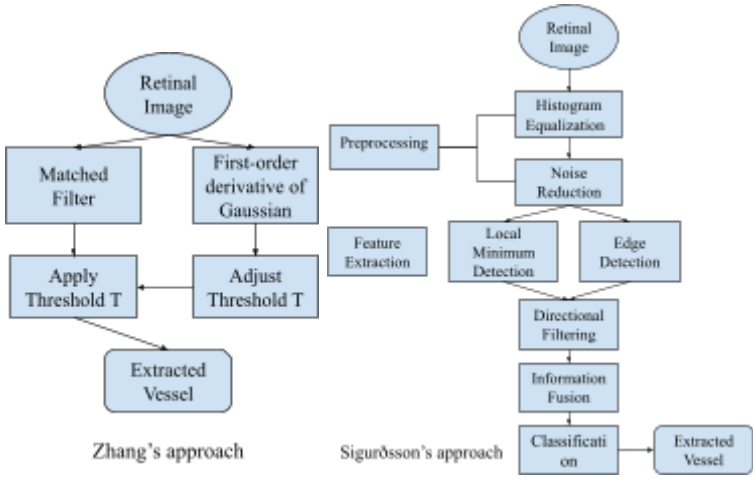
COMP 6771 Image Processing
Course Project

Zunao Hu
40232982

Part I: review of Two papers on Retinal vessel extraction

The automated extraction of vessels in retinal images holds great potential in facilitating computer-aided diagnosis and treatment of common eye diseases.[1]In their respective studies, Bob Zhang et al. and Eysteinn Már Sigurðsson et al.[2] proposed distinct methods for retinal vessel extraction. Zhang introduced a novel method utilising the matched filter(MF) with first-order derivative of Gaussian(FDOG), whereas Sigurðsson presented an information fusion method, combining the local minimum features and edge detection, to leverage the two essential vessel properties. Both of their researches demonstrated high accuracy and robust performances, offering innovative ideas to the field. Furthermore, compared to other methods, theirs maintained low computational costs.

The main process of their approaches can be seen in the following charts. Zhang used a matched filter on the retinal image, and used its response to the FDOG function to adjust and apply the threshold on the filtered image. After preprocessing, Sigurðsson used information fusion to combine the responses of images to local minimum detection and edge detection, and applied classification to obtain the results.



Both approaches included performance demonstrations on the DRIVE database, employing accuracy metrics presented in clear tables. Additionally, both studies include the presentation of Receiver Operating Characteristic (ROC) curves. Sigurðsson further provides tables displaying the outcomes of the parameter tuning process, showcasing the achieved optimized accuracy.

On the DRIVE database, Sigurðsson's method has achieved higher accuracy. It could be attributed to the presence of both healthy and pathological images in the DRIVE database, where Zhang's method may exhibit a lesser advantage in handling healthy cases. Another reason could be the difference in the emphasis of feature extraction. Zhang attempted to leverage the symmetric characteristics of vessels and their response to FDOG, and Sigurðsson concentrated on features related to darker appearances and edges. In the case of the DRIVE database, the latter might have suited better. In addition, the preprocessing step in Sigurðsson's method may mitigate the influence of noise on the final accuracy.

Reference:
[1] Bob Zhang, Lin Zhang, Lei Zhang, Fakhri Karray,
Retinal vessel extraction by matched filter with first-order derivative of Gaussian,
Computers in Biology and Medicine,
Volume 40, Issue 4,
2010,
Pages 438-445,
ISSN 0010-4825,
https://doi.org/10.1016/j.compbiomed.2010.02.008.
[2] Eysteinn Már Sigurðsson, Silvia Valero, Jón Atli Benediktsson, Jocelyn Chanussot, Hugues Talbot, Einar Stefánsson,
Automatic retinal vessel extraction based on directional mathematical morphology and fuzzy classification,
Pattern Recognition Letters,
Volume 47,
2014,
Pages 164-171,
ISSN 0167-8655,
https://doi.org/10.1016/j.patrec.2014.03.006.

Part II: re-implementation of the main algorithm of paper 1.

Result demonstration:
After implementing the custom functions for MF, FDOG filters and adaptive thresholding, five images from the DRIVE dataset were selected and processed.

