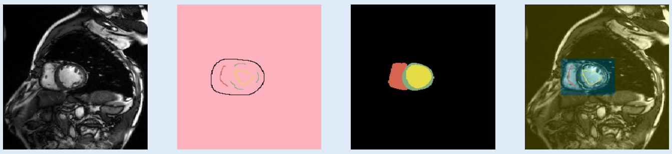
EFFDNet: A Scribble-Supervised Medical Image Segmentation Method with Enhanced Foreground Feature Discrimination

Jinhua Liu, Shu Yun Tan, Xulei Yang, Yanwu Xu, Si Yong Yeo*

*Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore; * Corresponding Author*

Motivation



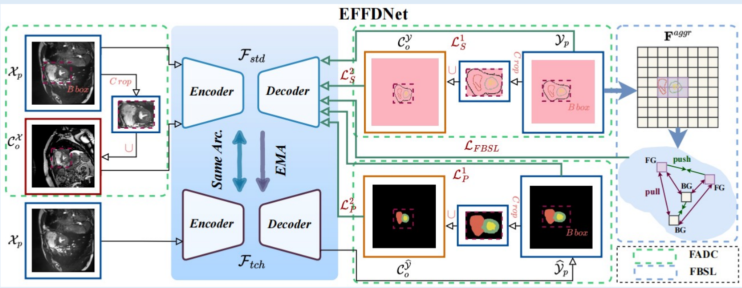
(a) Image (b) Scribble (c) Segmentation (d) Foreground-background semantics

- Merely treating scribble annotations as initial seed regions or sparse pixel-level annotations, without exploiting the underlying rich foreground-background semantics information, often fails to provide sufficient details for networks to effectively separate foreground anatomical structures from the background.
- Existing methods often fail to extract discriminative features to accurately identify the morphology of foreground anatomical structures.

Main idea

- We introduce the **Foreground-Background Separation Loss (FBSL)** to better utilize the foreground-background semantics contained in scribble annotations, thereby enhancing foreground discrimination in the feature space.
- We propose the **Foreground Augmentation with Diverse Context (FADC)** mechanism to enhance the network's foreground sensitivity and model generalization by using a new foreground augmentation strategy.

Method



- The pipeline of the Enhanced Foreground Feature Discrimination Network (EFFDNet). Here, "FG" denotes the foreground, and "BG" denotes the background. The "push" brings tensors closer in feature space, while "pull" distances them. $\mathcal{L}_S^{1,2}$ and $\mathcal{L}_P^{1,2}$ denote the supervised and unsupervised losses for original and augmented data, respectively, $\mathcal{L}_{FBSL}$ is the Foreground-Background Separation Loss.

Foreground-Background Separation Loss

- Local Region Aggregation :

$$F_i^{aggr}(r,s)=\frac{1}{\frac{H}{K}\cdot\frac{W}{K}}\sum_{m=0}^{\frac{H}{K}-1}\sum_{n=0}^{\frac{W}{K}-1}F_i^{(r\cdot\frac{H}{K}+m,\, s\cdot\frac{W}{K}+n)}$$

- Assigning Corresponding Region Labels R :

$$R_i(r,s)=\begin{cases}1 & \text{if } \sum_{m=0}^{\frac{H}{K}-1}\sum_{n=0}^{\frac{W}{K}-1}\mathcal{Y}_i^{(r\cdot\frac{H}{K}+m,\, s\cdot\frac{W}{K}+n)}\mathbb{I}_C>0,\\ 0 & \text{else}.\end{cases}$$

- Foreground-Background Separation Loss $\mathcal{L}_{FBSL}$:

$$\mathcal{L}_{FBSL}=-\frac{\sum_{e=0}^{\frac{H}{K}N}\sum_{f=0}^{\frac{W}{K}N}\sum_{u=1}^{\mathcal{M}_{(e,f)}}\log\left[\frac{\exp\langle z_{(e,f)},z^{+}_{(e,f,u)}\rangle/\tau}{\exp\langle z_{(e,f)},z^{+}_{(e,f,u)}\rangle/\tau+\sum_{v=1}^{\mathcal{N}_{(e,f)}}\exp\langle z_{(e,f)},z^{-}_{(e,f,v)}\rangle/\tau}\right]}{\frac{H}{K}N\cdot\frac{W}{K}N\cdot\mathcal{M}_{(e,f)}}$$

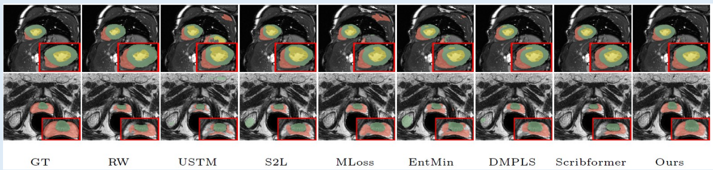
Foreground Augmentation with Diverse Context

