Approaches to Automated Musculoskeletal Segmentation

Fundamentals of Musculoskeletal MRI II

Anthony A. Gatti, Stanford University

May 12th, 2025

Wu Tsai Human Performance Alliance | Stanford University

Declaration of Financial Interests or Relationships

Speaker Name: Anthony Gatti

I have the following financial interest(s) or relationship(s) to disclose with regard to the subject matter of this presentation:

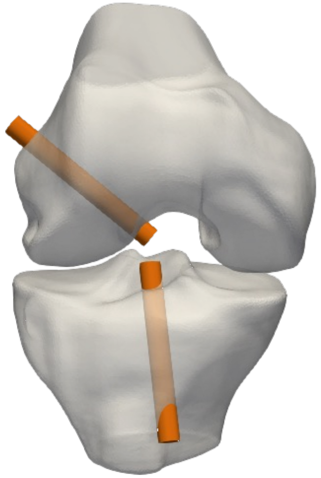
- Shareholder:
  - NeuralSeg
  - NodeAI
  - GeminiOV

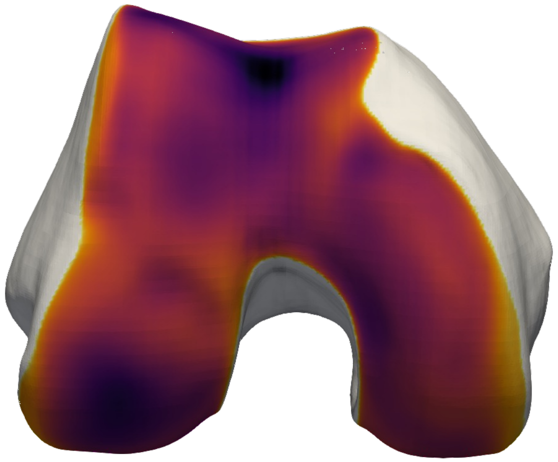
What to expect

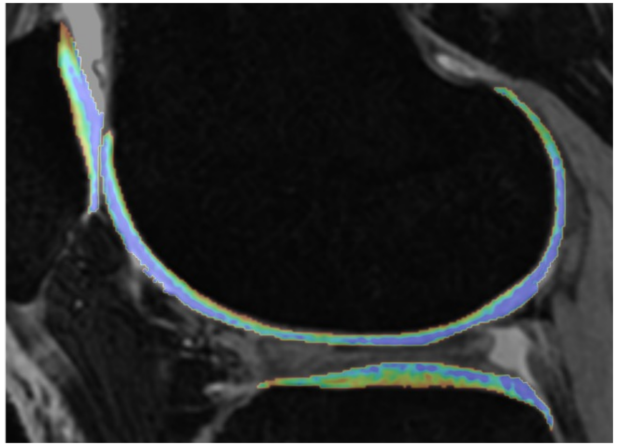
- Why automated segmentation??
- What auto MSK segmentation has enabled
- Evaluating auto segmentation
- Segmentation model overview
- Recommendations

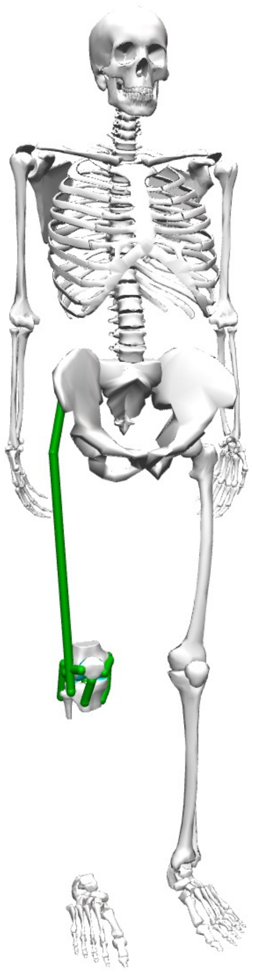
Why Automate Segmentation in MSK MRI?

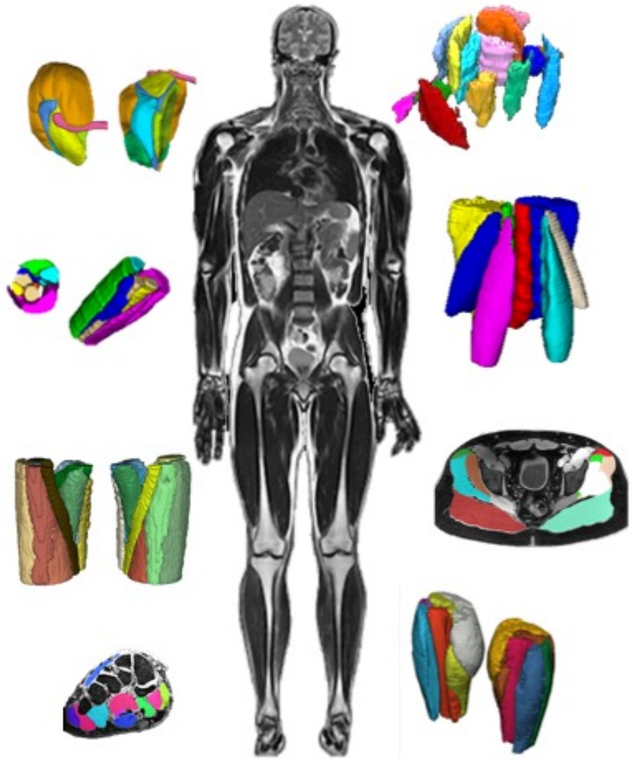
- Quantitative outcomes
  - Shape models
  - Compositional/quantitative MRI
- Digital Models
  - Surgical planning
  - Biomechanical simulations





