上海交通大學 SHANGHAI JIAO TONG UNIVERSITY
医疗机器人研究院 Institute of Medical Robotics

Pre-Foundation-Model-Era:
What can we study from the paradigm of self-supervised learning?

Yun Gu

Department of Automation
Institute of Medical Robotics
Shanghai Jiao Tong University

Tuning a foundation model is popular!

Google Scholar — MEDICAL

Articles — About 375 results (**0.05** sec)

Kirillov, Alexander, et al. "Segment Anything." (2023).

758 Citations by Oct 12th

>**300 “Medical” Citations**

- Any time
- Since 2023
- Since 2022
- Since 2019
- Custom range...

- Sort by relevance
- Sort by date

Create alert

Segment anything

Search within citing articles

Segment anything model for medical image analysis: an experimental study
MA Mazurowski, H Dong, H Gu, J Yang, N Konz… - **Medical** Image …, 2023 - Elsevier
… , **medical** image domains pose their own set of challenges. Here, we perform an extensive evaluation of SAM's ability to segment **medical** images on a collection of 19 **medical** … **medical** …
Save Cite Cited by 35 Related articles All 2 versions

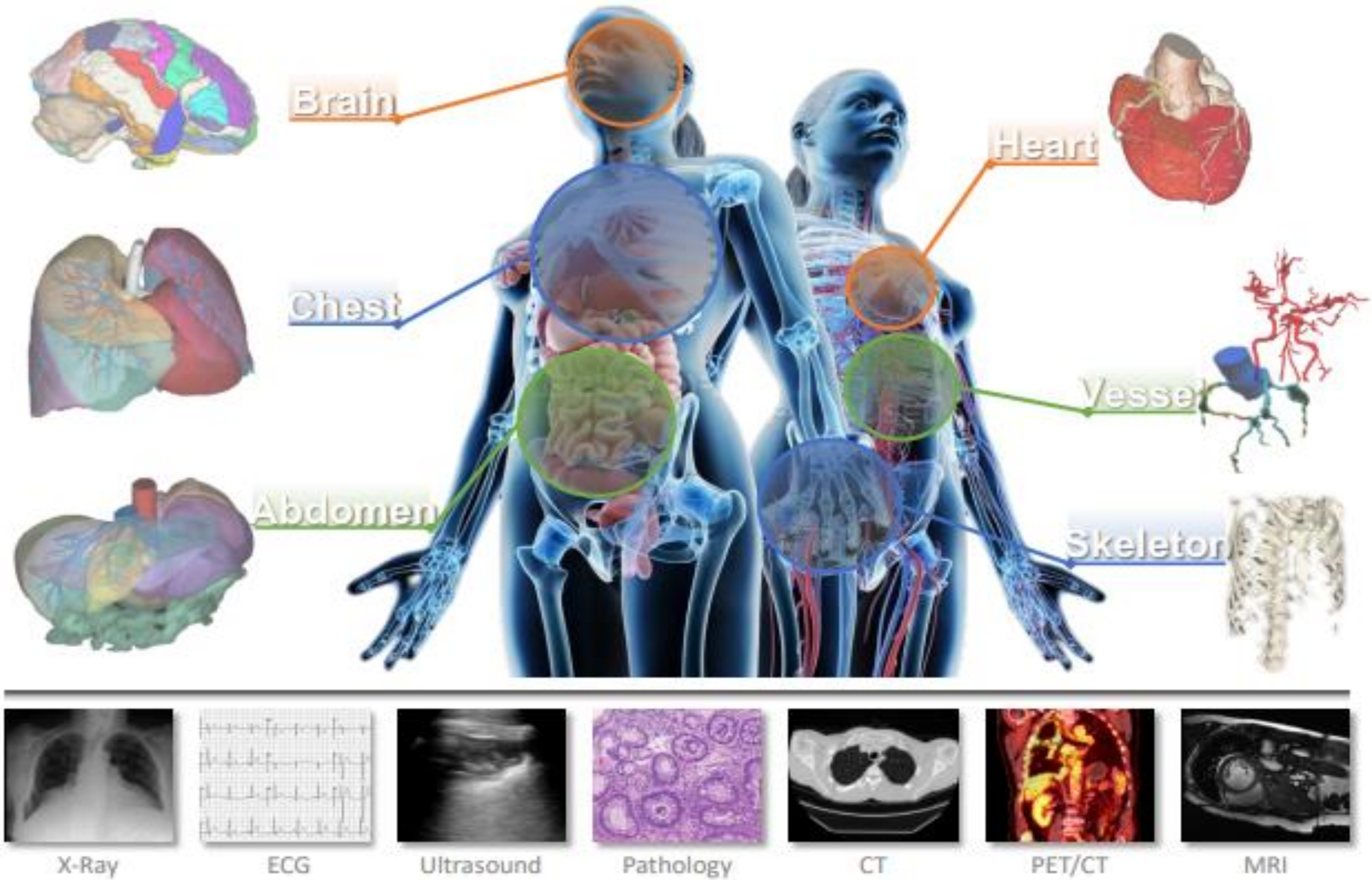
Segment anything in medical images — [PDF] arxiv.org
J Ma, B Wang - arXiv preprint arXiv:2304.12306, 2023 - arxiv.org
… of SAM to **medical** images, with the goal of creating a universal tool for the segmentation of various **medical** targets. Specifically, we first curate a large-scale **medical** image dataset, …
Save Cite Cited by 114 Related articles All 2 versions

Medical sam adapter: Adapting segment anything model for medical image segmentation — [PDF] arxiv.org
J Wu, R Fu, H Fang, Y Liu, Z Wang, Y Xu, Y Jin… - arXiv preprint arXiv …, 2023 - arxiv.org
… A **medical** image adapted SAM, which we have dubbed **Medical** SAM Adapter (MSA), shows superior performance on 19 **medical** image segmentation tasks with various image …
Save Cite Cited by 49 Related articles All 2 versions

[HTML] Large Language Models in Medical Education: Opportunities, Challenges, and Future Directions — [HTML] jmir.org
A Abd-Alrazaq, R AlSaad, D Alhuwail… - JMIR **Medical** …, 2023 - mededu.jmir.org
… (GPT) series, into **medical** education has the potential to … promise for revolutionizing **medical** curriculum development, … , and copyright concerns in **medical** education. As we navigate the …
Save Cite Cited by 30 Related articles All 5 versions

Segment anything model for medical images? — [PDF] arxiv.org
Y Huang, X Yang, L Liu, H Zhou, A Chang… - arXiv preprint arXiv …, 2023 - arxiv.org
… objects), so as to better develop a new generation of general **medical** segmentation models. In this report, we build a large **medical** image dataset named COSMOS 553K including …
Save Cite Cited by 34 Related articles All 3 versions

