Report on Practical Work: Segmentation of Retinal Vascular Networks

Pierre Billaud – Pascal Mahé

Contents

1 Introduction

This report aims to present the work carried out for Practical Work No. 2 by Pierre Billaud and Pascal Mahé.

2 Responses to Questions

2.1 Can you develop an optimized algorithm to segment a vascular network without machine learning or deep learning?

We were inspired by the article available here: Wiharto, 2019. The description of the method indicated in the article and our results constitute parts 2 to 5 of this report.

2.2 Explain the evaluation function provided in the Python script

The function calculates the performance of our algorithm.

```
def evaluate(img_out, img_GT):
    GT_skel  = thin(img_GT, max_num_iter = 15) # Assuming the
    maximum half
    img_out_skel  = thin(img_out, max_num_iter = 15) #
   thickness of a vessel is 15 pixels...
    TP = np.sum(img_out_skel & img_GT) # True Positives
    FP = np.sum(img_out_skel & ~img_GT) # False Positives
    FN = np.sum(GT_skel & ~img_out) # False Negatives
    ACCU = TP / (TP + FP) # Precision
    RECALL = TP / (TP + FN) # Recall
    return ACCU, RECALL, img_out_skel, GT_skel
```

It starts by normalizing the ground truth image and our result by skeletonizing them. Then, the two images are compared based on the respective pixel values:

- Pixels that are white in both images are true positives (TP)

- Pixels that are white in the generated image but black in reality are false positives (FP)
- Pixels that are black in the generated image but white in reality are false negatives (FN)

Finally, the precision and recall of our algorithm are calculated:

- Precision is the rate of true positives among the white pixels of the generated image: $ACCU = \frac{TP}{TP+FP}$
- Recall is the rate of true positives among all the white pixels in the original image: $RECALL = \frac{TP}{TP+FN}$

It is important to have balanced precision and recall values; otherwise, it may indicate that the algorithm is not performing properly.

2.3 Why do we use two metrics (Precision and Recall)?

Two metrics are used because we want to differentiate blood vessels from the background. With only one of these metrics, it would be easy to overlook a deficient algorithm. The following calculation seems to provide a better idea of the overall performance of the algorithm:

$$(Accuracy + Recall)/2$$

The article provides other metrics to measure its performance: specificity and the "area under the curve" (AUC) calculated as follows:

$$Specificity = \frac{TN}{TN + FP}$$

$$AUC = \frac{(Recall + Specificity)}{2}$$

(Note: in the article, recall is named sensitivity, but the calculation method is the same.)

2.4 What role does skeletonization play in this evaluation function?

Skeletonization is used to normalize the generated images compared to the ground truth. This allows us to focus the output of the algorithm on segmenting blood vessels, regardless of their thickness (bounded to 15 pixels in the skeletonization function). It brings us back to the original problem: generating a map of the vascular network. Skeletonization maps the different possible gray levels in the output to a binary response: whether a pixel is part of a blood vessel or not.

3 Recall of the Problem

The objective of the practical work is to extract an image of the retinal vascular network from laser scanning ophthalmoscopy images, a high-resolution retinal imaging technique (between 10 and 100 μm) and wide field – allowing most of the retina to be observed in a single image.

The analysis of the obtained fundus image allows for the diagnosis of several diseases, including high blood pressure, renal insufficiency, and various retinal diseases. The diagnosis requires precise segmentation of the network. This segmentation, performed by human experts, serves as our ground truth and will allow us to validate our automated segmentation.

Evaluation is done using the methods provided for the practical work, refer to part 1 above for more detail.

4 Method

The method is strongly inspired by the article "Blood Vessels Segmentation in Retinal Fundus Image using Hybrid Method of Frangi Filter, Otsu Thresholding and Morphology" (accessible here: Link to the document)

The method involves passing each image through three stages: pre-processing, segmentation, and morphological processing. Each stage consists of several operations, described below. The method also includes a testing phase that we have also implemented to serve as a point of comparison to the validation method provided with the practical work.

However, we have modified the method to achieve what we believe are better results. For this, we replaced the last step with another pruning step.