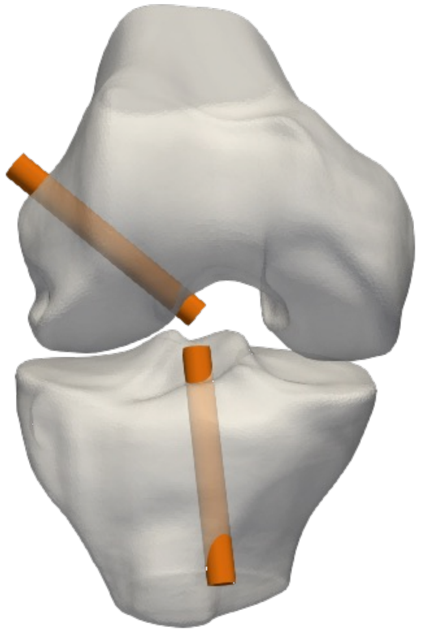




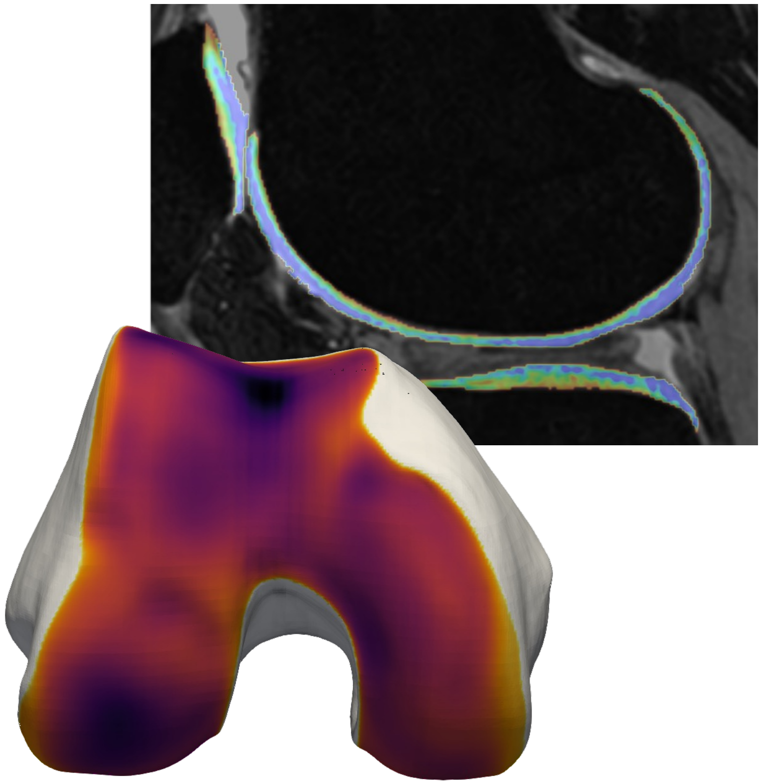


MSK Segmentation Beyond Research?

Surgical Planning



Clinical Trials



Improve subjective measures

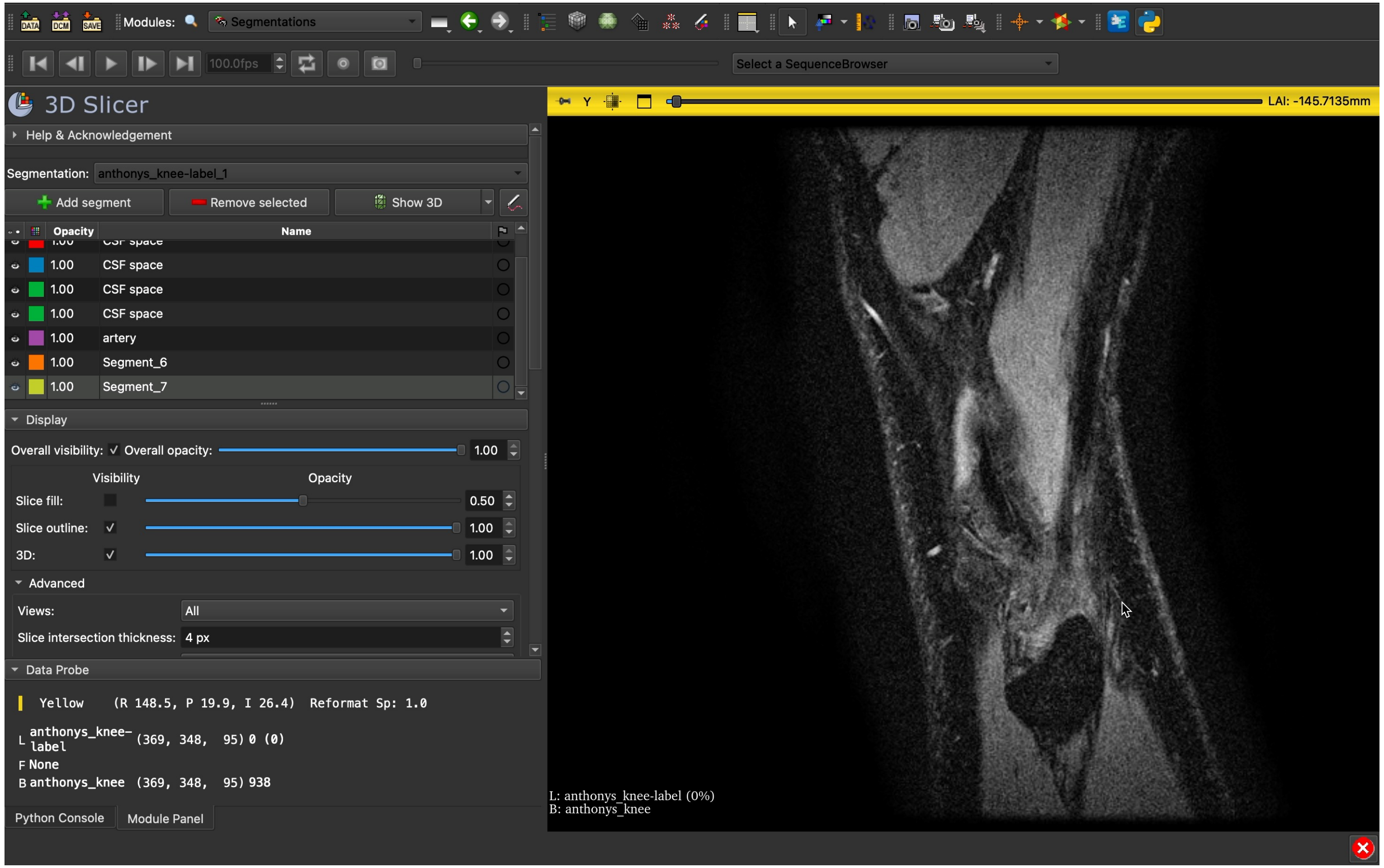
- Muscular dystrophy
- Arthritis
    - Osteo
    - Rheumatoid
- Etc.

