Data augmentation for deep learning based accelerated MRI reconstruction with limited data

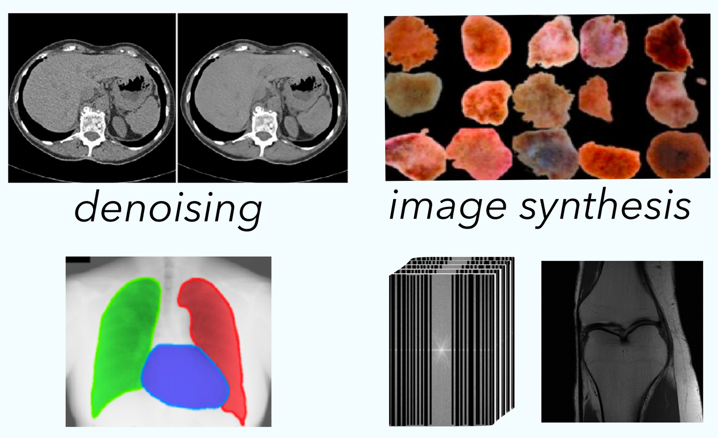
Zalan Fabian[1], Reinhard Heckel[2,3], Mahdi Soltanolkotabi[1]

USC[1], Technical University of Munich[2], Rice University[3]

USC Viterbi School of Engineering
Ming Hsieh Department of Electrical and Computer Engineering

Motivation

Limited data in medical imaging...



but DL models are **data-hungry!**

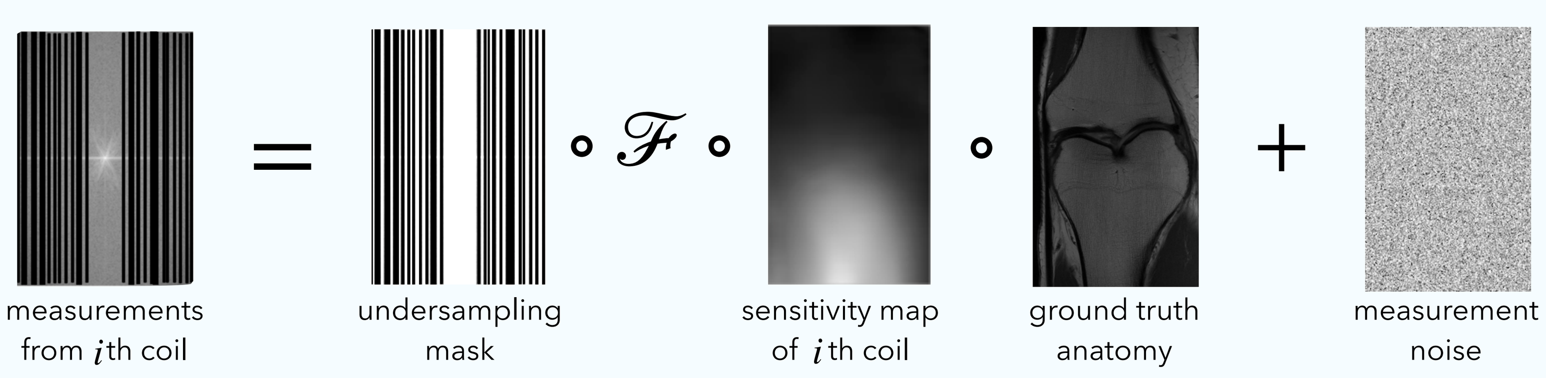
- Challenges of medical data collection
  - **Cost**: expensive instruments, time of experts
  - **Time**: long acquisition time (MRI: 60 mins / scan)
  - **Health**: ionizing radiation exposure (CT, PET)
  - **Data curation**: patient confidentiality, data compatibility

How do we train with limited data?

