

A New Bidirectional Unsupervised Domain Adaptation Segmentation Framework

Munan Ning^{1,2}, Cheng Bian¹, Dong Wei¹, Shuang Yu¹, Chenglang Yuan¹,
Yang Guo², Yaohua Wang², Kai Ma¹, Yefeng Zheng¹

¹ Tencent Jarvis Lab, Tencent, Shenzhen, China

² National University of Defense Technology, Changsha, China



Problem Definition

Most unsupervised domain adaptation (UDA) methods can only achieve satisfactory improvements in one adaptation direction (e.g., MRI to CT), but often perform poorly in the other (CT to MRI), limiting their practical usage.

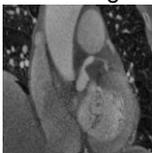
We propose a bi-directional UDA (BiUDA) framework based on disentangled representation learning for equally competent two-way UDA performances.

Problem Analysis

Characteristic of multi-modal images:

1. Images of different modals have different texture and contrast
2. Due to the variant image qualities, the difficulty of segmentation varies in two modals.

CT Image

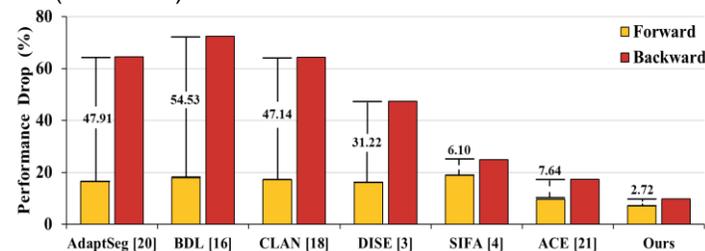


MRI Image

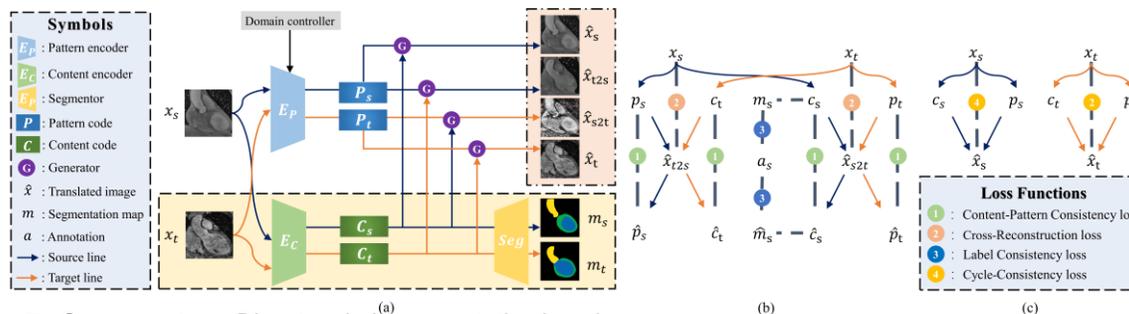


Domain drop phenomenon:

Existing methods work effectively for MRI to CT (forward), while suffering a dramatic drop in performance when reversing the adaptation direction (backward).



BiUDA Framework



Component 1 – Disentangled representation learning:

We make the model aware of the cross-domain commonalities and differences by a DRPL (disentangled representation learning) with a content encoder $E_C: \mathcal{X} \rightarrow \mathcal{C}$ and a pattern encoder $E_P: \mathcal{X} \rightarrow \mathcal{P}$. The framework is optimized with the following loss functions:

a) Content-pattern consistency loss (as shown in ① in Fig):

$$\mathcal{L}^{cp} = \mathbb{E}_{\hat{c} \sim \hat{C}, c \sim \mathcal{C}} [\|\hat{c} - c\|_1] + \mathbb{E}_{\hat{p} \sim \hat{P}, p \sim \mathcal{P}} [\|\hat{p} - p\|_1]$$

where the content code \hat{c} and pattern code \hat{p} for the transferred images and the codes c and p from the origin images are constrained to be consistent and the potential distortions in DRPL are penalized.

