1. Implementation Details

We have conducted our experiments with FocalClick SegFormerB3-S2 [3, 9] as the off-the-shelf pre-trained interactive segmentation model. To update student model parameters, we have used Adam optimizer with $\beta_1 = 0.9, \beta_2 = 0.999$, batch size 1 and learning rate $5 \times 10^{-5}$ (for DRIONS-DB [1] and DAVIS [5] learning rate is $10^{-4}$ and $10^{-5}$ correspondingly). We set $\gamma = 2$ in both $L_C$ and $L_I$. In case of continual adaptation a new optimizer is used for each new dataset $D_l$. We set $\alpha = 0.999$ for exponential moving average (EMA) updates of the teacher model parameters. Also, during the adaptation on each dataset $D_l$ parameter change regularizer $L_B$ uses the teacher model parameters obtained after adapting on dataset $D_{l-1}$ as initial parameters.

To verify that the proposed approach does not depend on the off-the-shelf pre-trained interactive segmentation model and deteriorates catastrophic forgetting we use RITM [7] with HRNet-18 backbone [2]. The student model is updated using Adam optimizer with $\beta_1 = 0.9, \beta_2 = 0.999$, batch size 1 and learning rate $5 \times 10^{-6}$ (for DRIONS-DB [1] and DAVIS [5] learning rate is $10^{-5}$ and $10^{-6}$ correspondingly). The teacher update rule is the same.

We use a GeForce RTX 2080 for our experiments.

2. More Qualitative Results

More qualitative results are provided to demonstrate the effectiveness of the proposed method to tackle catastrophic forgetting in continual adaptation. Figures 1 and 2 compare adaptation results of the baseline and our methods to show that it takes fewer click to achieve a higher IOU from our method. Figures 3, 4, 5 and 6 demonstrate the effectiveness of the proposed method for adaptation on new datasets comparing with SOTA interactive segmentation method FocalClick [3].

References

Figure 1. Comparison between the baseline and our method. Images are taken from GrabCut dataset. To illustrate the effectiveness of the proposed method to decrease catastrophic forgetting, the baseline and our method have been continuously adapted on DRIONS-DB [1] → Rooftop [8] → GrabCut [6]. Green and red points represent positive and negative clicks correspondingly. Blue horizontal boxes include results of the baseline method. Green horizontal boxes include results for the proposed teacher-student approach. IOU@Number of Clicks is reported. Red vertical boxes include the image and the ground truth.
Figure 2. Comparison between the baseline and our method. Images are taken from Berkeley dataset. To illustrate the effectiveness of the proposed method to decrease catastrophic forgetting, the baseline and our method have been continuously adapted on DRIONS-DB [1] → Rooftop [8] → Berkeley [4]. Green and red points represent positive and negative clicks correspondingly. Blue horizontal boxes include results of the baseline method. Green horizontal boxes include results for the proposed teacher-student approach. IOU@Number of Click is reported. Red vertical boxes include the image and the ground truth.
Figure 3. Comparison between FocalClick [3] and our method. Images are taken from \textit{DRIONS – DB} dataset. Green and red points represent positive and negative clicks correspondingly. Blue horizontal boxes include results of FocalClick. Green horizontal boxes include results for the proposed teacher-student approach. IOU@Number of Click is reported. Red vertical boxes include the image and the ground truth.
Figure 4. Comparison between FocalClick [3] and our method. Images are taken from Rooftop dataset. Green and red points represent positive and negative clicks correspondingly. Blue horizontal boxes include results of FocalClick. Green horizontal boxes include results for the proposed teacher-student approach. IOU@Number of Click is reported. Red vertical boxes include the image and the ground truth.
Figure 5. Comparison between FocalClick [3] and our method. Images are taken from Heart dataset. Green and red points represent positive and negative clicks correspondingly. Blue horizontal boxes include results of FocalClick. Green horizontal boxes include results for the proposed teacher-student approach. IOU@Number of Click is reported. Red vertical boxes include the image and the ground truth.
Figure 6. Comparison between FocalClick [3] and our method. Images are taken from Spleen dataset. Green and red points represent positive and negative clicks correspondingly. Blue horizontal boxes include results of FocalClick. Green horizontal boxes include results for the proposed teacher-student approach. IOU@Number of Click is reported. Red vertical boxes include the image and the ground truth.