Memorial Sloan Kettering Cancer Center

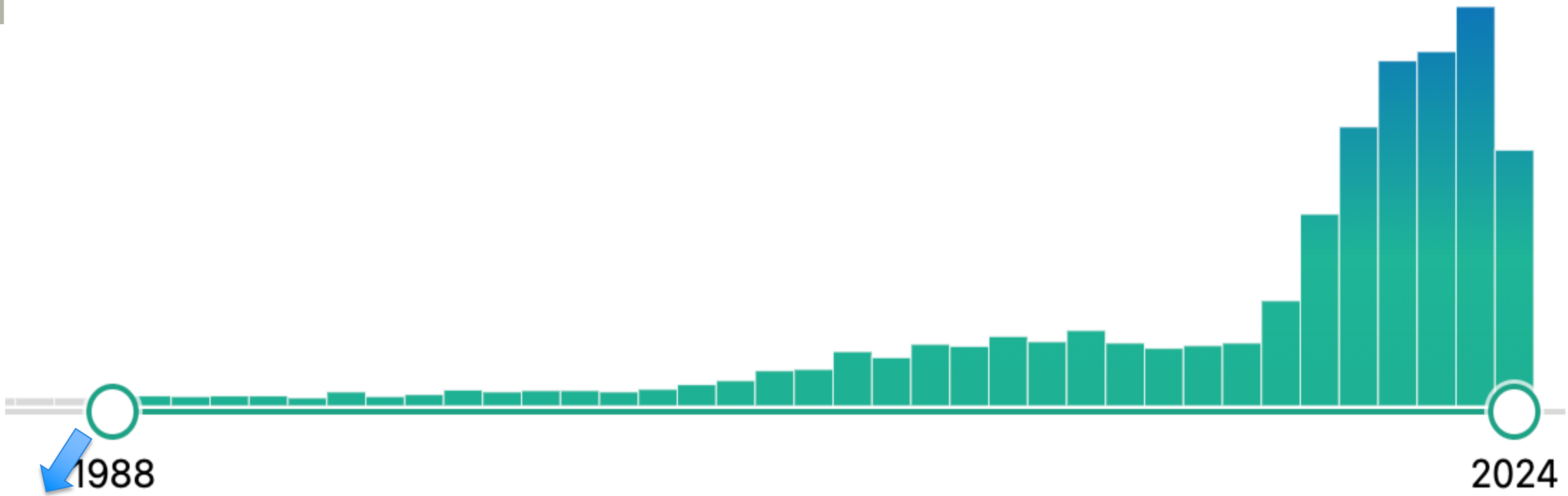
# Developing and deploying trustworthy AI models for cancer treatments

Harini Veeraraghavan, PhD
Director, AI for Image Guided Therapies
Associate Attending Computer Scientist
Memorial Sloan Kettering Cancer Center

ICCR 2024
XX[th] International Conference on the use of Computers in Radiation therapy

# *Use of AI in radiotherapy*

Pubmed search for papers "AI + radiotherapy"



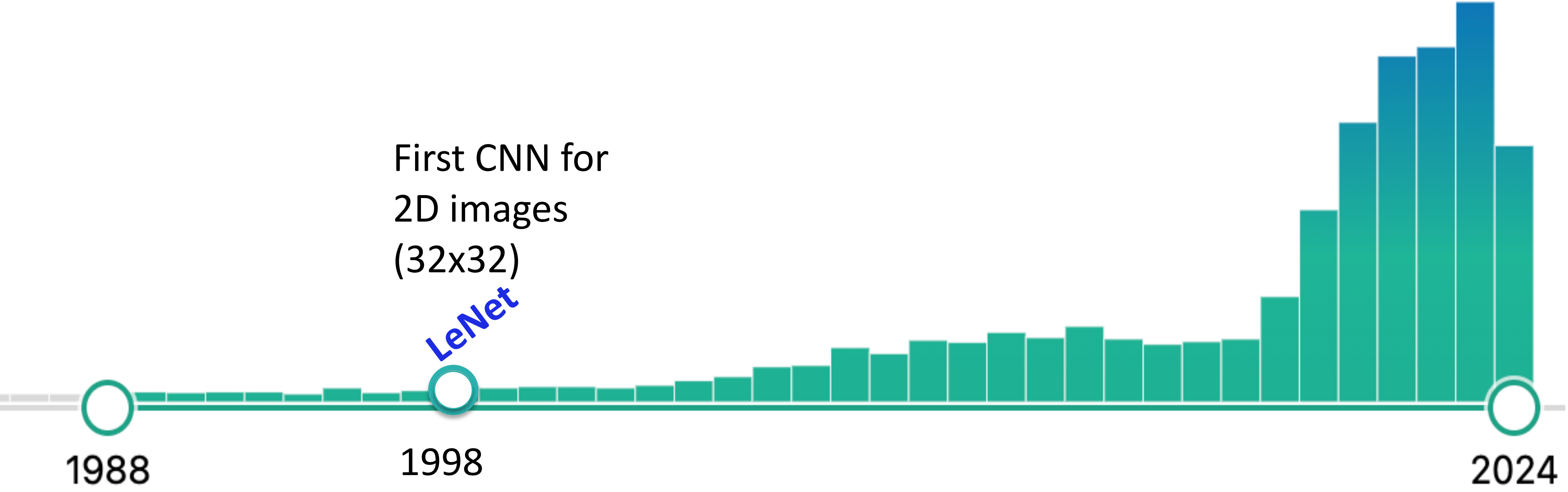

Applications of data bases and AI/expert systems in radiation therapy.

Laramore GE, Altschuler MD, Banks G, Kalet IJ, Pajak TF, Schultheiss TE, Zink S.

Am J Clin Oncol. 1988 Jun;11(3):387-93. doi: 10.1097/00000421-198806000-00015.

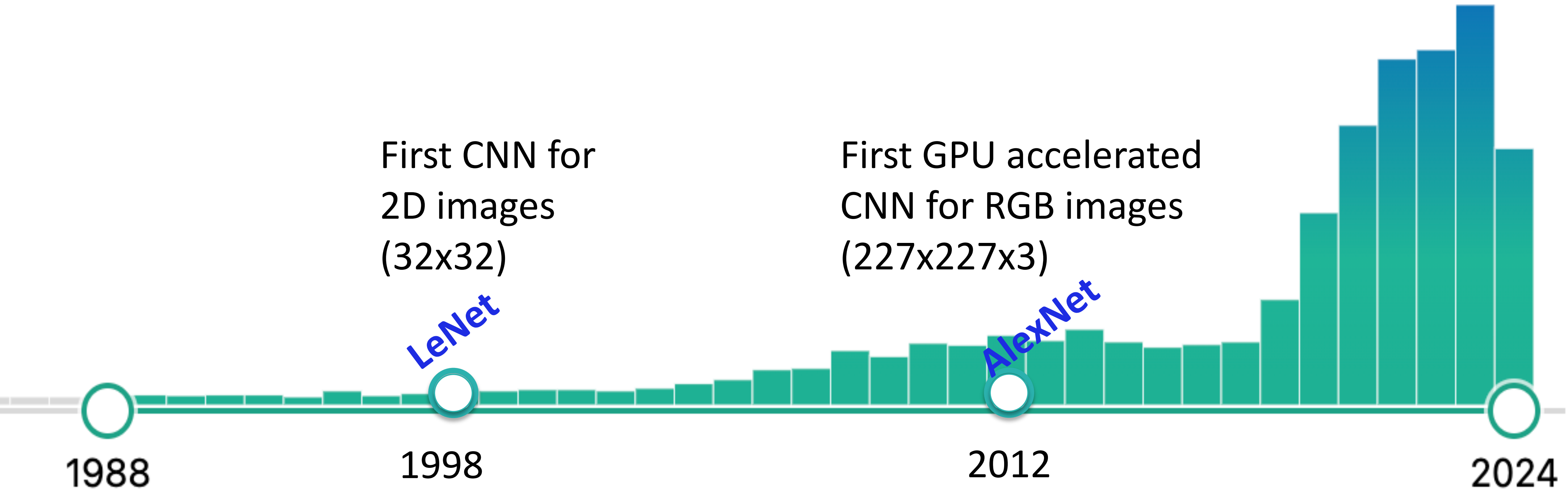
PMID: 3289369 Review. No abstract available.

# *Key innovations in deep learning*



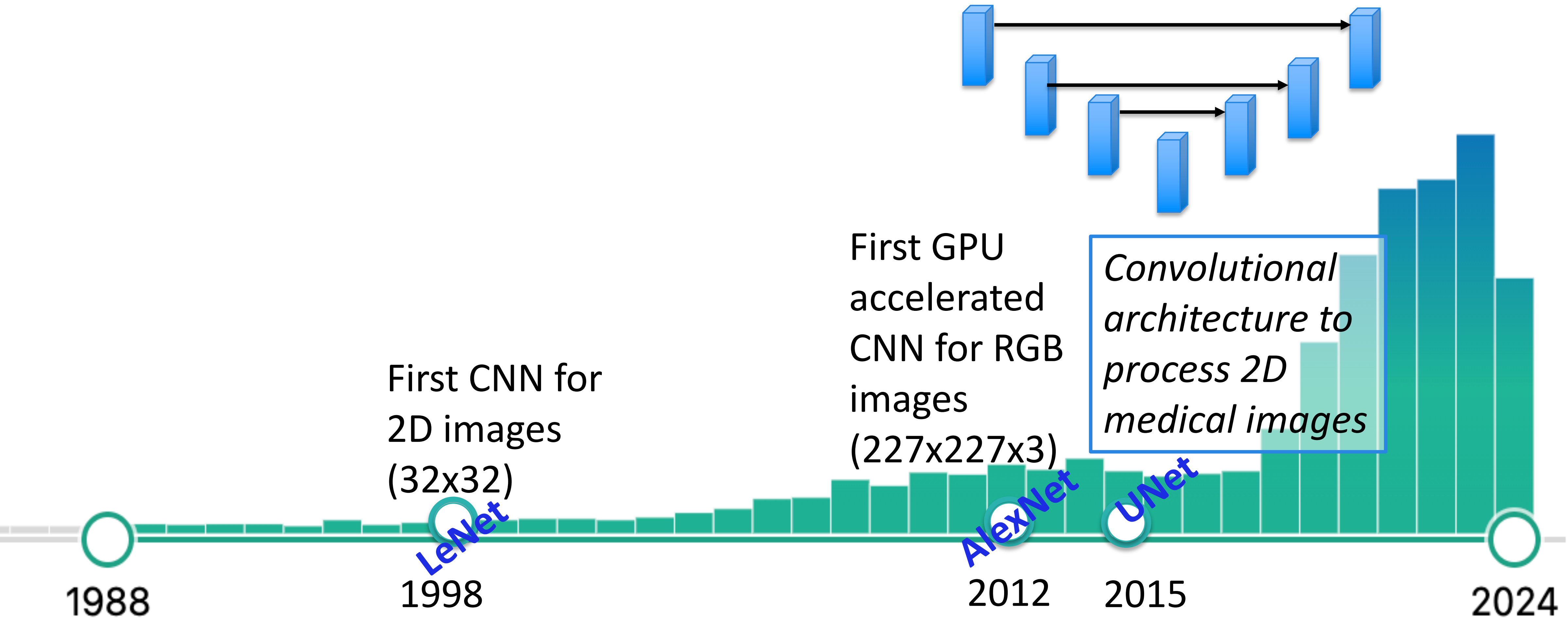

Memorial Sloan Kettering Cancer Center

# *Key innovations in deep learning*




Memorial Sloan Kettering
Cancer Center

# *Key innovations in deep learning*



# AI is applicable in almost every aspect of radiotherapy

Contents lists available at ScienceDirect

Radiotherapy and Oncology

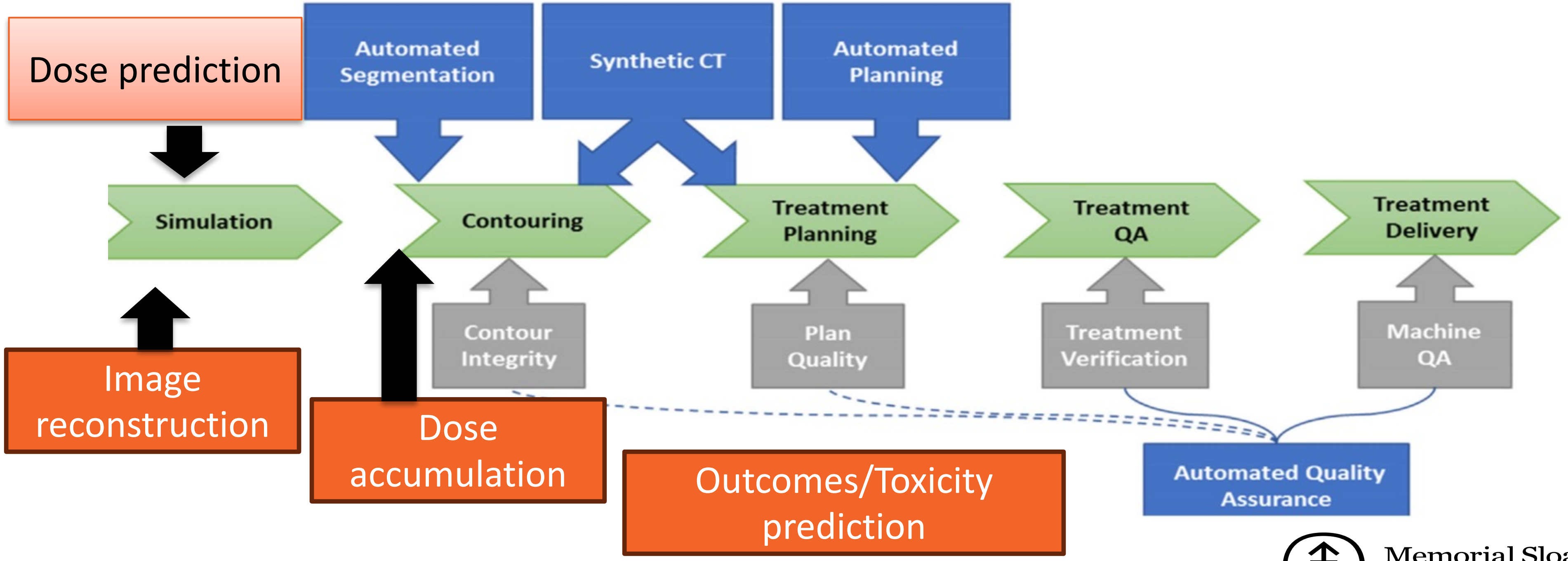
journal homepage: www.thegreenjournal.com

Review Article

## Overview of artificial intelligence-based applications in radiotherapy: Recommendations for implementation and quality assurance

Liesbeth Vandewinckele [a,b,1], Michaël Claessens [c,d,1], Anna Dinkla [e,1,*], Charlotte Brouwer [f], Wouter Crijns [a,b], Dirk Verellen [c,d], Wouter van Elmpt [g]

[a] Department Oncology, Laboratory of Experimental Radiotherapy, KU Leuven; [b] Department of Radiation Oncology, UZ Leuven; [c] Faculty of Medicine and Health Sciences, University of Antwerp; [d] Department of Radiation Oncology, Iridium Cancer Network, Wilrijk (Antwerp); [e] Department of Radiation Oncology, Amsterdam University Medical Center, University of Amsterdam; [f] University of Groningen, University Medical Center Groningen, Department of Radiation Oncology; and [g] Department of Radiation Oncology (Maastro), GROW School for Oncology, Maastricht University Medical Centre+



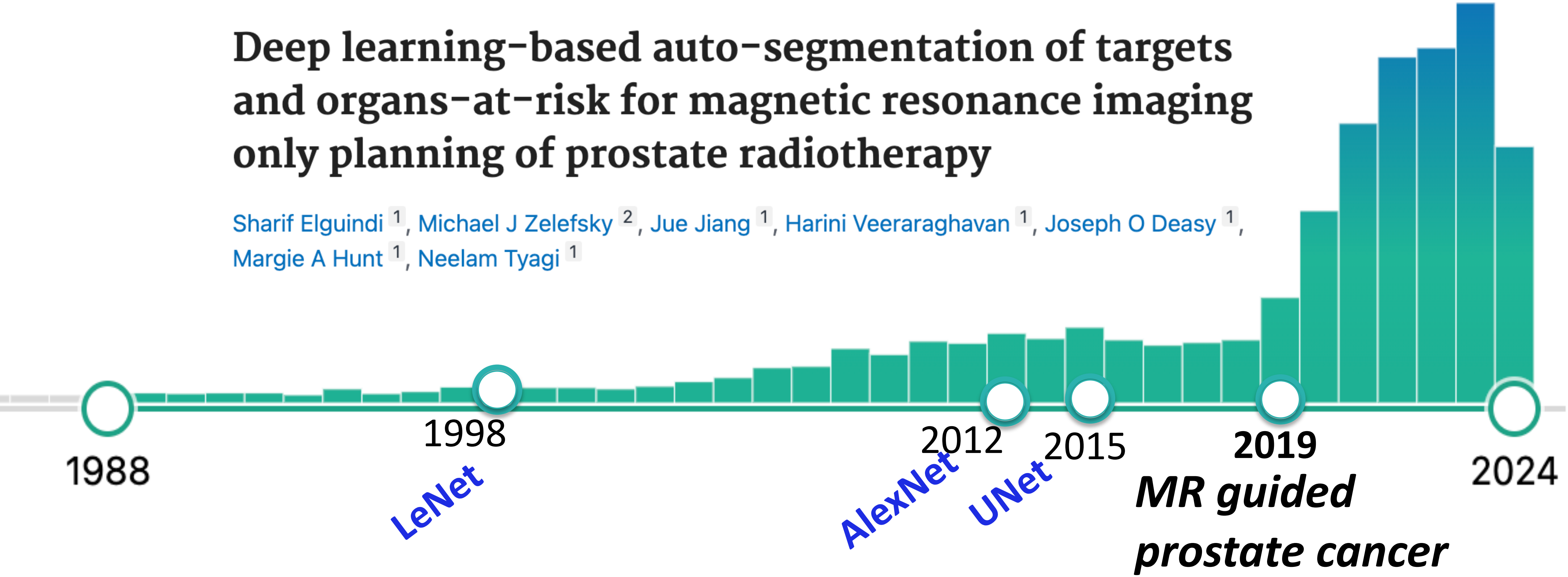

Memorial Sloan Kettering Cancer Center

# *Clinical implementation of deep learning at MSK*

> Phys Imaging Radiat Oncol. 2019 Oct:12:80-86. doi: 10.1016/j.phro.2019.11.006.
Epub 2019 Dec 12.

