华东师范大学
EAST CHINA NORMAL UNIVERSITY

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

# Bidirectional Copy-Paste for Semi-Supervised Medical Image Segmentation

Yunhao Bai[1] Duowen Chen[1] Qingli Li[1] Wei Shen[2] Yan Wang[1*]
[1]Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University
[2]MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University
{yhbai@stu., duowen_chen@stu., qlli@cs., ywang@cee.}ecnu.edu.cn, wei.shen@sjtu.edu.cn

WED-AM-313

## Semi-supervised segmentation

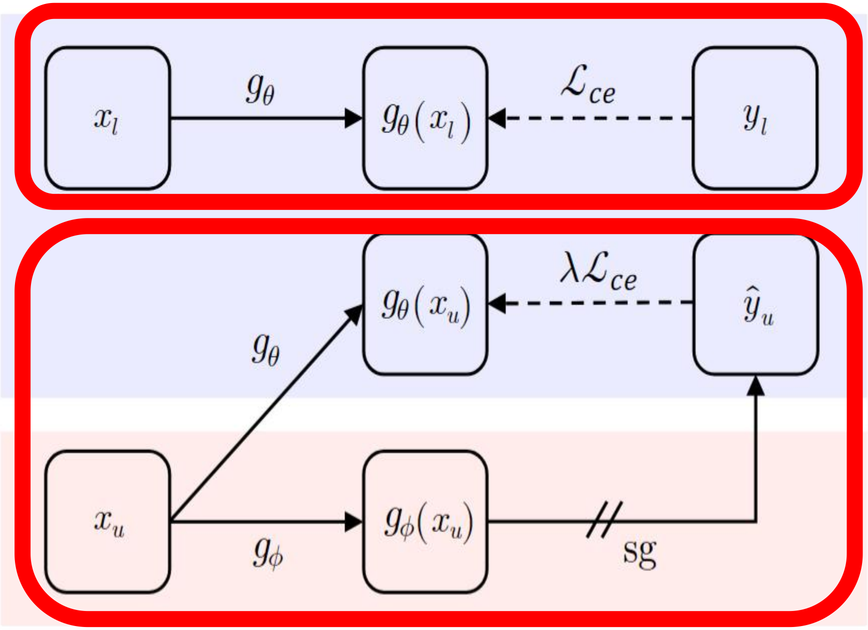



Figure of paper: *Learning from Future: A Novel Self-Training Framework for Semantic Segmentation (NIPS 2022)*

