Contrastive Image Synthesis and Self-supervised Feature Adaptation for Cross-Modality Biomedical Image segmentation

Xinrong Hu[1], Corey Wang[2], Yiyu Shi[1]

[1] University of Notre Dame, Notre Dame, IN 46556, USA

{xhu7, yshi4}@nd.edu

[2]Northwestern University, Evanston, IL 60208, USA

COMPUTER SCIENCE AND ENGINEERING

ICCV CVAMD 2023

IEEE COMPUTER SOCIETY

INTRODUCTION

- This work focuses on unsupervised domain adaptation (UDA) segmentation task for medical images. We have pixel-wise labels on source domain images while no annotation is provided for target domain.
- One branch of approaches to UDA is extracting common features shared by different domains with on encoder[1,2], but this feature adaptation is coarse-grained. On the other hand, style translation methods [3] are alternative solutions, most of which are based on a complex CycleGAN[4] framework.
- We present a novel framework CISFA (Contrastive Image synthesis and Self-supervised Feature Adaptation). We simplify the translation path with patch-wise shape constraint, and utilize a novel contrastive loss for feature adaptation.

[1]Yi-Hsuan Tsai, et al. Learning to adapt structured output space for semantic segmentation. CVPR 2018.

[2] Qiming Zhang, et al. .Category anchor-guided unsupervised domain adaptation for semantic segmentation. 2019

[3] Cheng Chen, et al. Unsupervised bidirectional cross-modality adaptation via deeply synergistic image and feature alignment for medical image segmentation.