The comparison between the results obtained is made in the results section. The article summarizes the method as follows:

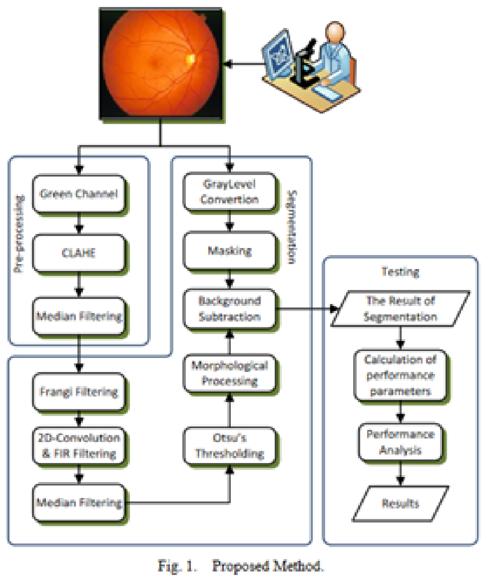

Figure 1: Proposed method

For examples, we will use the images with the best precision, *star08_OSN.jpg*, and the one with the worst precision, *star02_OSC.jpg*.

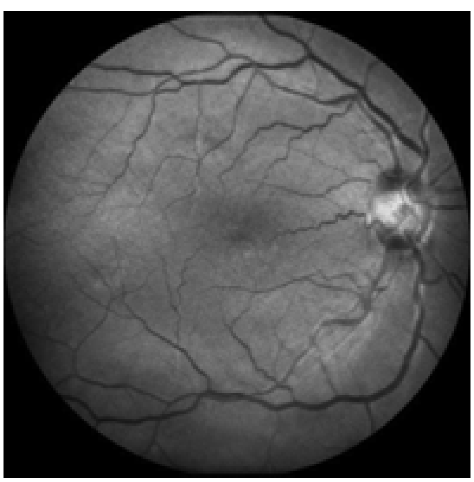

Figure 2: star02_OSC.jpg

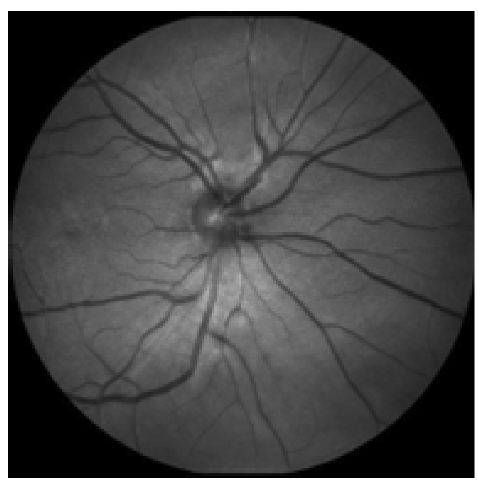

Figure 3: star08_OSN.jpg

5 Pre-processing

The goal of the pre-processing step is to improve the image quality. It includes two steps: contrast enhancement with CLAHE (Contrast-Limited Adaptive Histogram Equalization) and the application of a median filter. The article begins with a color separation step, which does not apply here since we are working with grayscale images.
CLAHE is a technique that involves dividing the image into several blocks and then calculating the gray level histogram for each block before interpolating each pixel from the 4 nearest pixels.
The median filter applied next reduces the noise generated by CLAHE.

After pre-processing, the example images yield the following results:

