Face Parsing via Recurrent Propagation

Sifei Liu[1]
sliu32@ucmerced.edu
Jianping Shi[2]
shijianping@sensetime.com
Ji Liang[3]
liangji20040249@gmail.com
Ming-Hsuan Yang[1]
mhyang@ucmerced.edu

[1] University of California, Merced
[2] SenseTime, Inc.
[3] Momenta, Inc.

1 More Details

More qualitative results. We show more qualitative results in Fig. 1, 2 and 3 for the three datasets as supplements to Fig.3,4 and 5 in the paper.

Model size. The model size with respect to each stage and model is listed in Table 1, which demonstrates that the proposed models are weighted significantly less than the state-of-the-art deep model. In addition, to facilitate efficient performance, the proposed model can be easily ported to mobile devices for real-time applications.

Table 1: Model size (MB) with respect to each model/stage.

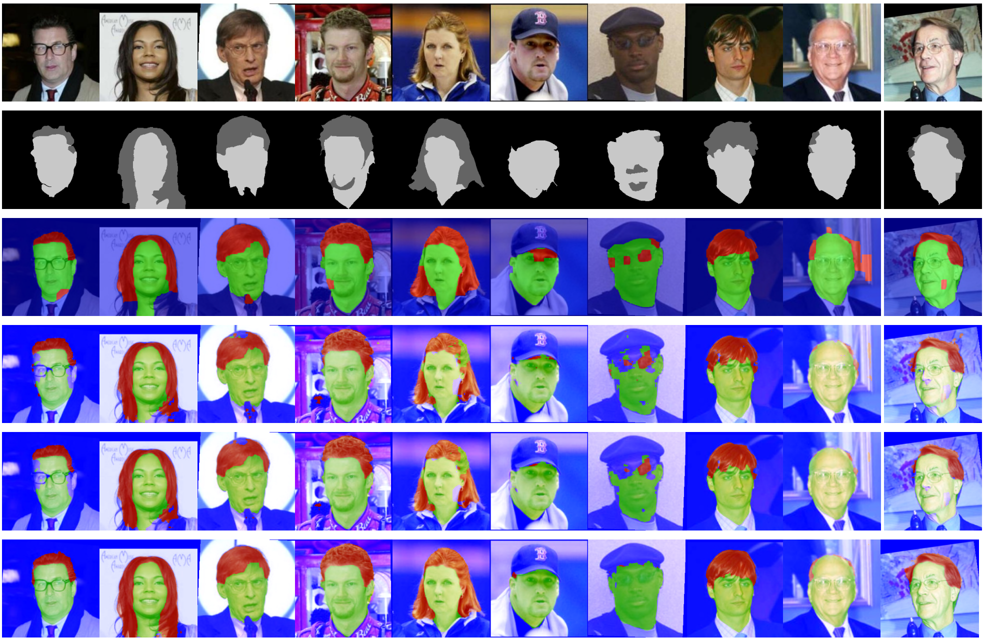
(%)	MO-GC [3]	LFW(stage-1)	HELEN(stage-1&2)	Multi-Face
SIZE	38	0.138	0.797	0.177

2 Applications

Based on the proposed fast face parsing algorithm, automatic facial editing applications can be constructed. We take (a) eyebrow type switching, (b) eyelash editing, (c) lip color adjustment, (d) facial skin beautification, and (e) facial makeup transferring as examples to demonstrate the applications of the proposed algorithm.

2.1 Eyebrow Editing

The application of eyebrow editing can be carried out by removing the existing eyebrow and appending a new type based on the accurate boundaries generated by face parsing. Given the input image in Fig. 4 (a), and eyebrow parsing results in (b), we apply Poisson image editing [4] to remove the existing eyebrow, shown in (c). In particular, we slightly dilate the mask to accommodate for any error caused by an uncovered boundary. New types of eyebrows can be easily generated through alpha blending. Fig. 4 (d) and (e) show two different eyebrow types enabled by our mask.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j)

Figure 1: Face parsing results on the LFW-PL dataset. Top row: input images. Second row: ground-truth annotations. Third row: results from [3]. Forth row: results from CNN-S. Fifth row: results from CNN with dense CRF. Last row: results by RNN-G. Best viewed in color.

