ECCV'20 ONLINE
23-28 AUGUST 2020

16TH EUROPEAN CONFERENCE ON
COMPUTER VISION

WWW.ECCV2020.EU

Laboratory of Image Science and Technology

Deep Complementary Joint Model for Complex Scene Registration and Few-shot Segmentation on Medical Images

Yuting He[1], Tiantian Li[1], Guanyu Yang[1,2*], Youyong Kong[1,2], Yang Chen[1,2], Huazhong Shu[1,2], Jean-Louis Coatrieux[2,3], Jean-Louis Dillenseger[2,3], Shuo Li[4]

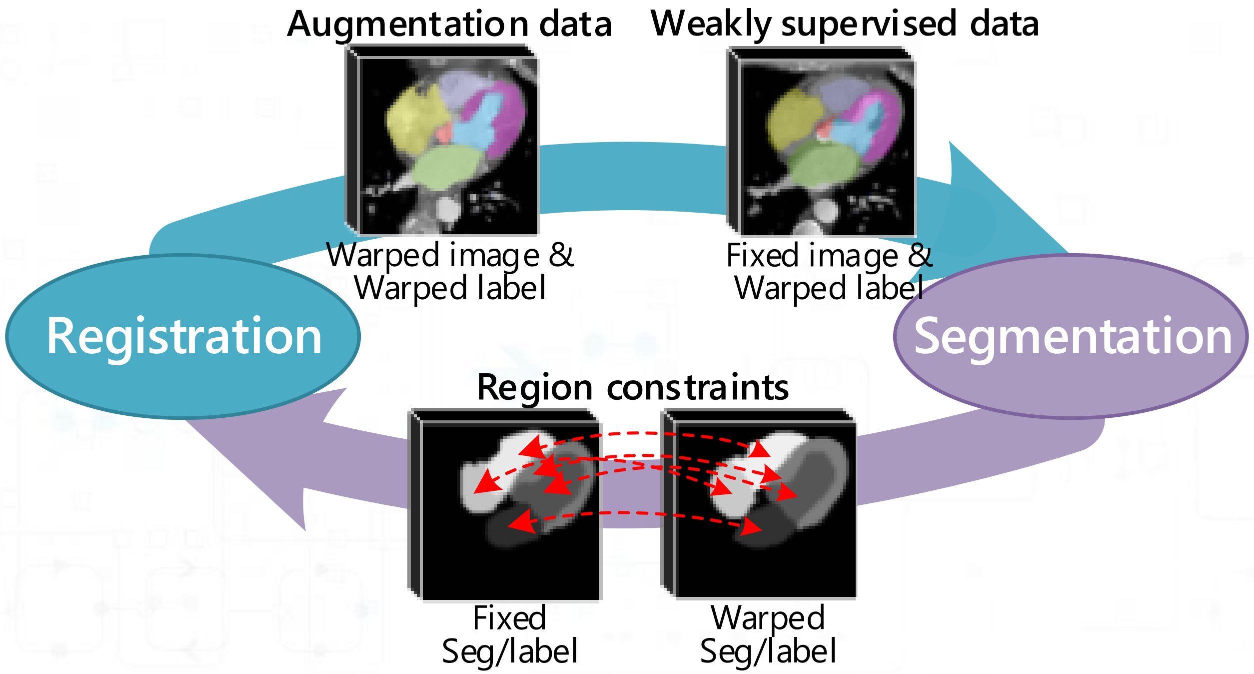
[1]Laboratory of Image Science and Technology, Southeast University, Nanjing 210096, China
[2]Centre de Recherche en Information Biom´ edicale Sino-Fran¸cais (CRIBs)
[3]Univ Rennes, Inserm, LTSI - UMR1099, Rennes, F-35000, France
[4]Dept. of Medical Biophysics, University of Western Ontario, London, ON, Canada
***yang.list@seu.edu.cn**

Western

Complementarity topology

Registration and segmentation tasks has great complementarity.