## Deep learning-based auto-segmentation of targets and organs-at-risk for magnetic resonance imaging only planning of prostate radiotherapy

Sharif Elguindi [1], Michael J Zelefsky [2], Jue Jiang [1], Harini Veeraraghavan [1], Joseph O Deasy [1], Margie A Hunt [1], Neelam Tyagi [1]



# *Challenges for clinical implementation*

Expectation

AI model Development

Training

*Optimize parameters*

Validation

*Select hyper-parameters*

*Data released after model building*

Prevents data leakage and ensures validity of model results

Reality due to data limitations

AI model Development

Training + Validation

Training + Validation + Testing

Requires nested cross validation to avoid data leak; Results are sub-optimal

## *Challenges for clinical implementation*

**Expectation**

AI model

Clinical test data resembles training data

In distribution data (ID) is used

**Reality**

AI model

Covariate shift due to data distribution differences

Concept shift due to pathological differences

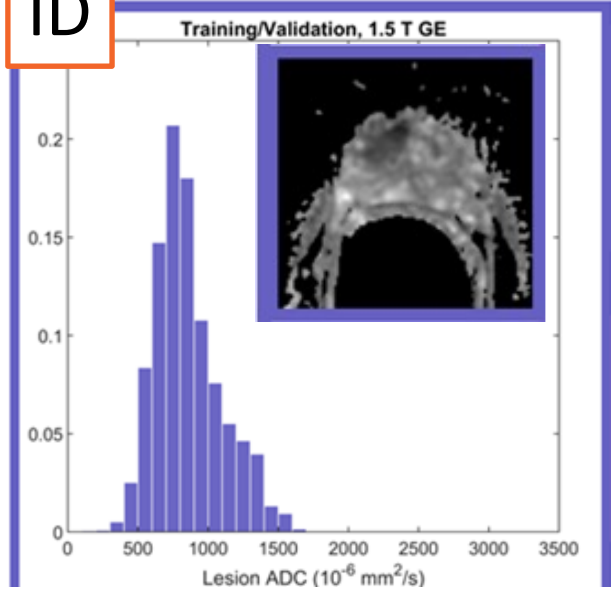
Some data is ID, a lot of it is out of distribution (OOD)

# *Clinical Implementation Challenges: Covariate shift*

Training + Validation

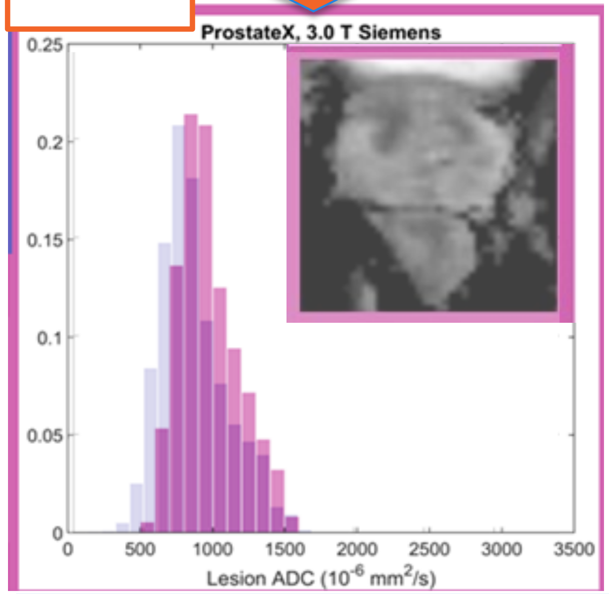
AI model
(Intra-prostatic lesion)

ID



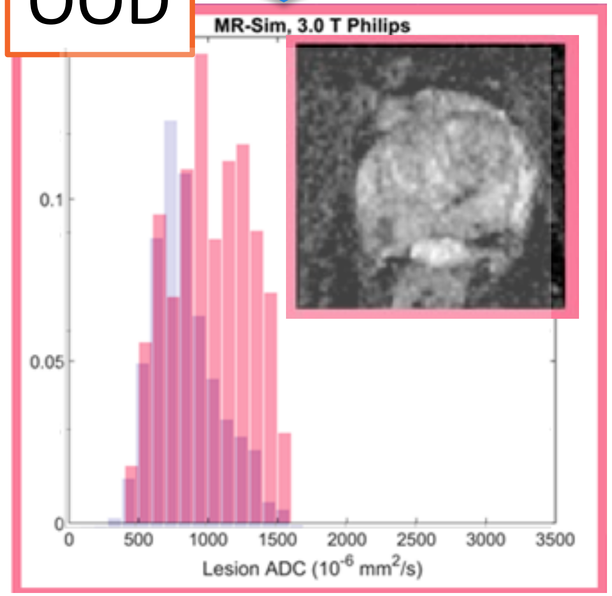

1.5 Tesla GE with endorectal coil **test**

OOD



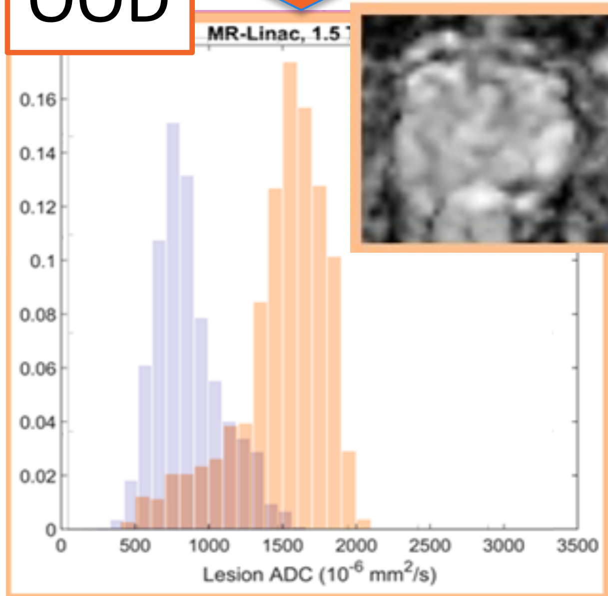

3.0 Tesla Siemens **test**

OOD




3.0 Tesla Phillips **test**

OOD




1.5 Tesla Phillips **test**

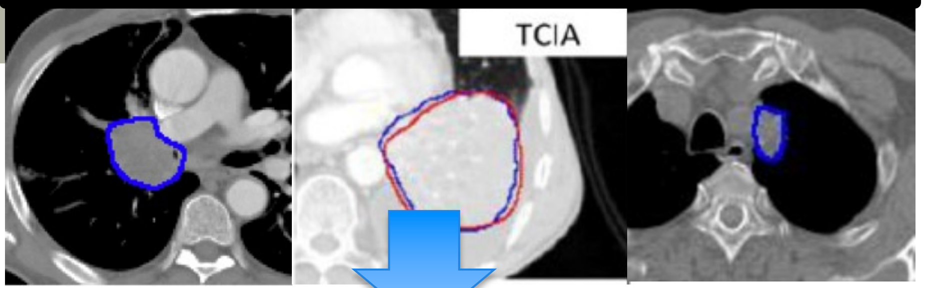
*Differences in Apparent diffusion coefficient distributions of **1,277** prostate cancers taken from **four different testing datasets***

**Simeth … Veeraraghavan et.al**

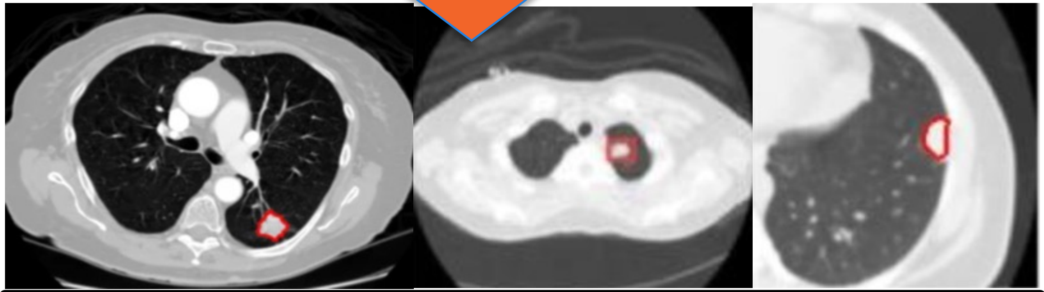
Memorial Sloan Kettering Cancer Center

# *Clinical Implementation Challenges: Concept drift*

*Training: Stage I – III non-small cell lung cancer*



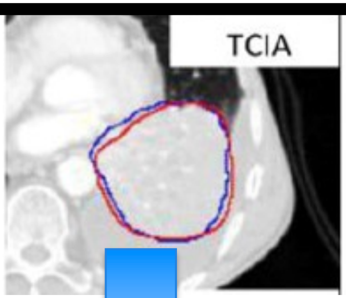

AI model



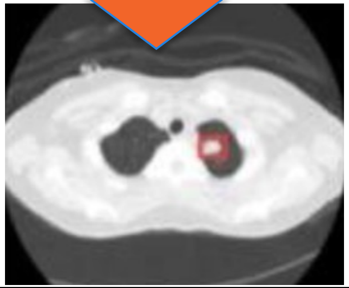
*Testing: Pre & non-cancerous lesions*

# Clinical Implementation Challenges: Concept drift

*Training: Stage I – III non-small cell lung cancer*

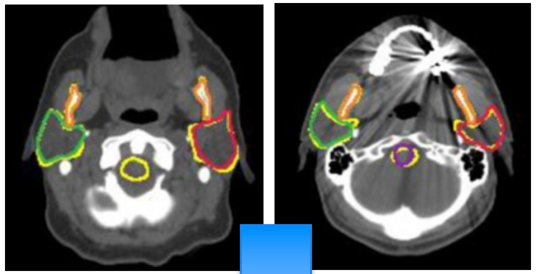



AI model

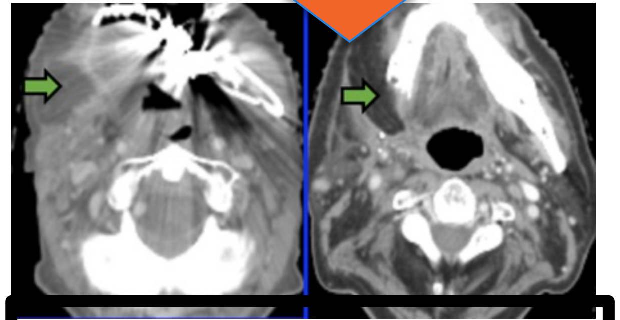


*Testing: Pre & non-cancerous lesions*

*Training: normal anatomy*



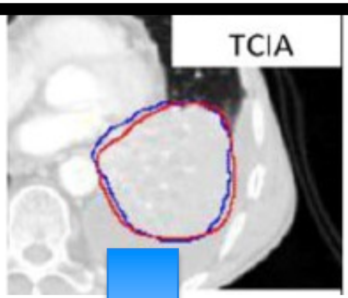
AI model



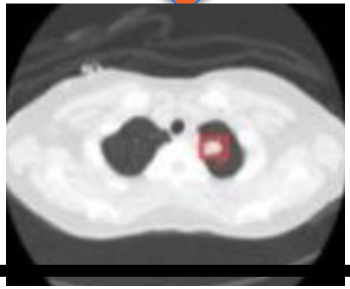
*Testing: Abnormal anatomy*

# *Clinical Implementation Challenges: Concept drift*

*Training: Stage I – III non-small cell lung cancer*

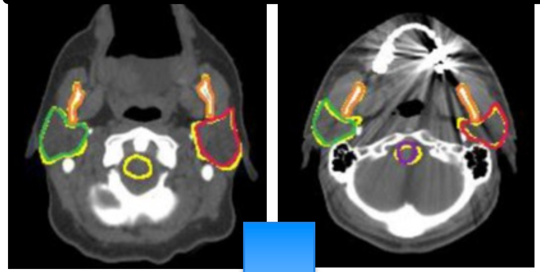



AI model

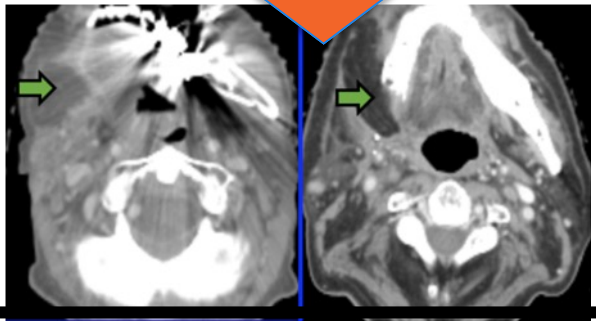


*Testing: Non cancers*

*Training: normal anatomy*

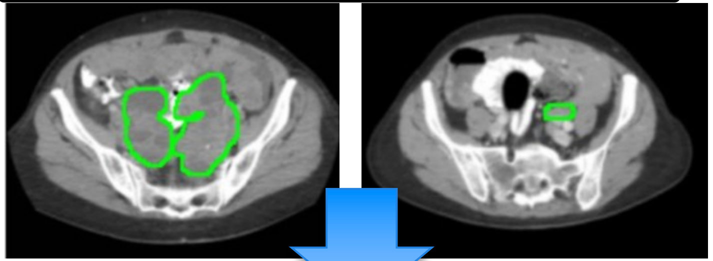


AI model

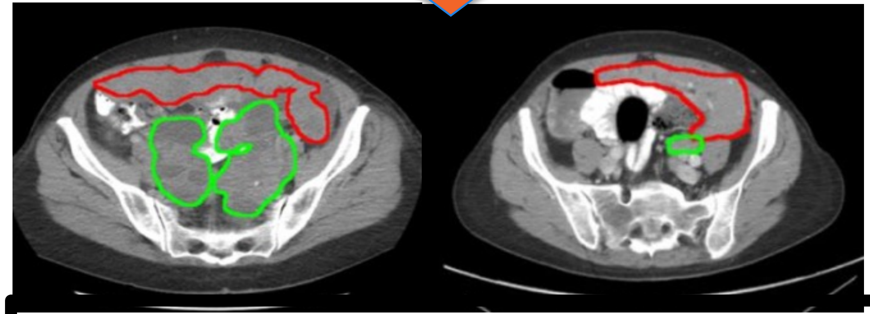


*Testing: Abnormal anatomy*

*Training: Incomplete segmentations*



AI model



*Testing: Complete segmentations*

# *Increasing trustworthiness of AI models in clinic*

*Online and routine Quality Assurance*

*Training* → AI model → *Clinical Deployment*

***Concept Drift:***
*Different disease; abnormal anatomy; Different output*

***Solutions:***
- *Maximally use labeled data*
- *Increase data diversity through domain adaptation*
- *Assess performance drifts*

***Covariate shift:***
*Different distribution; Contrast differences;..*

# Improving robustness of AI models

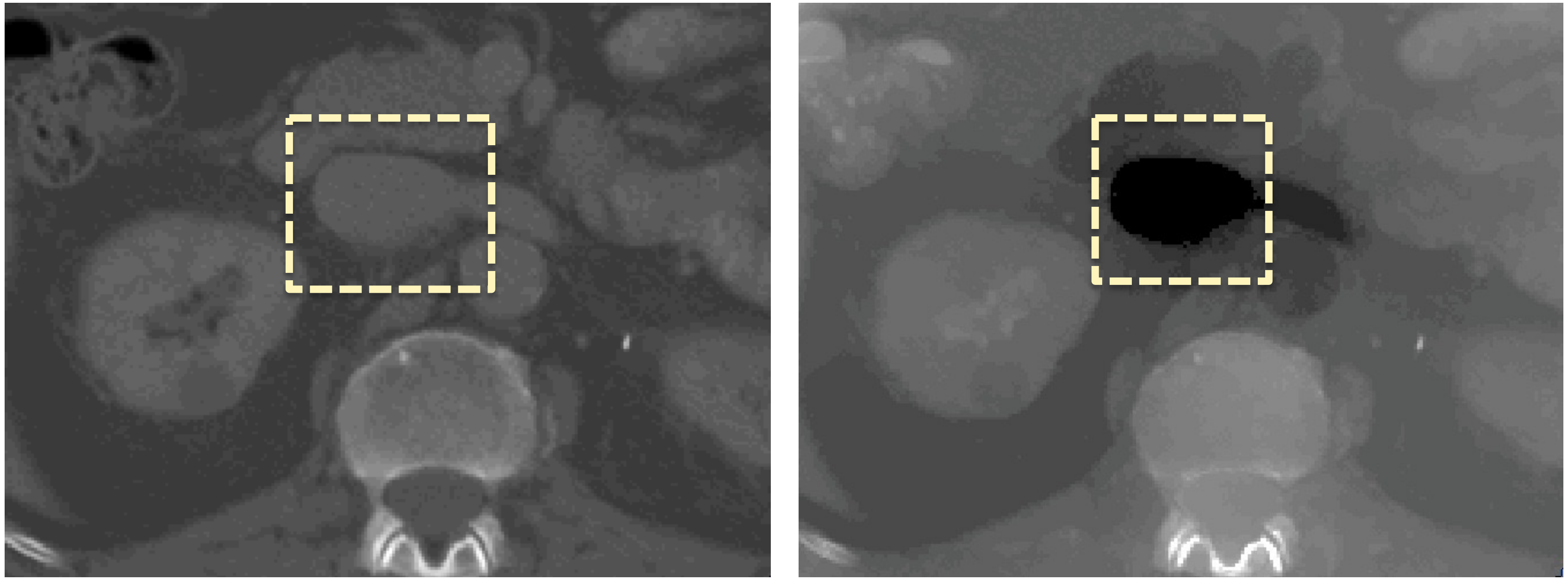
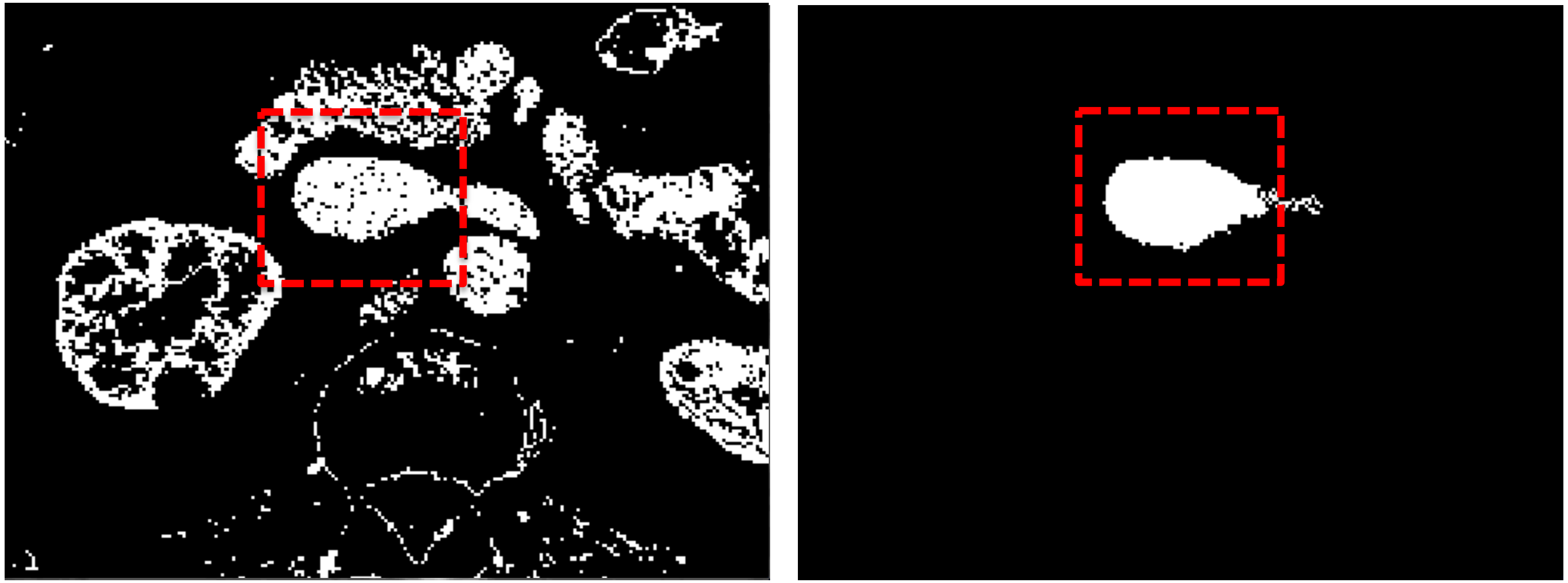


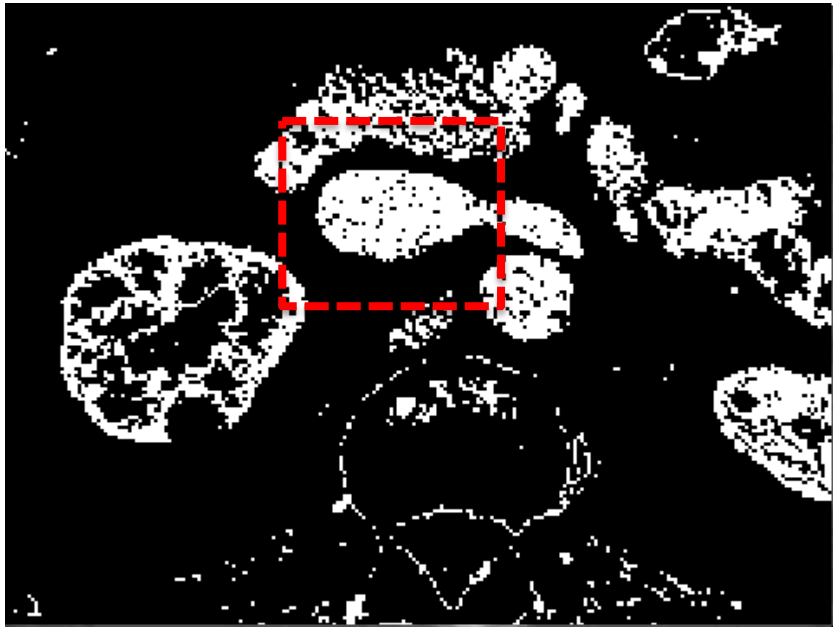
**Image representation is key to good segmentation**

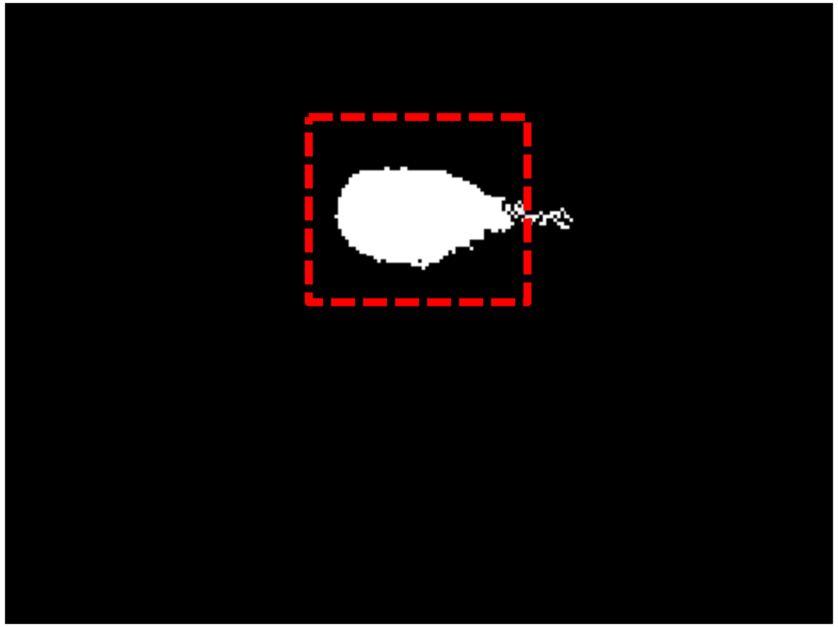
# “Learning” is about extracting useful features



