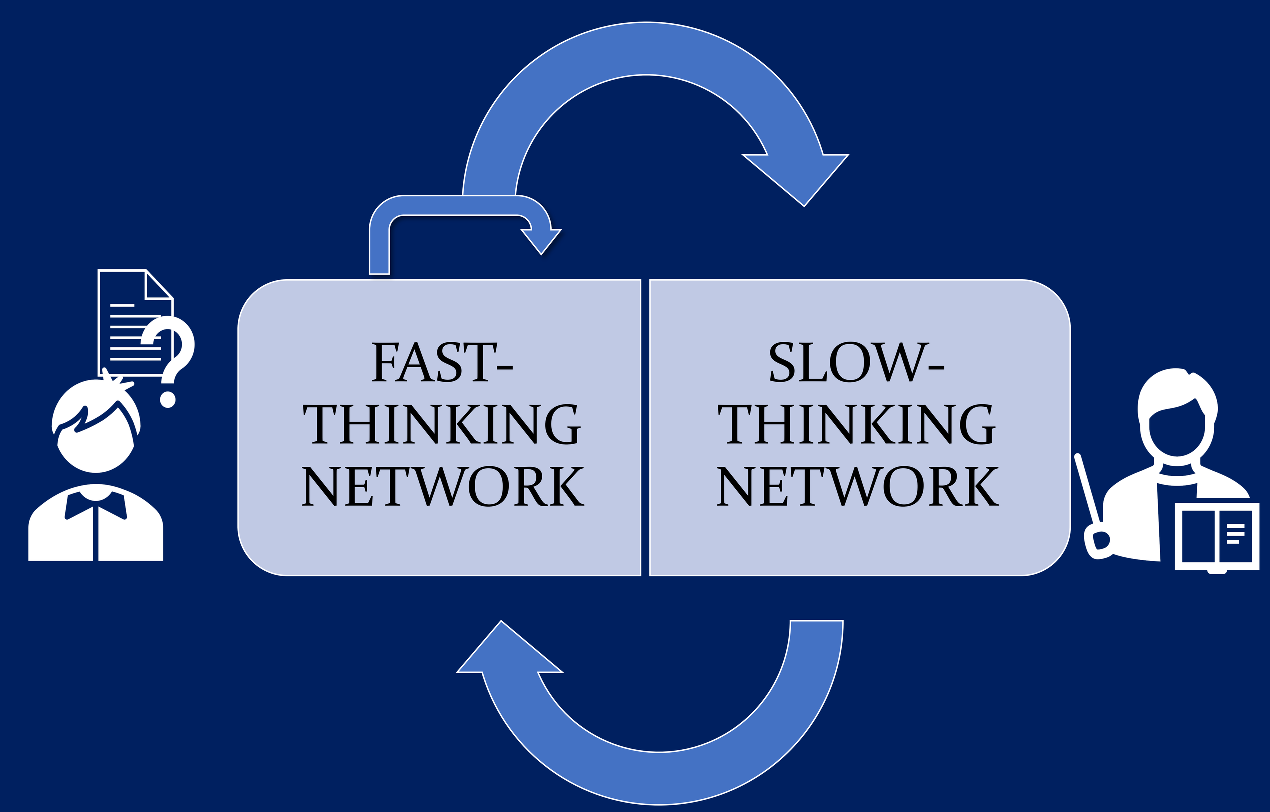
A cooperative training method to improve model **cross-domain** performance using **single-domain** data only

Generate hard examples via **latent space masking** to enhance training



Learn the shape knowledge to guide optimization and inference

Cooperative Training and Latent Space Data Augmentation for Robust Medical Image Segmentation

Chen Chen[1], Kerstin Hammernik[1,2], Cheng Ouyang[1], Chen Qin[3], Wenjia Bai[4,5], Daniel Rueckert[1,2]

1. BioMedIA Group, Imperial College London
2. Klinikum rechts der Isar, Technical University of Munich
3. Institute for Digital Communications, University of Edinburgh
4. Data Science Institute, Imperial College London
5. Department of Brain Sciences, Imperial College London

