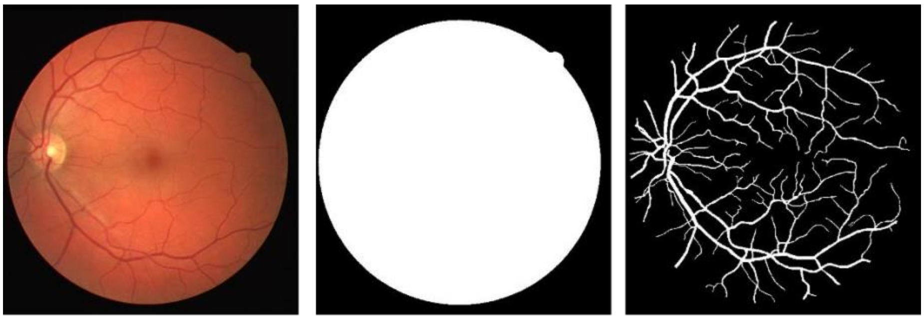
Coursework start date: 31/10/2023
Coursework end date: 8/11/2023

MRes Medical Robotics & Image Guided Intervention
The Hamlyn Centre
Imperial College London

SURG70004 Image Guided Intervention
IGI Coursework 1

Due date Wednesday 8th November 2023 (via Blackboard)

During this exercise you will investigate the detection of edges, image filtering and segmentation of blood vessels on retinal images. The retinal image dataset used for this coursework is part of the DRIVE database [1] and it includes a retinal image ("Retinal Image.tif" file), its corresponding mask image delineating its FOV ("Mask Image.gif" file) and a ground truth segmentation map for the image ("Ground Truth.gif" file) as shown in the figure below. For the processing of the retinal image, only the part highlighted as white in the mask image needs to be considered.



Retina Image　　Mask Image　　Ground Truth Map

Figure: Images included in the dataset.

The questions are provided in black font. Please use blue for your answers in the spaces indicated.

Task 1 – Edge Detection (30%)

(a) A binary image contains straight lines oriented horizontally, vertically, and diagonally at 45° and -45°. Give a set of 3x3 kernels which can be used to detect one pixel breaks in these lines. Assume that the lines are one pixel thick, they do not intersect and that the intensities of the pixels belonging to the lines and the background are 1 and 0, respectively. (max. 100 words)

(a) The kernels shown below [Figure 1a] can be used to detect lines at various orientations:

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix}$$

Horizontal　　Vertical　　Diagonal at 45°　　Diagonal at -45°

***Figure 1a**: Four kernels which respond maximally to horizontal, vertical and diagonal at 45° and -45° single-pixel wide lines.*

(b) Use the Laplacian of Gaussian operator to detect the edges of the provided skin lesion image. Investigate how the size of the LOG kernel affects the edge detection result. (max. 100 words)

(b) The edge detections using various kernel sizes are shown in [Figure 1b].

