Paper ID: 1935

Synergistic Image and Feature Adaptation:

Towards Cross-Modality Domain Adaptation for Medical Image Segmentation

Cheng Chen[1], Qi Dou[1], Hao Chen[1,2], Jing Qin[3] and Pheng-Ann Heng[1,4]

[1] The Chinese University of Hong Kong

[2] ImSight Medical Technology Co., Ltd.

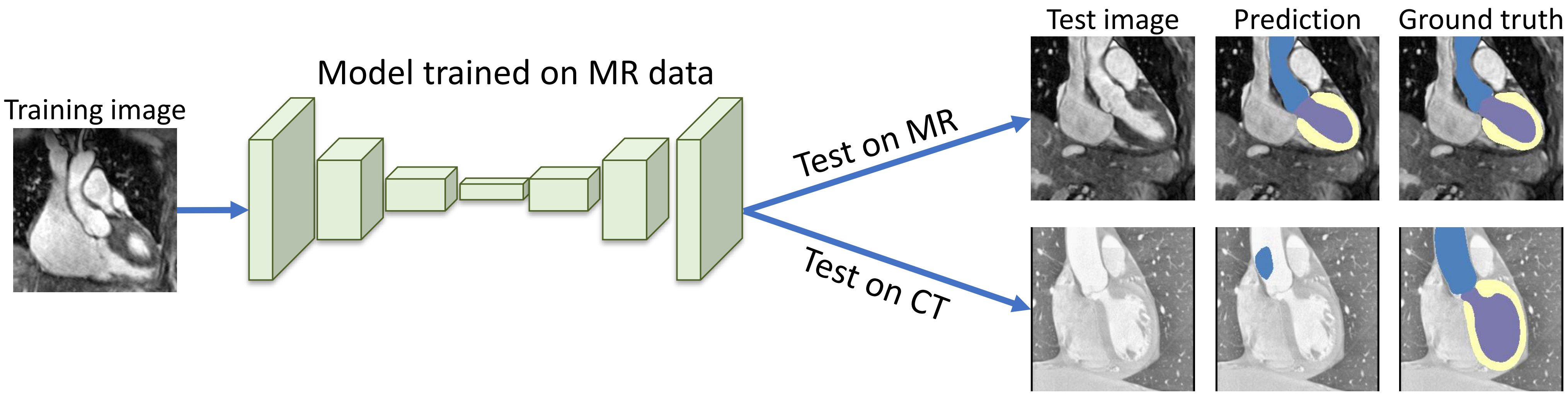
[3] The Hong Kong Polytechnic University

[4] Shenzhen Institutes of Advanced Technology

Problem & Challenges

- Success of deep neural networks
  - Assumption: training and testing data are drawn from the ***same distribution***.

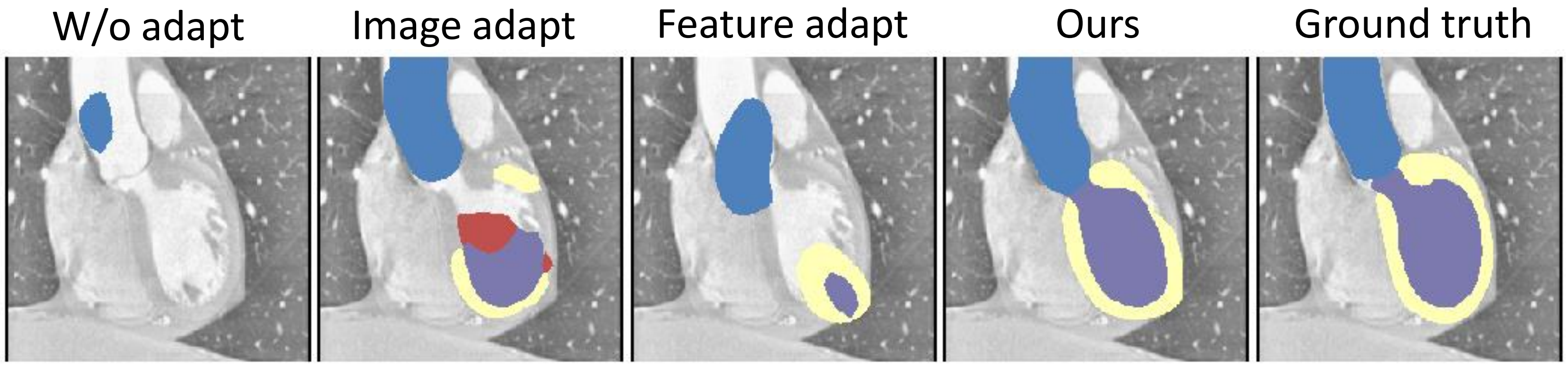
- Performance degradation
  - Domain shift: training and testing data present ***different distributions***.
  - Common situation in medical applications.