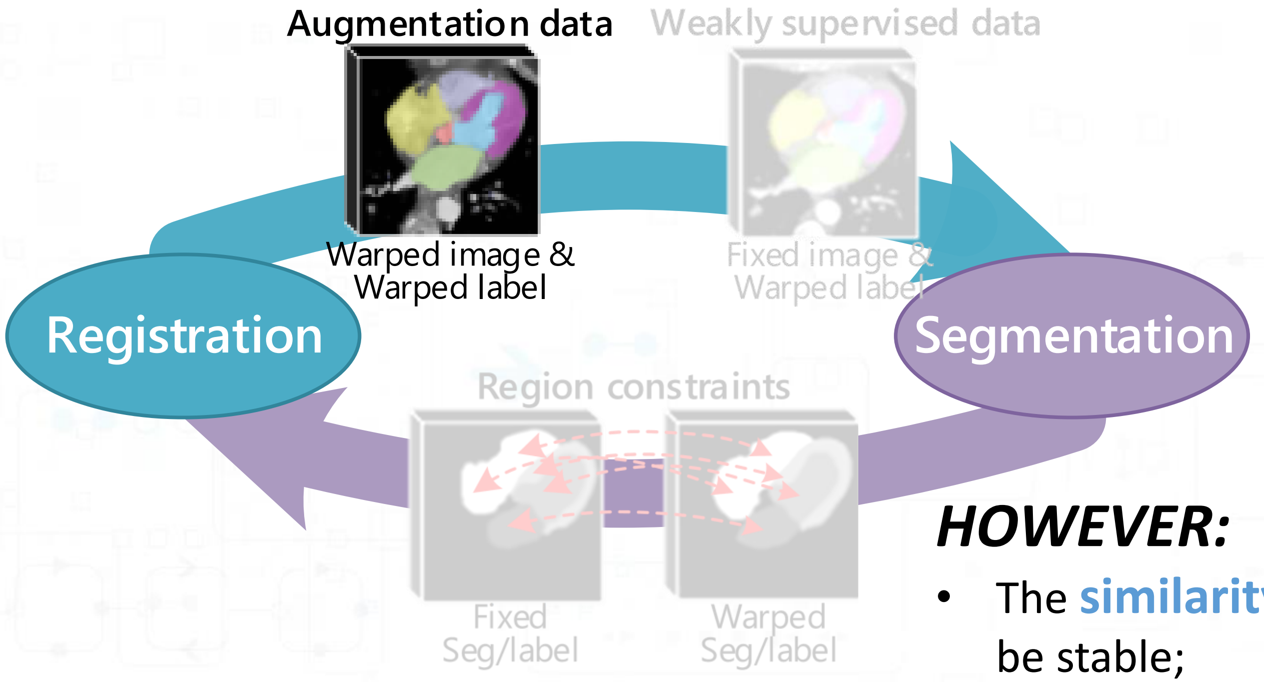
Complementarity topology

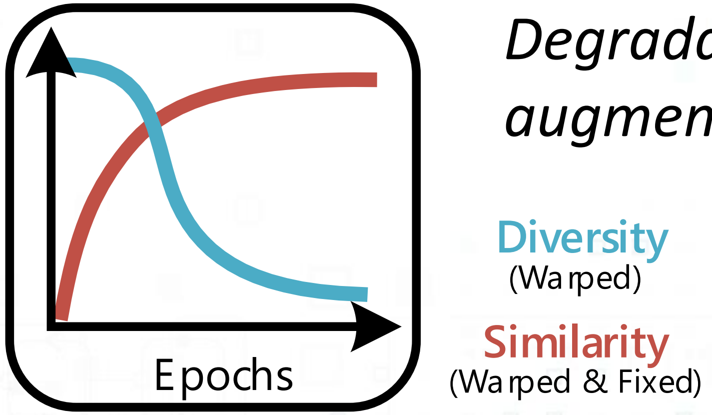


Registration and segmentation tasks has great complementarity.

- Registration model provides diverse ***augmentation data*** or ***weakly supervised data*** for segmentation model, reducing the requirement of labels and enhancing the segmentation generalization in few-shot situation.
- Segmentation model feeds back ***region constraints*** so that additional attention on ROIs is paid for finer registration in complex scene.

Limitation 1





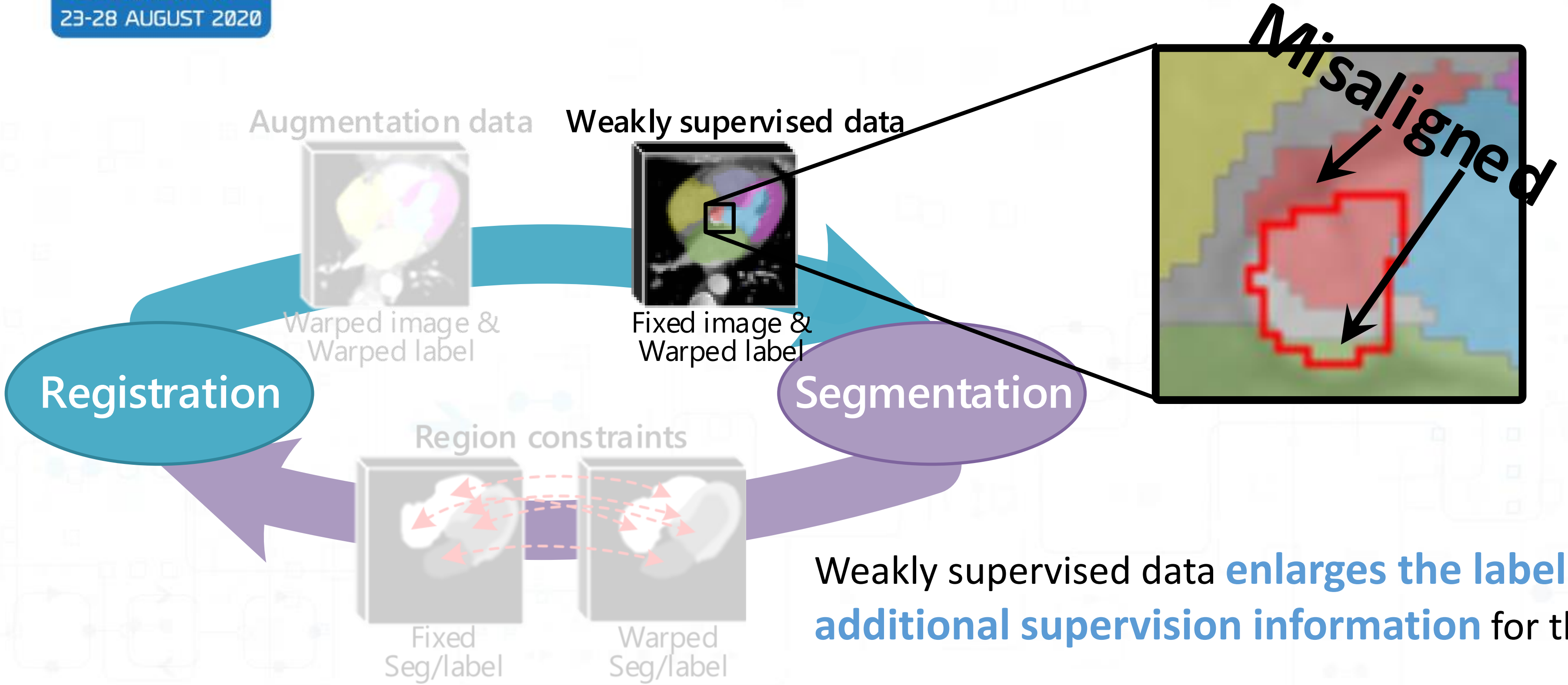
Degradation of data augmentation capability

HOWEVER:

- The **similarity** between warped and fixed images **increases** and tends to be stable;
- The **diversity** of warped images is **reduced** as the similarity stabilizes.

