CVPR
JUNE 17-21, 2024
SEATTLE, WA

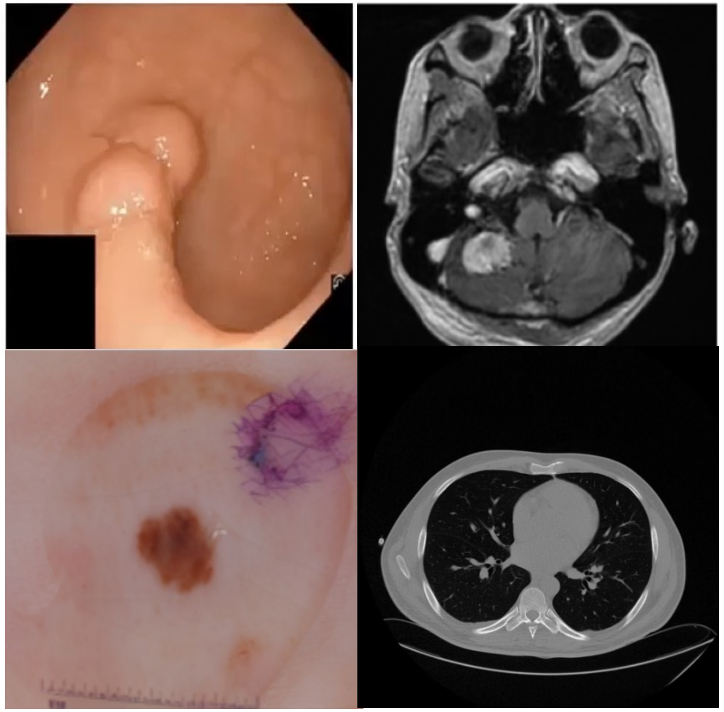
Towards Generalizable Tumor Synthesis

Qi Chen, Xiaoxi Chen, Haorui Song, Zhiwei Xiong,
Alan Yuille, Chen Wei, and Zongwei Zhou

Code

Paper

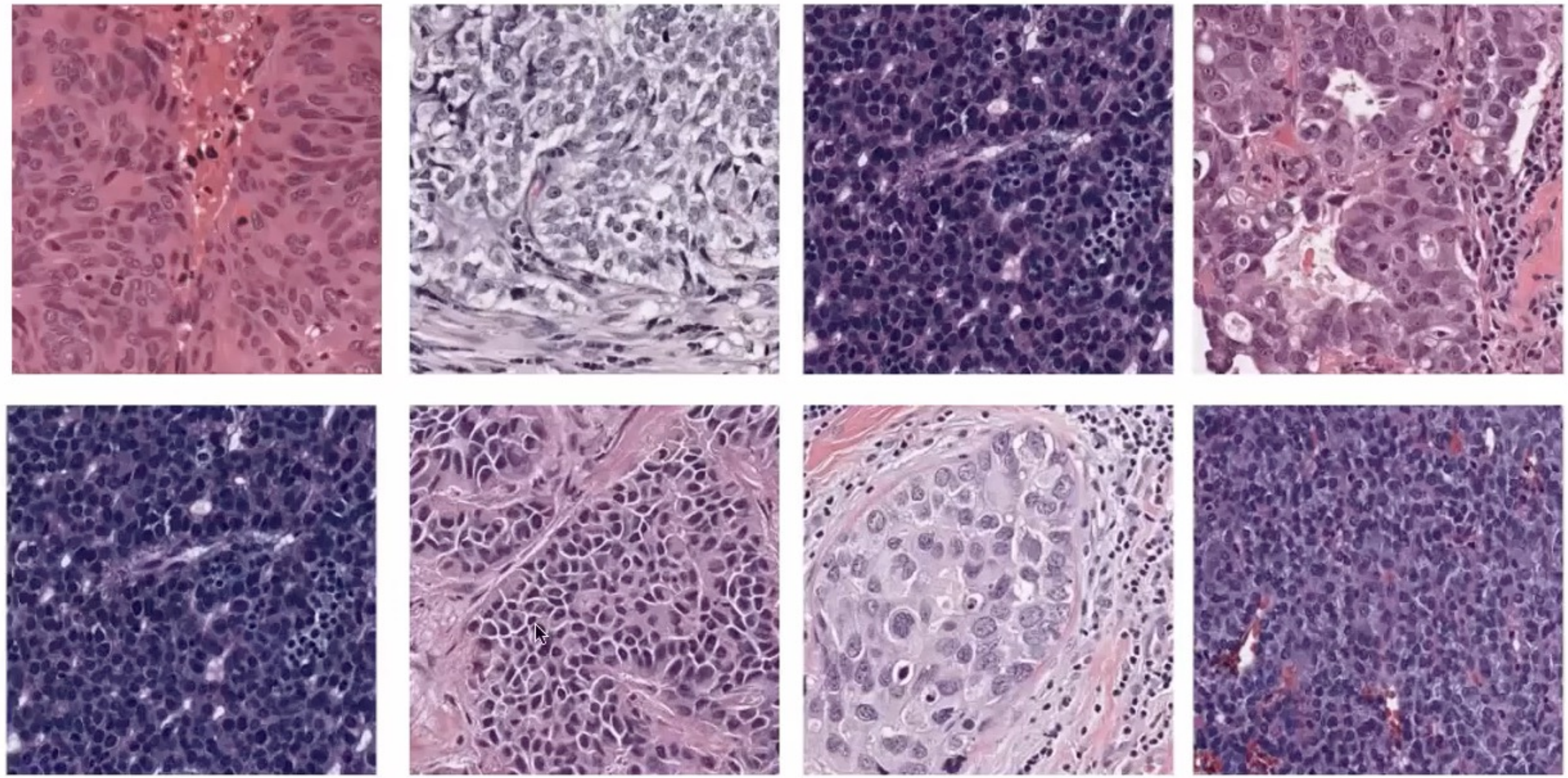
AI Generated Images



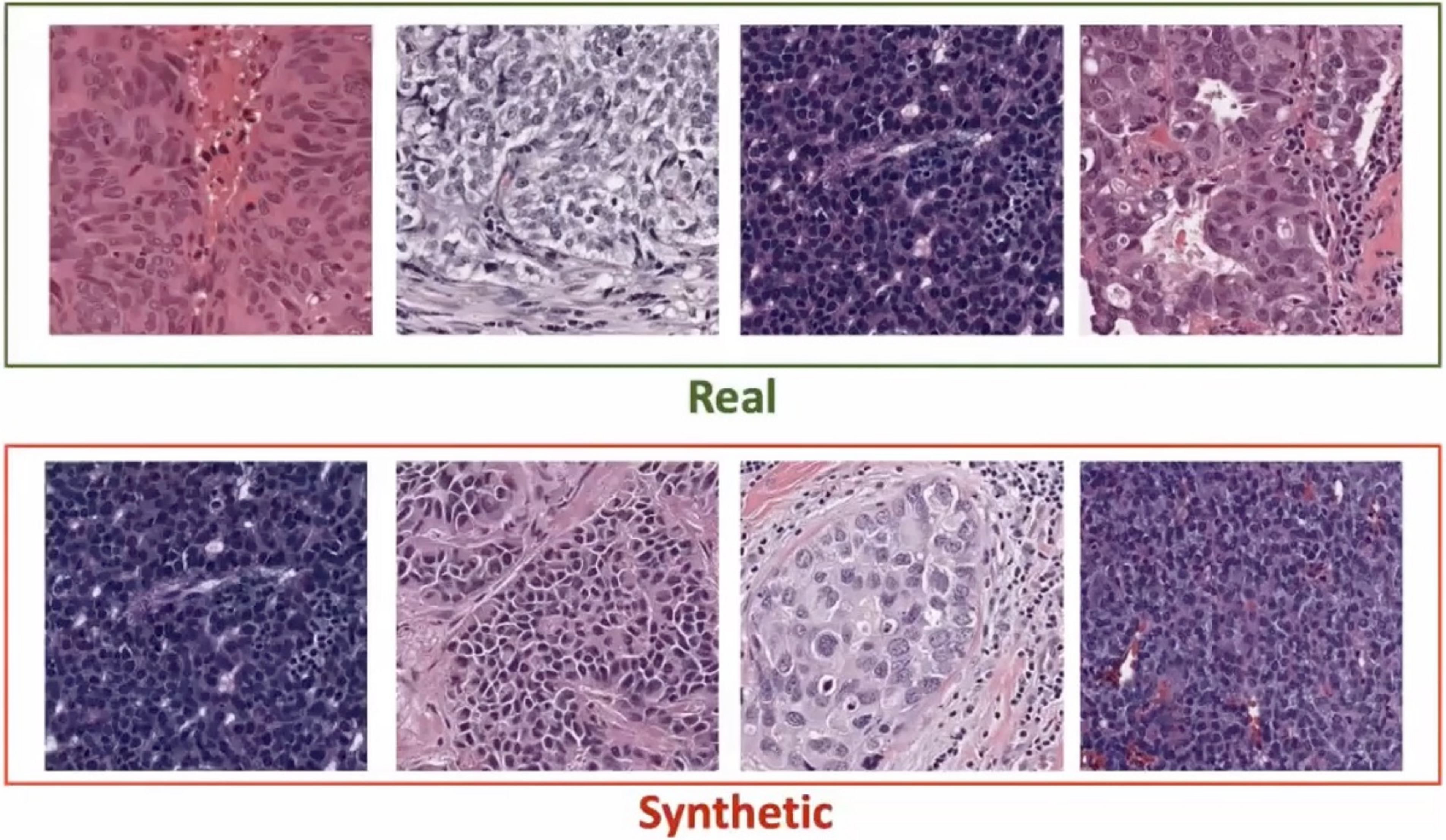
[1] https://pixlr.com/image-generator/
[2] Khader, et al. "Denoising diffusion probabilistic models for 3D medical image generation." *Scientific Reports*.
[3] Du, Shiyi, et al. "Boosting dermatoscopic lesion segmentation via diffusion models with visual and textual prompts." *arXiv 2310.*

Which ones are real/synthetic?

