Context-Aware Image Matting for Simultaneous Foreground and Alpha Estimation

Portland State University

Qiqi Hou and Feng Liu

ICCV 2019 Seoul, Korea

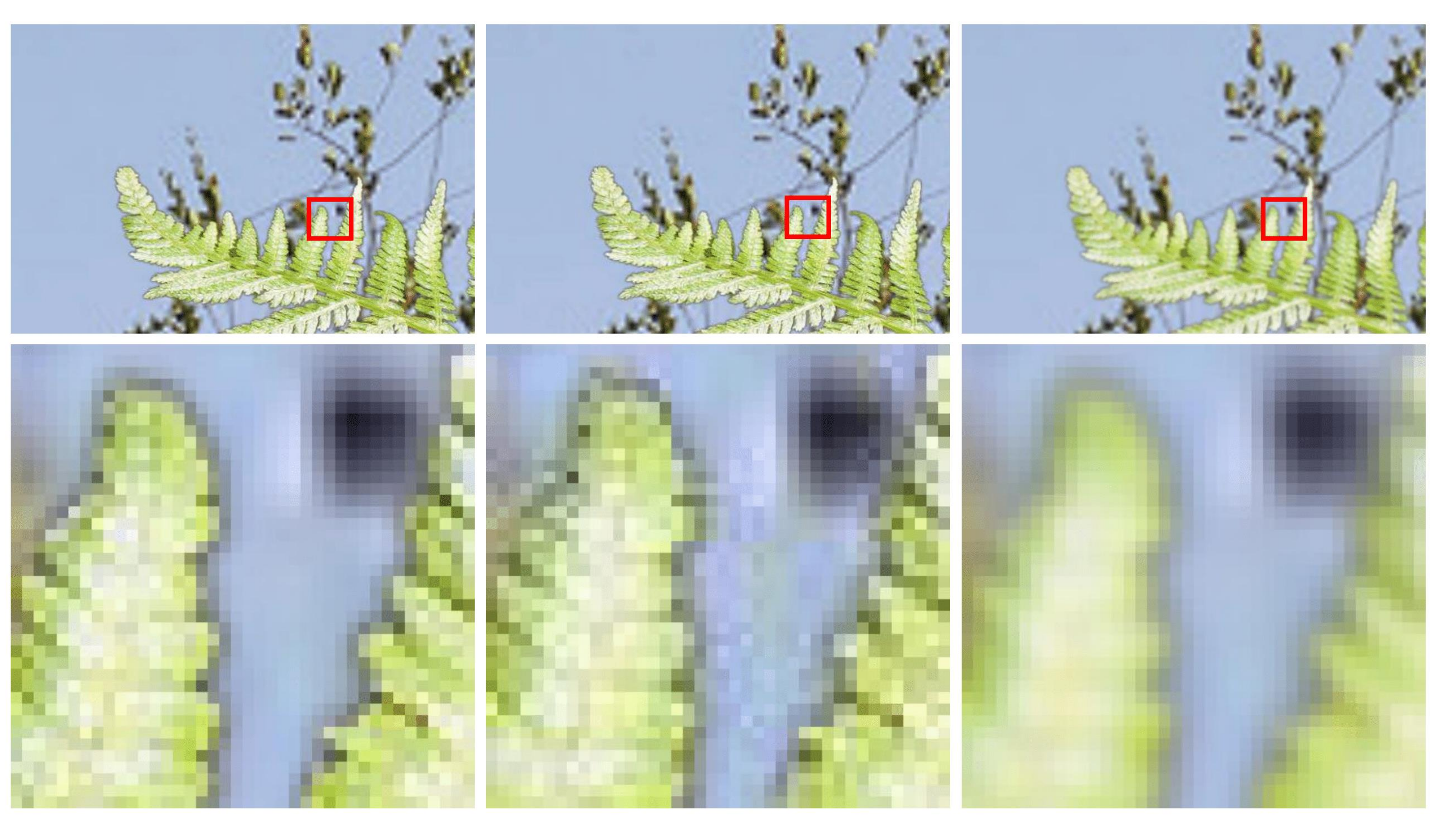
Overview

This paper presents a context-aware natural image matting method for simultaneous foreground and alpha matte estimation. Our method employs a two-encoder network to extract essential information for matting. Particularly, we use a matting encoder to learn local features and a context encoder to obtain more global context information. We concatenate the outputs from these two encoders and feed them into decoders to simultaneously estimate the foreground and alpha map. Both our qualitative and quantitative experiments demonstrate that our method is able to generate state-of-the-art results on challenging real-world examples.