2.2 Eyelash Editing

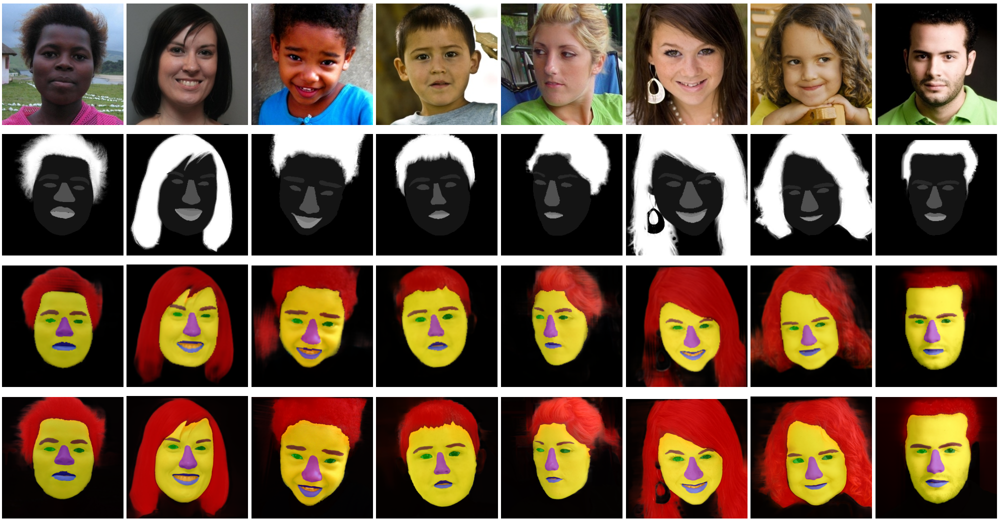
We can append eyelashes on the eye regions based on the accurate eye boundaries generated from parsing results. Given a face image in Fig. 5 as input, we adopt the thin plate splines (TPS) algorithm [1] to map the eyelash template to the eye boundary. The results are obtained through alpha blending.

2.3 Lip Color Adjustments

With the accurate mouth region in Fig. 6, we can directly change the color tone of the lips. Two examples are shown in Fig. 6 (b) and (c).

2.4 Skin Smoothing

Equipped with the facial skin mask, we can smooth the facial area without affecting the details in other components. We first apply the rolling guidance filter (RGF) [6] for edge-preserving smoothing. However, the filtered image does not contain fine details and is visually inauthentic. We denote the edges located by the RFG as edge region, and the other part as the smooth region. To preserve the fine details (e.g., skin poles), we preserve the high-frequency image texture on the corresponding smooth region, which can be obtained through the residue of a standard Gaussian filter ($\delta = 2$). The final result is illustrated in Fig. 7 (c).



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 2: Face parsing results on the HELEN [5] dataset. Top row: input images. Second row: ground-truth annotations. Third row: the results from [3]. Forth row: our results with 11-class pixel-wise parsing. Our model perform better on small facial components such as eye corner or mouth. Best viewed in color.