Background

- Accelerated multi-coil MRI acquisition

$$k_i = M\mathcal{F}S_i x^* + z_i \qquad i = 1,..,N$$

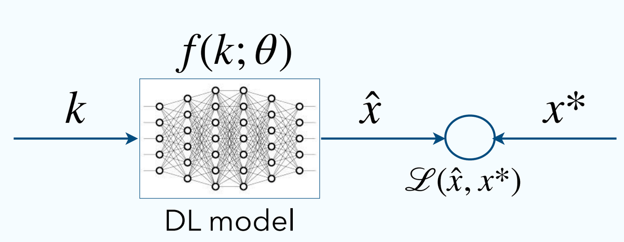


Compressed sensing reconstruction

$$\hat{x} = \arg\min_x \|\mathcal{A}(x) - k\|^2 + \mathcal{R}(x)$$

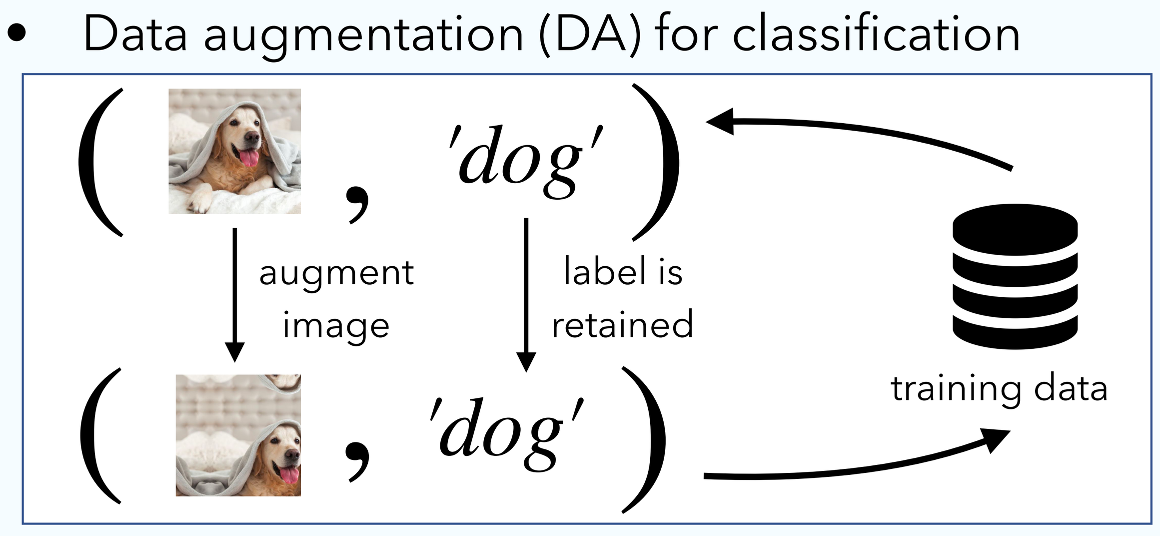
data consistency — prior knowledge

End-to-end DL reconstruction



Data augmentation for reconstruction

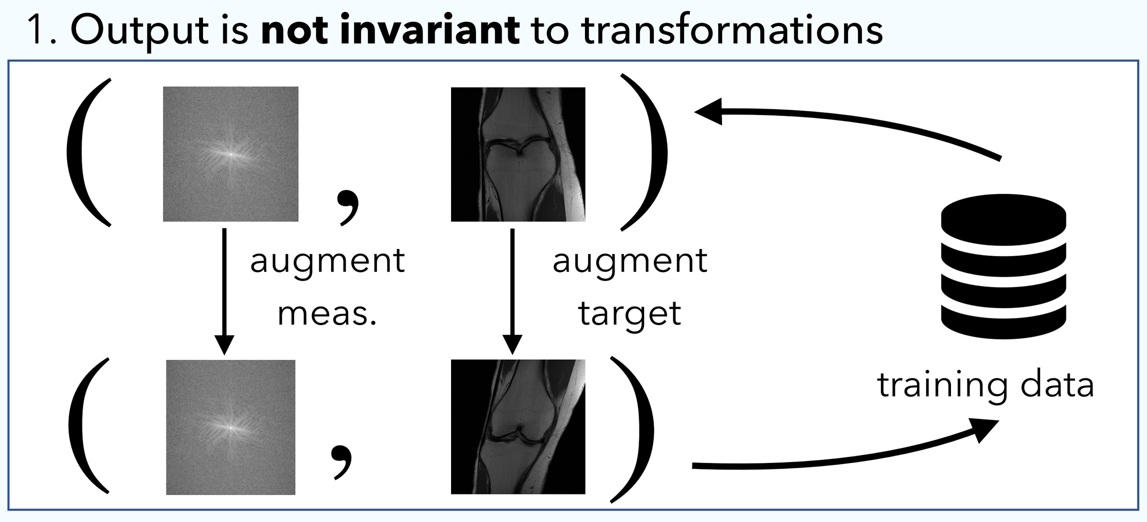
- Data augmentation (DA) for classification



