Benchmarking Robustness of Multilabel Segmentation Model

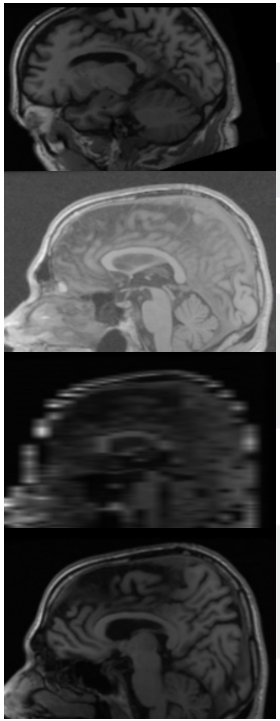
Yunni Qu
Department of Computer Science, University of Toronto

Aug 02, 2022

Agenda

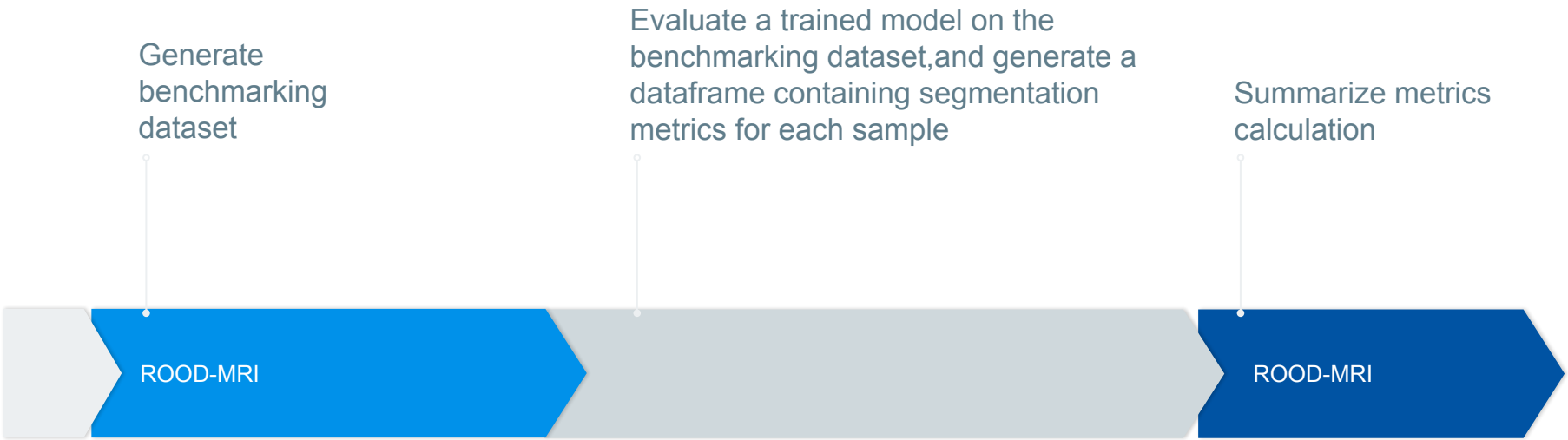
1. Background
2. Methods
3. Results

Background - Motivations



- Impressive performance on segmentation and classification tasks of deep learning model
- Difference in images from different sites
- Testing robustness of model

