



MMS: Morphology-mixup Stylized Data Generation for Single Domain Generalization in Medical Image Segmentation

Xiaochen He¹, Baoyao Yang¹, Fei Lyu²

¹ School of Computers, Guangdong University of Technology

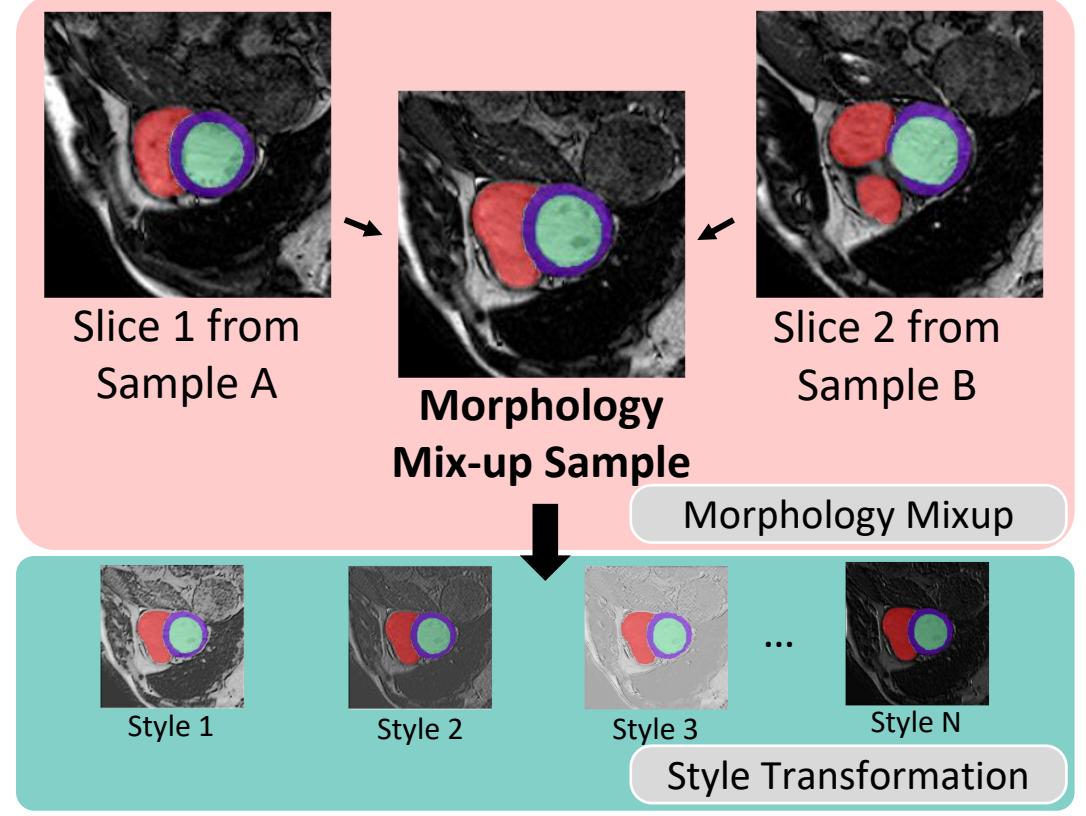
² Department of Computer Science, Hong Kong Baptist University

Representer: Yang Lu, Xiamen University

➤ Overview

We focus on increase the diversity and informativeness of training samples for single-source domain generalization.

1. We propose a date generation method considering the morphology and style.
2. We design a new model that simulate probable morphological structures by morphology mix-up.
3. The proposed method *outperform existing benchmarks* in three cross-domain medical segmentation datasets.





Background

- Deep Learning Challenges in Medical Image Segmentation

1. Data Limitation

- i. Training requires extensive, diverse labeled datasets^[1]
- ii. Difficult to compile due to data privacy, annotation costs^[2]

2. Domain Shift

- i. Performance drops when test data distribution differs from training data^[3]
- ii. Common in medical context due to variations in imaging techniques, equipment settings, and scanner models^[4]

[1] Enhancing MR image segmentation with realistic adversarial data augmentation, MIA, 2022

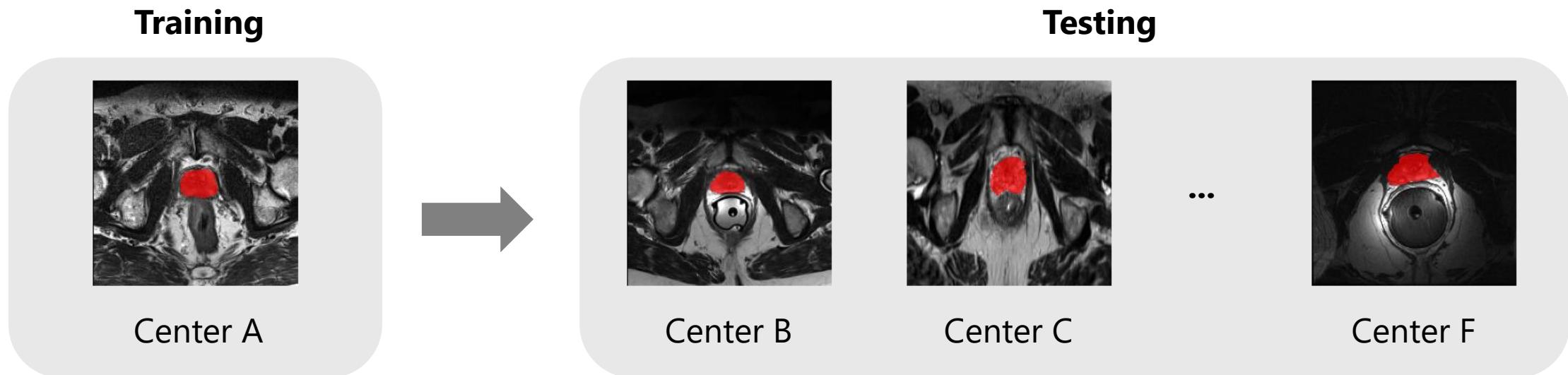
[2] Generalizing Deep Learning for Medical Image Segmentation to Unseen Domains via Deep Stacked Transformation, TMI, 2020