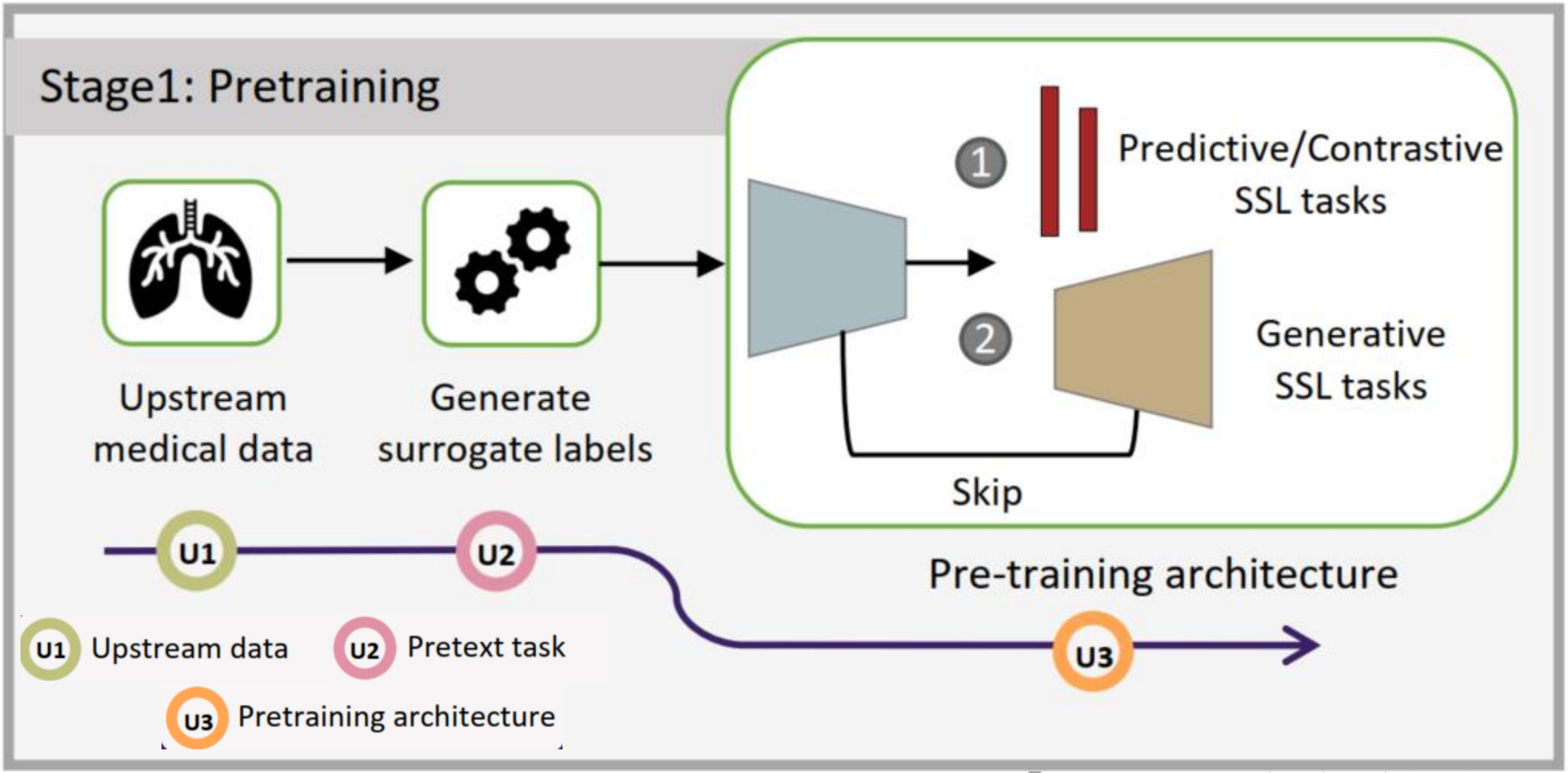
Foundation Models in Medical Imaging



S. Zhang et.al, *ON THE CHALLENGES AND PERSPECTIVES OF FOUNDATION MODELS FOR MEDICAL IMAGE ANALYSIS, 2023*

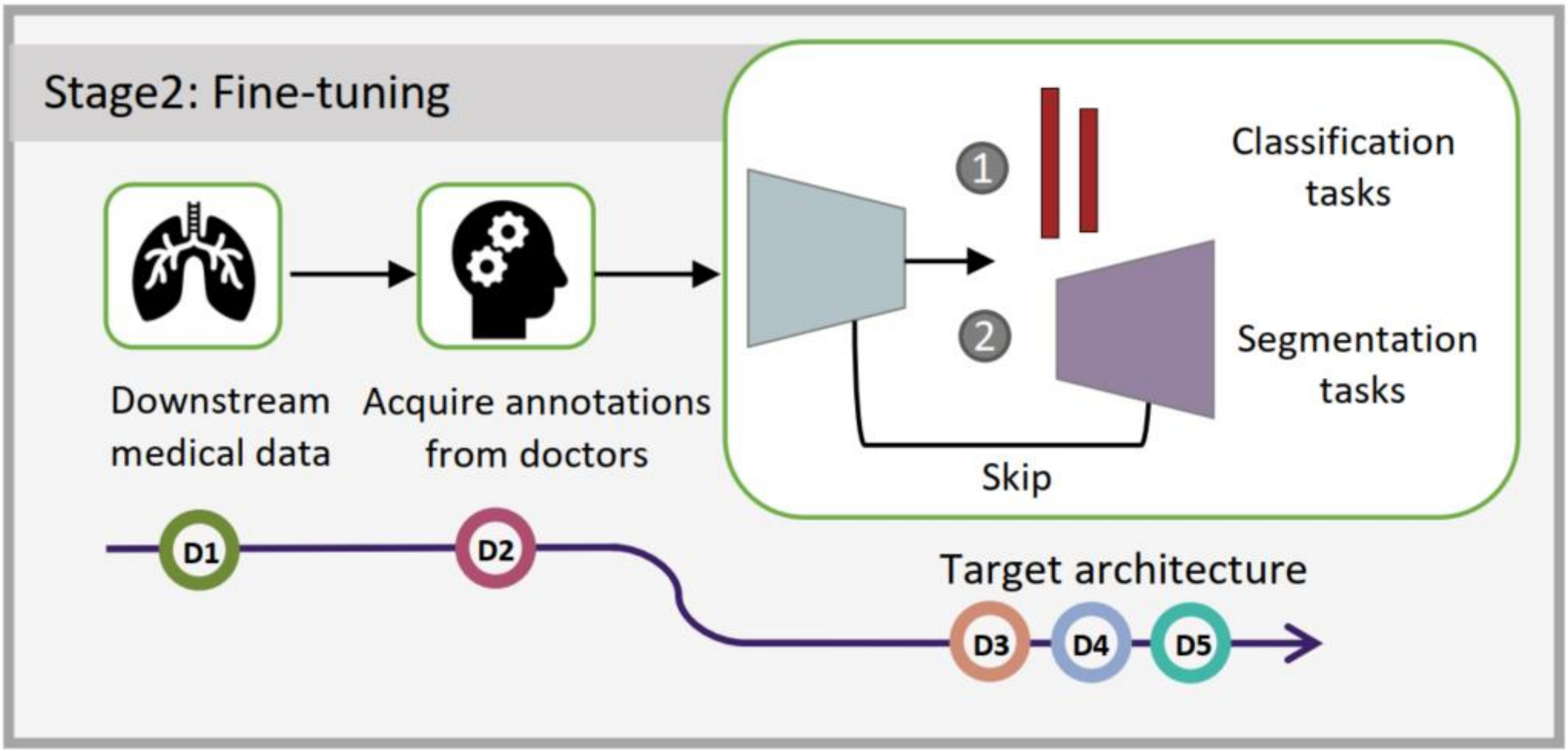
Pretrain-then-Finetune

- **Pretraining allows good initialization of model/feature representations via:**
  - Fully-supervised Pretraining
  - **Partially-supervised Pretraining**
  - **Unsupervised/Self-Supervised Pretraining**



Pretrain-then-Finetune

- **Finetuning aligns the knowledge of pretrained models and down-stream tasks by considering:**
  - **Data Domain Gap**
  - **Task Domain Gap**
  - **…**



Pretrain-then-Finetune

Key problems

- How to pretrain a model?
- How to pick the best pretrained model?
- How to fine-tune the pretrained model?

Guidance

- **Dive into the Details of Self-supervised Learning**

Pretraining

- Partially-Supervised Pretraining via ST-UNet, *arxiv, 2023*

Transferability

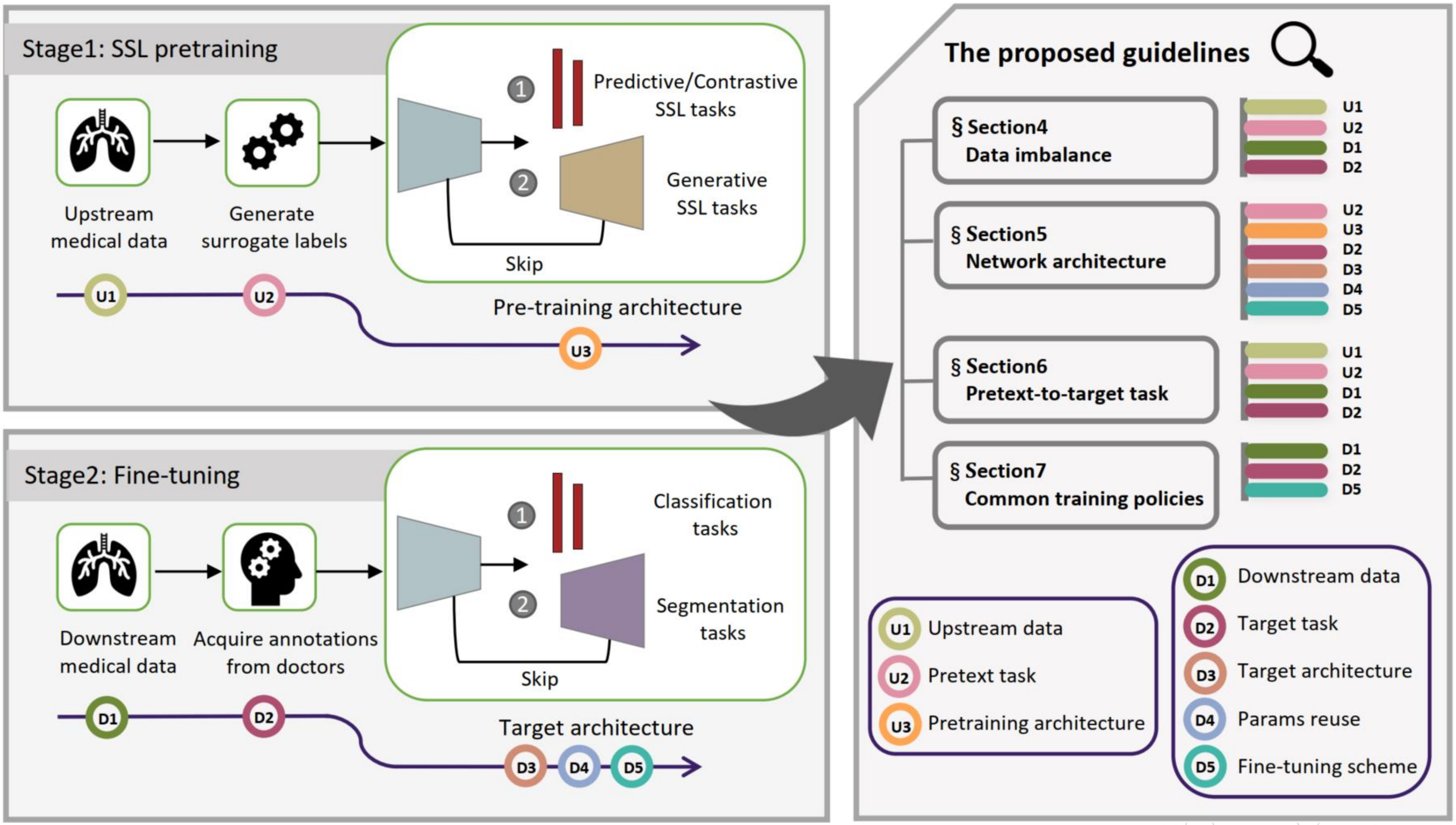
- Class-Consistency and Feature Variety, *MICCAI, 2023*

Finetuning

- These are open opportunities to you again!

Dive into Self-supervised Learning

- **Empirical Analysis:**
  - **Is Self-supervised Learning (SSL)** good for medical image analysis?



Chuyan Zhang et.al, Dive into the Details of Self-supervised Learning for Medical Image Analysis, ***Medical Image Analysis***, 2023

Dive into Self-supervised Learning

- **We have seen dozens of works on SSL for Medical Image Analysis**

Research topic	Literature
	Li et al. (2021), Zhu et al. (2020), Bai et al. (2019), Tan et al. (2021)
Predictive SSL	Blendowski et al. (2019), Zhuang et al. (2019), Taleb et al. (2020) Nguyen et al. (2020), Tajbakhsh et al. (2019), Ouyang et al. (2020)
Generative SSL	Tajbakhsh et al. (2019), Chen et al. (2019), Zhou et al. (2021b) Tao et al. (2020), Zhao et al. (2021), Xu and Adalsteinsson (2021)
Contrastive SSL	Sowrirajan et al. (2021), Sriram et al. (2021), Vu et al. (2021) Azizi et al. (2021), Chaitanya et al. (2020), Zhou et al. (2020)
Multi-SSL	Haghighi et al. (2020), Zhang et al. (2021), Dong et al. (2021) Zhou et al. (2021a), Taher et al. (2022), Haghighi et al. (2022)
–	Xu (2021), Chowdhury et al. (2021) Shurrab and Duwairi (2022), Chen et al. (2022)

Chuyan Zhang et.al, Dive into the Details of Self-supervised Learning for Medical Image Analysis, ***Medical Image Analysis***, 2023

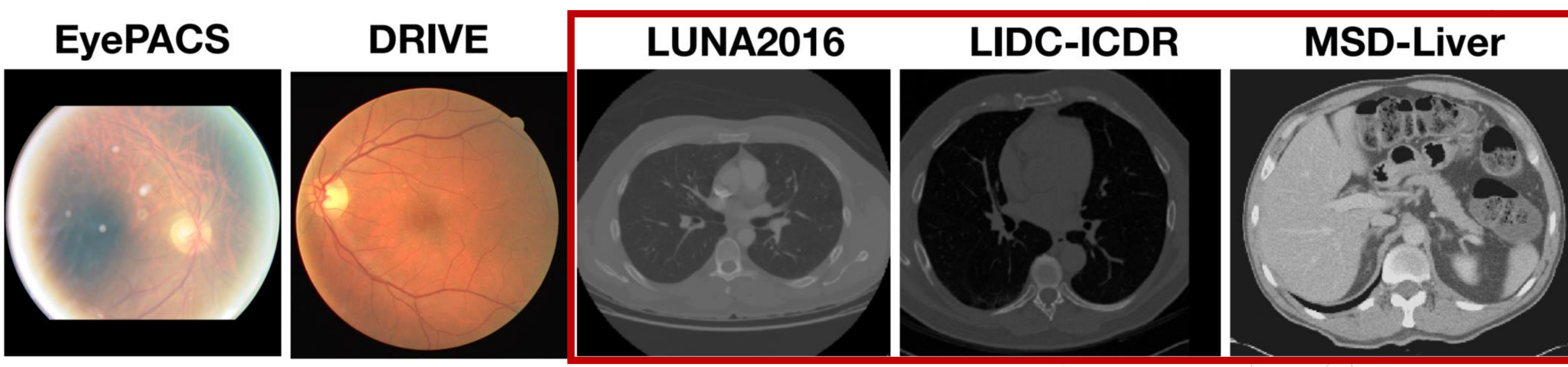
Dive into Self-supervised Learning

- **Fair comparisons with our benchmark:**

https://github.com/EndoluminalSurgicalVision-IMR/Medical-SSL

(a) 3D tasks: pretraining on LUNA 2016

Pretraining	Method	NCC		NCS			LCS	
		AUC (%)	ACC (%)	DSC (%)	mIoU (%)	mIoU+ (%)	DSC (%)	IoU (%)
From scratch	Random init (He et al., 2015)	98.53	97.00	73.79	80.86	62.58	93.83	88.50
Predictive SSL	ROT (Taleb et al., 2020)	99.32	**99.09**	73.25	80.64	62.15	94.49	89.65
	RPL (Taleb et al., 2020)	99.29	97.87	76.10	81.36	63.51	94.86	90.18
	Jigsaw (Taleb et al., 2020)	98.64	98.81	75.23	81.90	64.66	94.36	89.42
	RKB (Zhuang et al., 2019)	99.41	99.03	74.22	80.88	62.61	94.93	90.43
	RKB+ (Zhu et al., 2020)	98.88	98.48	74.31	80.82	62.51	**95.46**	**91.41**
Generative SSL	AE	97.76	97.62	74.26	81.15	63.17	93.77	88.92
	MG (Zhou et al., 2021b)	98.01	96.72	75.69	82.13	65.08	94.24	89.17
	PCRL (Zhou et al., 2021a)	98.97	97.88	75.60	81.97	64.55	93.87	88.56
Contrastive SSL	SimCLR (Chen et al., 2020b)	99.29	97.91	75.96	82.00	64.82	94.56	89.74
	BYOL (Grill et al., 2020)	**99.52**	97.65	**76.13**	**82.27**	**65.37**	94.43	89.56



