Foreground color prediction through inverse compositing
Supplementary materials

Sebastian Lutz
V-SENSE
Trinity College Dublin
lutzs@scss.tcd.ie

Aljosa Smolic
V-SENSE
Trinity College Dublin
smolica@scss.tcd.ie

1. Visualization of the iterative prediction process

To visualize our iterative prediction process, in figure 1, we show our predictions for $t = [1, 2, 3, 4, 5]$.

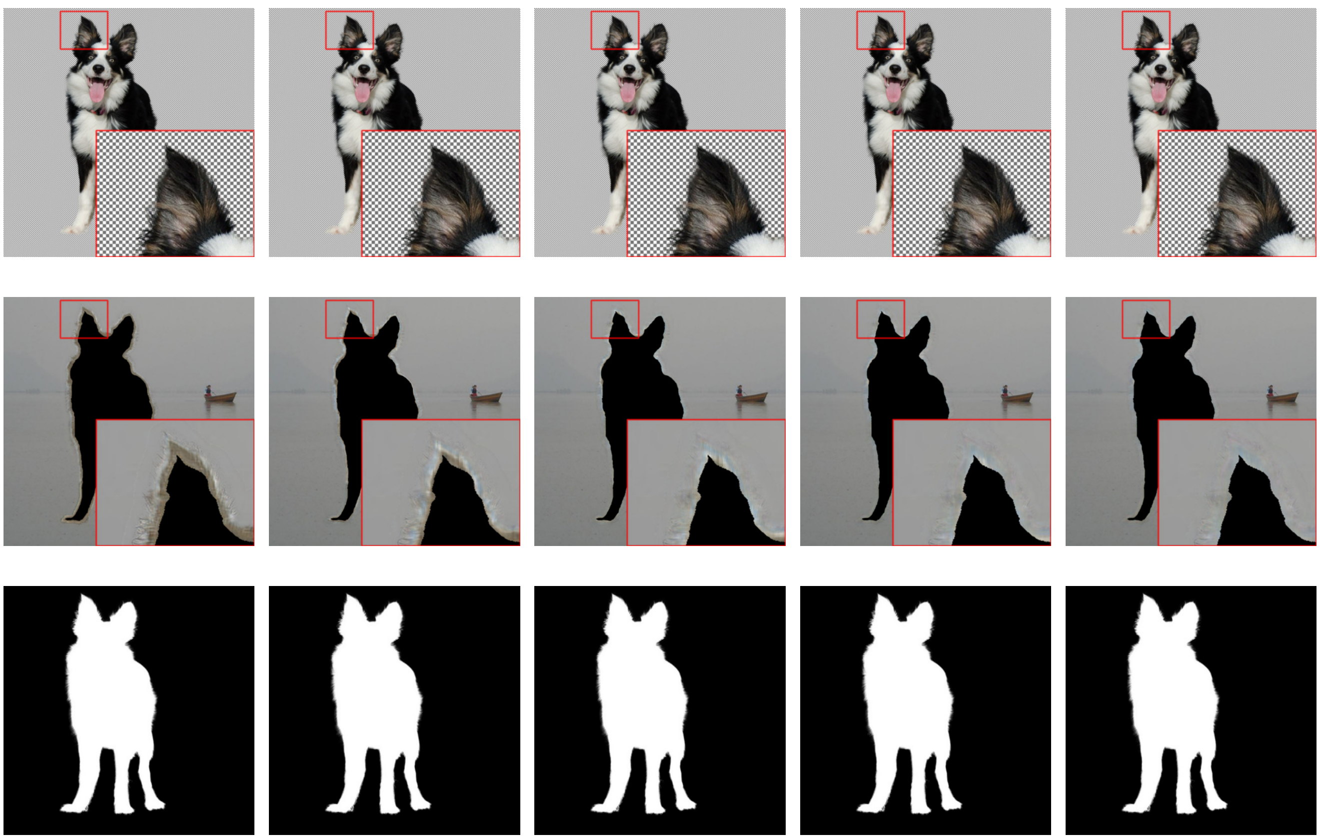

Figure 1: Our prediction over $t = 1, \ldots, 5$. While our alpha prediction barely changes from the initial input, both the foreground and background predictions get consecutively more accurate, as seen especially in the background prediction.