[3] A theory of learning from different domains, Machine Learning, 2010

[4] Machine learning with multi-site imaging data: An empirical study on the impact of scanner effects, ArXiv, 2019

Problem

- Single Source Domain Generalization in Medical Segmentation



Examples from Prostate Dataset^[1]

[1] Shape-aware meta-learning for generalizing prostate mri segmentation to unseen domains, MICCAI, 2020



Related Works

- Adversarial Data Augmentation^{[1][2][3]}
- Random Style Transformation^{[4][5][6]}

Limitations

1. Not faithful enough: not consider characteristic of medical image
2. Not diverse enough: only consider style transformation

[1] Learning to Diversity for Single Domain Generalization, ICCV, 2021

[2] Progressive Domain Expansion Network for Single Domain Generalization, CVPR, 2021

[3] Enhancing MR image segmentation with realistic adversarial data augmentation, MIA, 2022

[4] Generalizing Deep Learning for Medical Image Segmentation to Unseen Domains via Deep Stacked Transformation, TMI, 2020

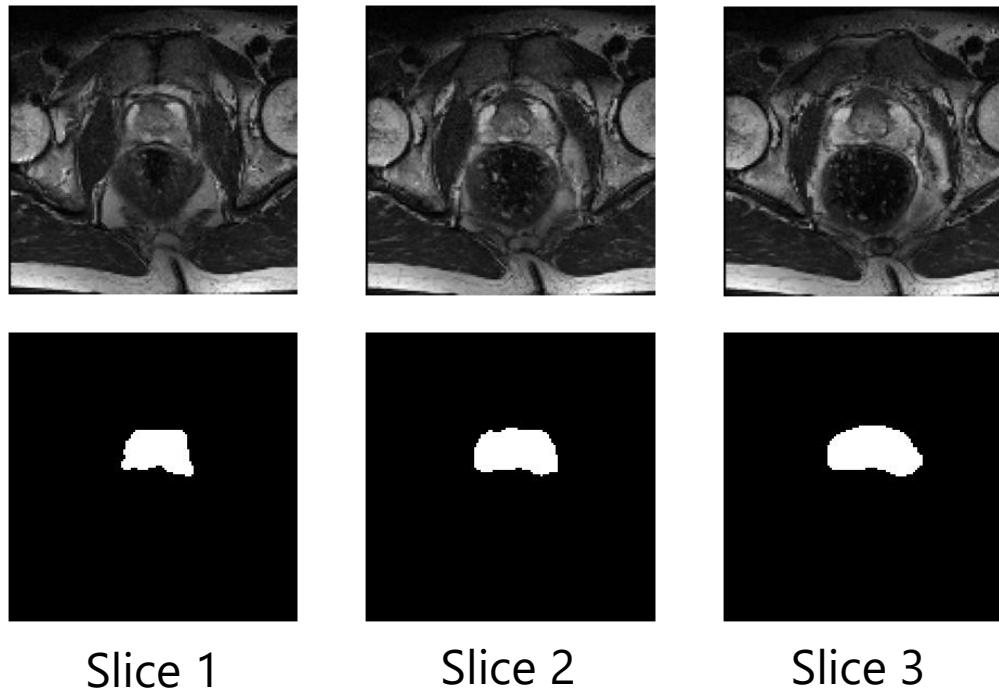
[5] Causality-inspired Single-source Domain Generalization for Medical Image Segmentation, TMI, 2022

[6] Rethinking Data Augmentation for Single-source Domain Generalization in Medical Image Segmentation, AAAI, 2023



Motivation

- Morphology-continuous Changes in Adjacent 2D Slices of Medical Image^[1]



