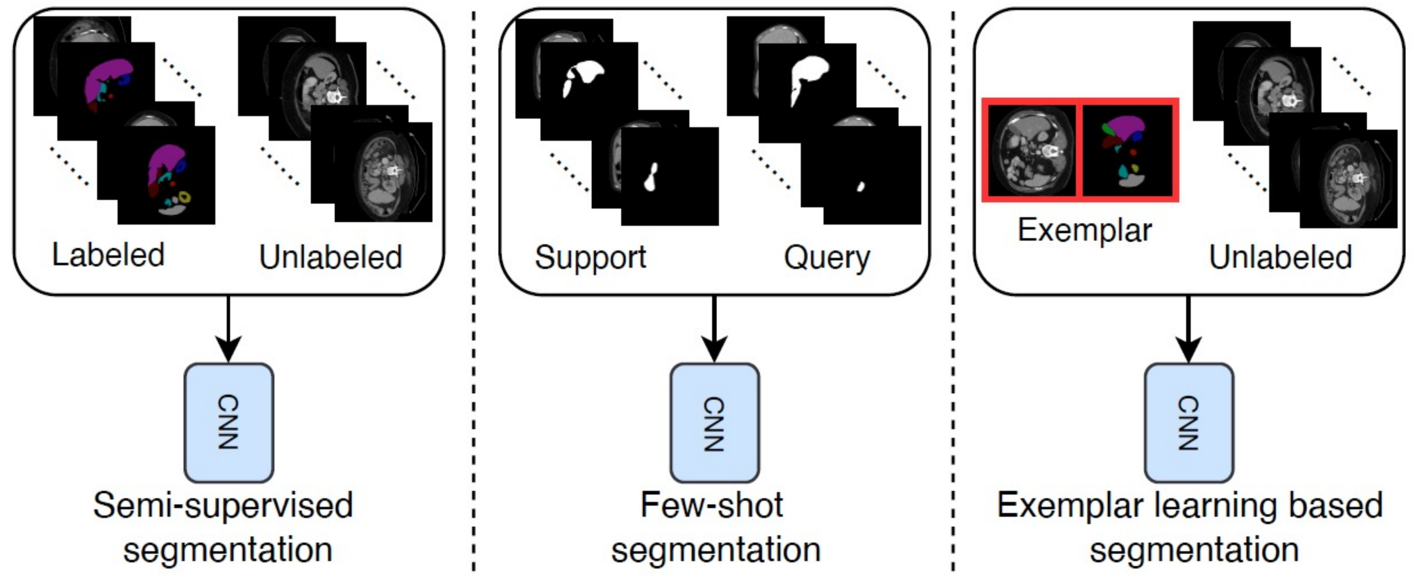
Exemplar Learning for Medical Image Segmentation

Qing En[1], Yuhong Guo[1,2]. [1]Carleton University, [2]Canada CIFAR AI Chair, Amii

Exemplar Learning

➢ We propose a novel learning scenario, Exemplar Learning (EL), to explore automated learning processes for medical image segmentation from a single annotated image example

