Reverse Engineering Breast MRIs: Predicting Acquisition Parameters Directly from Images

Nicholas Konz[1] Maciej Mazurowski[1,2,3,4]

[1]Department of Electrical and Computer Engineering, [2]Department of Radiology, [3]Department of Computer Science, [4]Department of Biostatistics & Bioinformatics, Duke University, Durham, North Carolina, USA

Code: (github.com/mazurowski-lab/MRI-IAP-prediction) **My GitHub:** **Twitter:** @nick_konz

Link to paper: https://arxiv.org/abs/2303.04911

Introduction

- The image acquisition parameters (**IAPs**, see Table 1) used to take MRIs are central to defining their appearance.
- Deep neural nets trained on MRIs taken with certain IAPs may not generalize well to images from other IAPs: a **domain shift problem**.
- Can we train a neural network to **predict the IAPs that generated an image, using only the image?**
- This would allow us to predict an image's domain to determine if it's applicable to some downstream model, and could also be used for domain adaptation/image harmonization, etc.

Contributions

1. We introduce a **neural network model for predicting many categorical and continuous IAPs of an MR image in one forward pass**, trained via multi-task learning.
2. We show that our model **predicts many complex IAPs of MRI scans of new patients to high accuracy**, over a large test set of MR slice images. We predict six out of ten categorical IAPs to $> 97\%$ top-1 accuracy on the test set, and all but two with $> 95\%$ top-2 accuracy.
3. We show that our method achieves **good accuracy** ($> 84\%$ top-1 accuracy, $> 95\%$ top-2) **on IAPs that are more challenging to predict, such as contrast agent type.**
4. We demonstrate a realistic application of our model: using it to sort new unlabeled data into different domains to determine which models to apply to the data for a downstream task.

IAP Prediction Network

- Task: simultaneously predict K categorical IAPs and M continuous IAPs of an MR image.
  - K **classification** tasks and M **regression** tasks.
- Model: ResNet-18 encoder with final fully-connected layer modified to predict all $K + M$ IAPs.

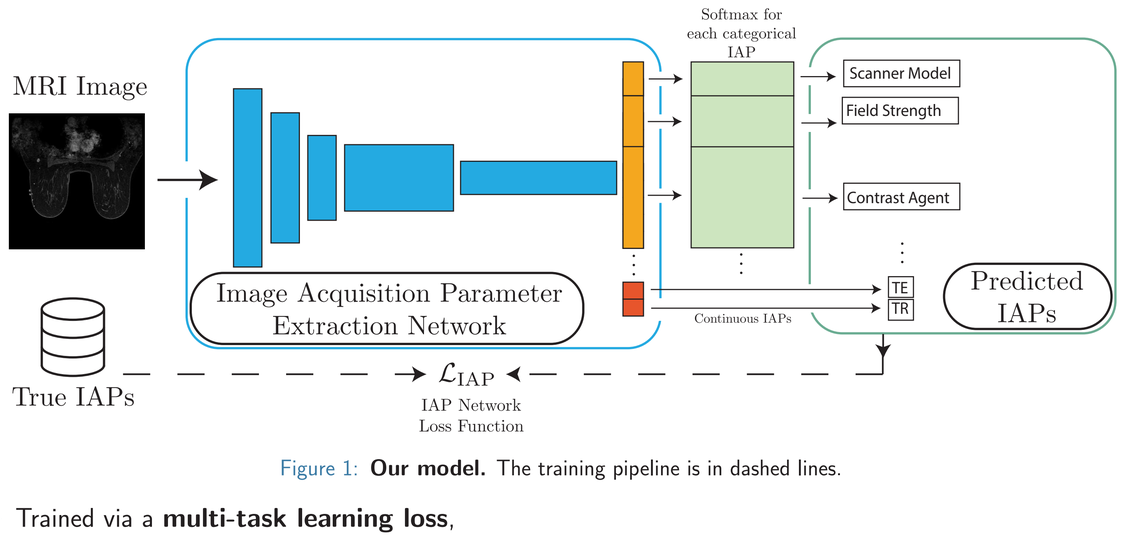


Figure 1: **Our model.** The training pipeline is in dashed lines.

- Trained via a **multi-task learning loss**,

$$\mathcal{L}_{\text{IAP}} = \lambda \sum_{k=1}^{K} \mathcal{L}_{\text{CE}}(\hat{y}_k, y_k) + \eta \sum_{m=1}^{M} \mathcal{L}_{\text{MSE}}(\hat{y}_m, y_m), \tag{1}$$

with:

- $\mathcal{L}_{\text{CE}}$: cross-entropy loss of predicted class $\hat{y}_k$ and true class y_k for the k^{th} categorical IAP.
- $\mathcal{L}_{\text{MSE}}$: MSE loss of predicted value $\hat{y}_m$ and true value y_m for the m^{th} continuous IAP.
- Hyperparameters: $\lambda = \eta = 1$

Dataset

- We use the Duke Breast Cancer (DBC) MRI dataset [1]: contains DCE-MRIs of 922 biopsy-confirmed breast cancer patients.
- Each scan has values for 12 categorical and continuous IAPs (Table 1).
- We use a subset of $14,000$ randomly sampled 2D slices from 3D fat-saturated scan volumes, split into train/validation/test sets as $9,952/2,064/1,984$, **with no patient appearing in multiple splits**.