The Bottleneck: Manual Segmentation

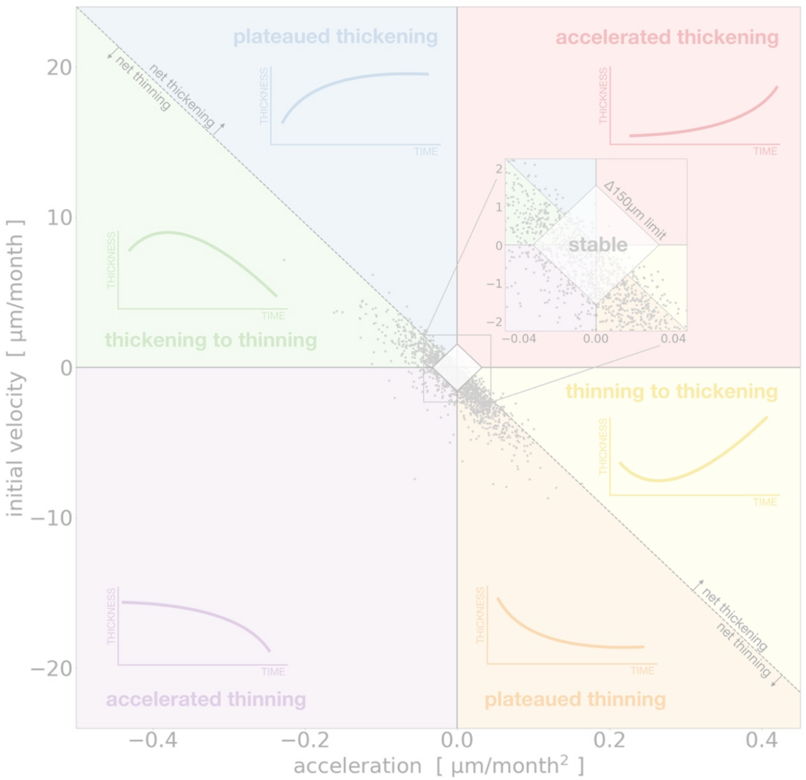
Patella

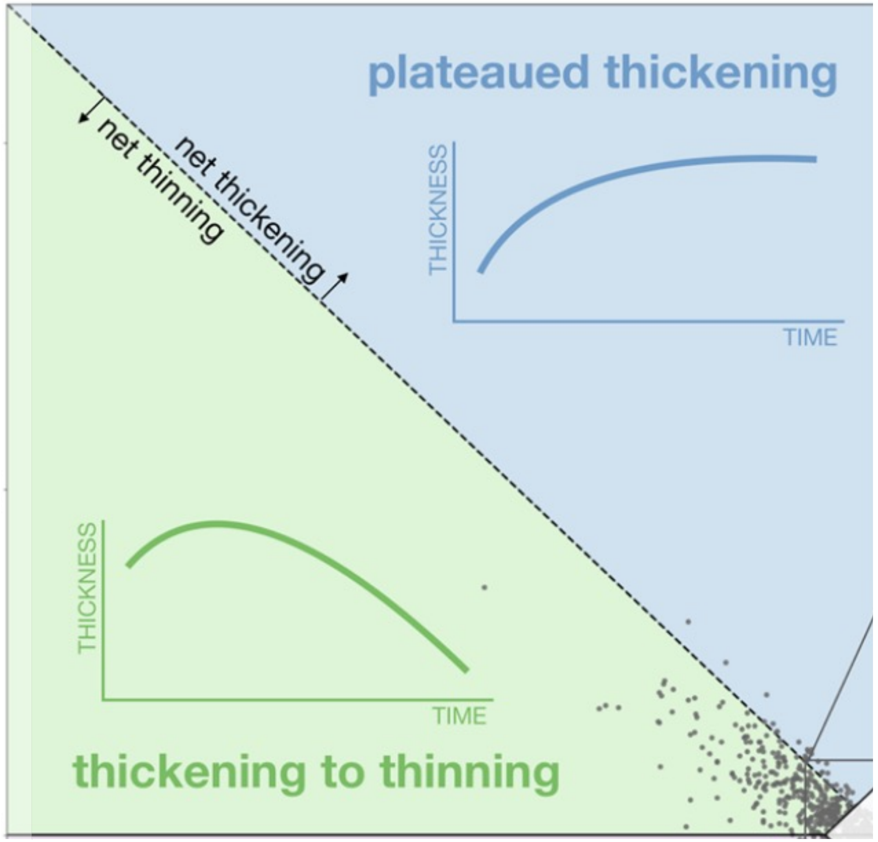
Femur

Tibia



What has Automated MSK Segmentation Enabled?

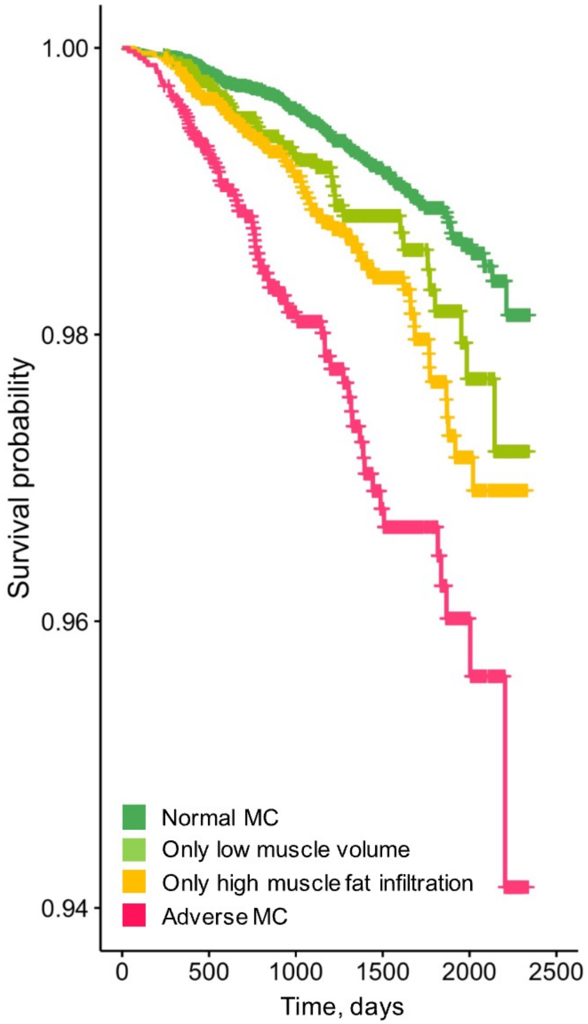




~4,000 knees
8-years

Iriondo et al. JOR, 2020

What has Automated MSK Segmentation Enabled?



