Multi-task accelerated MR reconstruction schemes for jointly training multiple contrasts

Stanford Children's Health

Caltech

Victoria Liu, Kanghyun Ryu, Cagan Alkan, John Pauly, Shreyas Vasanawala

California Institute of Technology, Stanford Medical School

1. Objectives

There is inherent variability in MR datasets. The current paradigm is to train separate models for each dataset. However, this demanding process cannot exploit information that may be shared amongst datasets. In response, we propose multi-task learning (MTL) schemes to jointly reconstruct multiple datasets. As a proof of concept, we jointly reconstruct non-routinely acquired and routinely acquired knee contrasts. We will:

- Leverage abundant dataset for information
- Find a suitable MTL scheme for our datasets

2. Background

UNN solves an inverse problem for image $\mathbf{m}$:

$$\hat{\mathbf{m}} = \arg\min_{\mathbf{m}} \|\mathbf{E}\mathbf{m} - \mathbf{k}\|_2^2 + R(\mathbf{m}) \quad (1)$$

MTL assumes that a network trained on one task has extracted useful features to learn a related task. Possible MTL schemes are summarized in Table 1.

Table 1: MTL schemes illustrated in Figure 2 and used to generate our result MRIs are bolded in orange.

MTL Schemes	Examples
hard sharing	**shared blocks at beginning** shared blocks in middle shared encoder & split decoder
soft sharing	attentional network cross-stitch network
loss function	**naive weighting** **DWA** uncertainty weighting pareto weighting
optimization	**shuffled dataloader** stratified dataloader gradient accumulation

We use two STL baselines in Figure 1:

- Joint training of scarce and abundant datasets
- Individual training of scarce and abundant datasets

