Focused and Collaborative Feedback Integration
for Interactive Image Segmentation
- Supplementary Material -

Qiaoqiao Wei    Hui Zhang    Jun-Hai Yong
School of Software, BNRist, Tsinghua University, Beijing, China
wqq18@mails.tsinghua.edu.cn    {huizhang,yongjh}@tsinghua.edu.cn

1. Network Architecture for the Collaborative Feedback Fusion Module

The precise layouts of the convolutional blocks $\Phi_1$, $\Phi_2$, and $\Phi_3$ in the proposed Collaborative Feedback Fusion Module (CFFM) are shown in Fig. 1, Fig. 3, and Fig. 2, respectively. The stride is set to 1 for all convolutional layers, and the padding parameter is set to $(\text{#kernel size} - 1)/2$. Inspired by Xception [3] and FocalClick [2], we apply depthwise separable convolutions in the convolution block $\Phi_2$ to update the feedback and the deep features.

Inspired by Xception [3] and FocalClick [2], we apply depthwise separable convolutions in the convolution block $\Phi_2$ to update the feedback and the deep features.

2. Additional Qualitative Results

In this section, we provide more qualitative results of our method. We visualize the interaction process and the resulting segmentation masks in Fig. 4. The examples are also illustrated in f-BRS [7]. Our method achieved 90% IoU with less than 3 clicks in those examples and has surpassed f-BRS in most cases. Even if our method failed in the first interaction, it regained satisfactory segmentation results in the second interaction (see the fourth example in Fig. 4).

Fig. 5 illustrates some challenging examples of existing interactive image segmentation methods. The challenges include multimodality [4], occlusions, motion blur, thin structures, etc. In some cases, they can be addressed by providing more clicks to specify the position of the target object.

References

Figure 4. Visualization of the interaction process. Throughout this material, the green clicks denote foreground clicks, and the red ones denote background clicks.
Figure 5. A comparison among f-BRS [7], FCA-Net [6], CDNet [1], FocusCut [5], and our method on some of the hard examples for existing interactive image segmentation methods. The number of clicks is 20 for all the visualized examples. The best results are in bold.