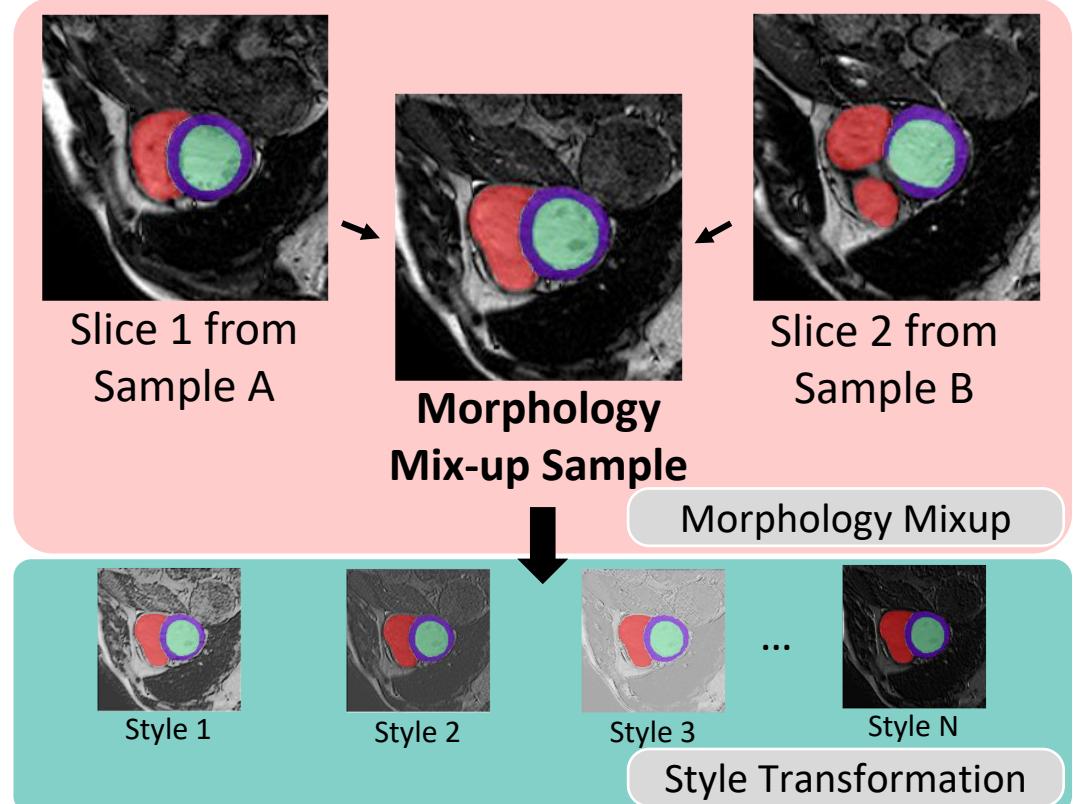
Main Idea

- Main Idea

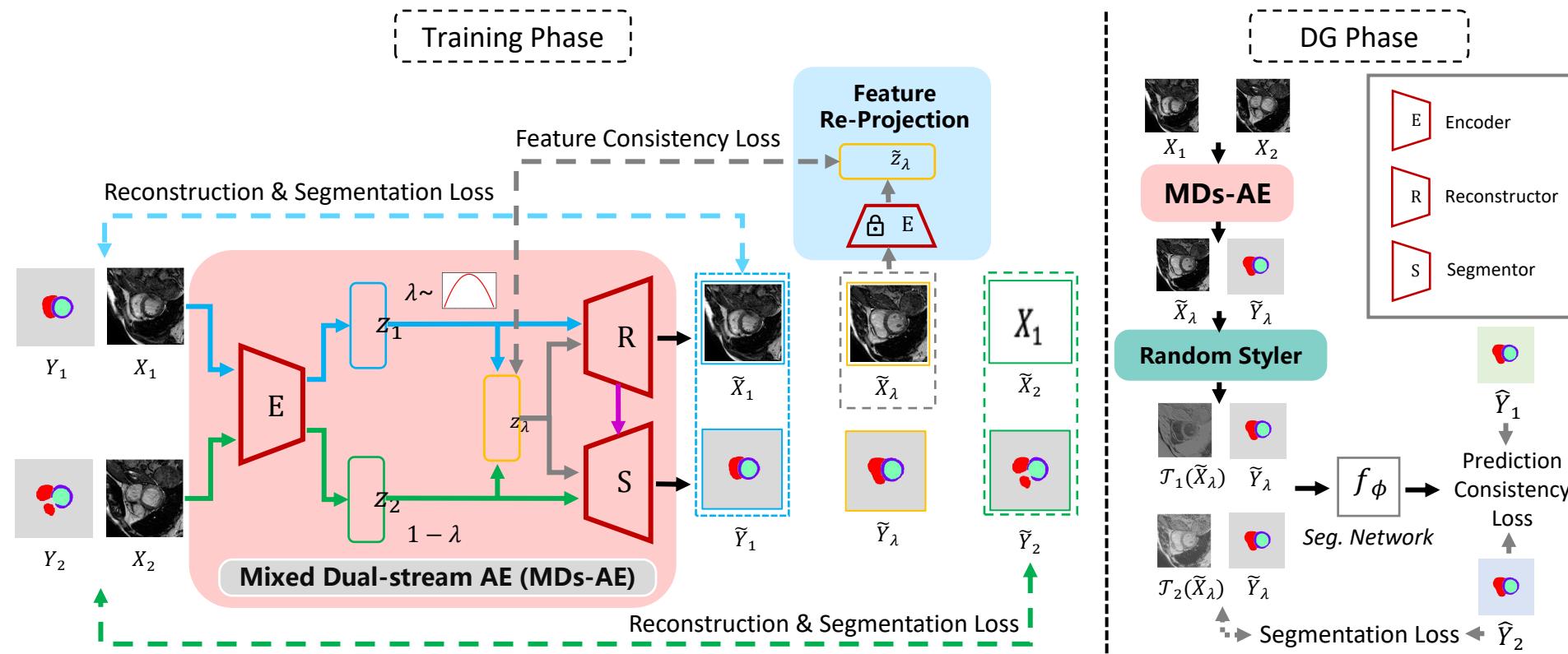
generate novel data with novel morphologies and styles