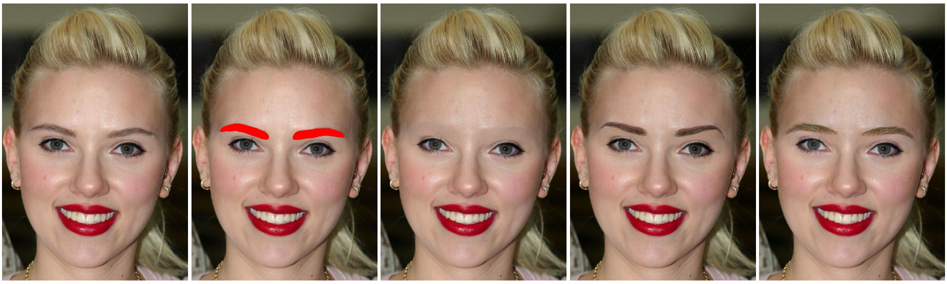
(a) Input Ground truth (c) Our result (d) Input (e) Ground-truth (f) Our result

Figure 3: Parsing results on the Multi-Face dataset. We can successfully process multiple face with our network. Best viewed in color.

2.5 Makeup Transfer

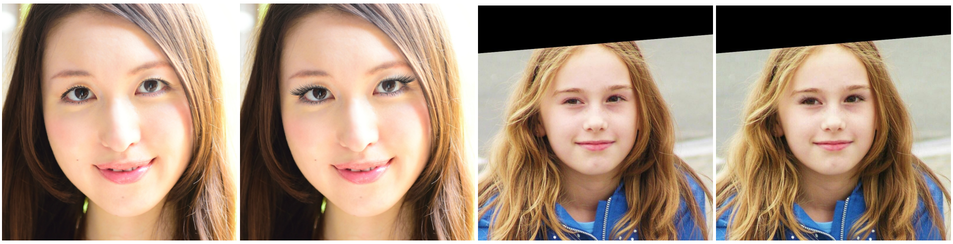
We combine the above-mentioned applications for full makeup editing. Specifically, we can transfer the makeup from a set of reference images with distinct makeup styles (see first row of Fig. 8) to a test image (see second row of Fig. 8). This is carried out by first accurately aligning the face on the reference image to the one on the test image through TPS with the semantic boundaries extracted from the parsing result, and then applying the facial fine texture and color from the corresponding components on the reference face. Given the fine segments, the proposed processing is simpler compare to the work by Guo and Sim [2].

For makeup transfer, we first apply the facial texture and color by linearly blending the gradients of facial skin of two images in the brightness channel, where the weights with respect to each image can be adjusted by user according to the final visual effect. Second, we reconstruct the brightness channel with the new facial gradient map through Poisson image editing [4]. Third, we apply the chromatic channels of the reference image with respect to facial skin and lips. The final virtual makeups, as shown in the second row of Fig. 8, are realistic and accurate on the facial component boundaries due to the fine precision of the face parsing results.



(a) input (b) mask (c) removal (d) new #1 (e) new #2

Figure 4: Swithcing of eyebrow types given the parsed facial components.



input with eyelash input with eyelash

Figure 5: Eyelash editing. Best viewed by zooming in.

(a) input (b) color #1 (c) color #2

Figure 6: Lip Color Adjustments. Best viewed in color.

(a) input (b) parsing skin area (c) smoothed

Figure 7: Smoothing the skin region. Best viewed by zooming in.

Figure 8: Facial makeup transfer. First row: reference model images with specific-stylized makeup. Second row: virtual makeup by applying facial detail and color from the models in the first row. Best viewed in color through zooming in.

References

[1] Fred L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on pattern analysis and machine intelligence*, 11(6):567–585, 1989.

[2] Dong Guo and T. Sim. Digital face makeup by example. In *CVPR*, 2009.

[3] Sifei Liu, Jimei Yang, Chang Huang, and Ming-Hsuan Yang. Multi-objective convolutional learning for face labeling. In *CVPR*, 2015.

[4] Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. In *ACM TOG*, volume 22, pages 313–318, 2003.

[5] Brandon Smith, Li Zhang, Jonathan Brandt, Zhe Lin, and Jianchao Yang. Exemplar-based face parsing. In *CVPR*, 2013.

[6] Qi Zhang, Xiaoyong Shen, Li Xu, and Jiaya Jia. Rolling guidance filter. In *ECCV*, 2014.