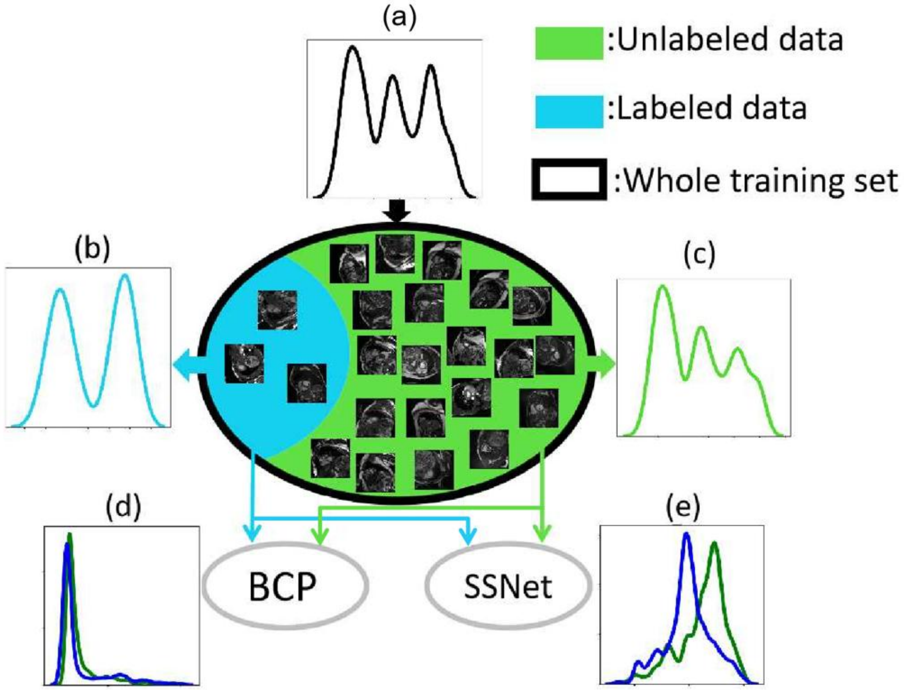
| Method | Scans used | | Metrics | | | |
|---|---|---|---|---|---|---|
| | Labeled | Unlabeled | Dice↑ | Jaccard↑ | 95HD↓ | ASD↓ |
| U-Net | 3(5%) | 0 | 47.83 | 37.01 | 31.16 | 12.62 |
| U-Net | 7(10%) | 0 | 79.41 | 68.11 | 9.35 | 2.70 |
| U-Net | 70(All) | 0 | 91.44 | 84.59 | 4.30 | 0.99 |
| UA-MT | 3(5%) | 67(95%) | 46.04 | 35.97 | 20.08 | 7.75 |
| SASSNet | | | 57.77 | 46.14 | 20.05 | 6.06 |
| DTC | | | 56.90 | 45.67 | 23.36 | 7.39 |
| URPC | | | 55.87 | 44.64 | 13.60 | 3.74 |
| MC-Net | | | 62.85 | 52.29 | 7.62 | 2.33 |
| SS-Net | | | 65.83 | 55.38 | 6.67 | 2.28 |
| Ours | | | **87.59**↑21.76 | **78.67**↑23.29 | **1.90**↓4.77 | **0.67**↓1.61 |
| UA-MT | 7(10%) | 63(90%) | 81.65 | 70.64 | 6.88 | 2.02 |
| SASSNet | | | 84.50 | 74.34 | 5.42 | 1.86 |
| DTC | | | 84.29 | 73.92 | 12.81 | 4.01 |
| URPC | | | 83.10 | 72.41 | 4.84 | 1.53 |
| MC-Net | | | 86.44 | 77.04 | 5.50 | 1.84 |
| SS-Net | | | 86.78 | 77.67 | 6.07 | 1.40 |
| Ours | | | **88.84**↑2.06 | **80.62**↑2.95 | **3.98**↓2.09 | **1.17**↓0.23 |

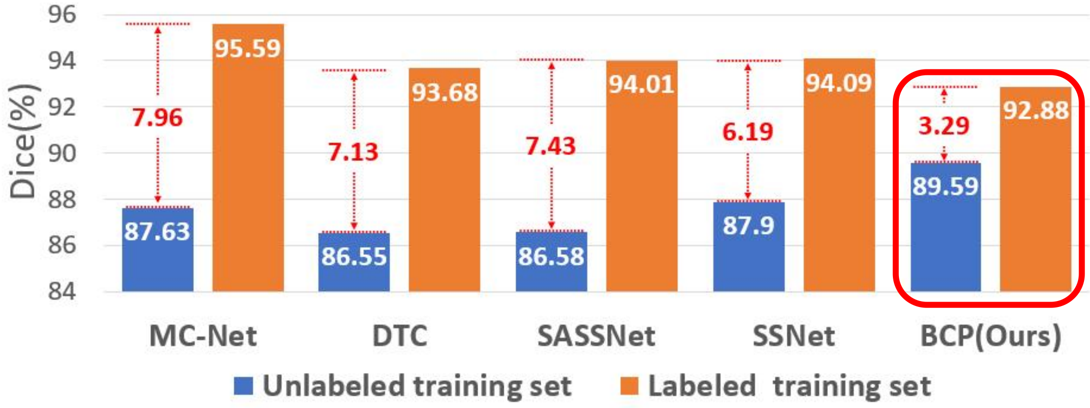
Table 3. Comparisons with state-of-the-art semi-supervised segmentation methods on the ACDC dataset.