- Morphology Mix-up

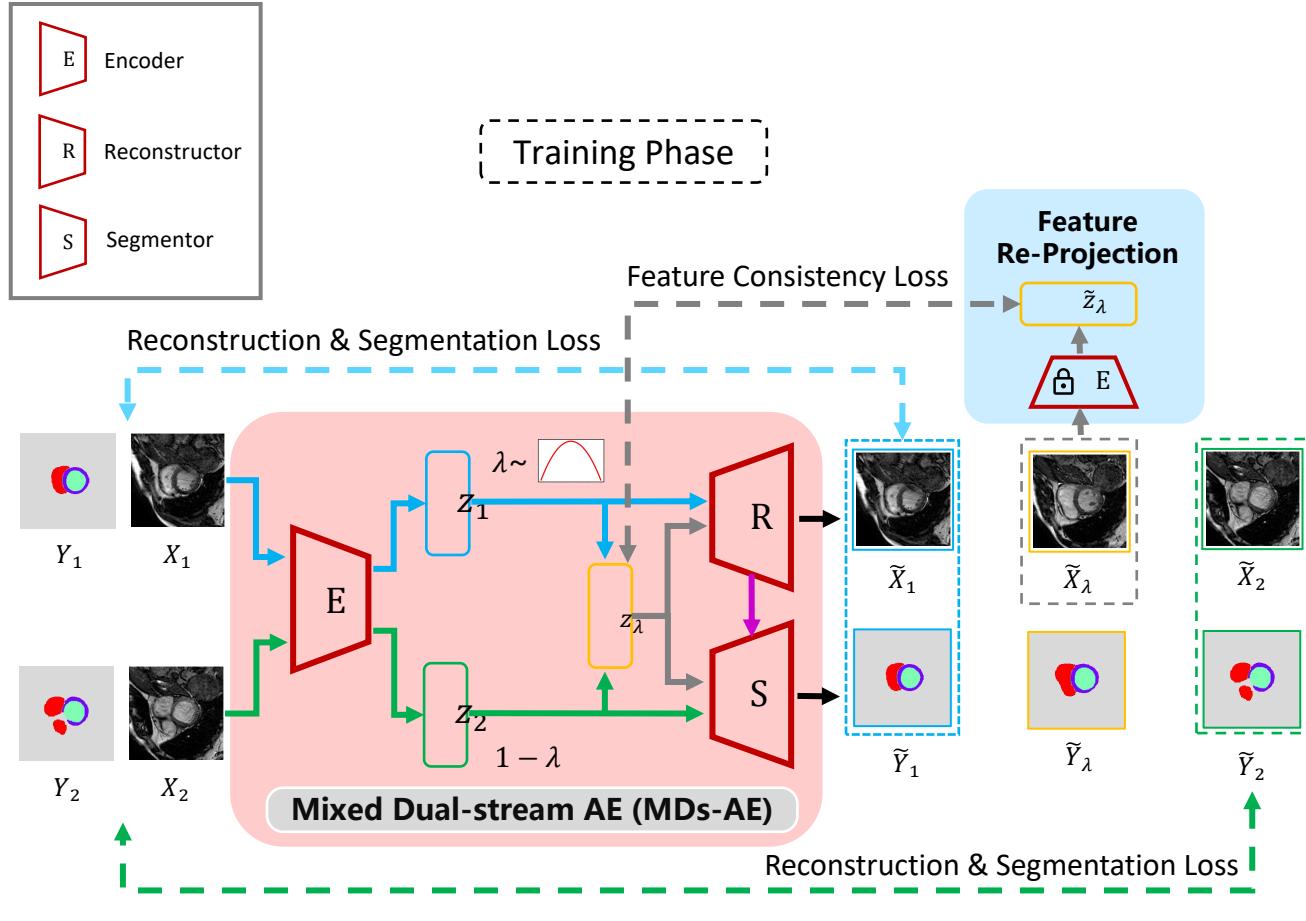
Simulate probable morphological structures between similar slices from different patients



Method



Method



- Mixed Dual-stream Auto-encoder

- Extract morphology-related features

Extract the same bottom feature z from reconstruction and segmentation branches
 $\mathcal{L}_{rec}(X, \tilde{X}) = \mathcal{L}_{MSE}(X, \tilde{X})$