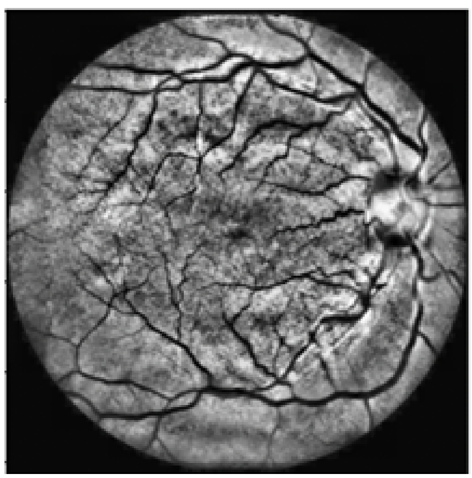

Figure 4: star02_OSC after pre-processing

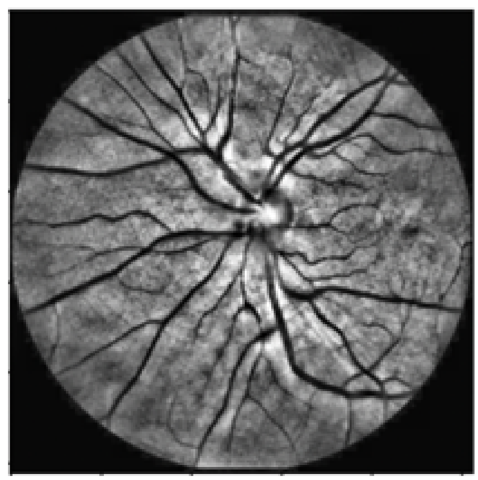

Figure 5: star08_OSN after pre-processing

We observe that the high-frequency regions are clearer and the contrast of the two images is higher.

6 Segmentation

Segmentation is the main step and aims to separate the blood vessels from the image background. It includes 4 operations:

1. Use of the Frangi filter

2. Filtering using a 2D convolution: a 2D FIR (Finite Impulse Response) filter

3. Use of Otsu's thresholding

The Frangi filter is a filter to detect and enhance the quality of blood vessels in images, thus typically the heart of our problem. The filter uses the Hessian matrix of the filter function and extracts its eigenvalues. Note that the filter takes a parameter, sigma, which describes the scale of the blood vessels. For us, the best results were obtained with a sigma of 0.5.

The convolution filter applies a 2D convolution to a kernel which is itself calculated by performing a circular average on each pixel. This results in an FIR filter.

Otsu's thresholding, originally developed by Nobuyuki Otsu in his paper "A Threshold Selection Method from Gray-Level Histograms", differentiates the foreground and background of an image by applying a threshold on the image's gray level histogram. The complete algorithm of the method seeks the best threshold by calculating the probability density and the mean for each possible class. The filtering returns an image where each pixel is a binary value, depending on whether it is judged to be part of the foreground or background.

After segmentation, the example images yield the following results:

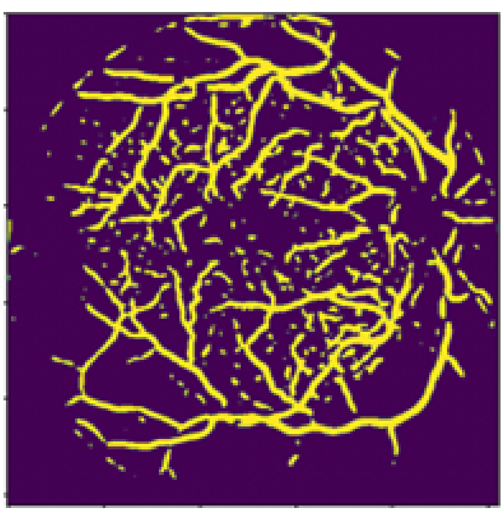

Figure 6: star02_OSC after segmentation

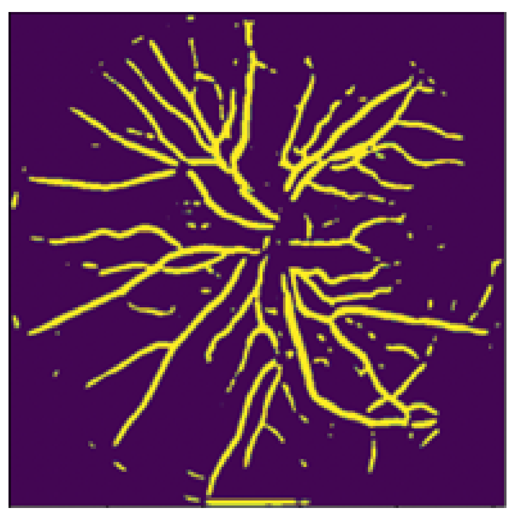

Figure 7: star08_OSN after segmentation

Segmentation clearly separates the blood vessels from the background. We still see some noise, especially near the edge of the mask, but the vessels are clearly visible.