As is demonstrated in Zhang's paper, two sets of filters were applied to extract wide and thin vessels separately. The results were then combined using the OR operation.

To enhance accuracy calculations, croppings was performed, focusing solely on the region of interest(RoI). In this case, the RoI was defined as a circle with a radius of 250 pixels, excluding the surrounding non-retinal regions in the resulting images.

The processed images and resulting images of the 21st to 23rd training image in DRIVE dataset are displayed as follows:

(1) Train_21
The first and second row depict the response maps generated by MF and FDOG filters, designed for wide and thin vessels, respectively. After adaptive thresholding, uncropped result images are obtained.

In the third row, the result images are combined using OR operation to contain both wide and thin vessel information.

The fourth row illustrates our region of interest, which will only include the retinal structures in the image. It demonstrates how the mask image and result image were cropped based on that region.

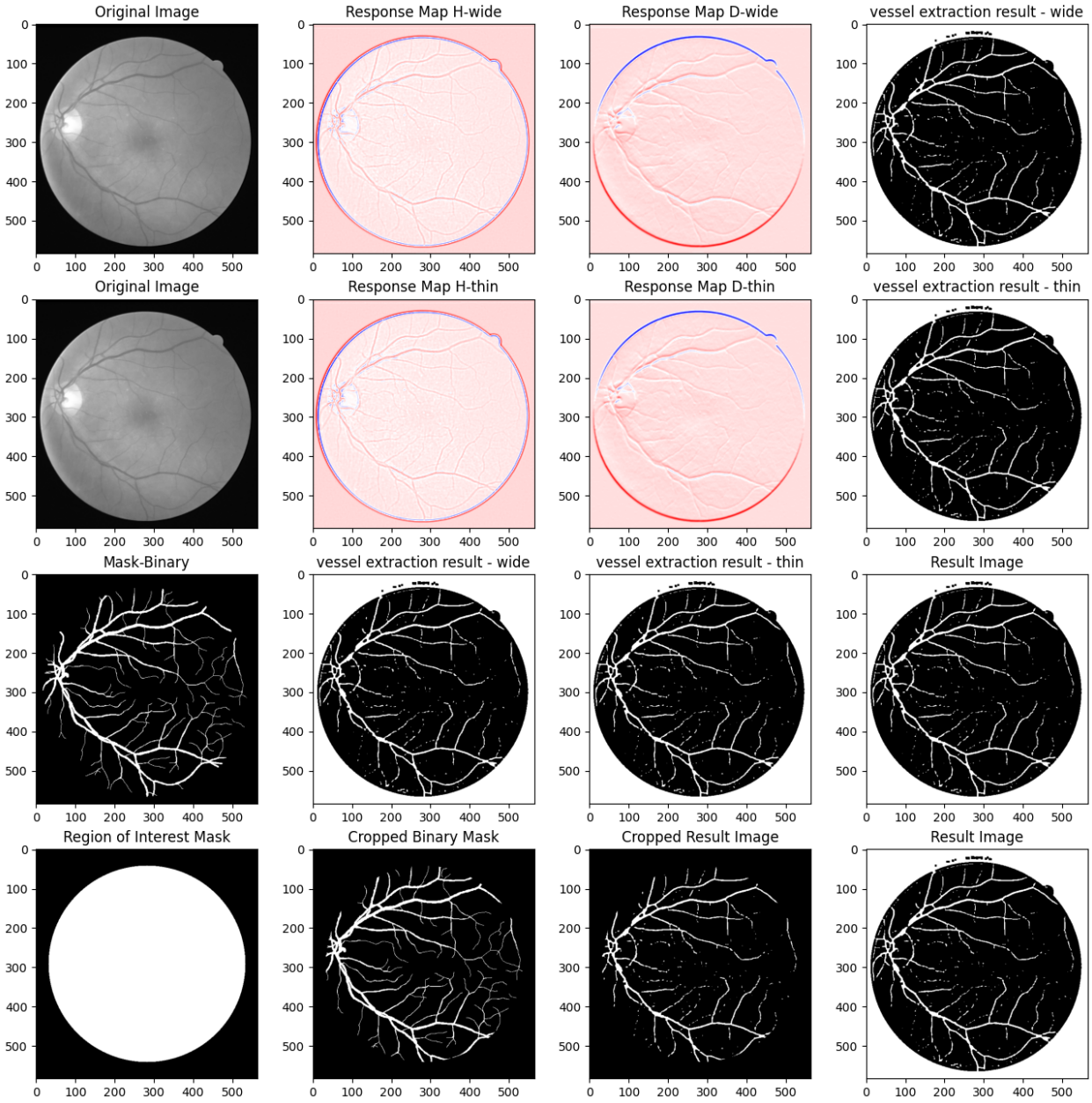


Figure 1. Processing of Training Image 21

(2) Train_22

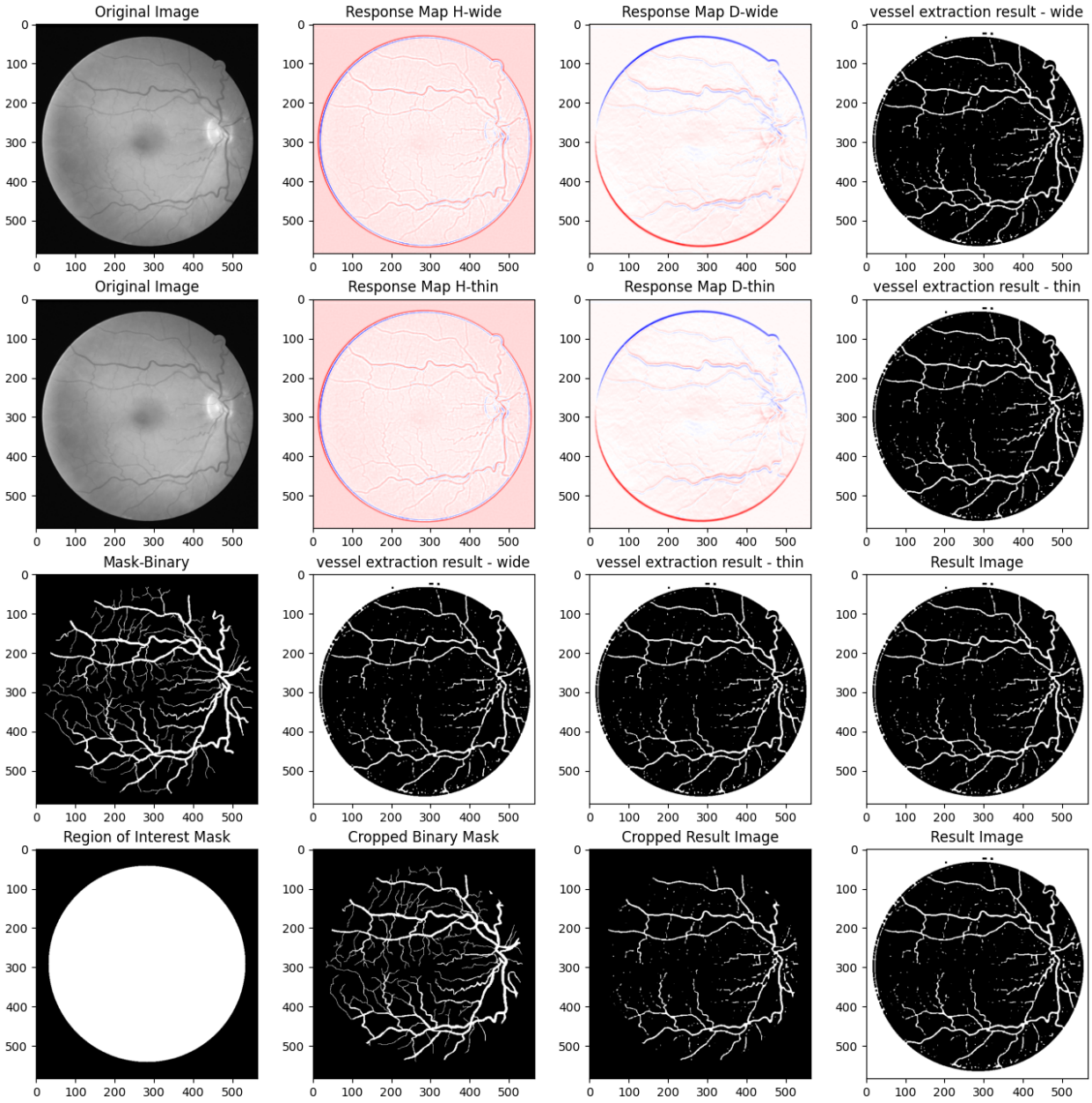


Figure 2. Processing of Training Image 22

(3) Train_23

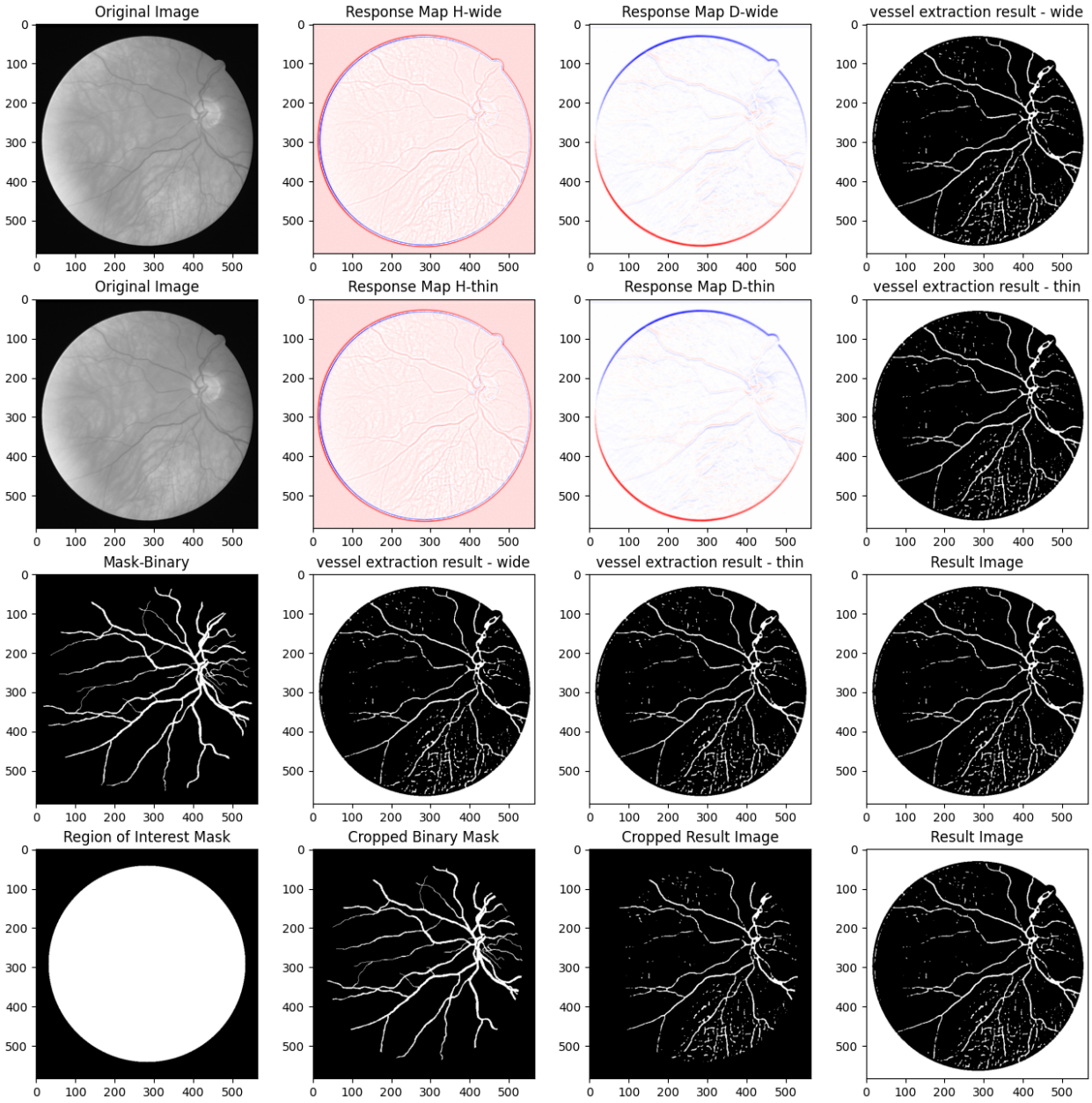


Figure 3. Processing of Training Image 22

Validation:
(1) Performance evaluation:
After ensuring that our result images and masks are both binary images, the performance of our extraction can be assessed by comparing each of their pixel values.