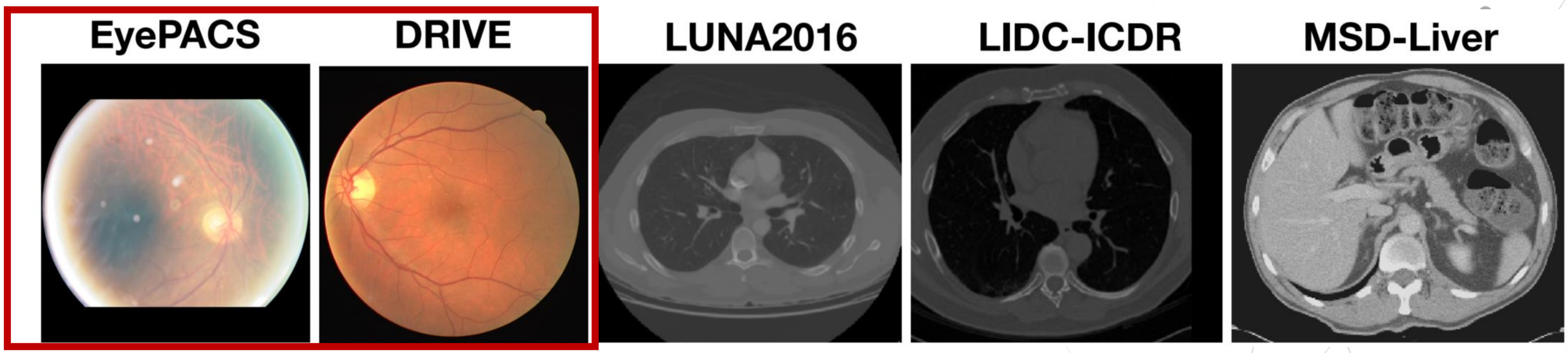
Dive into Self-supervised Learning

- **Fair comparisons with our benchmark:**

https://github.com/EndoluminalSurgicalVision-IMR/Medical-SSL

(b) 2D tasks: pretraining on EyePACS

Pretraining	Method	EPC		DVS	
		Kappa (%)	ACC (%)	DSC (%)	SEN(%)
From scratch	Random init (He et al., 2015)	76.32	77.33	78.57	75.45
Predictive SSL	ROT (Taleb et al., 2020)	77.85	77.16	77.56	75.67
	RPL (Taleb et al., 2020)	77.47	77.63	79.21	79.05
	Jigsaw (Taleb et al., 2020)	78.29	76.76	78.72	77.89
Generative SSL	AE	70.28	74.33	79.28	77.23
	MG (Zhou et al., 2021b)	**78.66**	**78.41**	80.00	78.63
	PCRL (Zhou et al., 2021a)	77.92	78.34	79.49	76.93
Contrastive SSL	SimCLR (Chen et al., 2020b)	78.46	77.91	**80.75**	**79.81**
	BYOL (Grill et al., 2020)	76.55	77.58	78.42	77.33



Dive into Self-supervised Learning

- **Detail ONE: Imbalanced DATA in medical imaging**

Example: Nodule Classification with LUNA16 (Normal v.s. Malignant: 1:100)

Question: How does SSL pretraining affect class-imbalanced learning?

Takeaway 1

- The prior probability ratio related to both upstream and downstream data affects the target performance, namely the inherent data imbalance of pretraining data would impair downstream learning.

Takeaway 2

- SSL methods improve the rare class more than the frequent class

Chuyan Zhang et.al, Dive into the Details of Self-supervised Learning for Medical Image Analysis, ***Medical Image Analysis***, 2023

Dive into Self-supervised Learning

- **Detail ONE: Imbalanced DATA in medical imaging**

Example: Nodule Classification with LUNA16 (Normal v.s. Malignant: 100:1)

Question: How does SSL pretraining affect class-imbalanced learning?

With that said,, we can assume that Z conditioned on $Y = +1$ and $Y = -1$ follows different Gaussian distributions. Formally, $Z|Y = +1 \sim N(\mu_1, \sigma^2)$ and $Z|Y = -1 \sim N(\mu_2, \sigma^2)$. Note that we ignore the variance difference between the two classes for simplicity. Assuming $\mu_1 > \mu_2$, the optimal Bayesian classifier can be expressed as $f(Z) = sign(Z - \frac{\mu_1+\mu_2}{2} + \frac{\sigma^2 \ln(\lambda)}{\mu_1-\mu_2})$, i.e. X is classified as $Y = +1$ when $f(\phi(X)) > 0$. Our estimation of f is composed of two terms: $\theta_1 = \frac{\mu_1+\mu_2}{2}$ and $\theta_2 = \frac{\sigma^2 \ln(\lambda)}{\mu_1-\mu_2}$. For the former term, the estimation is naturally constructed as $\hat{\theta}_1 = \frac{\sum_{i=1}^{N^+} Z_i^+/N^+ + \sum_{i=1}^{N^-} Z_i^-/N^-}{2}$, where N^+ and N^- are the number of positive class and negative class in D_{down} respectively. According to the Gaussian concentration inequality, we have:

Theorem 1. *Consider the above setup. For any $t > 0$, with probability at least $1 - e^{\frac{-2t^2}{\sigma^2}\frac{N^+N^-}{N^++N^-}}$ our estimated $\hat{\theta}_1$ satisfies:*

$$|\hat{\theta}_1 - \frac{\mu_1 + \mu_2}{2}| \le t \qquad (13)$$

- Imbalance degree of labelled training data affects the chance of obtaining a good estimate.

We then consider a large λ, i.e. there is less class imbalance in upstream and downstream data. Based on the understanding in Interpretation 1, it still have a high probability to get an accurate estimate of the optimal classifier. We further seek the error rate of such a good classifier. Suppose that the Bayesian decision boundary splits the whole feature space into two parts: Γ_+ and Γ_-, the error probability of each class can be computed as: $\epsilon_+ = \int_{\Gamma_-} p(Z|Y = +1)dZ$, $\epsilon_- = \int_{\Gamma_+} p(Z|Y = -1)dZ$. Then, we can derive the *Chernoff* error upper bound according to Fukunaga (2013):

Theorem 2. *Consider the above setup. Given the ratio of prior probabilities λ and the Bhattacharyya distance of two classes $D_B = \frac{1}{8}\frac{(\mu_2-\mu_1)^2}{\sigma^2}$, the error probability of each class of Bayes's classifier satisfies:*

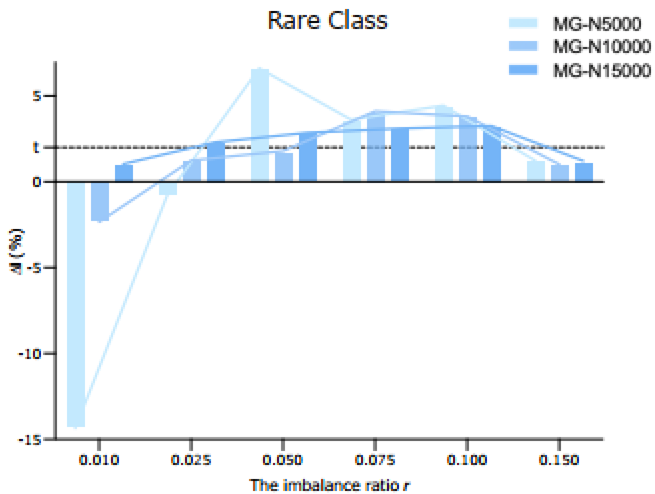
$$\begin{cases} \epsilon_+ \le \frac{1}{\sqrt{\lambda}} e^{-D_B} \\ \epsilon_- \le \sqrt{\lambda} e^{-D_B} \end{cases} \qquad (11)$$

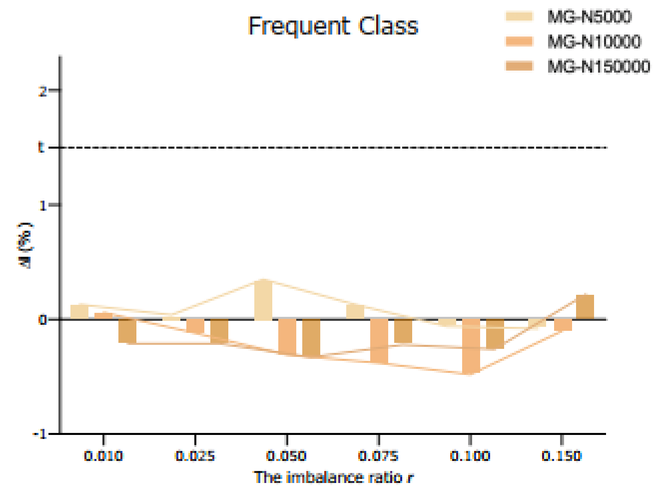
- SSL methods improve the rare class more than the major class.

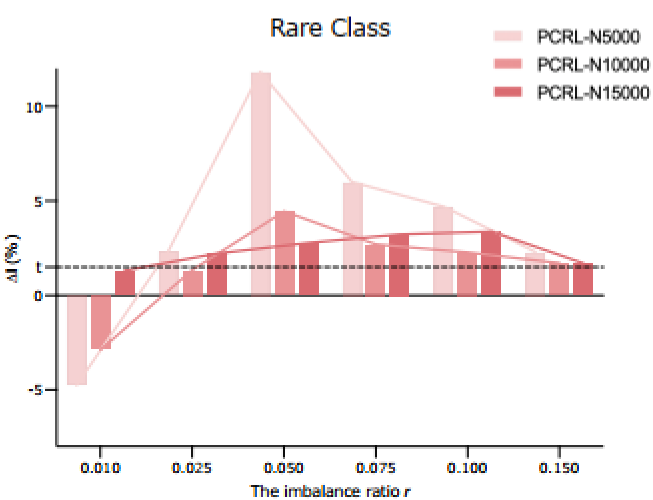
Dive into Self-supervised Learning

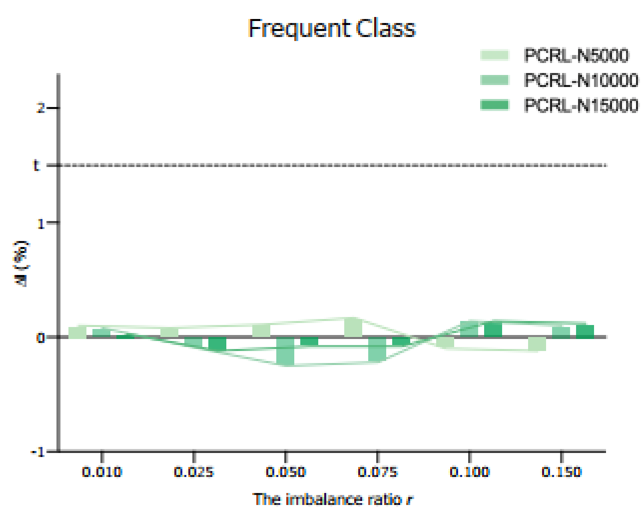
- **Detail ONE: Imbalanced DATA in medical imaging**

Nodule Classification Task (NCC) Based on *Model Genesis* (MedIA 21) and *PCRL* (TPAMI23)