Background - Existing Platform

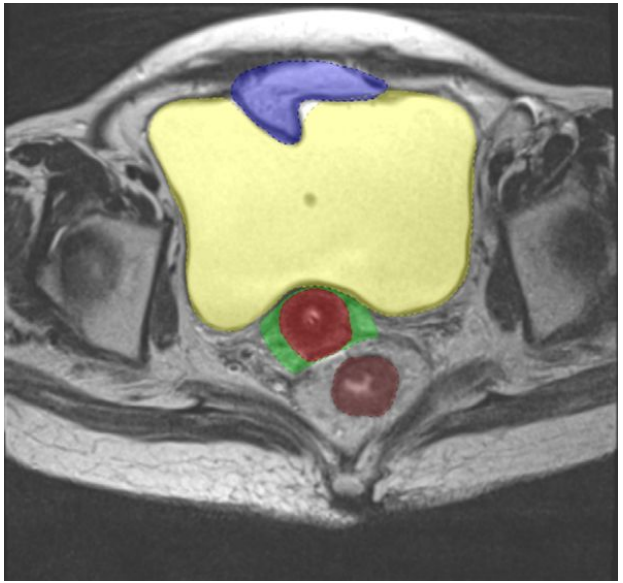


Background - Data

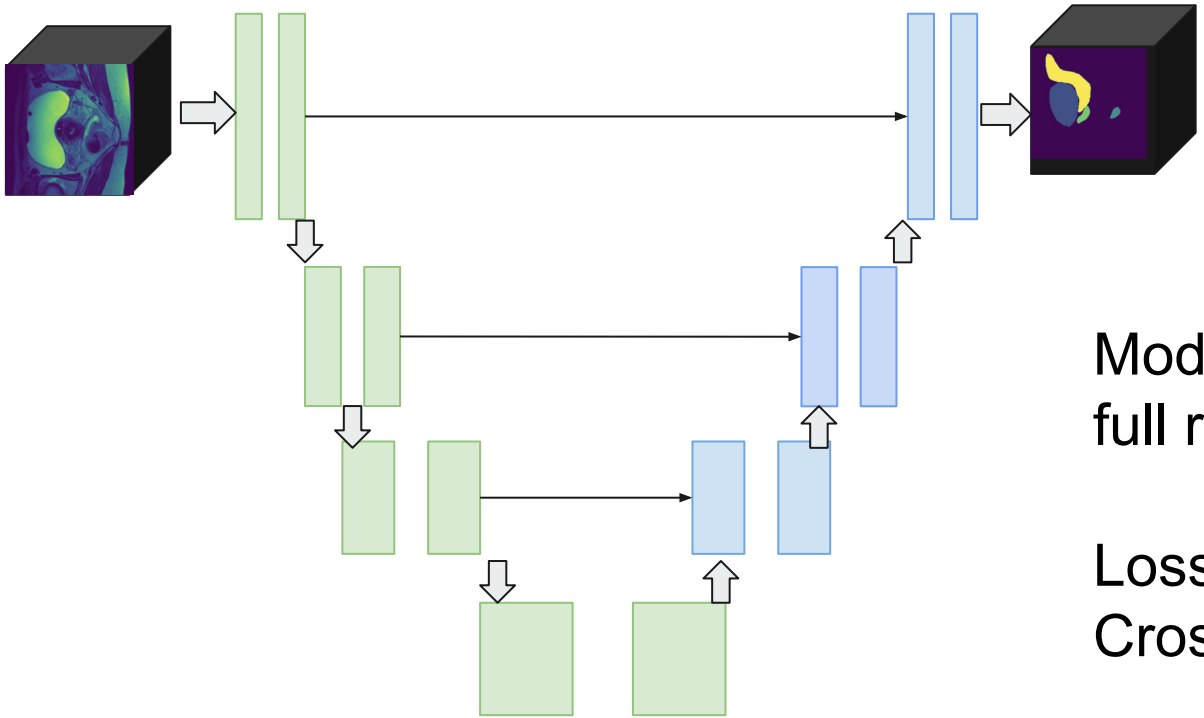
The data set contains 131 T2-weighted MRI scans with clinically used contours

Train-Test Split: 119 for training and 11 for testing

Regions of Interests Bladder, Rectum, Sigmoid, and Small Bowel



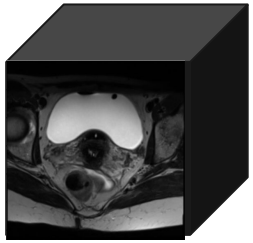
Background - Model



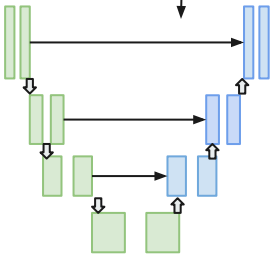
Model: nnU-net with 3D full resolution U-Net

Loss function: Dice + Cross Entropy loss

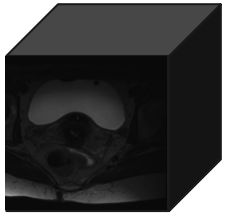
Methods - Overview



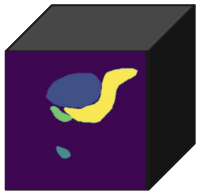
Training set

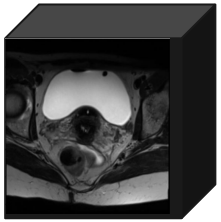


Segmentation Model



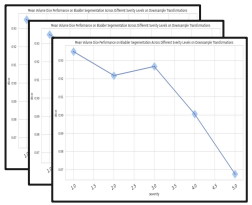
Loss function





Testing set

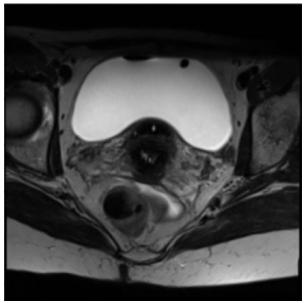
Noise transforms (Severity 1-5)