[4] Jun-Yan Zhu, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. ICCV 2017

Methods

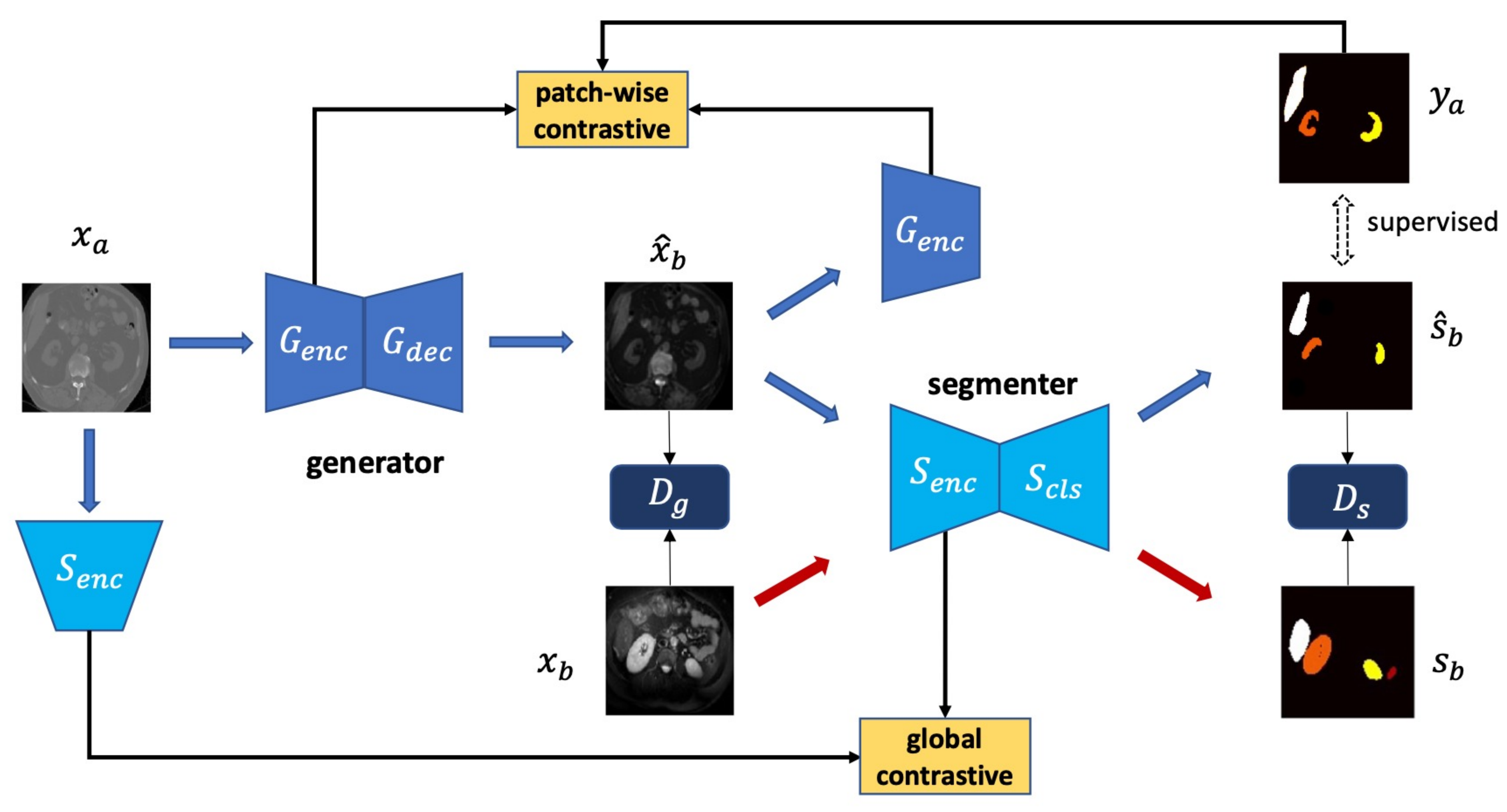

Fig 1. Overview of CISFA framework

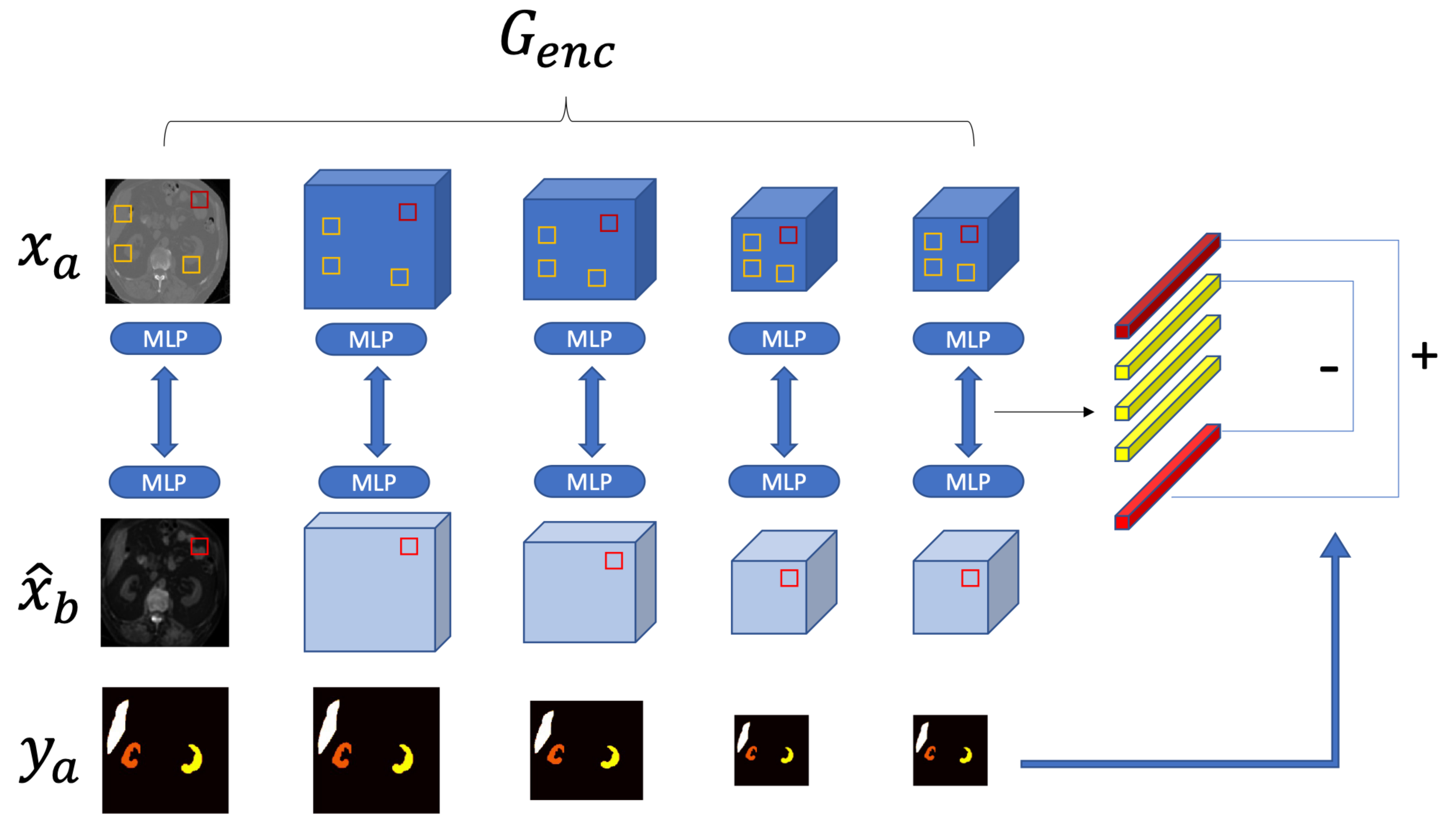

Fig 2. Illustration of patch-wise contrastive loss. The label masks are downsampled to the same resolution as each layer of feature maps and increase the weights for non-background patches