3. Single-task learning (STL)

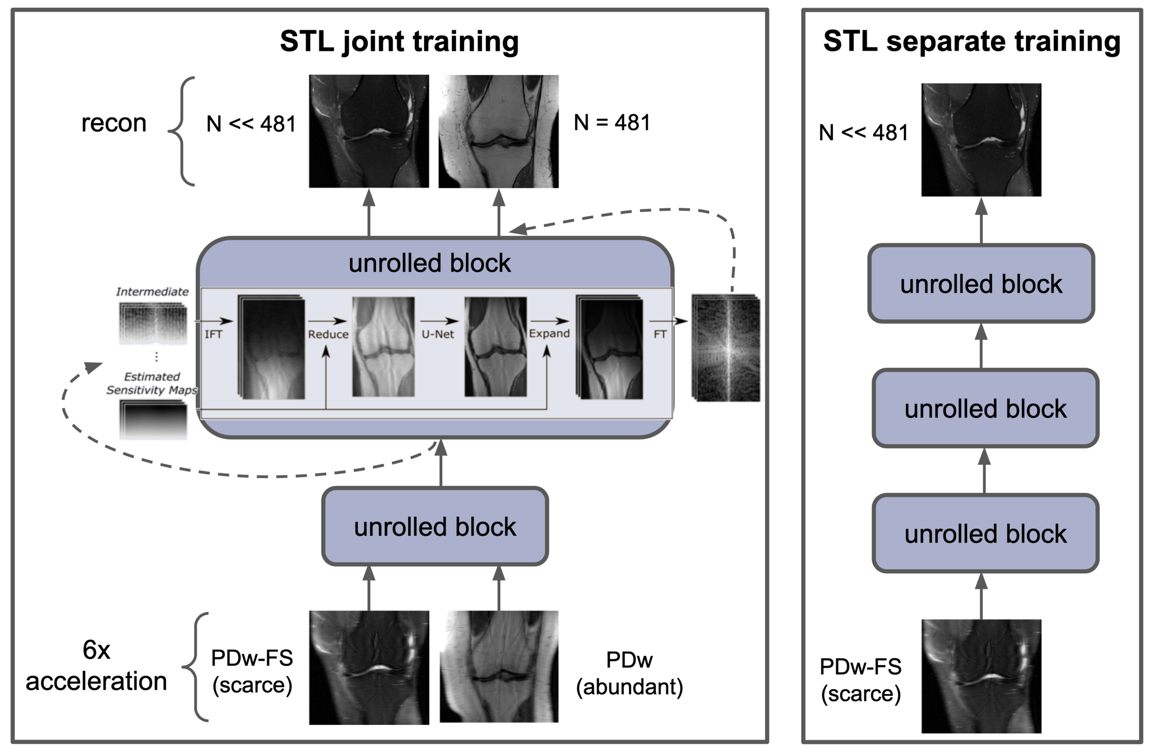

Figure 1: Abundant task = knee coronal PDw; all 481 slices are used. Scarce task = knee coronal PDw-FS ; a percentage of the total 492 slices is used.

4. Multi-task learning (MTL)

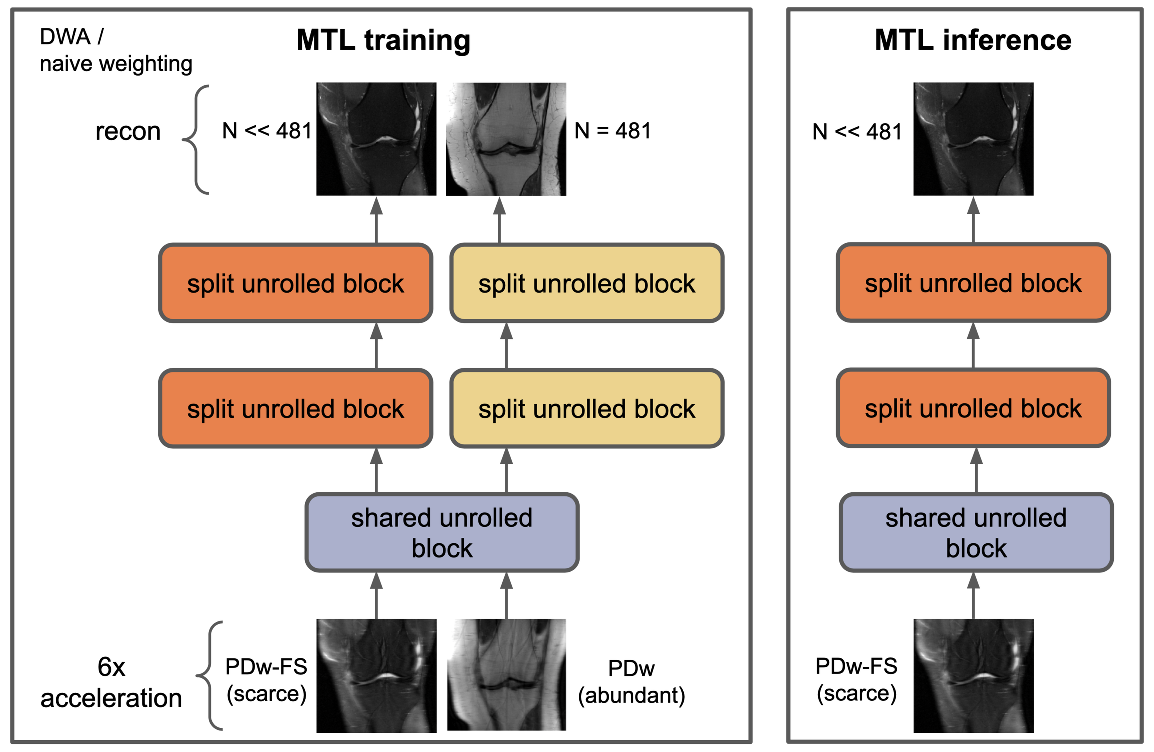

Figure 2: Hard sharing MTL scheme with shared blocks at the beginning. Split blocks are task-specific. Loss weighting is done using DWA or naively.