Related works

Unsupervised domain adaptation: adapt from labeled source domain to unlabeled target domain

- **Image adaptation**: align image appearance [Zhu et al. ICCV 2017; Zhang et al. MICCAI 2018]
- **Feature adaptation**: align feature distributions [Tzeng et al. CVPR 2017; Dou et al. IJCAI 2018]
- **Joint adaptation**: combine image and feature adaptation [Hoffman et al. ICML 2018; Zhang et al. CVPR 2018]

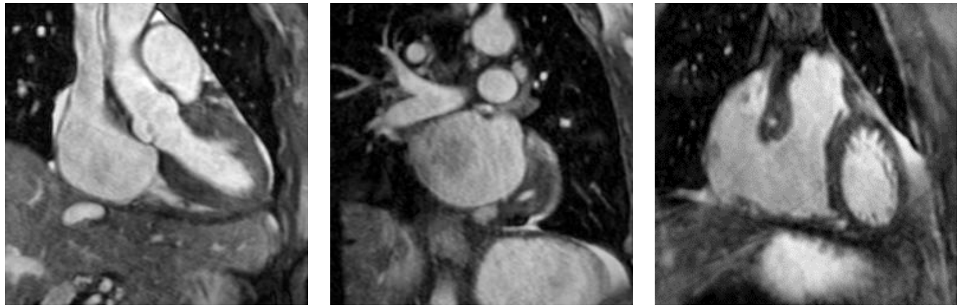


Proposed method

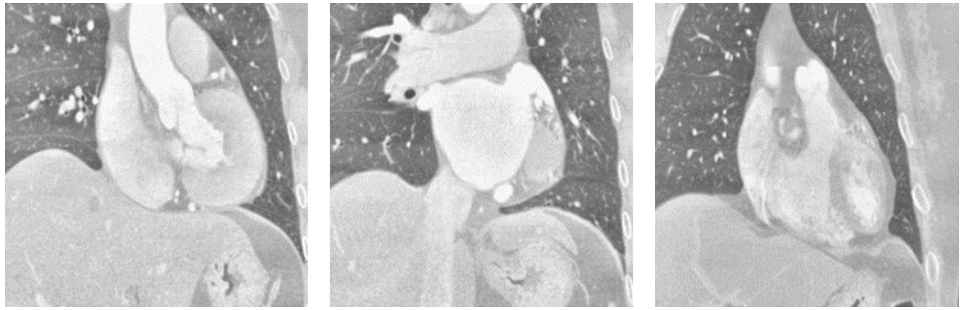
Synergistic image and feature adaptation (SIFA) for unsupervised domain adaptation

- Synergistic fusion of adaptations from both image and feature perspectives
- Feature encoder is shared to simultaneously transform image appearance and extract domain-invariant representations
- Challenging cross-modality medical image segmentation under severe domain shift

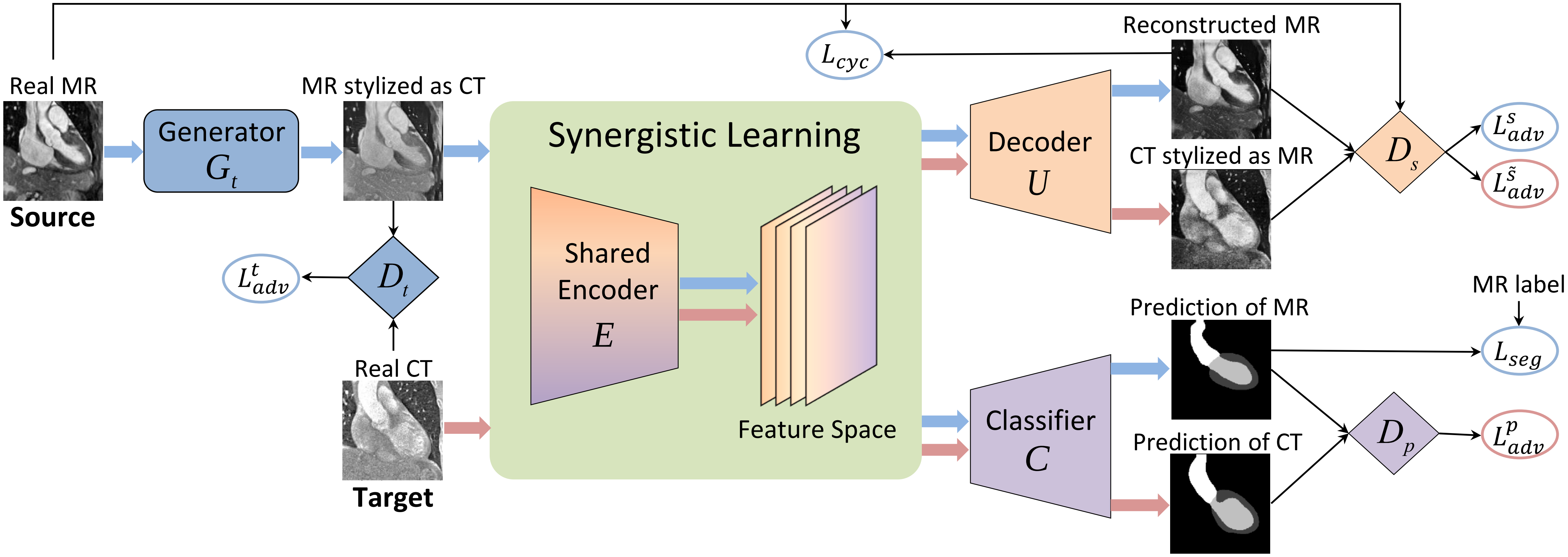
MR images (Source)