- Foreground Augmentation :

$$\mathcal{C}_o^{\mathcal{X}}=\mathcal{X}_o\setminus Bbox(\mathcal{X}_o)\cup Crop(\mathcal{X}_p,Bbox(\mathcal{X}_p));$$
$$\mathcal{C}_o^{\mathcal{Y}}=\mathcal{Y}_o\setminus Bbox(\mathcal{Y}_o)\cup Crop(\mathcal{Y}_p,Bbox(\mathcal{Y}_p));$$
$$\mathcal{C}_o^{\hat{\mathcal{Y}}}=\hat{\mathcal{Y}}_o\setminus Bbox(\hat{\mathcal{Y}}_o)\cup Crop(\hat{\mathcal{Y}}_p,Bbox(\hat{\mathcal{Y}}_p)).$$

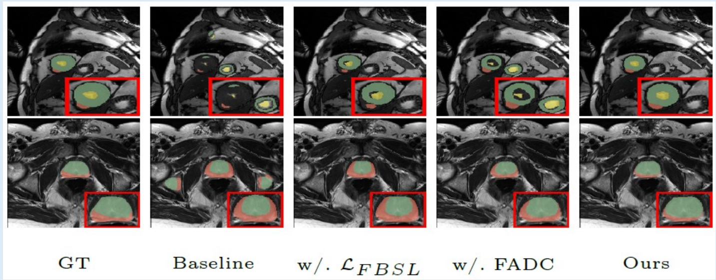
Experiments and Results

Comparison Visualization with State-of-the-art Models



- Improved foreground shape preservation: Our method achieves superior segmentation by enhancing foreground discrimination ability, leading to more accurate structures boundaries, complete foreground structures, and reduced over-segmentation.

Ablation Visualization



- Improved foreground discrimination ability: The FBSL enhances foreground-background separation, reducing over-segmentation and improving the accuracy of segmentation morphology. The FADC further strengthens foreground discrimination and generalization by introducing images of diverse augmented foreground, reducing inter-class confusion.

Quantitative evaluation

Type	Method	RV (ACDC)	Myo (ACDC)	LV (ACDC)	PZ (NCI-ISBI)	CG (NCI-ISBI)
	pCE	56.44 (11.51)*	56.63 (3.60)*	69.00 (10.12)*	22.03 (6.69)*	45.33 (4.61)*
	RW	81.54 (4.29)*	71.02 (4.08)*	84.75 (3.37)*	72.72 (5.12)	78.94 (3.58)*
	USTM	79.27 (4.18)*	74.07 (3.40)*	76.60 (7.85)*	65.57 (3.62)*	36.20 (9.82)*
	S2L	83.68 (2.54)*	81.87 (2.83)*	87.44 (6.67)	69.79 (4.57)*	55.66 (4.82)*
	MLoss	83.37 (2.56)*	82.56 (2.55)*	90.68 (4.01)	70.87 (4.17)*	81.39 (1.58)*
WSL	EntMin	83.21 (3.03)*	80.99 (2.82)*	88.73 (4.57)*	59.19 (4.48)*	42.74 (5.45)*
	DMPLS	86.22 (2.71)	83.82 (2.38)*	91.46 (3.27)	69.14 (4.18)*	56.43 (7.94)*
	Scribformer	86.24 (3.11)	84.01 (2.13)	91.07 (3.63)	69.37 (2.49)*	74.20 (3.54)*
	Baseline	83.13 (3.92)	80.16 (2.37)	87.71 (4.39)	65.91 (9.17)	65.92 (15.11)
	w/. LBSL	85.90 (2.84)	84.60 (1.61)	**92.17 (3.40)**	72.39 (4.24)	85.90 (0.28)
	w/. FADC	85.88 (3.12)	82.55 (2.97)	89.49 (3.48)	71.15 (7.28)	82.18 (1.66)
	Ours	**86.54 (2.47)**	**85.67 (2.71)**	92.15 (2.70)	**73.00 (3.95)**	**86.37 (0.62)**
FSL	FullSup	89.49 (1.90)	89.07 (1.88)	93.95 (2.76)	77.23 (3.96)	87.90 (0.73)

- Improved segmentation accuracy: The FBSL and FADC strategies respectively improve the network's capability of foreground–background separation and enhance foreground diversity, thereby yielding substantial performance gains and ultimately establishing state-of-the-art results on the ACDC and NCI-ISBI datasets.

Conclusion

- We propose EFFDNet, which enhances foreground discrimination ability. By analyzing the foreground-background semantics in scribbles, we introduce a novel loss function, FBSL, to improve the network's ability to distinguish between foreground and background regions in feature space. Additionally, we design a new data augmentation mechanism, FADC, which enhances the network's sensitivity of foreground regions and mitigates overfitting.