Yunhao Bai, Duowen Chen, Qingli Li, Wei Shen, and Yan Wang


# 1. Motivation

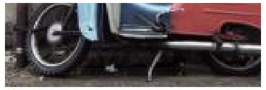
- a) Empirical distribution mismatch:



# 1. Motivation

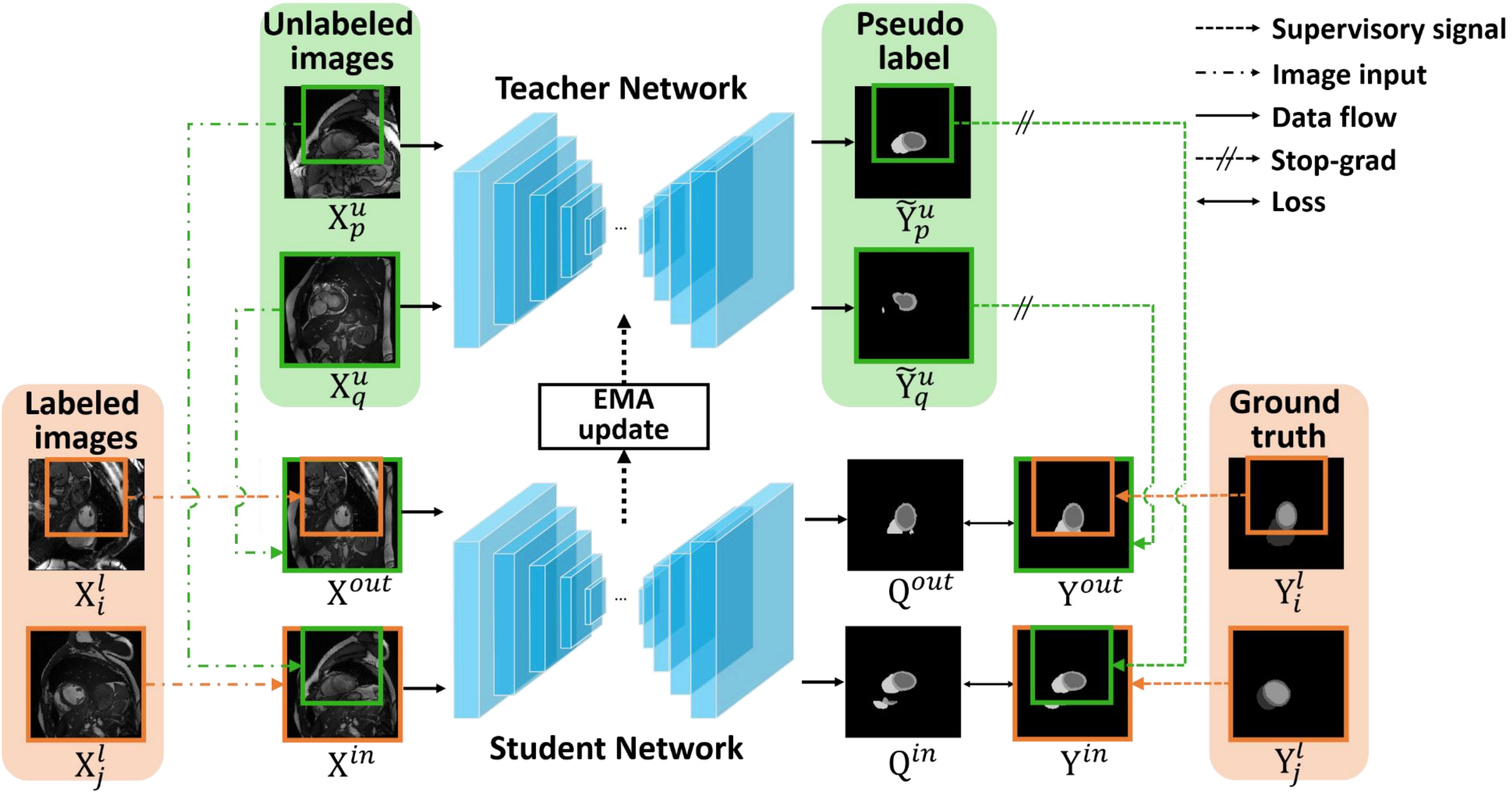
- b) CutMix (**C**opy-**P**aste, CP) could be used better:

WACV 2021: *DACS: Domain Adaptation via Cross-domain Mixed Sampling*

CVPR 2022: *Perturbed and Strict Mean Teachers for Semi-supervised Semantic Segmentation*