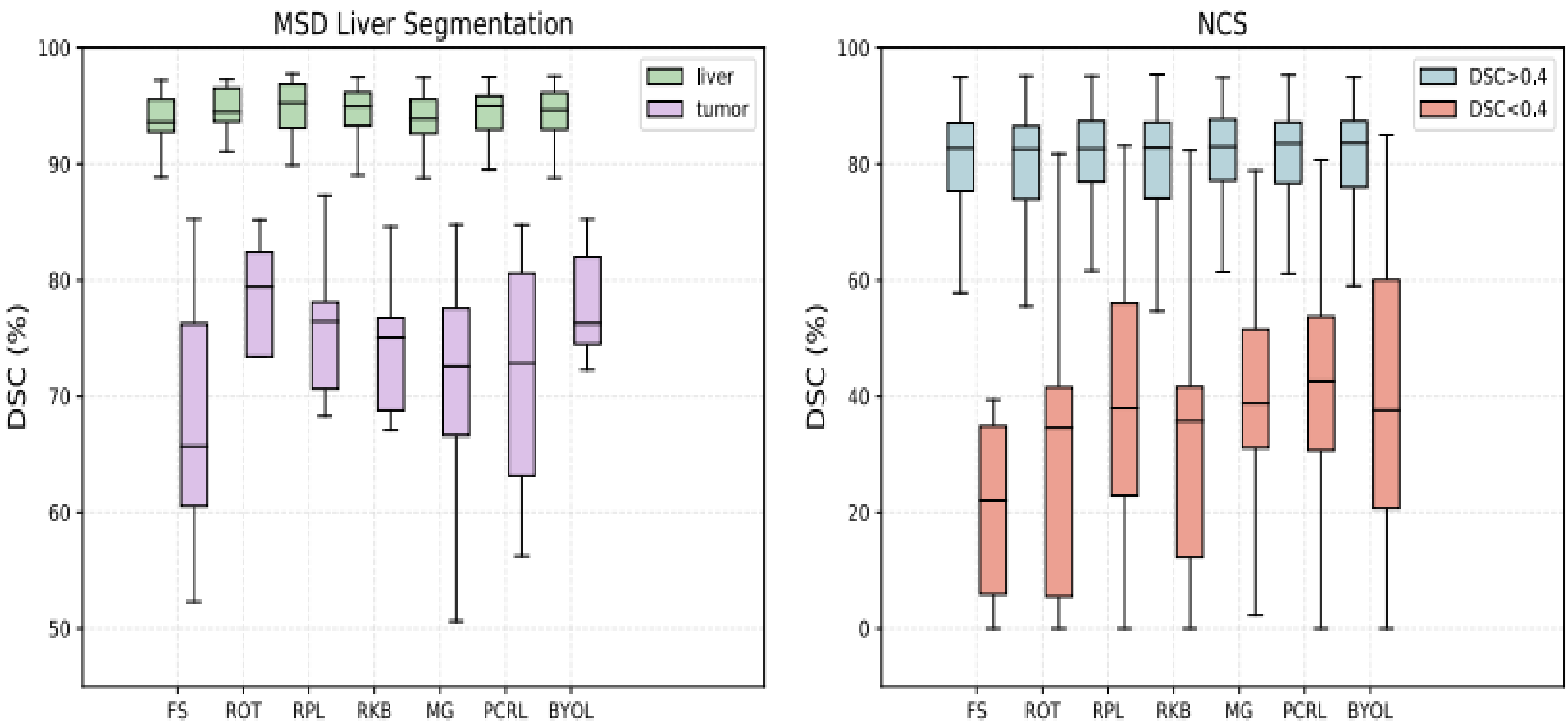




Dive into Self-supervised Learning

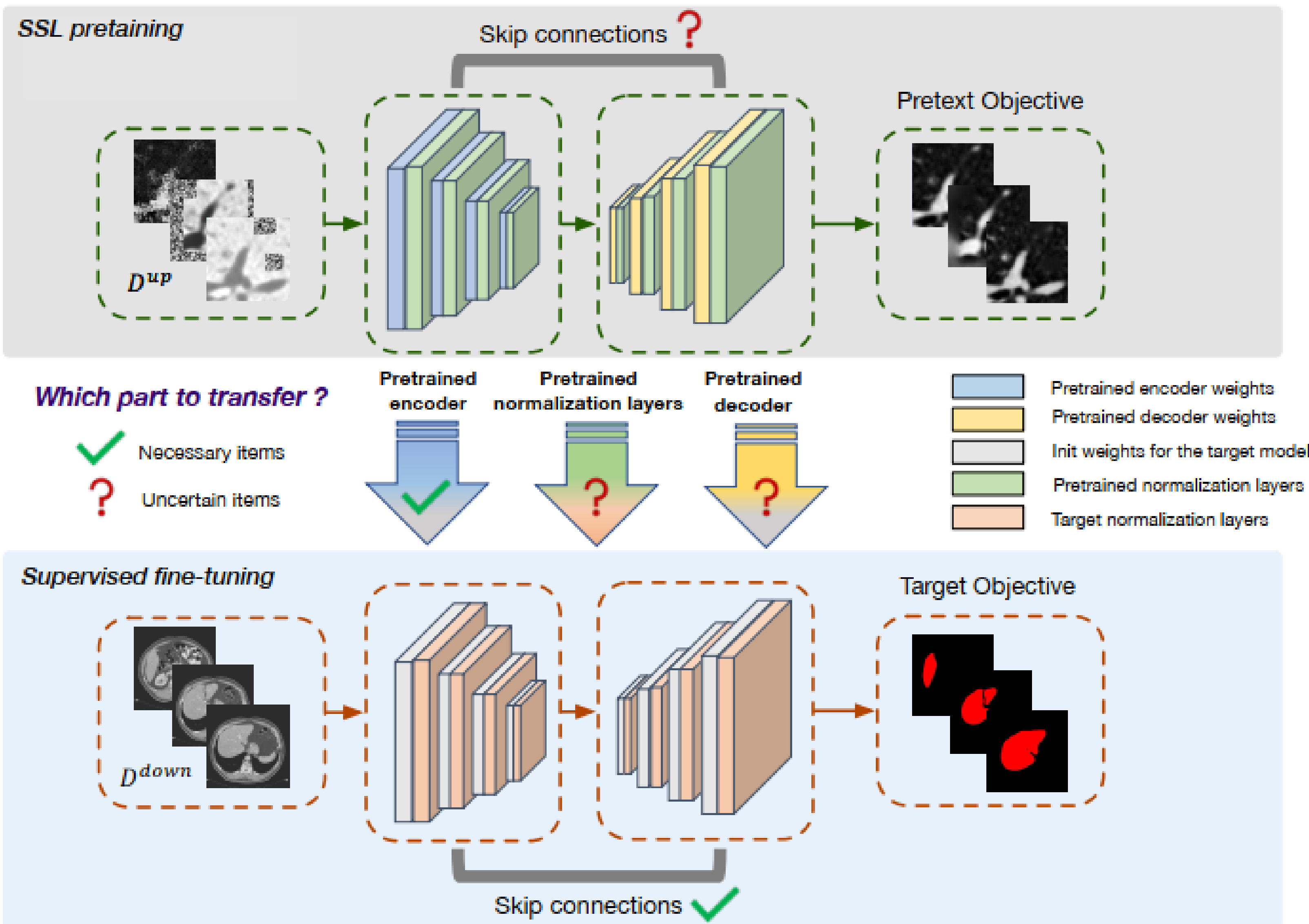
- **Detail ONE: Imbalanced DATA in medical imaging**

Node Segmentation/Liver Segmentation Tasks



Dive into Self-supervised Learning

- **Detail TWO: Network Modules and Designs**



Dive into Self-supervised Learning

- **Detail TWO: Network Modules and Designs**

Question: The impacts of modules in U-Shape Neural Networks

Takeaway 1

- The decoder risks overfitting to the reconstruction task, thus offering little benefits for downstream tasks. To mitigate this issue, removing the skip connections in pretraining could lead to better representations.

Takeaway 2

- The benefit of SSL pretraining may come from an alleviation of over-parameterization.

Takeaway 3

- Recollect the target BN statistics for inference.

Takeaway 4

- Full fine-tuning is more advantageous than warm-up finetuning.

Dive into Self-supervised Learning

- **Detail TWO: Network Modules and Designs**

Question: The impacts of modules in U-Shape Neural Networks

Model	Architecture			Pretext task	Target task		
	w/o encoder	w/o decoder	w/o skip	MAE	NCS	LCS	NCC
Model Genesis					75.69	94.24	98.01
		✓		0.0053	75.70	94.08	—
	✓				74.77	93.60	—
			✓	0.0095	77.09	94.72	99.00
PCRL					75.60	93.87	98.97
		✓		0.0402	74.02	93.25	—
	✓				74.43	93.56	—
			✓	0.0431	74.30	93.18	98.34

- The decoder risks overfitting to the reconstruction task, thus offering little benefits for downstream tasks. To mitigate this issue, removing the skip connections in pretraining could lead to better representations.

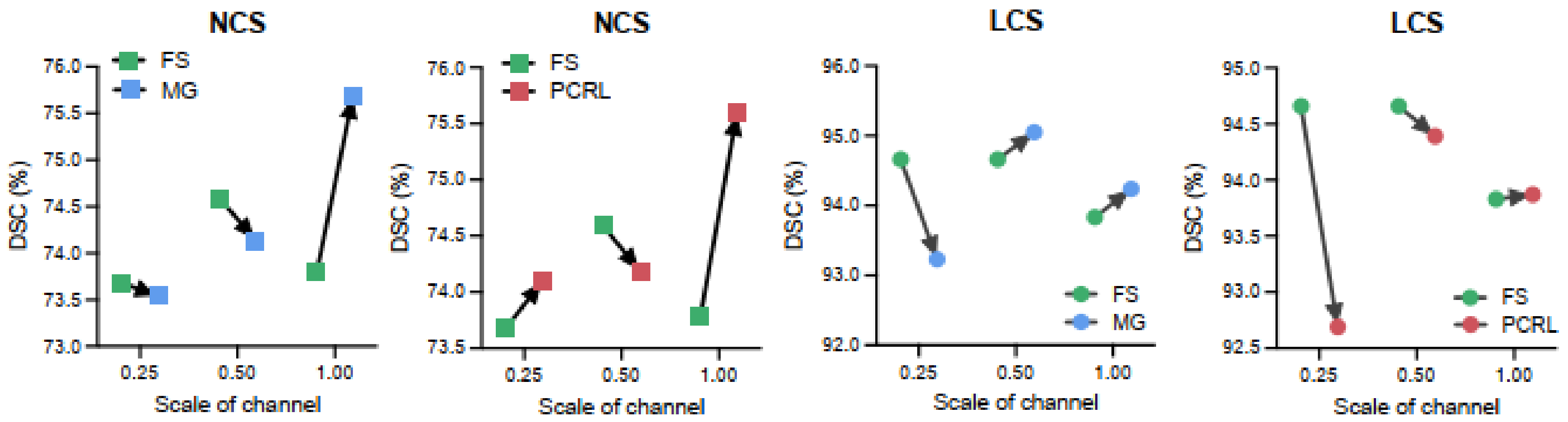
- The benefit of SSL pretraining may come from an alleviation of over-parameterization.

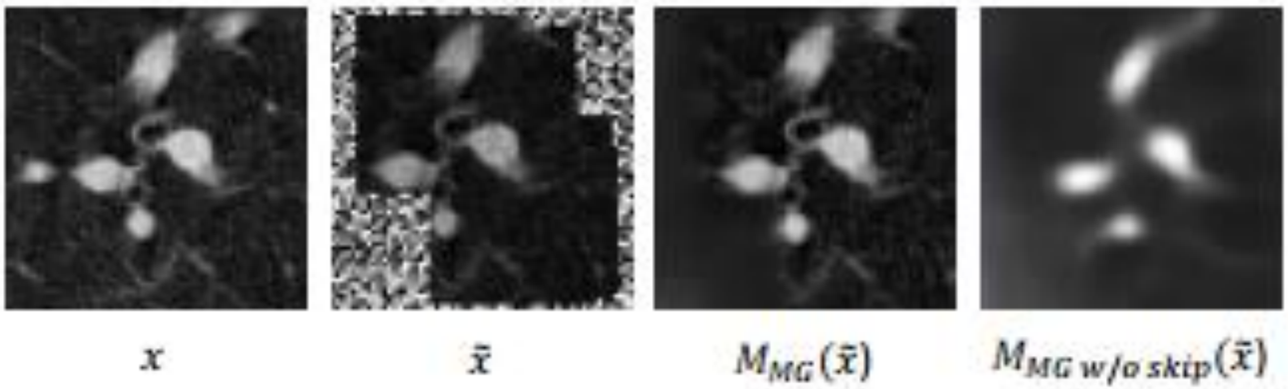
Dive into Self-supervised Learning

- **Detail TWO: Network Modules and Designs**

Question: The impacts of modules in U-Shape Neural Networks

Over-Param!



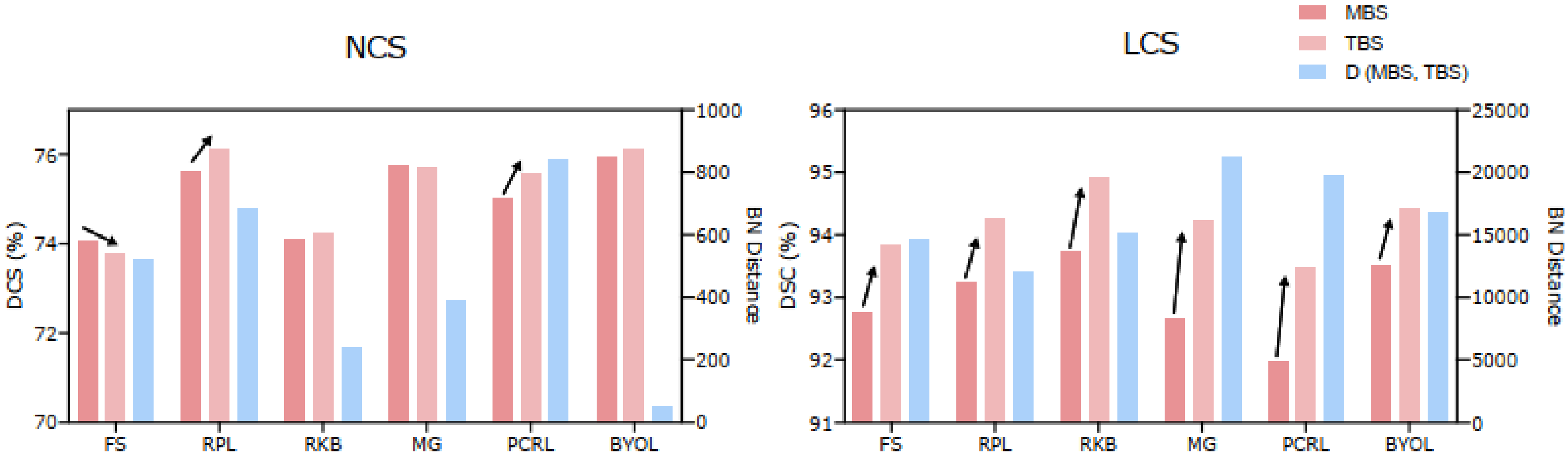


Skip-Connection!