Output is **invariant** to input transformations

- DA in reconstruction problems

1. Output is **not invariant** to transformations



DA pipeline has to generate **both** the augmented meas. **and** target

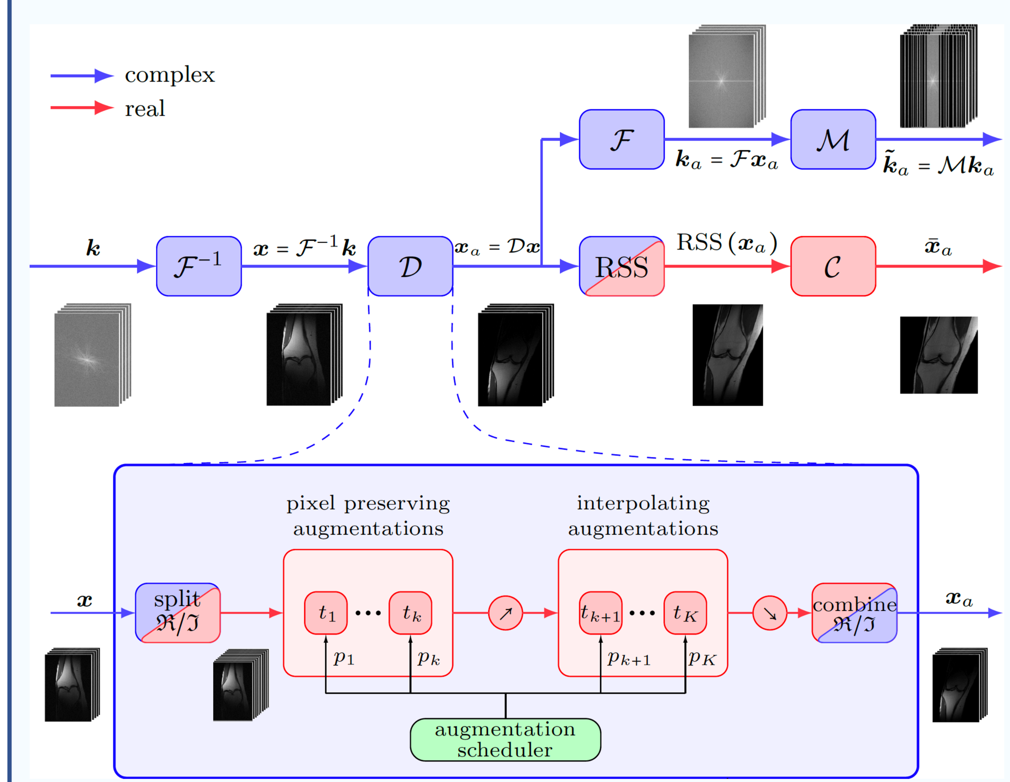
2. **Distribution shift** due to augmenting noise

$$x = x^* + n \longrightarrow x_{aug.} = \mathcal{D}x^* + \mathcal{D}n$$

augmented signal — augmented noise

Mismatch between train and test noise distribution results in poor generalization!

