Supplementary Material: Semantics-Adding Flaw-Erasing Network for Semantic Human Matting

Jiayu Sun[1,2]
jiayusun666@gmail.com
Zhanghan Ke[2]
kezhanghan@outlook.com
Ke Xu[2]
kkangwing@gmail.com
Fan Shao[3]
shaofan075@126.com
Lihe Zhang[1]
zhanglihe@dlut.edu.cn
Huchuan Lu[1,4]
lhchuan@dlut.edu.cn
Rynson W.H. Lau[2]
rynson.lau@cityu.edu.hk

[1] School of Information and Communication Engineering Dalian University of Technology Dalian, China
[2] Department of Computer Science City University of Hong Kong, Hong Kong SAR, China
[3] Wonxing Technology, Inc. Beijing, China
[4] Peng Cheng Laboratory Shenzhen, China

In this supplemental, we show some equations used in FEM. Moreover, we also provide the tables about our human matting dataset. Finally, we provide more visual results.

1 FEM

The detail of the inferior pixels indicated by $\mathcal{FM}$ can be enhanced via the following operations:

$$\begin{aligned}
i^t &= \sigma(W_{xi} * \tilde{F}_{in,c,s,d} + W_{hi} * h^{t-1} + W_{ci} \circ c^{t-1} + b_i),\\
f^t &= \sigma(W_{xf} * \tilde{F}_{in,c,s,d} + W_{hf} * h^{t-1} + W_{cf} \circ c^{t-1} + b_f),\\
c^t &= f^t \circ c^{t-1} + i^t \circ \tanh(W_{xc} * \tilde{F}_{in,c,s,d} + W_{hc} * h^{t-1} + b_c),\\
o^t &= \sigma(W_{xo} * \tilde{F}_{in,c,s,d} + W_{ho} * h^{t-1} + W_{co} \circ c^{t-1} + b_o),\\
h^t &= o^t \circ \tanh(c^t),
\end{aligned} \tag{1}$$

where i^t, f^t and o^t stand for the input, forget and output gates, respectively, and c^t stores the earlier information. $\circ$ denotes the Hadamard product. h^0 and c^0 are initialized to 0. At the last step N, refined features $F^N_{refine} = h^N$ are generated with the memory of the previous refined features. The final alpha matte is generated by $\alpha = \sigma(W_*(F^N_{refine}))$, where we use convolutional layers to reduce the dimension. We set $N = 3$ in practice.

2 Human Matting Dataset

Table 1 compares our dataset with the existing human matting datasets:

Table 1: Configurations of different human matting datasets.

Dataset	Publicly available	Train Set		Test Set	
		Human	Image	Human	Image
Shen et al. [8]	✓	1700	1700	300	300
Trimap+DIM [8]	✓	202	20200	11	220
Distinctions646 [8]	✓	362	36200	11	220
SHM [8]	×	34493	34493	1020	1020
BSHM [8]	×	9324	93240	125	1250
BSHM (Coarse)	×	+10597	+105970	-	-
Ours	✓	4494	44940	235	2350

Table 2 shows details of the three subsets of our dataset:

Table 2: Foreground configurations of the proposed dataset.

Sub-dataset	Train Set	Test Set	Total
Single Frontal Subset	2177	106	2283
Single Pose-varied Subset	1901	110	2011
Multiple Subset	416	19	435
Total	4494	235	4729

3 More results

Fig. 1, we provide more visual results:

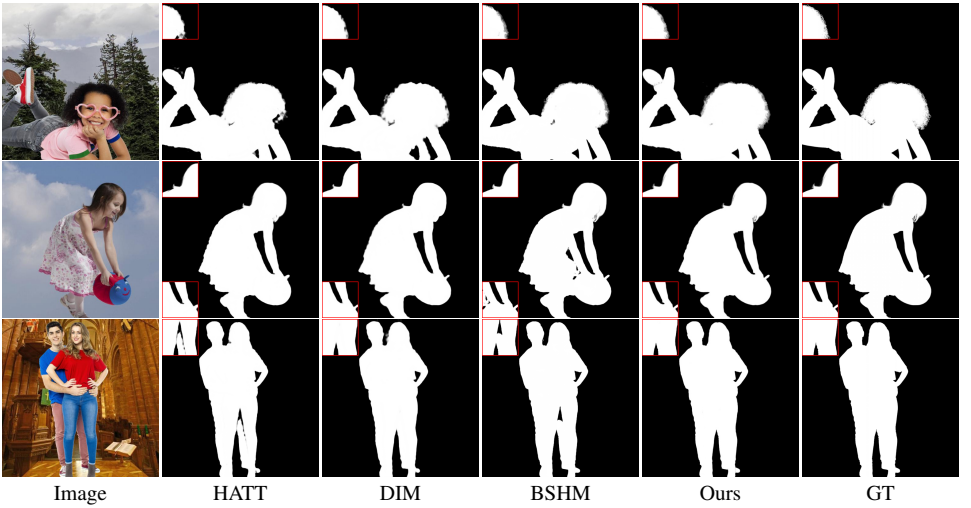

Figure 1: Qualitative comparison on the proposed dataset.

References

[1] Quan Chen, Tiezheng Ge, Yanyu Xu, Zhiqiang Zhang, Xinxin Yang, and Kun Gai. Semantic human matting. In ACM MM, pages 618–626, 2018.

[2] Jinlin Liu, Yuan Yao, Wendi Hou, Miaomiao Cui, Xuansong Xie, Changshui Zhang, and Xian-sheng Hua. Boosting semantic human matting with coarse annotations. In CVPR, pages 8563–8572, 2020.

[3] Yu Qiao, Yuhao Liu, Xin Yang, Dongsheng Zhou, Mingliang Xu, Qiang Zhang, and Xiaopeng Wei. Attention-guided hierarchical structure aggregation for image matting. In CVPR, pages 13676–13685, 2020.

[4] Xiaoyong Shen, Xin Tao, Hongyun Gao, Chao Zhou, and Jiaya Jia. Deep automatic portrait matting. In ECCV, pages 92–107. Springer, 2016.

[5] Ning Xu, Brian Price, Scott Cohen, and Thomas Huang. Deep image matting. In CVPR, pages 2970–2979, 2017.