Within the region of interest, the criteria of True Positive Rate(TPR), False Positive Rate(FPR) and accuracy can be calculated.

Image number	**TPR**	**FPR**	**Accuracy**

21	60.90%	1.12%	94.45%

Table 1. TPR, FPR and Accuracy of Training Image 21

(2) Parameter tunings
Tweaking of the values of different parameters was performed to see how they impact the performance of our system. Attempts to tweak the sigma(s) and L have been done, but it turns out that the change is not quite obvious. Therefore, three parameters were taken into consideration: threshold adjusting constant c, size of mean filter W in adaptive thresholding, and the number of rotating matched filters.

Nine sets of parameter tuning are performed in total and the results are as follows:

Parameter	**Image 21**	**22**	**23**	**24**	**25**	**Average**
c = 0.4 W = 31 Rotation = 16	94.45%	92.75%	93.66%	90.16%	90.52%	92.31%
c = 1 W = 31 Rotation = 16	84.50%	83.27%	81.61%	88.71%	82.56%	84.13%
c = 2 W = 31 Rotation = 16	15.34%	18.29%	13.17%	21.23%	18.87%	17.38%
c = 0.4 W = 19 Rotation =16	94.47%	92.77%	93.66%	90.19%	90.53%	92.32%
c = 0.4 W = 31 Rotation = 16	94.45%	92.75%	93.66%	90.16%	90.52%	92.31%
c = 0.4 W = 43 Rotation = 16	94.46%	92.76%	93.64%	90.19%	90.52%	92.31%
c = 0.4 W = 31 Rotation = 8	58.12%	60.94%	57.63%	69.52%	58.23%	60.89%
c = 0.4	94.45%	92.75%	93.66%	90.16%	90.52%	92.31%

W = 31 Rotation = 16						
c = 0.4 W = 31 Rotation = 32	94.45%	92.75%	93.66%	90.16%	90.52%	92.31%

Table 2. Results of Parameter Tuning

The constant c shows an optimal value around 0.4, which is lower than the recommended value of 2.3 in Zhang's paper. This variance could be due to the fact that adjustments and optimizations in Zhang's implementation are mainly based on the STARE dataset. The distinct characteristics of the DRIVE dataset, including variations in patient composition, formatting, and data acquisition sources, might result in a higher response to the FDOG filter, requiring a lower value of c for adjustment.

The tweaking of W does not significantly impact accuracy, potentially because the pixel values in the response map D are closely distributed without distinct features at different scales.

It can be observed that fewer rotations results in lower accuracy, probably because vessels of fewer orientations are extracted due to the reduced rotations of MF.

Moreover, an increase in the number of rotations beyond 16 does not yield a substantial improvement in accuracy, suggesting that 16 rotations are adequate for extracting most vessels with different orientations. To conserve computational resources, setting the number of rotations to 16 is preferable.

(3) Comparison:
Using our average accuracy with optimal parameter settings, the comparison with the methods mentioned in Zhang's paper is as follows:

Method	**Accuracy**
2nd Humanobserver	0.9473
Staal	0.9442
Soares	0.9466
Matched filter	0.9284
Martinez-Perez	0.9344
Zhang's implementation	0.9382
Our implementation	0.9231

Table 3. Comparison with Other Methods

Our implementation demonstrates comparable performance to Zhang's implementation, achieving a high accuracy score, indicating its success.

Nevertheless, it is noteworthy that certain methods, such as Staal, still surpass our implementation. This difference in performance could be attributed to the observation mentioned by Zhang, suggesting that MF-FDOG may exhibit less robust performance on healthy cases compared to pathological images.