➔ Identical warped images are generated in different epochs, resulting in **the reduction of augmentation data diversity**. Thus, the data augmentation ability is degraded and the further enhancement of segmentation will be limited.

Limitation 2

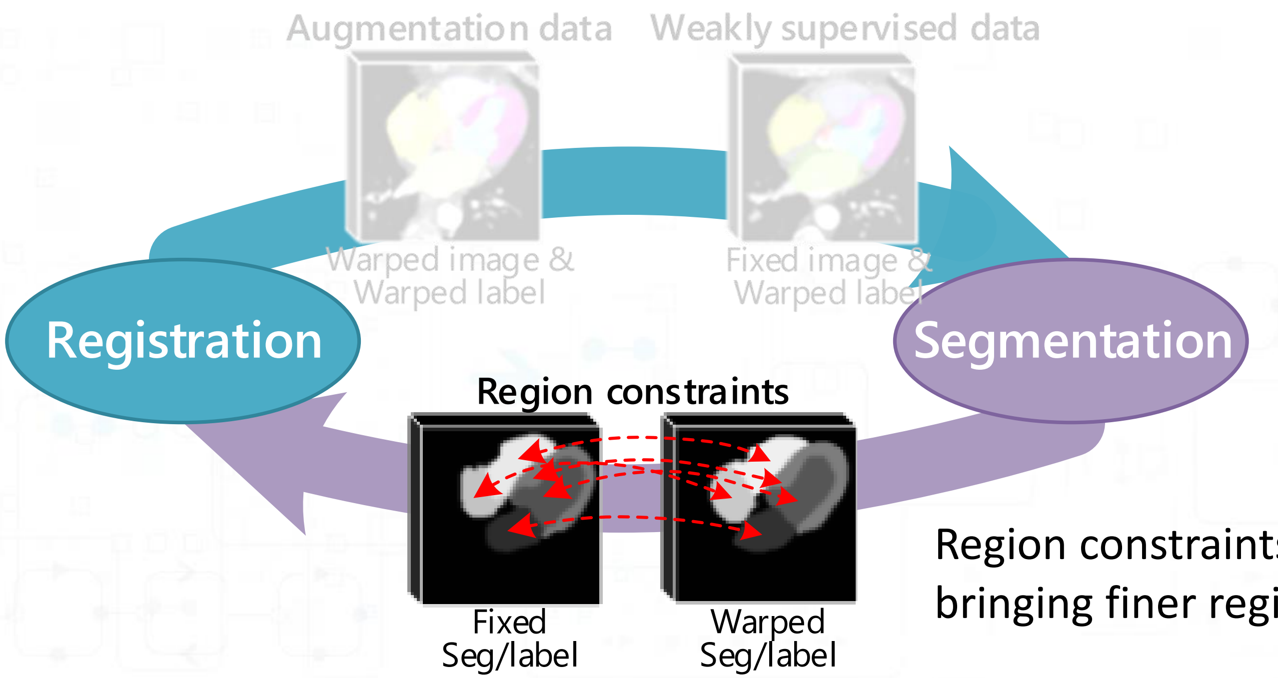


Misaligned regions in weakly supervised data

Weakly supervised data **enlarges the labeled dataset and provide additional supervision information** for the segmentation model.

HOWEVER: Large misaligned regions in these data will **produce incorrect optimization targets and it will disturb the training process** leading to serious mis-segmentation if used directly.

Limitation 3



Lack of label-based region constraints

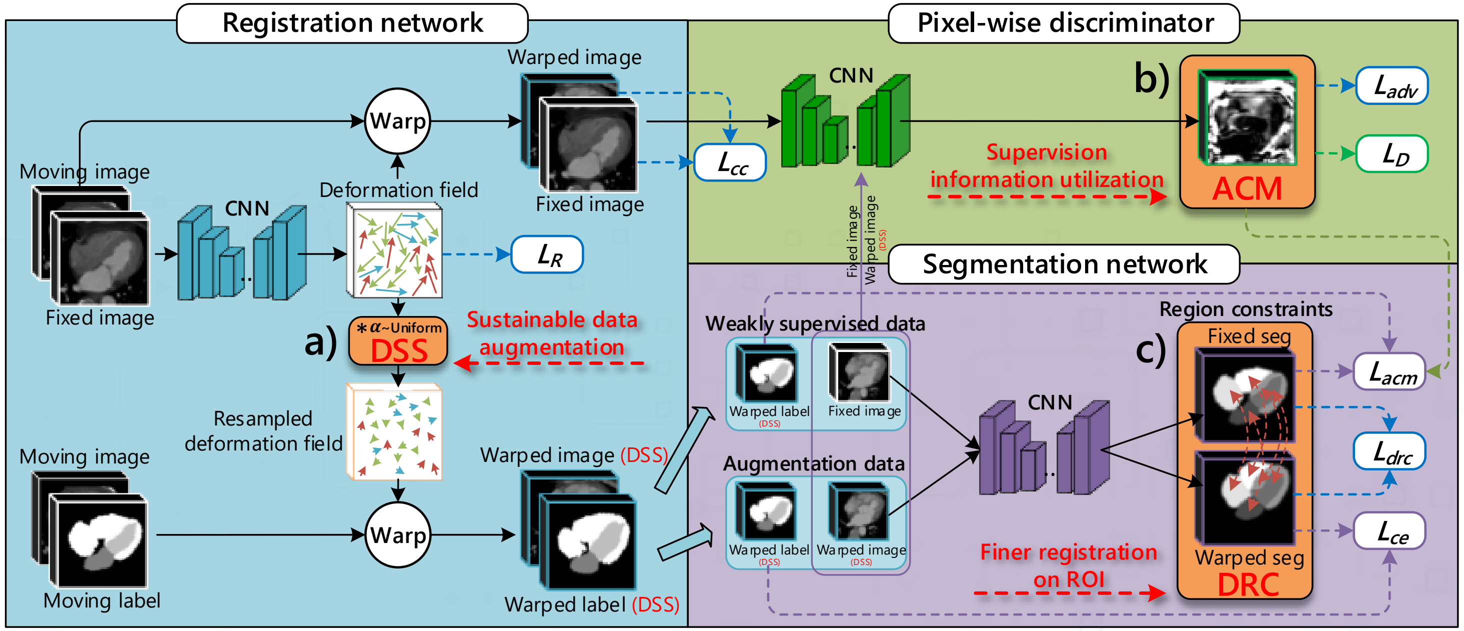
Region constraints provide **specific alignment information for regions** bringing finer registration optimization.