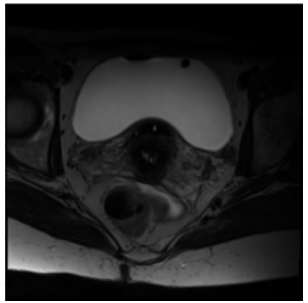
DSC

Robustness Metrics

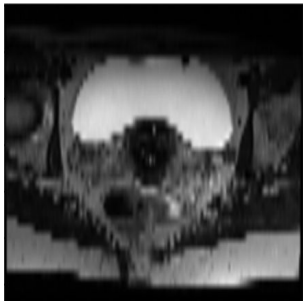
Methods - Transformations



Original

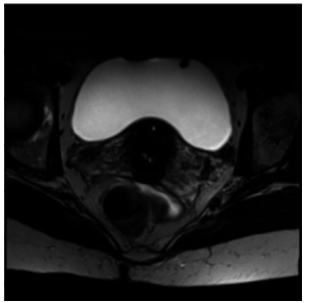


Biasfield

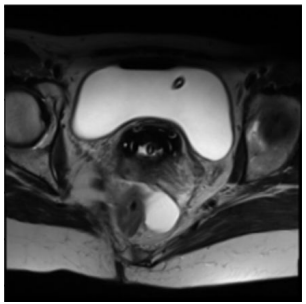


Down Sample

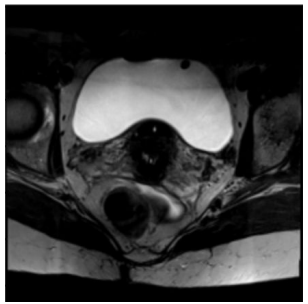
Contrast Compression



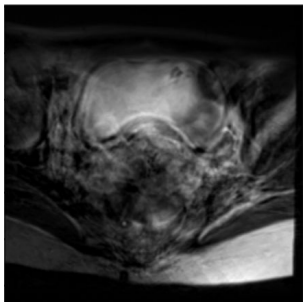
Contrast Expansion



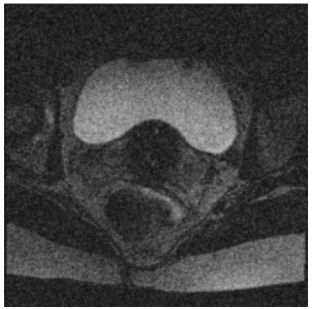
Elastic Deformation



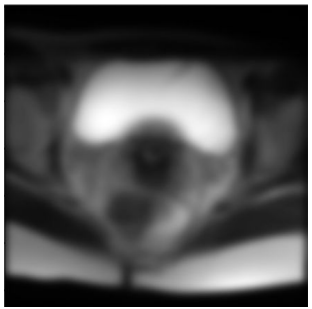
Ghosting



Random Motion

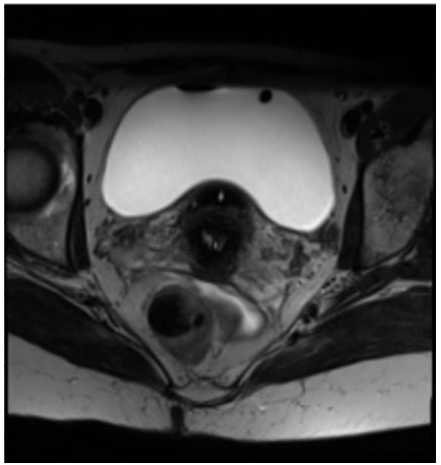


Rician Noise



Smoothing

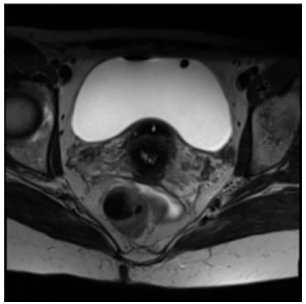
Methods - Signal to Noise Ratio Transformations