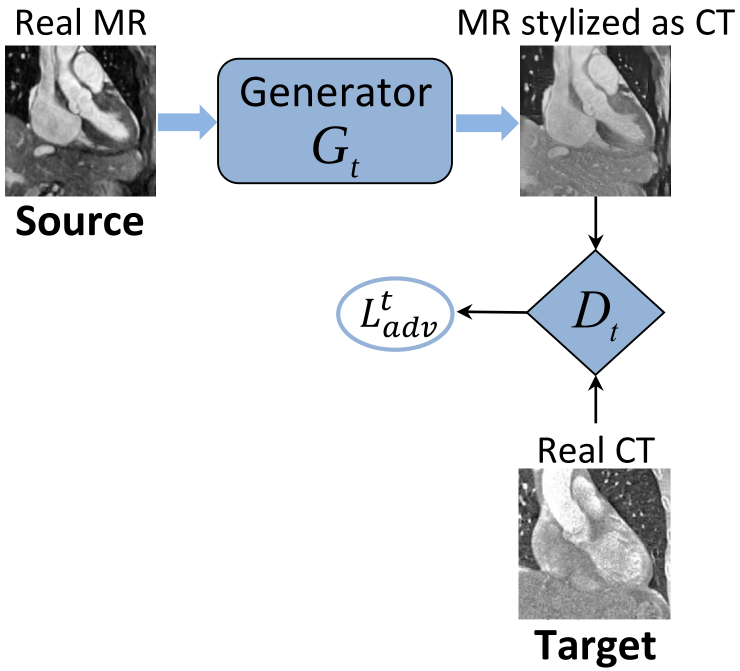
CT images (Target)



Proposed method

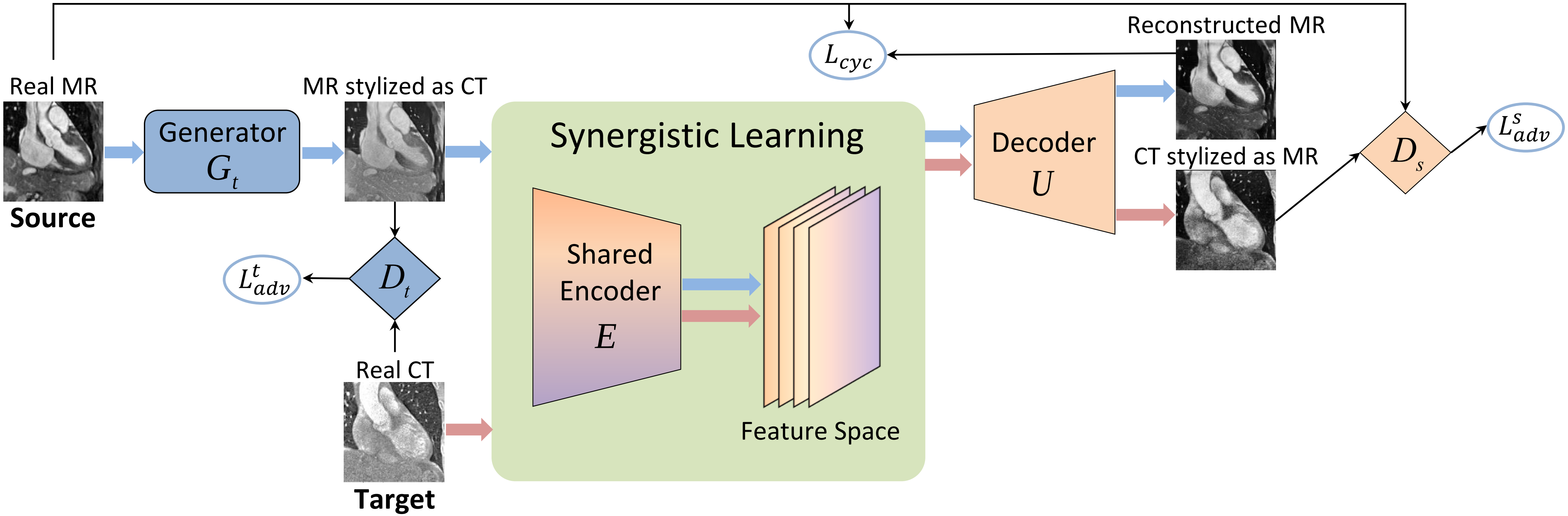


Proposed method



$$\mathcal{L}^t_{adv}(G_t, D_t) = \mathbb{E}_{x^t \sim X^t}[\log D_t(x^t)] + \mathbb{E}_{x^s \sim X^s}[\log(1 - D_t(G_t(x^s)))]$$