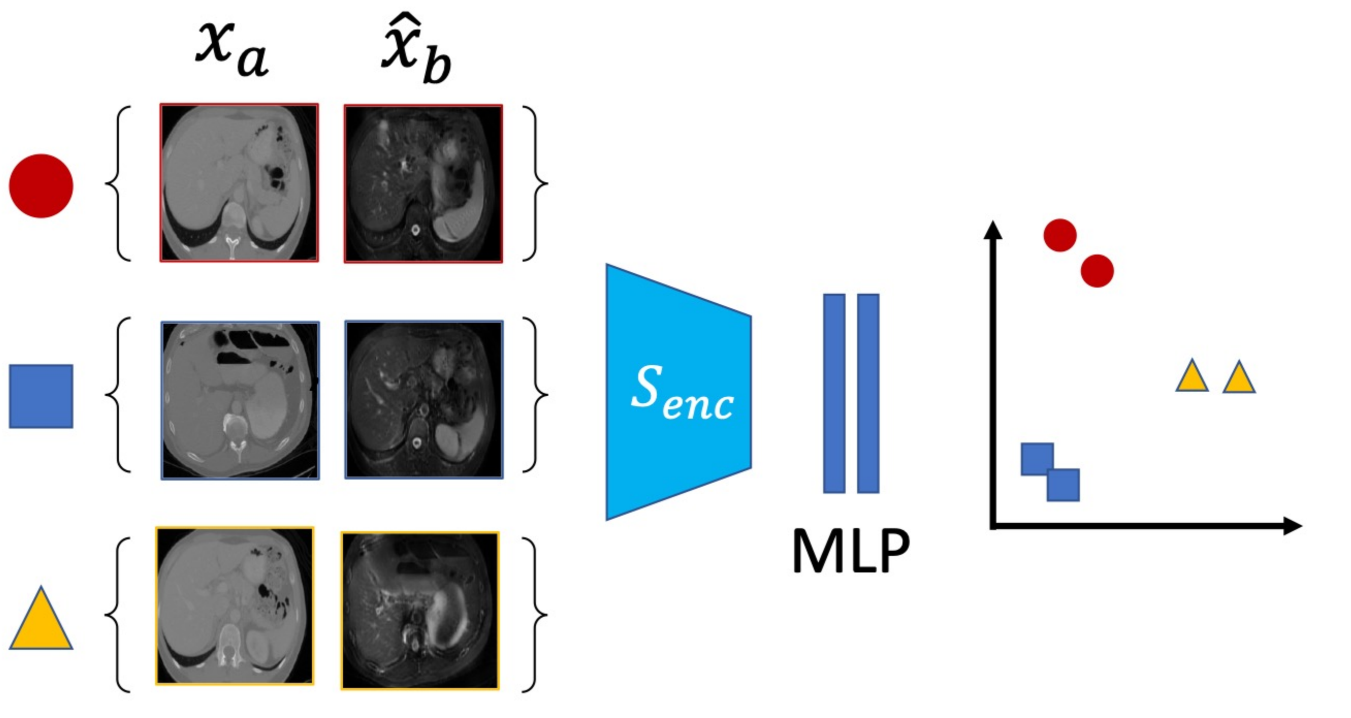

Fig 3 Illustration of global contrastive loss between input images and the corresponding translated fake target images

Results

Table 1. Comparison between state-of-the-art methods and our CISFA on abdominal images, and the translation direction is CT->MRI

Methods	Dice% ↑				
	liver	LK	RK	spleen	avg
Supervised	89.00±1.08	87.19±2.49	83.31±5.05	88.08±1.82	86.90±2.19
W/o adaptation	10.15±3.94	3.67±3.57	4.04±2.95	7.15±6.81	6.25±1.26
CUT	38.17±6.33	32.20±10.69	34.01±9.32	35.83±10.44	35.05±8.19
VarDA	41.63±1.77	32.95±6.47	34.53±4.14	32.23±4.72	35.33±2.60
SASAN	67.23±9.98	61.41±12.95	67.94±14.63	62.63±13.65	64.80±11.48
SIFA	77.24±2.03	68.03±5.60	68.99±5.16	66.79±4.87	70.26±3.69
CISFA(no weight)	76.14±10.72	72.12±4.52	**74.94±4.14**	73.18±3.11	74.10±1.84
CISFA	**80.13±2.21**	**74.45±5.67**	74.51±5.16	**75.86±5.28**	**76.24±2.17**