2. Additional comparison results on the Composition-1k test dataset

We show further comparisons of our method to Context-Aware Matting [4] and Samplenet [9] in figures 2 and 3

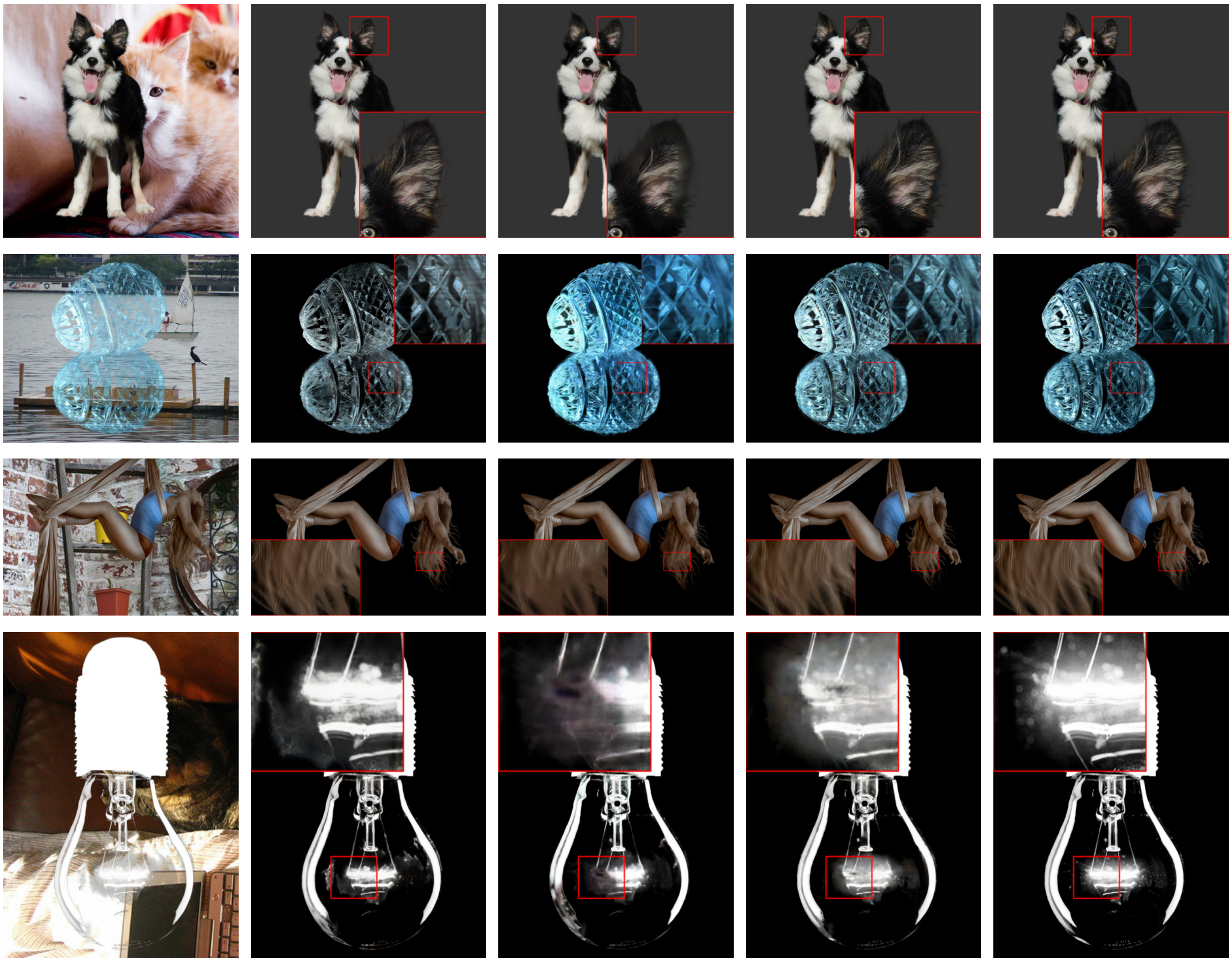

Figure 2: Visual comparison on the Composition-1k dataset. From left to right: Input image, Compositions from Context-Aware Matting [4], Samplenet [9], Ours, Ground-truth.

Figure 3: Visual comparison on the Composition-1k dataset. From left to right: Input image, Compositions from Context-Aware Matting [4], Samplenet [9], Ours, Ground-truth.

