Supplementary Material

Anonymous

		Semantic				Object		
Method	Staining	Dice	Accuracy	Precision	Recall	AJI	Dice	Hausdorff
Proposed	Ki67	70.27	79.86	74.46	66.60	41.91	54.43	5.77
Removing \mathcal{L}_C	Ki67	70.34	80.66	71.23	69.64	40.58	52.24	6.11
Imposing $\mathbf{r}{=}1$	Ki67	68.53	78.40	75.46	62.94	40.50	54.17	5.73
Removing $\mathcal{L}_{\mathcal{R}}$	Ki67	18.46	50.06	22.46	15.74	7.67	5.64	12.43
Proposed	HER2	62.59	77.05	70.67	56.51	39.29	60.01	4.68
Removing \mathcal{L}_C	HER2	60.53	77.20	63.92	57.61	37.09	44.18	5.48
Imposing $r=1$	HER2	59.13	73.30	77.35	48.02	35.42	49.17	4.64
Removing $\mathcal{L}_{\mathbf{R}}$	HER2	7.35	50.01	9.34	6.10	2.74	2.83	9.45

Table 1. Results of the ablation study. We computed the performances on both DeepLIIF (Ki67) and Warwick datasets (HER2), removing reconstruction loss (\mathcal{L}_R) , consistency loss (\mathcal{L}_C) and imposing a sharpening factor **r** of 1 instead of 60.



Fig. 1. Examples of prediction from the benchmarked methods on the DeepLIIF dataset. Each color represents a different segmented instance.



Fig. 2. Examples of prediction from the benchmarked methods on the Warwick dataset. Each color represents a different segmented instance.



Fig. 3. Examples of prediction from the benchmarked methods on the BCDataset dataset. Each color represents a different segmented instance. Since masks were not available for BCDataset, the center of each nuclei has been higlighted on the input image. Both proposed method and fully supervised Unet were trained on DeepLIIF training set.