7 Morphological Processing

The morphological processing indicated in the article consists of three operations:

1. A closure
2. Diagonal filling
3. Connecting non-connected pixels

Closure is a simple closure as seen in the course: a dilation followed by erosion.

Diagonal filling aims to eliminate noise by filling holes in the structure of the vessels. It involves checking if a pixel is a hole between two diagonally placed pixels. If so, the hole is filled.

Finally, connecting pixels performs a similar operation by checking if a pixel is a hole between pixels, this time checking two pixels away.

After these operations, the example images yield the following results:

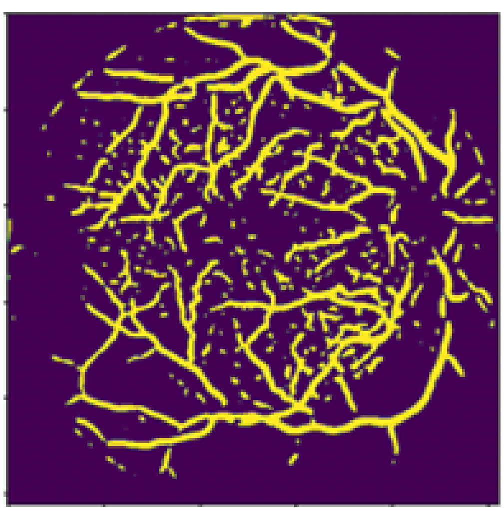

Figure 8: star02_OSC after morphological processing

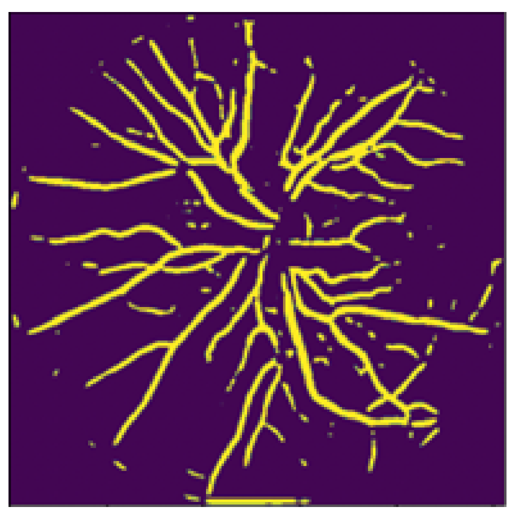

Figure 9: star08_OSN after morphological processing

The difference compared to the previous step is not striking.

8 Pruning

However, as indicated above, we replaced the last step with one that seems more effective to us. It consists of two pruning operations. The first removes small objects (using the `remove_small_object` function from skimage), that is, it replaces image artifacts smaller than a certain size with empty pixels. The second operation prunes small branches by performing an unbalanced closure: the erosion is done with a disc filter of size 1 while the dilation is done with a disc filter of size 3.

The example images yield the following results:

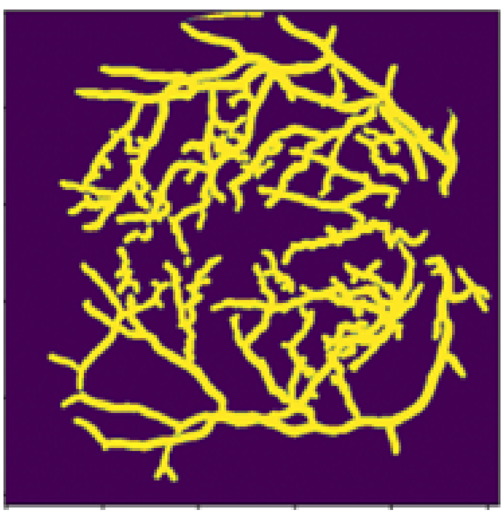

Figure 10: star02_OSC after pruning

Figure 11: star08_OSN after pruning

We see the vessels much more clearly; they are thicker. At the same time, the noise has significantly decreased. We have thus gained in sharpness while losing details.

9 Results

9.1 Precision and Recall

By taking the metrics proposed in the validation method for each image, we obtain the following comparison table: (Cells indicating a better result for our method are in *italic*, those indicating a better result for the article are in **bold**.)