HOWEVER: In few-shot situation, the label-based region constraints are lacked with few labels. Especially, in complex scene, the registration model will **take rough optimization and the complex backgrounds will limit the registration performance** on ROIs.

Proposed Solution: DeepRS

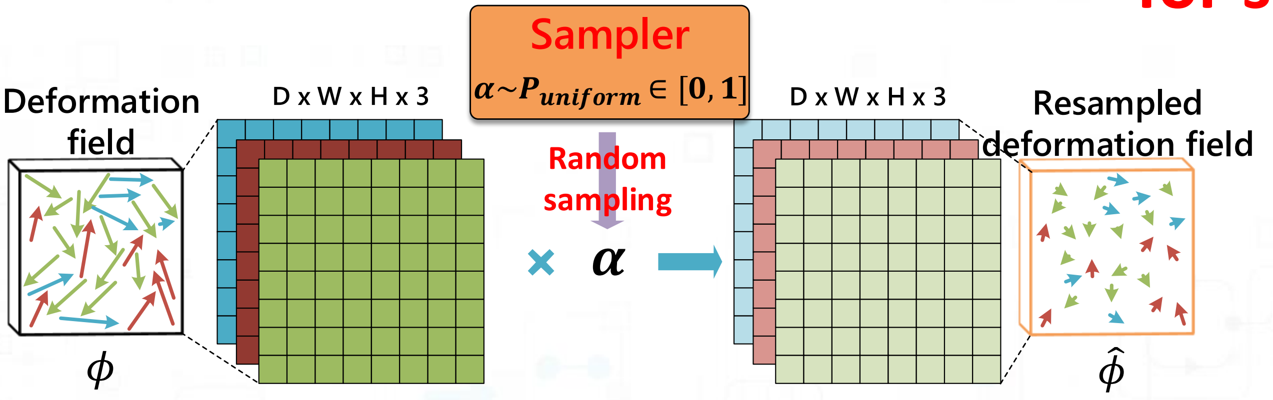
DeepRS minimizes background interference in complex scene registration, and greatly reduces the label requirements of few-shot segmentation via:

- **Deep Structure Sampling** block for sustainable data augmentation;
- **Alignment Confidence Map** method for supervision information utilization;
- **Deep-based Region Constraint** strategy for finer registration on ROIs.