***Model training** should focus on extracting features that robustly differentiate foreground (structure/organ of interest) from background*

# "Learning" is about extracting useful features



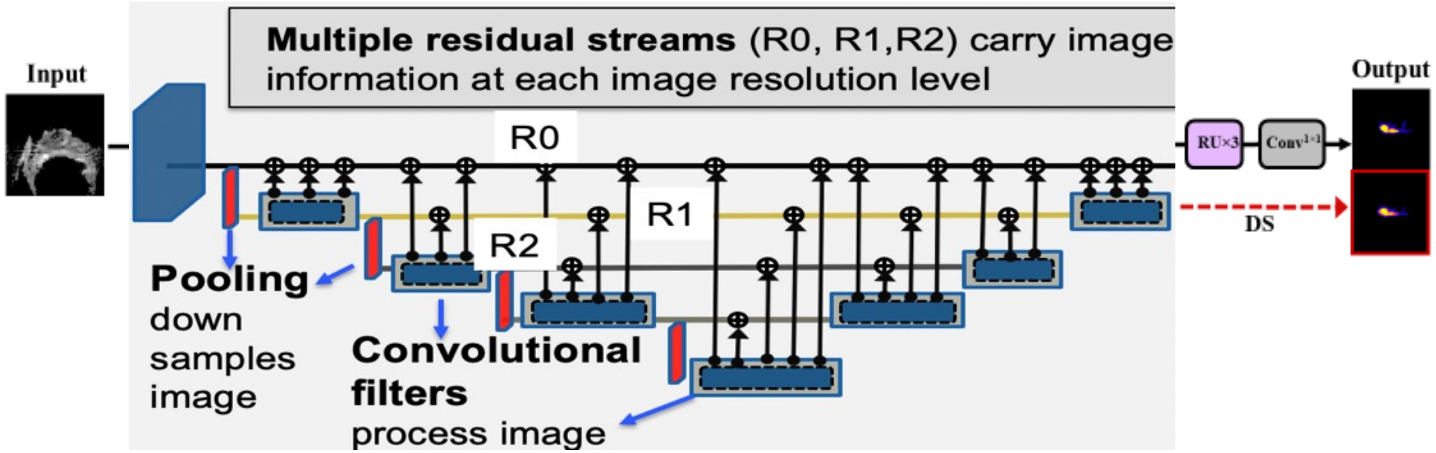


- *Increase training data variability*
- *Regularize training to extract "useful" features*

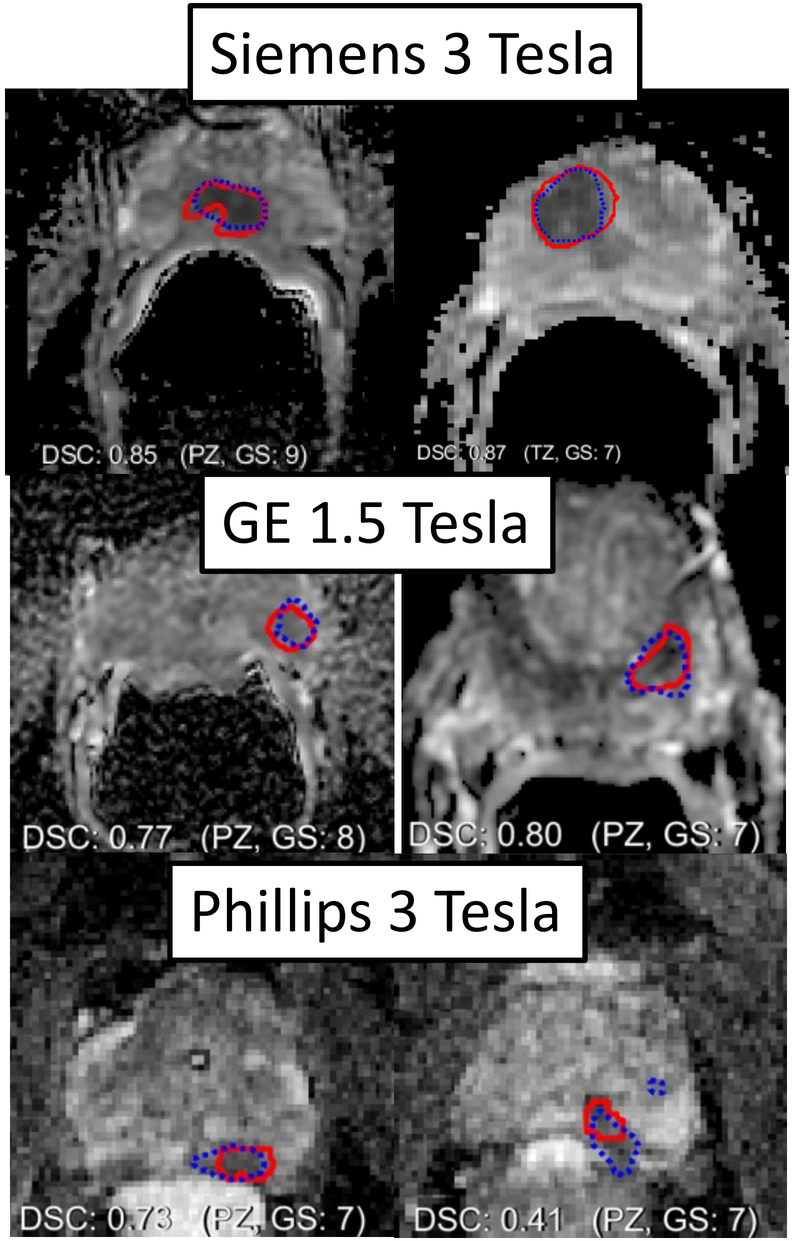
> Med Phys. 2023 Aug;50(8):4854-4870. doi: 10.1002/mp.16320. Epub 2023 Mar 13.

# Deep learning-based dominant index lesion segmentation for MR-guided radiation therapy of prostate cancer

Josiah Simeth [1], Jue Jiang [1], Anton Nosov [2], Andreas Wibmer [2], Michael Zelefsky [3], Neelam Tyagi [1], Harini Veeraraghavan [1]



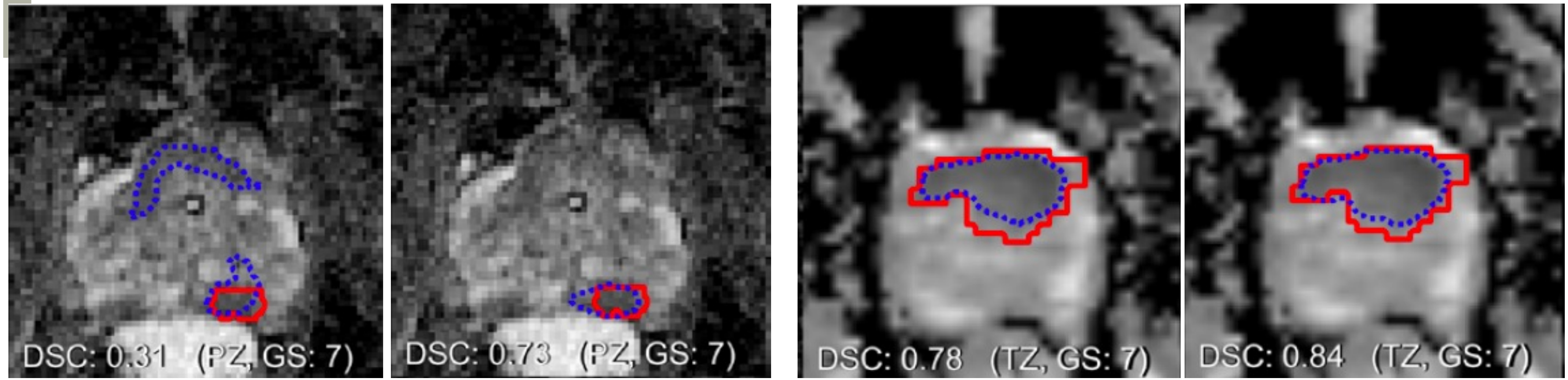

*Extract a richer feature representation by combining residual and dense connections in a deep network*




Memorial Sloan Kettering Cancer Center

# Improving segmentation accuracy on OOD data

Unet MRRN Unet MRRN



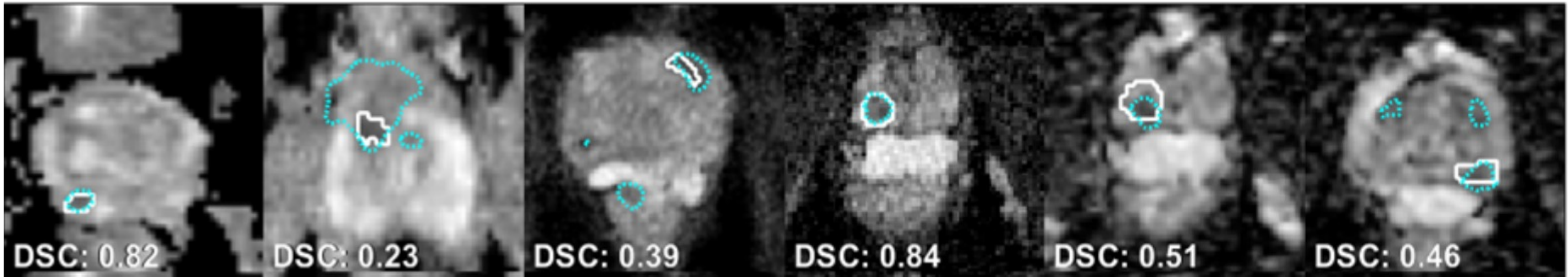

*Phillips 3 Tesla* *Siemens 3 Tesla*

*Extracting a richer feature representation by combining residual connections and dense features increases capability to handle data variations in MRRN*

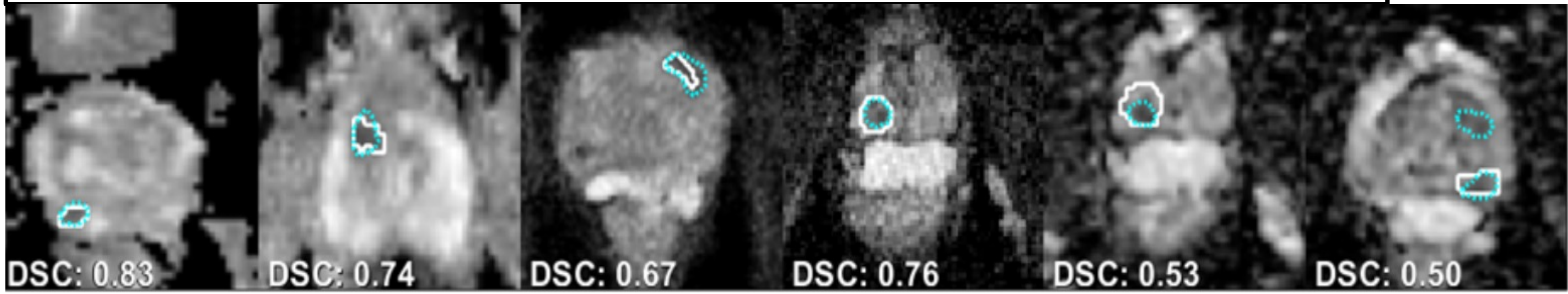
# *Implicit feature augmentation with mixup to further regularize network training*

*MRRN*

3T Siemens | 3T Philips | 1.5T Philips



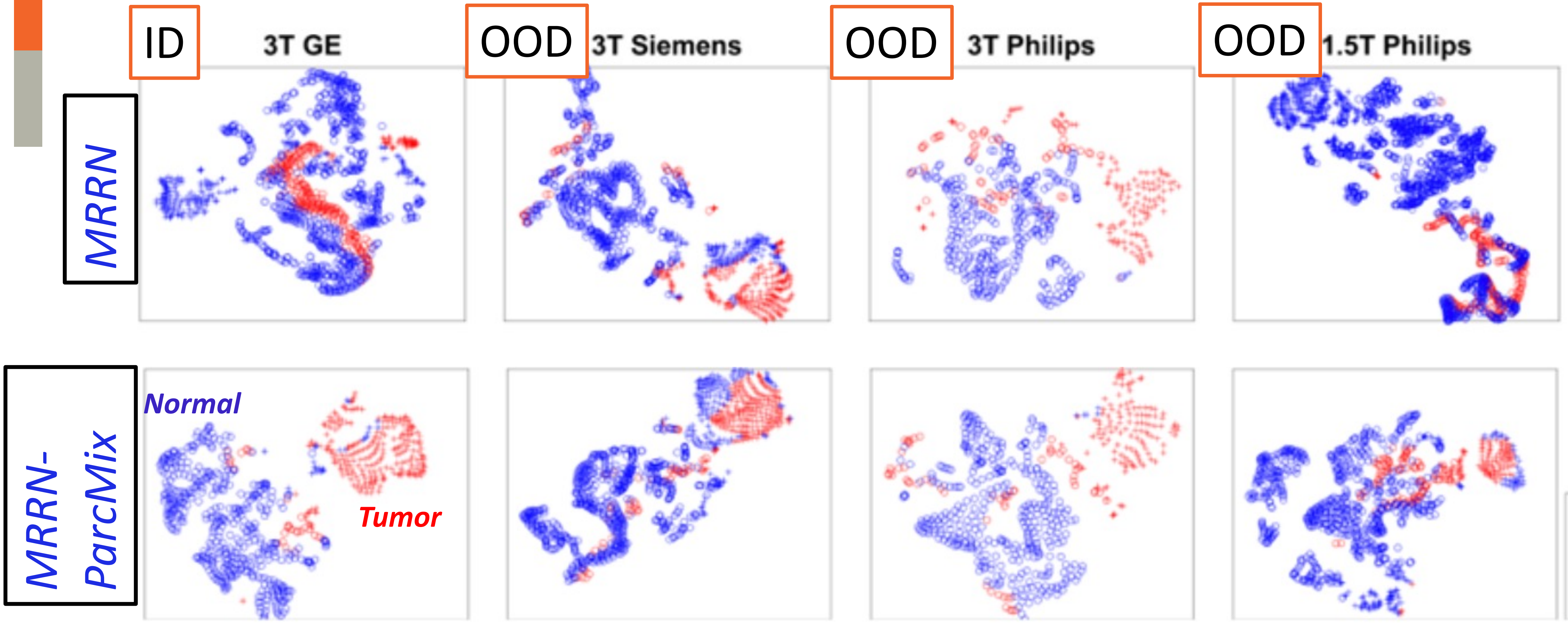

*MRRN with **Par**allel **C**oherent **mix**up (Parc-mix)*




**Simeth, … Veeraraghavan**

Memorial Sloan Kettering Cancer Center

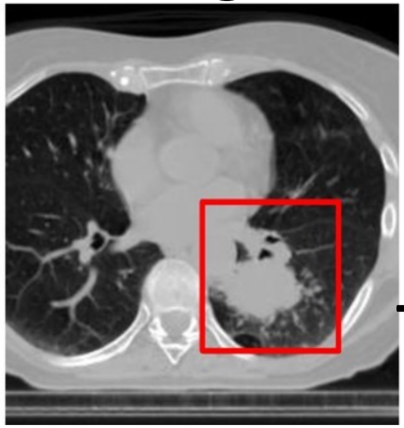
# *Improved regularization leads to better generalization across scanners*




**Analysis: Josiah Simeth, PhD**

# *Using domain adaptation for increasing data variability*

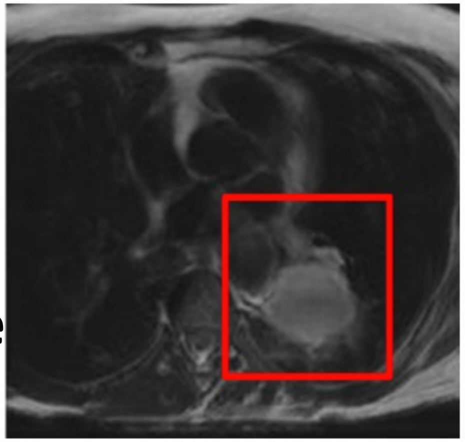
CT Image



Our approach

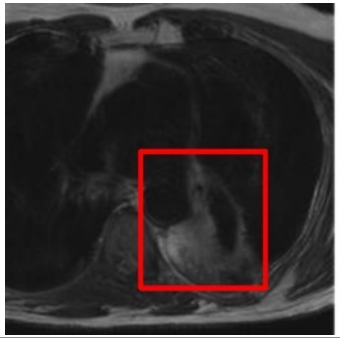
**Tumor-aware**

Generated MRI

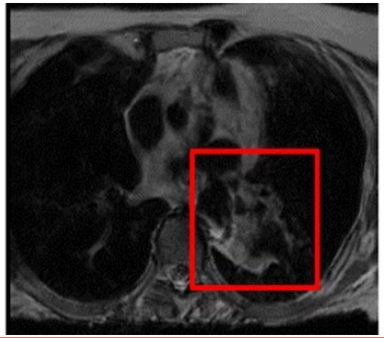


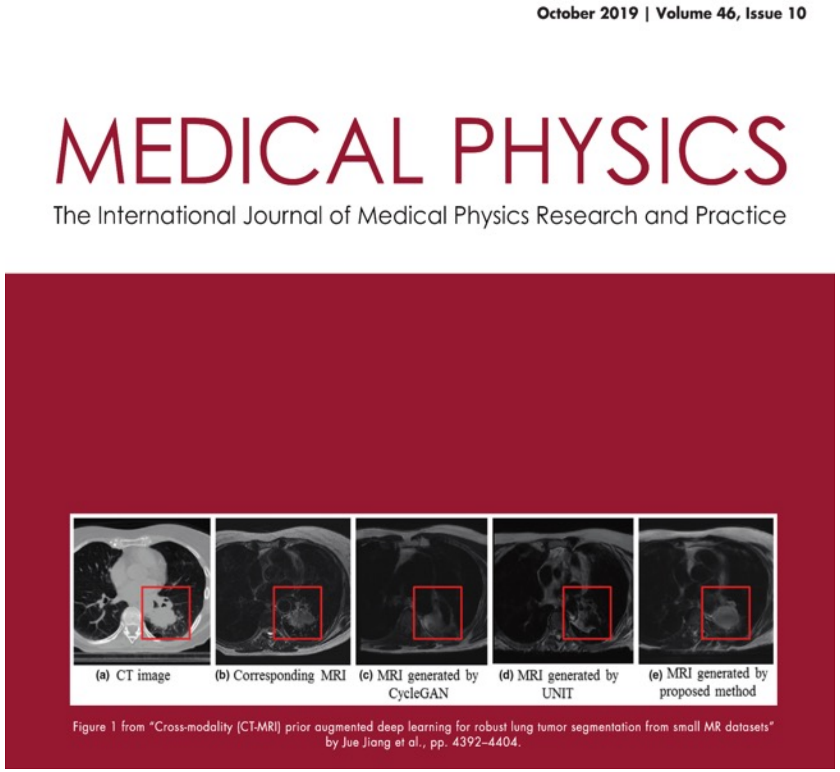
Comparison methods

Cycle GAN

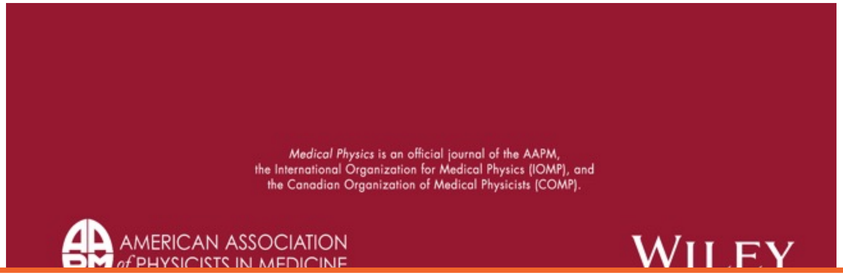


UNIT GAN






Jiang J, .. Deasy J, Veeraraghavan H 2018



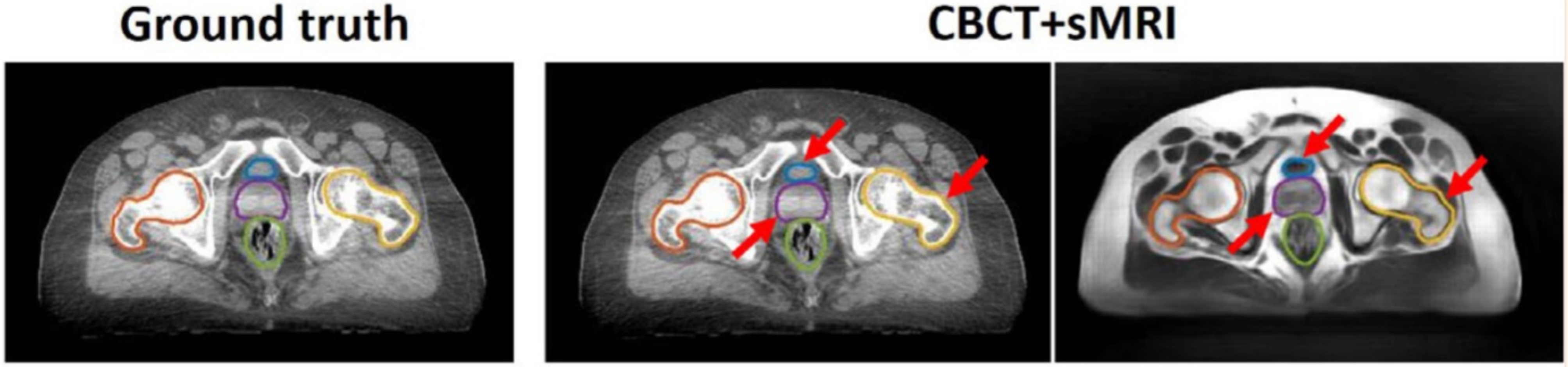

Good constraints are required to ensure preservation of structures like tumors

# *Cross domain adaptation for improving segmentation*

> Med Phys. 2020 Aug;47(8):3415-3422. doi: 10.1002/mp.14196. Epub 2020 May 11.