Table 1: Comparison of precision and recall between the paper's method and our method

	Accuracy		Recall	
	Paper's method	Our method	Paper's method	Our method
star01_OSC	74,50%	*87,18%*	**75,35%**	74,99%
star02_OSC	63,98%	*68,17%*	86,22%	*87,04%*
star03_OSN	84,21%	*87,01%*	**67,49%**	64,18%
star08_OSN	**89,72%**	89,39%	**67,82%**	66,25%
star21_OSC	69,18%	*78,28%*	**68,34%**	62,60%
star26_ODC	70,63%	*78,77%*	**72,20%**	68,83%
star28_ODN	74,84%	*77,81%*	**68,25%**	64,28%
star32_ODC	69,41%	*70,20%*	**78,07%**	69,17%
star37_ODN	71,09%	*73,04%*	**76,25%**	73,71%
star48_OSN	77,32%	*81,64%*	**76,78%**	73,22%
Moyenne	74,49%	*79,15%*	**73,68%**	70,43%

We can see that our method allows for better precision in general although with some loss in terms of recall. This confirms what the images showed: thickened vessels are more visible but precision is affected. Note that the average made by metric shows an increase in precision of 4.5 points for a loss of recall of 3 points. We can therefore consider our method as slightly more effective.

9.2 Skeletonization

Furthermore, skeletonization also allows us to get an idea of the result, the goal being well to have an image exploitable by the human eye. For each example, the skeletonization is as follows:

