Object Skeleton Extraction in Natural Images by Fusing Scale-associated Deep Side Outputs
Supplementary Material

Wei Shen[1], Kai Zhao[1], Yuan Jiang[1], Yan Wang[2], Zhijiang Zhang[1], Xiang Bai[3]
[1] Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Shanghai University
[2] Rapid-Rich Object Search Lab, Nanyang Technological University
[3] School of Electronic Information and Communications, Huazhong University of Science and Technology
wei.shen@shu.edu.cn, zeakey@outlook.com, jy9387@outlook.com, wyanny.9@gmail.com, zjzhang@shu.edu.cn, xbai@hust.edu.cn

1. Additional Results

1.1. Qualitative Results on the WH-SYMMAX Dataset

Qualitative comparisons on the WH-SYMMAX Dataset [3] are illustrated in Fig. 1.

Figure 1. Illustration of skeleton extraction results on the WH-SYMMAX dataset [3] for several selected images. The groundtruth skeletons are in yellow and the thresholded extraction results are in red. Thresholds were optimized over the whole dataset.

1.2. Qualitative Results on the SYMMAX300 Dataset

A large number of groundtruths are labeled on non-object parts in SYMMAX300 [4], which do not have organized structures as object skeletons. Our aim is to suppress those on non-object parts, so that the obtained skeletons can be used for other potential applications. In addition, the groundtruths for scale are not provided by SYMMAX300. Therefore, we do not evaluate our method quantitatively on SYMMAX300. Even so, Fig. 2 shows that our method can obtain good skeletons of some objects in SYMMAX300. We also observe that the results obtained by our method have significantly less noises on background.

1.3. Symmetric Part Segmentation

To verify the usefulness of the extracted skeletons, we follow the criteria in [2] for symmetric part segmentation. We evaluate the ability of our skeleton to find segmentation masks corresponding to object parts in a cluttered scene. Our network provides a predicted scale for each skeleton pixel (the fused skeleton score map for which scale has maximal response). With it we can recover object parts from skeletons. For each skeleton pixel x_j, we can predict its scale by $\hat{s}_j = \sum_{i=1}^{M} \Pr(z_j = i|X; \mathbf{\Theta}*, \mathbf{\Phi}*, \mathbf{a}_0*) r_i$. Then for a skeleton segment $\{x_j, j = 1, \ldots, N\}$, where N is the number of the skeleton pixels in this segment, we can obtain a segmented object part mask by $\mathcal{M} = \bigcup_{j=1}^{N} D_j$, where D_j is the disk of center x_j and diameter $\hat{s}_j$. A confidence score is also assigned to each object part mask for quantitative evaluation: $P_{\mathcal{M}} = \frac{1}{N}\sum_{j=1}^{N}(1 - \Pr(z_j = 0|X; \mathbf{\Theta}*, \mathbf{\Phi}*, \mathbf{a}_0*))$. We compare our segmented part masks with Lee's method [1] and Levinshtein's method [2] on their BSDS-Parts dataset [1], which contains 36 images annotated with ground-truth masks corresponding to the symmetric parts of prominent objects. The segmentation results are evaluated by the protocol used in [1]: A segmentation mask $\mathcal{M}_{seg}$ is counted as a hit if its overlap with the ground-truth mask $\mathcal{M}_{gt}$ is greater than 0.4, where overlap is measured by intersection-over-union (IoU). A precision/recall curve is obtained by varying a threshold over the confidence scores of segmented masks. The quantitative evaluation results are summarized in Fig. 3, which indicate a significant improvement over the other two methods. Some qualitative results on the BSDS-Parts dataset [1] are shown in Fig. 4.

References

[1] T. S. H. Lee, S. Fidler, and S. J. Dickinson. Detecting curved symmetric parts using a deformable disc model. In *Proc. IC-*