Dive into Self-supervised Learning

- **Detail TWO: Network Modules and Designs**

  Question: The impacts of modules in U-Shape Neural Networks



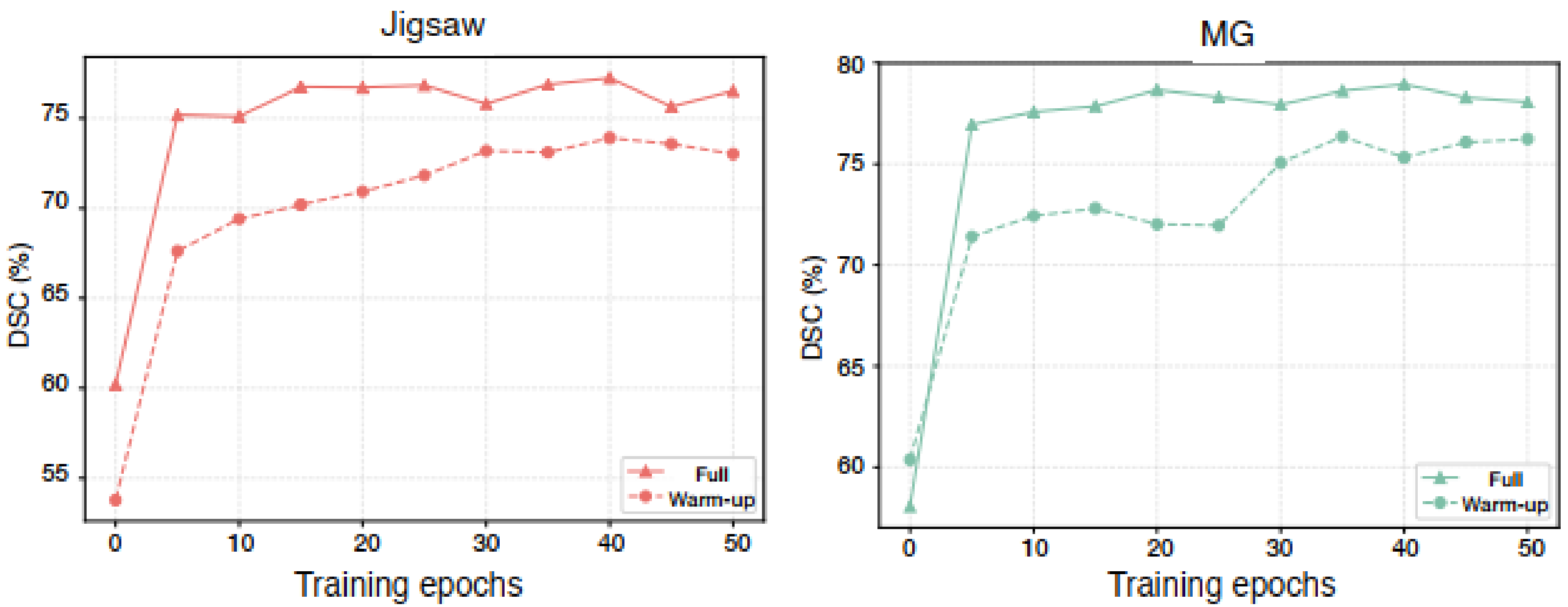
Mixed BN statistics usually induce performance degeneration at a large distribution distance so that recollecting the target BN statistics for inference is necessary

Chuyan Zhang et.al, Dive into the Details of Self-supervised Learning for Medical Image Analysis, ***Medical Image Analysis***, 2023

Dive into Self-supervised Learning

- **Detail TWO: Network Modules and Designs**

  Question: The impacts of modules in U-Shape Neural Networks



Full fine-tuning is more advantageous than warm-up finetuning due to the essential gap between self-training and downstream semantic tasks.

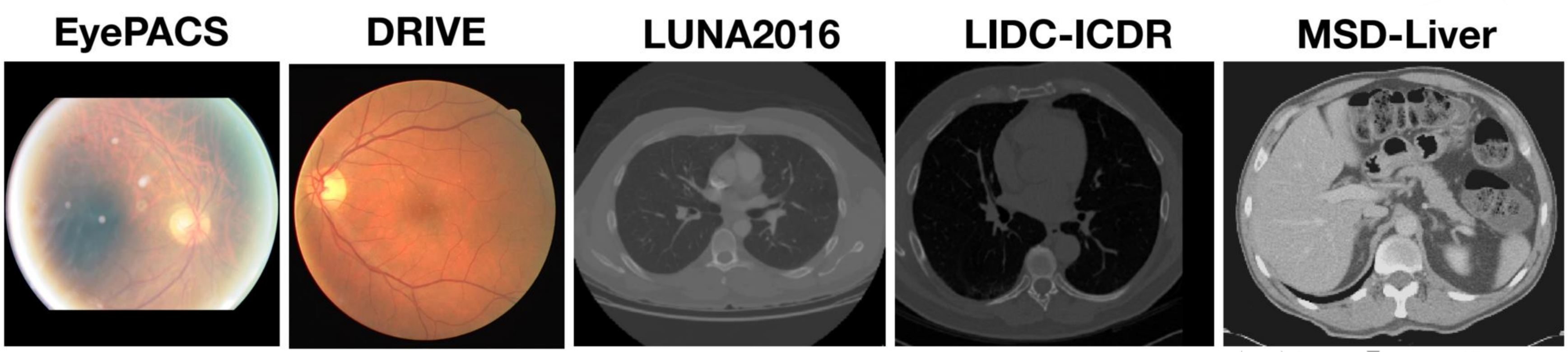
Chuyan Zhang et.al, Dive into the Details of Self-supervised Learning for Medical Image Analysis, ***Medical Image Analysis***, 2023

Dive into Self-supervised Learning

- **Detail THREE: Pretext-Target Task**

(a) 3D tasks: pretraining on LUNA 2016

Pretraining	Method	NCC		NCS			LCS	
		AUC (%)	ACC (%)	DSC (%)	mIoU (%)	mIoU+ (%)	DSC (%)	IoU (%)
From scratch	Random init (He et al., 2015)	98.53	97.00	73.79	80.86	62.58	93.83	88.50
Predictive SSL	ROT (Taleb et al., 2020)	99.32	**99.09**	73.25	80.64	62.15	94.49	89.65
	RPL (Taleb et al., 2020)	99.29	97.87	76.10	81.36	63.51	94.86	90.18
	Jigsaw (Taleb et al., 2020)	98.64	98.81	75.23	81.90	64.66	94.36	89.42
	RKB (Zhuang et al., 2019)	99.41	99.03	74.22	80.88	62.61	94.93	90.43
	RKB+ (Zhu et al., 2020)	98.88	98.48	74.31	80.82	62.51	**95.46**	**91.41**
Generative SSL	AE	97.76	97.62	74.26	81.15	63.17	93.77	88.92
	MG (Zhou et al., 2021b)	98.01	96.72	75.69	82.13	65.08	94.24	89.17
	PCRL (Zhou et al., 2021a)	98.97	97.88	75.60	81.97	64.55	93.87	88.56
Contrastive SSL	SimCLR (Chen et al., 2020b)	99.29	97.91	75.96	82.00	64.82	94.56	89.74
	BYOL (Grill et al., 2020)	**99.52**	97.65	**76.13**	**82.27**	**65.37**	94.43	89.56



Chuyan Zhang et.al, Dive into the Details of Self-supervised Learning for Medical Image Analysis, ***Medical Image Analysis***, 2023

Dive into Self-supervised Learning

- **Detail THREE: Pretext-Target Task**

Takeaway 1

- The generative SSL underperforms in the classification task.

Takeaway 2

- Take into account properties of the target task when selecting the pretext task for pretraining.

Takeaway 3

- For the adaptation under a large domain shift between upstream and downstream data (e.g. LCS), it is important to prevent potential overfitting to the upstream data in predictive or generative SSL.

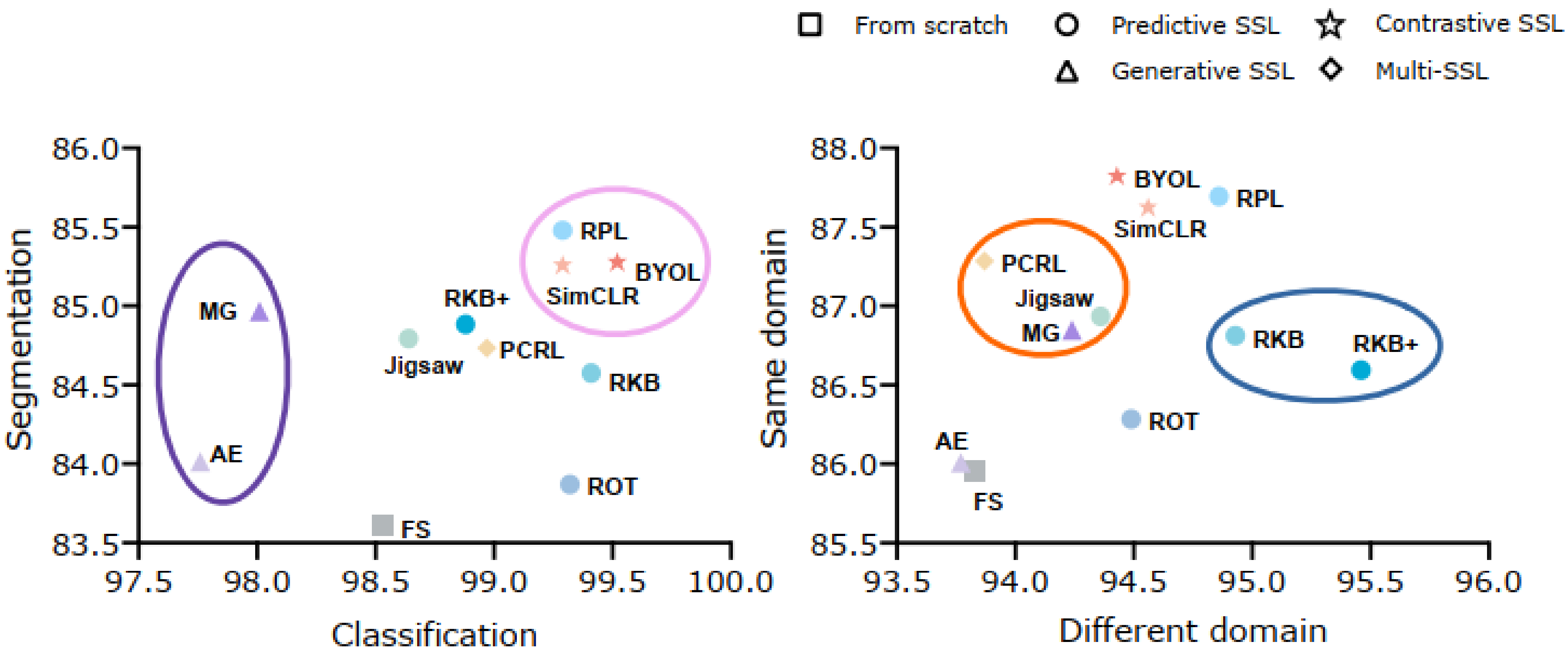
Takeaway 4

- Contrastive SSL and RPL could be the go-to solutions for researchers.