The comparison with baseline methods, namely binary thresholding and Canny edge detection, has also been conducted:

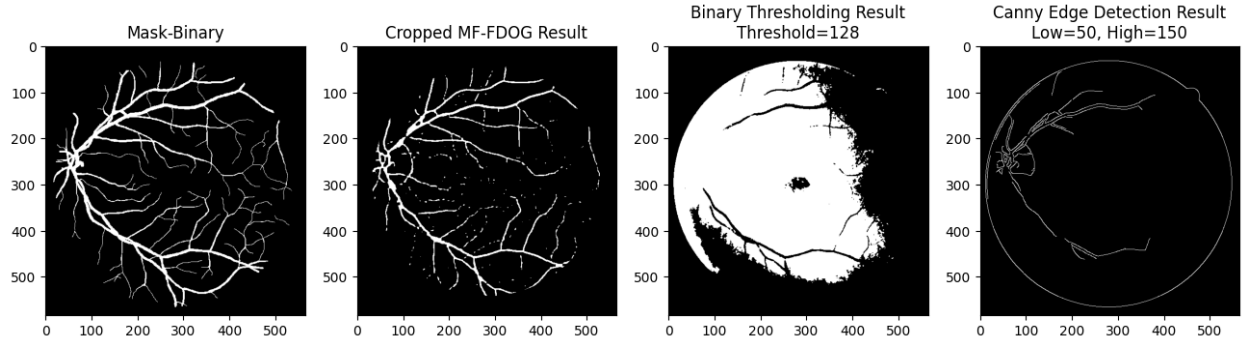


Figure 4. Comparison with Baseline Methods

It can be seen that the binary thresholding method and the Canny edge detection method do not perform as well as the MF-FDOG method we employed.

This discrepancy could be due to the limitation of binary thresholding, which mainly focuses on extracting pixels with low values. However, in vessel structures, pixel values can be irregular, and the method tends to neglect thin vessels.

While Canny edge detection performs slightly better than binary thresholding, it still falls behind the precision of MF-FDOG. Notably, Canny edge detection only captures segments of vessels with prominent edge-like features, indicating that adaptive thresholding in Zhang's method is a great strategy to preserve the integrity of vessel structures.

Discussion:
(1) The pros of the implemented method:
a. The method is concise and efficient, requiring no training. As a result, its runtime is shorter compared to methods involving neural networks.
b. The processes of applying MF filter and FDOG filter are pretty similar, which saves some work for the implementation.
c. It has achieved relatively high performances in case of accuracy. The algorithm has successfully segmented the structural information of retinal vessels.

(2) The cons of the implemented method:
a. The method currently lacks preprocessing steps on the images. The incorporation of additional preprocessing techniques, such as denoising or histogram equalization, has the potential to enhance robustness across different datasets.

b. The inclusion of post-processing steps, such as closing, could address issues observed in the processed and result images, such as holes and gaps in thin vessels. Implementing the closing method has the potential to improve overall accuracy.
c. The two sets of filters for wide and thin vessel extractions do not exhibit significant differences in performance when applied to the DRIVE dataset. Exploring alternative settings for thin vessel extraction may lead to improved accuracy.

(3)The main difficulties in the reimplementation process:
a. The paper lacks detailed information on the application of the matched filter. Although Zhang mentioned rotating the matched filter, the specific number of rotations for optimal results is not provided. To address this gap, I have to incorporate the rotation parameter into our parameter tuning process. Additionally, more comprehensive mathematical explanations of the principles behind the matched filter would enhance understanding.

b. Not all parameters recommended by Zhang yield optimal results on the test retinal images, requiring careful consideration. For instance, setting the thresholding control parameter (c) to the author's reference value results in predominantly black outcomes, indicating that the thresholding value is set too high and critical information is erased. Adjusting c to around 0.4 significantly improves results. Similarly, the two sets of matched filters designed for extracting wide and thin vessels respectively do not exhibit distinct performances in their intended tasks, and even with parameter tweaks, the outcomes remain largely unchanged.

c. The existence of multiple variations of the DRIVE dataset poses a challenge. Initially using the version on Kaggle in .png format resulted in an accuracy of 80%, lower than expected. It was later discovered that various versions of the DRIVE dataset exist in different formats, and the .png version had been downsampled, making it harder to extract concise features. Switching to the .tif version of the DRIVE dataset ultimately led to a reimplementation with performance comparable to Zhang's.