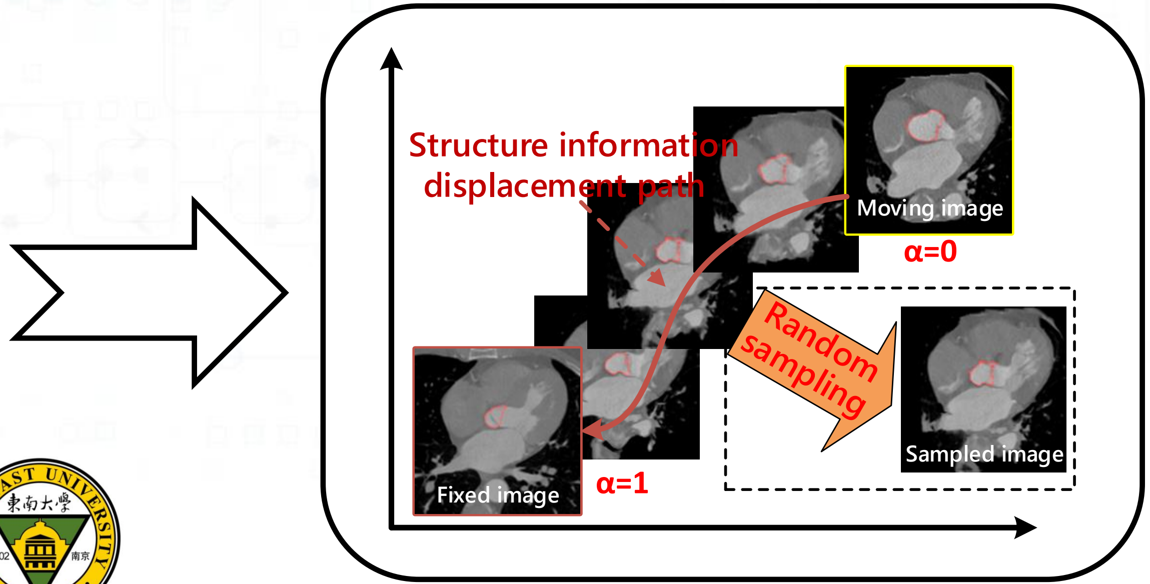
Solution 1: Deep Structure Sampling

for sustainable data augmentation



$$\hat{\emptyset} = \emptyset \times \alpha \sim P_{uniform} \in [0,1]$$

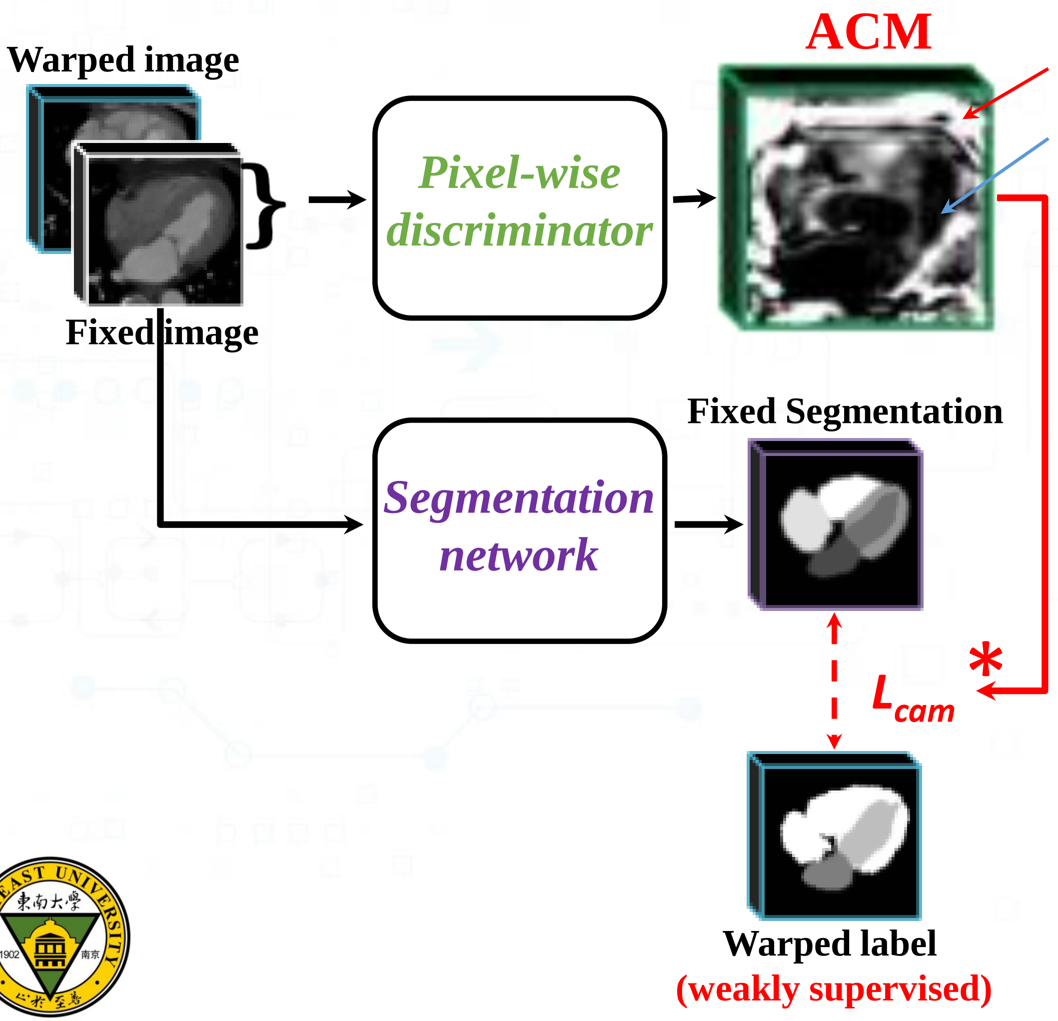
Sustainable data augmentation. The perturbation factor controls the deformation degree so that the registration network generates diverse augmentation data sustainably.



Registration makes the **structure information displacement** and our DSS **samples the information on its displacement** path.

Solution 2: Alignment Confidence Map

for supervision information utilization



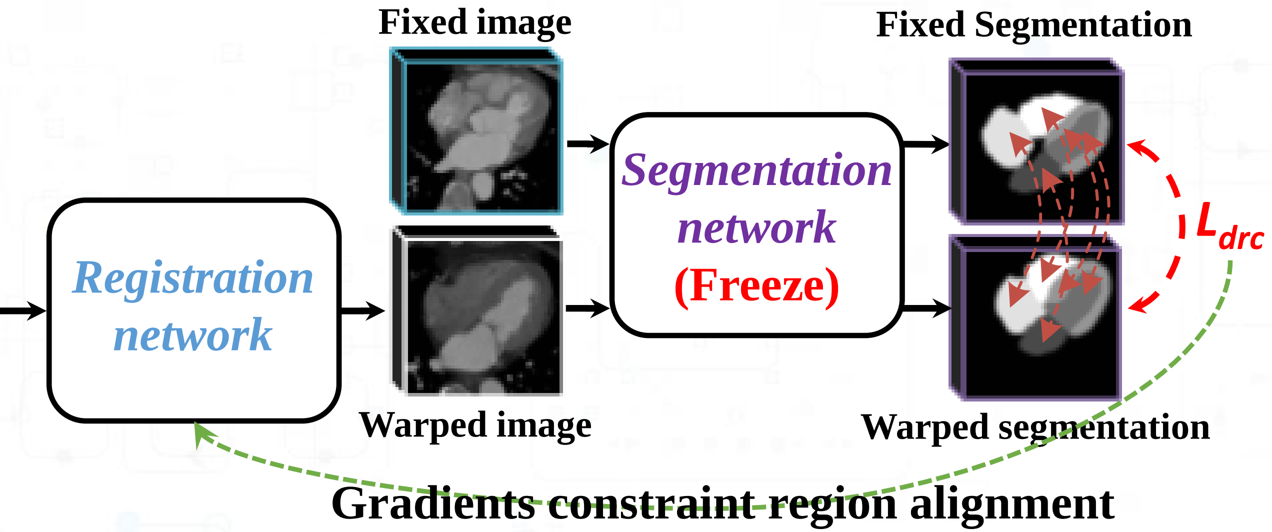
- The ACM maps evaluate the pixel-wise similarity between warped and fixed images and will **highlight the aligned regions**.
- ACMs weights the cross-entropy loss between warped labels and fixed segmentation, thus the **loss value in misaligned region will be suppressed.**