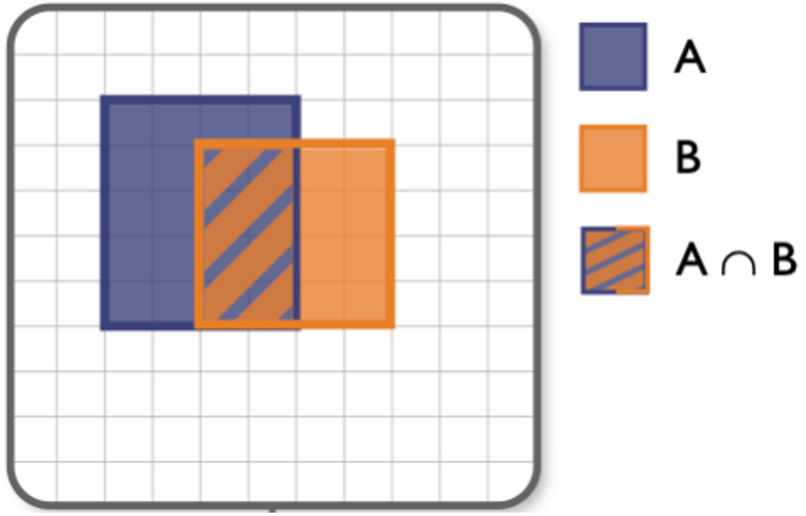
Linge et al. JCSM, 2021

~40,000 people
Green = Healthy muscle
Red = Unhealthy muscle

Poor muscle health leads to higher probability of death

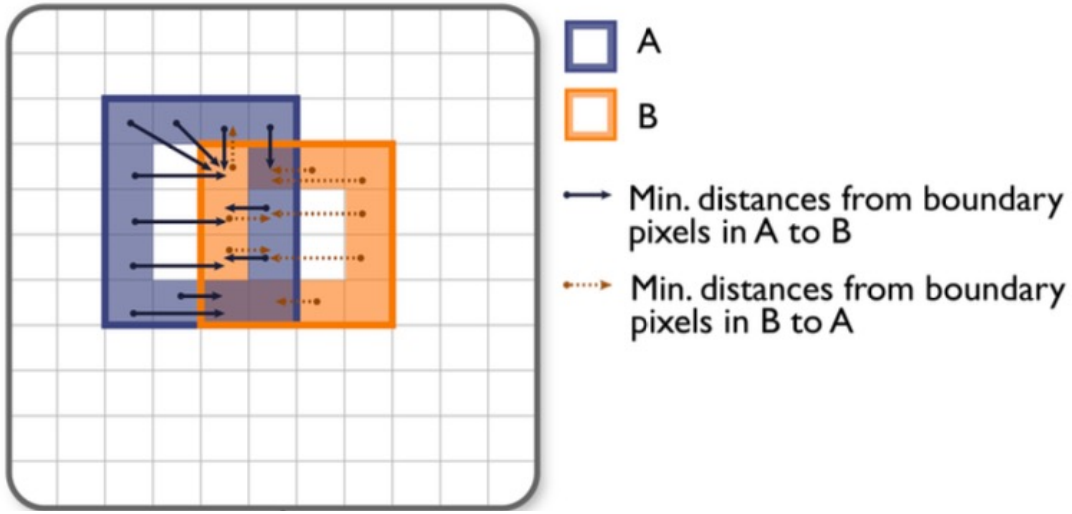
How Do We Evaluate?

Dice Similarity Coefficient (DSC)



small objects

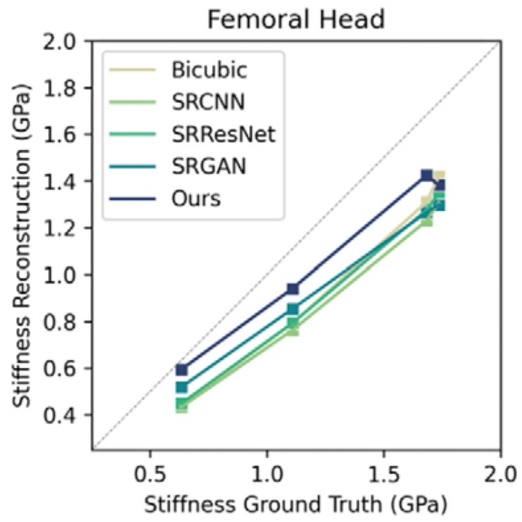
Average Symmetric Surface Distance (ASSD)



big objects

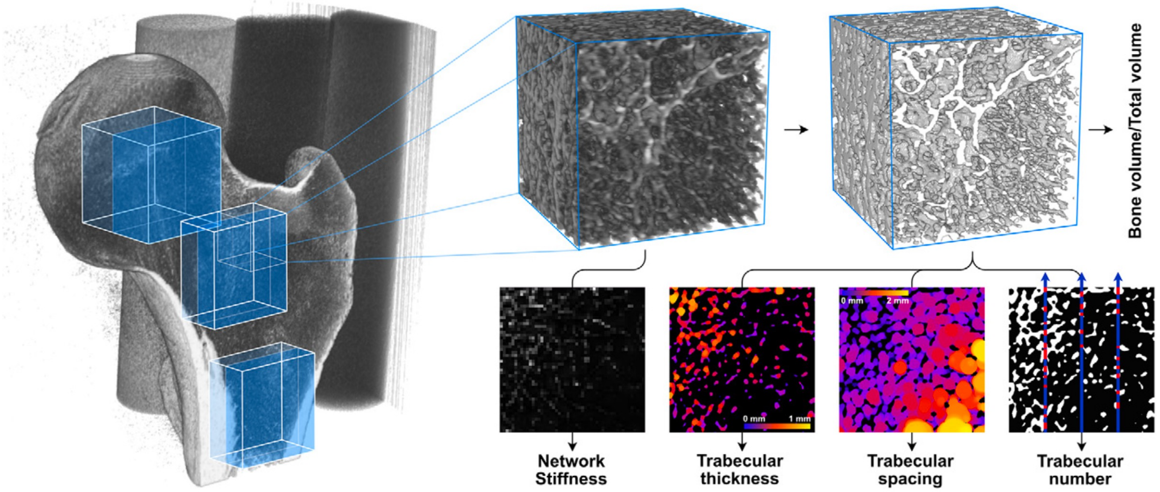
Reinke et al. arXiv, 2021.
Common Limitations of Image Processing Metrics: A Picture Story

How Do We Evaluate? Beyond Dice



Cartilage Thickness Reliability

	corFLASH			
	Manual		U-Net	
	SEM	SDC	SEM	SDC
Cartilage thickness (mm)				
MFTC	0.09	0.24	0.07	0.21
LFTC	0.10	0.27	0.13	0.36