Which ones are real/synthetic?



Which ones are real/synthetic?



AI Generated Videos/Volumes

A litter of golden retriever puppies playing in the snow. Their heads pop out of the snow, covered in.

Reflections in the window of a train traveling through the Tokyo suburbs.

[1] https://sora.aitubo.ai/videos/latest
[2] Hamamci, Ibrahim Ethem, Sezgin Er, Enis Simsar, Alperen Tezcan, Ayse Gulnihan Simsek, Furkan Almas, Sevval Nil Esirgun et al. "GenerateCT: text-guided 3D chest CT generation." arxiv2305.

2 - Synthetic Data for Vision Recognition

- Synthetic data for supervised pre-training
- Synthetic data for data augmentation
- Synthetic data for addressing class imbalance

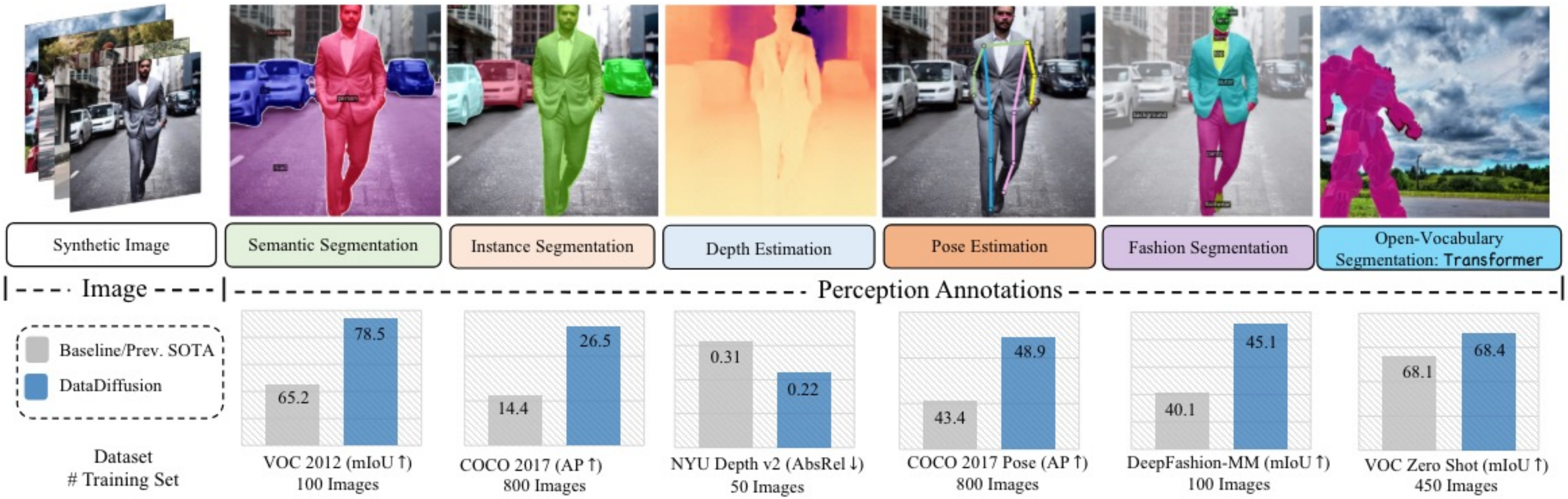
1/3 Synthetic data for supervised pre-training

Train Set	Number	Backbone	mIoU
*Train with Pure **Real** Data*			
VOC	R: 10.6k (all)	R50	77.3
	R: 10.6k (all)	Swin-B	84.3
	R: 5.0k	Swin-B	83.4
*Train with Pure **Synthetic** Data*			
DiffuMask	S: 60.0k	R50	57.4
	S: 60.0k	Swin-B	70.6
*Finetune on **Real** Data*			
VOC, **DiffuMask**	S: 60.0k + R: 5.0k	R50	77.6
	S: 60.0k + R: 5.0k	Swin-B	84.9

Table 1 – Result of Semantic Segmentation

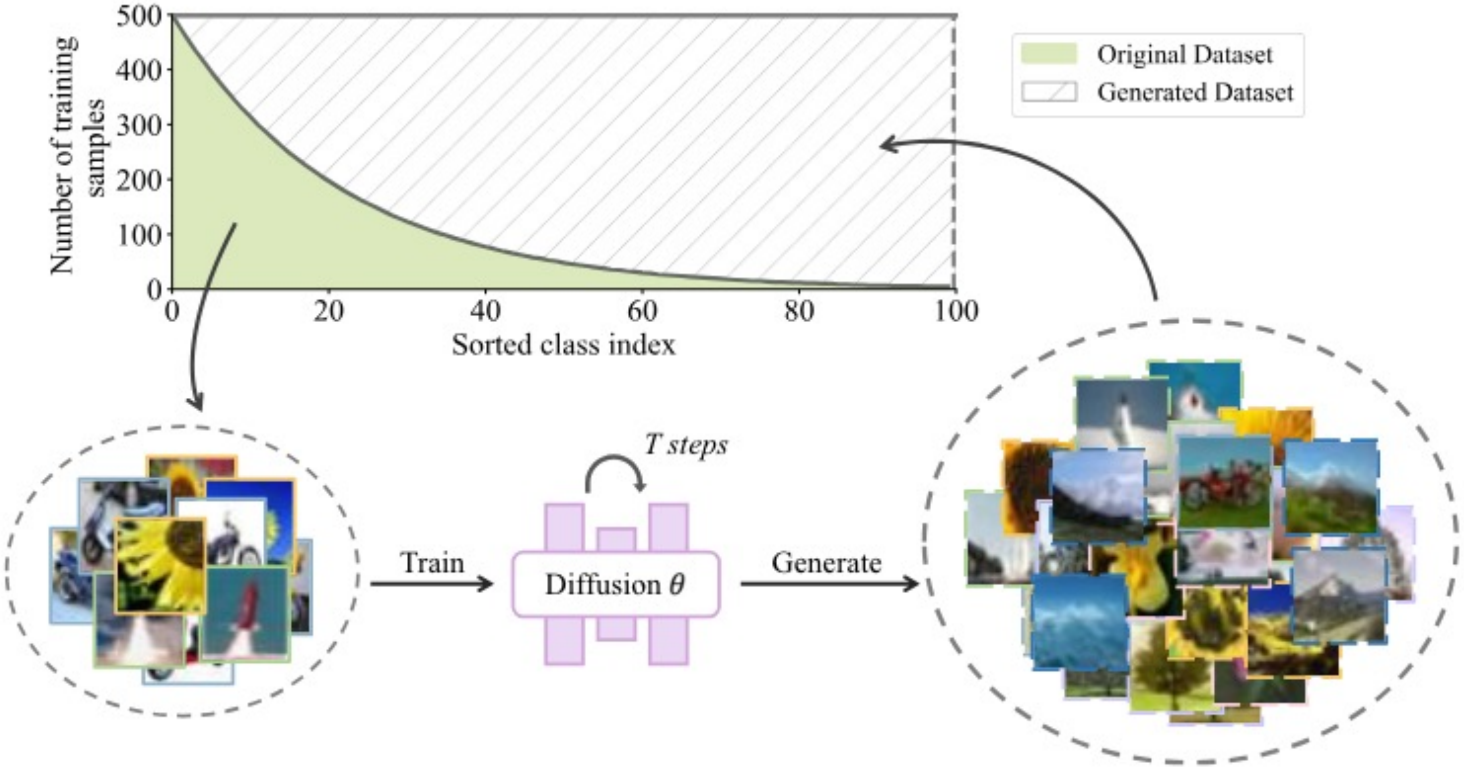
[1] Wu, Weijia, Yuzhong Zhao, Mike Zheng Shou, Hong Zhou, and Chunhua Shen. "DiffuMask: Synthesizing Images with Pixel-level Annotations for Semantic Segmentation Using Diffusion Models". ICCV 2023.