Table 2. Comparison between state-of-the-art methods and our CISFA on abdominal images, and the translation direction is MRI->CT

Methods	Dice% ↑				
	liver	LK	RK	spleen	avg
Supervised	89.03±.95	85.53±12.79	83.94±9.46	85.49±4.05	86.00±3.67
W/o adaptation	9.38±3.08	8.88±1.26	8.40±1.31	9.70±1.52	9.09±0.68
CUT[22]	17.78±8.74	28.34±8.05	21.16±11.83	19.29±10.60	21.64±8.66
VarDA[29]	32.78±2.29	38.11±4.17	31.71±4.32	30.26±3.33	33.22±2.38
SASAN[25]	75.36±4.24	67.33±6.43	67.25±6.08	58.70±15.24	67.13±4.32
SIFA[3]	74.03±1.13	65.21±9.88	63.17±10.91	63.53±11.85	66.49±5.61
CISFA(no weight)	**77.45±2.15**	66.91±7.16	64.92±4.57	65.40±13.12	68.67±2.03
CISFA	75.78±3.70	**69.30±7.77**	**70.15±4.77**	**66.57±12.40**	**70.45±2.81**

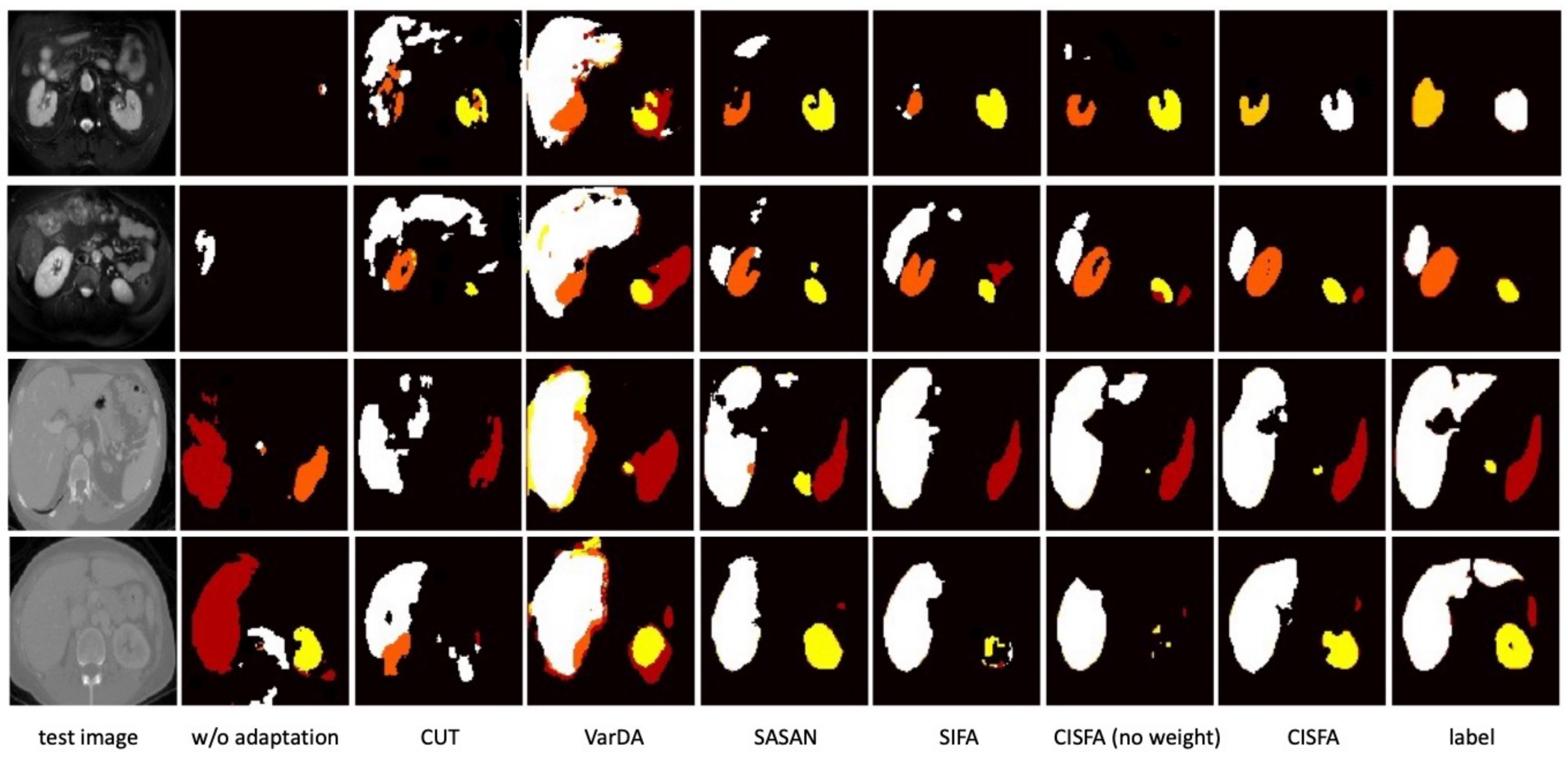

Fig 4. Qualitative results on abdominal dataset, for the first two rows, MRI is target domain, and for the lower two, target domain is CT.

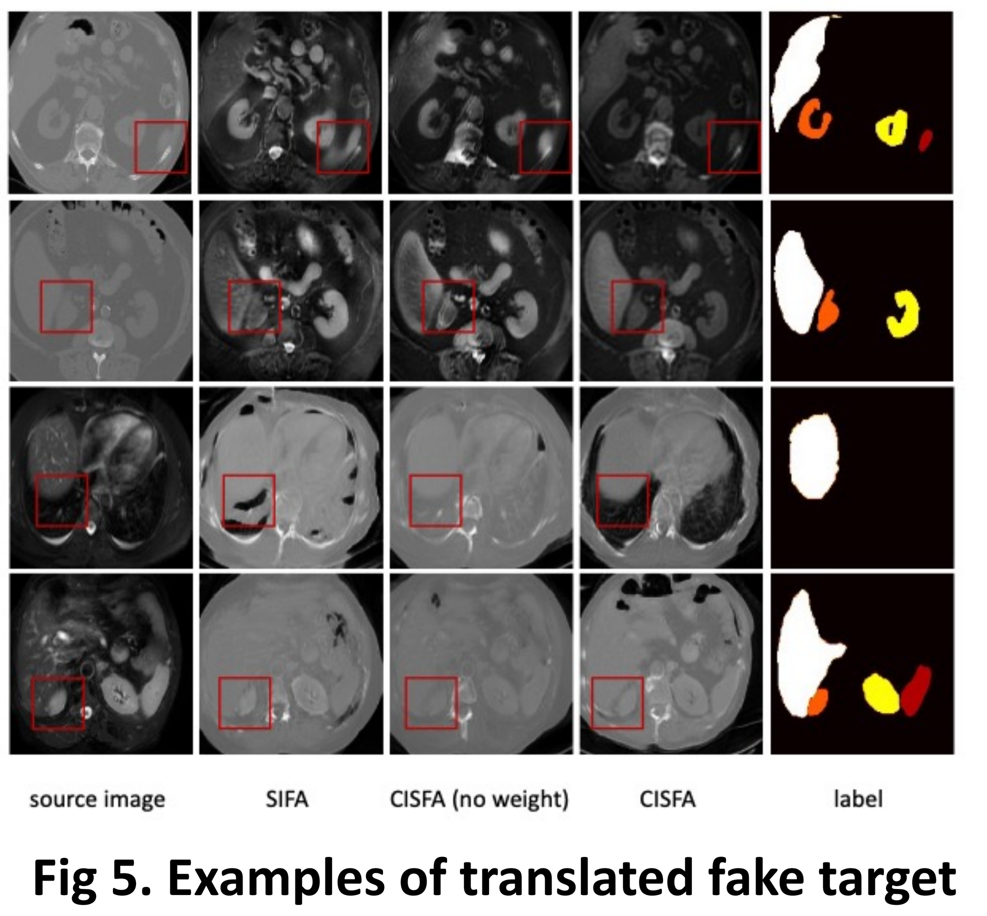

Fig 5. Examples of translated fake target domain images

Table 3. Comparison on MMWHS dataset, MRI->CT

Method	Dice%	ASSD
Supervised	89.78±1.26	0.33±0.05
W/o adaptation	3.13±1.99	-
CUT[22]	37.28±8.32	3.37±1.54
VarDA[29]	40.36±2.86	2.74±0.67
SASAN[25]	61.74±3.34	1.80±0.78
SIFA [3]	64.50±4.21	2.14±1.21
CISFA (ours)	**68.87±3.15**	**1.49±0.31**