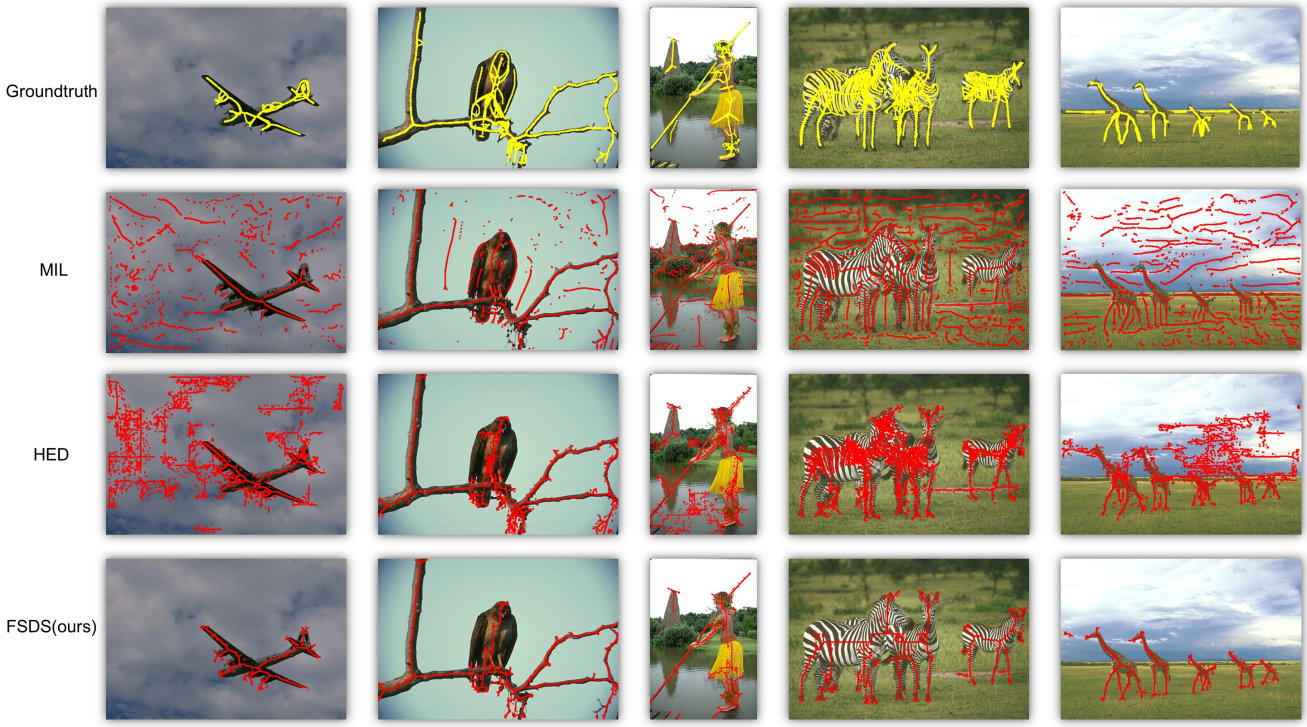

Figure 2. Illustration of skeleton extraction results on the SYMMAX300 dataset [4] for several selected images. The groundtruth skeletons are in yellow and the thresholded extraction results are in red. Thresholds were optimized over the whole dataset.

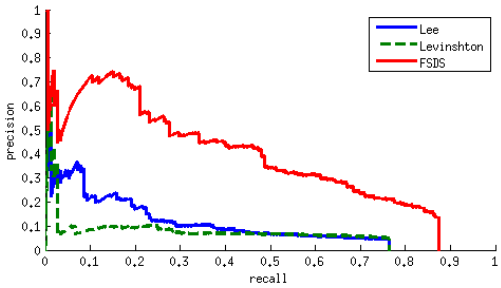

Figure 3. Symmetric part segmentation results on BSDS-Parts dataset [1].

CV, pages 1753–1760, 2013.

[2] A. Levinshtein, S. J. Dickinson, and C. Sminchisescu. Multi-scale symmetric part detection and grouping. In *Proc. ICCV*, pages 2162–2169, 2009.

[3] W. Shen, X. Bai, Z. Hu, and Z. Zhang. Multiple instance subspace learning via partial random projection tree for local reflection symmetry in nature images. *Pattern Recognition*, 52:266–278, 2016.

[4] S. Tsogkas and I. Kokkinos. Learning-based symmetry detection in natural images. In *Proc. ECCV*, pages 41–54, 2012.

Figure 4. Illustration of symmetric part segmentation results on the BSDS-Parts dataset [1] for several selected images. In each column, we show the orginal image, the segmentation groundtruth, the thresholded extracted skeleton (in green), the segmented masks recovered by the skeleton. Thresholds were optimized over the whole dataset.