2/3 Synthetic data for data augmentation



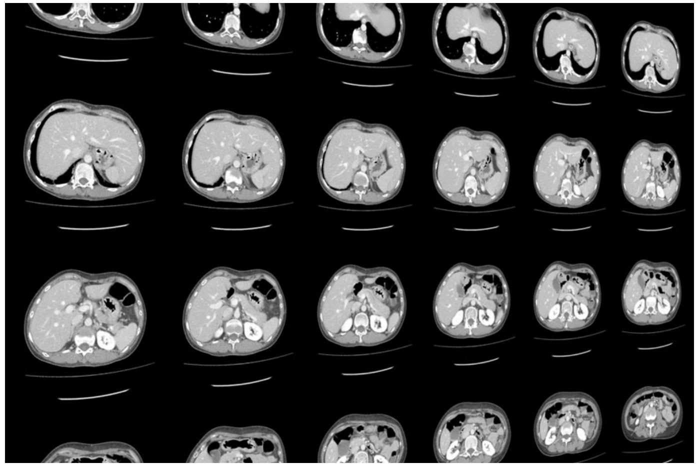
[1] Wu, Weijia, Yuzhong Zhao, Hao Chen, Yuchao Gu, Rui Zhao, Yefei He, Hong Zhou, Mike Zheng Shou, and Chunhua Shen. "DatasetDM: Synthesizing Data with Perception Annotations Using Diffusion Models". NIPS 2023.

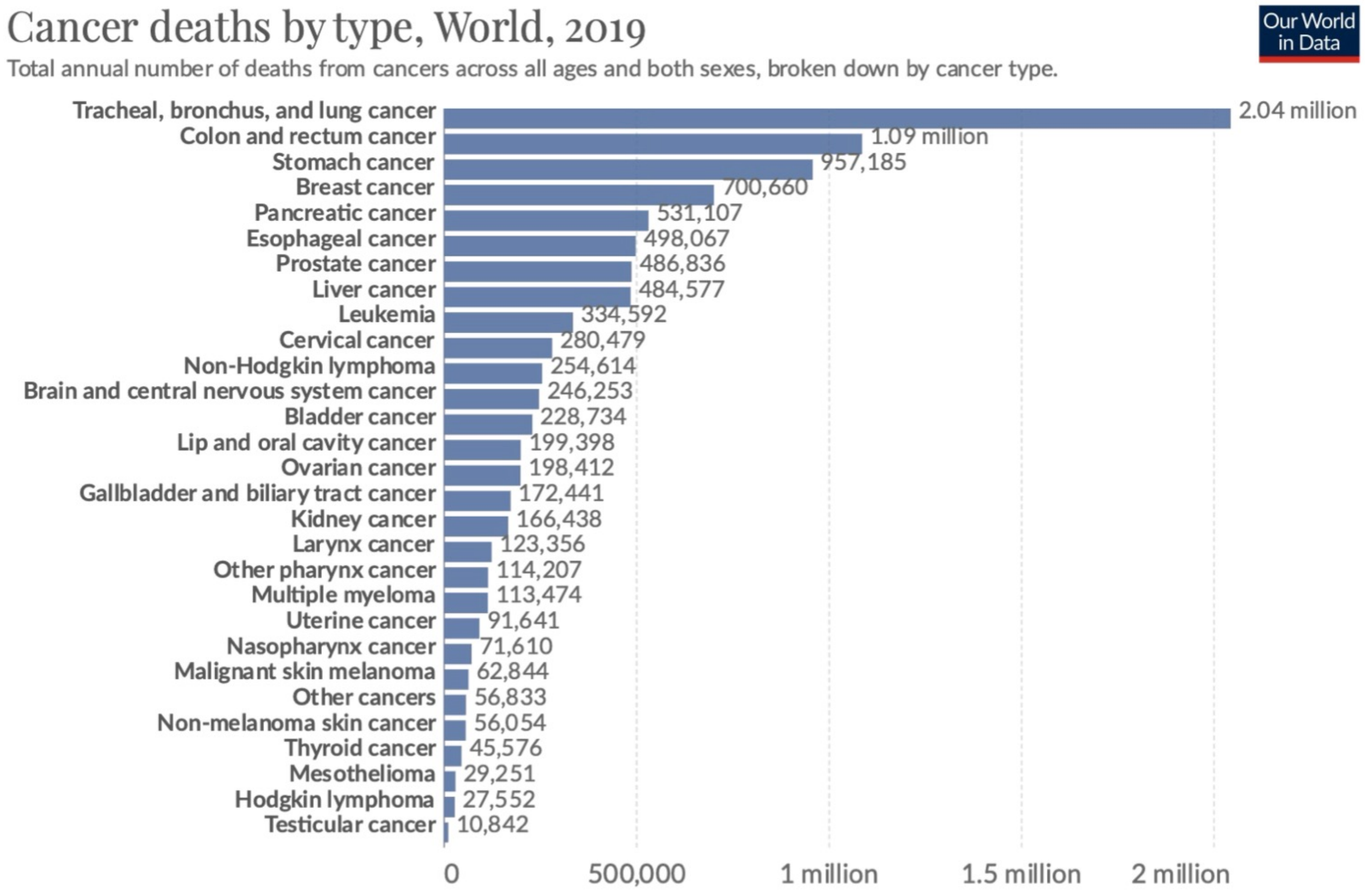
3/3 Synthetic data for addressing class imbalance



[1] Shao, Jie, Ke Zhu, Hanxiao Zhang, and Jianxin Wu. "DiffuLT: How to Make Diffusion Model Useful for Long-tail Recognition." *arXiv:2403.*

Cancer Diagnosis From CT scans





Cancer Research Challenges

- Data scarcity
- Generalizable to different organs
- Generalizable to different demographics

1/3 Data Scarcity

Challenges for releasing tumor data

- Patient Privacy and Consent
- Data Collection and Annotation

ImageNet: 14,197,122

Places: over 10,000,000

Microsoft COCO: 328,000

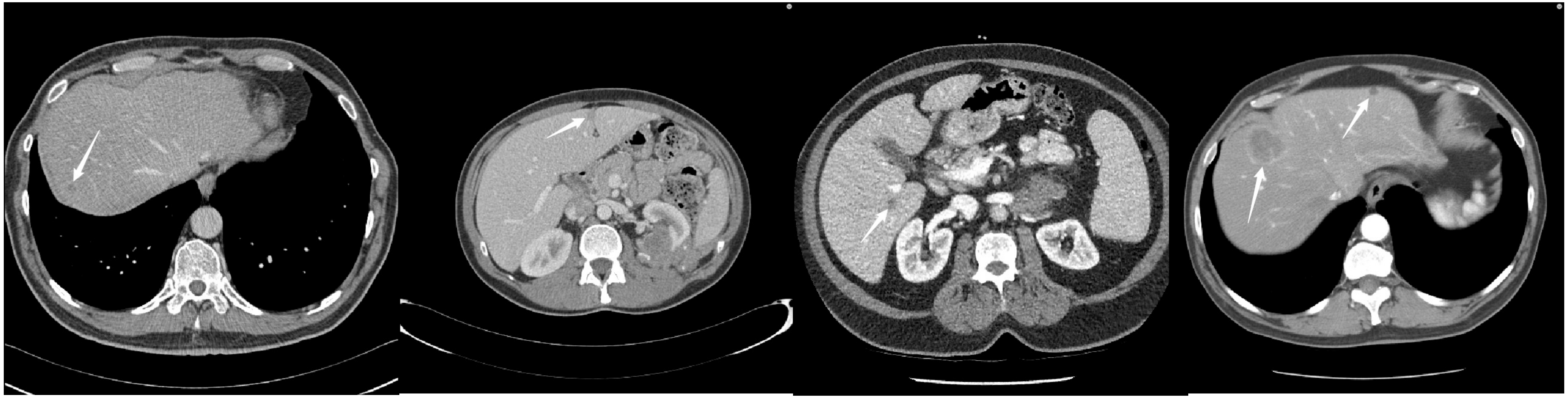
LiTS (liver tumor data) : 131 examples

MSD-Pancreas (pancreas tumor data): 282 examples

KiTS (kidney tumor data) : 300 examples

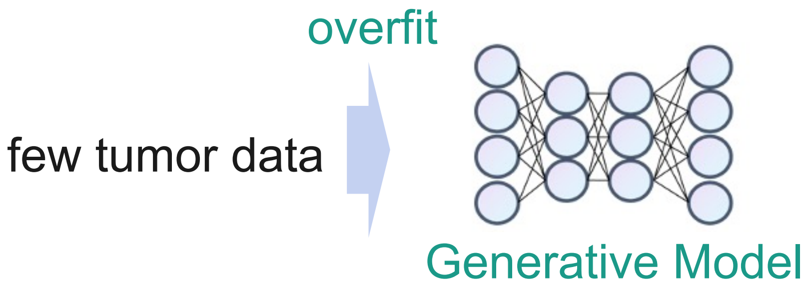
Generative Model, like Diffusion Model, requires large-scale annotated data