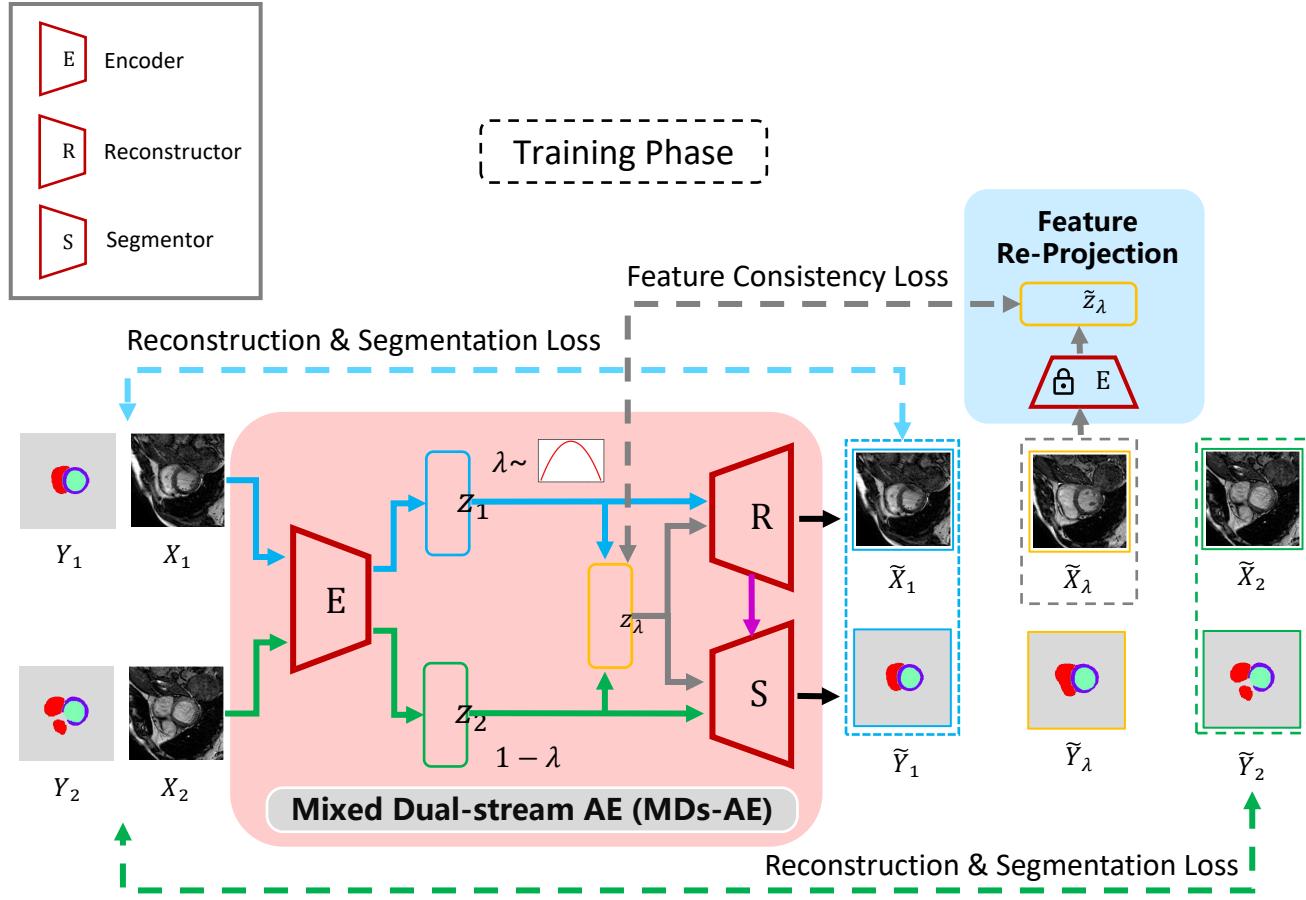
$$\mathcal{L}_{seg}(Y, \tilde{Y}) = \mathcal{L}_{Dice}(Y, \tilde{Y})$$

- Model the morphology shift

Mix-up the z and project it back to image and segmentation mask
 $z_k \xrightarrow{\text{Beta}(\alpha, \beta)} z_i + \lambda z_j$

$$\tilde{x}_k, \tilde{y}_k = R(z_k), S(z_k)$$

Method



- Feature Consistency Strategy

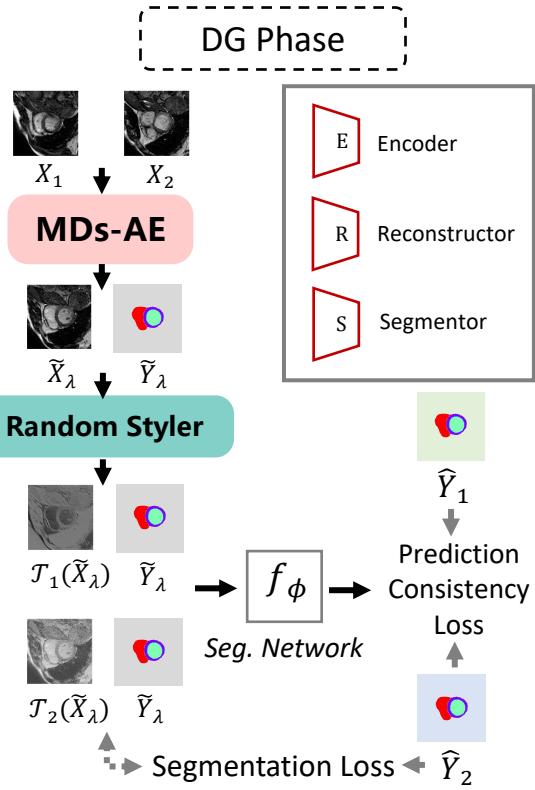
Keep the consistency features from generated image \tilde{z}_λ and the original mixed features z_λ .