## Pelvic multi-organ segmentation on cone-beam CT for prostate adaptive radiotherapy

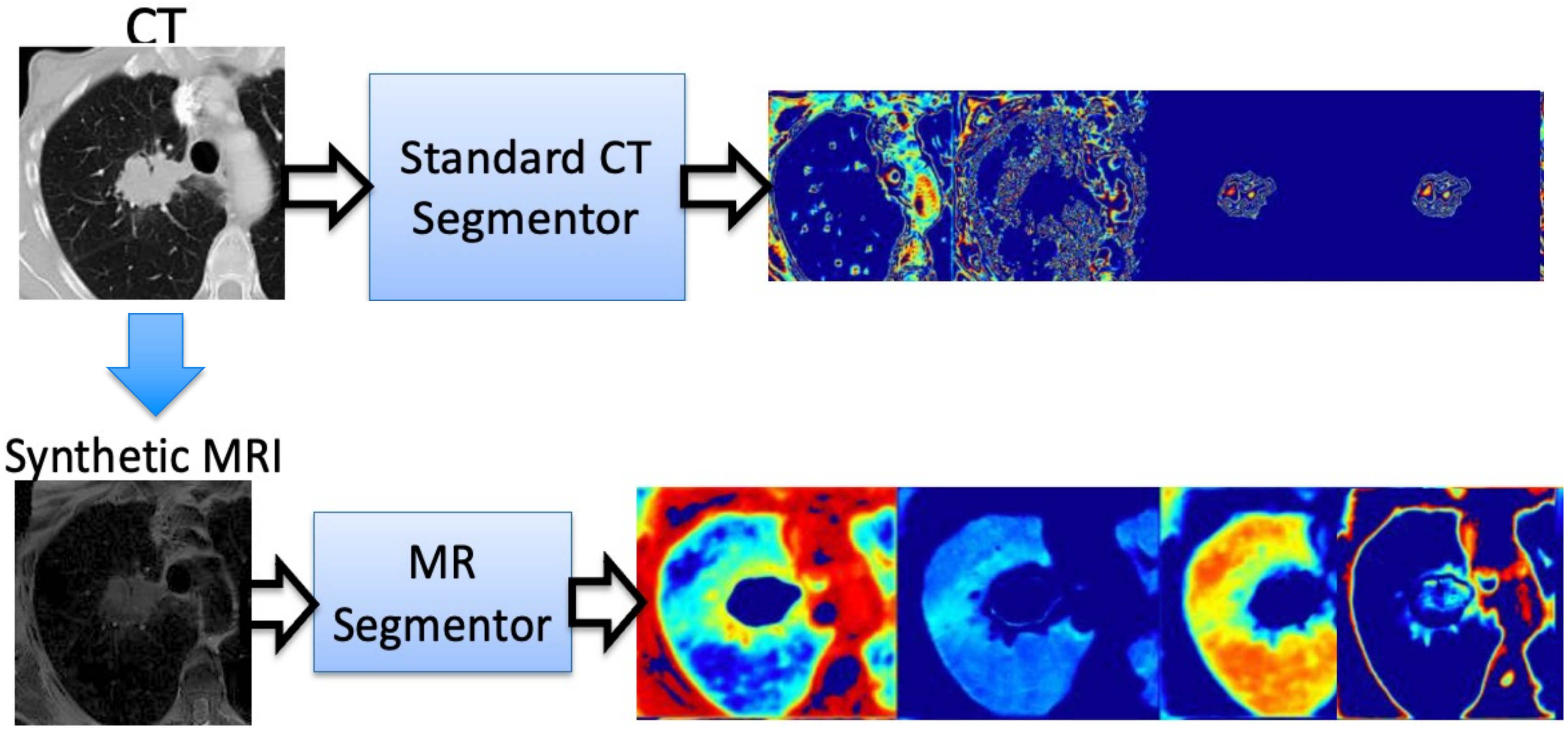
Yabo Fu [1], Yang Lei [1], Tonghe Wang [1], Sibo Tian [1], Pretesh Patel [1], Ashesh B Jani [1], Walter J Curran [1], Tian Liu [1], Xiaofeng Yang [1]




Uses high soft-tissue contrast in MRI to improve cone beam CT segmentation

# Cross modality distillation to learn "better" features

***Problem:*** *Features extracted directly from CT cannot sufficiently discriminate tumor from background*




Memorial Sloan Kettering Cancer Center

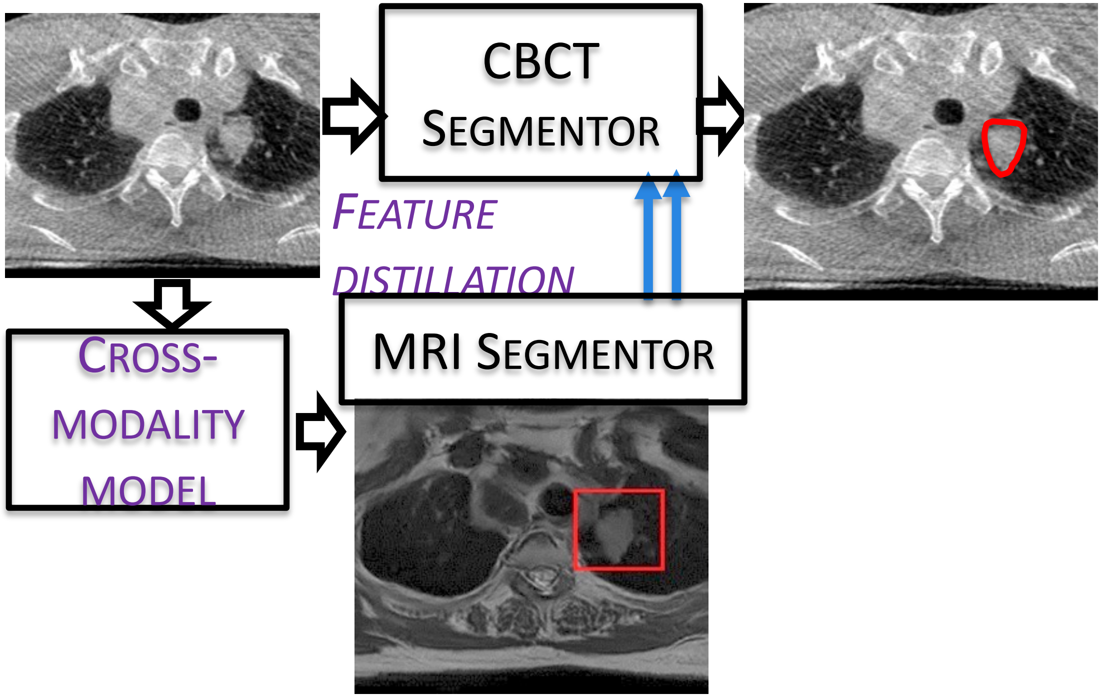
# *Leverage MRI to regularize CBCT network and extract "relevant" features*

> Med Phys. 2021 Apr 27. doi: 10.1002/mp.14902. Online ahead of print.

## Deep cross-modality (MR-CT) educed distillation learning for cone beam CT lung tumor segmentation

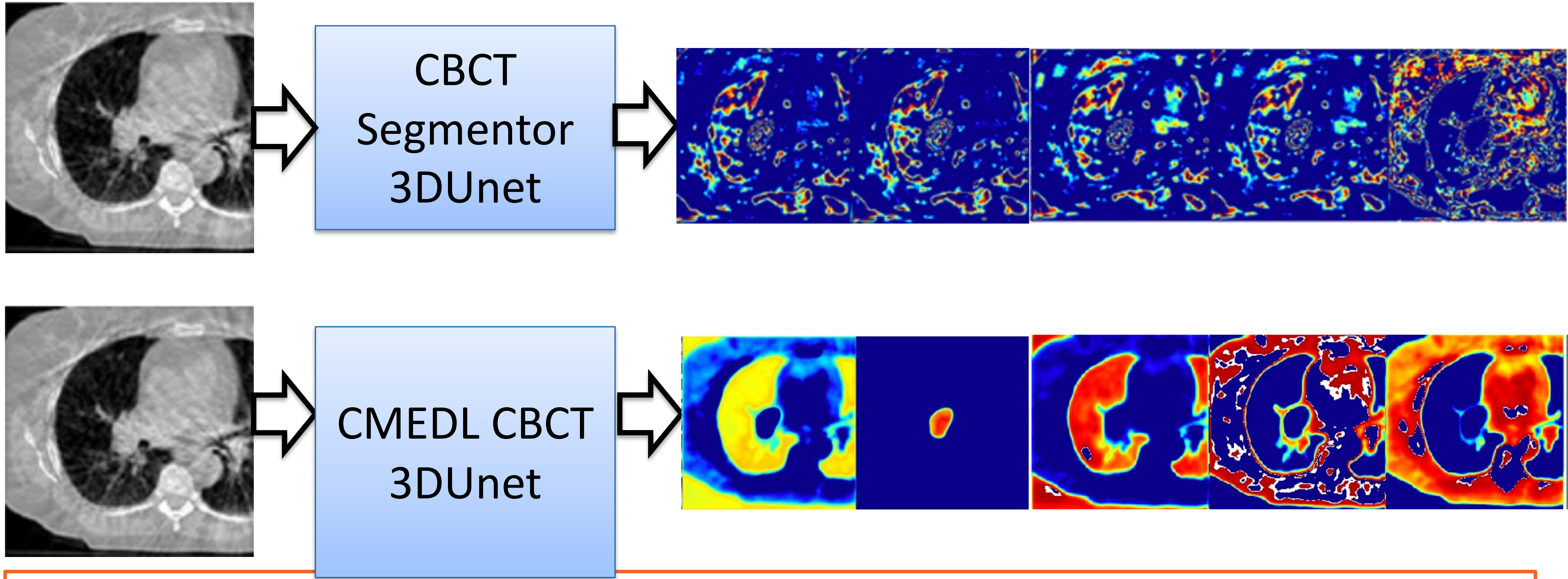
Jue Jiang [1], Sadegh Riyahi Alam [1], Ishita Chen [2], Perry Zhang [1], Andreas Rimner [2], Joseph O Deasy [1], Harini Veeraraghavan [1]

- <u>Practical</u> no need for paired multi-modality datasets
- <u>Efficient</u> only needs the CBCT network for testing
- <u>Accurate</u> because errors in synthesis do not propagate at testing



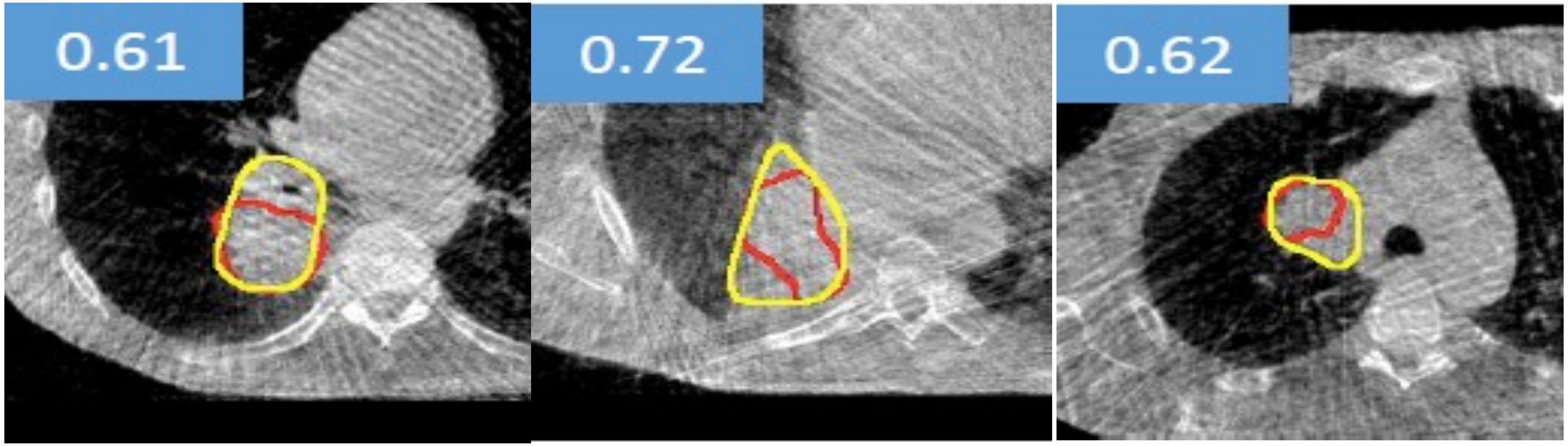

Memorial Sloan Kettering Cancer Center

# *Why does CMEDL regularization work?*

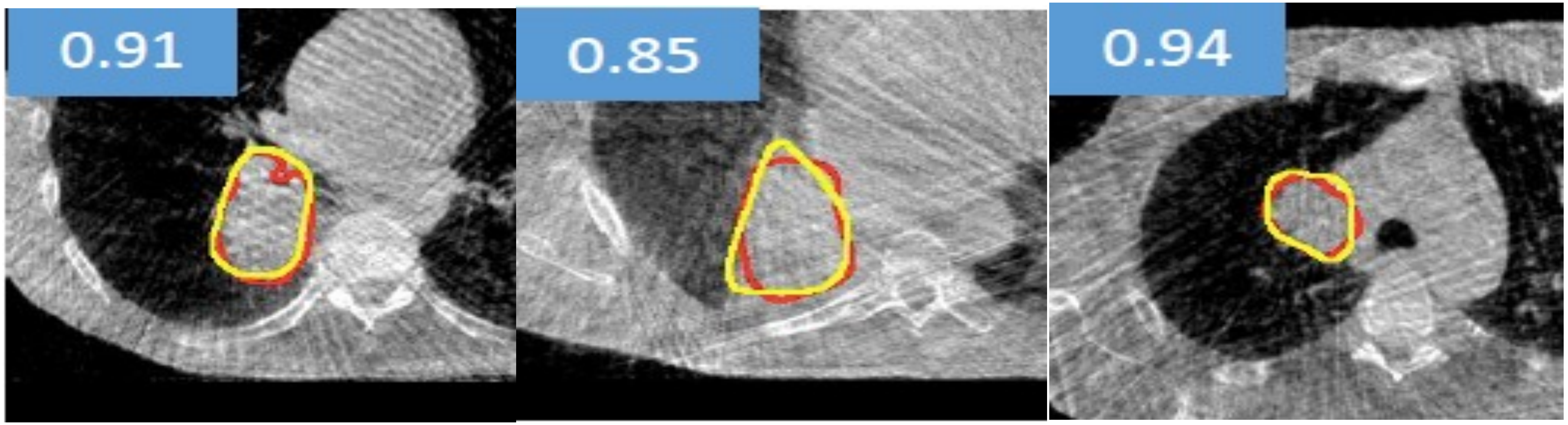



Cross modality distillation forces network to learn features signaling contrast between foreground and background

# *Does CMEDL improve accuracy?*



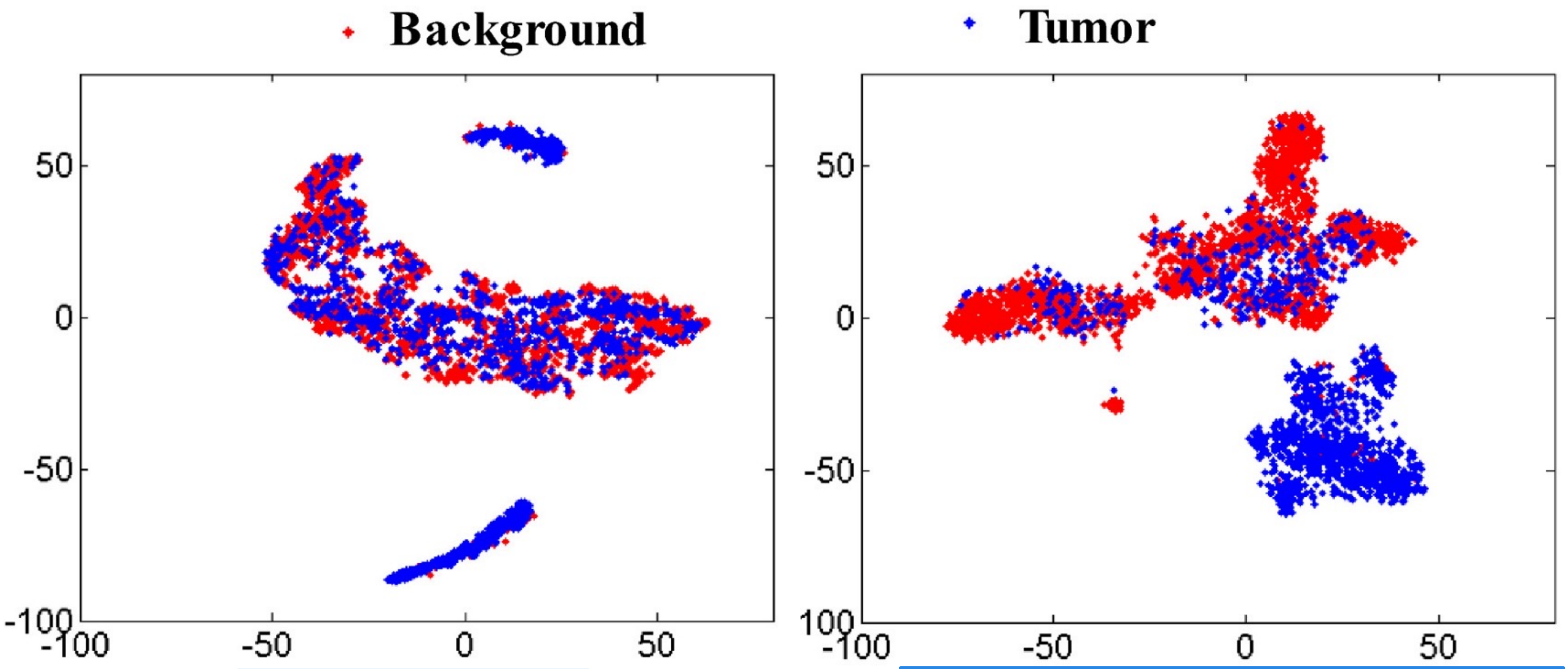

CBCT using 3D Unet




CBCT using CMEDL 3D Unet

Algorithm
Expert

# *Does CMEDL really work?*

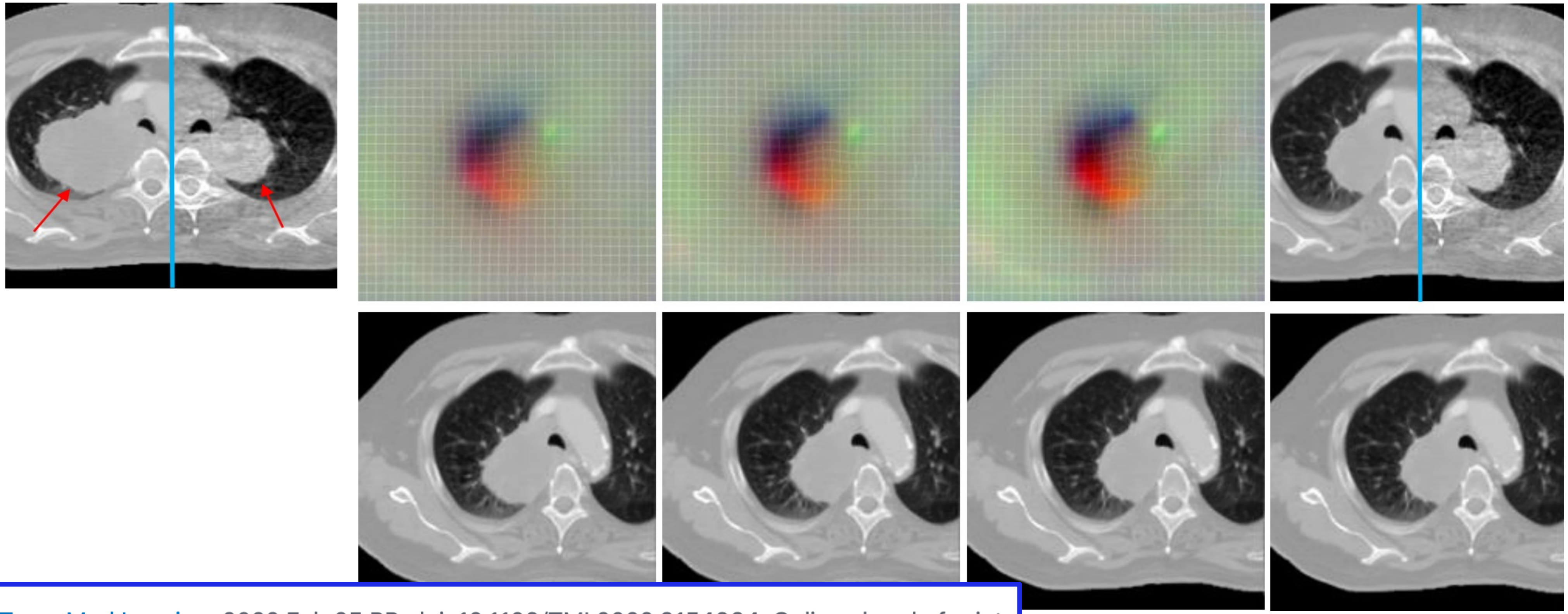



*CBCT 3D Unet*

*CBCT CMEDL 3D Unet*

*Results from mix of public and internal 38 test cases*

# *Patient specific segmentation combining registration*



> IEEE Trans Med Imaging. 2022 Feb 25;PP. doi: 10.1109/TMI.2022.3154934. Online ahead of print.