Proposed method

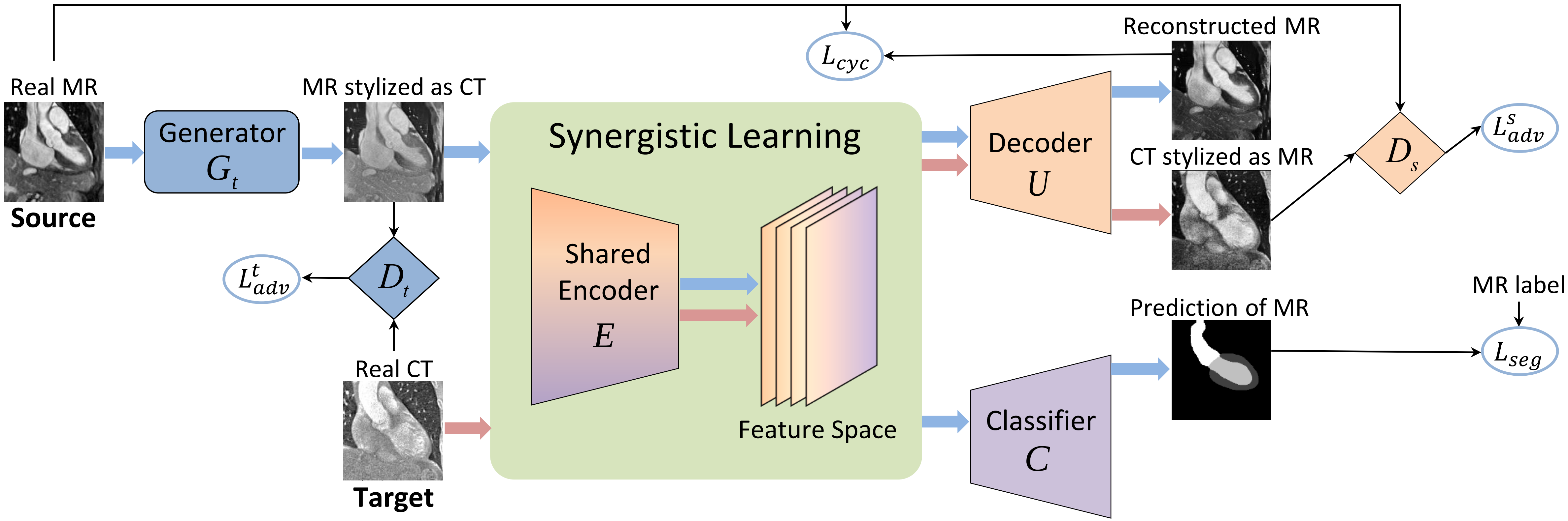


$$\mathcal{L}^t_{adv}(G_t, D_t) = \mathbb{E}_{x^t \sim X^t}[\log D_t(x^t)] + \mathbb{E}_{x^s \sim X^s}[\log(1 - D_t(G_t(x^s)))]$$

$$\mathcal{L}^s_{adv}(E, U, D_s) = \mathbb{E}_{x^s \sim X^s}[\log D_s(x^s)] + \mathbb{E}_{x^t \sim X^t}[\log(1 - D_s(U(E(x^t))))]$$