Chuyan Zhang et.al, Dive into the Details of Self-supervised Learning for Medical Image Analysis, ***Medical Image Analysis***, 2023

Dive into Self-supervised Learning

- **Detail THREE: Pretext-Target Task**



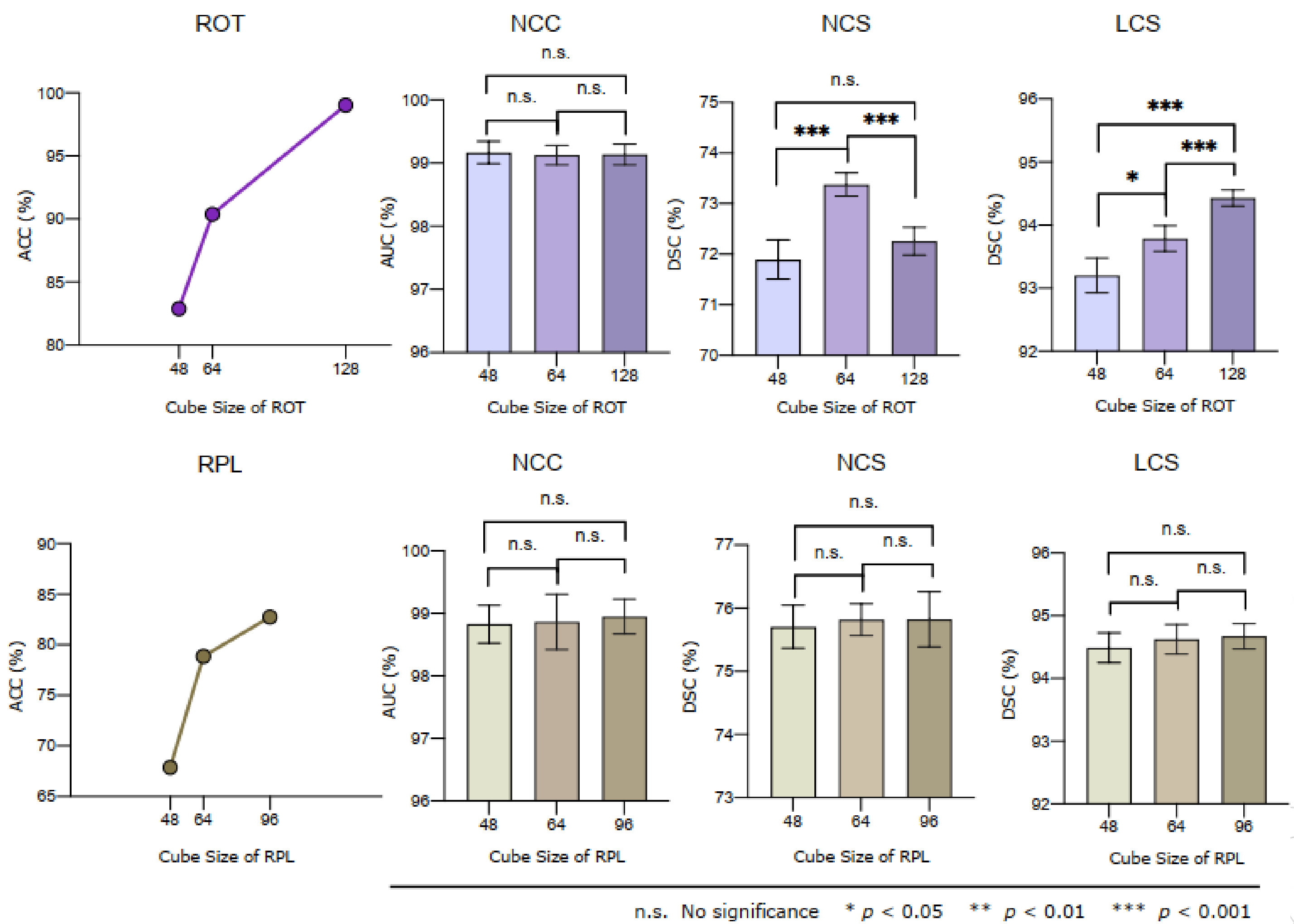
(a) Classification *vs.* Segmentation

(b) Inter-domain *vs.* Intra-domain

Chuyan Zhang et.al, Dive into the Details of Self-supervised Learning for Medical Image Analysis, ***Medical Image Analysis***, 2023

Dive into Self-supervised Learning

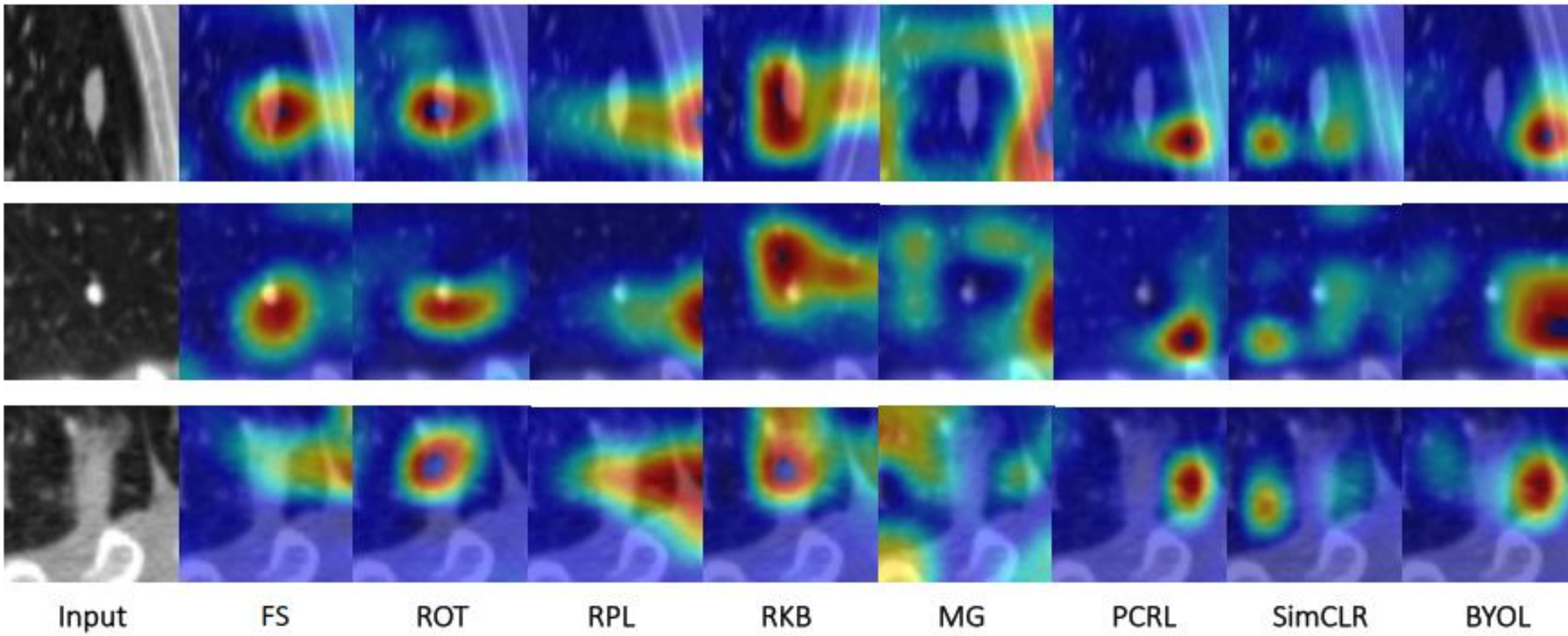
- **Detail THREE: Pretext-Target Task-Input Size Matters**



Chuyan Zhang et.al, Dive into the Details of Self-supervised Learning for Medical Image Analysis, ***Medical Image Analysis***, 2023

Dive into Self-supervised Learning

- **Detail THREE: Pretext-Target Task**



Chuyan Zhang et.al, Dive into the Details of Self-supervised Learning for Medical Image Analysis, ***Medical Image Analysis***, 2023

Dive into Self-supervised Learning

- **Detail FOUR: Common Training Policies**

Takeaway 1

- The combination of resampling and SSL techniques is preferred in severe class imbalance and low-data regimes, whereas solely resampling is preferred in slight class imbalance.

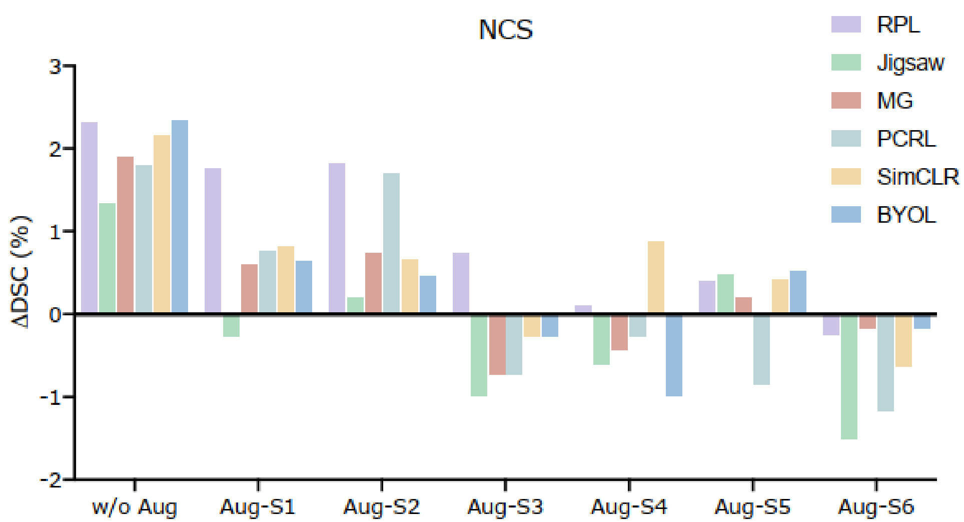
Takeaway 2

- Strong augmentation diminishes the value of SSL pretraining.

Dive into Self-supervised Learning

- **Detail FOUR: Common Training Policies**

Imbalance Ratio	N	Mean-Resampling		+MG		+PCRL	
		Δ_{AUC}	Δ_{Recall}	Δ_{AUC}	Δ_{Recall}	Δ_{AUC}	Δ_{Recall}
r=0.01	5k	-4.21	19.10	(+6.10) 1.89	(+13.28) 32.38	(+2.77) -1.44	(+24.24) 43.34
	10k	9.96	37.00	(+12.57) 22.53	(+17.31) 54.31	(+4.93) 14.89	(+22.53) 59.53
	15k	17.21	52.66	(+1.96) 19.17	(+17.75) 70.41	(+4.01) 21.22	47.52
r=0.05	5k	4.14	45.16	(+10.20) 14.34	(+3.93) 49.09	(+2.33) 6.47	38.38
	10k	22.94	47.25	22.98	47.01	21.39	41.78
	15k	21.05	44.13	21.25	(+8.09) 52.22	19.04	44.91
r=0.1	5k	-6.20	40.21	(+4.68) -1.52	34.99	(+2.94) -3.26	26.37
	10k	16.42	47.78	16.12	47.51	16.45	43.08
	15k	11.26	47.51	11.53	46.21	9.98	(+2.10) 49.61



Pretrain-then-Finetune

Key problems

- How to pretrain a model?
- How to pick the best pretrained model?
- How to fine-tune the pretrained model?