**One shot PACS: Patient specific Anatomic Context and Shape prior aware recurrent registration-segmentation of longitudinal thoracic cone beam CTs**

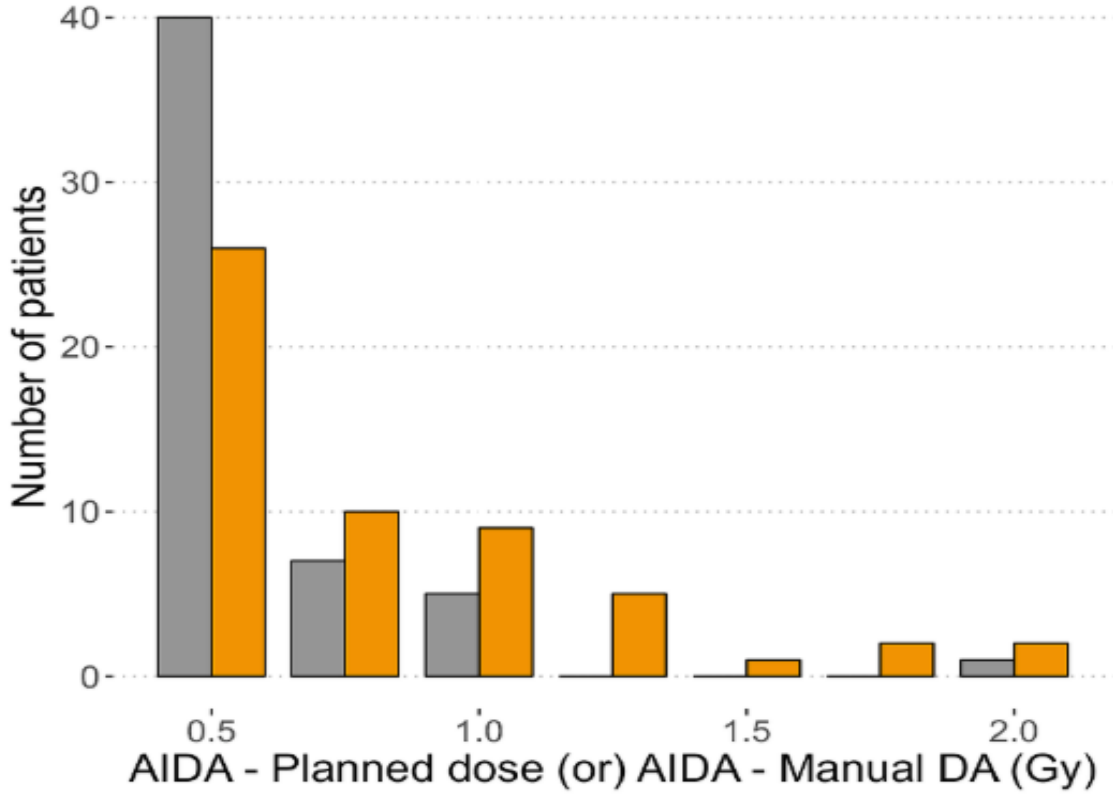
Jue Jiang, Harini Veeraraghavan

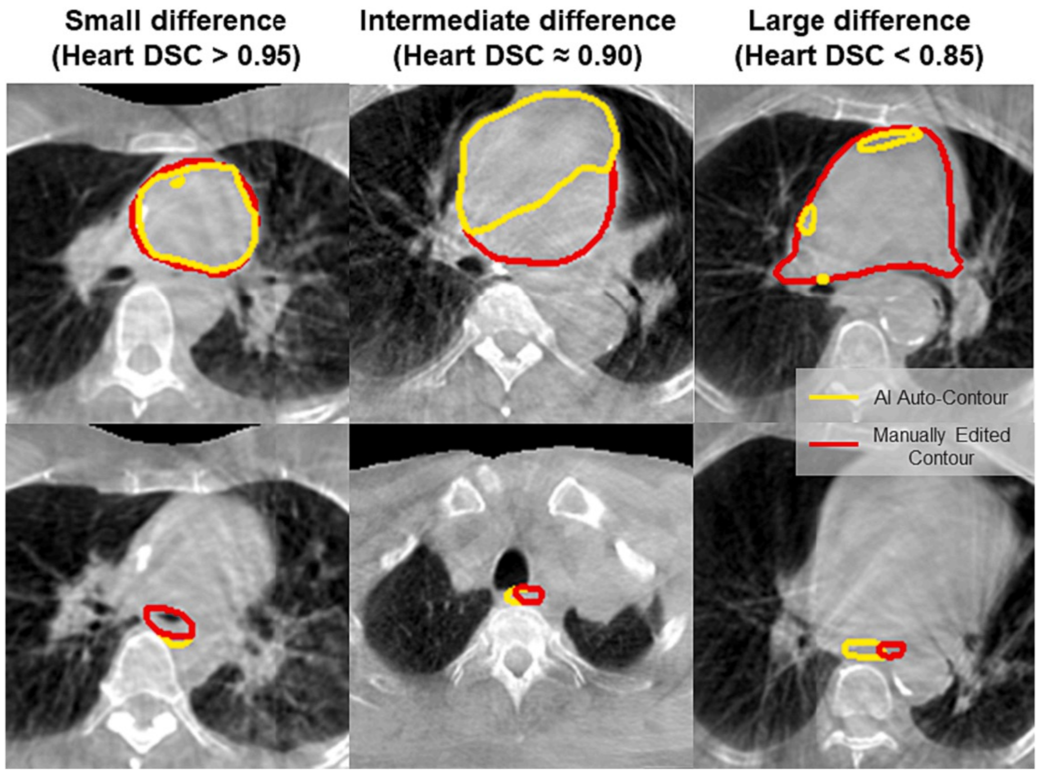
Memorial Sloan Kettering Cancer Center

> Phys Imaging Radiat Oncol. 2024 Feb 1:29:100542. doi: 10.1016/j.phro.2024.100542.
eCollection 2024 Jan.

# Artificial intelligence-based automated segmentation and radiotherapy dose mapping for thoracic normal tissues

Jue Jiang [1], Chloe Min Seo Choi [1] [2], Joseph O Deasy [1], Andreas Rimner [3], Maria Thor [1],
Harini Veeraraghavan [1]



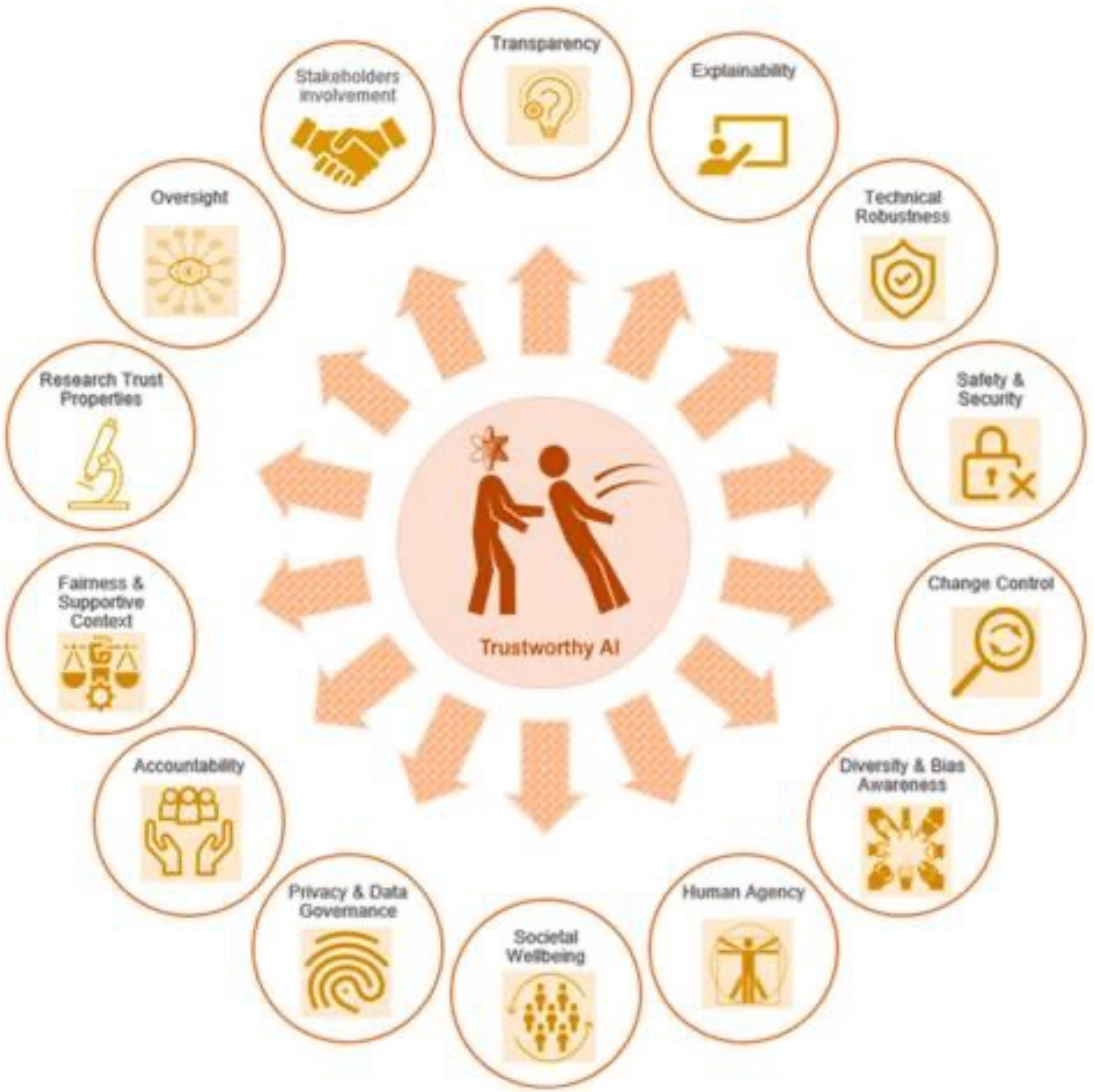

Memorial Sloan Kettering Cancer Center

***Models also need to be trustworthy to ensure Fair and Uninterrupted treatments***

# *Multiple dimensions of trustworthiness*

## Change control plan

- *Prespecified methods to update and evaluate models*
  - *Commissioning testing and testing data used for updates*

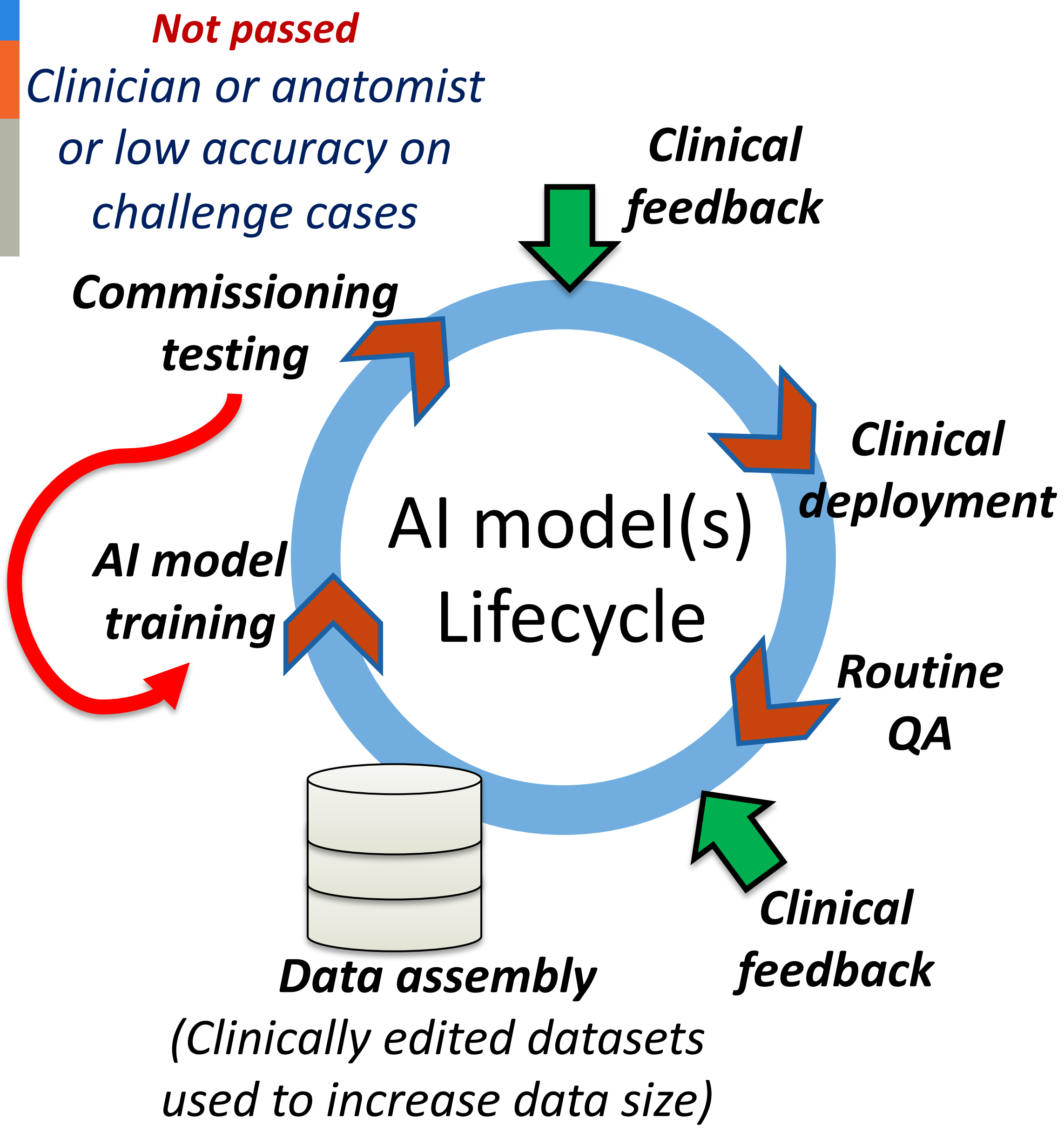



Review > PET Clin. 2022 Jan;17(1):1-12. doi: 10.1016/j.cpet.2021.09.007.

**Trustworthy Artificial Intelligence in Medical Imaging**

Navid Hasani [1], Michael A Morris [2], Arman Rhamim [3], Ronald M Summers [4], Elizabeth Jones [4], Eliot Siegel [5], Babak Saboury [6]

# *Deploying AI models in clinic*

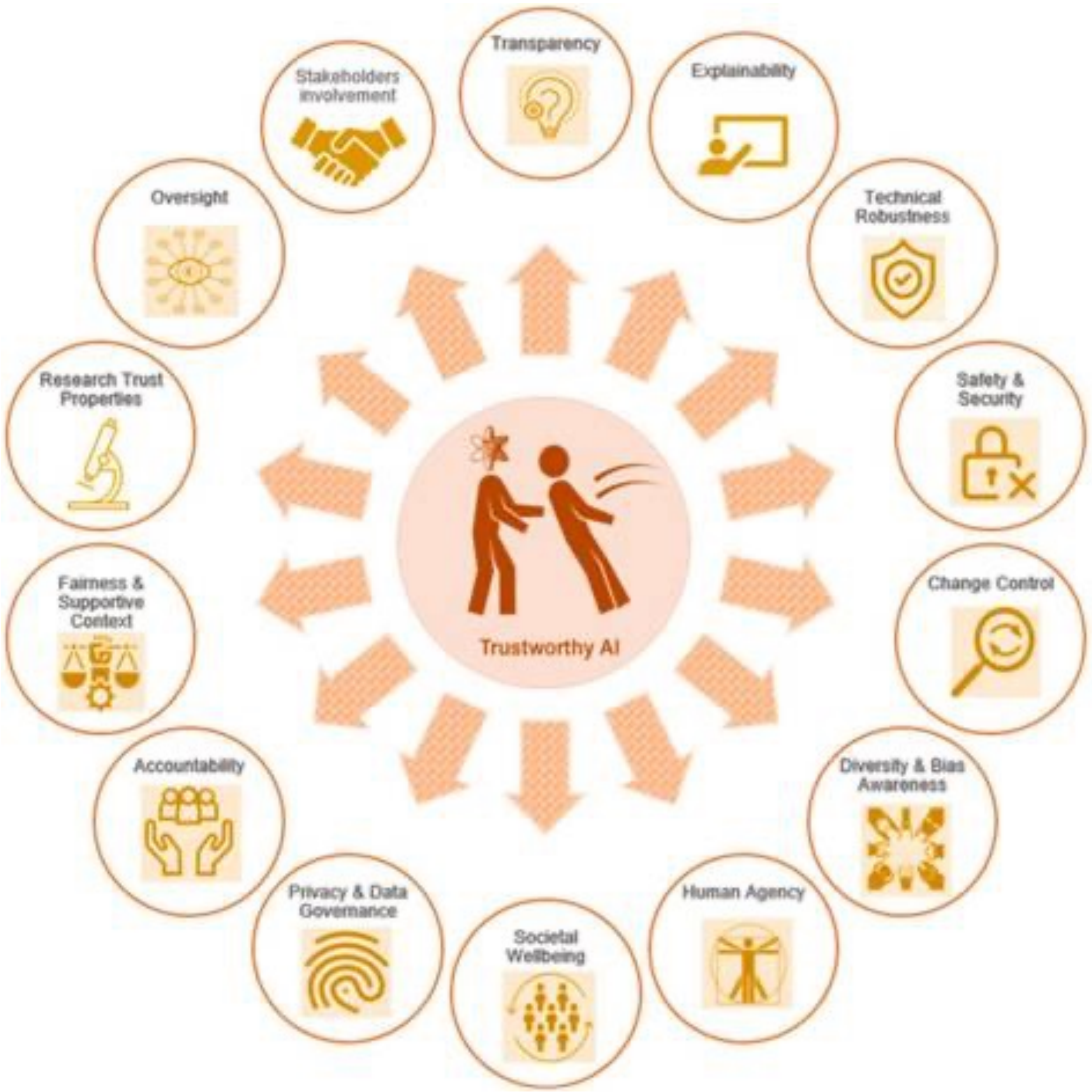



- *Defined metrics for success*
- *Standardized development and testing pipeline*
  - *Commissioning testing sets separate from training & validation + "challenge" cases from prospectively scanned cases*
- *Online Quality assessment*
- *Involve various stakeholders in the development, testing, and commissioning process*
  - *radiation oncologists, physicists, computer scientists, anatomists, radiologists*

# *Multiple dimensions of trustworthiness*

## Robustness and Fairness

- *Resilience to data variations and drifts*
- *Reduced bias to under-represented populations*



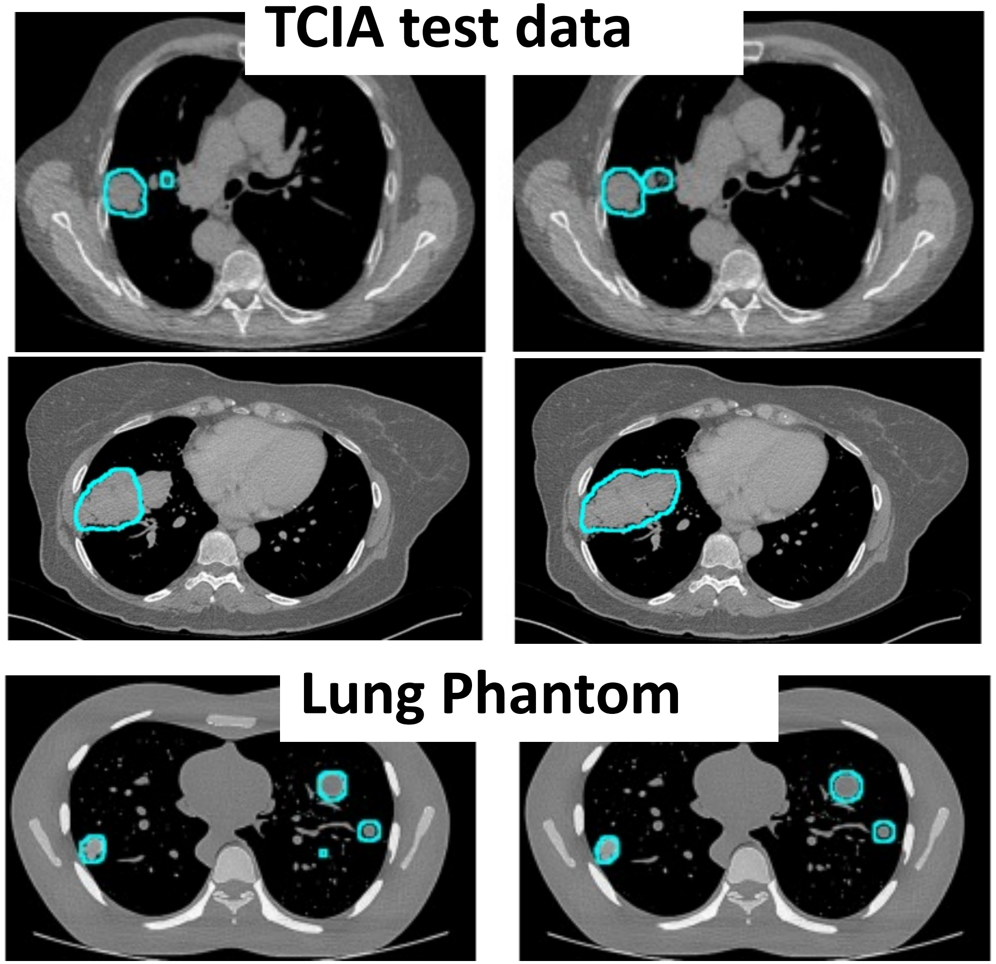

Review > PET Clin. 2022 Jan;17(1):1-12. doi: 10.1016/j.cpet.2021.09.007.