$$\mathcal{L}_{acm} = -D(W(x_m, \hat{\varnothing}), x_f) W(y_m, \hat{\varnothing}) \log S(x_f)$$

- The **contribution** of the weakly supervised data is **increasing** during the training, because the **registration network defeats the discriminator**.

Solution 3: Deep-based Region Constraint

for finer registration on ROIs



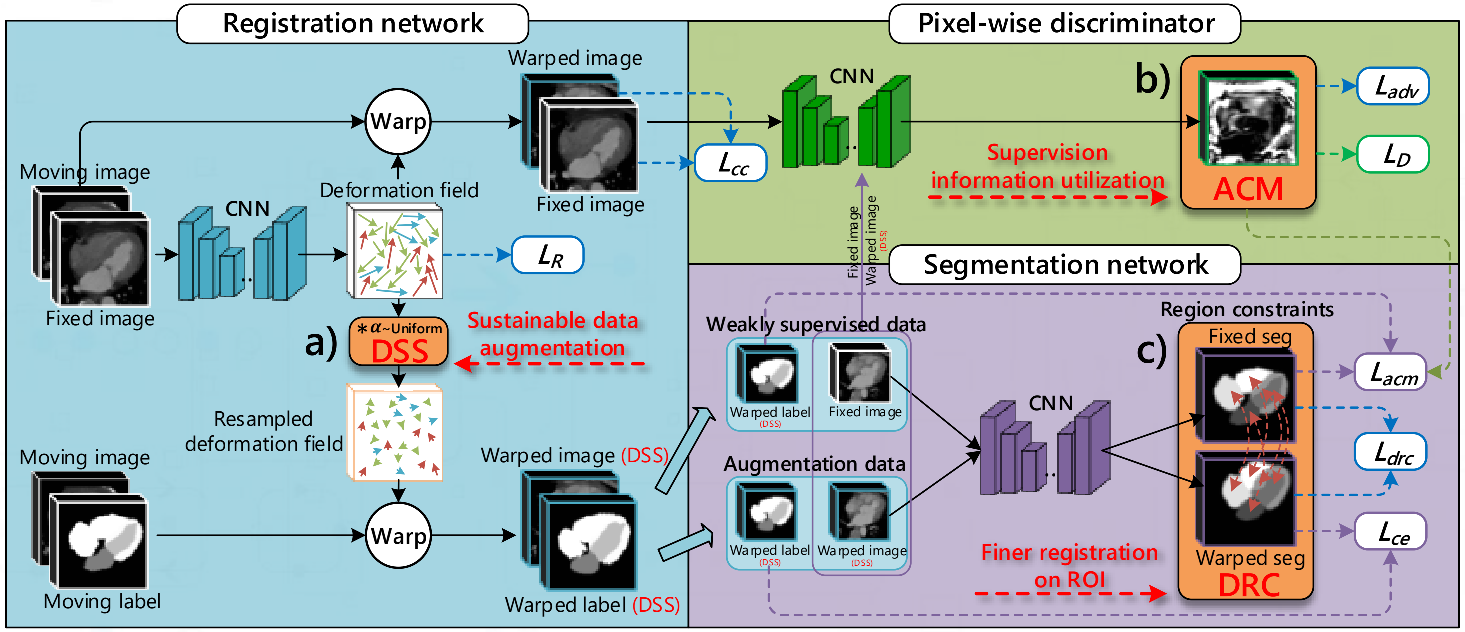
DRC strategy takes the alignment of the corresponding regions in warped and fixed images as the optimization target, so that

- label requirements of **label-based region constraints is freed** in few-shot situation;
- **additional region attention on the ROIs is paid** for finer registration.

Mean square error loss between these two seg-masks is calculated to optimize the registration network:

$$\mathcal{L}_{drc} = -(S(W(x_m, \hat{\varnothing})) - S(x_f))^2$$

Overview



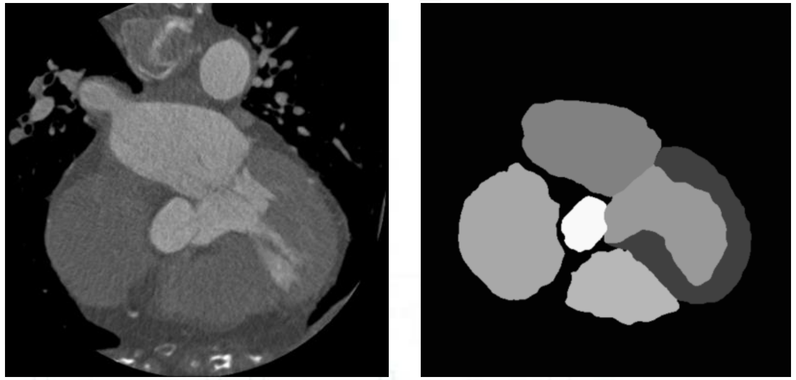
$$\mathcal{L}_{reg} = \lambda_{adv}\mathcal{L}_{adv} + \lambda_{drc}\mathcal{L}_{drc} + \lambda_{cc}\mathcal{L}_{cc} + \lambda_R\mathcal{L}_R$$

$$\mathcal{L}_{seg} = \lambda_{acm}\mathcal{L}_{acm} + \lambda_{ce}\mathcal{L}_{ce}$$

$$\mathcal{L}_D = -\log(D(x_r, x_f)) - \log(1 - D(x_w, x_f))$$

Experiment Setting

MICCAI 2017 Multi-Modality Whole Heart Segmentation CT dataset [1] with 20 labeled CT images and 40 unlabeled images. 40 unlabeled images and 4 labeled images as training set making few-shot situation, 16 labeled images as testing set.