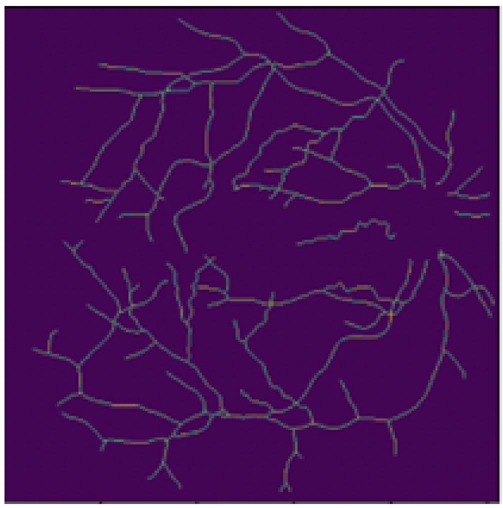

Figure 12: star02_OSC skeletonization of the ground truth

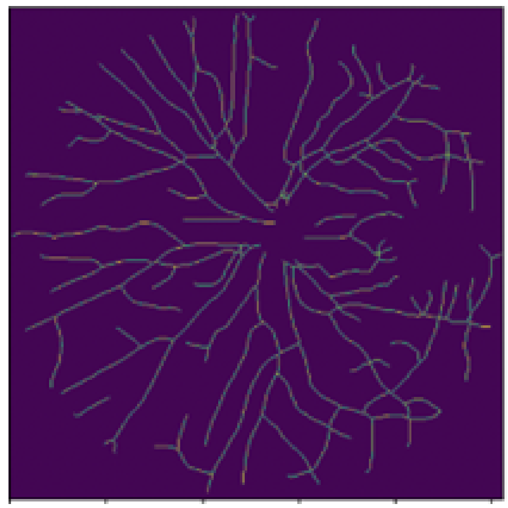

Figure 13: star08_OSN skeletonization of the ground truth

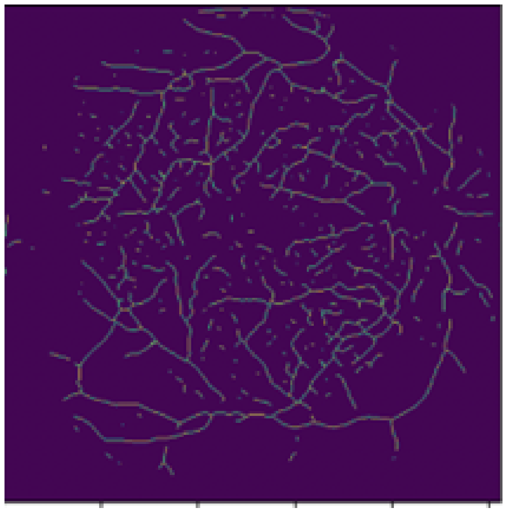

Figure 14: star02_OSC skeletonization after the article's method

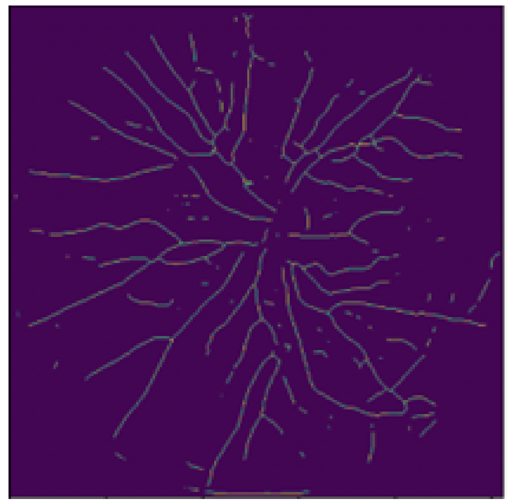

Figure 15: star08_OSN skeletonization after the article's method

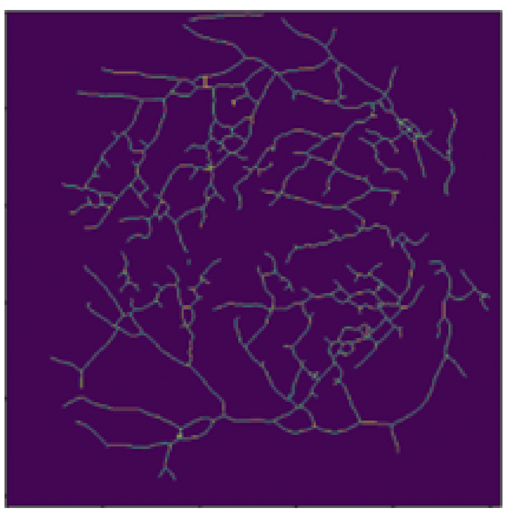

Figure 16: star02_OSC skeletonization after our method

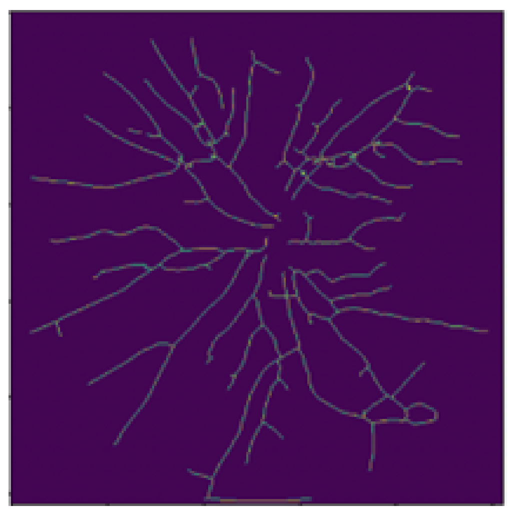

Figure 17: star08_OSN skeletonization after our method

Skeletonization further accentuates our observations: our method removes all traces of noise but the simplification causes us to lose too much information. Note that, contrary to what the numerical results seem to indicate, the visual result is better on star02_OSC than on star08_OSN. A bit of optimism could even make one think that the result thus obtained is more precise than the ground truth, the poor results would thus be due to the fact that the false positives are true positives not recognized as such.

10 Other Metrics

The article defines other metrics in the testing stage: specificity and the area under the curve (AUC). The comparison table for these metrics is as follows:

Table 2: Comparison table for Specificity and Area Under the Curve metrics.

	Specificity		Area Under the Curve	
	Article Method	Our Method	Article Method	Our Method
star01_OSC	**0.949**	0.913	**0.851**	0.832
star02_OSC	**0.907**	0.839	**0.885**	0.855
star03_OSN	**0.957**	0.903	**0.816**	0.772
star08_OSN	**0.963**	0.903	**0.820**	0.783
star21_OSC	**0.924**	0.881	**0.803**	0.754
star26_ODC	**0.928**	0.878	**0.825**	0.783
star28_ODN	**0.941**	0.890	**0.812**	0.766
star32_ODC	**0.904**	0.836	**0.843**	0.764
star37_ODN	**0.938**	0.886	**0.850**	0.811
star48_OSN	**0.961**	0.915	**0.865**	0.824

On these metrics, the result is clear: the article's method gives higher results.