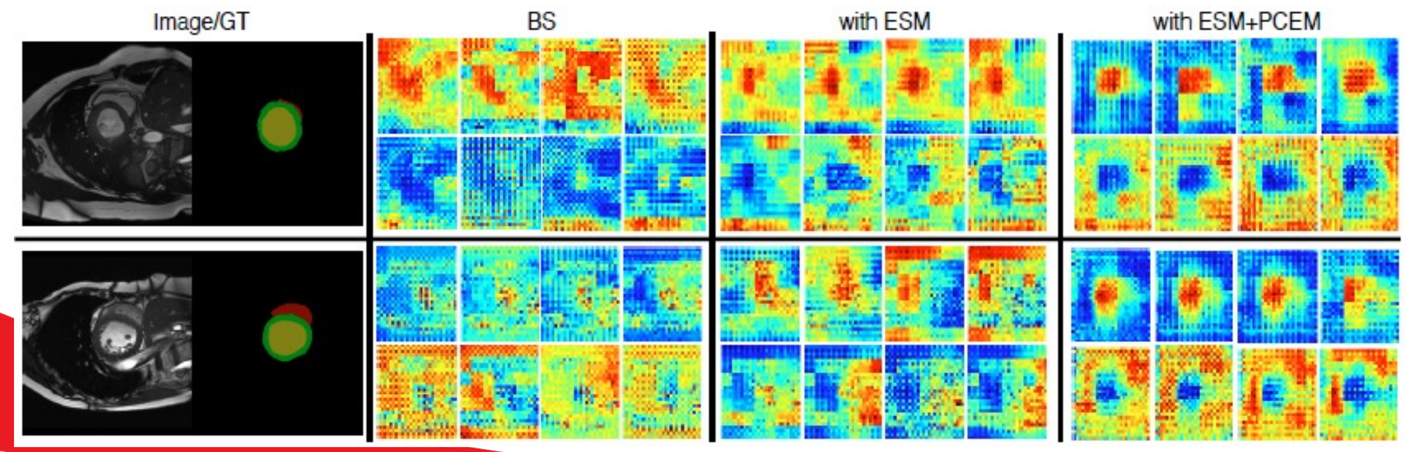


Contribution

➢ We propose a novel learning scenario, Exemplar Learning, which explores medical image segmentation from a single annotated image

➢ We propose a novel ELSNet framework to segment medical images in the EL scenario by creating exemplar-based data, learning pixel-prototype based contrastive embedding and exploring unlabeled data with pseudo-labels

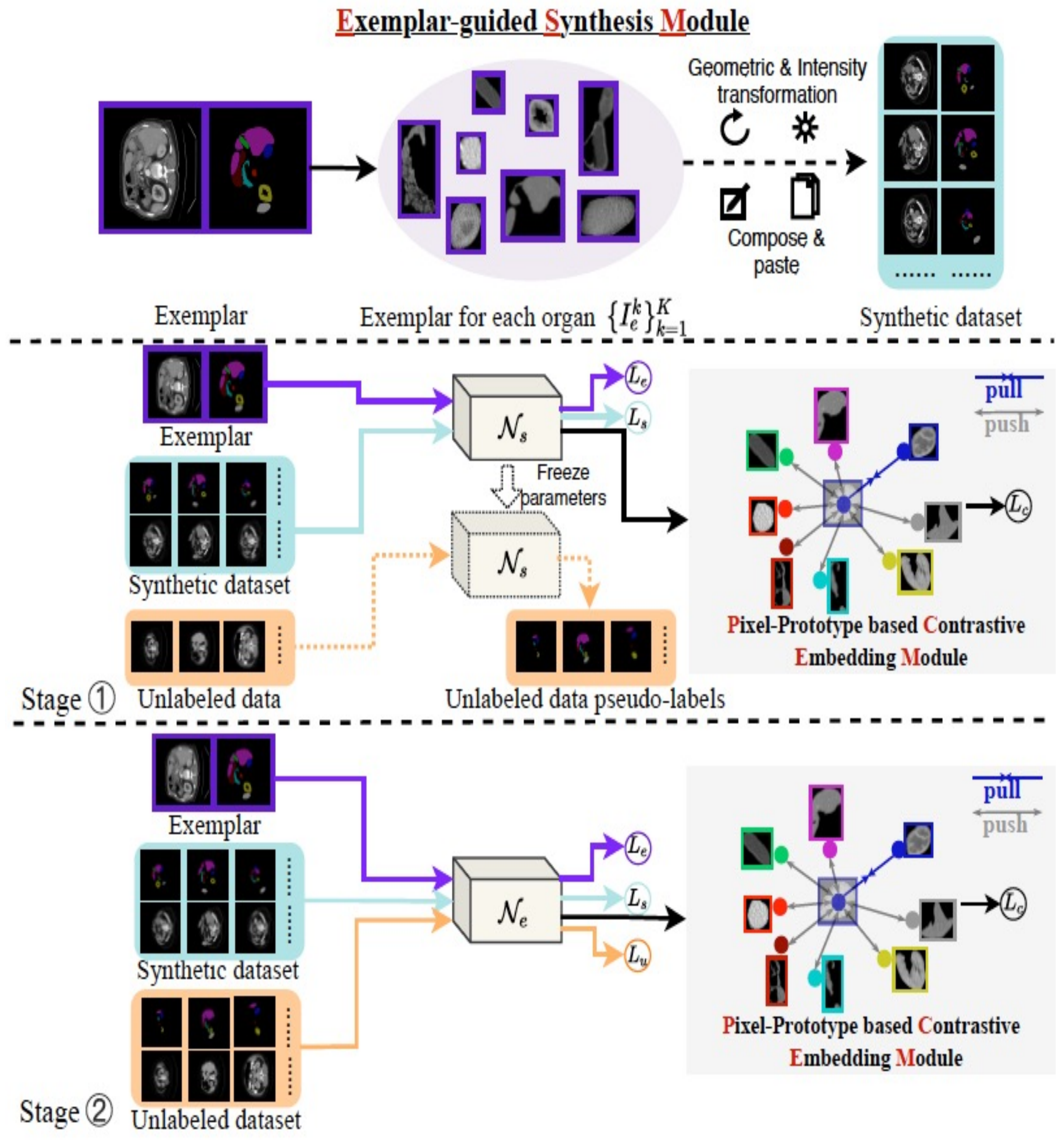
➢ Experimental results on two medical image segmentation datasets show that the proposed ELSNet can effectively perform the medical semantic segmentation task