$$\mathcal{L}_{fea-con} = \mathcal{L}_{mse}(z_{kr}, \tilde{z}_{kr})$$

• Total Loss for MDs-AE

$$\mathcal{L}_{IDS-AE} = \mathcal{L}_{rec} + \mathcal{L}_{Dice} + \lambda_{fea-con} \mathcal{L}_{fea-con}$$

Method



- Augment Data from Morphology and Style Perspective

$$(\tilde{X}_\lambda, \tilde{Y}_\lambda) = \text{MDs-AE}(X_1, X_1)$$

$$\mathcal{T}_1(\tilde{X}_\lambda) = \text{Random-Styler}(\tilde{X}_\lambda), \mathcal{T}_2(\tilde{X}_\lambda) = \text{Random-Styler}(\tilde{X}_\lambda)$$

(we adopt RandConv¹ as Random Styler)

- Train Segmentor with Consistency Loss

$$\hat{y}_{\lambda,k} = \text{Seg-NET}(\mathcal{T}_k(\tilde{X}_\lambda))$$

$$\mathcal{L}_{pred-con} = \text{KLD}(\hat{y}_{\lambda,1}, \hat{y}_{\lambda,2})$$

$$\mathcal{L}_{DG} = \mathcal{L}_{seg} + \mathcal{L}_{pred-con}$$



Experiment

- Datasets and Generation Setting

Details of cross-domain datasets used in this study

Name	Labels	View	Split	Domains	No. of 3-D scans	Origin
Cross-modality Abdominal	Liver, L-kidney, R-kidney, Spleen	Axial	Source Target	CT T2 MRI	30	[1]
					20	[2]
Cross-sequence Cardiac	L-ventricle, Myocardium, R-ventricle	Short-axis	Source Target	bssFP MRI LGE MRI	45 45	[3]
Cross-center Prostate	Prostate	Axial	1 Source 5 Targets	Prostate MRI from 6 centers	30, 30, 19 13, 12, 12	[4] [5]-[7]

[1] Miccai multi-atlas labeling beyond the cranial vault–workshop and challenge, MICCAI, 2015

[2] Chaos challenge combined (ct-mr) healthy abdominal organ segmentation, Medical Image Analysis, 2021

[3] Cardiac segmentation on late gadolinium enhancement mri: a benchmark study from multi-sequence cardiac mr segmentation challenge, Medical Image Analysis, 2020

[4] Shape-aware meta-learning for generalizing prostate mri segmentation to unseen domains, MICCAI, 2020

[5] Nciisbi 2013 challenge: automated segmentation of prostate structures, The Cancer Archive, 2015

[6] Computer-aided detection and diagnosis for prostate cancer based on mono and multi-parametric mri: a review, Computers in biology and medicine, 2015

[7] Evaluation of prostate segmentation algorithms for mri: the promise12 challenge, MIA, 2014



Experiment

- Implementation Details
 1. ResNet-18 backbone for Encoder, R and S are the reverse of E
 2. U-Net with an EfficientNet-b2 backbone as our task model
- Training Details
 1. Adam optimizer with learning rate decay used for training
 2. 578 generated images, the same as traditionally augmented samples to other methods.
- Pair Selection Details

measure MSE between slices from the whole dataset and pick the closest pair from different patients



Experiment

Segmentation Results on Three Cross-Domain Scenarios. Dice Score is the Evaluation Metric. The Highest are in red.

Method	Cardiac bSSFP-LGE				Abdominal CT-MRI					Prostate Cross-center Avg
	L-ventricle	Myocardium	R-ventricle	Avg	Liver	R-Kidney	L-Kidney	Spleen	Avg	
Upper bound	89.89	80.58	89.97	86.15	91.66	92.26	89.51	87.46	90.22	85.86
Baseline	74.98	53.45	63.09	63.84	81.95	85.75	75.19	65.60	77.12	51.90
Mixup [1]	81.38	68.41	71.89	73.89	77.17	82.61	83.25	59.62	75.66	58.25
MixStyle [2]	86.73	61.48	77.48	75.23	82.34	86.22	78.71	68.27	78.88	55.68
Pix2Pix[3]	82.44	69.10	79.11	76.89	78.83	84.12	80.21	69.12	78.07	60.12
AdvChain [4]	80.15	72.12	82.33	80.15	79.62	83.20	82.29	80.15	81.32	63.34
CSDG [5]	87.59	74.96	86.12	82.89	87.30	86.34	87.73	86.84	87.05	70.01
SLAug [6]	88.48	78.18	87.12	84.60	89.68	88.47	86.65	85.13	87.48	68.44
ours	88.83	79.05	87.93	85.27	87.74	89.13	89.66	86.25	88.19	71.96