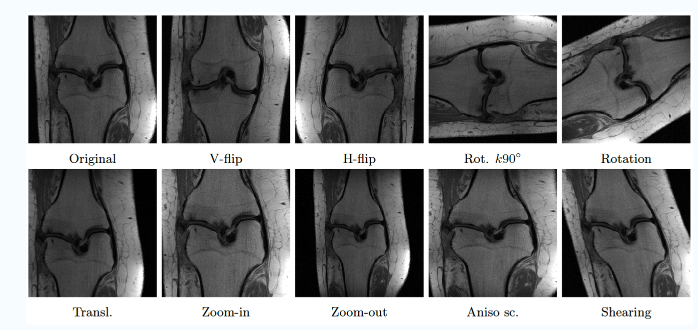
MRAugment pipeline



- Noise in MRI
  - i.i.d. Gaussian in real/imaginary parts
  - independent across coils

$$k_i = M\mathcal{F}S_i x^* + z_i$$

- Transformations



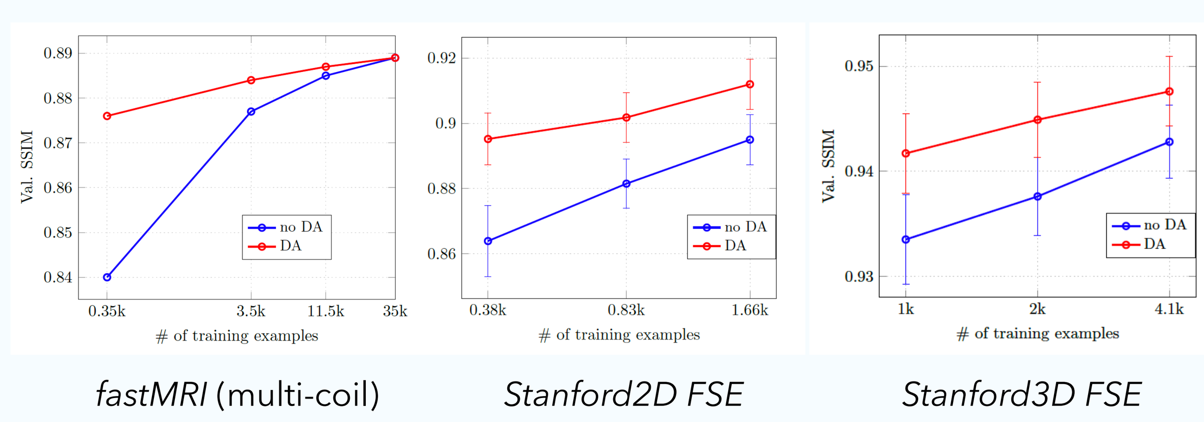
- applied coil-by-coil
- augmentation probability is scheduled over training iterations

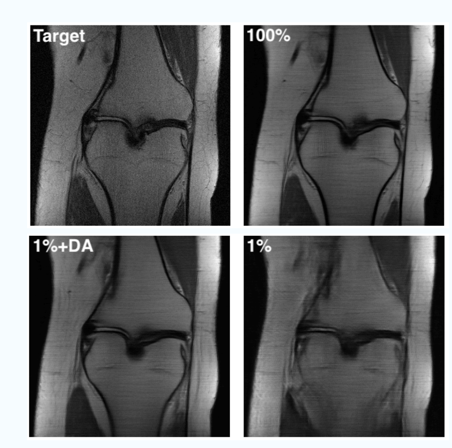
Experimental results

	# of slices	# of coils	Anatomies	Field strength
fastMRI knee	35k train + 7k val	15	knee	1.5 T, 3.0 T
Stanford3D FSE	5120	8	knee	3.0 T
Stanford2D FSE	2037	8-16	various	3.0 T

- **Goal**: study the effect of DA as a function of the size of training set
- **Model**: E2E VarNet
- **Acceleration**: 8x
- **Metric**: SSIM