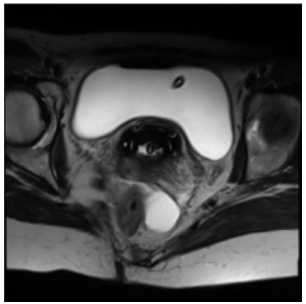
Original

Rician Noise

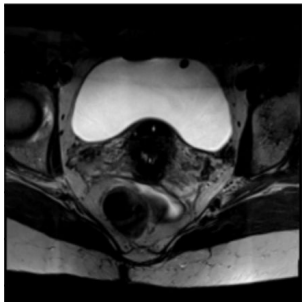
Methods - Motion Transformations



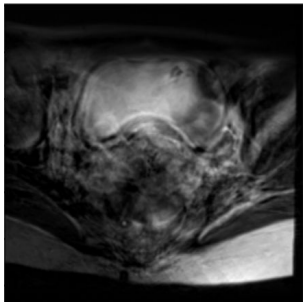
Original



Elastic Deformation

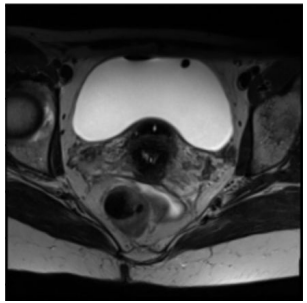


Ghosting

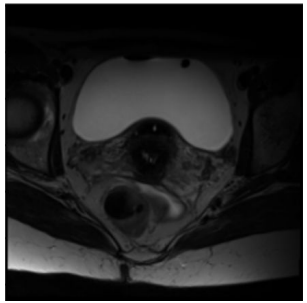


Random Motion

Methods - Contrast and Intensity Transformations

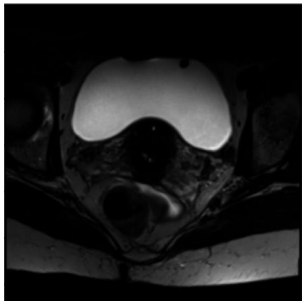


Original



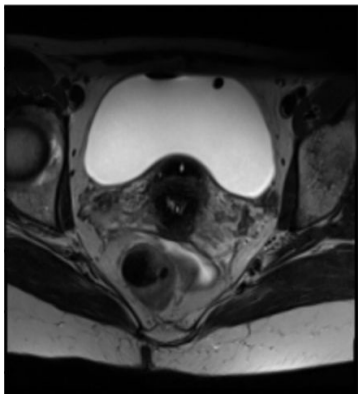
Biasfield

Contrast Compression

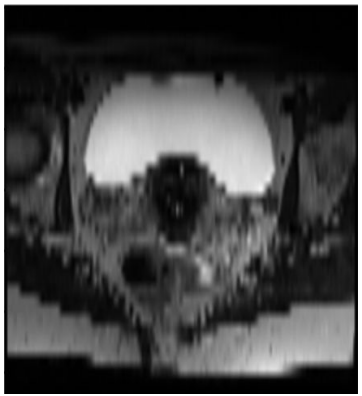


Contrast Expansion

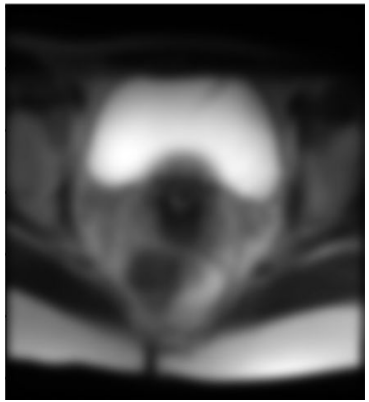
Methods - Image Resolution and Blurring Transformations



Original

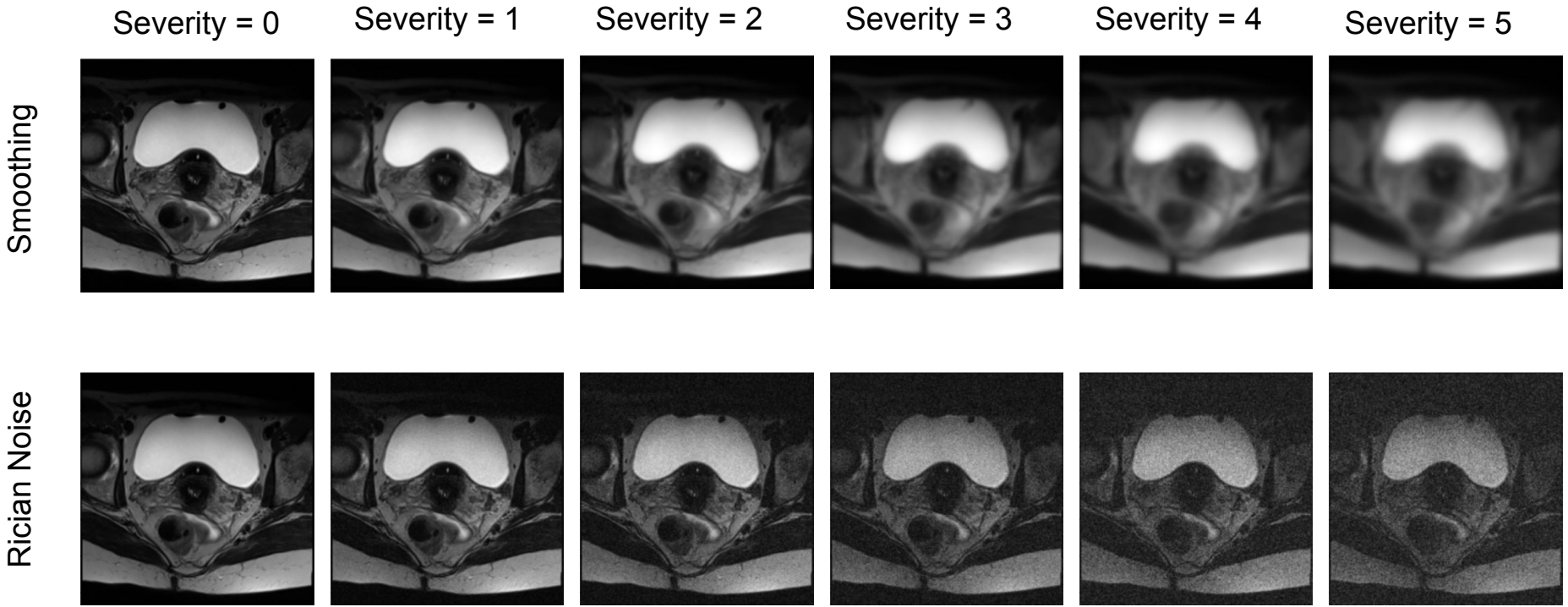


Down Sample



Smoothing

Methods - Severity Levels



BHK LAB

Results - Segmentation Masks of Smoothing Transformation

Original

Severity = 1

Severity = 2

Severity = 3

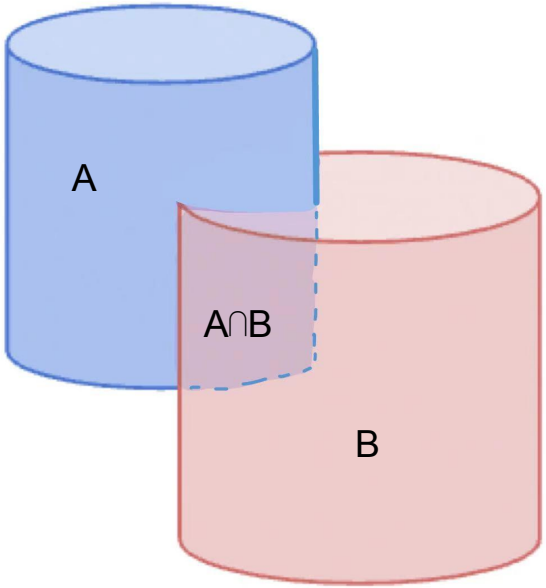
Severity = 4

Severity = 5

Bladder(blue), Rectum(teal), Sigmoid(green), Small Bowel(yellow)