Similarity Among Same Type of Tumors

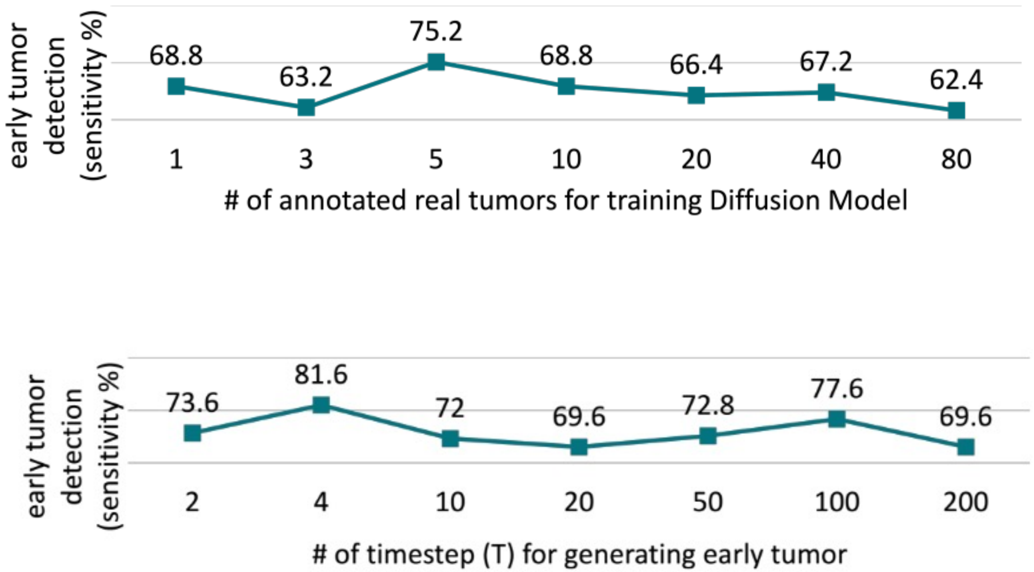


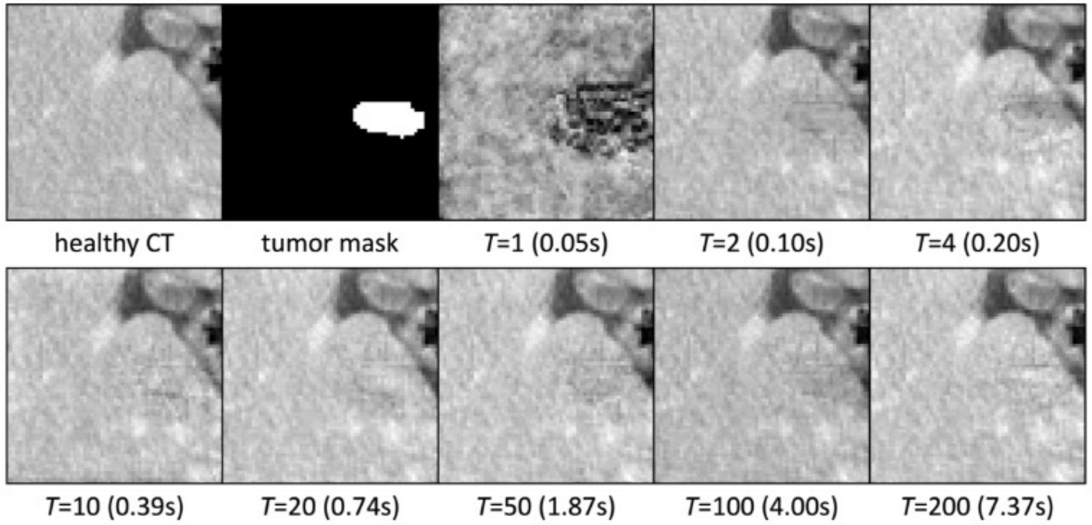
Liver Tumor

- Small
- Simple texture
- Similar shape - spheroid



Reduced Annotations and Accelerated Tumor Synthesis

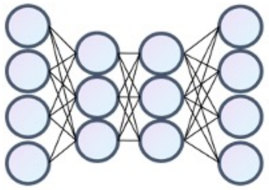




2/3 Generalizable to Different Organs

Paradigm1 (old)

CT scans of Liver tumor



Segmentation Model

Liver tumor mask

tumor types increase

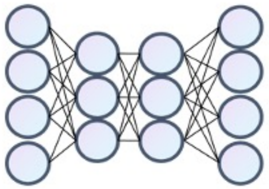
- Train a lot of segmentation models (for every tumor type)
- Need lots of different types of tumor data

2/3 Generalizable to Different Organs

Paradigm2 (new)

Training phase

Liver tumor

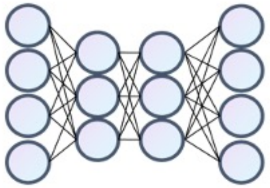


Generative Model

Liver tumor

Inference phase

Any Organ



Generative Model

Liver tumor

Pancreatic tumor

Kidney tumor

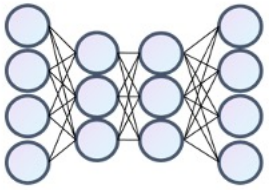
…

Liver tumor

Pancreatic tumor

Kidney tumor

…



Segmentation Model

Liver tumor mask

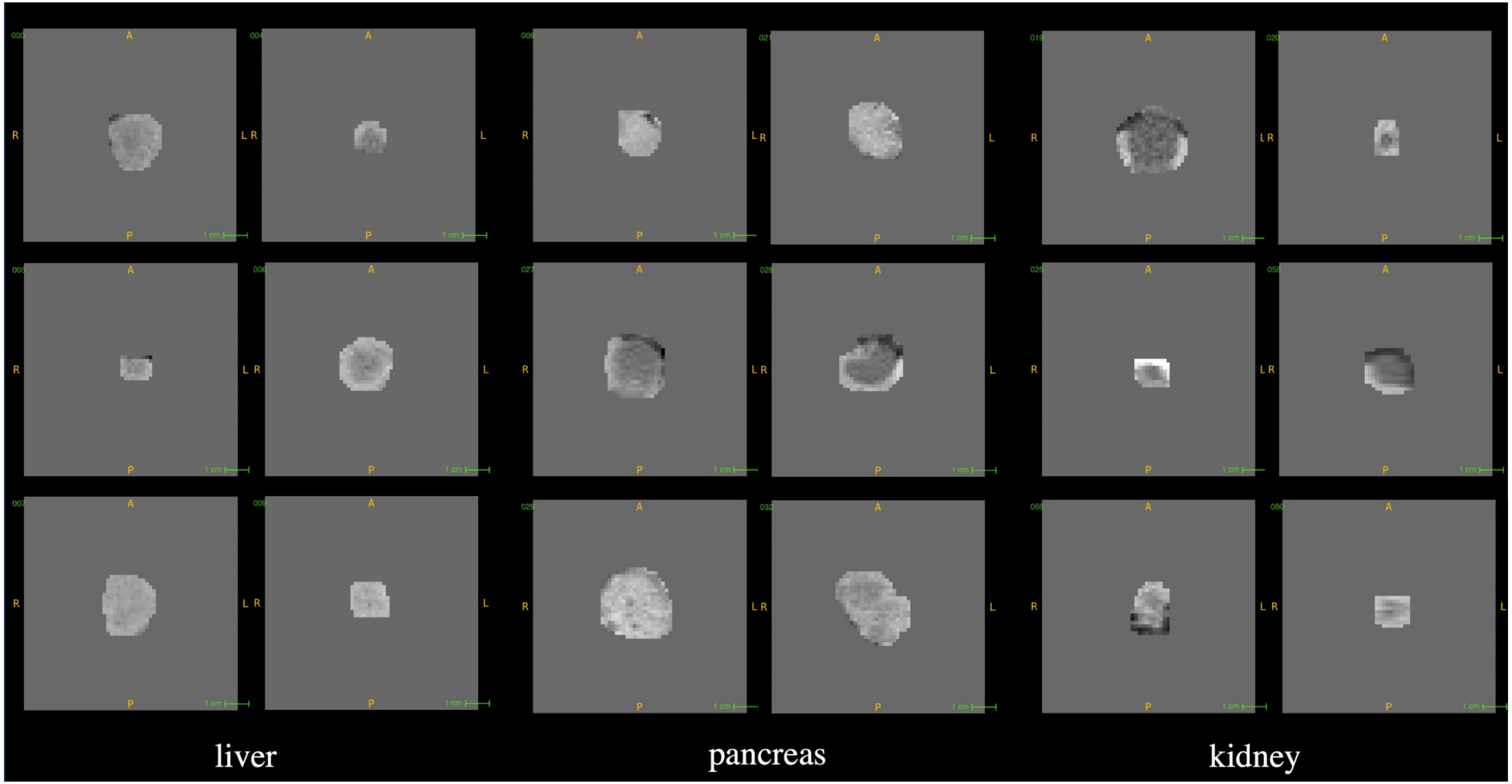
Pancreatic tumor mask

Kidney tumor mask

…

Key: Find the connection across organs!