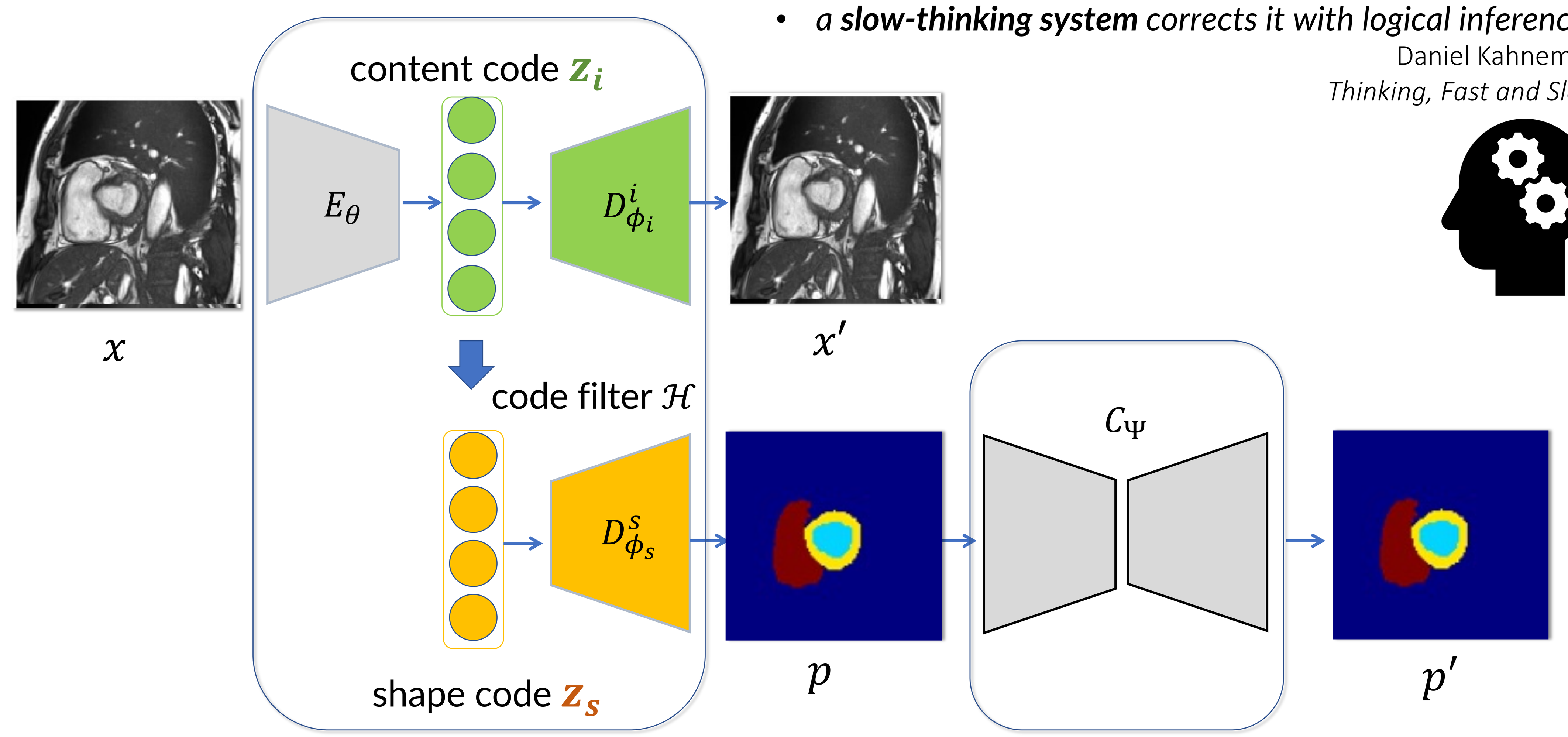
GitHub

Dual-Thinking Framework

Two networks cooperates both at training and testing for improved performance

When facing with unfamiliar scenarios, human have
- *a **fast-thinking system** makes intuitive judgment*
- *a **slow-thinking system** corrects it with logical inference.*

