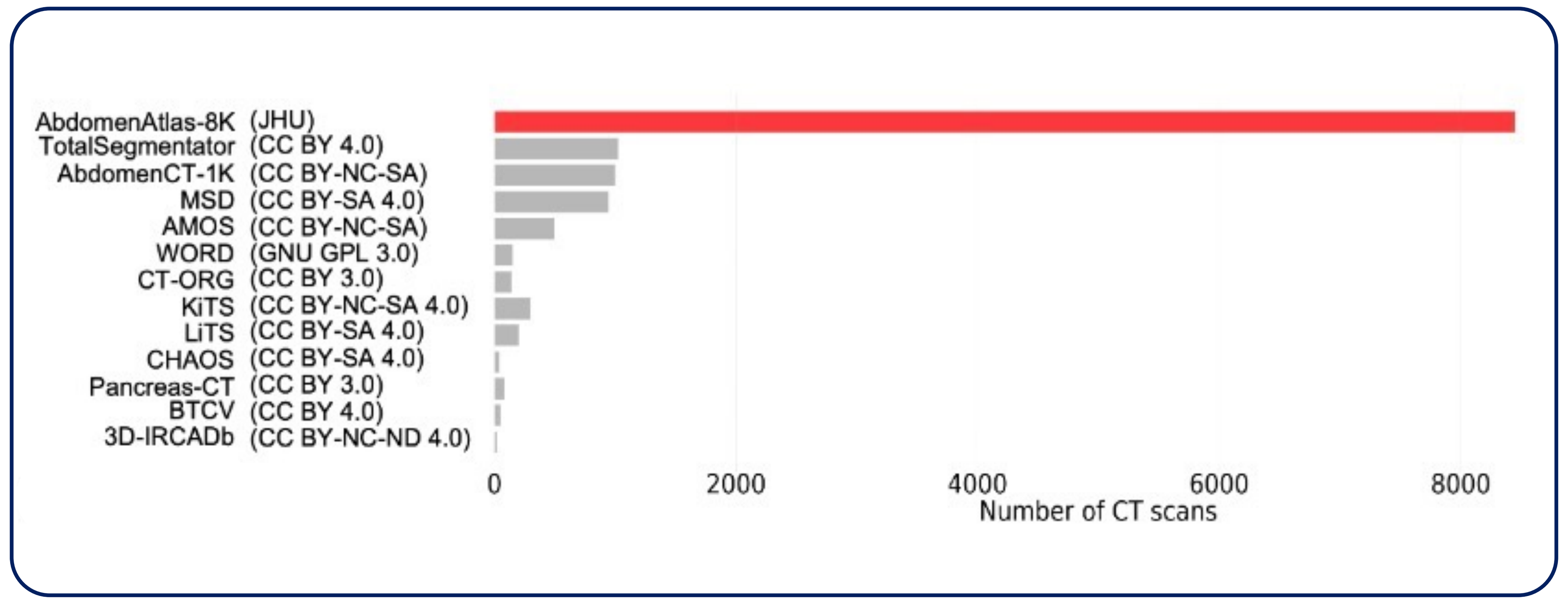
AbdomenAtlas-8K: Annotating 8,000 CT Volumes for Multi-Organ Segmentation in Three Weeks

Chongyu Qu, Tiezheng Zhang, Hualin Qiao, Jie Liu, Yucheng Tang, Alan L. Yuille, Zongwei Zhou*