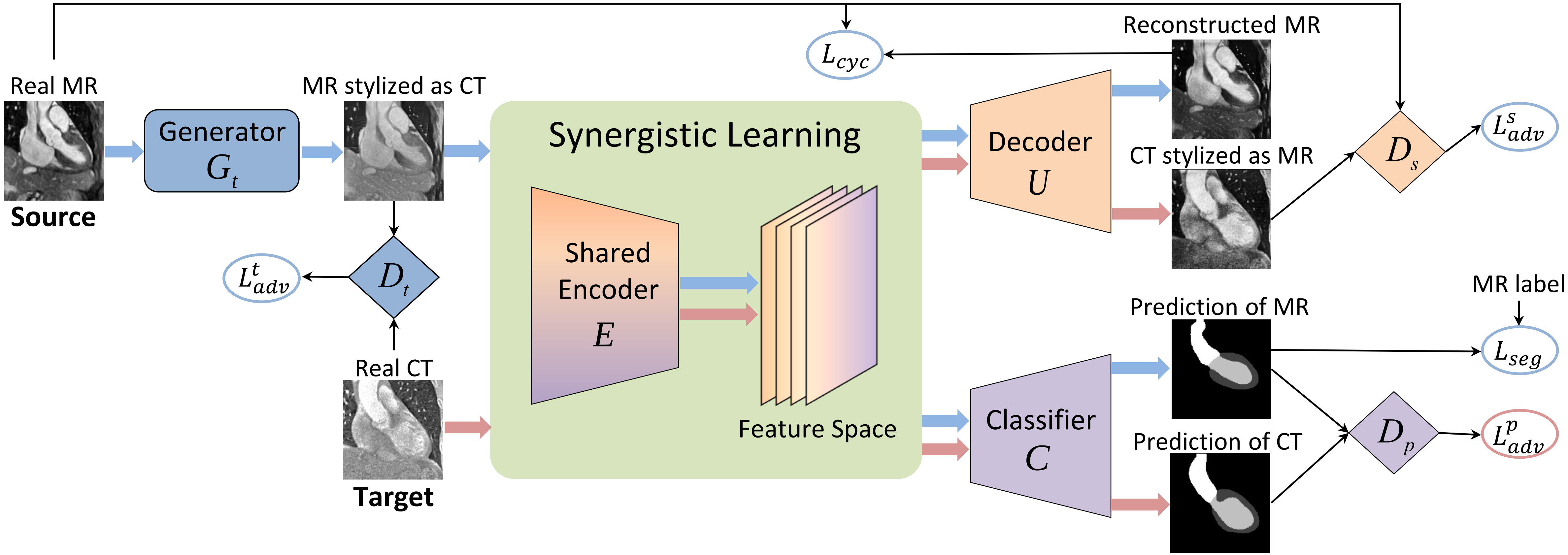
$$\mathcal{L}_{cyc}(G_t, E, U) = \mathbb{E}_{x^s \sim X^s}||U(E(G_t(x^s))) - x^s||_1 + \mathbb{E}_{x^t \sim X^t}||G_t(U(E(x^t))) - x^t||_1$$

Proposed method



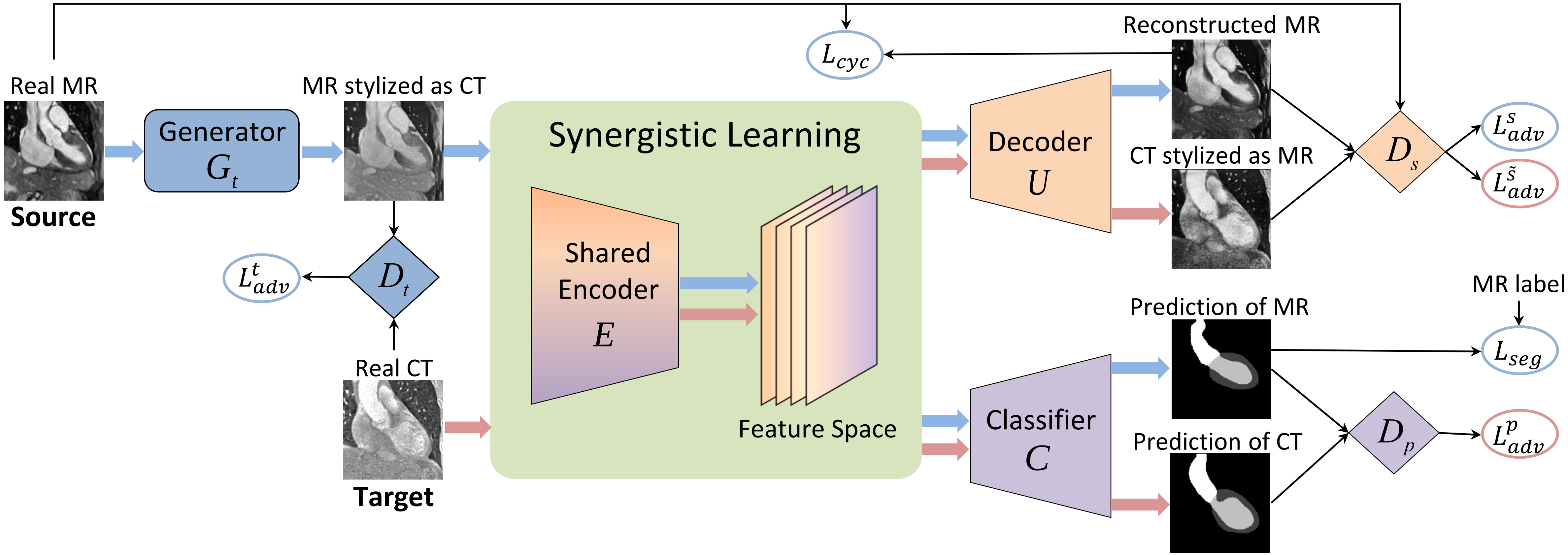
$$\mathcal{L}_{seg}(E, C) = H(y^s, \hat{y}^{s\to t}) + \alpha \cdot Dice(y^s, \hat{y}^{s\to t})$$