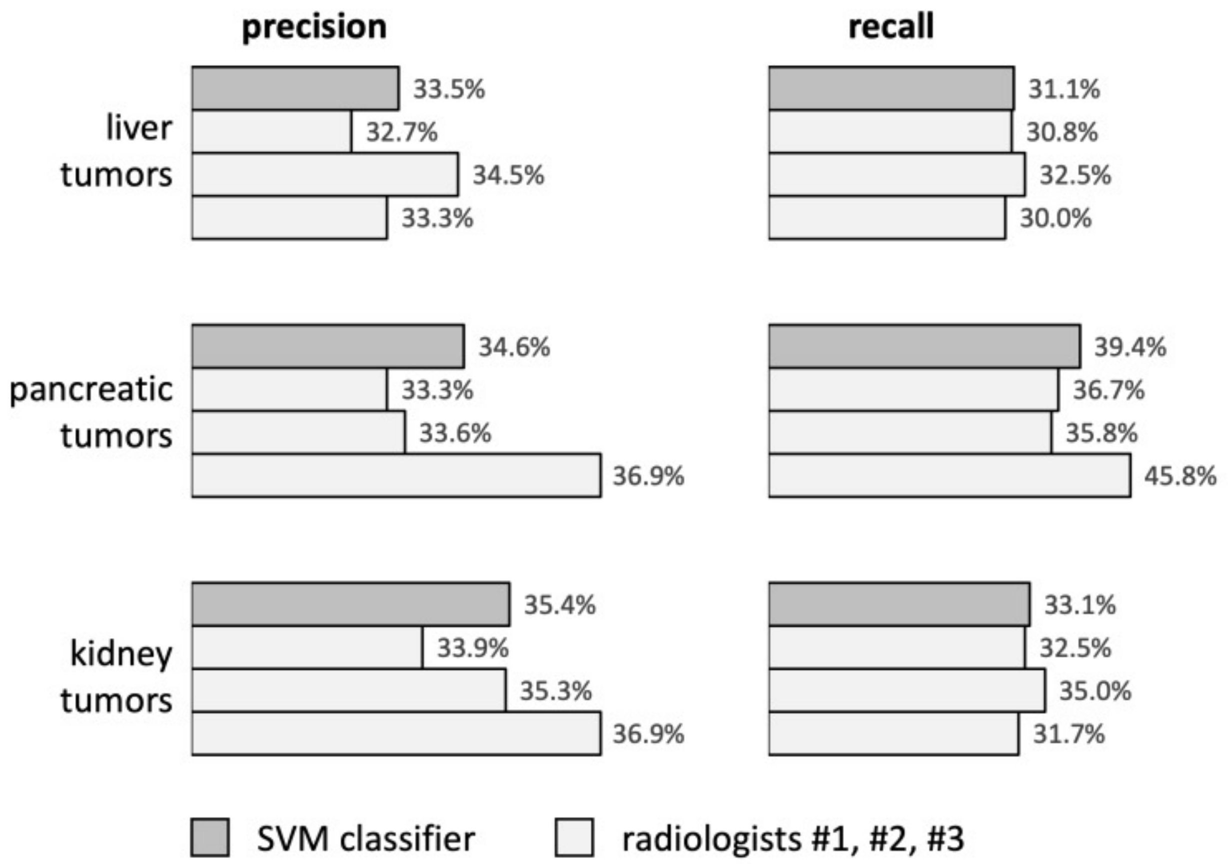
High Similarity Among Different Type of Tumors



- Early-stage tumors (< 2cm) tend to have similar imaging characteristics in computed tomography (CT).

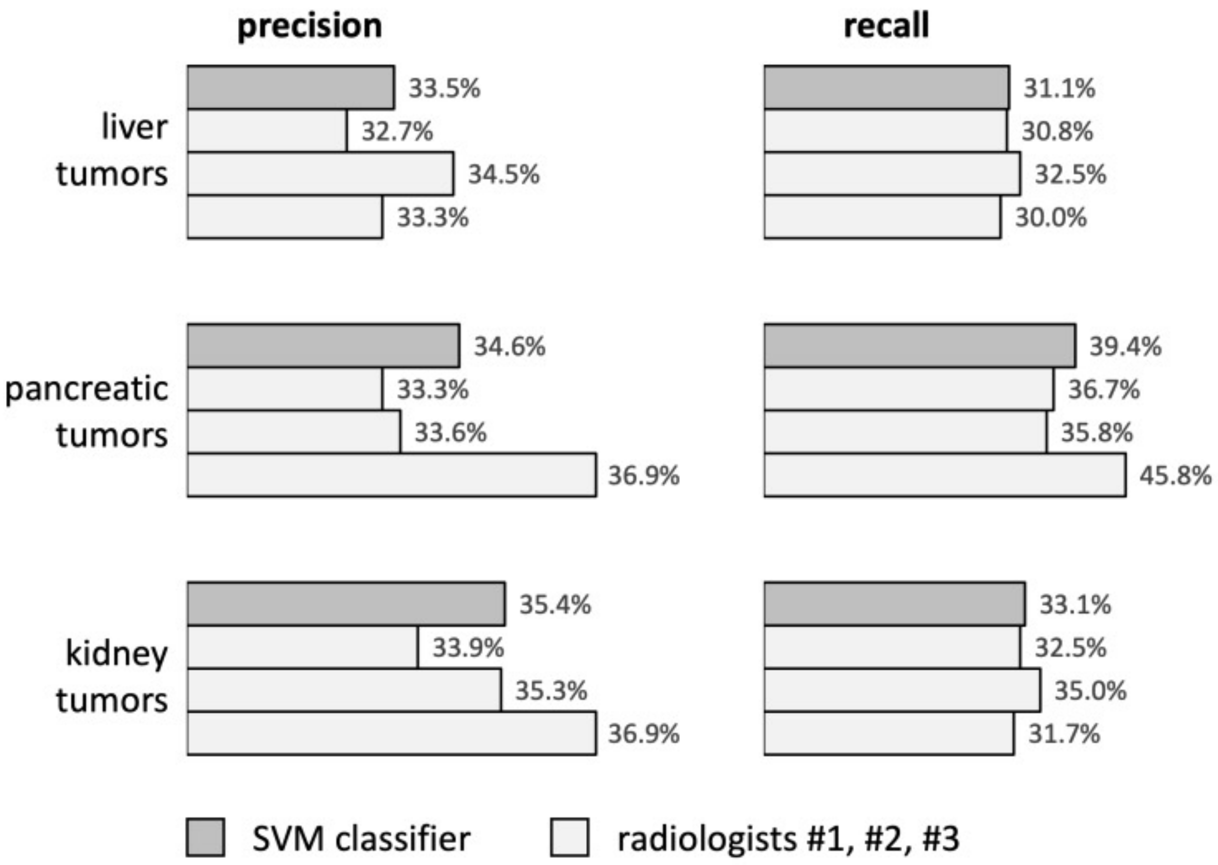
Concept Proof - Reader Study

- Early-stage tumors (< 2cm) tend to have similar imaging characteristics in computed tomography (CT).