Daniel Kahneman
Thinking, Fast and Slow

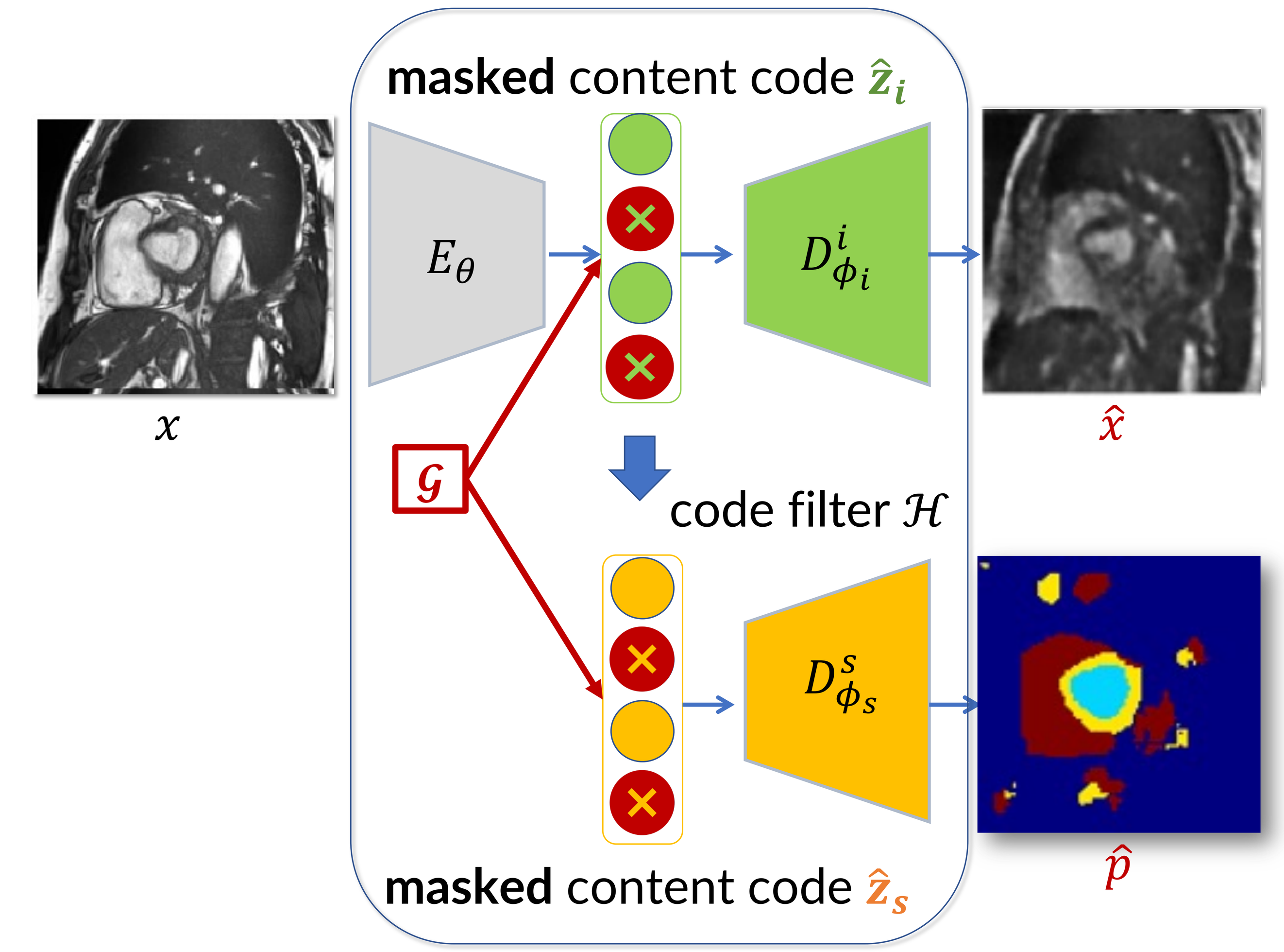


Fast-thinking network (FTN)
Image representation learning and segmentation

Slow-thinking network (STN)
Shape correction

Latent Space Masking-based Data Augmentation

Self-generates hard examples to train both FTN and STN