Guidance

- Dive into the Details of SSL, Medical Image Analysis, 2023

Pretraining

- SOOOOOOOO MANY WORKS! **My Ads here: ST-UNet**

Transferability

- Class-Consistency and Feature Variety, *MICCAI, 2023*

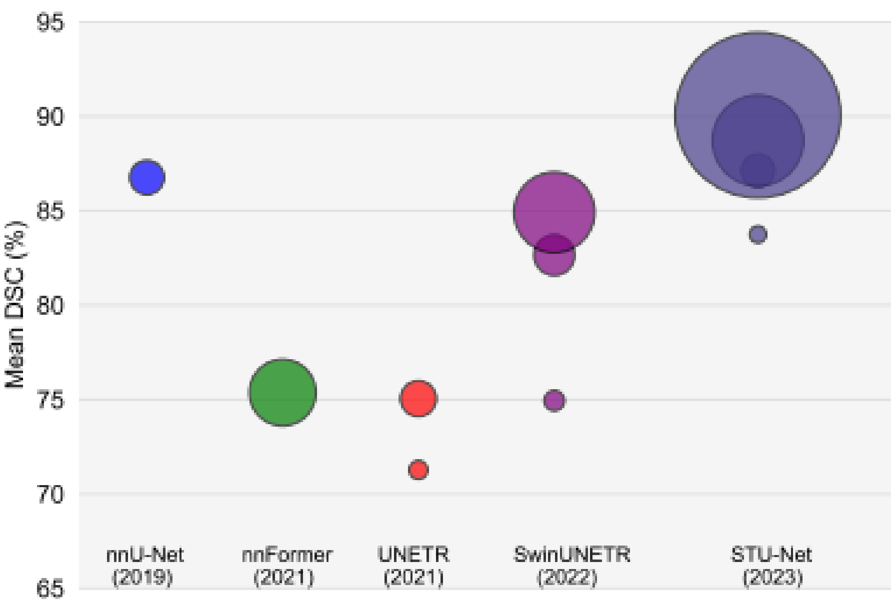
Finetuning

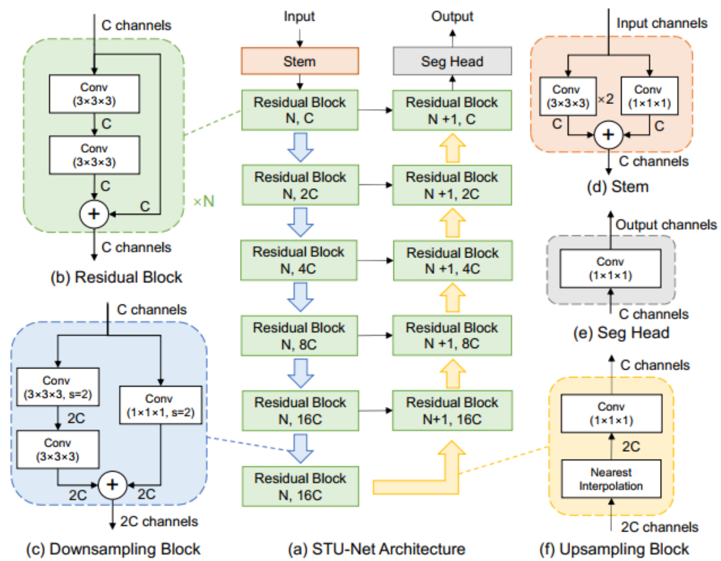
- These are open opportunities to you again!

Scalable and Transferable Medical Image Segmentation Models

- **Better Designs of UNet**

Settings	nnU-Net	3D U-Net	STU-Net (ours)
number of resolution stages	4-7	5	6
convolution kernels	3×3×3 or 3×3×1	3×3×3	3×3×3
up(down)-sample ratios	(2,2,2) or (2,2,1)	(2,2,2)	(2,2,2) or (2,2,1)
input patch size	task-specific	fixed	task-specific
input spacing	task-specific	fixed	task-specific
up-sample operation	transpose convolution	transpose convolution	interpolation with (1,1,1) convolution





Ziyan Huang et.al, STU-Net: Scalable and Transferable Medical Image Segmentation Models Empowered by Large-Scale Supervised Pre-training 2023

Pretrain-then-Finetune

Key problems

- How to pretrain a model?
- How to pick the best pretrained model?
- How to fine-tune the pretrained model?

Guidance

- Dive into the Details of SSL, Medical Image Analysis, 2023

Pretraining

- Partially-Supervised Pretraining via ST-UNet, arxiv, 2023

Transferability

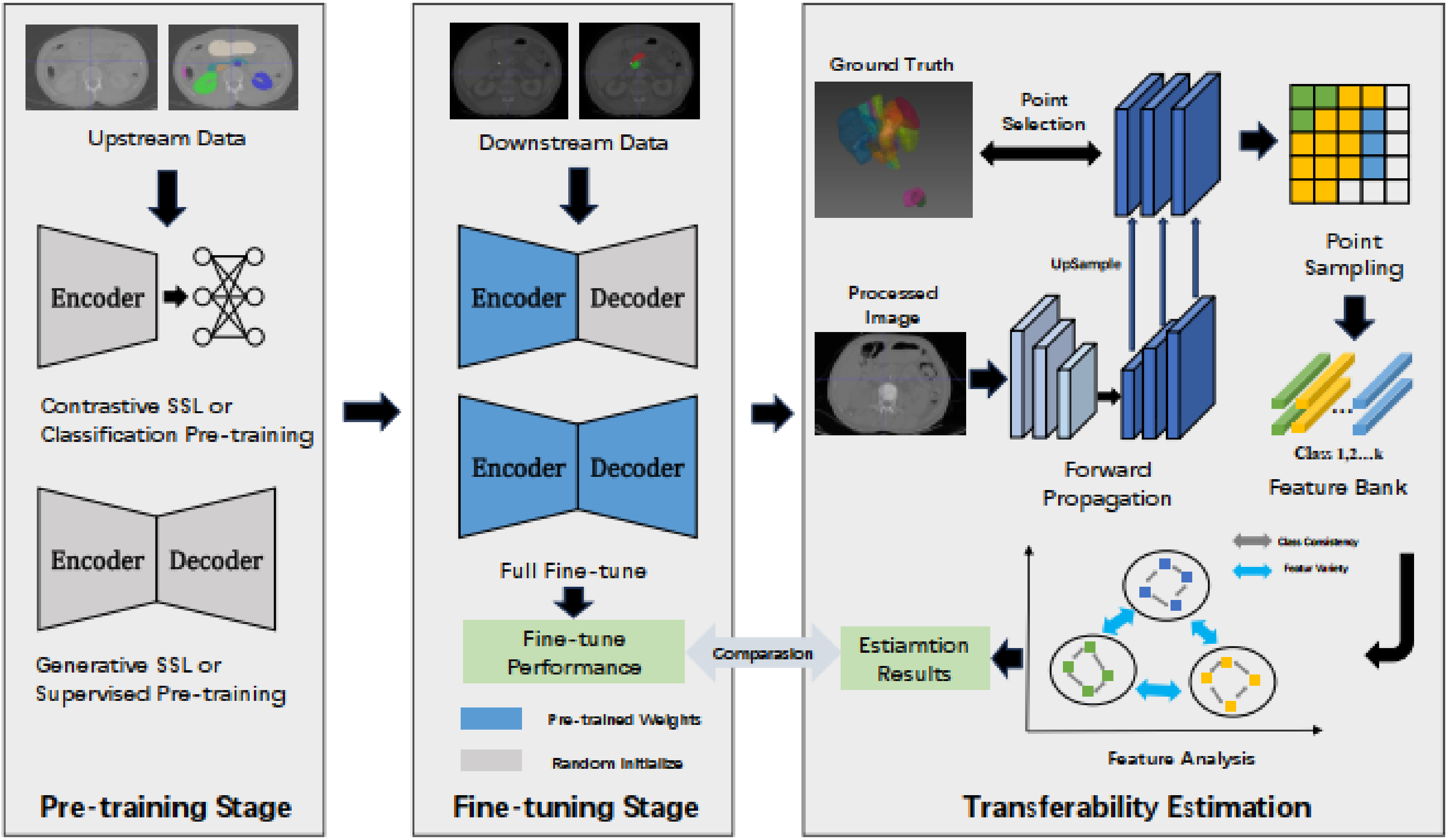
- **Class-Consistency and Feature Variety,** ***MICCAI, 2023***

Finetuning

- These are open opportunities to you again!

Pick the Best Pretrained Model

- **Given a model zoo, how to pick the best model for finetuning**
  - **Effectively (Accurate)**
  - **Efficiently (Fast)**

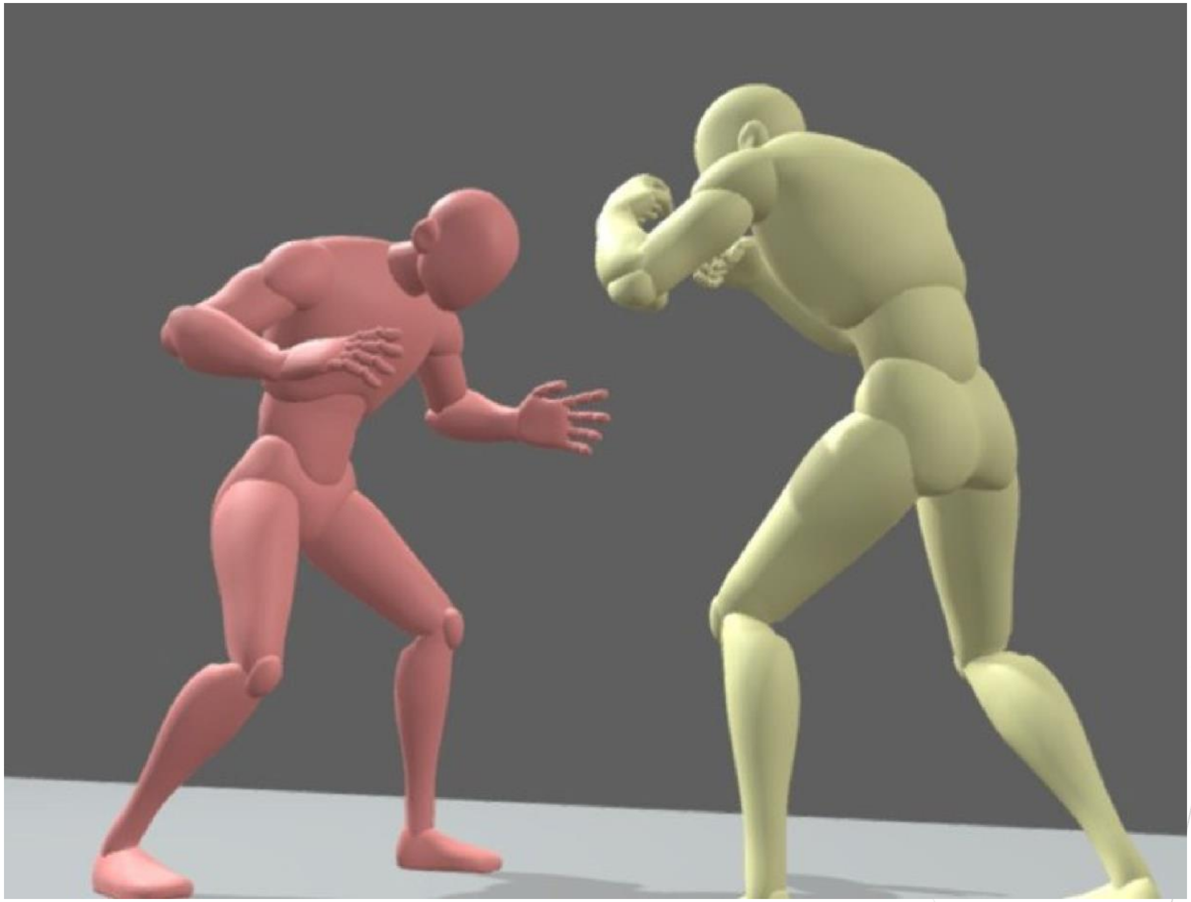


Yuncheng Yang et.al, Transferability Estimation For Medical Image Segmentation: Method and Observations, ***MICCAI 2023 (Early Accepted)***

Pick the Best Pretrained Model

Criteria

Given pretrained models, If **Model A** performs better than **Model B** **after the fine-tuning** on target task, the transferability metric of **Model A** is larger than **Model B** **without fine-tuning**.