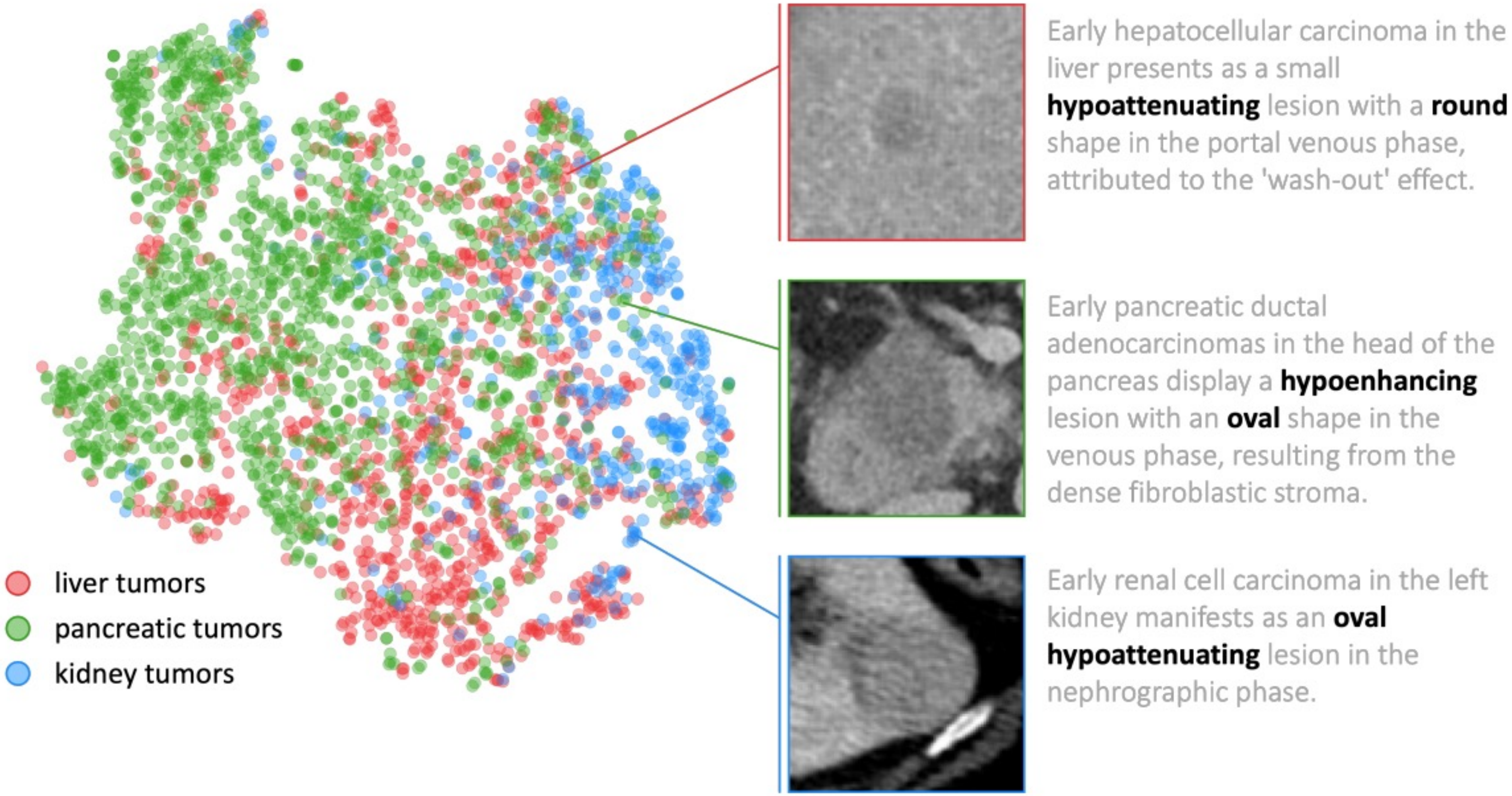
Concept Proof – Feature Analysis - SVM

- Early-stage tumors (< 2cm) tend to have similar imaging characteristics in computed tomography (CT).



Concept Proof – Feature Analysis – t-SNE

- Early-stage tumors (< 2cm) tend to have similar imaging characteristics in computed tomography (CT).



Main Results - Generalizable to Different Organs

Early-stage tumor detection performance (tumor-wise Sensitivity %).

source \ target		liver	pancreas	kidneys
liver	real tumors	75.6	0	2.4
	Hu *et al.* [37]	77.8	56.3	52.4
	DiffTumor	**82.2**	**56.3**	**76.2**
pancreas	real tumors	0.7	64.3	0
	Hu *et al.* [37]	74.1	67.0	52.4
	DiffTumor	**75.3**	**71.4**	**71.4**
kidney	real tumors	0.1	0	50.0
	Hu *et al.* [37]	**74.1**	56.3	66.7
	DiffTumor	68.8	**61.6**	**78.6**

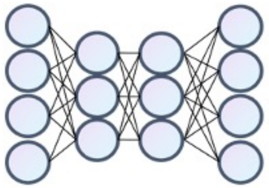
3/3 Generalization on Different Demographics

- Region
- Age
- Gender
- …

Generalizable Tumor Synthesis across Demographics

Training phase

Tumor data from
one hospital

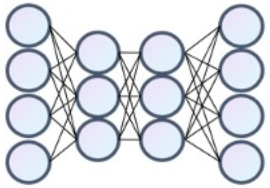


Segmentation Model

tumor mask

Inference phase

Tumor data from
another hospital

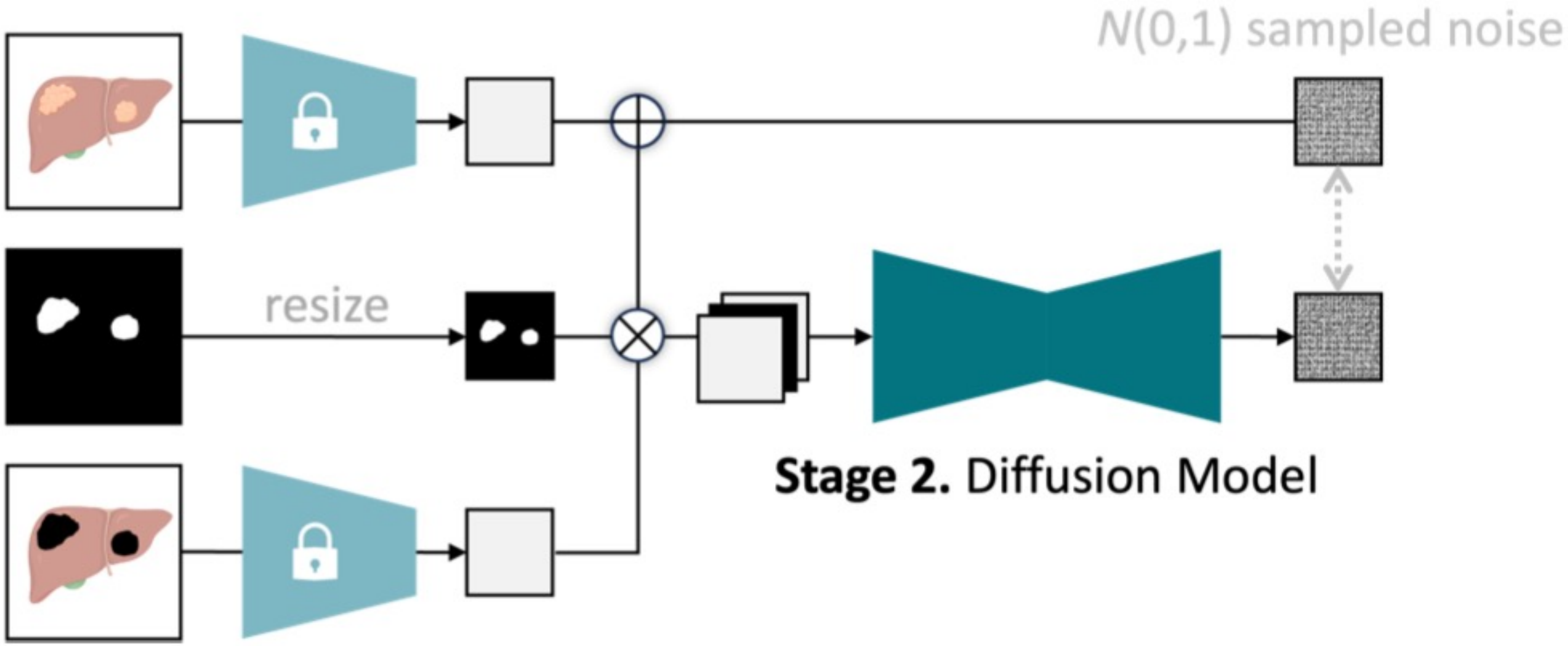


Segmentation Model

?

Learning to Generate Tumor by Denoising

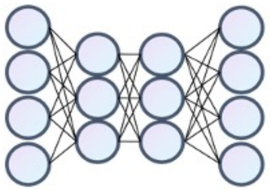
Inpainting task (mask and masked volume as condition)