Proposed method



$$\mathcal{L}^p_{adv}(E, C, D_p) = \mathbb{E}_{x^{s\to t}\sim X^{s\to t}}[\log D_p(C(E(x^{s\to t})))] + \mathbb{E}_{x^t\sim X^t}[\log(1 - D_p(C(E(x^t))))]$$

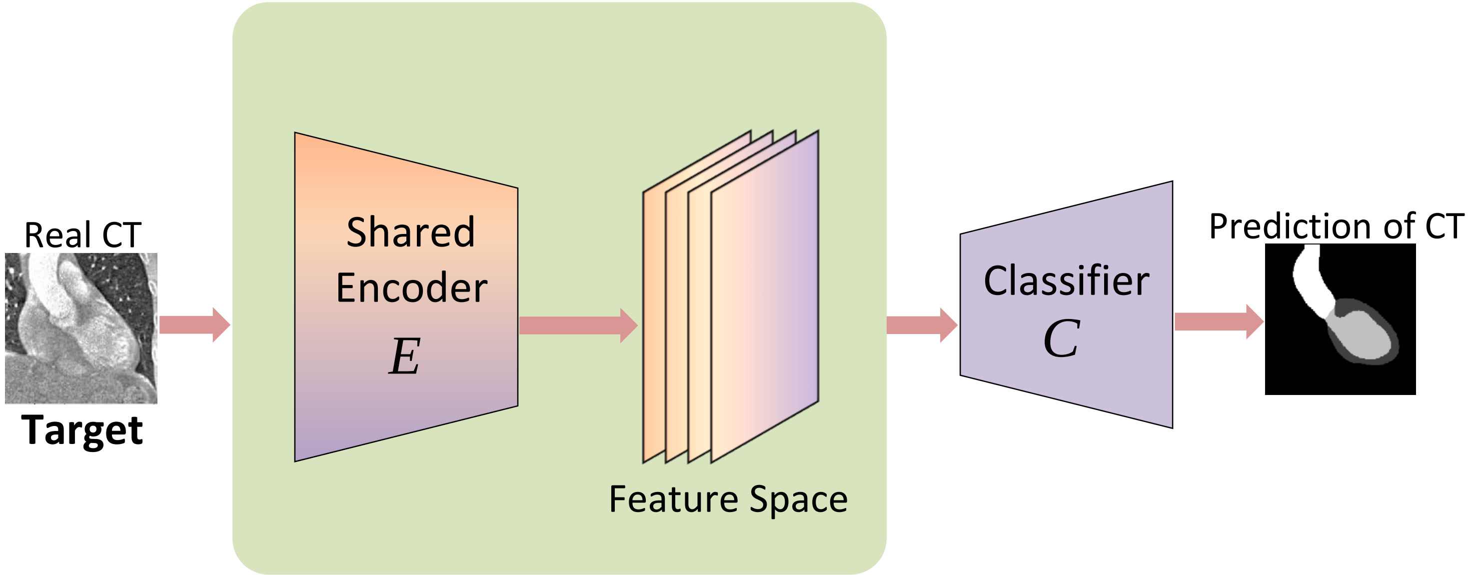
Proposed method



$$\mathcal{L}^p_{adv}(E, C, D_p) = \mathbb{E}_{x^{s\to t}\sim X^{s\to t}}[\log D_p(C(E(x^{s\to t})))] + \mathbb{E}_{x^t\sim X^t}[\log(1 - D_p(C(E(x^t))))]$$

$$\mathcal{L}^{\tilde{s}}_{adv}(E, D_s) = \mathbb{E}_{x^{s\to t}\sim X^{s\to t}}[\log D_s(U(E(x^{s\to t})))] + \mathbb{E}_{x^t\sim X^t}[\log(1 - D_s(U(E(x^t))))]$$

Inference


Real CT
Target
Shared Encoder
E
Feature Space
Classifier
C
Prediction of CT

Experiments

- **Dataset**
  - Multi-Modality Whole Heart Segmentation Challenge 2017 dataset
  - Unpaired 20 MR and 20 CT volumes collected at different clinical sites
  - Multi-class segmentation: ascending aorta (AA), the left atrium blood cavity (LAC), the left ventricle blood cavity (LVC), and the myocardium of the left ventricle (MYO).