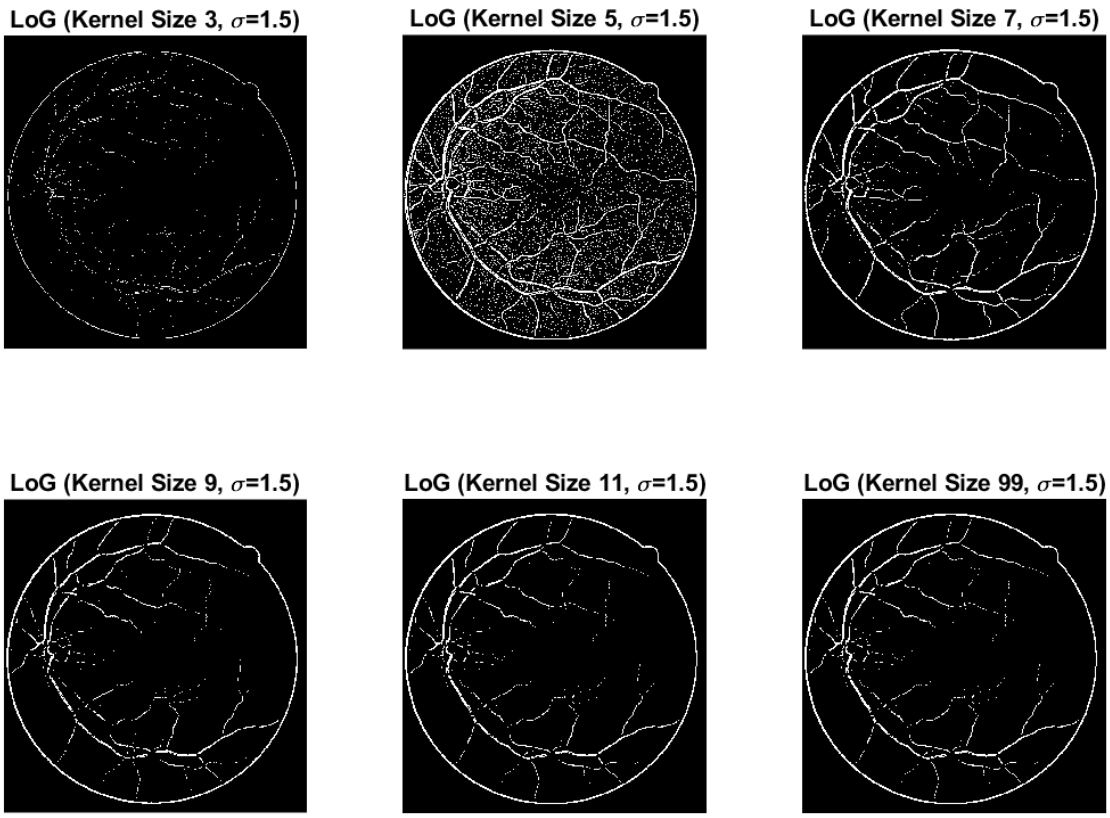


Figure 1b*: Edge detection based on different sizes of LoG kernel.*

To summarise, a smaller kernel size heightens sensitivity to finer edges within the image (e.g. KZ = 7). This sensitivity, however, may also intensify noise, resulting in a noisier edge map (KZ = 5). When the kernel size is too small concerning the σ value, severe truncation can occur (KZ =3). On the contrary, larger kernel sizes capture broader edges and are more robust to noise but might overlook finer image details.

Task 2 – Image Filtering (40%)

(a) Explain how the Fourier transform can be used to apply on a 2D image a high pass filter in the frequency domain. (max. 100 words)

(a) **First step: Convert the 2D image to the frequency domain**
Fourier transform converts a function into an alternative representation. In this case, it converts the spatial information of the image into its frequency representation.
Second step: Design and apply the high-pass filter
Create a filter to eliminate low frequencies and let high frequencies pass to enhance it (e.g. the edge-detection filter). Then multiply the frequency spectrum by the high-pass filter.
Third step: Inverse Fourier transform for reconstruction
Convert the modified spectrum back to the spatial domain of the image.

(b) Consider a filter defined in the frequency domain by the function $H(u, v)$ which is zero at the centre of the (centred) transform and 1 elsewhere. Explain what effect this filtering process will have on the intensities of the image. (max. 100 words)

(b) This filtering process performs as a high-pass filter, which enhances high-frequency components while reducing low-frequency details. It enhances edges, fine details, and texture, resulting in heightened contrast and sharper features.
In summary, this filter will improve the overall sharpness and edge details of the image.

(c) The Discrete Fourier Transform (DFT) is complex, and it can be expressed in polar form as:

$$F(u, v) = |F(u, v)|e^{j\varphi(u,v)}$$

where, $|F(u, v)|$ and $\varphi(u, v)$ are the magnitude and phase, respectively.
Prove mathematically that the DFT of a real $f(x, y)$, in which (x, y) is replaced by $(-x, -y)$, is $F^*(u, v)$, the complex conjugate of the DFT of $f(x, y)$. This property is mathematically described as follows:

$$f(-x, -y) \Leftrightarrow F^*(u, v)$$

where, the symbol $\Leftrightarrow$ indicates a Fourier transform pair.

(c) The DFT expressed in polar form:

$$F(u, v) = |F(u, v)|e^{\,j\varphi(u,v)},$$

where $|F(u, v)|$ is the modulus of the complex

As a property of complex Z, we can get:

$$Z^* = \frac{|Z|^2}{Z}$$

Then, the complex conjugate of $F(u, v)$ is:

$$F^*(u, v) = \frac{|F(u,v)|^2}{F(u,v)} = |F(u, v)|e^{\,-j\varphi(u,v)} \qquad \textit{Equation 1}$$