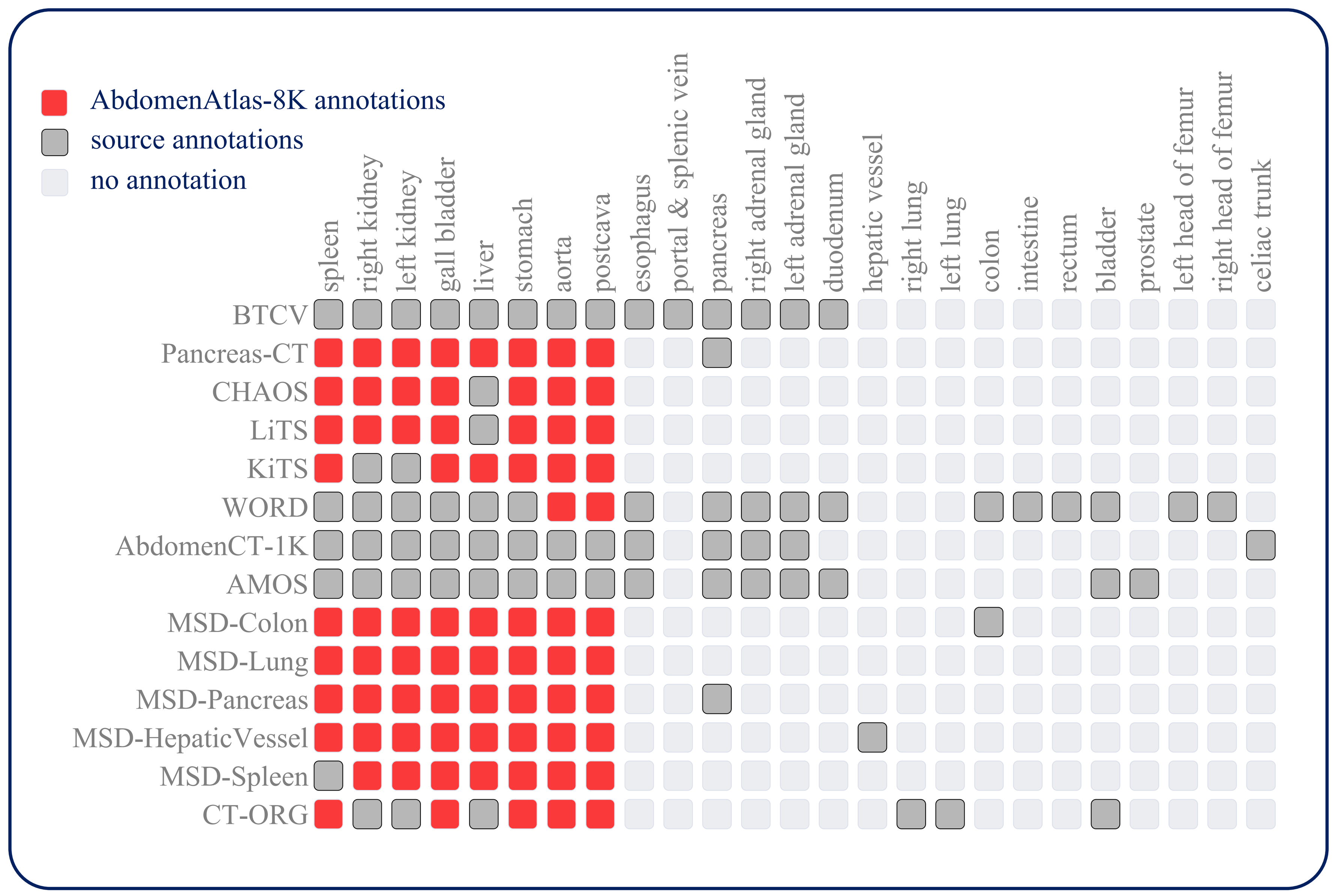
❖ Background

Medical segmentation, crucial for diagnosis, treatment, and radiotherapy planning, relies on substantial annotated datasets. However, creating large-scale 3D medical datasets with per-voxel annotations is impractical using conventional methods due to the high cost and time involved. Consequently, **publicly available datasets for multi-organ segmentation are:**

