Supplementary Material of ModelMix

Ke Zhang and Vishal M. Patel

Department of Electrical and Computer Engineering, The Johns Hopkins University

0.1 Supervision sensitivity

We verified the supervised sensitivity of the proposed ModelMix by training the model using 20% scribble annotated images and 80% unlabeled images, as well as 100% scribble labeled images. Table I and Table II summarizes the experiment results on MyoPS and MSCMRseg dataset.

Irregular pathology segmentation of MyoPS: One can observe that our ModelMix consistently outperforms semi-supervised benchmarks with both 20% and 100% scribble annotations on MyoPS datasets. As shown in Table I, by mixing with the model parameters of MSCMR dataset, our ModelMix obtains remarkable performance gain of 26.9% and 25.0% Dice with 20% and 100% scribble annotations, respectively. These results demonstrate the robustness of the proposed ModelMix in the situation of different supervision amount.

Table I: Supervision sensitivity: Irregular pathology segmentation on MyoPS dataset.

Methods	Ratio	20% scribbles			100% scribbles		
		Scar	Edema	Avg	Scar	Edema	Avg
PCE	×	.242±.170	.122±.077	.182±.144	.504±.213	.057±.022	.281±.271
CVIR	✓	.288±.191	.085±.034	.186±.170	.505±.214	.080±.031	.293±.263
nnPU	✓	.290±.166	.236±.078	.263±.131	.530±.241	.085±.035	.308±.282
w/ MSCMR	×	**.488±.263**	**.575±.147**	**.532±.215**	**.541±.268**	**.575±.214**	**.558±.240**
FullSup-UNet	-	.423±.253	.445±.149	.434±.205	.537±.232	.659±.135	.633±.202
FullSup-nnUNet	-	.496±.252	.563±.141	.529±.204	.610±.169	.651±.246	.630±.209

Regular ventrical segmentation of MSCMRseg: By changing the supervision amount from 20% to 100%, we also verify the supervision sensitivity of proposed ModelMix on MSCMRseg dataset. As illustrated in Table II, the proposed ModelMix succeeds in both scenarios, with evident improvement in average Dice by 9.7% and 2.7%, respectively.

0.2 The details of compared methods

We compare our ModelMix to ten methods, the detailed information are summarized as follows:

1) *PCE:* UNet backbone trained with the PCE loss ($\mathcal{L}_{\text{pce}}$).

Table II: Supervision sensitivity: regular ventrical segmentation on MSCMRseg dataset.

Methods	20% scribbles				100% scribbles			
	LV	MYO	RV	Avg	LV	MYO	RV	Avg
Mixup	.440±.102	.310±.127	.021±.013	.257±.200	.483±.09	.466±.080	.455±.134	.468±.102
Cutout	.315±.103	.307±.153	.166±.110	.263±.139	468±.076	.642±.132	.694±.146	.602±.154
CycleMix	.517±.086	.421±.108	.007±.007	.315±.237	.872±.060	.734±.048	.787±.073	.798±.083
ShapePU	.758±.191	.567±.168	.059±.026	.461±.331	.880±.046	.785±.080	.833±.087	.833±.082
WSL4	.809±.079	.653±.109	.599±.261	.687±.191	.902±.040	.815±.033	.828±.101	.848±.076
w/ MyoPS	**.875±.077**	**.754±.079**	**.722±.201**	**.784±.145**	**.919±.036**	**.842±.035**	**.865±.063**	**.875±.056**
FullSup-UNet	.775±.158	.604±.147	.572±.207	.651±.191	.917±.046	.813±.058	.750±.162	.827±.122
FullSup-nnUNet	.885±.085	.757±.147	.757±.201	.799±.160	.909±.049	.880±.027	.902±.047	.907±.044

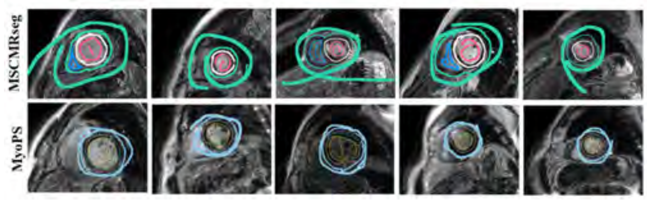

Fig. II: The visualization of scribbles from MSCMRseg and MyoPS dataset.

2) *Mixup:* UNet backbone trained with mixup augmentation and PCE loss.
3) *Cutout:* UNet backbone trained with cutout augmentation and PCE loss.
4) *WSL4*: We adopted code of WSL4 released by the authors via https://github.com/HiLab-git/PyMIC/blob/master/pymic.
5) *CycleMix*: We use the code released via https://github.com/BWGZK/CycleMix.
6) *ShapePU*: We leverage the code released via https://github.com/BWGZK/ShapePU.
7) *CVIR*: We use the released code from https://github.com/acmi-lab/PU_learning to implement the technique of Conditional Value Ignoring Risk (CVIR). Since CVIR is developed on the condition that mixture ratio is known, we provide it with ground truth ratio for model training. Given that CVIR is proposed for classification tasks, we apply it to each individual pixel classification task to achieve the pixel-level segmentation.
8) *nnPU*: The reproduced code via https://github.com/kiryor/nnPUlearning is adopted for our implementation of PU loss. Since nnPU is designed for classification tasks, we adapt it to segmentation tasks by applying it to each pixel.
9) *FullSup-UNet*: The fully supervised UNet trained with the cross-entropy loss calculated with full annotations.
10) *FullSup-nnUNet*: We adopot the code released via https://github.com/MIC-DKFZ/nnUNet.

0.3 Visualization of scribbles:

The typical scribble annotations of MSCMRseg and MyoPS dataset are visualized in Figure II.