Supplemental Material for the Paper
“Multiscale Centerline Detection by Learning a Scale-Space Distance Transform”

Amos Sironi¹, Vincent Lepetit², and Pascal Fua¹

¹Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne (EPFL)
²Institute for Computer Graphics and Vision, Graz University of Technology
{firstname.lastname}@epfl.ch, lepetit@icg.tugraz.at

Additional Results

In this appendix we present additional results that, because of space limitations, did not fit in our submission. For each dataset used in our experiments we will present the training and test images; the centerline detection and segmentation results of our method, compared against the results of other four methods; the Precision-Recall (PR) curves computed for all the values of the parameters $\rho$ and $\delta$, as described in section 4.2 of the paper; the filters used to extract the features. Moreover we will compare the performances obtained by our method using these features or other convolutional ones.

1. Centerline Detection and Segmentation Results

We compare the results obtained with our method against the results obtained with the other 4 methods mentioned in the paper: OOF [2], OOF+OFA [3], MDOF [4] and Classification, i.e. a variation of [1] for centerlines classification. For each dataset and each method, we compute the maximum projection of the response function along the radial component. We then performed Non-Maximum Suppression (NMS) on the projected images and thresholded them. To select the threshold,

Figure 1. The 7 training images of the Aerial dataset.
we first computed Precision-Recall curves with tolerance $\rho = 2$, as described in section 4.2 of the paper, and then took the threshold that gave the maximum F-measure [5]. Finally, the segmentation results are obtained using the radial estimation, with the procedure explained in section 4.2 of the paper. In this case the threshold was chosen to maximize the F-measure computed with $\delta = 0.4$, where $\delta$ is the tolerance factor introduced in section 4.2 of the paper. In Fig. 2, 3 and 4 are shown the results corresponding to the Aerial dataset. The results for the Brightfield dataset are presented in Fig. 11 and 12, the results for the VC6 dataset are depicted in Fig. 19 and the results for the Vivo2P dataset are shown in Fig. 26, Fig. 27 and Fig. 28. The training images used for the learning stage of our method and of the Classification approach are shown in Fig. 1, 10, 18 and 25.

2. Precision-Recall Curves

The Precision-Recall (PR) curves for centerline detection and radial estimation were computed using the evaluation methodology described in section 4.2 of the paper. We plot the curves for several values of $\rho$ between 1 and 5 and several values of $\delta$ between 0.1 and 0.6. The curves corresponding to the Aerial dataset are depicted in Fig. 5 and 6. The curves corresponding to the Brightfield dataset are shown in Fig. 13 and 14. The curves corresponding to the VC6 dataset are presented in Fig. 20 and 21 and those corresponding to the Vivo2P dataset in Fig. 29 and 30.

3. Features

To compute the features used by our method we first learn a bank of convolutional filters and then approximate them with separable filters, as described in section 3.3 of the paper. The full-rank filters learned on the Aerial dataset and the separable ones used to approximate them are shown in Fig. 7, the full-rank and separable filters learned on the Brightfield dataset are shown in Fig. 15, those learned on the VC6 dataset are shown in Fig. 22 and those learned on the Vivo2P dataset in Fig. 31. To prove the power of these filter banks, we compared their performances against 3 other filter banks:

- Gaussian Derivatives: A filter bank composed of Gaussian derivatives up to the fourth order, steered at 8 different orientation in the case of 2D images and 14 different orientations in the case of 3D volumes.
- OOF Filters: A filter bank composed of second derivatives of a ball. These are the filters on which the OOF response is based. We rotated them at 8 different orientation in the case of 2D images and 14 different orientations in the case of 3D volumes. The radius of the ball was chosen to match to the scale of the structure of interest;
- Random Filters: A filter bank of 121 filters generated from pseudorandom values drawn from the standard normal distribution.

Each of the 4 filter banks was used to extract features from the images via convolution. The results of the convolution were used as input at training and test time for our method. For these experiments, in order to have filter bank with comparable number of filters, we did not rotate the learned filters in the case of 3D volumes.

The PR curves obtained using the different sets of filters are shown in Fig. 8 and Fig. 9 for the Aerial dataset, in Fig. 16 and Fig. 17 for the Brightfield dataset, in Fig. 23 and Fig. 24 for the VC6 dataset and in Fig. 32 and Fig. 33 for the Vivo2P6 dataset. They demonstrate the advantage of using the learned filters instead of the hand-crafted ones.