Our main contributions are

- ✓ Integration of local visual features and global context information
- ✓ Combination of the Laplacian and feature loss
- ✓ Various effective data augmentation strategies to help generalizing our method to a wide variety of challenging real-world images

Data Augmentation

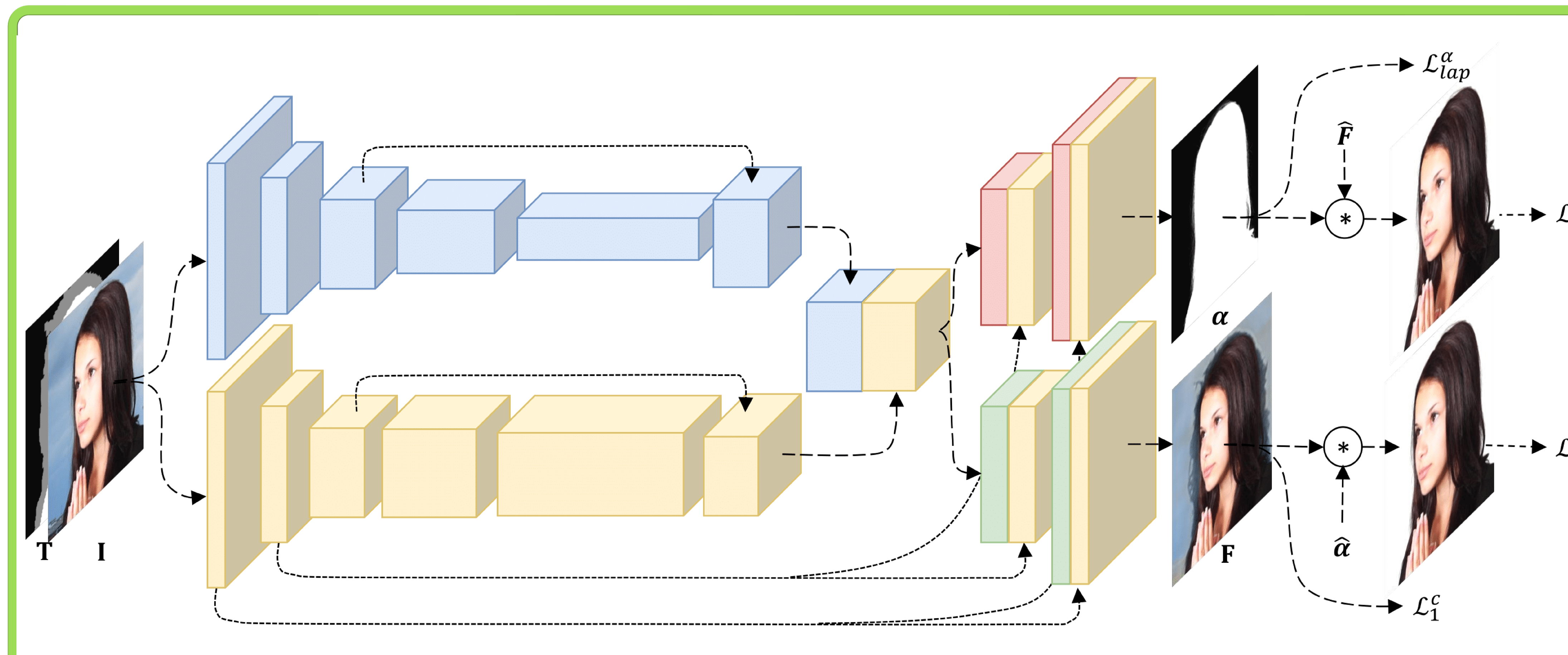


(a) No augmentation (b) Re-JPEGing (c) Gaussian Blur

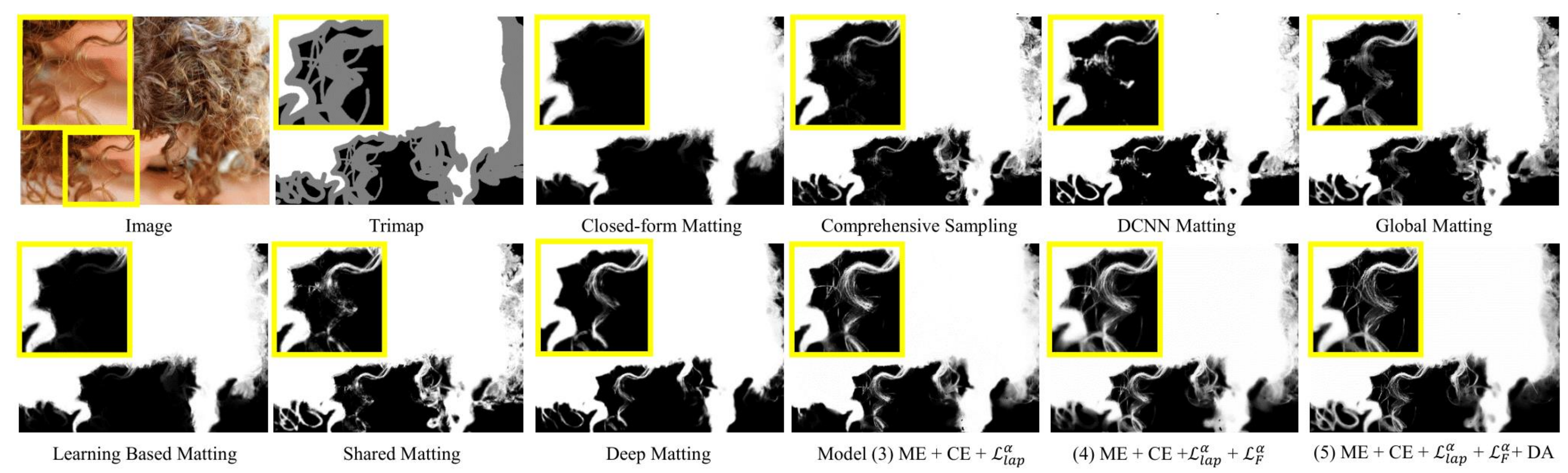
Many subtle artifacts, such as misaligned JPEG blocks, compression quantization artifacts, and resampling artifacts, can sometimes affect the network a lot despite the fact that the images look plausible to the human eyes.