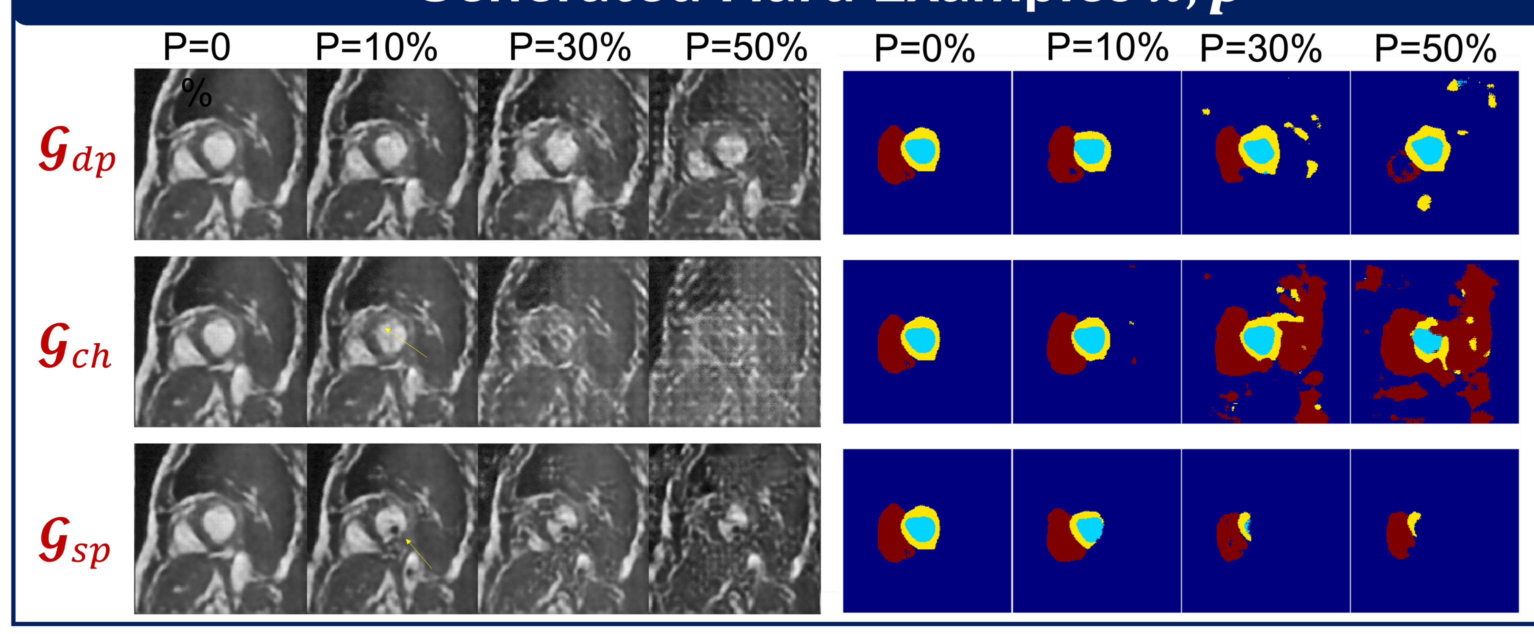


Three types of mask generators $\{\mathcal{G}\}$ to select:
- $\mathcal{G}_{dp}$: random channel-wise dropout
- $\mathcal{G}_{ch}$: targeted channel-wise soft masking*
- $\mathcal{G}_{sp}$: targeted spatial-wise soft masking*

*use gradient information to identify *salient* features to mask.