❖Limited in dataset size

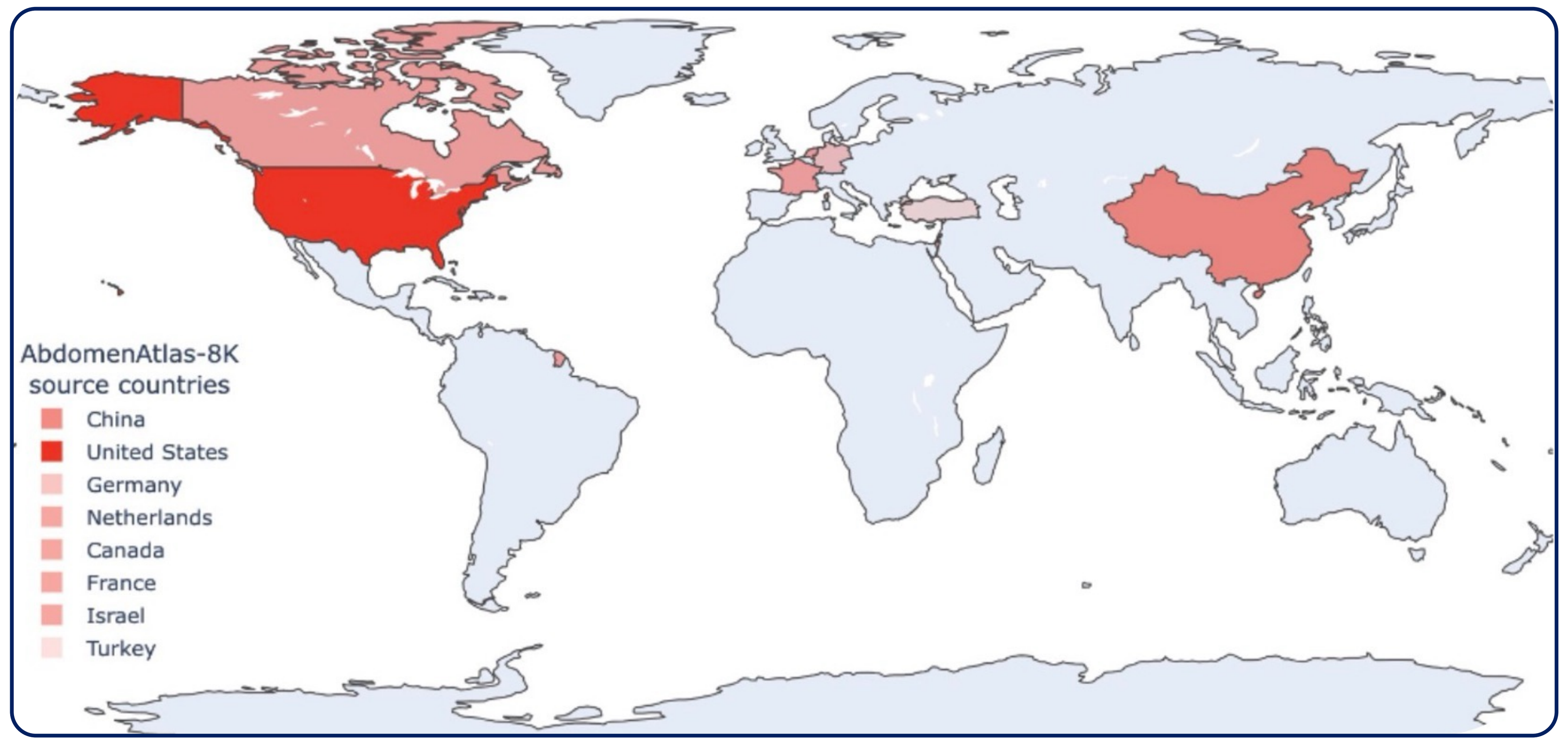


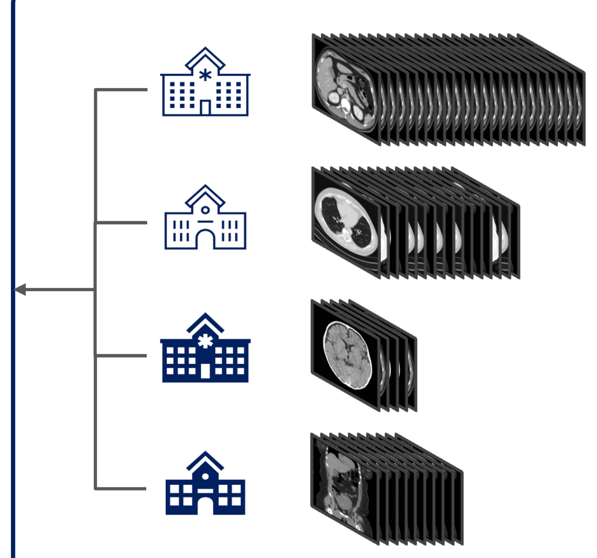
❖Restricted annotation sizes



❖ Dataset Properties

AbdomenAtlas-8K was a composite dataset that unified datasets from 26 different hospitals worldwide. In total, over **60.6 × 10^9** voxels were annotated in comparison with **4.3 × 10^9** voxels annotated in the public datasets. We scaled up the organ annotation **by a factor of 15**