- **Evaluation metrics**
  - Dice coefficient

$$Dice = \frac{2TP}{2TP + FP + FN}$$

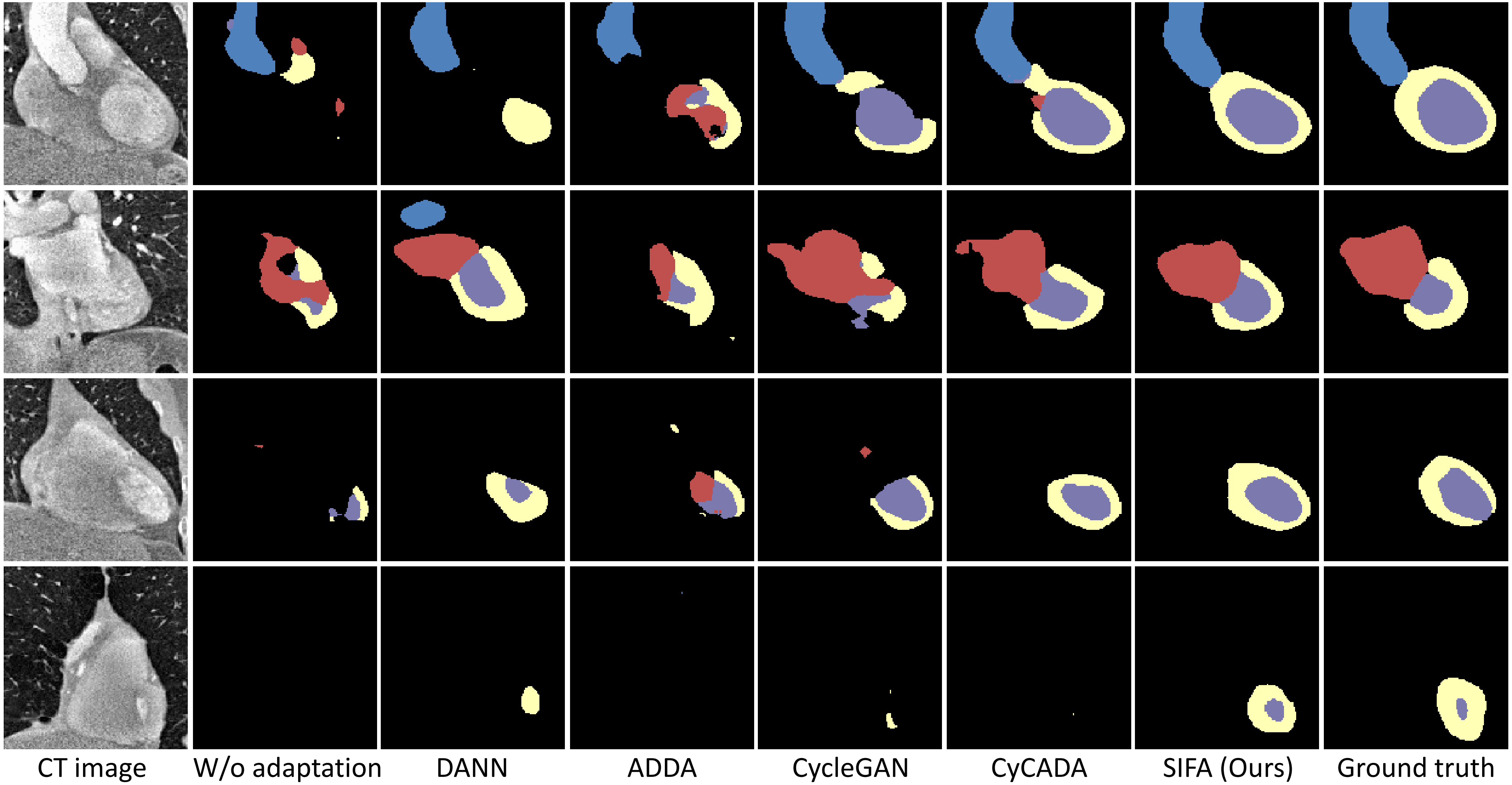
  - Average surface distance (ASD)