**Trustworthy Artificial Intelligence in Medical Imaging**

Navid Hasani [1], Michael A Morris [2], Arman Rhamim [3], Ronald M Summers [4], Elizabeth Jones [4], Eliot Siegel [5], Babak Saboury [6]

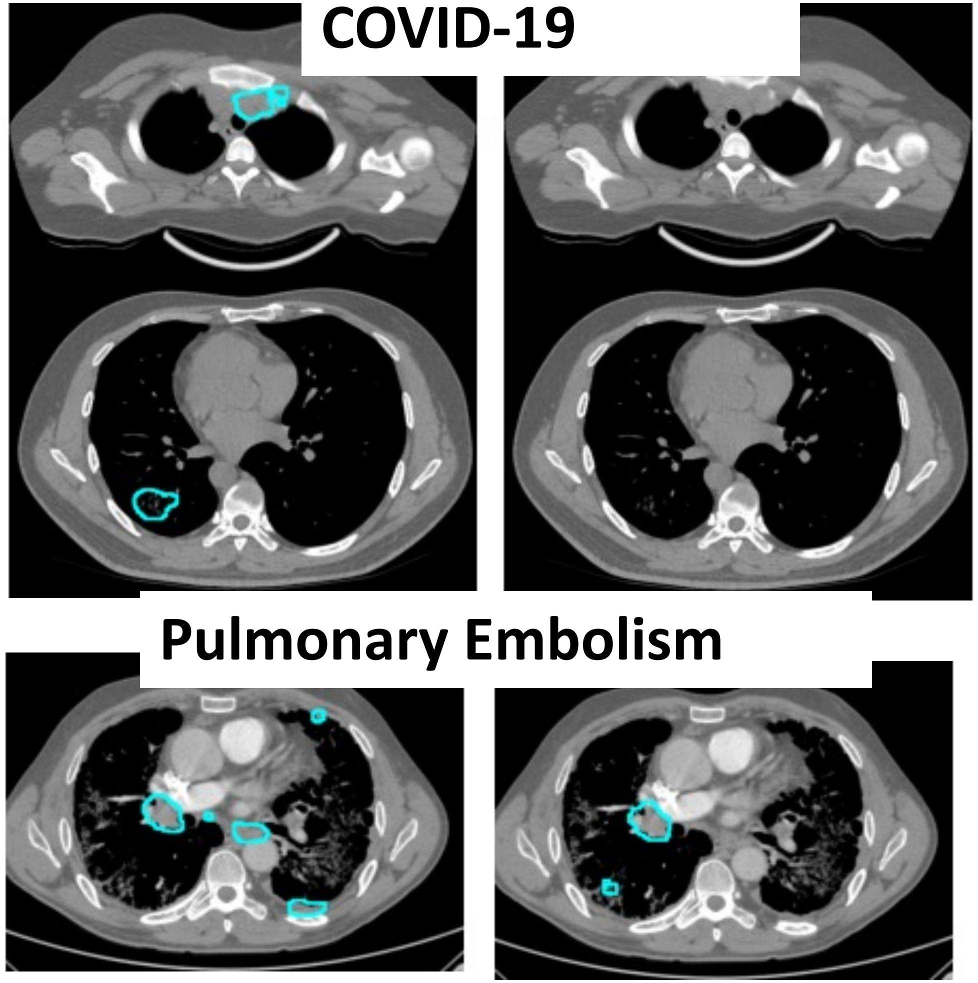
Memorial Sloan Kettering Cancer Center

# *Analyzing robustness of AI model to concept drift*




Swin UNETR SMIT

**In distribution**

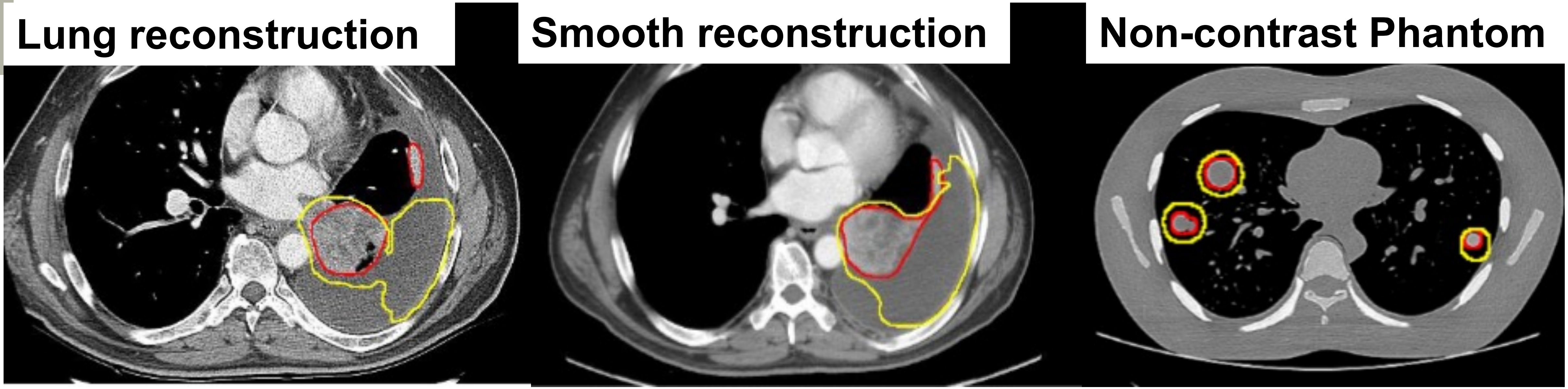



Swin UNETR SMIT

**Out of distribution**

Rangnekar, ... Veeraraghavan In submission.

# *Quantifying performance drifts is essential to make improvements*

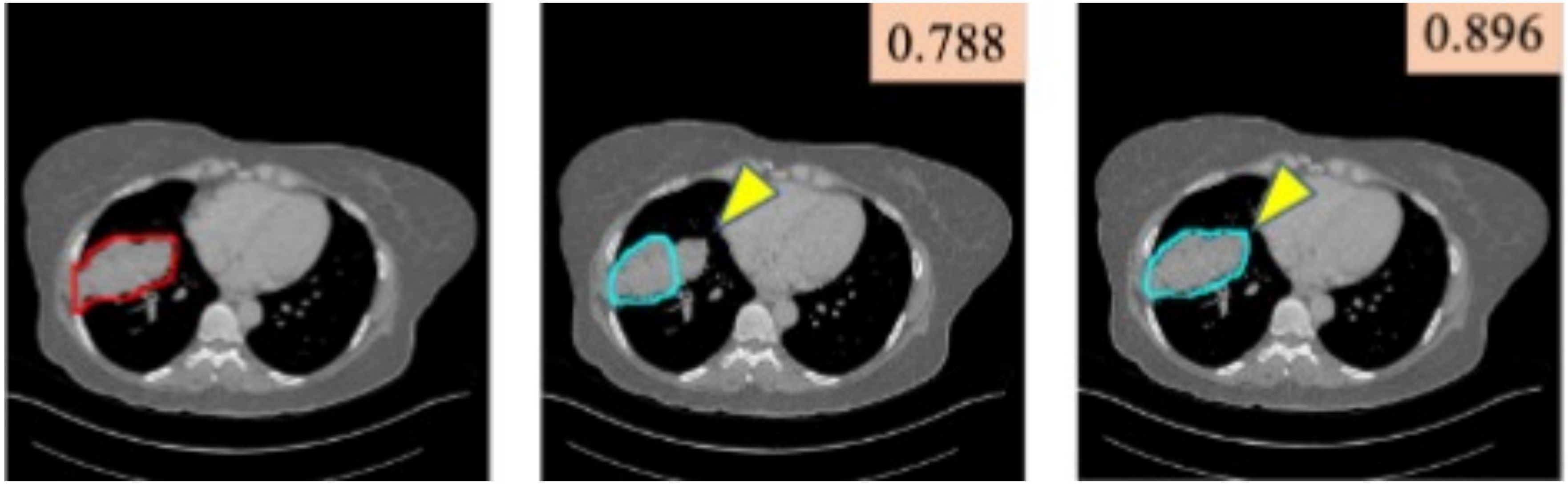



- *Same model can result in performance variations with changes in images even from same patient*
- *Models need to assessed not only for accuracy but also for performance drifts when employed in continuous clinical use*

Jiang, Rangnekar, Veeraraghavan

Memorial Sloan Kettering Cancer Center

# *Standard metrics may be insufficient to detect drifts*

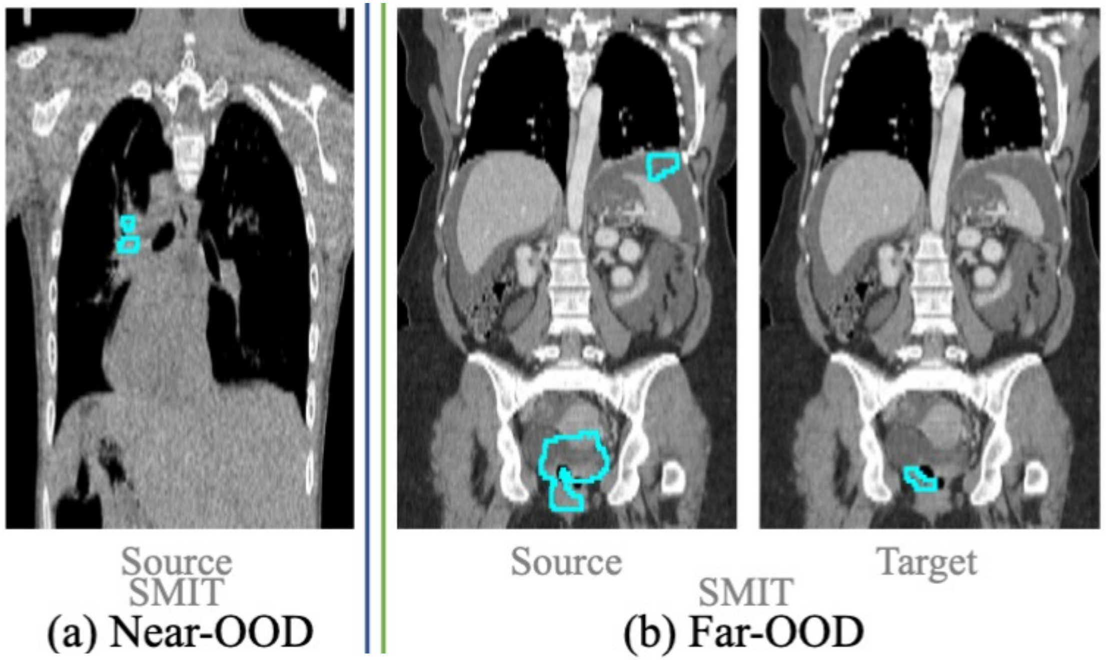



| Model | DSC (↑) | RVD (↓) | Pr (↑) | Rc (↑) |
|---|---|---|---|---|
| Swin UNETR | 0.783 ± 0.091 | 0.175 ± 0.329 | 0.035 | 0.578 |
| SMIT | 0.798 ± 0.075 | 0.157 ± 0.281 | 0.131 | 0.635 |

*Accuracy metrics show similar performance of two different networks*

Rangnekar, … Veeraraghavan https://arxiv.org/abs/2403.13113

# *Accuracy metrics to identify performance drift*




COVID-19 Chest CT

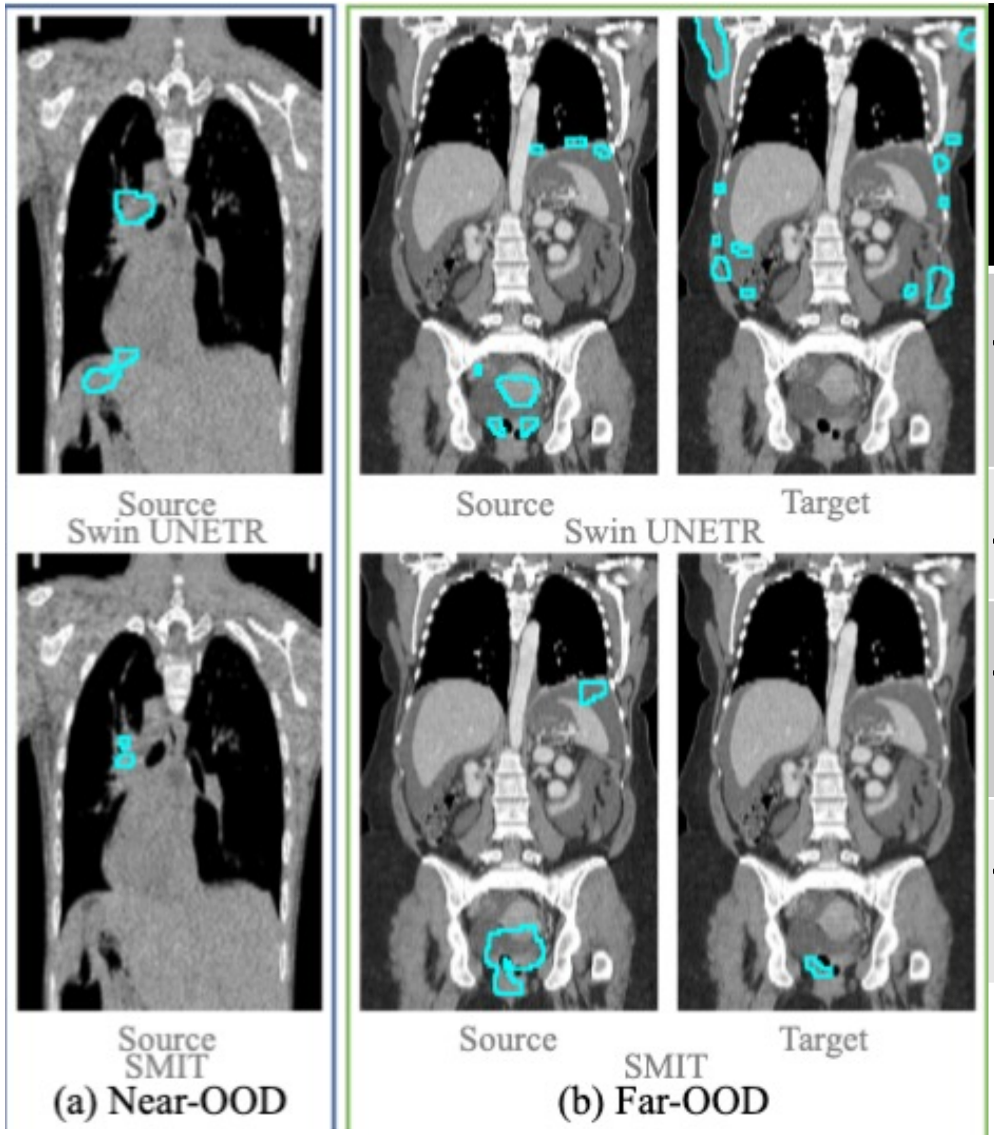
Abdomen CT

| Data | AUROC | FPR @ 95 |
|---|---|---|
| COVID-19 lung CT | 89.85 % | 34.62 |
| Abdomen CT | 99.02 % | 5.77 |

- *AUROC measures the accuracy of correctly detecting the lung tumors when they occur and not detecting other lesions as lung tumors*
- *FPR @ 95 or False positive rate at 95% measures the probability that a negative (out-of-distribution) example or non-lung tumor is segmented as positive (or in-distribution) lung tumor with a True Positive Rate as high as 95%*

**Rangnekar.. Veeraraghavan 2024, https://arxiv.org/abs/2403.13113**

# *Differences emerge when analyzing OOD performance*




COVID-19 Chest CT

Abdomen CT

| Model | Dataset | AUROC | FPR @ 95 |
|---|---|---|---|
| Swin UNETR | COVID-19 | 89.19 % | 34.62 |
| SMIT | COVID-19 | 89.85 % | 34.62 |
| Swin UNETR | Abdomen CT | 82.39 % | 38.46 |
| SMIT | Abdomen CT | 99.02 % | 5.77 |

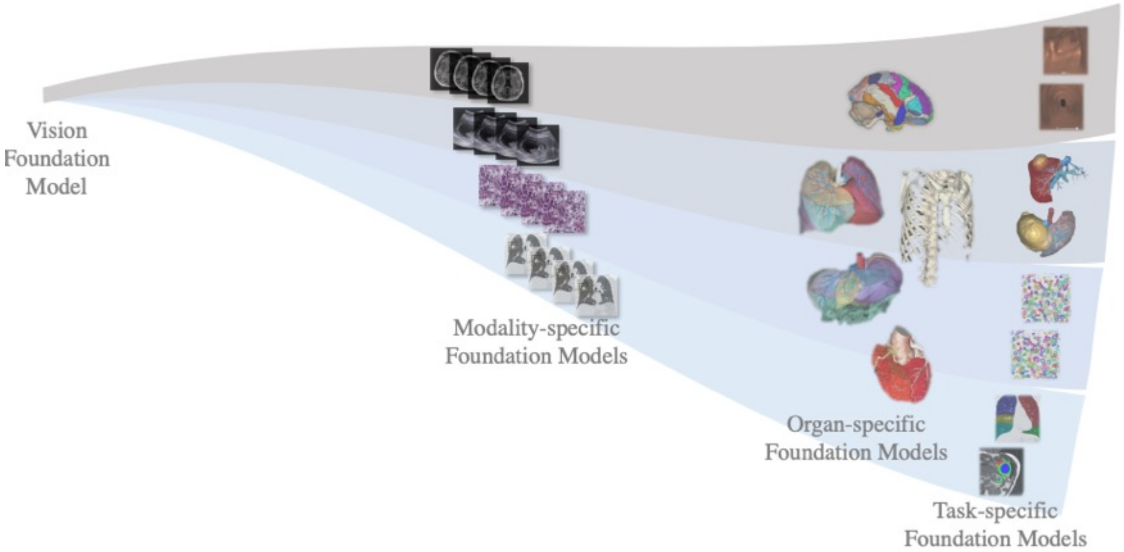
# *Increasing robustness to image variations*

## On the Challenges and Perspectives of Foundation Models for Medical Image Analysis

A Preprint

**Shaoting Zhang**
Shanghai AI Laboratory
Shanghai, China
zhangshaoting@pjlab.org.cn

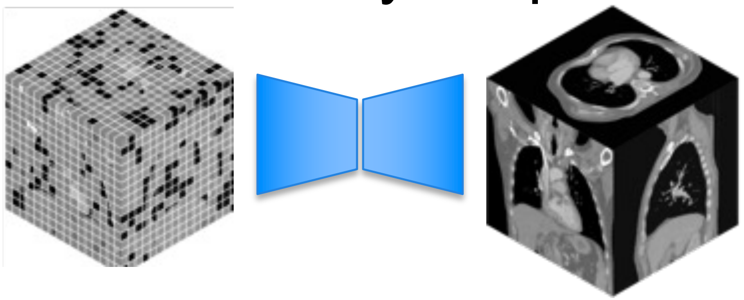
**Dimitris Metaxas**
Rutgers University
New Brunswick, NJ, USA
dnm@cs.rutgers.edu