Wirth et al. MAGMA, 2020
Eckstein et al. Quant Imaging Med Surg, 2024

Chan et al. Rad:AI, 2023

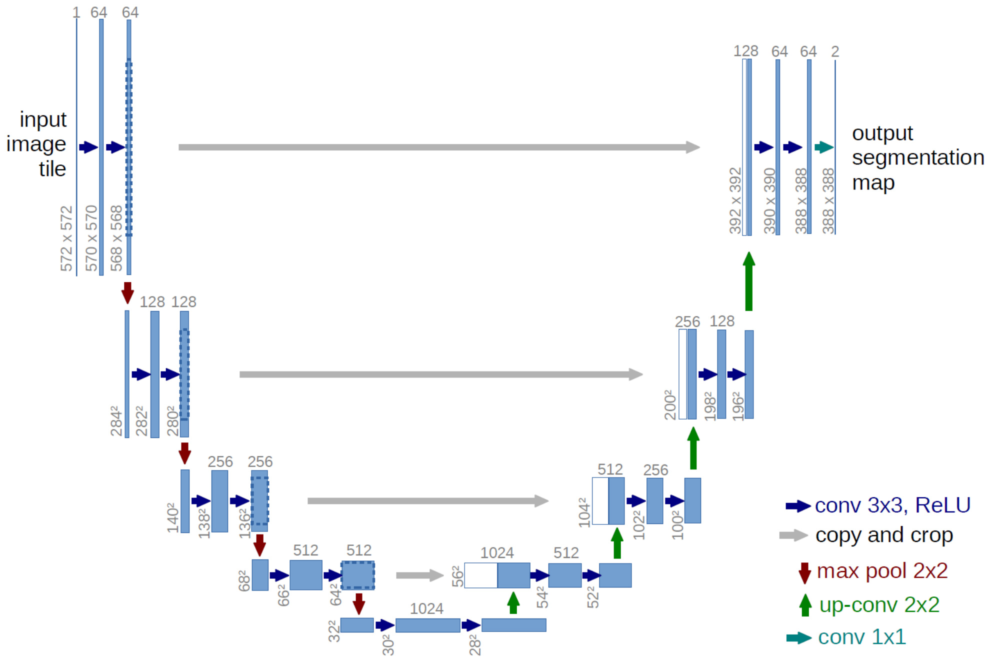
Deep Learning Workhorse: The U-Net

Invented MICCAI for medical imaging

Used throughout deep learning.

- Diffusion models
- Image-to-image translation

<u>**110,107 citations**</u>



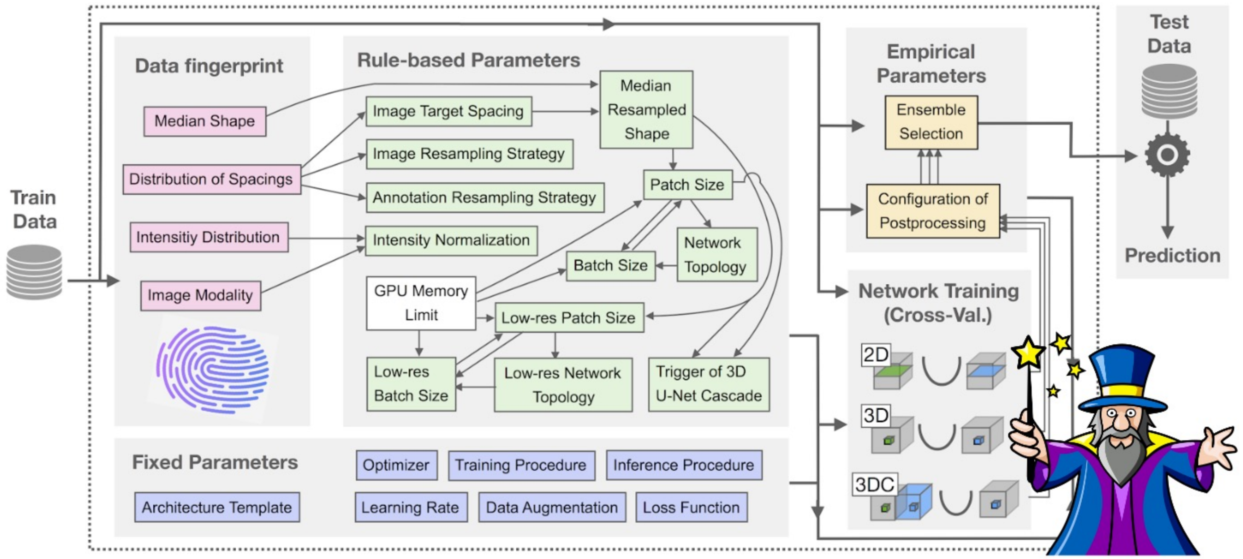
Ronneberger et al. MICCAI, 2015

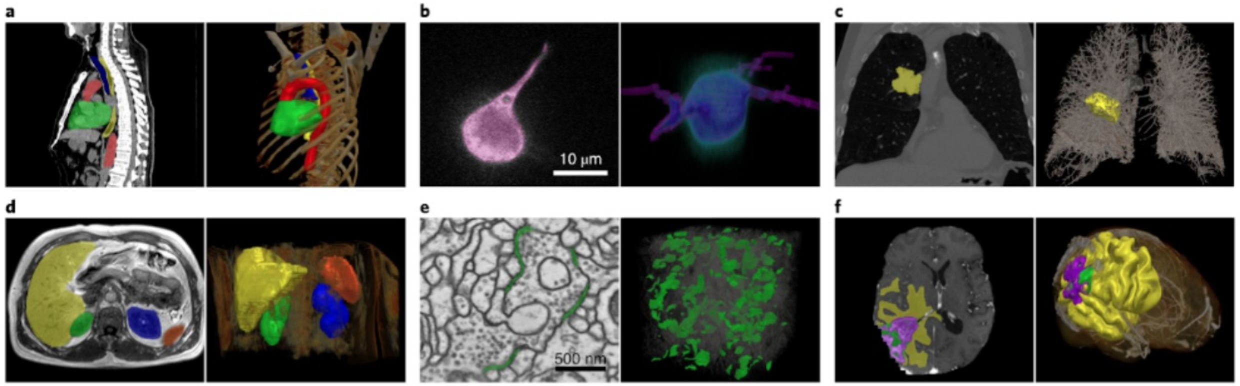
Making it Practical: nnU-Net ("No-New-U-Net")