Quantitative results

Compare with the state-of-the-art unsupervised domain adaptation methods

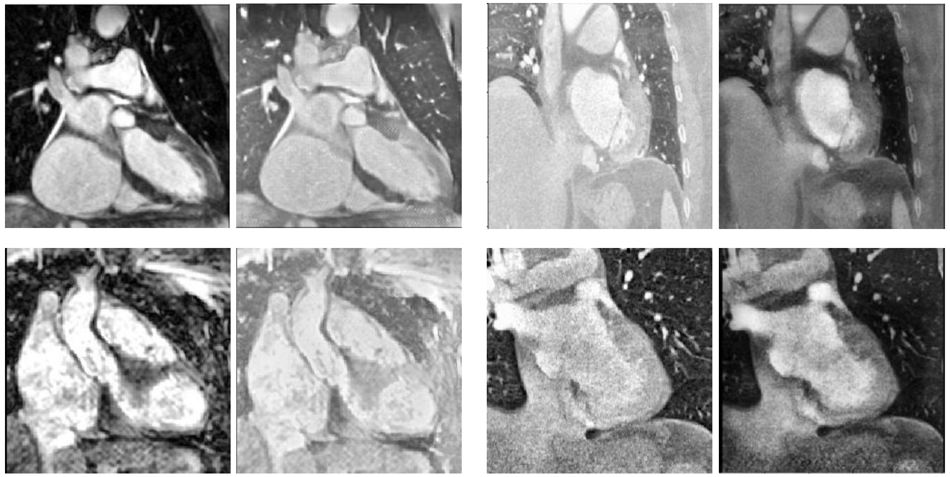
Methods	Adaptation		Dice					ASD				
	Image	Feature	AA	LAC	LVC	MYO	Average	AA	LAC	LVC	MYO	Average
W/o adaptation			28.4	27.7	4.0	8.7	17.2	20.6	16.2	N/A	48.4	N/A
DANN (Ganin et al. 2016)		✓	39.0	45.1	28.3	25.7	34.5	16.2	9.2	12.1	10.1	11.9
ADDA (Tzeng et al. 2017)		✓	47.6	60.9	11.2	29.2	37.2	13.8	10.2	N/A	13.4	N/A
CycleGAN (Zhu et al. 2017)	✓		73.8	75.7	52.3	28.7	57.6	11.5	13.6	9.2	8.8	10.8
CyCADA (Hoffman et al. 2018)	✓	✓	72.9	**77.0**	62.4	45.3	64.4	**9.6**	8.0	9.6	10.5	9.4
Dou et al. (Dou et al. 2018)		✓	74.8	51.1	57.2	47.8	57.7	27.5	20.1	29.5	31.2	27.1
Joyce et al. (Joyce et al. 2018)		✓	-	-	66	44	-	-	-	-	-	-
SIFA (Ours)	✓	✓	**81.1**	76.4	**75.7**	**58.7**	**73.0**	10.6	**7.4**	**6.7**	**7.8**	**8.1**

Visual comparison results

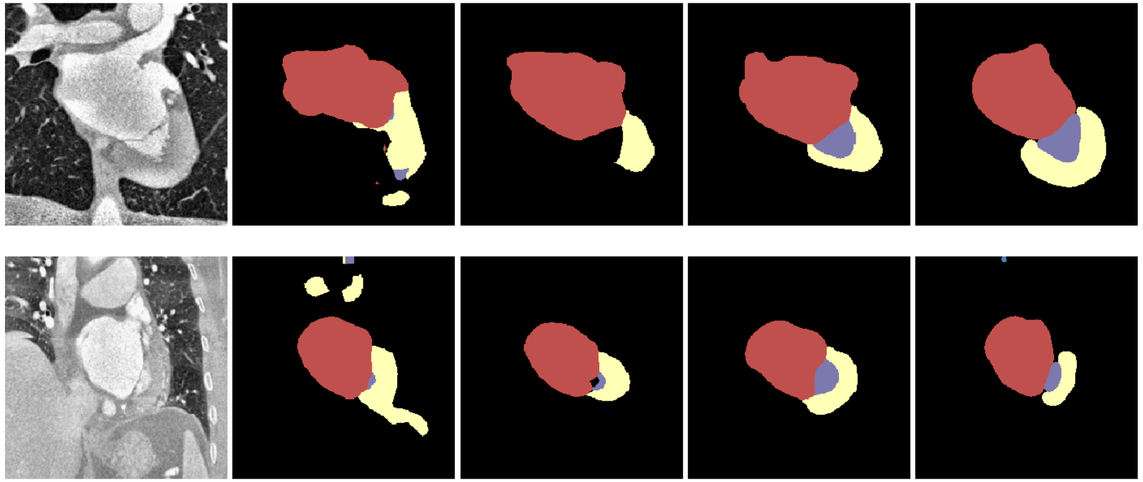


Evaluation of key components

Methods	IA	$\mathcal{L}^{p}_{adv}$	$\mathcal{L}^{\tilde{s}}_{adv}$	Average Dice
W/o adaptation				17.2
+ Image adaptation	✓			58.0
+ FA-P	✓	✓		65.7
+ FA-I	✓	✓	✓	**73.0**



MR MR stylized as CT CT CT stylized as MR



CT image IA IA with FA-P SIFA Ground truth

Conclusion & Future works

Conclusion

- Propose synergistic image and feature adaptation to achieve unsupervised domain adaptation from complementary perspectives.
- Without using any annotations from the target domain, recovers performance from 17.2% to 73.0%, outperforming the state-of-the-art methods.

Future works

- Efficient multi-modality domain adaptation
- Extend to 3D network architecture

Thank you!
Q & A

Code is available at:

https://github.com/cchen-cc/SIFA