The common complex number and its modulus are given by:

$$Z = x + yi = re^{i\theta},$$

$$r = |x + iy| = \sqrt[2]{x^2 + y^2},$$

where r is the modulus and x is the real part

The modulus (or magnitudes) of $Z(-x, -y)$ and $Z(x, y)$ are equivalent:

$$r = \sqrt[2]{(-x)^2 + (-y)^2} = \sqrt[2]{x^2 + y^2}$$

Thus, for the modulus of DFT

$$|F(u, v)| = |F(-u, -v)|$$

Considering $F(-u, -v)$:

$$F(-u, -v) = |F(-u, -v)|e^{j\varphi(-u,-v)}$$
$$= |F(u, v)|e^{j\varphi(-u,-v)}$$

Now trying to prove $e^{j\varphi(-u,-v)} = e^{-j\varphi(u,v)}$

The 2D DFT can be expressed as (from lecture notes):

$$F(u, v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x, y)\, e^{-2\pi j\left(\frac{ux}{M}+\frac{vy}{N}\right)} = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x, y)\, e^{2\pi j\left(-\frac{ux}{M}-\frac{vy}{N}\right)}$$

Thus, $\varphi(u, v)$ is an odd function:

$$\varphi(-u, -v) = -\varphi(u, v)$$

Now, considering $F(-u, -v)$ again:

$$F(-u, -v) = |F(u, v)|e^{j\varphi(-u,-v)}$$
$$= |F(u, v)|e^{-j\varphi(u,v)} \qquad \textit{Equation 2}$$

Comparing Equation 1 and 2, which are $F(-u, -v)$ and $F^*(u, v)$:

$$F(-u, -v) = F^*(u, v)$$

Using Fourier transform pair:

$$f(-x, -y) \Leftrightarrow F(-u, -v)$$

Thus, we can get:

$$f(-x, -y) \Leftrightarrow F^*(u, v)$$

(d) Consider a binary image which illustrates the digit “5” in the centre. If you compute the DFT of this image, multiply the phase of this transform by -1 and compute the inverse DFT, you will obtain a transformed image. Explain the effect of this operation on the original image and sketch the transformed image.

(d) The multiplication by -1 of the phase in the DFT essentially causes a 180-degree phase shift of the original image. The effect on the transformed image will be an alteration of its spatial structure without changing its overall magnitude or shape.

The original and transformed images are displayed in [Figure 2d]. The transformed image is rotated 180 degrees from the original image. Due to the rotational symmetry of the digit '5', the transformed image retains the upside down '5' in the centre.

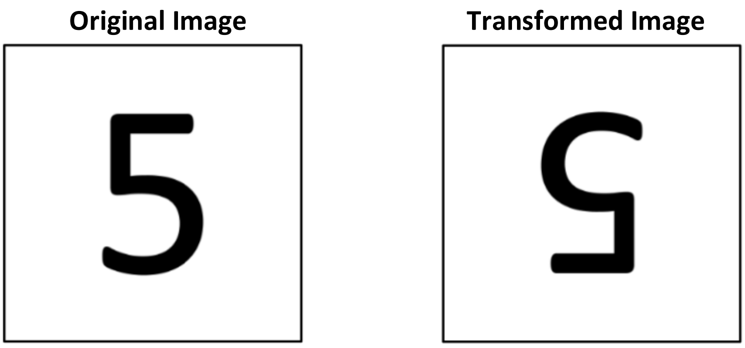


Figure 2d: Original and transformed images.

(e) Corrupt the retinal image by Gaussian noise of zero mean and standard deviation of 0.1. Apply image smoothing using 2D convolution between the retinal image and a gaussian kernel of standard deviation $\sigma = 1$ and size $[(2 * ceil(2 * \sigma) + 1) \times (2 * ceil(2 * \sigma) + 1)]$. Plot the smoothed image. Apply the same filtering process in the frequency domain. Explain the steps you followed. Plot the Fourier representation (magnitude) of the Gaussian filter and of the image. By presenting the result of filtering, show what is the effect of changing the value of σ on the smoothed image? (max. 100 words)

(e) [Figure 2e.1] shows the original, noised and smoothed image. The noised image is made by applying Gaussian noise.