b) Cross-reconstruction loss (as shown in ② in Fig):

$$\mathcal{L}^{GAN} = \mathbb{E}_{\hat{x}_{s2t} \sim \hat{\mathcal{X}}_{s2t}} [\log(1 - D(\hat{x}_{s2t}))] + \mathbb{E}_{x_t \sim \mathcal{X}_t} [\log D(x_t)]$$

where a discriminator D is introduced to align the transferred image and corresponding image (e.g., \hat{x}_{s2t} and x_t) and produce high-quality recomposed image. Symmetry loss is calculated between \hat{x}_{t2s} and x_s .

c) Cycle-consistency loss (as shown in ④ in Fig):

$$\mathcal{L}^{cycle} = \mathbb{E}_{\hat{x} \sim \hat{\mathcal{X}}, x \sim \mathcal{X}} [\|\hat{x} - x\|_1]$$

where the input images x and their reconstructions \hat{x} are constrained to be close.

Component 2 – Anatomical consistency based segmentation learning:

The source-domain image x_s and source-to-target transferred image \hat{x}_{s2t} should contain the same anatomical structures. So we introduce a label consistency loss to supervise the segmentation learning:

$$\mathcal{L}^{lc} = \mathcal{L}^{seg}(a_s, m_s) + \mathcal{L}^{seg}(a_s, \hat{m}_s)$$

where the m_s and \hat{m}_s denotes segmentation of x_s and \hat{x}_{s2t} respectively, and they share the consistent label a_s . Label consistency loss provides supplementary target-domain training data and help relieve the domain drop.

Experiments

Table 1. Performance comparison of our proposed BiUDA framework with SOTA UDA algorithms on the MMWS dataset using the average performance drops (lower is better) in Dice and F1 score.

Method	AA		LABC		LVBC		LVM		Mean	
	Dice ⁺	F1 ⁺								
M2M*	81.68	82.11	85.25	85.39	93.01	93.07	84.68	84.71	86.16	86.39
C2C*	96.17	96.21	93.25	93.28	89.67	89.92	84.26	84.54	90.84	91.07
AdaptSegNet [20]	55.80	52.50	48.94	48.67	23.20	19.62	34.14	33.41	40.52	37.62
BDL [16]	55.31	48.06	48.91	47.88	32.33	25.42	44.47	44.15	45.25	40.15
CLAN [18]	56.90	51.46	49.17	47.47	23.18	19.58	33.49	32.76	40.68	36.68
DISE [3]	38.48	34.89	43.70	34.39	12.61	10.66	32.62	27.31	31.85	25.19
SIFA [4]	16.29	15.02	22.72	21.58	20.61	19.42	27.74	27.05	21.84	20.49
ACE [21]	13.39	11.48	11.43	10.81	5.08	4.76	24.00	23.31	13.47	12.44
Ours	8.70	8.54	6.68	6.37	3.50	3.48	15.07	14.76	8.49	8.29

*Upper-bound performances are reported as the original Dice and F1 scores.

Table 2. Performance comparison of our proposed BiUDA framework with SOTA UDA algorithms on the MMAS dataset. Dice⁺: average performance drop in Dice; F1⁺: average performance drop in F1 score.

Method	Liver		R.kidney		L.kidney		Spleen		Mean	
	Dice ⁺	F1 ⁺								
M2M*	93.89	93.98	93.34	93.16	92.30	92.13	91.95	92.6	92.87	92.97
C2C*	96.25	96.06	90.99	90.68	91.86	92.27	93.72	92.96	93.21	92.99
AdaptSegNet [20]	14.29	17.87	27.33	32.58	39.52	43.36	27.36	32.46	27.12	31.57
BDL [16]	21.94	26.81	27.93	32.59	44.18	48.79	27.29	30.91	30.33	34.77
CLAN [18]	19.90	24.80	27.63	32.61	40.37	44.07	27.37	32.57	28.82	33.51
DISE [3]	8.00	10.04	9.38	11.24	10.01	12.34	8.27	10.71	8.92	11.08
SIFA [4]	5.84	5.93	5.75	5.93	9.94	10.94	8.55	8.96	7.52	7.94
ACE [21]	5.36	5.90	5.82	7.16	5.02	5.99	4.89	5.57	5.27	6.15
Ours	5.04	5.79	4.44	5.13	3.69	4.42	4.38	5.42	4.39	5.19

*Upper-bound performances are reported as the original Dice and F1 scores.