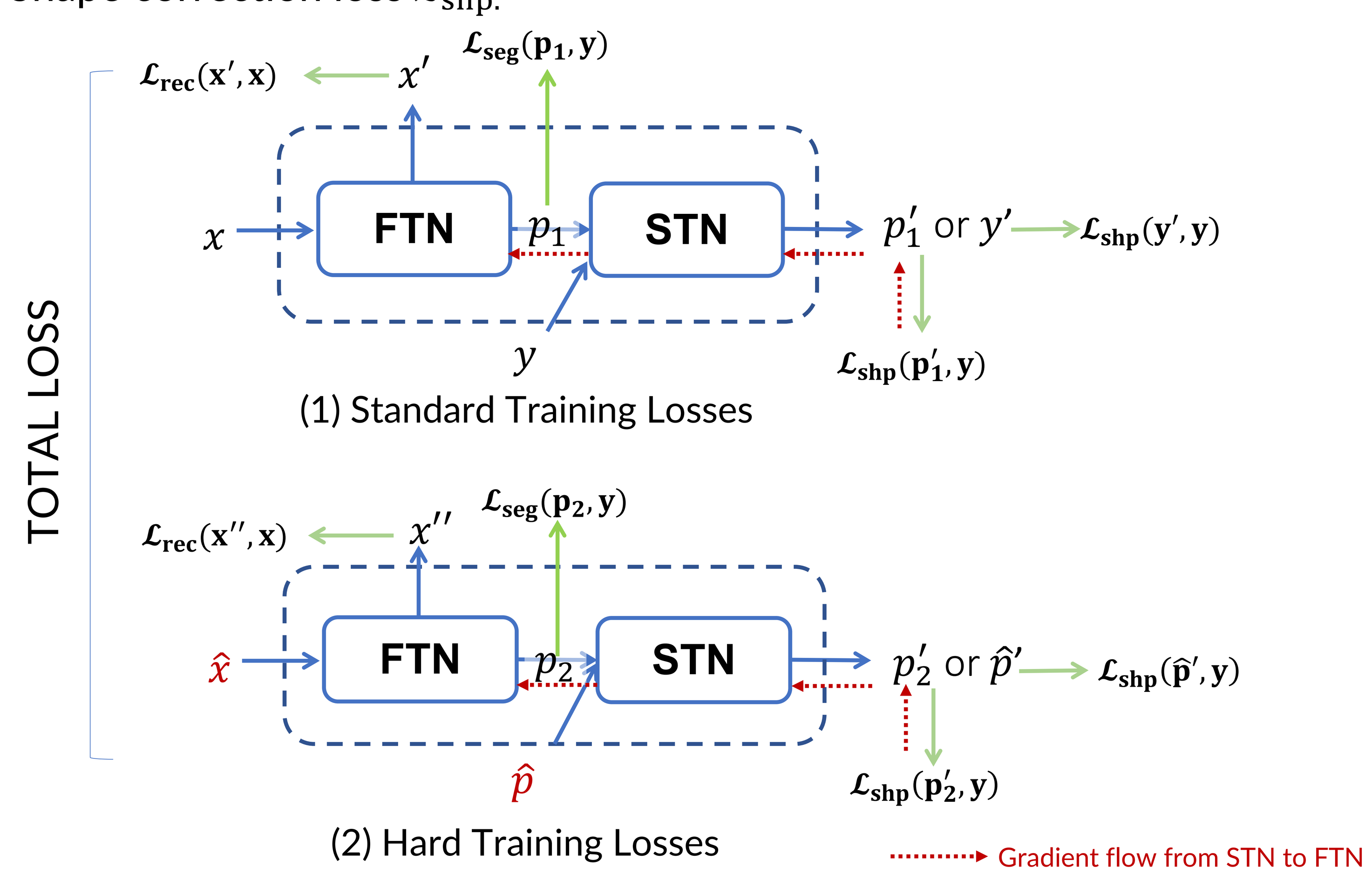
MORE

Generated Hard Examples $\hat{x}, \hat{p}$



Cooperative Training

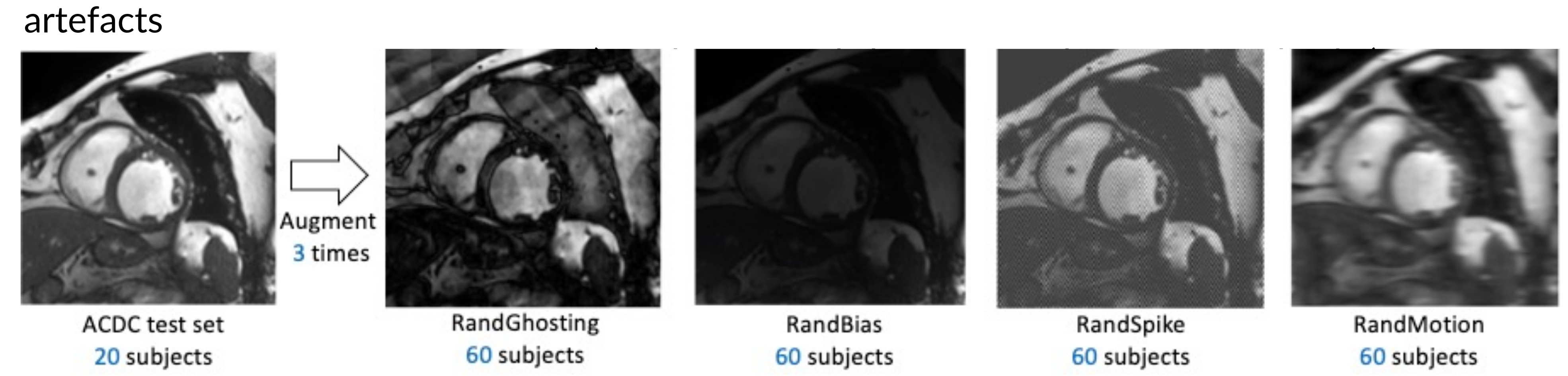
Multi-task loss computed on original and hard examples for joint optimization: a) Image reconstruction loss $\mathcal{L}_{\text{rec}}$; b) Segmentation loss $\mathcal{L}_{\text{seg}}$; c) Shape correction loss $\mathcal{L}_{\text{shp}}$.



Experiments and Results

More results are available online

- Task: Cardiac MR segmentation
- Training data: 10 ACDC[1] subjects (**single** site)
- Test data:
  - 1 Intra-domain test set: 20 ACDC subjects
  - 2 Unseen domain test sets :
    - M&Ms[2] for *cross-domain* test: 150 subjects collected from **multiple** sites
    - ACDC-C sets for *robustness* test: 240 corrupted ACDC test subjects with 4 types of common artefacts



[1] Olivier Bernard, TMI'18 [2] Víctor M. Campello, TMI'21

Segmentation results across different test sets

Method	ACDC	M&Ms	RandBias	RandGhosting	RandMotion	RandSpike	AVG (FTN)	AVG (FTN+STN)