5. MTL reconstructions reduce errors

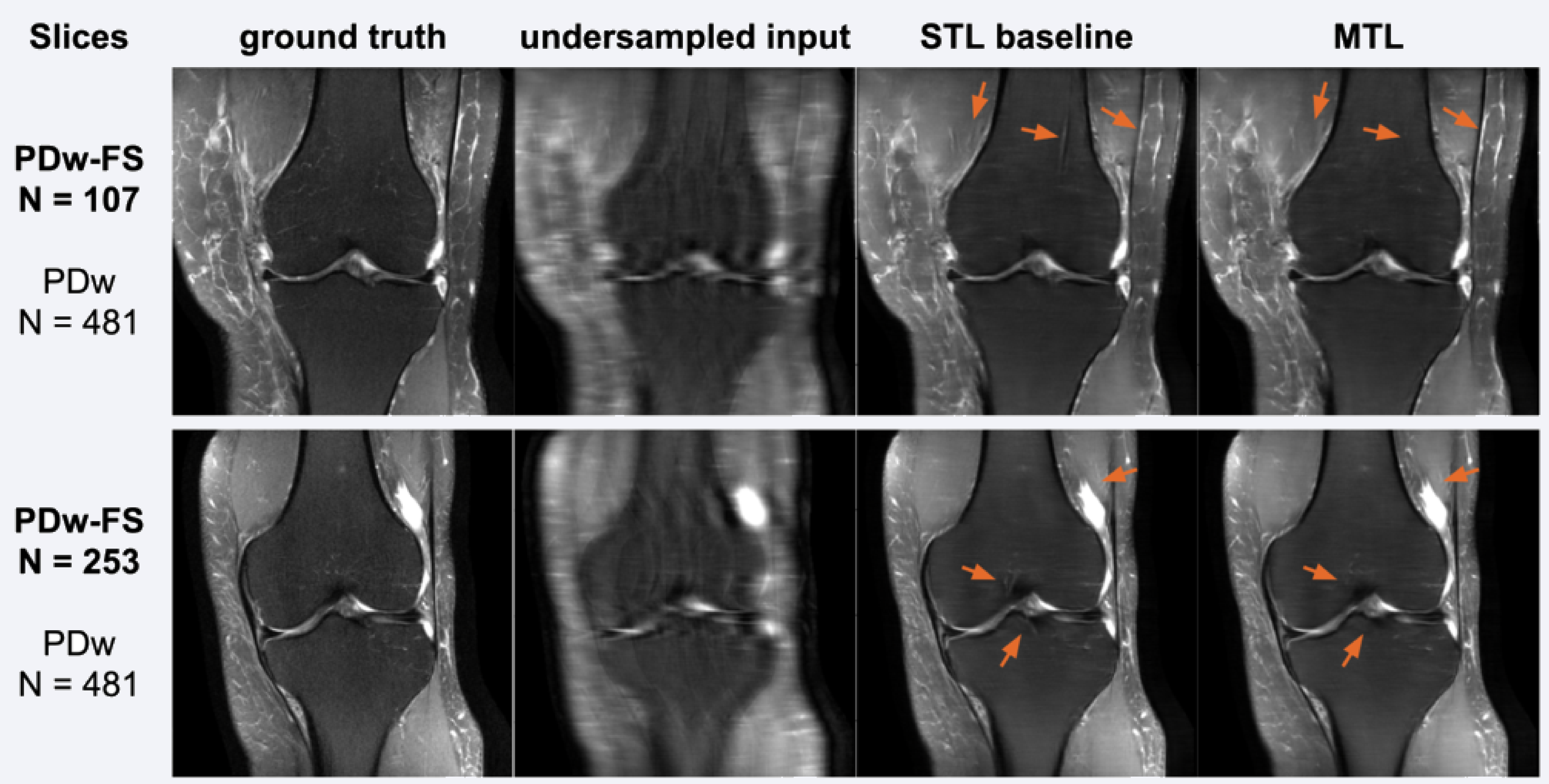

Figure 3: When PDw-FS slice count is 25 % or 50 % of PDw's, MTL reduces errors from STL.