# 2. Method


Unlabeled images
$X_p^u$
$X_q^u$
Teacher Network
Pseudo label
$\tilde{Y}_p^u$
$\tilde{Y}_q^u$
Supervisory signal
Image input
Data flow
Stop-grad
Loss
Labeled images
$X_i^l$
$X_j^l$
EMA update
$X^{out}$
$X^{in}$
Student Network
$Q^{out}$
$Q^{in}$
$Y^{out}$
$Y^{in}$
Ground truth
$Y_i^l$
$Y_j^l$

# 3. Experiments

| Method | Scans used | | Metrics | | | |
|---|---|---|---|---|---|---|
| | Labeled | Unlabeled | Dice↑ | Jaccard↑ | 95HD↓ | ASD↓ |
| V-Net | 4(5%) | 0 | 52.55 | 39.60 | 47.05 | 9.87 |
| V-Net | 8(10%) | 0 | 82.74 | 71.72 | 13.35 | 3.26 |
| V-Net | 80(All) | 0 | 91.47 | 84.36 | 5.48 | 1.51 |
| UA-MT | 4(5%) | 76(95%) | 82.26 | 70.98 | 13.71 | 3.82 |
| SASSNet | | | 81.60 | 69.63 | 16.16 | 3.58 |
| DTC | | | 81.25 | 69.33 | 14.90 | 3.99 |
| URPC | | | 82.48 | 71.35 | 14.65 | 3.65 |
| MC-Net | | | 83.59 | 72.36 | 14.07 | 2.70 |
| SS-Net | | | 86.33 | 76.15 | 9.97 | 2.31 |
| Ours | | | **88.02**↑1.69 | **78.72**↑2.57 | **7.90**↓2.07 | **2.15**↓0.16 |
| UA-MT | 8(10%) | 72(90%) | 87.79 | 78.39 | 8.68 | 2.12 |
| SASSNet | | | 87.54 | 78.05 | 9.84 | 2.59 |
| DTC | | | 87.51 | 78.17 | 8.23 | 2.36 |
| URPC | | | 86.92 | 77.03 | 11.13 | 2.28 |
| MC-Net | | | 87.62 | 78.25 | 10.03 | 1.82 |
| SS-Net | | | 88.55 | 79.62 | 7.49 | 1.90 |
| Ours | | | **89.62**↑1.07 | **81.31**↑1.69 | **6.81**↓0.68 | **1.76**↓0.14 |

Table 1. Comparisons with state-of-the-art semi-supervised segmentation methods on LA dataset. Improvements compared with the second best results are highlighted.

| Method | Scans used | | Metrics | | | |
|---|---|---|---|---|---|---|
| | Labeled | Unlabeled | Dice↑ | Jaccard↑ | 95HD↓ | ASD↓ |
| V-Net | 12(20%) | 50(80%) | 69.96 | 55.55 | 14.27 | **1.64** |
| DAN | | | 76.74 | 63.29 | 11.13 | 2.97 |
| ADVNET | | | 75.31 | 61.73 | 11.72 | 3.88 |
| UA-MT | | | 77.26 | 63.82 | 11.90 | 3.06 |
| SASSNet | | | 77.66 | 64.08 | 10.93 | 3.05 |
| DTC | | | 78.27 | 64.75 | 8.36 | 2.25 |
| CoraNet | | | 79.67 | 66.69 | 7.59 | 1.89 |
| Ours | | | **82.91**↑3.24 | **70.97**↑4.28 | **6.43**↓1.16 | 2.25↑0.61 |

Table 2. Comparisons with state-of-the-art semi-supervised segmentation methods on the Pancreas-NIH dataset.