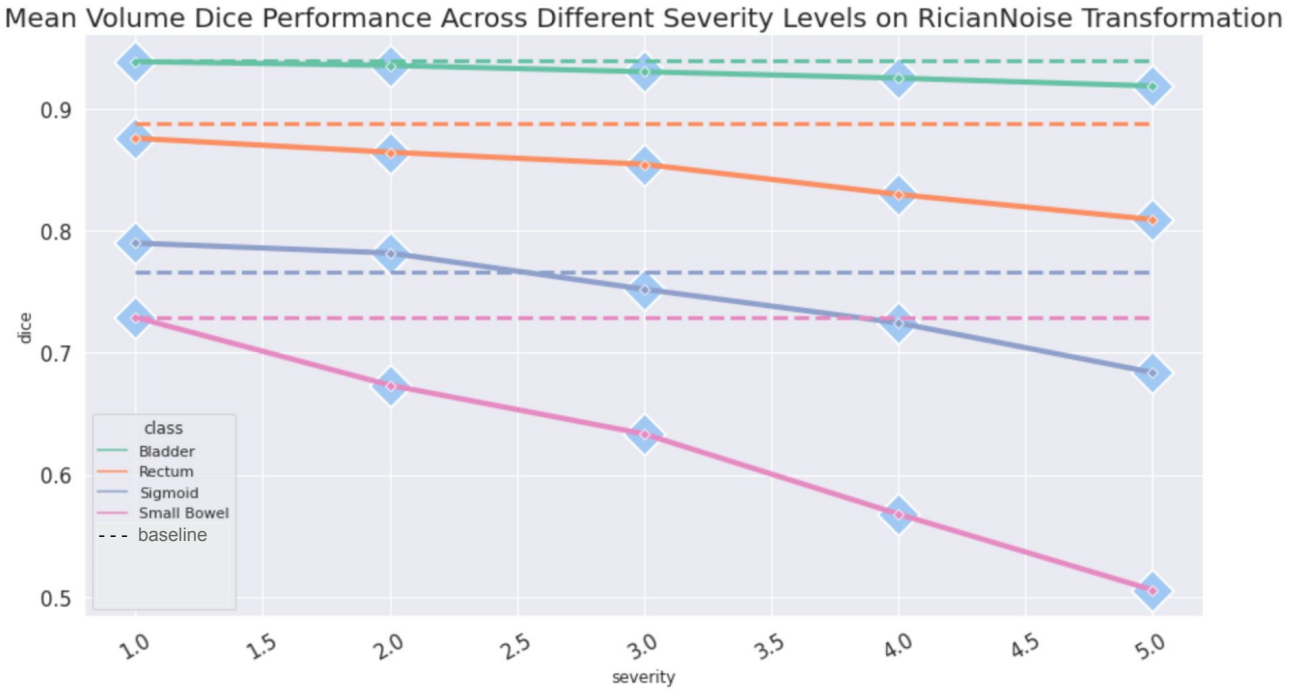
BHK LAB

Results - Volumetric Dice Similarity Coefficient

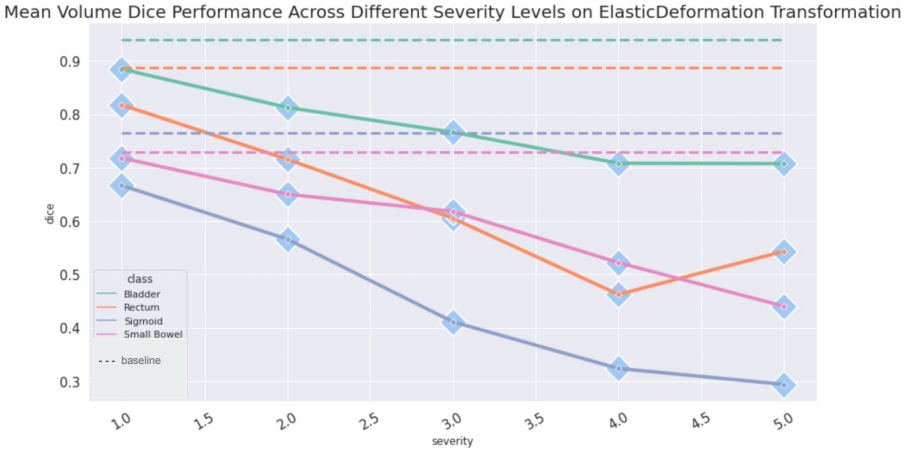


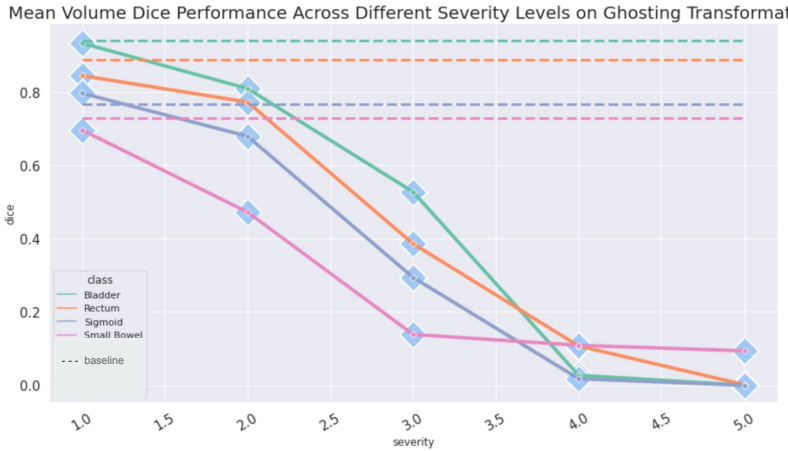
$$vDSC = 2(A \cap B)/(A+B)$$

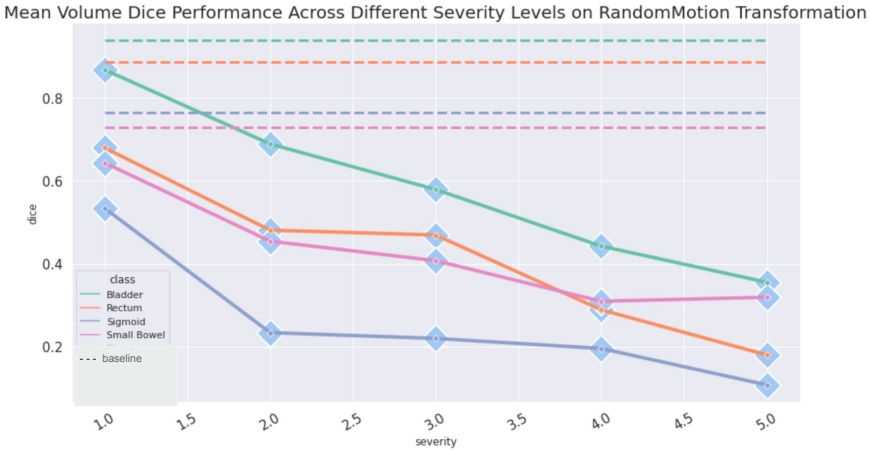
Results - Signal to Noise Ratio Transformations