3. Visualization of the manual editing

We show the automatic predictions, the manual edits done and the refined results in figure 4.

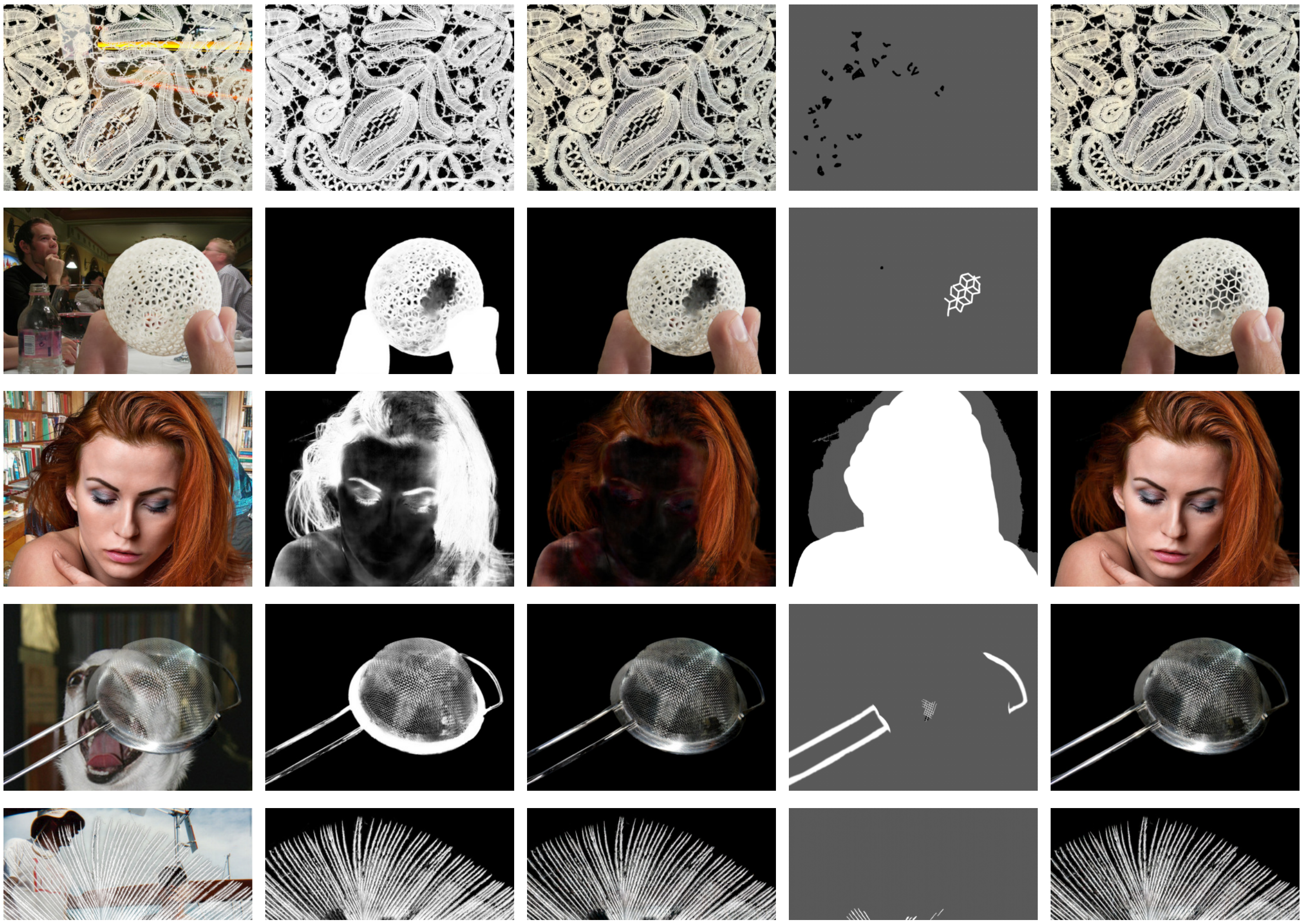

Figure 4: Visualization of the manual editing process. From left to right: Input image, predicted alpha, composition from automatically predicted alpha and foreground, editing mask for the alpha, composition from the updated alpha and newly predicted foreground.
As can be seen, the faulty automatic alpha predictions lead to unappealing compositions. However, a small amount of manual editing is sufficient to recover foreground color predictions that lead, alongside the new alpha, to much better compositions. Please note that the amount of manual editing was deliberately kept low.