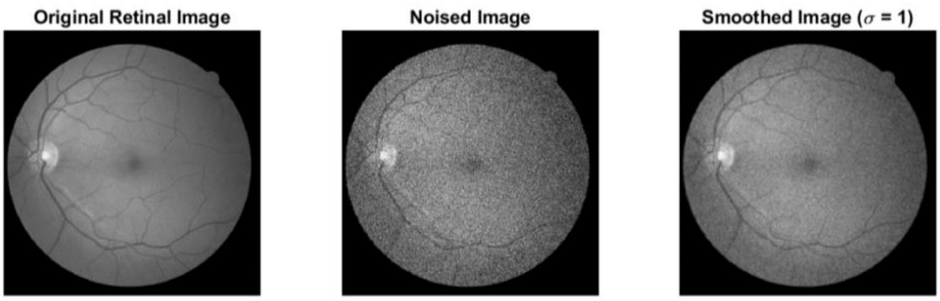


Figure 2e.1: The original, noised and smoothed image.

[Figure 2e.2] shows the Fourier representations respectively.

First, the image is transformed from the time to frequency domain using Fourier Transform (FT). Gaussian noise is introduced in frequency domain. FT converts the Gaussian kernel to the frequency domain, simplifying convolution. Multiplication in frequency domain occurs between the transformed image and kernel. Finally, a logarithmic function enhances image visualisation, and Fourier representation is plotted.

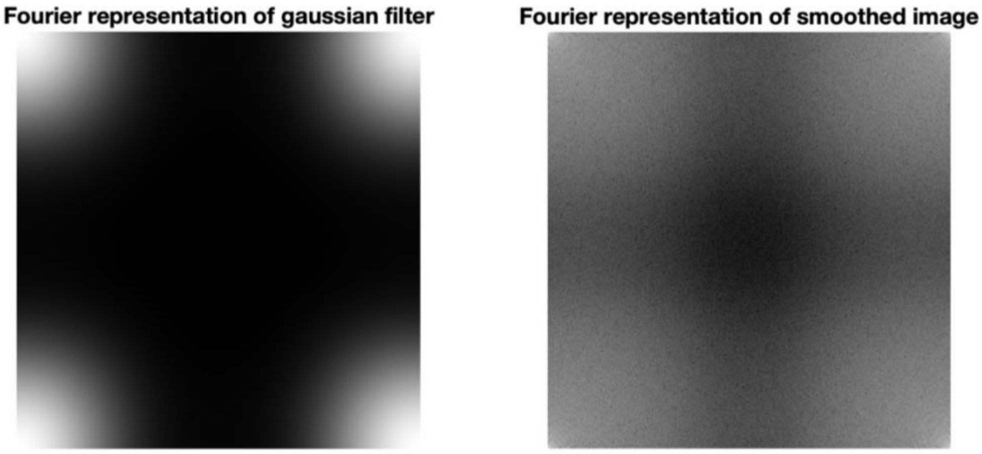


Figure 2e.2: The Fourier representations (in magnitudes).

The smoothed images shown in [Figure 2e.3] are filtered with $\sigma = 1,3,5,7$. Larger σ values result in blurring and smoother image, while smaller σ values detect finer details and features. σ = 1 contains the most details, while σ = 7 is more blurry.