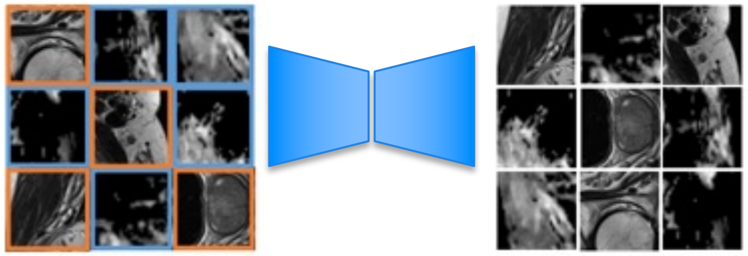

Medical foundation models or **large pretrained models** could allow to solve a wide range of tasks by accelerating the development of accurate models, while reducing need for large amounts of labeled training data

# *Self-supervised learning background*

- Machine learning approach that uses unsupervised learning for tasks that conventionally require supervised learning
  - Instead of relying on supervised labels, generate implicit labels from unstructured data
  - Can be more time efficient and effective for massive datasets
  - Tasks are designed such that loss function can use unlabeled data as “ground truth” to extract meaningful representations
  - Pretext tasks yield pseudo labels



Predict masked portions of an image
Jiang et.al MICCAI 2022



Jigsaw puzzle recovery of images
Taleb et.al TMI 2020

Memorial Sloan Kettering Cancer Center

# Foundation models approach

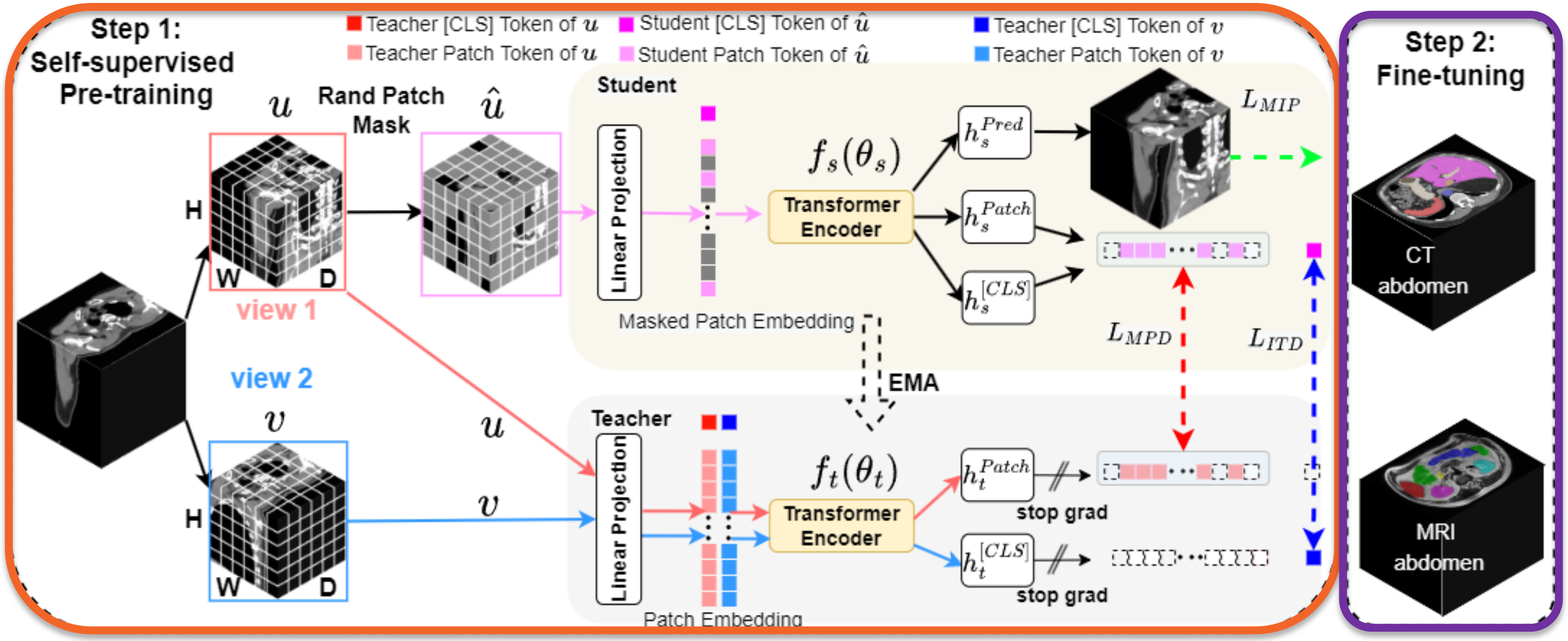
Self-supervised 3D Anatomy Segmentation Using Self-distilled Masked Image Transformer (SMIT)

Jue Jiang, Neelam Tyagi, Kathryn Tringale, Christopher Crane & Harini Veeraraghavan

Conference paper | First Online: 16 September 2022

1287 Accesses

Part of the Lecture Notes in Computer Science book series (LNCS,volume 13434)



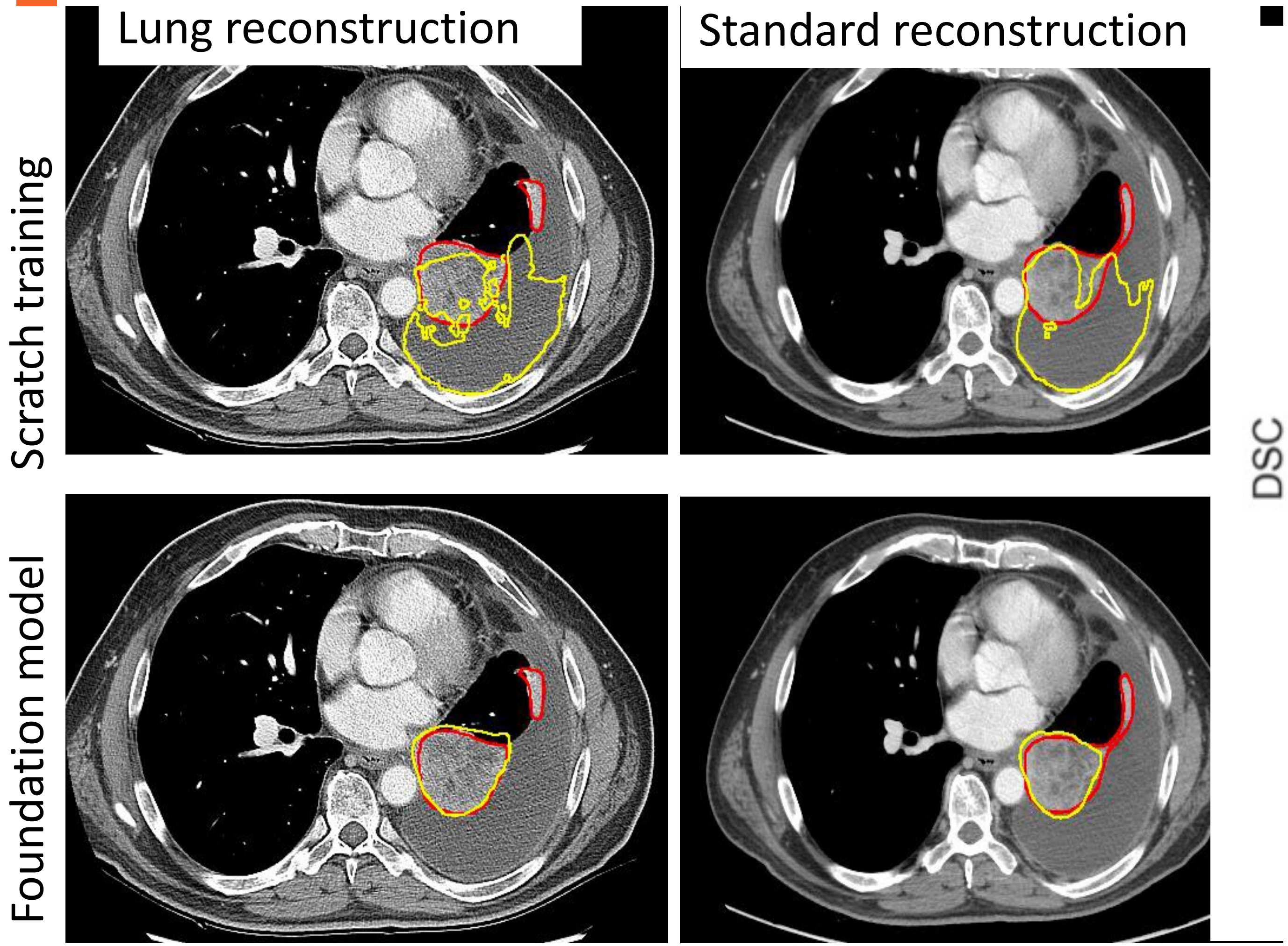

1) *Self-supervised pre-training with uncurated 3D CTs from The Cancer Imaging Archive and Internal datasets*

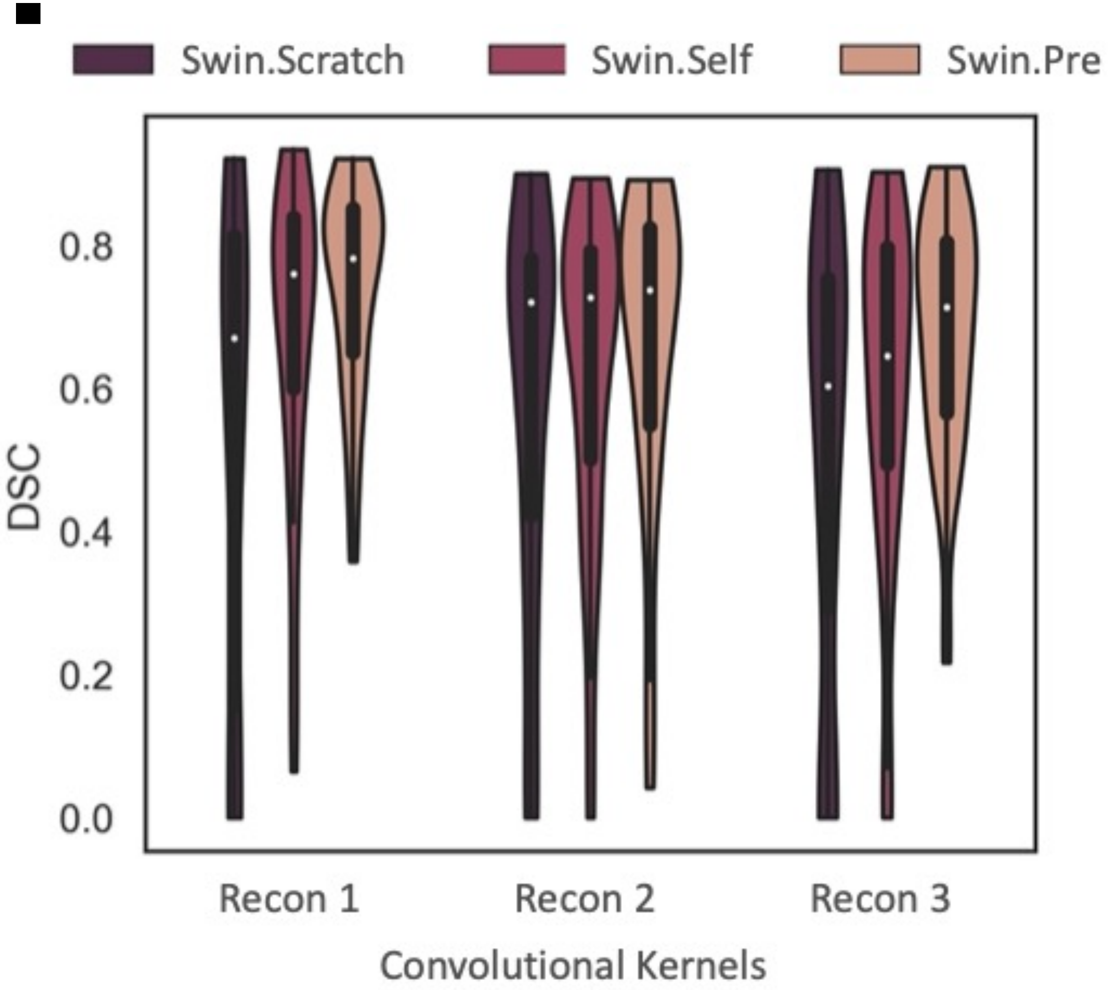
➢ *Current model uses ~ 10,000 3D CTs*

2) *Fine tune/transfer learn on task specific datasets*

Memorial Sloan Kettering Cancer Center

# *Foundation models improve robustness to CT variations*



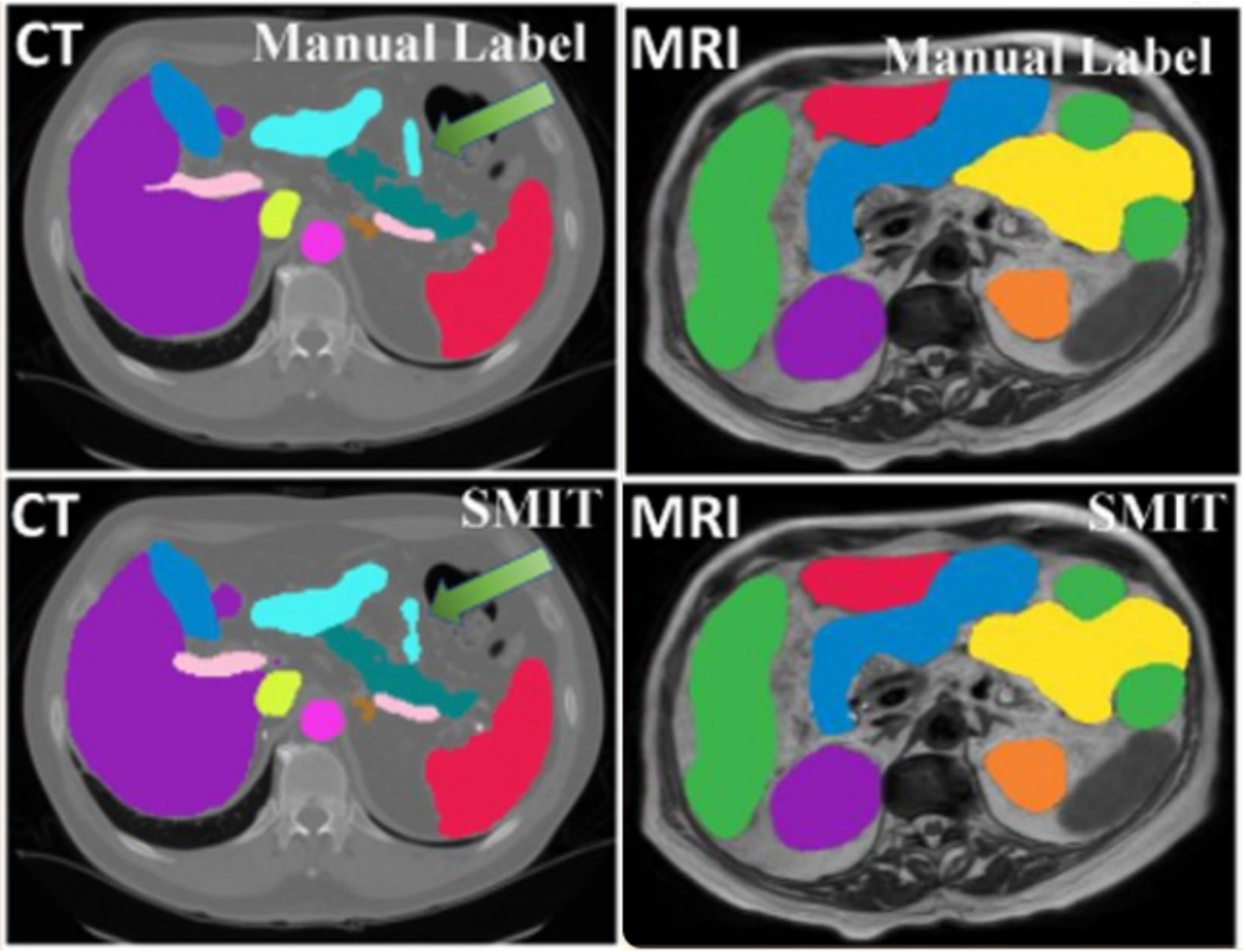



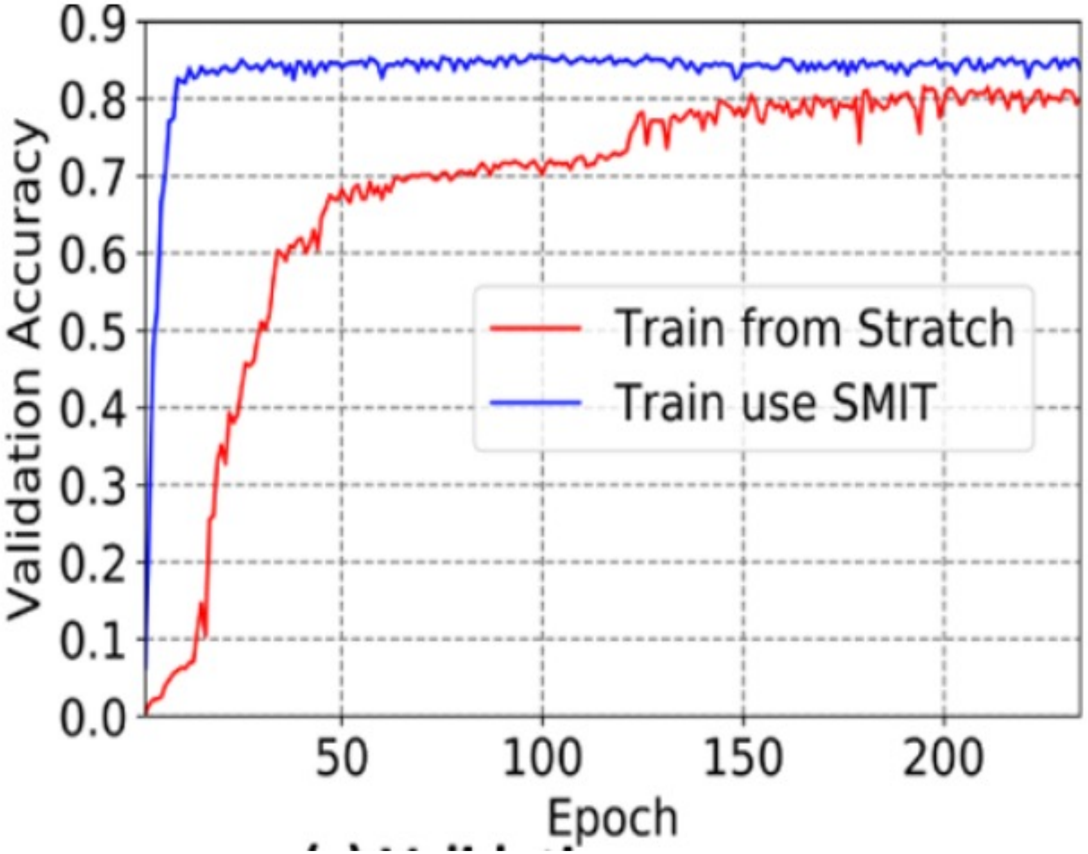

Jiang and Veeraraghavan MIDL 2024

Memorial Sloan Kettering Cancer Center

# SMIT is transferable to multiple imaging modalities

- *Fine-tuning SSL pretrained model is more accurate than supervised learning alone*