Exemplar Learning-based Synthesis Net

➢ The Exemplar-guided Synthesis Module (ESM) is proposed to enrich and diversify the training set by synthesizing annotated samples from the given exemplar

➢ The Pixel-prototype based Contrastive Embedding Module (PCEM) is proposed to enhance the discriminative capacity of the base segmentation model via contrastive self-supervised learning



ESM and PCEM

➢ Create the synthetic dataset by ESM

$$(I_s, Y_s) = \Omega(\mathcal{T}_g(\mathcal{T}_i(I_e)), \mathcal{T}_g(Y_e), \mathcal{T}_g(\mathcal{T}_i(I_b)))$$

➢ Train the synthetic segmentation network by PCEM

$$v_k = \frac{\sum_{(i,j)} X^{(i,j)} \mathbb{1}[\hat{Y}_x^{(k,i,j)} \neq 0]}{\sum_{(i,j)} \mathbb{1}[\hat{Y}_x^{(k,i,j)} \neq 0]}$$

$$L_c = -\sum_{n=1}^{N}\sum_{k=1}^{K} \log \frac{exp(v_k^n \cdot v_k^m/\tau)}{exp(v_k^n \cdot v_k^m/\tau) + \sum_{j\neq k}\sum_{i\neq n} exp(v_k^n \cdot v_j^i/\tau)}$$

➢ Generate the pseudo-labels of unlabeled data

$$D_U = \{(I_u^t, \mathcal{N}_s(I_u^t))\}_{t=1}^{T}$$

Two-Stage Training

L_e: Exemplar loss
L_s: Synthetic segmentation loss
L_c: Prototype contrastive loss
L_u: unlabeled segmentation loss

$$L_e = \mathcal{L}_{seg}(\hat{Y}_e, Y_e),\ L_s = \mathcal{L}_{seg}(\hat{Y}_s, Y_s) \text{ and } L_u = \mathcal{L}_{seg}(\hat{Y}_u, Y_u)$$

$$\mathcal{L}_{seg}(\hat{Y}, Y) = 0.5 * l_{ce}(\hat{Y}, Y) + 0.5 * l_{dice}(\hat{Y}, Y)$$

➢ The exemplar and the synthetic dataset are used as training data to train the synthetic segmentation network

$$L_{s1} = L_e + \lambda_s L_s + \lambda_c L_c$$

➢ Make use of the exemplar, synthetic dataset and the unlabeled dataset to train the exemplar learning segmentation network by PCEM

$$L_{s2} = \lambda_e L_e + \lambda_s L_s + \lambda_c L_c + \lambda_u L_u$$

Comparison with the State-of-the-Art Methods

- Datasets：ACDC and Synapse dataset
- Evaluation metrics: the Dice Similarity Coefficient (DSC) and the 95% Hausdorff Distance (HD95)

➢ Comparison results on the ACDC Dataset

Method	DSC.Avg↑	RV	Myo	LV	HD95.Avg↓	RV	Myo	LV
UNet[ ]	0.142	0.140	0.112	0.174	43.30	63.76	35.60	30.80
MT-UNet[ ]	0.142	0.119	0.126	0.182	74.20	83.91	61.48	77.22
MLDS[ ]	0.189	0.144	0.165	0.258	50.03	72.13	30.20	47.77
ELSNet	**0.410**	**0.293**	**0.374**	**0.563**	**26.64**	**47.63**	**16.58**	**15.73**