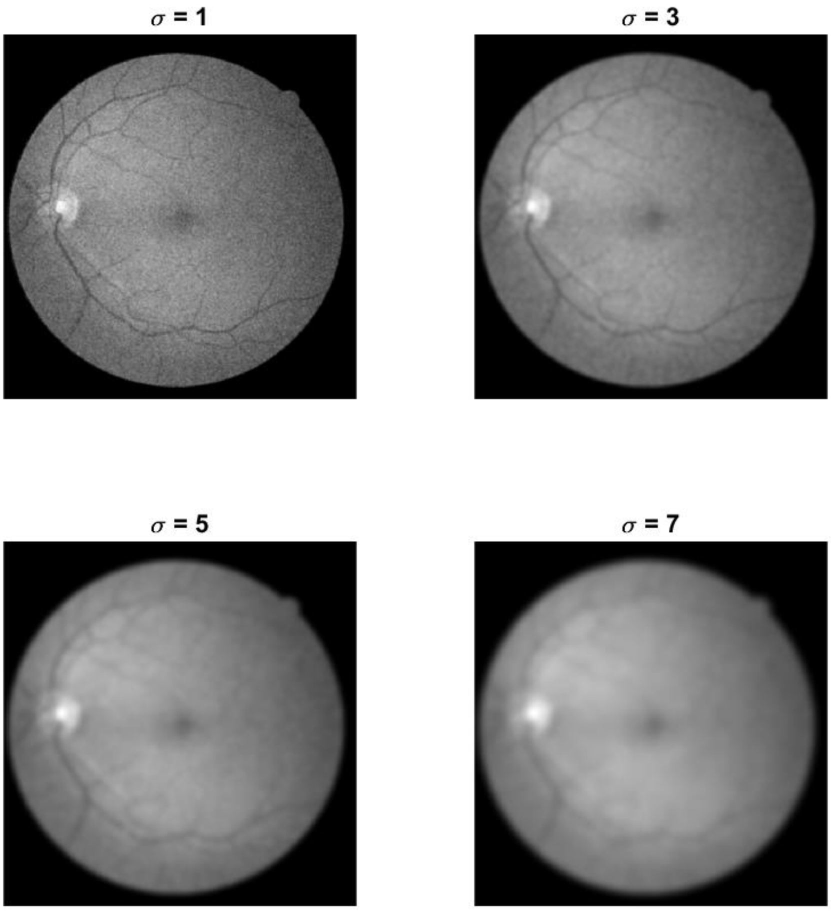


Figure 2e.3: Smoothed image using various values of σ.

Task 3 – Image Segmentation (30%)

(a) Apply threshold-based segmentation to identify blood vessels in the provided retinal image. Explain your threshold selection and investigate how the selection of the threshold affects the accuracy of the segmentation result. You need to compare your results to the provided ground truth segmentation map (max. 100 words).

(a) [Figure 3a] shows the ground truth and segmentations with different thresholds (Thr).
Numerous thresholds ranging from 0 to 1 have been chosen for the segmentation process. By comparing these segmented outputs to the ground truth, three specific thresholds have been selected to demonstrate their impact on accuracy. It is observed that a threshold value of 0.1 yields a more accurate segmentation result.

High thresholds merge regions, losing information, while low ones induce over-segmentation, adding noise. Greater thresholds filter noise but risk losing details, whereas lower ones capture more detail yet include noise. Adaptive thresholds tailored to image specifics are crucial.