Generalizable Tumor Synthesis across Demographics

Training phase

Tumor data from
one hospital

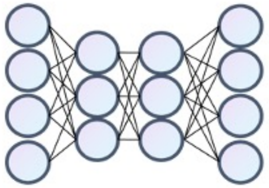


Generative Model

Inference phase

Large-scale
unlabeled CT

From different **hospital**
with different **demographics**



Generative Model

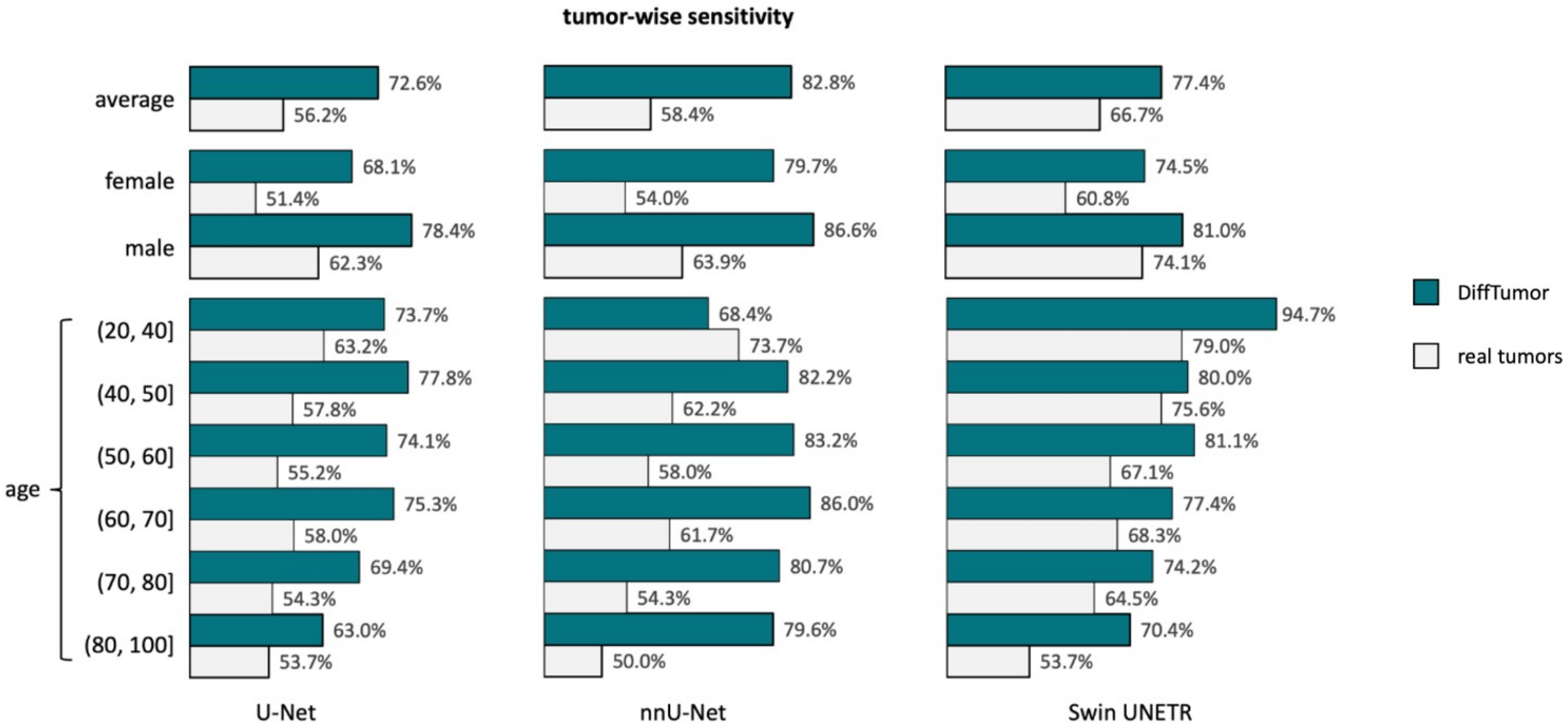
Liver Tumor

Pancreatic Tumor

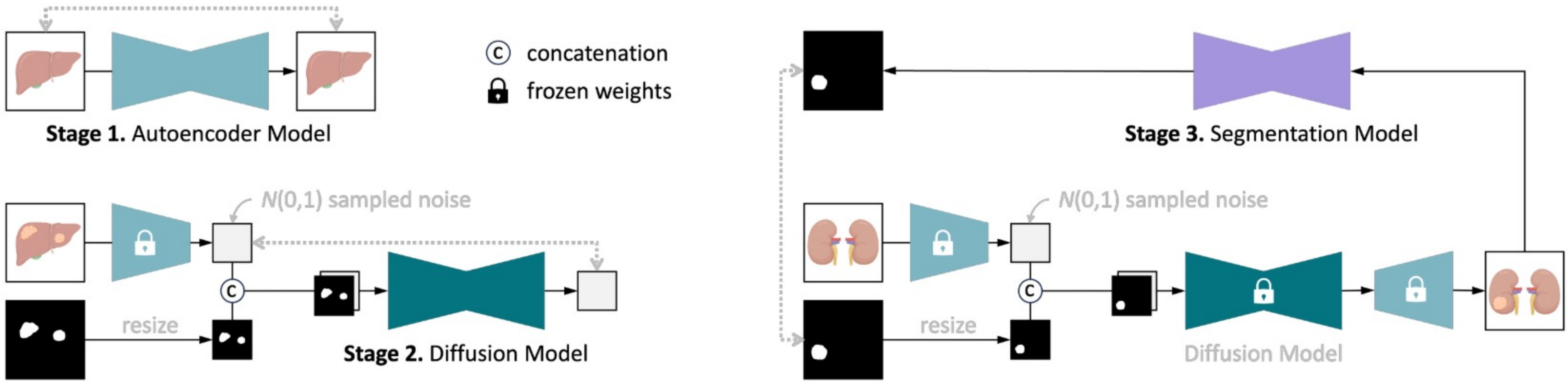
Kidney Tumor

…

Main Results - Generalizable to Different Demographics

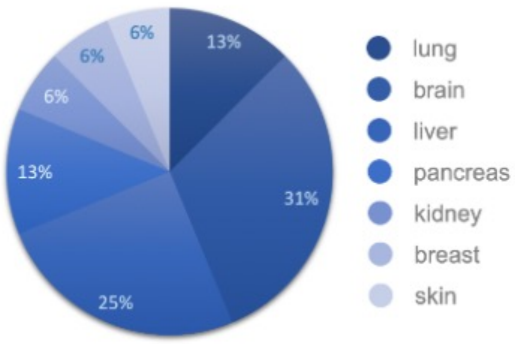


DiffTumor

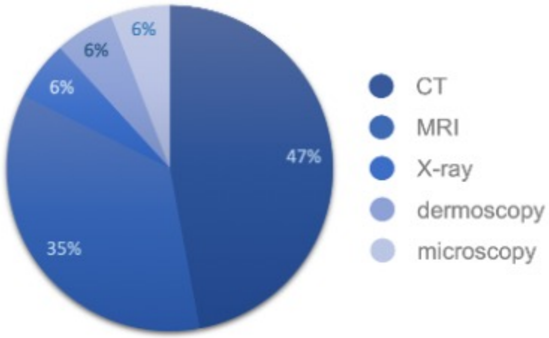

concatenation
frozen weights
Stage 1. Autoencoder Model
N(0,1) sampled noise
resize
Stage 2. Diffusion Model
Stage 3. Segmentation Model
N(0,1) sampled noise
resize
Diffusion Model

Future Thought 1

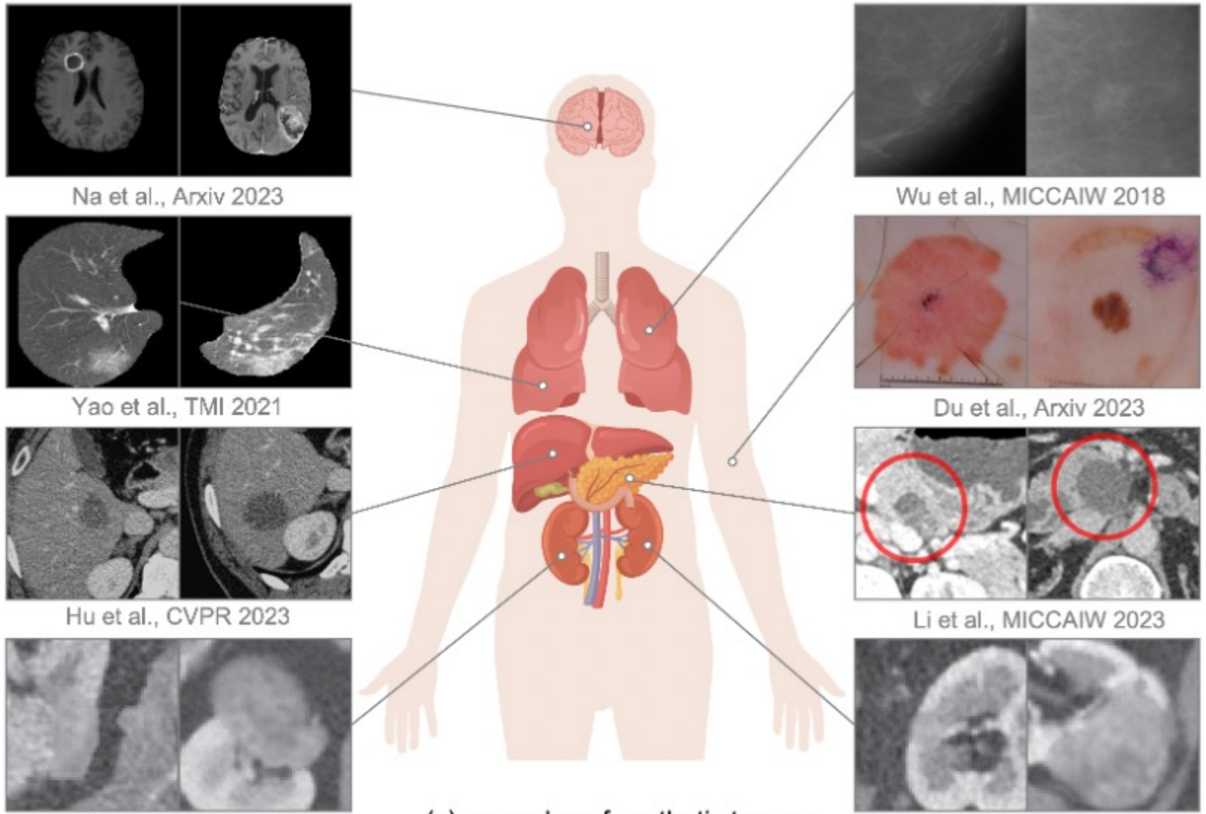
- More tumor types & More image modality (CT, MRI, X-ray)