1. Significant improvement in the low-data regime





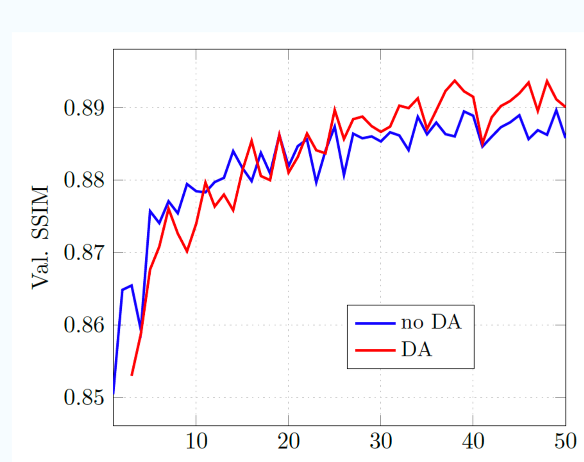
2. Improved model robustness

- Improved SSIM on **unseen scanners**

2% train	no DA	DA
$3T \to 3T$	0.8646	**0.9049**
$3T \to 1.5T$	0.8241	**0.8551**
$1.5T \to 3T$	0.8174	**0.8913**

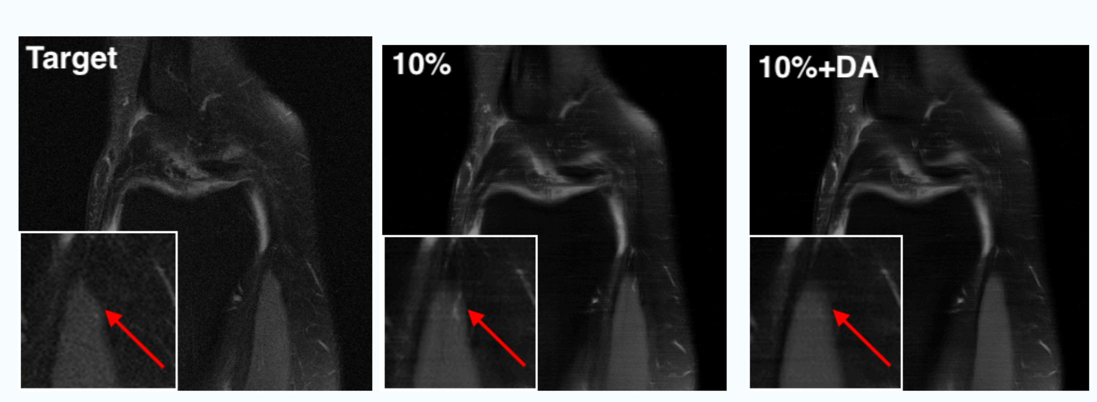
100% train	no DA	DA
$3T \to 3T$	0.9177	**0.9185**
$3T \to 1.5T$	0.8686	**0.8690**
$1.5T \to 3T$	0.9043	**0.9062**

- Improved SSIM on **unseen anatomies**



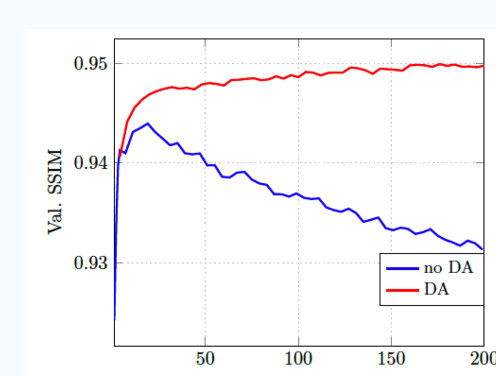
- We train on the full *fastMRI knee* dataset with and without DA
- We evaluate the models on *fastMRI brain* data

- Avoids **hallucinations**



Feature is not present on ground truth but appears on low-data reconstruction

Hallucinated feature is avoided with DA



DA prevents the model from overfitting to the training dataset

More details

Paper

ICML International Conference On Machine Learning

Code