References

Figure 2. Centerline detection and Segmentation on the Aerial dataset for the different methods (Images 1 and 2). **First and fourth rows:** Original image and maximum projection of the score function returned by the different methods; **Second and fifth rows:** Ground truth centerlines and centerlines detected with the different methods; **Third and sixth rows:** Segmentation ground truth and segmentations obtained with the different methods.
Figure 3. Centerline detection and Segmentation on the Aerial dataset for the different methods (Images 3 and 4).
Figure 4. Centerline detection and Segmentation on the Aerial dataset for the different methods (Images 5, 6 and 7).
Figure 5. Precision-Recall curves for Centerline Detection on the Aerial dataset. For all the values $\rho$ our method outperforms the others.

Figure 6. Precision-Recall curves for Segmentation accuracy on the Aerial dataset. For all the values $\delta$ our method outperforms the others.
Figure 7. The filters learned on the Aerial dataset.

Figure 8. Filter Banks Performance Comparison. Precision-Recall curves for Centerline Detection on the Aerial dataset. For all the values of $\rho$ the learned filters lead to better performance.
Figure 9. Filter Banks Performance Comparison. Precision-Recall curves for segmentation results on the Aerial dataset. For all the values of $\delta$ the learned filters lead to better performance.

Figure 10. The 3 training images of the Brightfield dataset.
Figure 11. Centerline detection and segmentation on the Brightfield dataset for the different methods (Image 1). **First row**: Original image and maximum projection of the score function returned by the different methods; **Second row**: Ground truth centerlines and centerlines detected with the different methods; **Third row**: Segmentation ground truth and segmentations obtained with the different methods.
Figure 12. Centerline detection and Segmentation on the Brightfield dataset for the different methods (Image 2). **First row:** Original image and maximum projection of the score function returned by the different methods; **Second row:** Ground truth centerlines and centerlines detected with the different methods; **Third row:** Segmentation ground truth and segmentations obtained with the different methods.
Figure 13. Precision-Recall curves for Centerline Detection on the Brightfield dataset. For all the values $\rho$ our method outperforms the others.

Figure 14. Precision-Recall curves for Segmentation accuracy on the Brightfield dataset. For all the values $\delta$ our method outperforms the others.
Figure 15. The filters learned on the Brightfield dataset.

Figure 16. Filter Banks Performance Comparison. Precision-Recall curves for Centerline Detection on the Brightfield dataset. For all the values of $\rho$ the learned filters lead to better performance.
Figure 17. Filter Banks Performance Comparison. Precision-Recall curves for segmentation results on the Brightfield dataset. For all the values of $\delta$ the learned filters lead to better performance.

Figure 18. The 3 training images of the VC6 dataset.
Figure 19. Centerline detection and Segmentation on the VC6 dataset for the different methods. **First and fourth rows:** Original image and maximum projection of the score function returned by the different methods; **Second and fifth rows:** Ground truth centerlines and centerlines detected with the different methods; **Third and sixth rows:** Segmentation ground truth and segmentations obtained with the different methods.
Figure 20. Precision-Recall curves for Centerline Detection on the VC6 dataset. For all the values $\rho$ our method outperforms the others.

Figure 21. Precision-Recall curves for Segmentation accuracy on the VC6 dataset. For all the values $\delta$ our method outperforms the others.
Figure 22. The filters learned on the VC6 dataset.

Figure 23. Filter Banks Performance Comparison. Precision-Recall curves for Centerline Detection on the VC6 dataset. For all the values of $\rho$ the learned filters lead to better performance.
Figure 24. Filter Banks Performance Comparison. Precision-Recall curves for segmentation results on the VC6 dataset. For all the values of $\delta$ the learned filters lead to better performance.

Figure 25. The 2 training images of the Vivo2P dataset.
Figure 26. Centerline detection and segmentation on the Vivo2P dataset for the different methods (Images 1, 2, and 3).
Figure 27. Centerline detection and Segmentation on the Vivo2P dataset for the different methods (Images 4, 5 and 6).
Figure 28. Centerline detection and Segmentation on the Vivo2P dataset for the different methods (Images 7, 8 and 9).
Figure 29. Precision-Recall curves for Centerline Detection on the Vivo2P dataset. For all the values $\rho$ our method outperforms the others.

Figure 30. Precision-Recall curves for Segmentation accuracy on the Vivo2P dataset. For all the values $\delta$ our method outperforms the others.
Figure 31. The filters learned on the Vivo2P dataset.

Figure 32. Filter Banks Performance Comparison. Precision-Recall curves for Centerline Detection on the Vivo2P dataset. For all the values of $\rho$ the learned filters lead to better performance.
Figure 33. Filter Banks Performance Comparison. Precision-Recall curves for segmentation results on the Vivo2P dataset. For all the values of $\delta$ the learned filters lead to better performance.