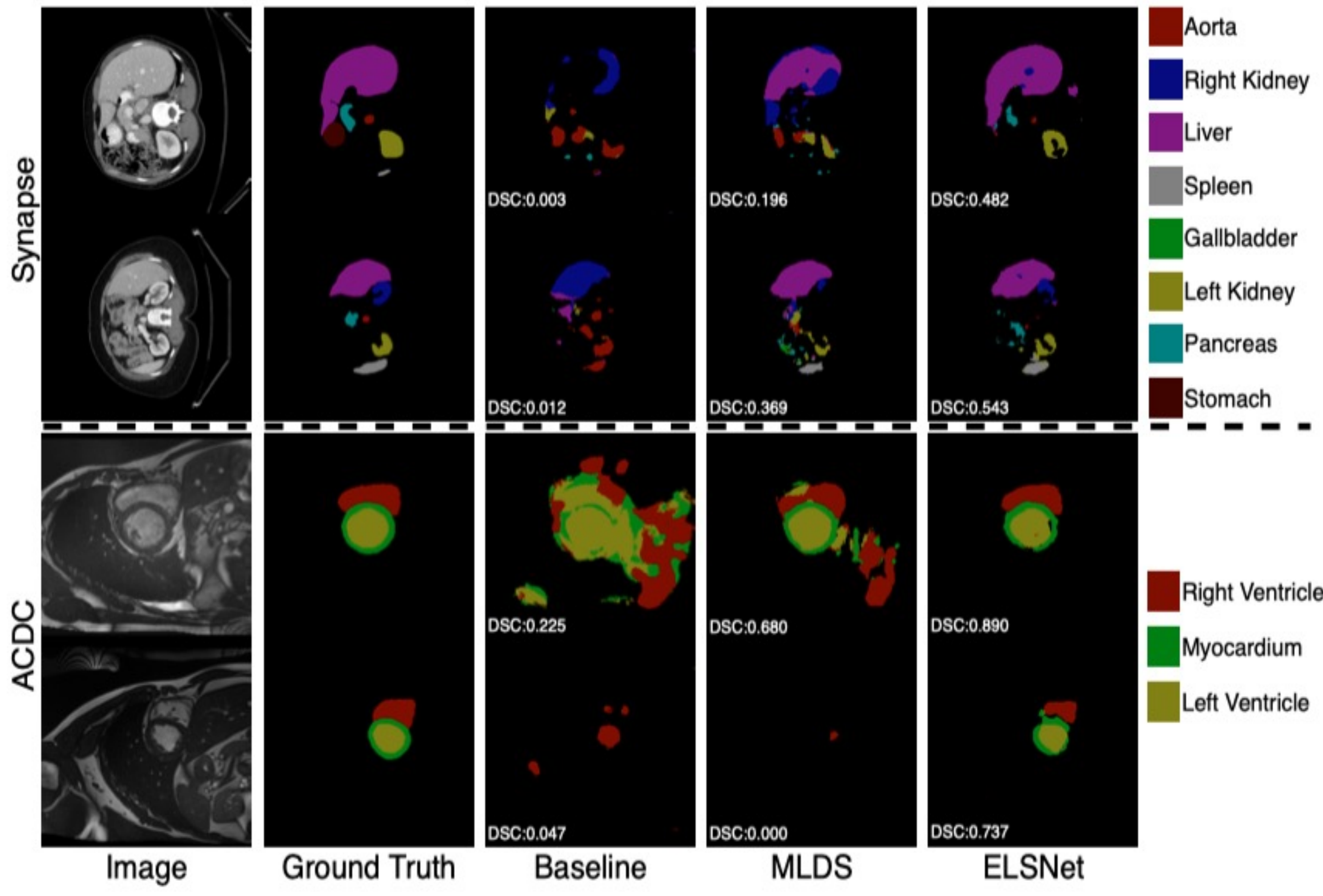
➢ Comparison results on the Synapse Dataset

Method	HD95↓	DSC↑	Aor	Gal	Kid(L)	Kid(R)	Liv	Pan	Spl	Sto
UNet[ ]	132.42	0.160	0.026	0.167	0.177	0.154	0.649	0.015	0.059	0.033
MTUNet[ ]	154.60	0.112	0.066	0.108	0.155	0.053	0.352	0.008	0.046	0.102
MLDS[ ]	159.26	0.221	0.057	0.147	0.306	0.183	0.638	0.038	**0.306**	0.090
ELSNet	**109.70**	**0.315**	**0.319**	**0.372**	**0.381**	**0.219**	**0.784**	**0.067**	0.276	**0.104**

➢ Impact of the proposed modules

Method	Synapse				ACDC			
	DSC↑	ΔDSC	HD95↓	ΔHD95	DSC↑	ΔDSC	HD95↓	ΔHD95
BS	0.112	-	154.60	-	0.142	-	74.20	-
+ESM	0.234	+0.122	120.59	-34.01	0.273	+0.131	38.35	-35.85
+ESM+PCEM_S1	0.264	+0.152	**101.00**	-53.60	0.355	+0.213	40.80	-33.40
+ESM+PCEM_S1&2	**0.315**	+0.203	109.70	-44.90	**0.410**	+0.268	**26.64**	-47.56

➢ Qualitative Evaluation



Carleton University