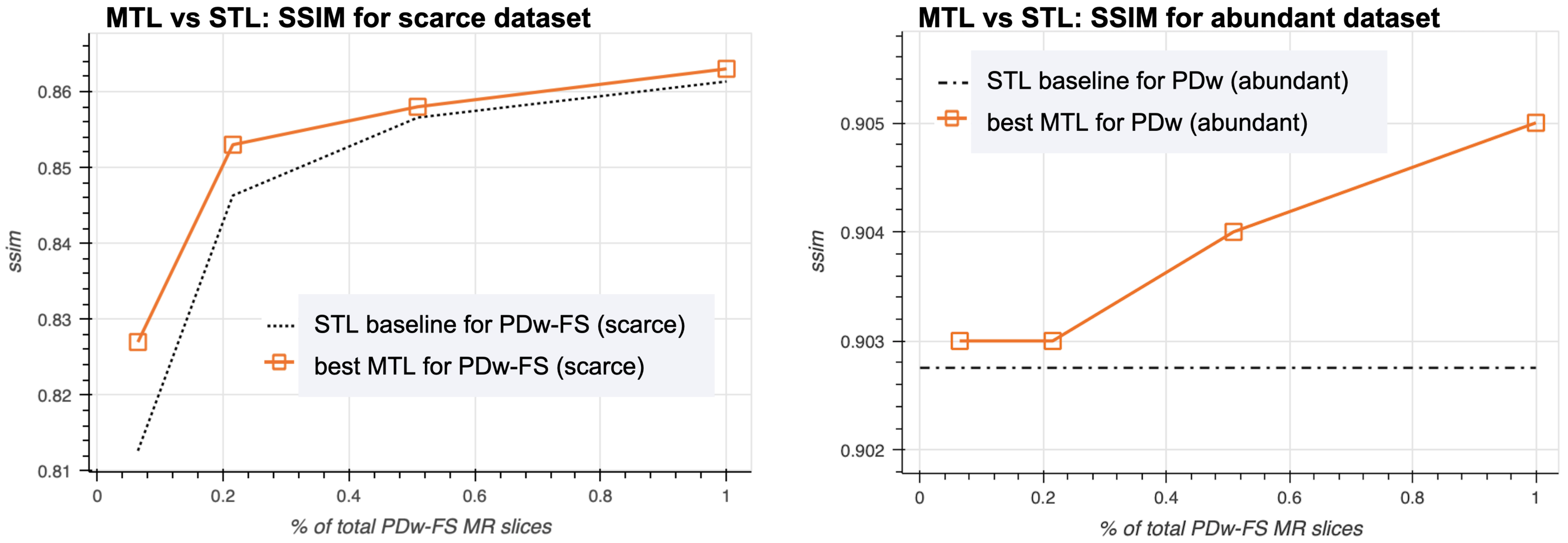

Figure 4: PDw-FS and PDw MTL recons both have better SSIMs than STL recons.

6. Conclusions

MTL improves SSIM & PSNR and qualitatively reduces STL errors. For knee coronal PDw vs PDw-FS, effective MTL schemes:

- Share blocks at beginning
- Use DWA or naive weighting

Our framework successfully introduces inductive biases in the network by enforcing the sharing of useful information between tasks.

7. Challenges

Transfer learning outperforms current MTL models. In addition, negative transfer can cause MTL to perform worse than STL:

Table 2: PSNR for PDw-FS N = 107.

Network	PSNR
STL: baseline	33.11
MTL: DWA, shared blocks at beginning	**33.40**
MTL: Uncertainty, shared blocks in middle	**32.71**

8. Future work

- Explore other datasets, including differing anatomies.
- Expand the number of datasets

9. References

[1] Michael Crawshaw. Multi-task learning with deep neural networks: A survey, 2020.

Acknowledgements

This project is supported by NIH R01 EB009690 and NIH R01 EB026136, as well as GE Healthcare and the Marcella Bonsall Fellowship.