Crop the hearts → Affine transformation → Normalization → Model training

Crop the hearts: Crop the heart regions to unify the information in different images.

Affine transformation: Perform affine alignment on the images constructing $n*(n-1)$ image pairs

Normalization: $(X - \text{mean})/\text{std}$

Model training: Put the image into the segmentation and registration joint model for model training.

[1] X. Zhuang and J. Shen, "Multi-scale patch and multi-modality atlases for whole heart segmentation of mri," Medical Image Analysis, vol. 31, pp. 77 – 87, 2016. [Online]. Available: http: //www.sciencedirect.com/science/article/pii/S1361841516000219

Results: Metrics

Table 1. The proposed DeepRS model achieves the state-of-the-art performance both in registration (R) and segmentation (S) tasks on cardiac CT data.

Method	R-Dice	S-Dice
Affine only	64.6±10.7	-
VoxelMorph-2[2]	71.7±10.6	-
Adv-Reg[7]	68.8±10.7	-
3D U-Net[3]	-	78.8±9.2
3D U-Net-aug[3]	-	80.0±12.0
3D FCN[24]	-	71.4±11.3
V-Net[25]	-	69.8±10.9
DeepAtlas[37]	71.3±10.5	81.8±7.5
HybridCNN[22]	69.2±10.3	78.8±7.9
DeepRS(Ours)	**77.6±7.9**	**85.7±7.7**

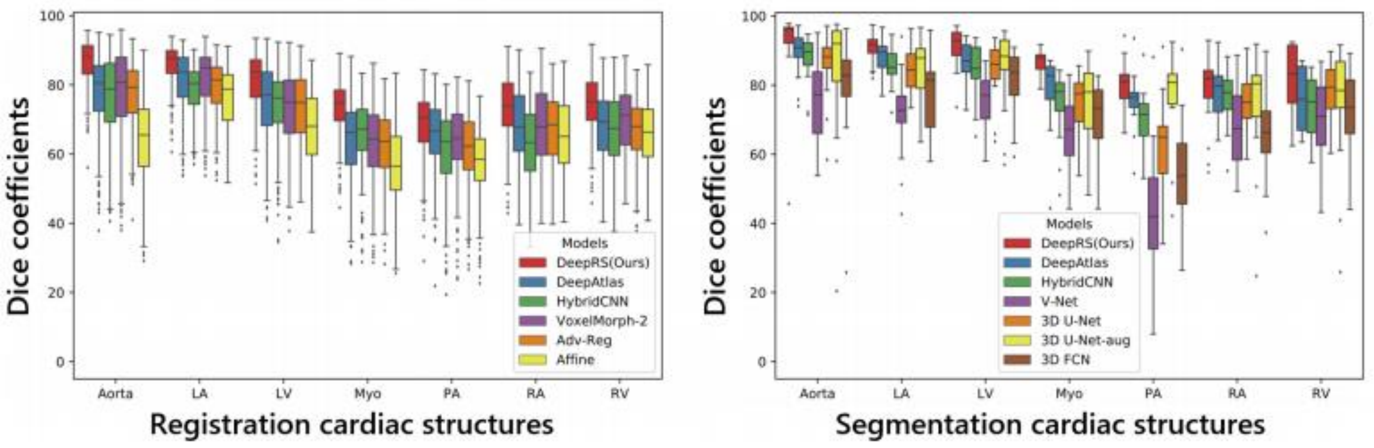


Fig. 5. Our DeepRS achieves excellent dice coefficients on each structure. The box plots shows the proposed DeepRS (red box) model achieves the state-of-the-art performance in complex scene registration (*Left*) and few-shot segmentation (*Right*).

Results: Visual Analysis

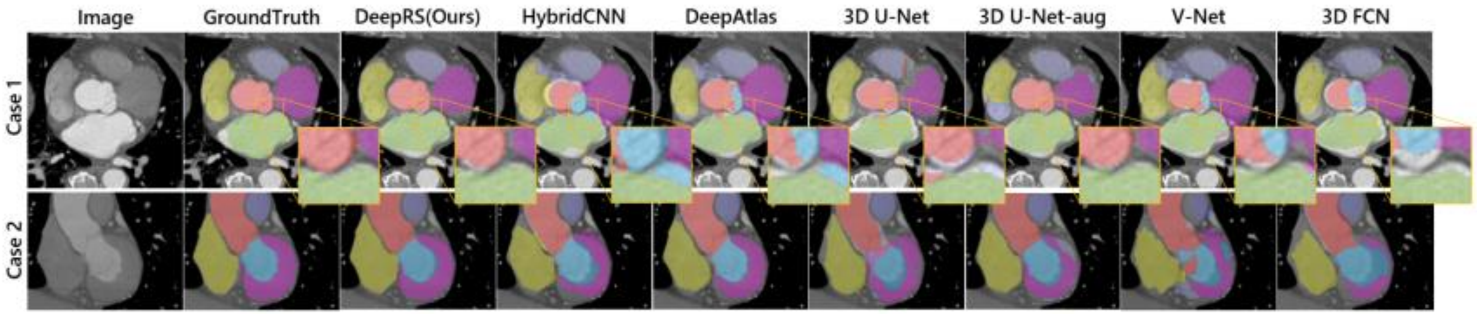


Fig. 7. Our DeepRS brings higher segmentation generalization ability trained on 4 labeled images. Yellow boxes show the excellent generalization ability in detail. The example slices from 3D CT image show the regions of Aorta (red), RA (yellow), RV (purple), Myo (pink), LV (green) and LV (blue).