5671 citations

	BTCV n=30	ACDC n=200	LiTS n=131	BraTS n=1251	KiTS n=489	AMOS n=360	VRAM [GB]	RT [h]	Arch.	nnU
nnU-Net (org.) [21]	83.08	91.54	80.09	91.24	86.04	88.64	7.70	9	CNN	Yes
nnU-Net ResEnc M	83.31	91.99	80.75	91.26	86.79	88.77	9.10	12	CNN	Yes
nnU-Net ResEnc L	83.35	91.69	81.60	91.13	88.17	89.41	22.70	35	CNN	Yes
nnU-Net ResEnc XL	83.28	91.48	81.19	91.18	**88.67**	89.68	36.60	66	CNN	Yes
MedNeXt L k3 [31]	84.70	**92.65**	82.14	91.35	88.25	89.62	17.30	68	CNN	Yes
MedNeXt L k5 [31]	**85.04**	92.62	**82.34**	**91.50**	87.74	**89.73**	18.00	233	CNN	Yes
STU-Net S [20]	82.92	91.04	78.50	90.55	84.93	88.08	5.20	10	CNN	Yes
STU-Net B [20]	83.05	91.30	79.19	90.85	86.32	88.46	8.80	15	CNN	Yes
STU-Net L [20]	83.36	91.31	80.31	91.26	85.84	89.34	26.50	51	CNN	Yes
SwinUNETR [32]	78.89	91.29	76.50	90.68	81.27	83.81	13.10	15	TF	Yes
SwinUNETRV2 [17]	80.85	92.01	77.85	90.74	84.14	86.24	13.40	15	TF	Yes
nnFormer [41]	80.86	92.40	77.40	90.22	75.85	81.55	5.70	8	TF	Yes
CoTr [37]	81.95	90.56	79.10	90.73	84.59	88.02	8.20	18	TF	Yes
No-Mamba Base	83.69	91.89	80.57	91.26	85.98	89.04	12.0	24	CNN	Yes
U-Mamba Bot [26]	83.51	91.79	80.40	91.26	86.22	89.13	12.40	24	Mam	Yes
U-Mamba Enc [26]	82.41	91.22	80.27	90.91	86.34	88.38	24.90	47	Mam	Yes
A3DS SegResNet [1,28]	80.69	90.69	79.28	90.79	81.11	87.27	20.00	22	CNN	No
A3DS DiNTS [1,18]	78.18	82.97	69.05	87.75	65.28	82.35	29.20	16	CNN	No
A3DS SwinUNETR [1,32]	76.54	82.68	68.59	89.90	52.82	85.05	34.50	9	TF	No

Isensee et al. MICCAI, 2024





Isensee et al. Nat. Methods, 2021

Generalist Segmentation Models

TotalSegmentator

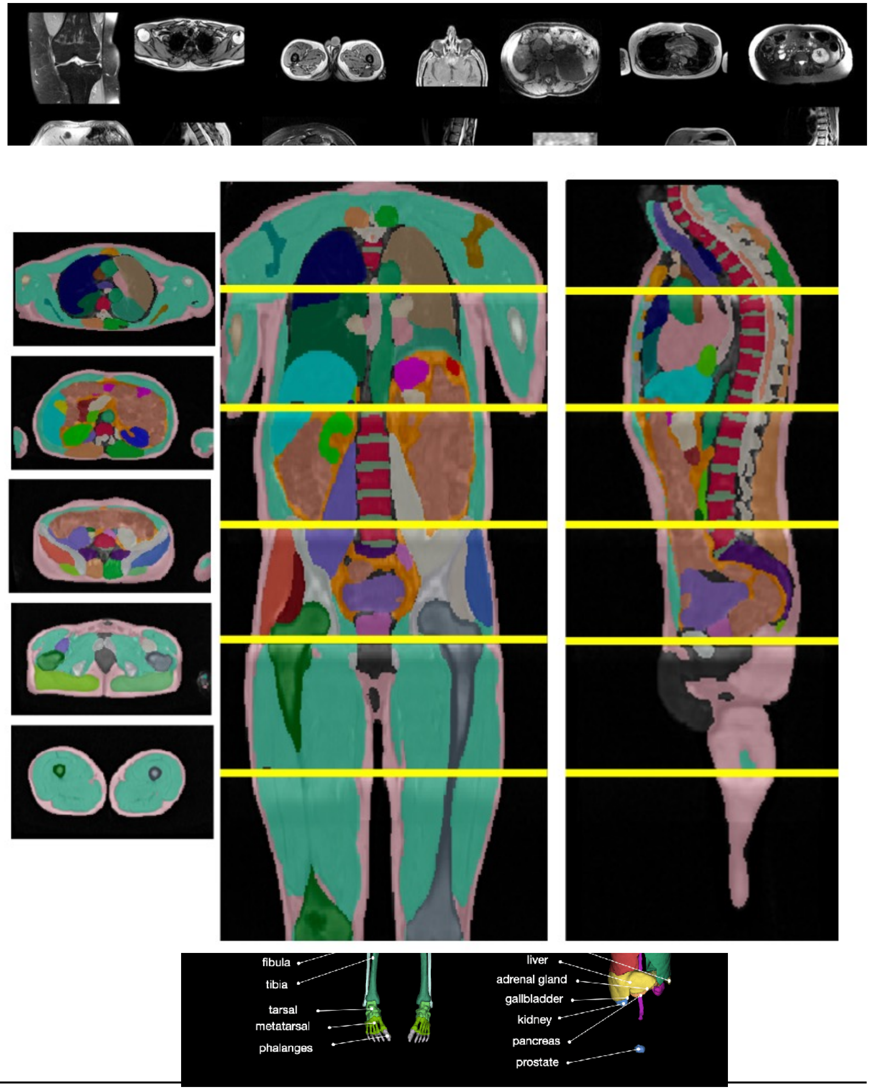
Trained ~1200 CTs to segment 104 tissues

TotalSegmentator MRI

Train ~1100 scans (600 MRI) to segment 80 tissues

UKBOB

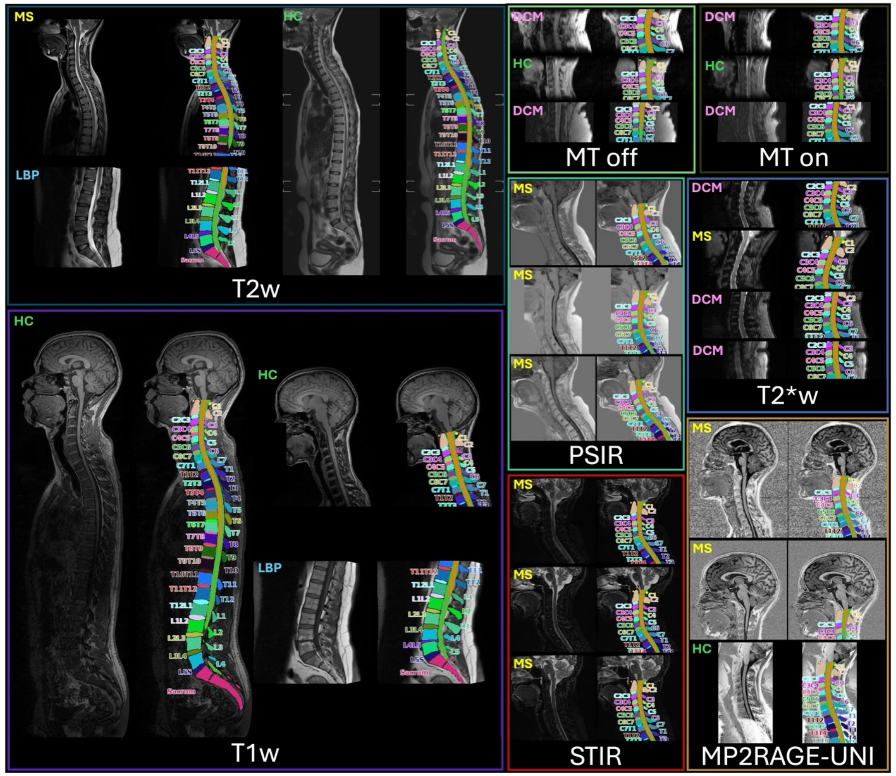
TotalSegmentator segmentations for 50k wholebody MRIs from UK Biobank
Train ViT-based segmentation foundation model