Standard training	0.7681	0.3909	0.4889	0.6964	0.7494	0.4901	0.5970	0.6018
Rand MWM [1]	0.7515	0.3984	0.4914	0.6685	0.7336	0.5713	0.6024	0.6131
Rand Conv [2]	0.7604	0.4544	0.5538	0.6891	0.7493	0.4902	0.6162	0.6404
Adv Noise [3]	0.7678	0.3873	0.4903	0.6829	0.7543	0.6244	0.6178	0.6276
Adv Bias [4]	0.7573	0.6013	0.6709	0.6773	0.7348	0.3840	0.6376	0.6604
Proposed w. $\hat{x}$	0.7497	0.5154	0.5921	0.6921	0.7417	0.6633	0.6591	0.6709
Proposed w. $\hat{x}, \hat{p}$	0.7696	0.5454	0.6174	0.7073	0.7643	0.6226	0.6711	0.6901

- STN consistently improves the segmentation performance on **EVERY** test set
- Our latent space data augmentation (DA) outperforms SOTA image-space DA methods [1-4], achieving **TOP** average performance across all test sets

[1] Zongwei Zhou, MICCAI'19, MedIA'21 [2] Zhenlin Xu, ICLR'21 [3] Takeru Miyato, TPAMI'18 [4] Chen Chen, MICCAI'20

@cherise_go

MICCAI 2021 Imperial College London BioMedIA

Acknowledgement: This work was supported by the SmartHeart EPSRC Programme Grant (EP/P001009/1)

smartheart