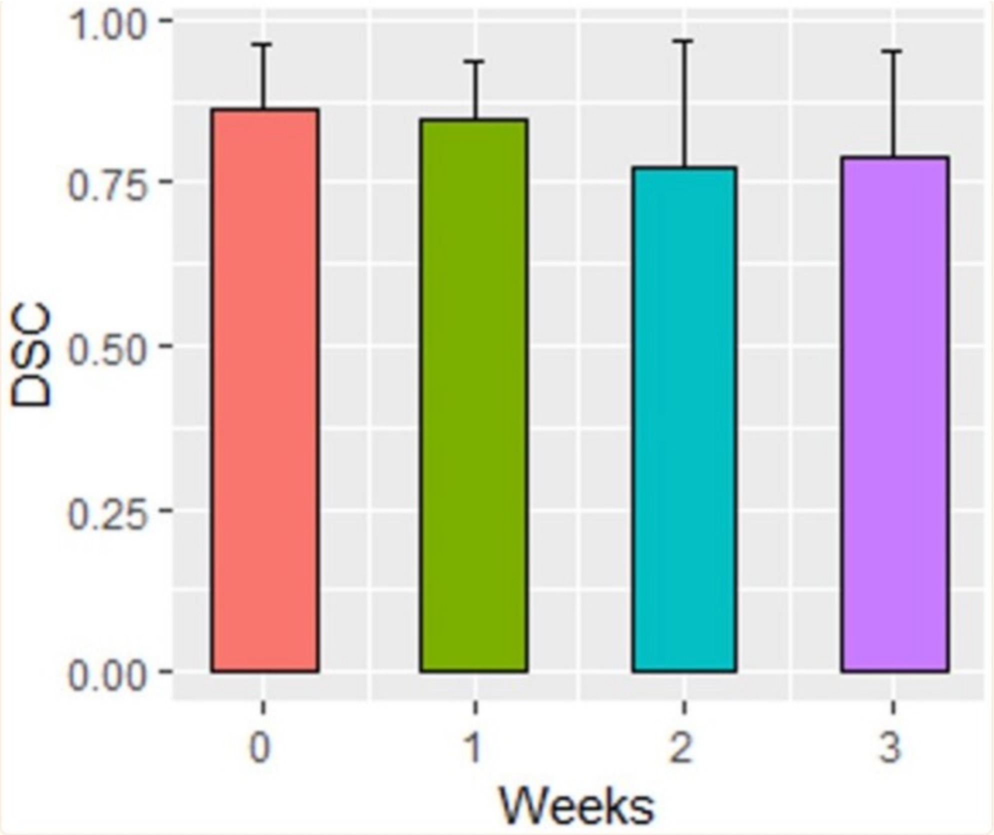

**Code + Model: https://github.com/harveerar/SMIT**

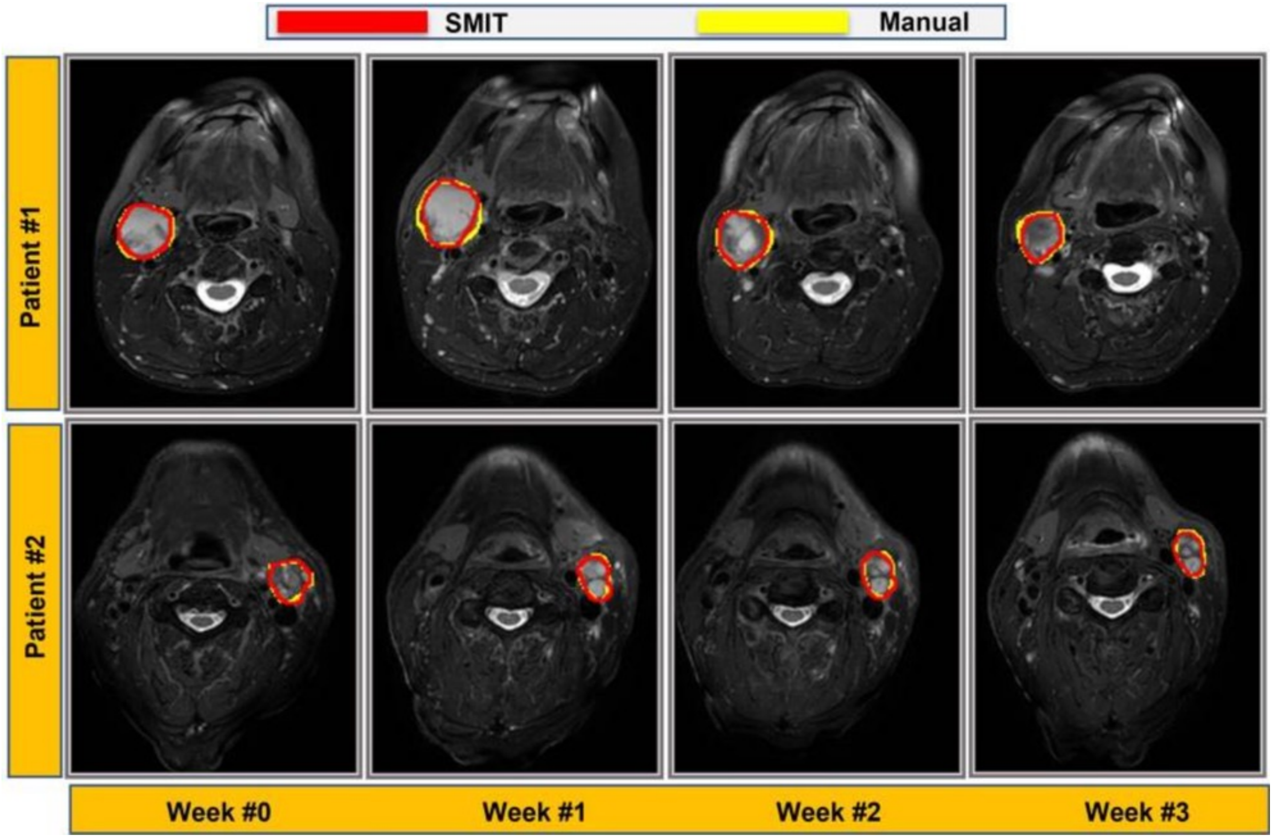
Memorial Sloan Kettering Cancer Center

**Jiang, … Veeraraghavan et.al MICCAI 2022**

# Auto-segmentation of neck nodal metastases using self-distilled masked image transformer on longitudinal MR images

Ramesh Paudyal, PhD, Jue Jiang, PhD, James Han, MD, Bill H Diplas, MD, Nadeem Riaz, MD, Vaios Hatzoglou, MD, Nancy Lee, MD, Joseph O Deasy, PhD, Harini Veeraraghavan, PhD ✉, Amita Shukla-Dave, PhD ✉



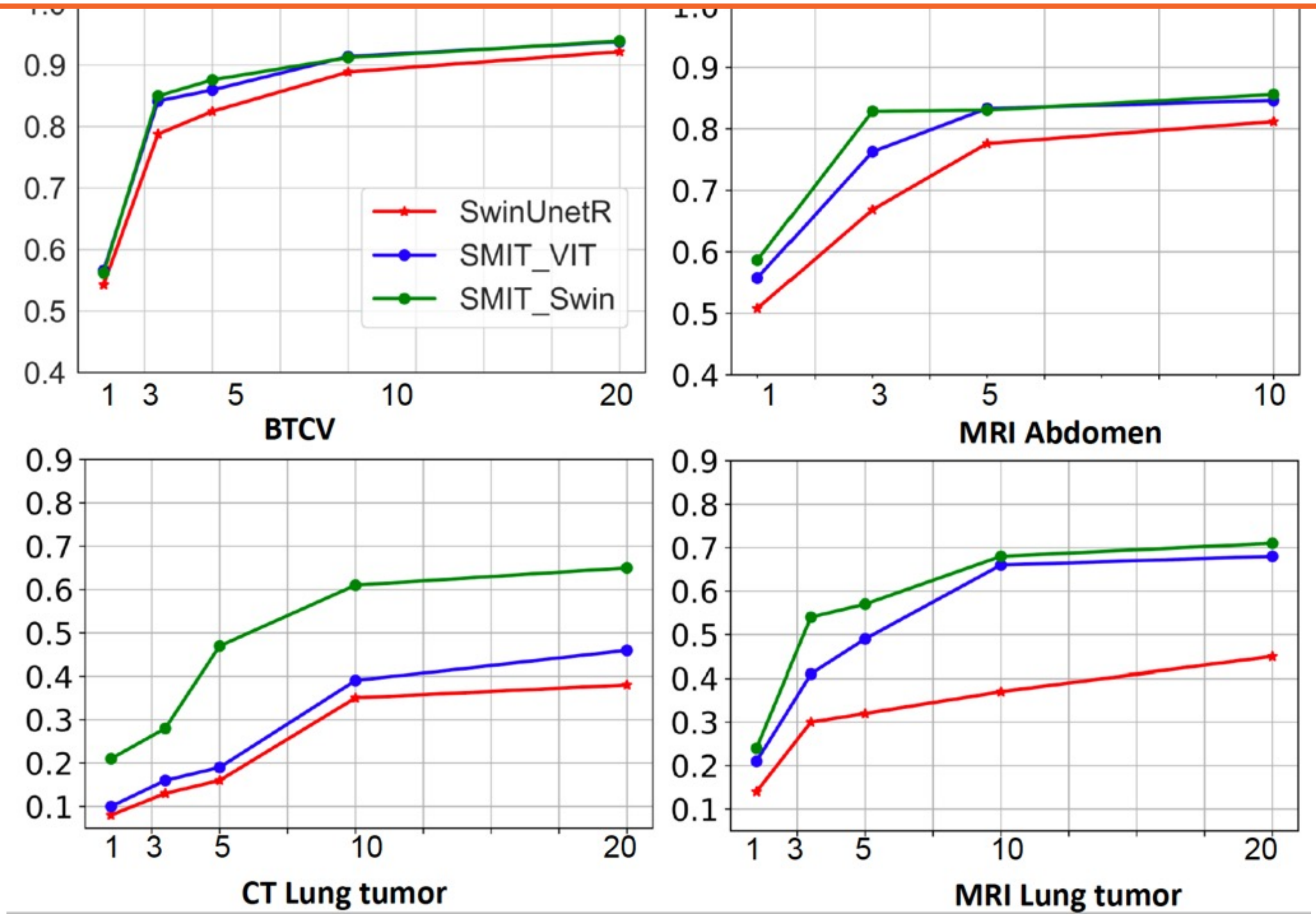

*SMIT shows capability to track nodal metastases changes from MRI*

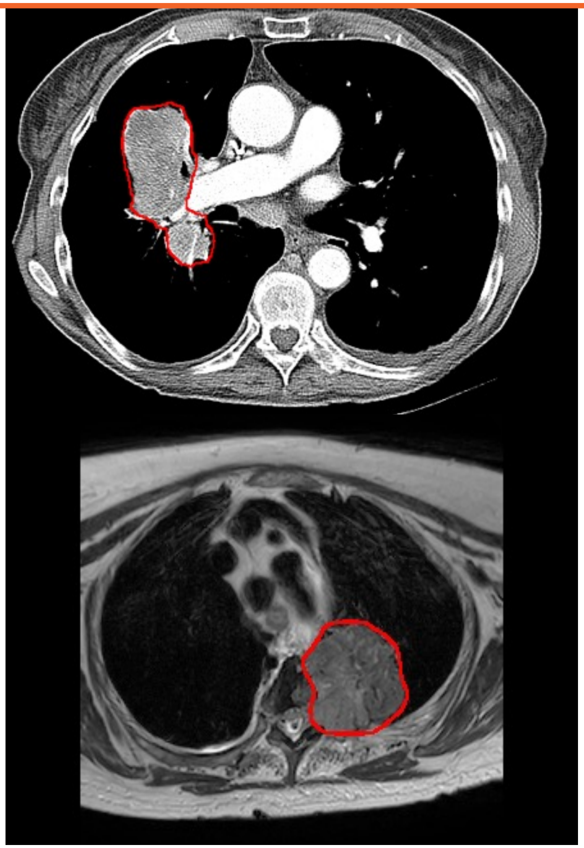
Memorial Sloan Kettering Cancer Center

# *SMIT can be applied in few-shot settings*

- *Models show capability for few shot training*

**Pretrained model using 10,000 3D CT from TCIA and institutional cases from cancer and non-cancer images**



Jiang, … Veeraraghavan in submission

Memorial Sloan Kettering Cancer Center
1884

# *Image Foundation models used in MSK AI*

- "One model to segment them all" approach simplifies clinical implementation and maintenance
- Clinical models at MSK using foundation models used for thorax, abdomen/pelvis, and brain

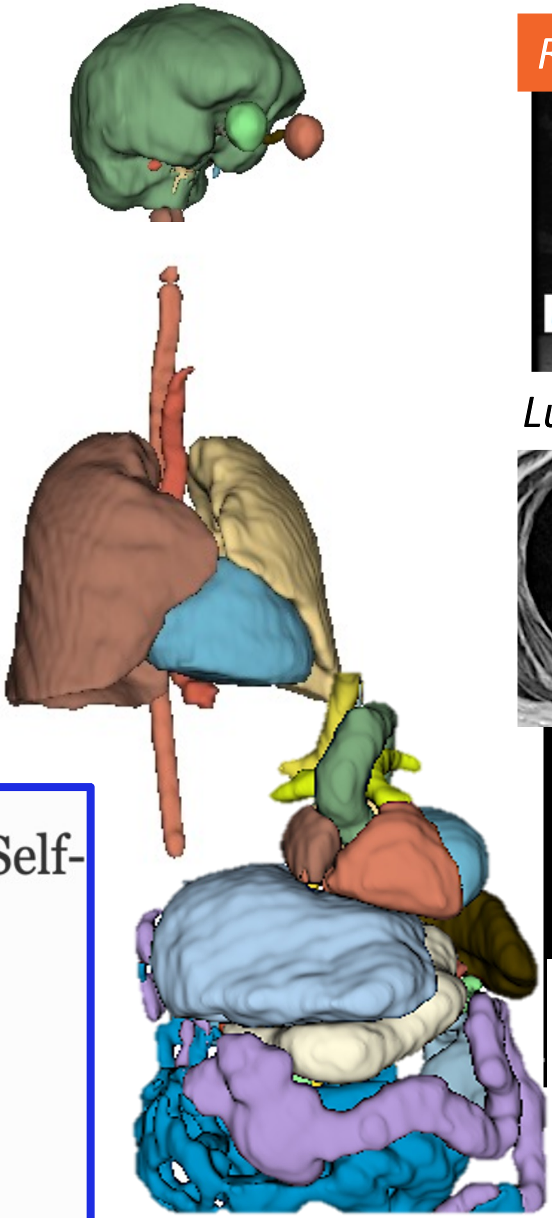
Self-supervised 3D Anatomy Segmentation Using Self-distilled Masked Image Transformer (SMIT)

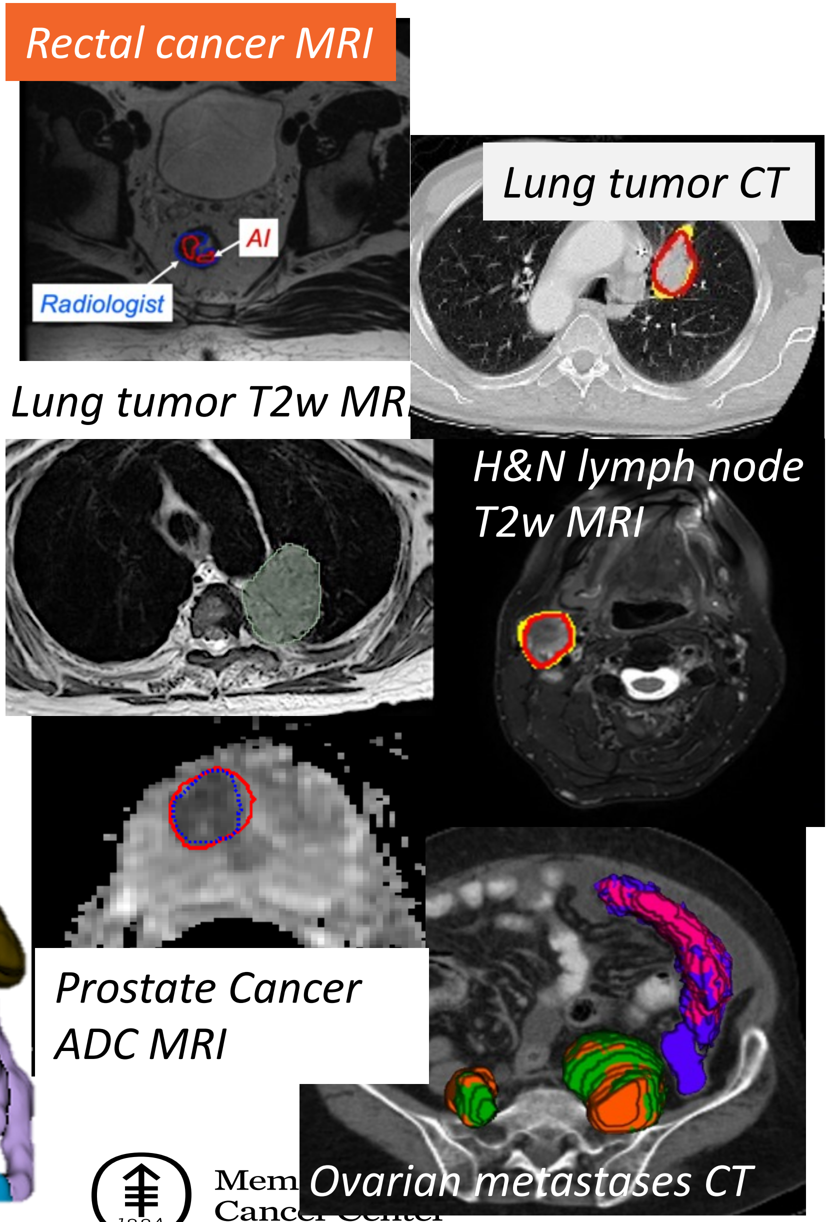
Jue Jiang, Neelam Tyagi, Kathryn Tringale, Christopher Crane & Harini Veeraraghavan

Conference paper | First Online: 16 September 2022

1287 Accesses

Part of the Lecture Notes in Computer Science book series (LNCS,volume 13434)





# pyCERR library of AI segmentation models

| Image modality | Site | Network architecture | Segmented Structures | Jupyter notebook |
|---|---|---|---|---|
| Planning CT Axial | Lung | DeepLab (V3+) https://doi.org/10.1016/j.phro.2020.05.009 | Aorta, SVC, IVC, PA, LA, LV, RA, RV, atria, ventricles, and pericardium | https://github.com/cerr/pyCERR-Notebooks/blob/main/autosegment_CT_Heart_OARs.ipynb |
| Planning CT Axial | Lung | Incremental MRRN doi.org/10.48550/arXiv.2005.13690 | Right lung, left lung, heart, esophagus, cord, PBT | https://github.com/cerr/pyCERR-Notebooks/blob/main/autosegment_CT_Lung_OARs.ipynb |
| Planning T2w MR Axial | Prostate | DeepLab https://doi.org/10.1016/j.phro.2019.11.006 | CTV, bladder, penile bulb, rectum, urethra foley, rectal spacer, large bowel | https://github.com/cerr/pyCERR-Notebooks/blob/main/autosegment_MR_Prostate_OARs.ipynb |
| Planning CT Axial | Lung | SMIT https://doi.org/10.48550/arXiv.2205.10342 | Lung tumor | https://github.com/cerr/pyCERR-Notebooks/blob/main/autosegment_CT_Lung_SMIT.ipynb |
| Planning T2w MR Axial (longitudinal) | Pancreas | ProRSeg https://doi.org/10.1002/mp.16527 | Liver, large bowel , small bowel, duo-stomach<br><br>Deformable vector field (first to last scan) | https://github.com/cerr/pyCERR-Notebooks/blob/main/auto_register_segment_MR_Pancreas_OARs.ipynb |
| Planning CT Axial | Head and neck | Nested Block Self Attention https://doi.org/10.48550/arXiv.1909.05054 DeepLab (V3+) https://doi.org/10.1088/1361-6560/ac4000 | Left parotid, right parotid, left submandibular gland, right submandibular gland, mandible, spinal cord, brain stem, oral cavity | |

Memorial Sloan Kettering Cancer Center

# AI GUIDED RT (AIGRT) IS ~~THE FUTURE~~ HERE

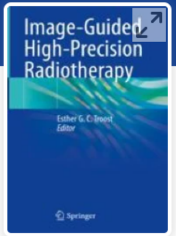
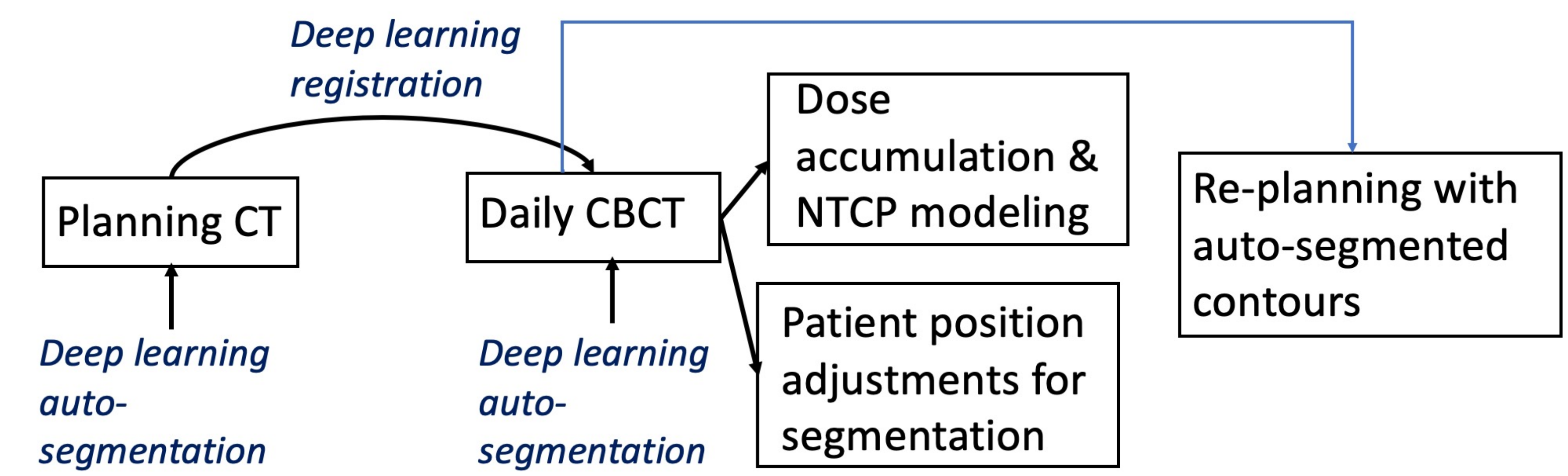



Image-Guided High-Precision Radiotherapy pp 249–267 | Cite as

## Artificial Intelligence in Radiation Oncology: A Rapidly Evolving Picture

Harini Veeraraghavan & Joseph O. Deasy

Chapter
178 Ac




AI guided radiation treatment (AIGRT)