We propose two data augmentation methods that are particularly helpful for our network to generalize to real-world images although it is trained on a synthetic dataset:

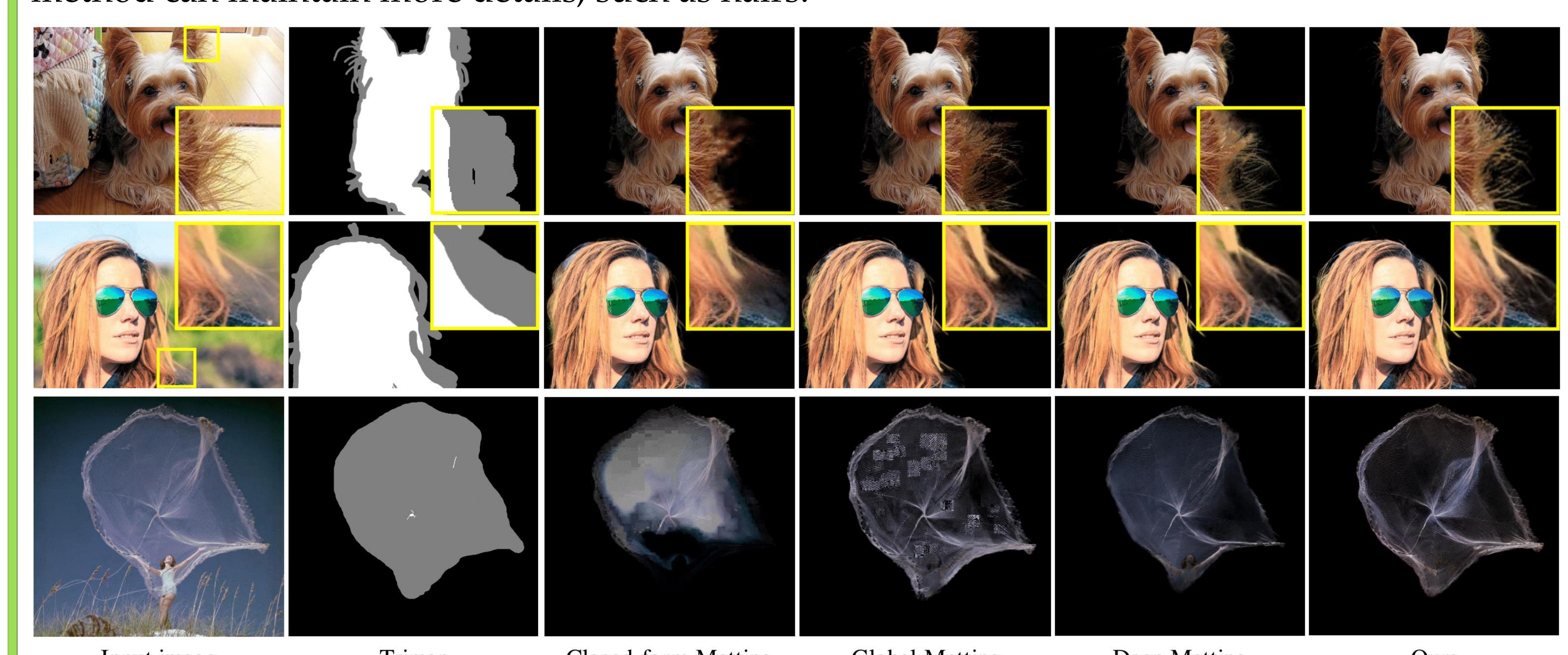
- ✓ **Re-JPEGing:** introduce compression artifacts to the foreground and background.
- ✓ **Gaussian blur:** smoothen high-frequency artifacts.