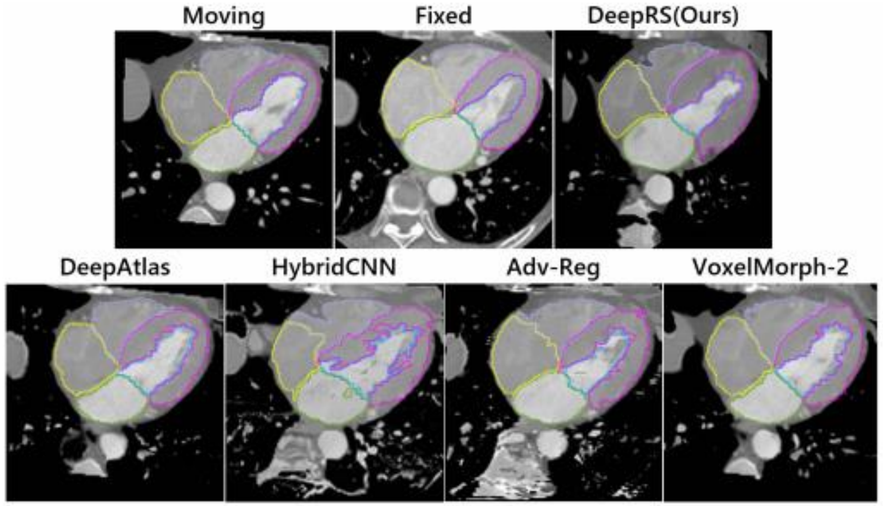


Fig. 6. Our DeepRS gets finer registration on ROIs. The example slices from 3D CT image show the overlaid boundaries of the LV (green), RA (yellow), RV (purple), LV (blue) and Myo (pink). Our model makes these structures in moving image alike structures in fixed image.

Results: Few-shot segmentation

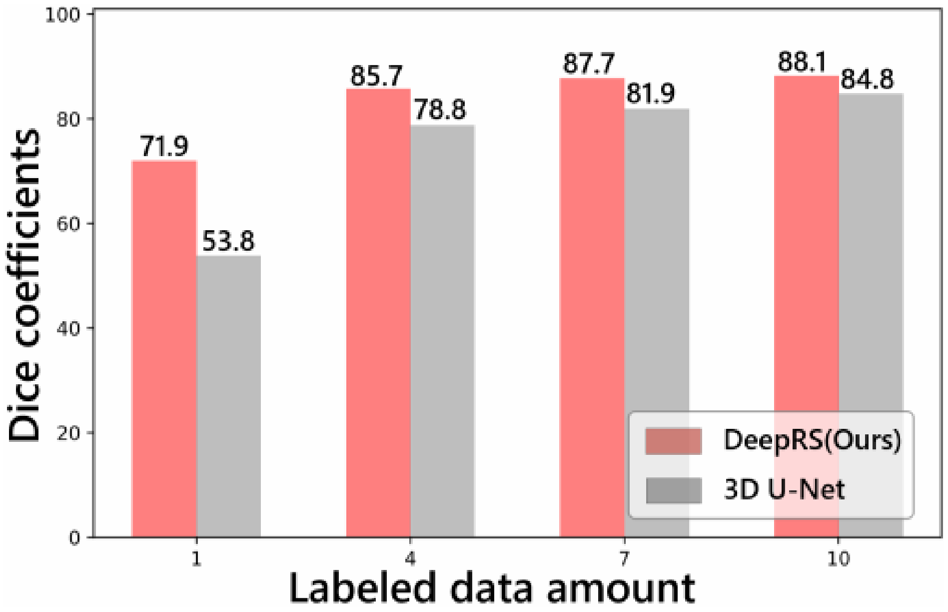


Fig. 8. Especially in few-shot situation, the segmentation network in our DeepRS model achieves much higher mean dice coefficients of all structures than 3D U-Net[3].

Table 2. The ablation study analyses the contributions of our innovations.

R	S	DSS	ACM	DRC	R-Dice	S-Dice
✓					72.2±10.3	-
	✓				-	78.8±9.2
✓	✓				72.9±10.4	80.5±10.2
✓	✓	✓			72.9±9.6	83.9±8.3
✓	✓		✓		72.5±10.1	84.1±8.3
✓	✓			✓	75.9±9.1	82.5±9.2
✓	✓	✓	✓	✓	**77.6±7.9**	**85.7±7.7**

Conclusion

- To the best of our knowledge, we build a novel complementary topology of registration and segmentation for the first time, and propose the DeepRS model utilizing the data generation ability of registration for few-shot segmentation, and the label-free region constraint ability of segmentation for complex scene registration.
- We propose a deep structure sampling (DSS) block adding a random perturbation factor to the registration for sustainable data augmentation ability.
- We propose an alignment confidence map (ACM) method which efficiently utilizes the supervision information in weakly supervised data thus bringing powerful segmentation generalization.
- We propose a deep-based region constraint (DRC) strategy which frees up the label requirements of label-based methods achieving finer registration on ROIs.