[1] mixup: BEYOND EMPIRICAL RISK MINIMIZATION, ICLR, 2018

[2] DOMAIN GENERALIZATION WITH MIXSTYLE, ICLR, 2021

[3] Image translation for medical image generation: Ischemic stroke lesion segmentation, BSPC, 2022

[4] Enhancing MR image segmentation with realistic adversarial data augmentation, MIA, 2022

[5] Causality-inspired Single-source Domain Generalization for Medical Image Segmentation, TMI, 2022¹⁴

[6] Rethinking Data Augmentation for Single-source Domain Generalization in Medical Image Segmentation, AAAI, 2023



Experiment

Dices Score in Cross-Center Prostate Dataset. Dice Score is the Evaluation Metric. The Highest are in red.

	A	B	C	D	E	F	Avg
Upper bound	91.71	86.18	83.08	87.53	87.93	84.72	85.86
Baseline	70.68	56.82	47.71	64.50	42.62	29.07	51.90
Mixup [1]	72.03	60.52	53.98	64.78	54.12	44.08	58.25
MixStyle [2]	67.93	56.31	51.39	68.25	43.78	46.41	55.68
Pix2Pix[3]	68.33	62.23	56.13	69.13	55.91	47.22	60.12
AdvChain [4]	72.98	65.26	59.40	71.46	50.23	60.73	63.34
CSDG [1]	81.73	65.27	63.90	73.61	65.25	70.63	70.01
SLAug [2]	80.87	62.31	63.47	71.93	60.97	71.07	68.44
Ours	82.31	69.34	67.47	74.20	66.98	71.43	71.96

[1] mixup: BEYOND EMPIRICAL RISK MINIMIZATION, ICLR, 2018

[2] DOMAIN GENERALIZATION WITH MIXSTYLE, ICLR, 2021

[3] Image translation for medical image generation: Ischemic stroke lesion segmentation, BSPC, 2022

[4] Enhancing MR image segmentation with realistic adversarial data augmentation, MIA, 2022

[5] Causality-inspired Single-source Domain Generalization for Medical Image Segmentation, TMI, 2021¹⁵

[6] Rethinking Data Augmentation for Single-source Domain Generalization in Medical Image Segmentation, AAAI, 2023



Experiment

- Ablation Study

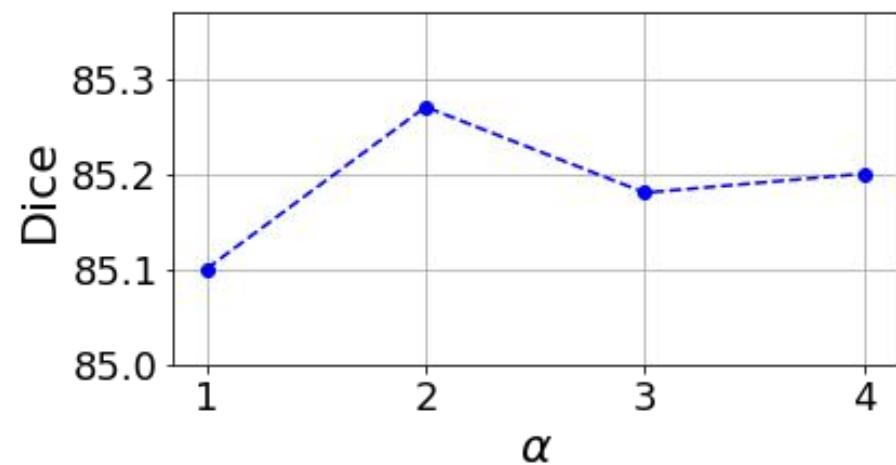
Ablation Study on MMSR in Cardiac Dataset

MDsAE	Random Styler	L-ventricle	Myocardium	R-ventricle	Avg
	✓	74.98	53.45	63.09	63.84
	✓	86.21	73.39	85.79	81.78
	✓	87.57	73.96	86.56	82.70
✓	✓	88.83	79.05	87.93	85.27

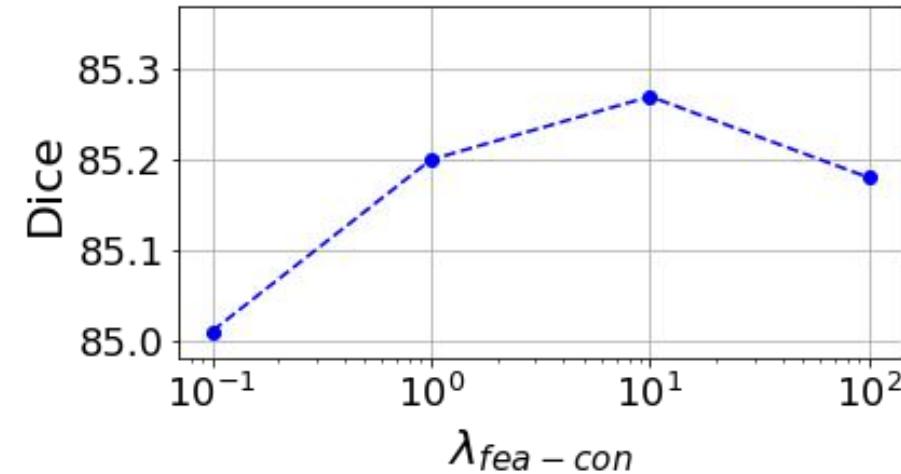
Experiment

- Hyper-parameters Analysis

Performances of Using Different values of α and $\lambda_{fea-con}$ in Cardiac dataset