| Method | Scans used | | Metrics | | | |
|---|---|---|---|---|---|---|
| | Labeled | Unlabeled | Dice↑ | Jaccard↑ | 95HD↓ | ASD↓ |
| U-Net | 3(5%) | 0 | 47.83 | 37.01 | 31.16 | 12.62 |
| U-Net | 7(10%) | 0 | 79.41 | 68.11 | 9.35 | 2.70 |
| U-Net | 70(All) | 0 | 91.44 | 84.59 | 4.30 | 0.99 |
| UA-MT | 3(5%) | 67(95%) | 46.04 | 35.97 | 20.08 | 7.75 |
| SASSNet | | | 57.77 | 46.14 | 20.05 | 6.06 |
| DTC | | | 56.90 | 45.67 | 23.36 | 7.39 |
| URPC | | | 55.87 | 44.64 | 13.60 | 3.74 |
| MC-Net | | | 62.85 | 52.29 | 7.62 | 2.33 |
| SS-Net | | | 65.83 | 55.38 | 6.67 | 2.28 |
| Ours | | | **87.59**↑21.76 | **78.67**↑23.29 | **1.90**↓4.77 | **0.67**↓1.61 |
| UA-MT | 7(10%) | 63(90%) | 81.65 | 70.64 | 6.88 | 2.02 |
| SASSNet | | | 84.50 | 74.34 | 5.42 | 1.86 |
| DTC | | | 84.29 | 73.92 | 12.81 | 4.01 |
| URPC | | | 83.10 | 72.41 | 4.84 | 1.53 |
| MC-Net | | | 86.44 | 77.04 | 5.50 | 1.84 |
| SS-Net | | | 86.78 | 77.67 | 6.07 | 1.40 |
| Ours | | | **88.84**↑2.06 | **80.62**↑2.95 | **3.98**↓2.09 | **1.17**↓0.23 |

Table 3. Comparisons with state-of-the-art semi-supervised segmentation methods on the ACDC dataset.

# 3. Experiments

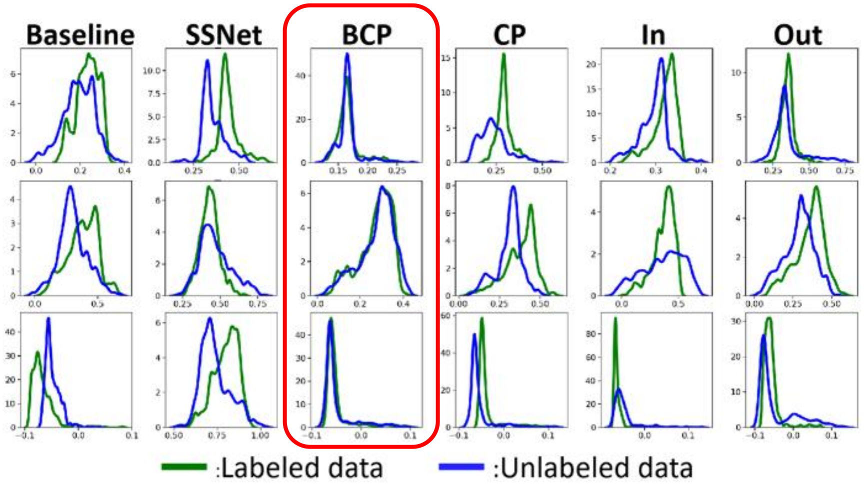



Figure 5. Kernel dense estimations of different methods, trained on 10% labeled ACDC dataset. Top to bottom are kernel dense estimations of features belong to three different class of ACDC: right ventricle, myocardium and left ventricle. Baseline: Only labeled data are used to train the network. *CP*, *In* and *Out* are same as Table 4. It can be seen that our BCP could make the features of labeled data and unlabeled data align better. Furthermore, the outstanding performance of our method compared with *In* and *Out* demonstrates the necessity of *bidirectional* copy-paste.

| BCP | nms | Pre-Train | Dice↑ | Jaccard↑ | 95HD↓ | ASD↓ |
|---|---|---|---|---|---|---|
| | | | 47.62 | 36.61 | 29.02 | 11.46 |
| ✓ | | | 83.26 | 72.71 | 23.90 | 7.49 |
| ✓ | ✓ | | 82.33 | 72.76 | 9.78 | 4.74 |
| ✓ | ✓ | ✓ | **87.59** | **78.67** | **1.90** | **0.67** |

Table 10. Ablation on ACDC dataset with 5% labeled data, $\alpha = 0.5$ across all experiments. nms: Post-processing the pseudo-labels for unlabeled data. Pre-Train: Initialized from a pre-trained model with copy-paste on labeled data.