Experimental Settings

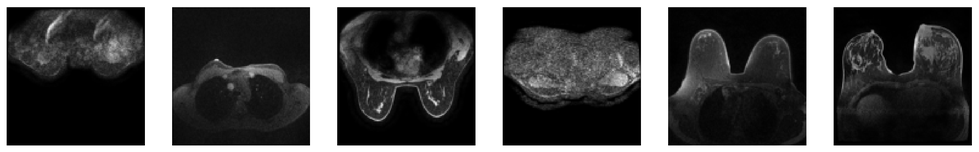
- Trained with a batch size of 512 for 100 epochs, using Adam with a learning rate of 0.001 and weight decay strength of 0.0001, on a 48 GB NVIDIA A6000.
- Images resized to 224×224 and normalized to $[0, 255]$. Final model chosen based on validation set performance.

Results: Our Model's IAP Prediction Performance on the Test Set

Table 1: **Quantitative Summary: IAP Prediction Performance on the Test Set.** $*$ denotes prediction MSEs for models with categorical IAPs trained instead as continuous.

	MRI acquisition parameter (IAP)	No. categories	Examples	Top-1 pred. acc. (%)	Top-2 pred. acc. (%)	Pred. MSE
1	Scanner Manufacturer	2	GE, Siemens	99.74	N/A	N/A
2	Scanner Model	8	Avanto, Signa HDx	97.78	99.29	N/A
3	Scan Options	9	PFP/FS, PFP/SFS	99.40	99.60	N/A
4	Field Strength	5	1.5 T, 3 T	98.19	99.70	N/A
5	Patient Position	2	FFP, HFP	97.73	N/A	N/A
6	Contrast Agent Type	6	Gadavist, MultiHance	84.73	95.46	N/A
7	Acquisition Matrix	10	448×448, 384×360	91.53	99.14	N/A
8	Slice Thickness	21	1.3 mm, 2 mm	76.66	87.05	0.157 mm*
9	Flip Angle	4	10°, 12°	99.65	99.75	0.073°*
10	FOV Computed	27	320 cm, 360 cm	51.21	69.30	164 cm*
11	Repetition Time (TR)	N/A	4.27 ms, 5.34 ms	N/A	N/A	0.0305 ms
12	Echo Time (TE)	N/A	2.4 ms, 1.5 ms	N/A	N/A	0.0116 ms

Figure 2: **Example Predictions of Acquisition Parameters for MRIs in the Test Set.** The symbols "✓" and "×" indicate correct and incorrect predictions, respectively (TE and TR predictions are treated as "correct" if the relative error is $< 2\%$).



Predicted IAPs					
Manufacturer:	GE ✓	SIEMENS ✓	GE ✓	GE ✓	SIEMENS ✓	SIEMENS ✓
Model:	Signa HDxt ✓	Avanto ✓	Signa HDxt ✓	SIGNA HDx ✓	Avanto ✓	Skyra ✓
Scan Options:	1 ✓	5 ✓	1 ✓	2 ✓	5 ✓	6 ✓
Patient Position:	FFP ✓	FFP ✓	FFP ✓	FFP ✓	FFP ✓	FFP ✓
Field Strength:	3 T ✓	1.5 T ✓	3 T ✓	1.5 T ✓	1.5 T ✓	3 T ✓
Contrast Agent:	MULTIHANCE ✓	MULTIHANCE ✓	MAGNEVIST ✓	MAGNEVIST ✓	MAGNEVIST ✓	MULTIHANCE ✓
Acquisition Matrix:	384 X 384 ✓	448 X 448 ✓	350 X 350 ✓	340 X 340 ✓	448 X 448 ✓	448 X 381 ✓
Slice Thickness:	1.2 mm ✓	2 mm ✓	2 mm ✓	2 mm ✓	1.1 mm ✓	2 mm ✓
Flip Angle:	10 deg ✓	10 deg ✓	10 deg ✓	10 deg ✓	12 deg ✓	10 deg ✓
FOV Computed:	340 cm ✓	340 cm ✓	340 cm ✓	320 cm ✓	360 cm ✗	350 cm ✓
TE (Echo Time):	2.537 ms ✓	1.304 ms ✓	2.382 ms ✓	2.373 ms ✓	1.324 ms ✓	1.445 ms ✓
TR (Rep. Time):	5.963 ms ✓	4.415 ms ✗	5.686 ms ✓	5.171 ms ✗	4.197 ms ✓	3.722 ms ✓

Application: Sorting Unlabeled Data into Domains for Downstream Task Model Selection

- Consider **two cancer classification networks** trained on data from different scanner manufacturers: **GE** and **Siemens**.
- How can we tell which network to use on new breast MRIs from **unknown scanners**?
- We use our IAP prediction model on the new images to determine which scanner type they were taken with.
  - Doing so **greatly improves cancer classification performance on an unlabeled, shuffled test set of both GE and Siemens images**, rather than just guessing the scanner type (table below).

Table 2: **Using our IAP prediction model to sort unlabeled data for cancer classification model selection.** Values shown are **cancer classification accuracies** on the test set of GE and Siemens images, unless otherwise stated.

GE Model	**Siemens Model**	GE Model (on only GE images)	Siemens Model (on only Siemens images)	**Model chosen according to predicted IAPs**	Model chosen according to true IAPs
68.82%	56.50%	80.37%	71.75%	76.95%	77.43%

Future Work

- Use our model to guide image translation/harmonization networks, as an "IAP discriminator".
- Use our model to probe the relationship between IAPs, image quality, and domain.
- Extend our model to full 3D MR volumes using 2.5D or 3D CNNs.

References

[1] Ashirbani Saha, Michael R Harowicz, Lars J Grimm, Connie E Kim, Sujata V Ghate, Ruth Walsh, and Maciej A Mazurowski. A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 dce-mri features. *British journal of cancer*, 119(4):508–516, 2018.