11 Conclusion

In conclusion, we revisited the method indicated in the article "Blood Vessels Segmentation in Retinal Fundus Image using Hybrid Method of Frangi Filter, Otsu Thresholding and Morphology", which allowed us to achieve a precision of almost 75%. A slight modification of the method gave a slightly higher precision, at 79%. However, this method causes a drop in recall from approximately 73.5% to 70%. Thus, the modification has a cost. On the metrics defined in the article, the cost is even more pronounced since both decrease.

That said, the result after skeletonization is more encouraging. There might be a need for further research to obtain even better results.

Bibliography

1. Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, 62 - 66.

2. Wiharto. (2019). Blood Vessels Segmentation in Retinal Fundus Image using Hybrid Method of Frangi Filter, Otsu Thresholding and Morphology. International Journal of Advanced Computer Science and Applications, 417-422. Accessible here.

Appendix 1 – Unsuccessful Path Explored

Before settling on the aforementioned article, we attempted a more empirical approach by trying various filtering operations, morphological treatments, and segmentation entirely from those offered by scikit-image.

We had arrived at the following method:

- A filtering operation with a disk
- Coarse histogram equalization (with a 20-pixel disk filter)
- High-frequency extraction
- An erosion
- A final filtering with a disk

The final images for the examples were as follows:

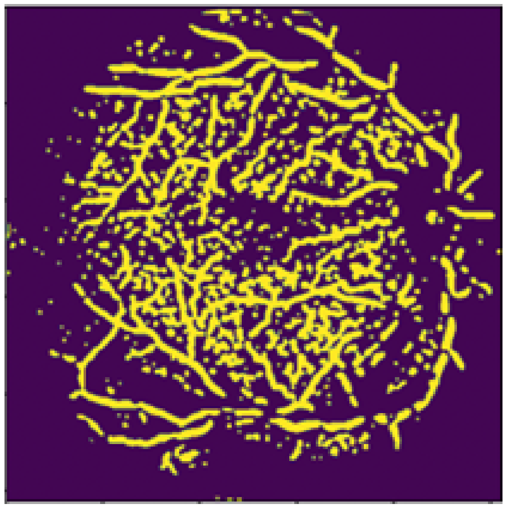

Figure 18: star02_OSC after empirical segmentation

Figure 19: star08_OSN after empirical segmentation

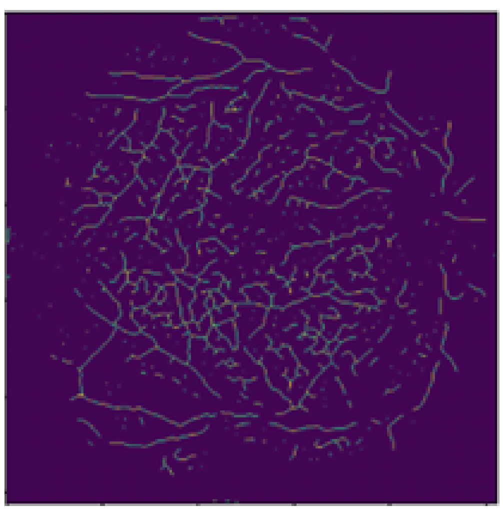

Figure 20: star02_OSC empirical skeletonization

Figure 21: star08_OSN empirical skeletonization

The numerical results were also less good:

Table 3: Precision and Recall Comparison

	Precision	Recall
star01_OSC	59.5%	65.0%
star02_OSC	52.1%	78.8%
star03_OSN	56.5%	75.3%
star08_OSN	60.3%	78.8%
star21_OSC	45.1%	80.8%
star26_ODC	47.5%	79.8%
star28_ODN	56.9%	74.5%
star32_ODC	38.8%	89.5%
star37_ODN	34.5%	95.5%
star48_OSN	30.7%	91.2%

Seeing that this empirical approach only brought little results for the time invested, we preferred to look for a better method...