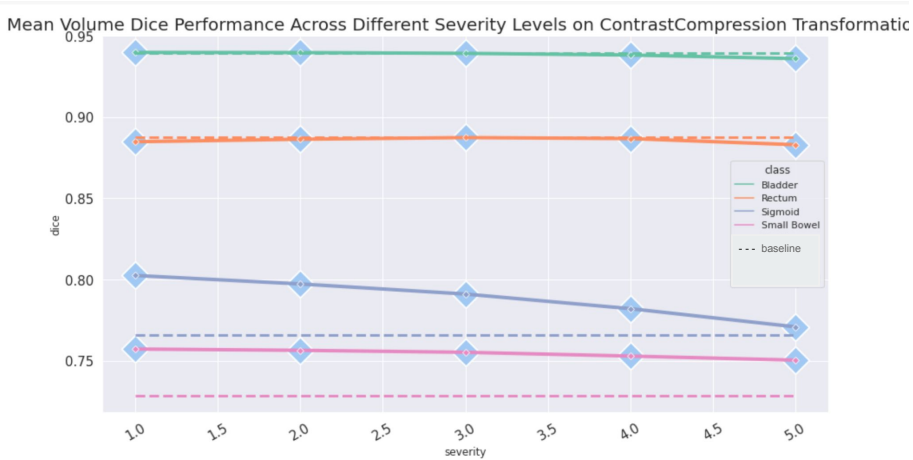
Results - Motion Transformations

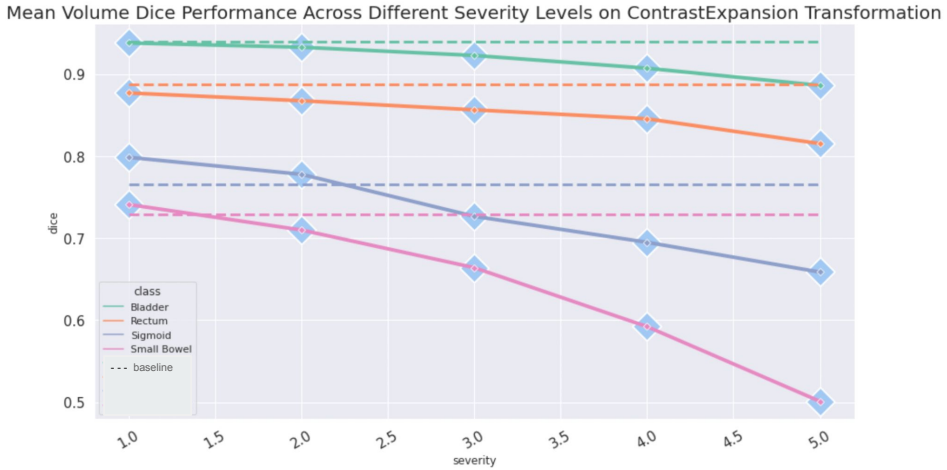


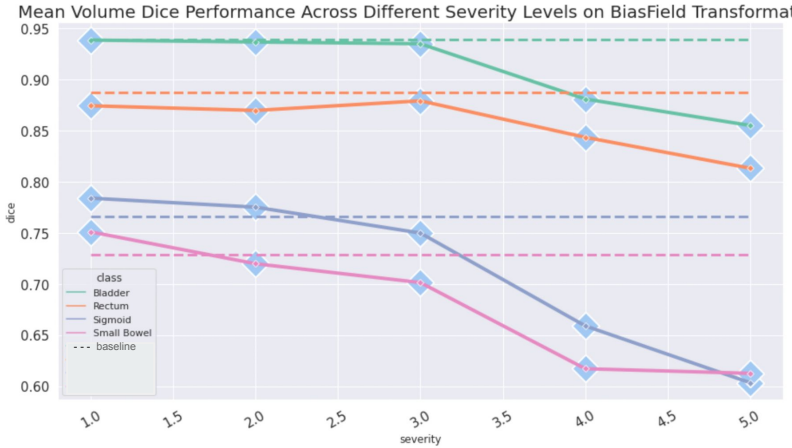




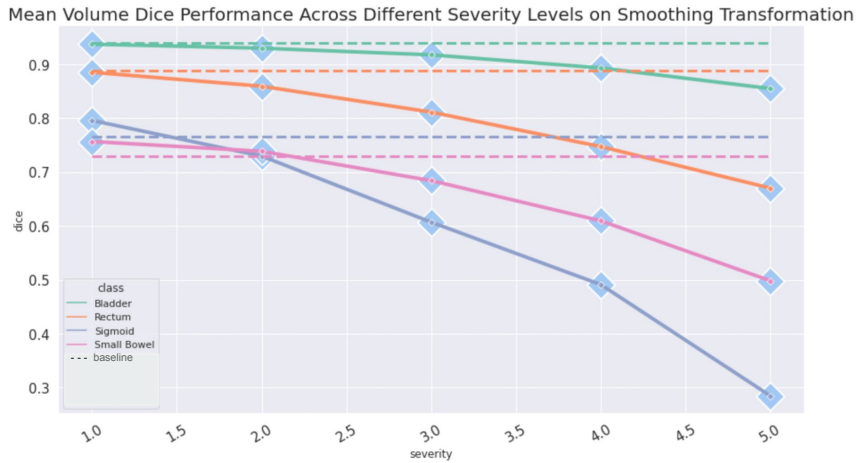
Results - Contrast and Intensity Transformations