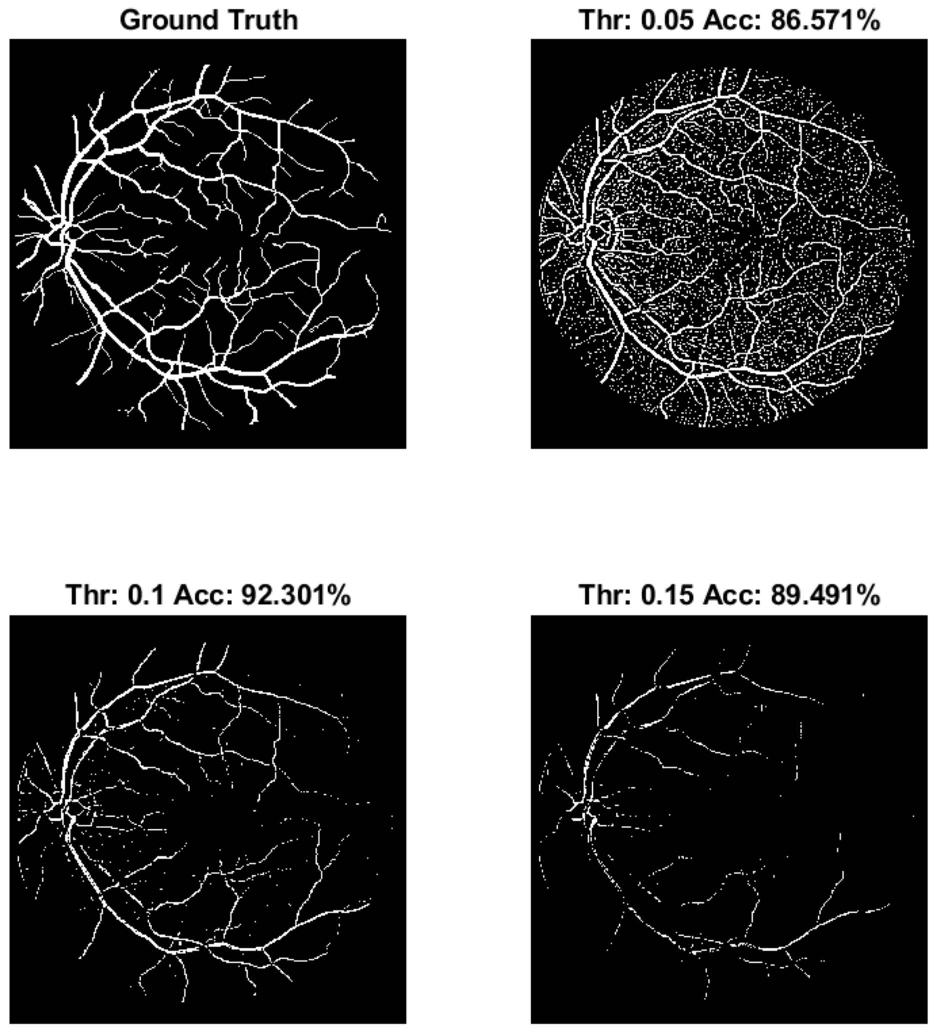


Figure 3a: Threshold-based segmentation and accuracy.

(b) Use ROC analysis to select the optimal threshold value for your segmentation and display the best segmentation map. (max. 100 words).

(b) [Figure 3b.1] shows the ROC analysis. The Area Under the ROC Curve (AUC) has been determined to be 0.7497. Through assessing the distances between the point (0,1) and various points along the curve, the minimum distance of 0.402 has been identified. This particular point, representing the minimum distance, has been highlighted on the graph, signifying the location of the optimal threshold. By leveraging the coordinated False Positive Rate (FPR) and True Positive Rate (TPR) values, the optimal threshold of 0.101 can be precisely identified.

[Figure 3b.2] illustrates the best segmentation map with optimal threshold (0.101).

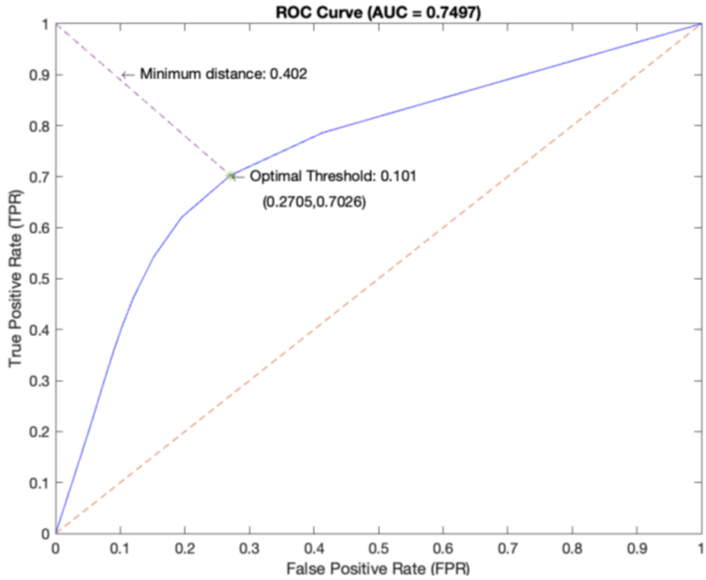


Figure 3b.1: ROC Curve analysis and AUC value.

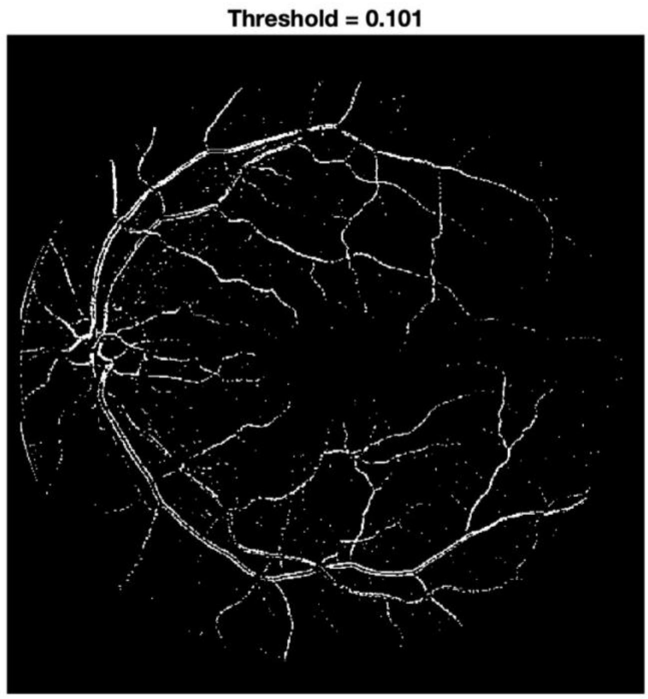


Figure 3b.2: Best segmentation map.