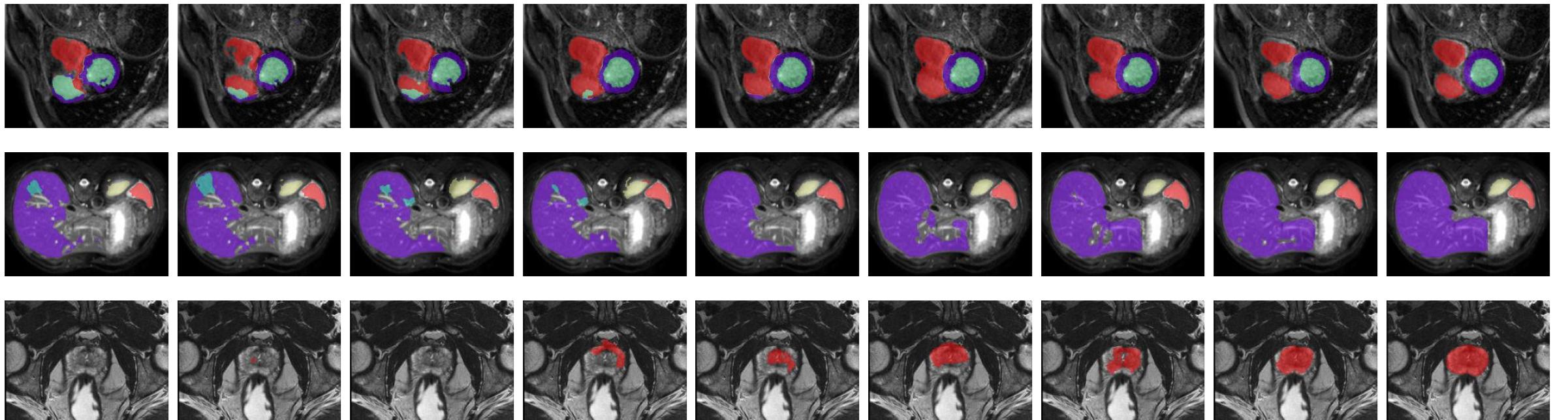
$$z_k = z_i + (1 - \lambda)z_i \quad \lambda \sim Beta(\alpha, \alpha)$$



$$\mathcal{L}_{IDS-AE} = \mathcal{L}_{rec} + \mathcal{L}_{Dice} + \lambda_{fea-con} \mathcal{L}_{fea-con}$$

Experiment

- Comparison Results Visualization



Baseline

Mixup

MixStyle

Pix2Pix

AdvChain

CSDG

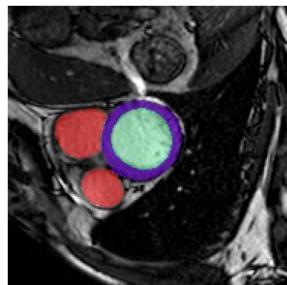
SLAug

Ours

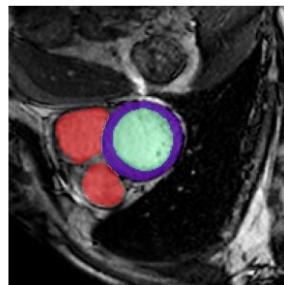
Ground Truth

Experiment

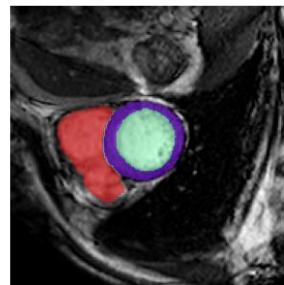
- Visualization of Morphology Shift from MDs-AE



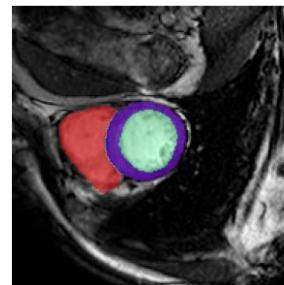
Sample A



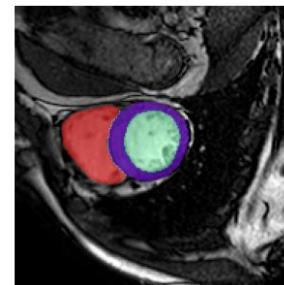
$\lambda = 0.25$



$\lambda = 0.5$



$\lambda = 0.75$



Sample B



Thanks for Listening

MMS: Morphology-mixup Stylized Data Generation for Single Domain Generalization in Medical Image Segmentation

Contact info:

Baoyao Yang (ybaoyao@gdut.edu.cn)