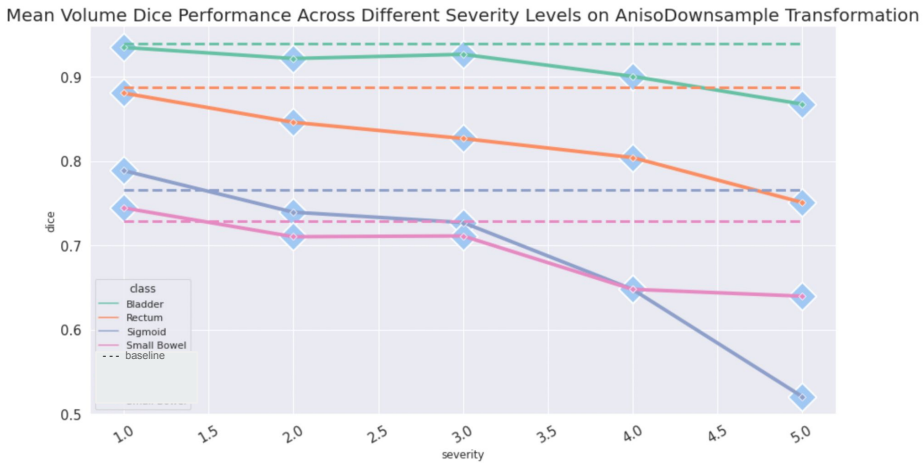




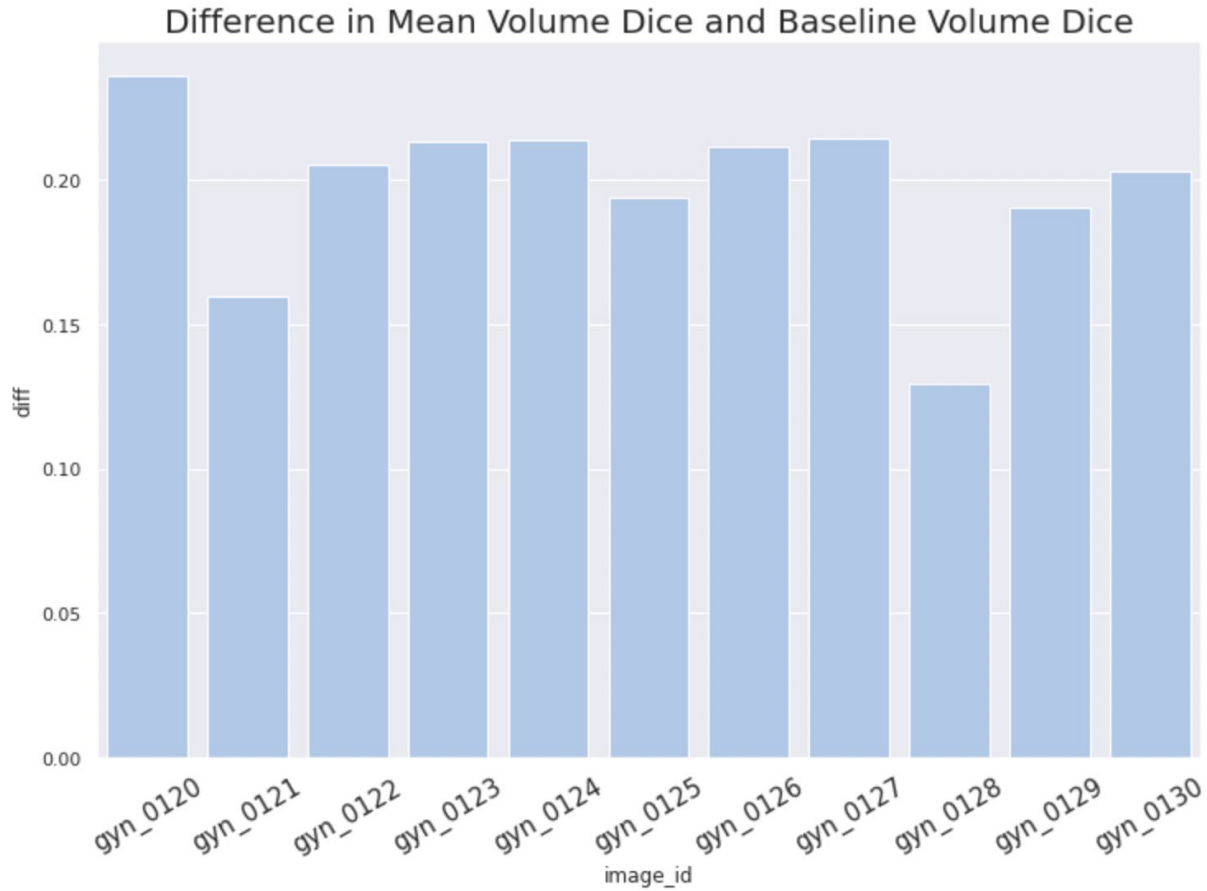


Results - Image Resolution and Blurring Transformations





Results - Performance by Image



Results - Mean vDSC

	Mean vDSC Baseline	Mean vDSC
Bladder	0.94	0.76
Rectum	0.89	0.67
Sigmoid	0.77	0.54
Small bowel	0.73	0.56

Results - Robustness Metrics

Weighted mean DSC:

$$\mathrm{wmDSC}_T = \frac{1}{\sum_{s=0}^{5} w_s} \sum_{s=0}^{5} w_s \cdot \mathrm{mDSC}_{T,s},$$

Weighted DSC sd:

$$\mathrm{wsDSC}_T = \frac{1}{\sum_{s=0}^{5} w_s} \sum_{s=0}^{5} w_s \cdot \mathrm{sDSC}_{T,s},$$

Mean-based DSC degradation:

$$\mathrm{mDDeg}_T = \frac{1}{\sum_{s=1}^{5} w_s} \sum_{s=1}^{5} w_s (\mathrm{mDSC}_{\mathrm{clean}} - \mathrm{mDSC}_{T,s}),$$

Variance-based DSC degradation:

$$\mathrm{vDDeg}_T = \frac{1}{\sum_{s=1}^{5} w_s} \sum_{s=1}^{5} w_s (\mathrm{sDSC}_{T,s} - \mathrm{sDSC}_{\mathrm{clean}}),$$

(Boone et al., 2022)

Results - Robustness Metrics

	wmDSC	wsDSC	mDDeg	vDDeg
Bladder	0.86	0.08	0.12	0.04
Rectum	0.78	0.06	0.17	0.05
Sigmoid	0.66	0.14	0.17	0.00
Small Bowel	0.65	0.15	0.13	0.01

Discussions

Baseline U-Net Model from Rood-MRI

	mDDeg	vDDeg
Hippocampi	0.13	0.06
Ventricles	0.13	0.07
WMHs	0.19	0.06

Performing good on:
Rician Noise (SNR)
Contrast Compression (contrast & intensity)
Biasfield (contrast & intensity)

Needs Improvement:
Elastic deformation (motion)
Ghosting (motion)