Collected from

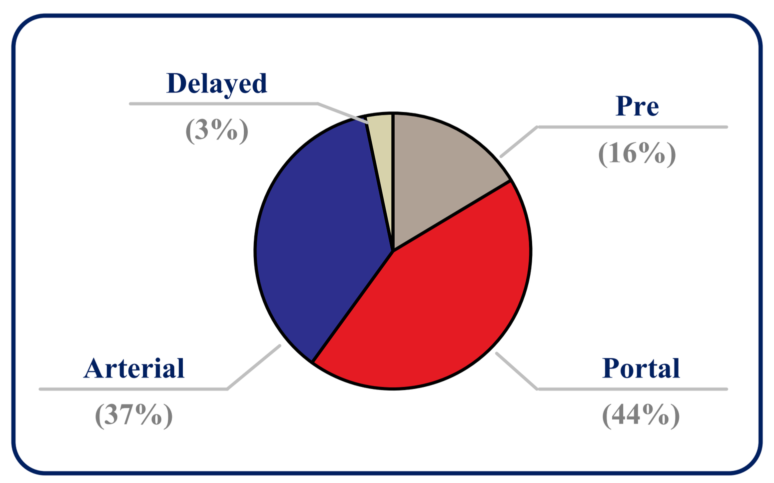
26 hospitals

worldwide

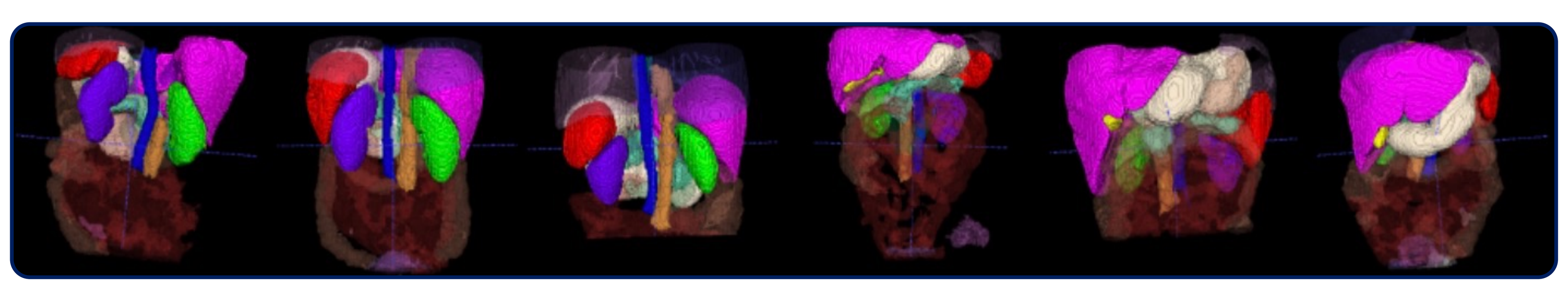
Annotated **8** organs

Integrated **15** datasets

Annotated **8,448** CT volumes

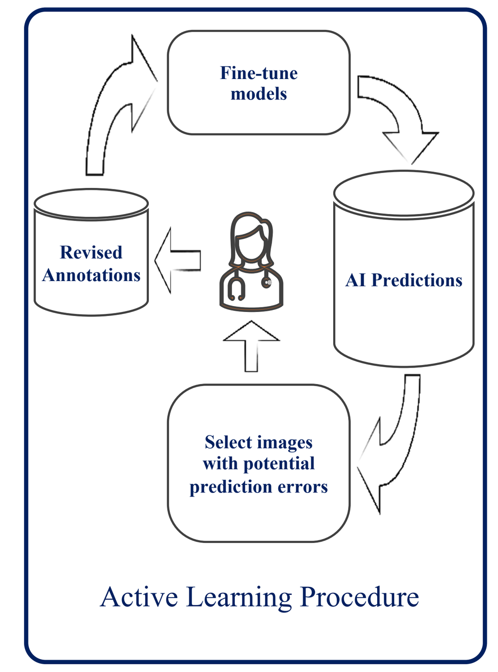


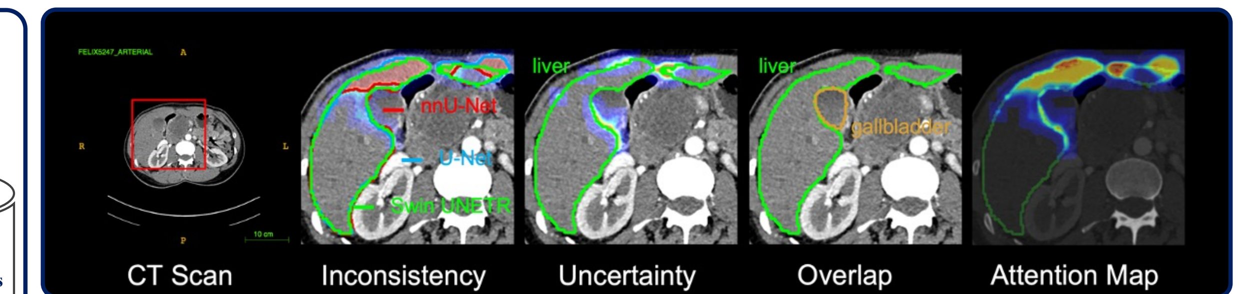
Annotated **3.2M** images



❖ Active Learning

Uncertainty and **diversity** are key criteria in active learning. Our active learning considers **anatomical priors**, **AI prediction uncertainty**, and **data diversity**. It focuses on prospective applications, and the criteria produce an **attention map** for precise detection of high-risk prediction errors.

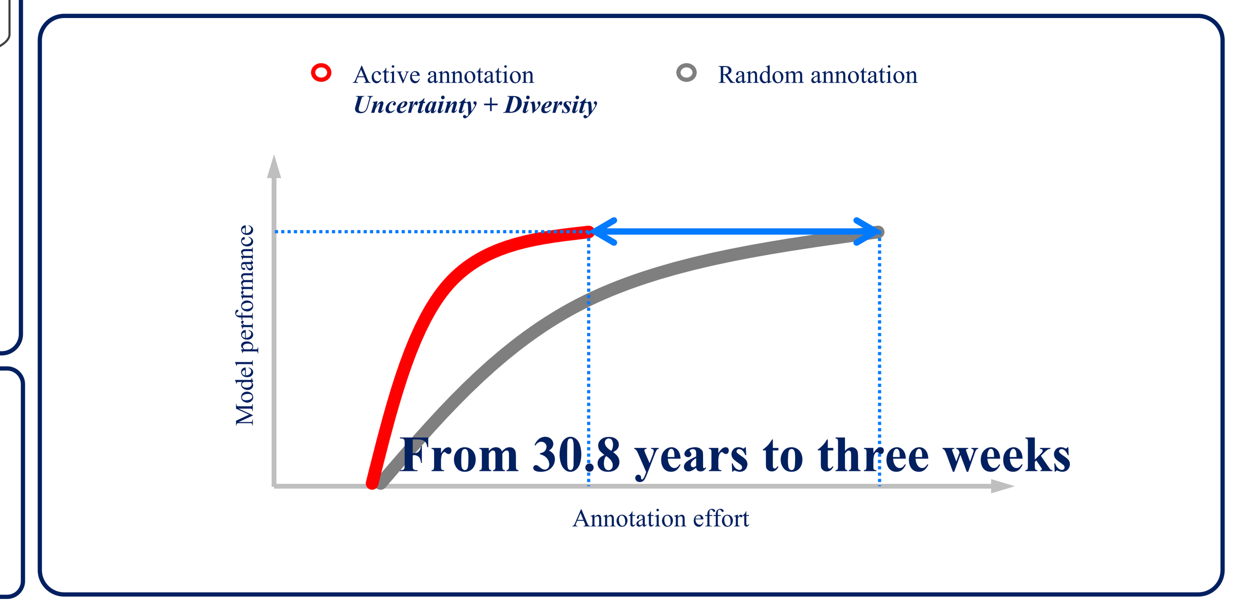




Up to

533x faster

than previous strategies



❖ Benchmarking Results

We applied four AI models trained on the AbdomenAtlas-8K (public) directly to the JHH dataset (unseen, private) and compared their performance with AI models trained on JHH.

AI Models	Trained on JHH (private)		Trained on AbdomenAtlas-8K (public)	
	mDSC (%)	mNSD (%)	mDSC (%)	mNSD (%)
SwinUNETR	84.8 ± 12.6	66.5 ± 14.8	86.5 ± 7.5	60.8 ± 10.9
UNETR	78.6 ± 13.0	53.9 ± 14.1	86.6 ± 6.5	59.5 ± 10.4
U-Net	84.8 ± 11.2	64.2 ± 14.8	87.3 ± 6.0	61.0 ± 10.2
SegResNet	87.6 ± 6.60	64.9 ± 10.9	87.0 ± 6.1	60.4 ± 10.1

This study is a concrete demonstration of how AbdomenAtlas-8K can be used to train AI models that can be generalized to many CT volumes from novel hospitals and be **adapted to address a range of clinical problems.**