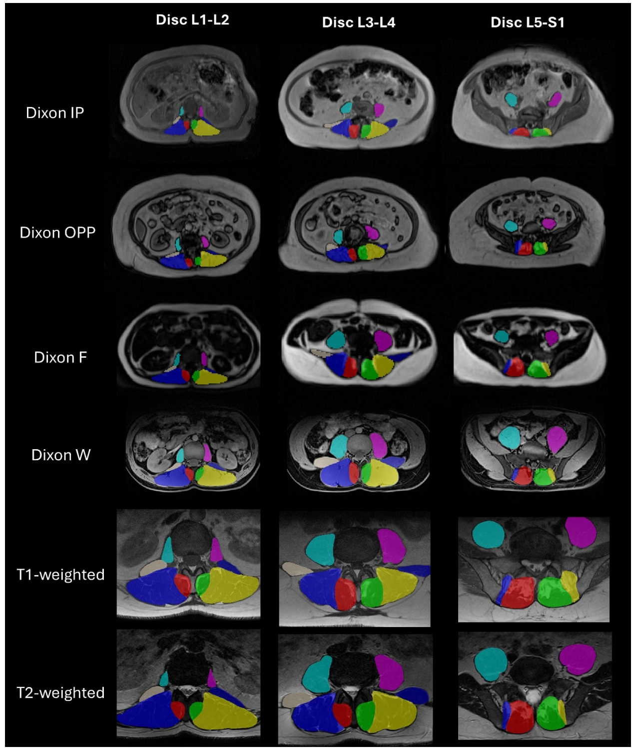
Generalist Segmentation Models

3:45 Towards Foundation Models in MRI

0624: TotalSpineSeg



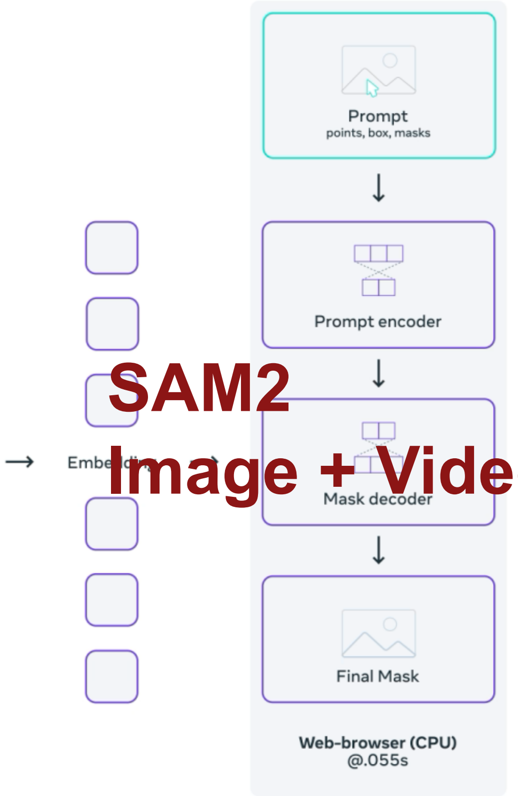
0622: SegmentAnyMuscle



Stanford University

Foundation Models & Prompting

Segment Anything Model (SAM)



SAM2
Image + Video

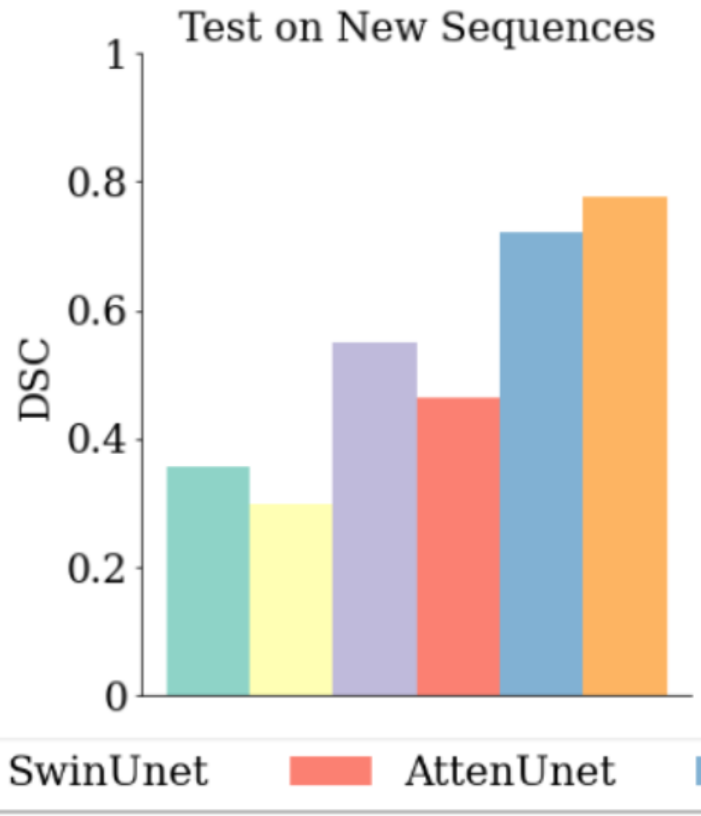
Ogier et al. Frontier in Neurology 2021; Hu, M. et al. JMRI 2024; Kirillov et al. arXiv 2023; Ravi et al. arXiv 2024