Random Motion (motion)

Acknowledgements

Benjamin Haibe-Kains

Ruiyan Ni

The BHK Lab

Princess Margaret Cancer Centre

Funding:

The Gattuso Slaight Personalized Cancer Medicine Fund

The Princess Margaret Cancer Foundation UHN

Canadian Cancer Society

NSERC CRSNG

CIHR IRSC

The Cancer Research Society

OICR

The Terry Fox Research Institute

STAND UP TO CANCER standup2cancer.ca

Thank you for your attention

References

(PDF) bias field correction for MRI images - researchgate. (n.d.). Retrieved August 2, 2022, from https://www.researchgate.net/publication/221512036_Bias_Field_Correction_for_MRI_Images

Boone, L., Biparva, M., Forooshani, P. M., Ramirez, J., Masellis, M., Bartha, R., Symons, S., Strother, S., Black, S. E., Heyn, C., Martel, A. L., Swartz, R. H., & Goubran, M. (2022, March 11). *Rood-MRI: Benchmarking the robustness of deep learning segmentation models to out-of-distribution and corrupted data in MRI*. arXiv.org. Retrieved August 1, 2022, from https://arxiv.org/abs/2203.06060

Smoothing¶. Smoothing - Introduction to MRI. (n.d.). Retrieved August 1, 2022, from http://jpeelle.net/mri/image_processing/smoothing.html

Zaitsev, M., Maclaren, J., & Herbst, M. (2015, October). *Motion artifacts in MRI: A complex problem with many partial solutions*. Journal of magnetic resonance imaging : JMRI. Retrieved August 1, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4517972/