Supplementary Material
Efficient and Generic Interactive Segmentation Framework

1 Ablation Study

We conducted studies to determine the effect of various hyper-parameters for our method.

1. **Optimum iterations:** We obtained the optimum number of iterations of back-propagation for obtaining a dice coefficient of 0.95 for each segmentation network. As seen in the Fig. 1, the optimum number of iterations was 80, 100, 130, 140 and 130 for the 2018 Data Science Bowl [5], CoNSeP [8], LiTS Challenge [6], CHAOS dataset [13] and BraTS' 15 [20] segmentation respectively.

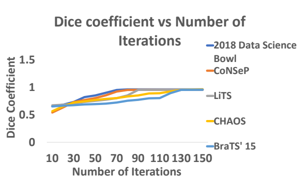

Fig. 1: Optimum Iterations Determination

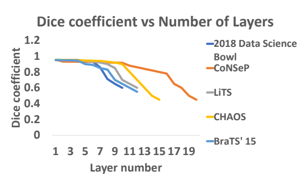

Fig. 2: Optimum Layer Determination

Table 1: Number of Mouse Clicks(N), User Time **(UT)** and Machine Time **(MT)** for various user inputs (in mins).

Method	**N**	**UT**	**MT**
Point	100	13	16
Box	34	10	15
Scribble	10	5	10

2. **Optimum layer:** Once, the optimum number of iterations are determined, our next step is to determine the optimum layer number upto which back-propagation needs to be performed for each segmentation network. We observe that we obtain the best possible dice coefficient for 4, 6, 4, 3 and 5 layers for the 2018 Data Science Bowl [5], CoNSeP [8], LiTS Challenge [8], CHAOS dataset [13] and BraTS' 15 [20] segmentation respectively as seen in the right panel of Fig. 2.
3. **Optimum user input type:** Our method has the unique and remarkable capability of being able to work with any type of user input such as points, boxes and scribbles. We first performed experiments to determine the most suitable user input modality for segmentation correction. We found that scribbles required the least number of user interactions (30% lesser mouse clicks), as well as user and machine time (Table 1). Hence, the experiments in the paper were done with scribbles only.
4. **Optimum scribble length:** We also evaluated the effect of scribble length while using our method. We observed that without region growing, we needed more user interactions to correct the segmentation as the scribble length

reduced. However, with region growing, there was hardly any change in the number of user interactions required as seen in Fig. 3 (obtained for LiTS challenge, similar behavior observed for other datasets, but were not able to provide owing to space restrictions).

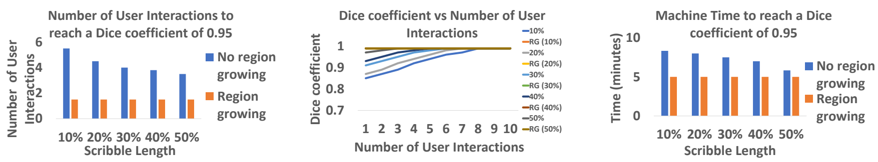

Fig. 3: Effect of scribble length on the increase in the number of user interactions (RG - Region Growing).

2 Evaluation of our method by medical expert

We evaluated our interactive segmentation method with the help of medical experts. We have provided the user interaction time and machine time required for the 2018 Data Science Bowl (2018 DSB), LiTS and BraTS' 15 challenges here. It was possible to obtain a reduction in user annotation time as well as machine time as seen in Table 2.

Table 2: User Interaction Time (**UT**) and Machine Time (**MT**) in minutes for separating structures by a medical expert (**F:** Full Human Annotation, **R:** Our method - Region Growing, **N:** Our Method - No Region Growing. All the semi-automated methods [19, 21, 25, 28, 32, 33] were applied till a dice coefficient of 0.95 was reached.

Dataset	**User Interaction Time**									**Machine Time**							
	F	**R**	**N**	[25]	[28]	[19]	[21]	[32]	[33]	**R**	**N**	[25]	[28]	[19]	[21]	[32]	[33]
2018 DSB	55	4	6	15	14	14	-	-	-	7	12	19	18	15	-	-	-
LiTS	100	6	7	-	-	-	14	15	16	9	10	-	-	-	12	14	15
BraTS' 15	150	9	12	-	-	-	65	75	90	50	80	-	-	-	120	126	145