Comparison



Data augmentation and feature loss can greatly improve the generalization of our trained networks on real-world images. With the feature loss and data augmentation strategies, our method can maintain more details, such as hairs.



Compared to other methods, our method can capture the fine structure when the foreground color and the background color are similar. In the last example, our method not only keeps the delicate edges of the lace, but also is free from the color bleeding problem.

Experiments

Table 1. Alpha map results on the Composition-1K testing set.

Methods	SAD	MSE(10^3)	Grad	Conn
Shared Matting [16]	128.9	91	126.5	135.3
Learning Based Matting [54]	113.9	48	91.6	122.2
Comprehensive Sampling [42]	143.8	71	102.2	142.7
Global Matting [19]	133.6	68	97.6	133.3
Closed-Form Matting [27]	168.1	91	126.9	167.9
KNN Matting [6]	175.4	103	124.1	176.4
DCNN Matting [8]	161.4	87	115.1	161.9
Three-layer Graph [29]	106.4	66	70.0	-
Deep Matting [52]	50.4	14	31.0	50.8
Information-flow Matting [2]	75.4	66	63.0	-
AlphaGan-Best[1] [33]	52.4	30	38.0	-
(1) ME + $\mathcal{L}_{deepmatting}$	49.1	13.4	26.7	49.8
(2) ME + $\mathcal{L}^{\alpha}_{lap}$	43.9	11.8	20.6	41.6
(3) ME + CE + $\mathcal{L}^{\alpha}_{lap}$	**35.8**	**8.2**	17.3	**33.2**
(4) ME + CE + $\mathcal{L}^{\alpha}_{lap}$ + $\mathcal{L}^{\alpha}_{F}$	38.8	9.0	19.0	36.0
(5) ME + CE + $\mathcal{L}^{\alpha}_{lap}$ + $\mathcal{L}^{\alpha}_{F}$ + DA	71.3	23.6	38.8	72.0
(6) ME + CE + $\mathcal{L}^{\alpha}_{lap}$ + $\mathcal{L}^{\alpha}_{F}$ + $\mathcal{L}^{c}_{1}$ + $\mathcal{L}^{c}_{F}$	38.0	8.8	**16.9**	35.4
(7) ME + CE + $\mathcal{L}^{\alpha}_{lap}$ + $\mathcal{L}^{\alpha}_{F}$ + $\mathcal{L}^{c}_{1}$ + $\mathcal{L}^{c}_{F}$ + DA	84.1	29.1	39.2	-
(8) ME + CE + $\mathcal{L}^{\alpha}_{lap}$ + $\mathcal{L}^{\alpha}_{F}$ + $\mathcal{L}^{c}_{1}$ + $\mathcal{L}^{c}_{F}$ + DA - ReJPEGing	55.1	15.5	24.6	54.7
(9) ME + CE + $\mathcal{L}^{\alpha}_{lap}$ + $\mathcal{L}^{\alpha}_{F}$ + $\mathcal{L}^{c}_{1}$ + $\mathcal{L}^{c}_{F}$ + DA - GaussianBlur	69.1	23.5	39.6	69.1

Laplacian loss and our two-encoder structure bring in a major performance improvement. Feature loss and data augmentation are important to obtain high-quality matting results and help networks generalize to real-world images.

Table 2. The foreground result on the Composition-1k dataset.

Methods	SAD	MSE(10^3)
Global Matting [19]	220.39	36.29
Closed-Form Matting [27]	254.15	40.89
KNN Matting [6]	281.92	36.29
(6) ME + CE + $\mathcal{L}^{\alpha}_{lap}$ + $\mathcal{L}^{\alpha}_{F}$ + $\mathcal{L}^{c}_{1}$ + $\mathcal{L}^{c}_{F}$	61.72	3.24
(7) ME + CE + $\mathcal{L}^{\alpha}_{lap}$ + $\mathcal{L}^{\alpha}_{F}$ + $\mathcal{L}^{c}_{1}$ + $\mathcal{L}^{c}_{F}$ + DA	94.41	8.67
(8) ME + CE + $\mathcal{L}^{\alpha}_{lap}$ + $\mathcal{L}^{\alpha}_{F}$ + $\mathcal{L}^{c}_{1}$ + $\mathcal{L}^{c}_{F}$ + DA - ReJPEGing	73.79	4.96
(9) ME + CE + $\mathcal{L}^{\alpha}_{lap}$ + $\mathcal{L}^{\alpha}_{F}$ + $\mathcal{L}^{c}_{1}$ + $\mathcal{L}^{c}_{F}$ + DA - GaussianBlur	85.8	7.10

Compared to three non-deep learning matting methods, our method reduces the errors by a large margin.

Our code, models and results can be downloaded at: https://github.com/hqqxyy/Context-Aware-Matting