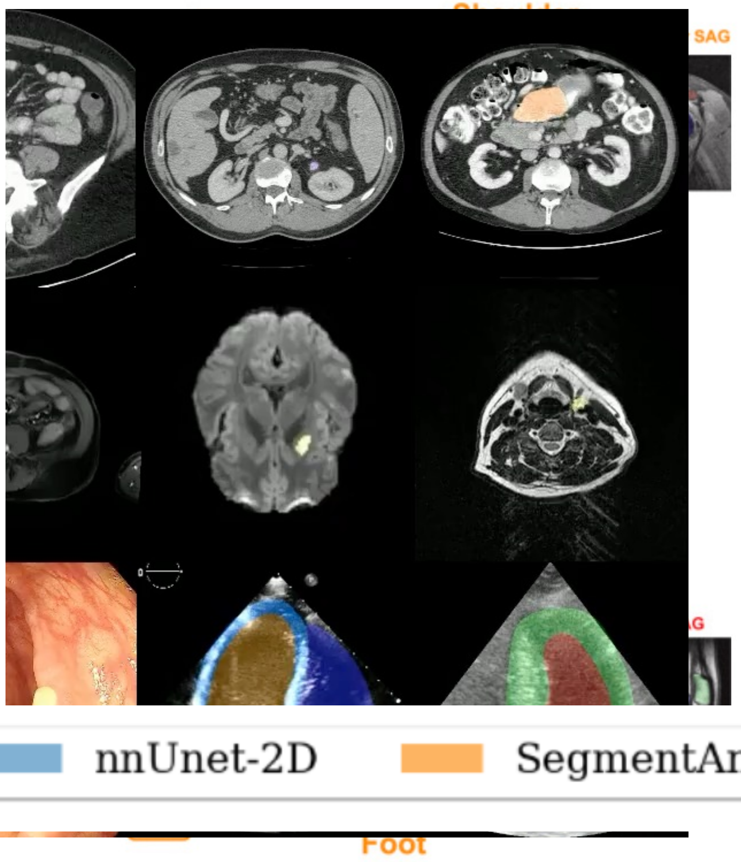
Medical Specific - Foundation Models

Segment Anything Model (SAM)

- 3D medical image and video segr
- 455,000+ 3D image-mask pairs
- 76,000+ annotated video frames

- SegmentAnyBone
- mskSAM





Ma et al. arXiv, 2025; Hoyer et al. arXiv, 2025; Gu et al. Med Image Anal, 2025

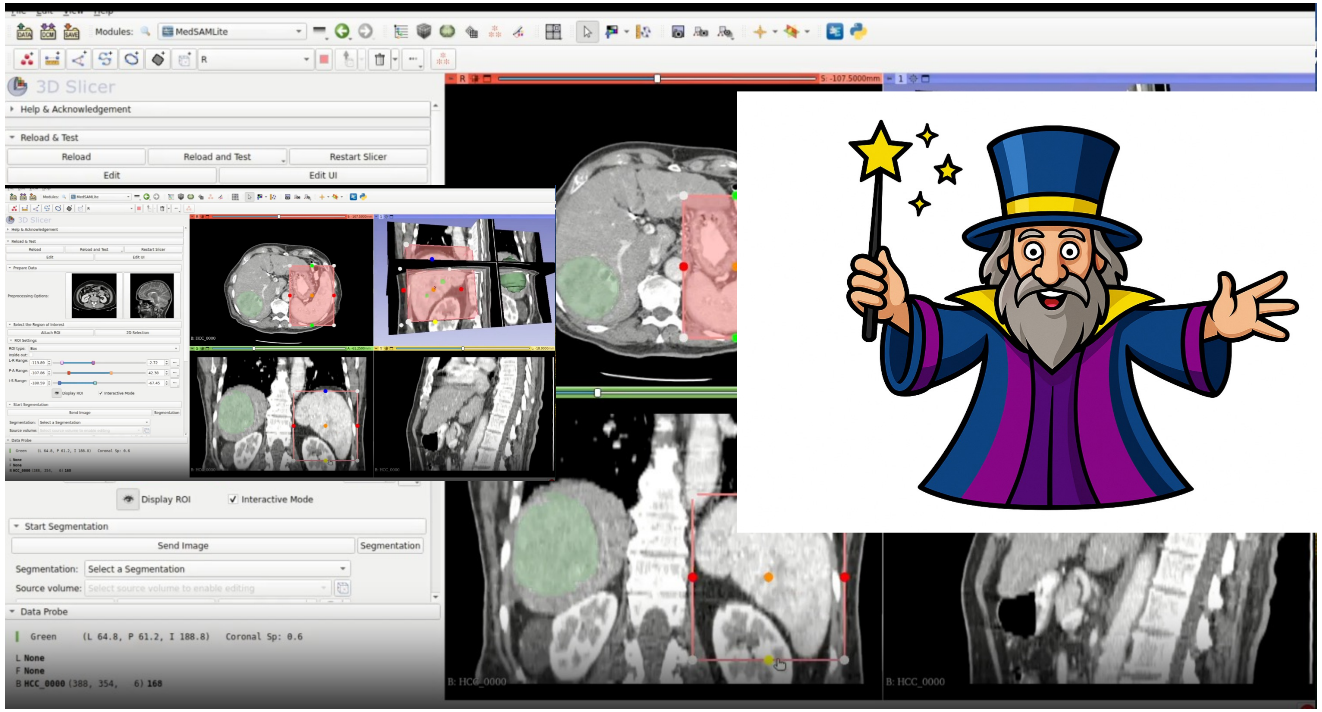
Training Recommendations

Labeling New Data

SAM-based models

Training local model

nnUNet



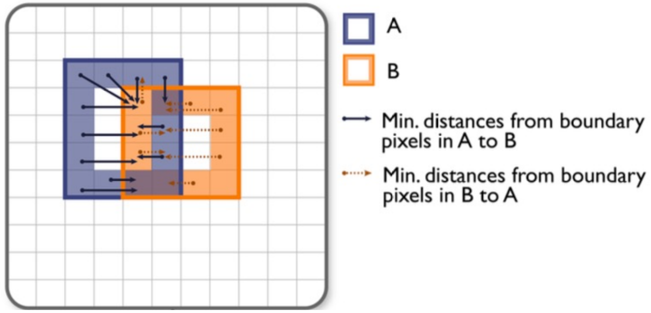
State of the art!!

… test the other methods first!

Evaluation Recommendations

<u>Big Objects</u>

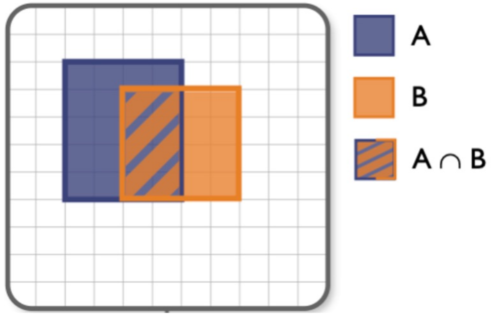
Surface-based errors



ASSD

<u>Small Objects</u>

Overlap measures



DSC

<u>Always! Go-beyond Dice</u>

Why do **<u>you</u>** want a segmentation?

Evaluate on **<u>your application</u>**!

Future Outlook

- Still in research mode

- Need to show new clinical utilities
    - Models in the clinic were also segmented manually
    - Demonstrate clinical utility & who will / wants to pay