(a) tumor synthesis in different organs



(b) tumor synthesis in different modalities



(c) examples of synthetic tumors

Future Thought 2

- More effective synthetic data (Data filter/Extra control)

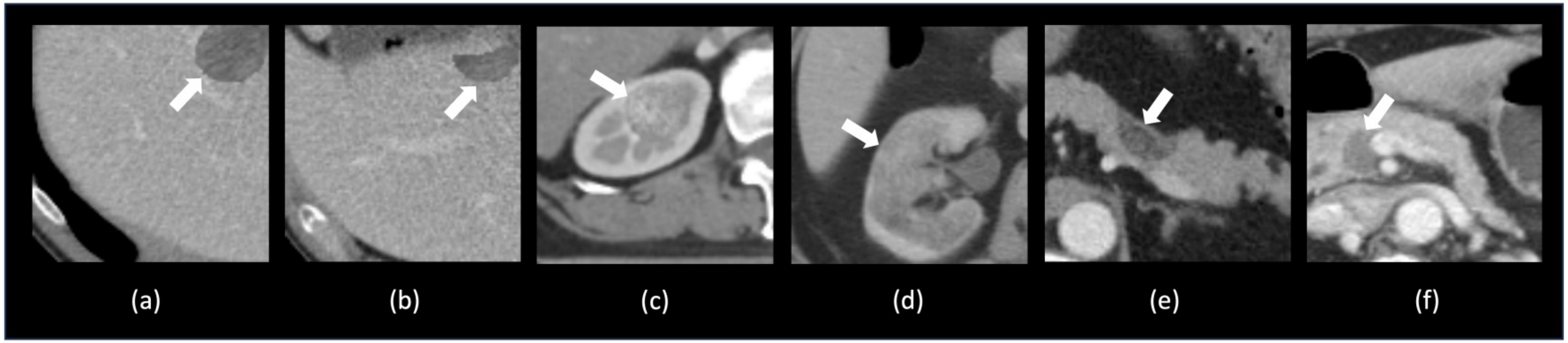
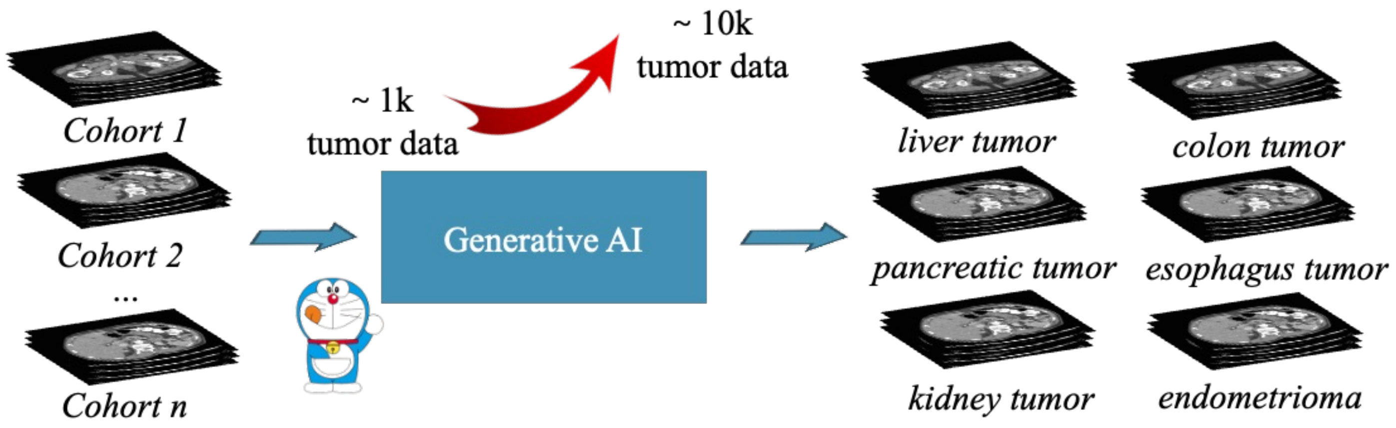
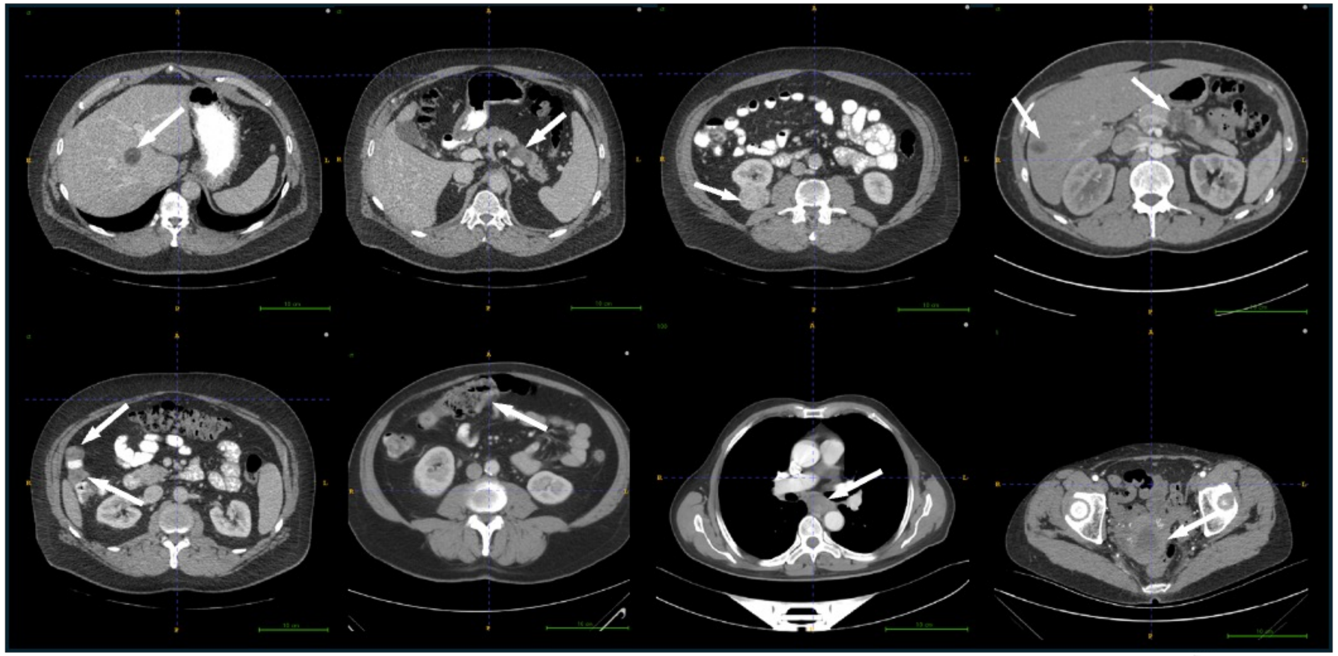


Figure 15. **Unrealistic generation cases.** **(a)** A synthetic liver tumor with an edge that is too sharp for a malignant tumor, and the inner noise has a discrepancy from the surrounding tissue. **(b)** A synthetic tumor with a sickle shape, which is unrealistic for a liver tumor. **(c)** A synthetic kidney tumor in the corticomedullary junction zone exhibiting a round nodular shape. However, this lesion fails to display a mass effect, as the healthy surrounding renal structure shows no deformation due to the tumor's inherent volume. Additionally, the texture and noise inside the tumor do not match the CT background. **(d)** A synthetic kidney tumor has a shape that matches the kidney "perfectly." However, solid tumors tend to grow expansively into a round nodular shape rather than precisely overlapping with the organ. **(e)** A synthetic tumor with a sickle shape, which is unrealistic for a pancreatic tumor. Additionally, the attenuation of this lesion is too low for a solid pancreatic tumor. **(f)** A synthetic pancreatic tumor adjacent to extra-pancreatic vessels. This tumor shows no mass effect, leaving the vessels without displacement or infiltration.

AbdomenAtlasX

- More tumor types
- Data filter





Thanks & QA

Qi Chen
qic@mail.ustc.edu.cn

Master of USTC