4. User study results

We show some example images from the user study we conducted in figure 5.

Figure 5: Example images from the user study. From left to right: Input image, Context-Aware Matting [4], our result with the alpha prediction from Context-Aware as input, Samplenet [9], our result with the alpha prediction from Samplenet as input.

As can be seen in the comparison with Samplenet, the color predictions from Samplenet are somewhat smoothed out on the edges, which can lead to more unappealing compositions. For the comparison with Context-Aware Matting, the differences in color predictions are more difficult to spot.

5. Alpha matte prediction

We show the results on the commonly used evaluation metrics for alpha matting in table 1.

Methods	MSE	SAD	Grad	Conn
KNN [3]	0.078	112.60	67.68	113.47
KNN [3] + Ours	0.078	112.81	67.75	113.52
IF [1]	0.066	75.41	63.39	75.48
IF [1] + Ours	0.066	75.47	63.38	75.48
AlphaGAN [7]	0.031	68.71	50.97	70.42
AlphaGAN [7] + Ours	0.031	68.72	50.97	70.40
Deep Image Matting [10]	0.014	50.4	31.0	50.8
IndexNet [6]	0.013	45.8	25.9	43.7
VDRN [8]	0.011	45.3	30.0	45.6
AdaMatting [2]	0.010	41.7	16.8	–
SampleNet [9]	0.010	46.79	22.50	45.64
SampleNet [9] + Ours	0.010	46.82	22.51	45.66
GCA [5]	0.009	**35.28**	16.92	32.53
GCA [5] + Ours	**0.009**	35.31	**16.91**	**32.53**

Table 1: Quantitative results of the alpha prediction on the Composition-1k dataset. Best results are emphasized in bold. Note that not all image could be predicted for KNN Matting and Information-flow Matting due to trimaps incompatible with these methods.

References

[1] Yagiz Aksoy, Tunç Ozan Aydin, and Marc Pollefeys. Designing effective inter-pixel information flow for natural image matting. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 228–236. IEEE Computer Society, 2017.

[2] Shaofan Cai, Xiaoshuai Zhang, Haoqiang Fan, Haibin Huang, Jiangyu Liu, Jiaming Liu, Jiaying Liu, Jue Wang, and Jian Sun. Disentangled image matting. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 8819–8828, 2019.

[3] Qifeng Chen, Dingzeyu Li, and Chi-Keung Tang. KNN matting. In *2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, June 16-21, 2012*, pages 869–876. IEEE Computer Society, 2012.

[4] Qiqi Hou and Feng Liu. Context-aware image matting for simultaneous foreground and alpha estimation. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*, pages 4129–4138. IEEE, 2019.

[5] Yaoyi Li and Hongtao Lu. Natural image matting via guided contextual attention. *CoRR*, abs/2001.04069, 2020.

[6] Hao Lu, Yutong Dai, Chunhua Shen, and Songcen Xu. Indices matter: Learning to index for deep image matting. *CoRR*, abs/1908.00672, 2019.

[7] Sebastian Lutz, Konstantinos Amplianitis, and Aljosa Smolic. Alphagan: Generative adversarial networks for natural image matting. In *British Machine Vision Conference 2018, BMVC 2018, Northumbria University, Newcastle, UK, September 3-6, 2018*, page 259. BMVA Press, 2018.

[8] Huan Tang, Yujie Huang, Ming-e Jing, Yibo Fan, and Xiaoyang Zeng. Very deep residual network for image matting. In *2019 IEEE International Conference on Image Processing, ICIP 2019, Taipei, Taiwan, September 22-25, 2019*, pages 4255–4259. IEEE, 2019.

[9] Jingwei Tang, Yagiz Aksoy, Cengiz Öztireli, Markus H. Gross, and Tunç Ozan Aydin. Learning-based sampling for natural image matting. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*, pages 3055–3063. Computer Vision Foundation / IEEE, 2019.

[10] Ning Xu, Brian L. Price, Scott Cohen, and Thomas S. Huang. Deep image matting. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 311–320. IEEE Computer Society, 2017.