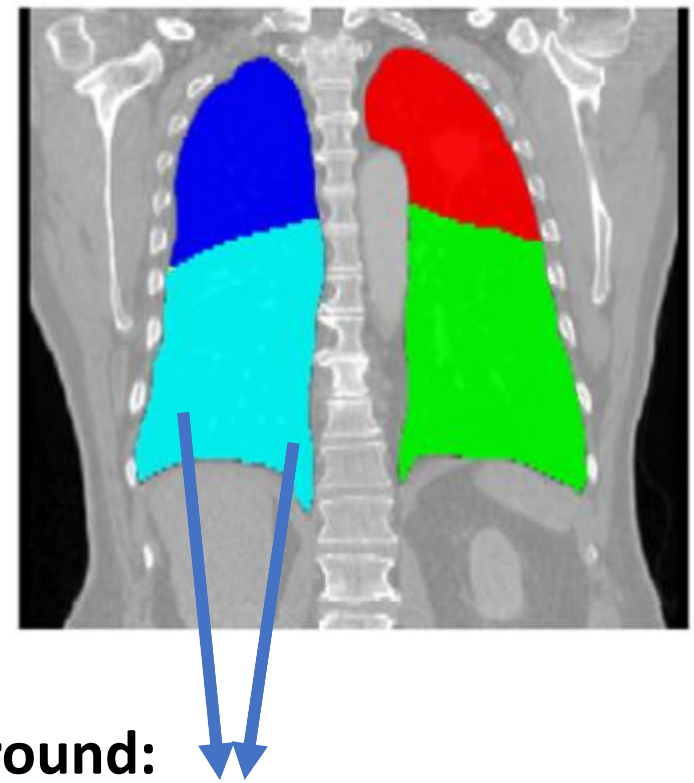


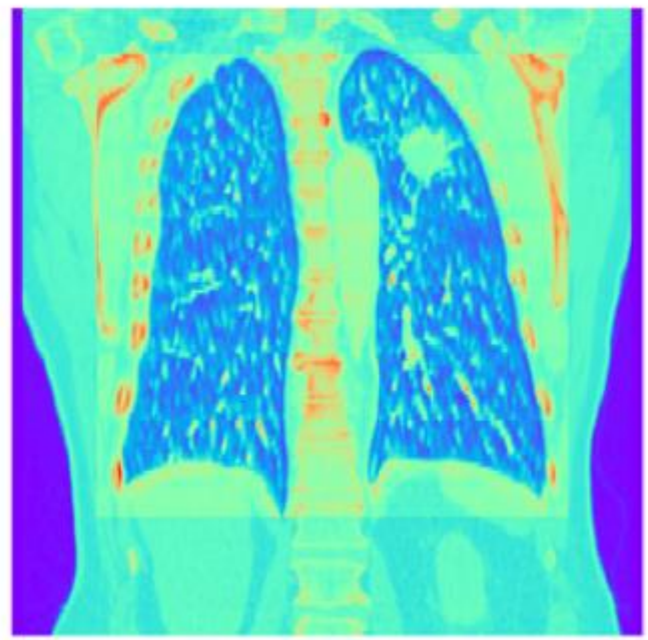
https://sketchfab.com/3d-models/rapid-punching-animation-898c066dc84b4fe7b19fa2f40f4fd145

Segmentation Tasks

Criteria

Given pretrained models, If **Model A** performs better than **Model B** **after the fine-tuning** on target task, the transferability metric of **Model A** is larger than **Model B** **without fine-tuning**.





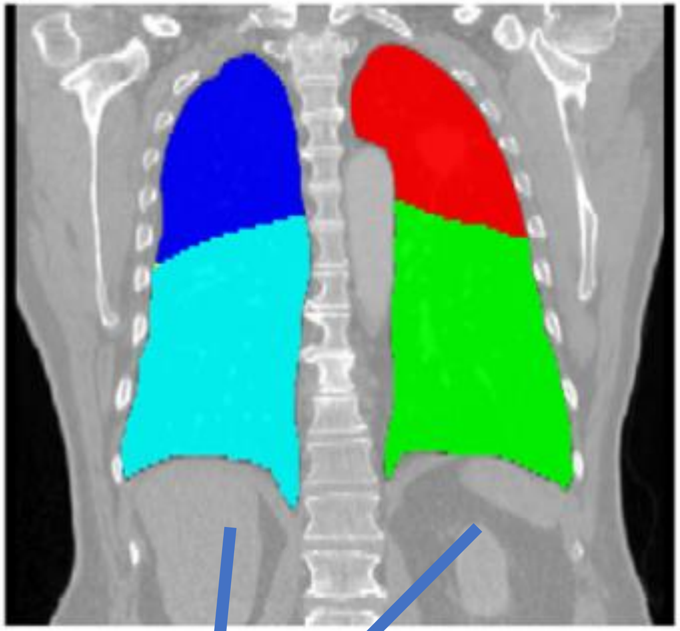
Foreground:

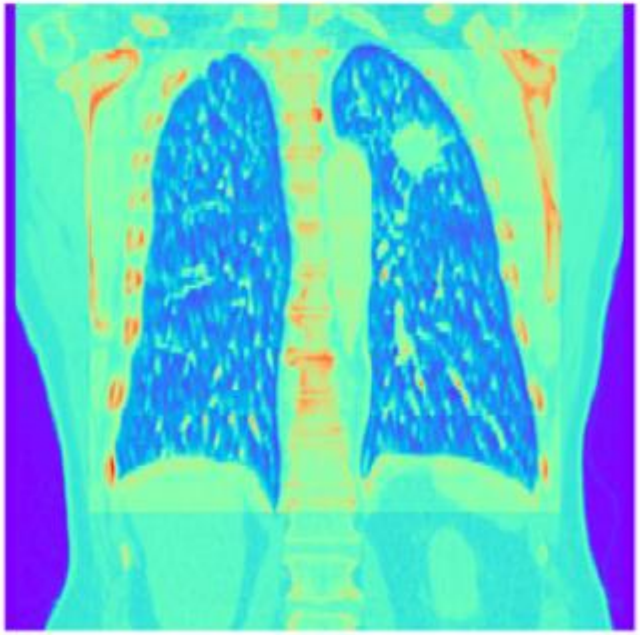
Are they similar in the feature space of pretrained model?

Segmentation Tasks

Criteria

Given pretrained models, If **Model A** performs better than **Model B** **after the fine-tuning** on target task, the transferability metric of **Model A** is larger than **Model B** **without fine-tuning**.





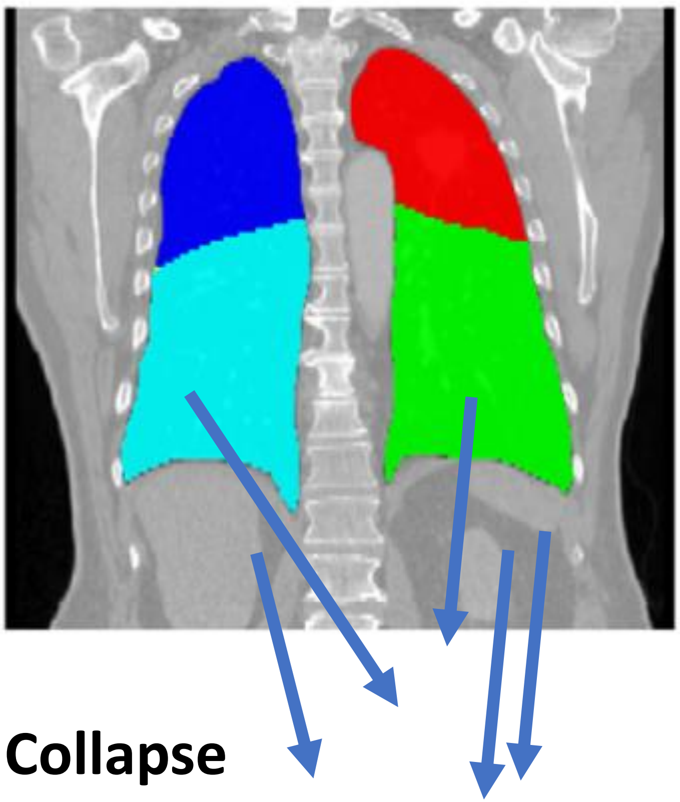
Background:

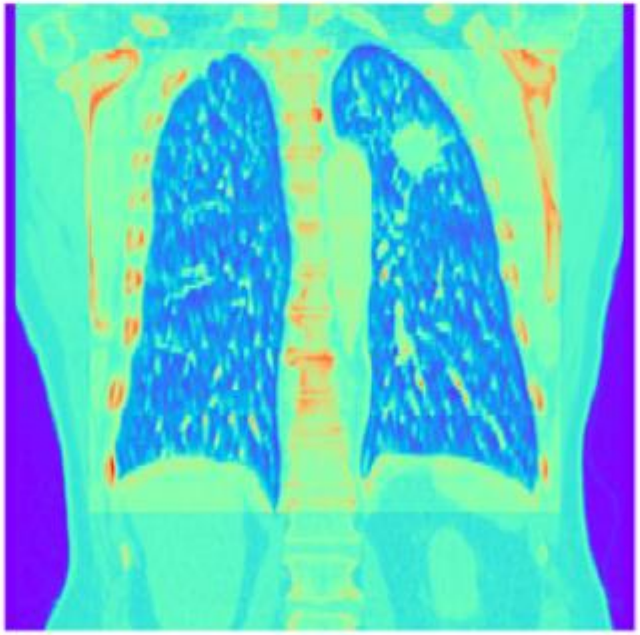
Should we also follow the idea of foreground regions?

Segmentation Tasks

Criteria

Given pretrained models, If **Model A** performs better than **Model B** **after the fine-tuning** on target task, the transferability metric of **Model A** is larger than **Model B** **without fine-tuning**.





Model Collapse

If all of them are similar?

Pick the Best Pretrained Model

Criteria

Given pretrained models, If **Model A** performs better than **Model B** **after the fine-tuning** on target task, the transferability metric of **Model A** is larger than **Model B** **without fine-tuning**.

Class-Consistency (CC) only for foreground classes

$$\mathcal{W}_2^2(F_j^k, F_{j'}^k) = \left\| \boldsymbol{\mu}_{F_j^k} - \boldsymbol{\mu}_{F_{j'}^k} \right\|^2 + \mathrm{Tr}\left(\boldsymbol{\Sigma}_{F_j^k}\right) + \mathrm{Tr}\left(\boldsymbol{\Sigma}_{F_{j'}^k}\right) - 2\,\mathrm{Tr}\left(\left(\boldsymbol{\Sigma}_{F_j^k}\boldsymbol{\Sigma}_{F_{j'}^k}\right)^{1/2}\right) \quad (1)$$

$$C_{cons} = \frac{1}{N(N-1)} \sum_{k=1}^{C} \sum_{i \neq j} \mathcal{W}_2(\boldsymbol{F}_i^k, \boldsymbol{F}_j^k) \quad (2)$$

Feature-Variety (FV) for background classes

$$\boldsymbol{E}_s(\boldsymbol{v}) = \sum_{i=1}^{L} \sum_{j=1, j\neq i}^{L} e_s(\|\boldsymbol{v}_i - \boldsymbol{v}_j\|) = \begin{cases} \sum_{i\neq j} \|\boldsymbol{v}_i - \boldsymbol{v}_j\|^{-s}, & s > 0 \\ \sum_{i \neq j} \log\left(\|\boldsymbol{v}_i - \boldsymbol{v}_j\|^{-1}\right), & s = 0 \end{cases} \quad (3)$$

$$\boldsymbol{F}_v = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{E}_s^{-1}(\boldsymbol{v})$$

Yuncheng Yang et.al, Transferability Estimation For Medical Image Segmentation: Method and Observations, ***MICCAI 2023 (Early Accepted)***

Pick the Best Pretrained Model

Criteria

Given pretrained models, If **Model A** performs better than **Model B** **after the fine-tuning** on target task, the transferability metric of **Model A** is larger than **Model B** **without fine-tuning**.

Class-Consistency (CC) and Feature-Variety (FV)

$$\mathcal{T}_i = \frac{1}{D}\sum_{i=1}^{D} \log \frac{F_v^i}{C_{cons}^i}$$

Yuncheng Yang et.al, Transferability Estimation For Medical Image Segmentation: Method and Observations, ***MICCAI 2023 (Early Accepted)***

Pick the Best Pretrained Model

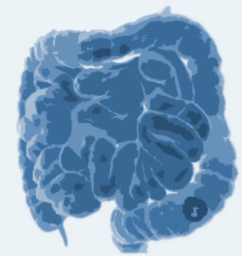
https://github.com/EndoluminalSurgicalVision-IMR/CCFV

Liver Tumours

Target: Liver and tumour
Modality: Portal venous phase CT
Size: 201 3D volumes (131 Training + 70 Testing)
Source: IRCAD Hôpitaux Universitaires
Challenge: Label unbalance with a large (liver) and small (tumour) target

Lung Tumours

Target: Lung and tumours
Modality: CT
Size: 96 3D volumes (64 Training + 32 Testing)
Source: The Cancer Imaging Archive
Challenge: Segmentation of a small target (cancer) in a large image



Colon Cancer

Target: Colon Cancer Primaries
Modality: CT
Size: 190 3D volumes (126 Training + 64 Testing)
Source: Memorial Sloan Kettering Cancer Center
Challenge: Heterogeneous appearance

Spleen

Target: Spleen
Modality: CT
Size: 61 3D volumes (41 Training + 20 Testing)
Source: Memorial Sloan Kettering Cancer Center
Challenge: Large ranging foreground size

Pancreas Tumour

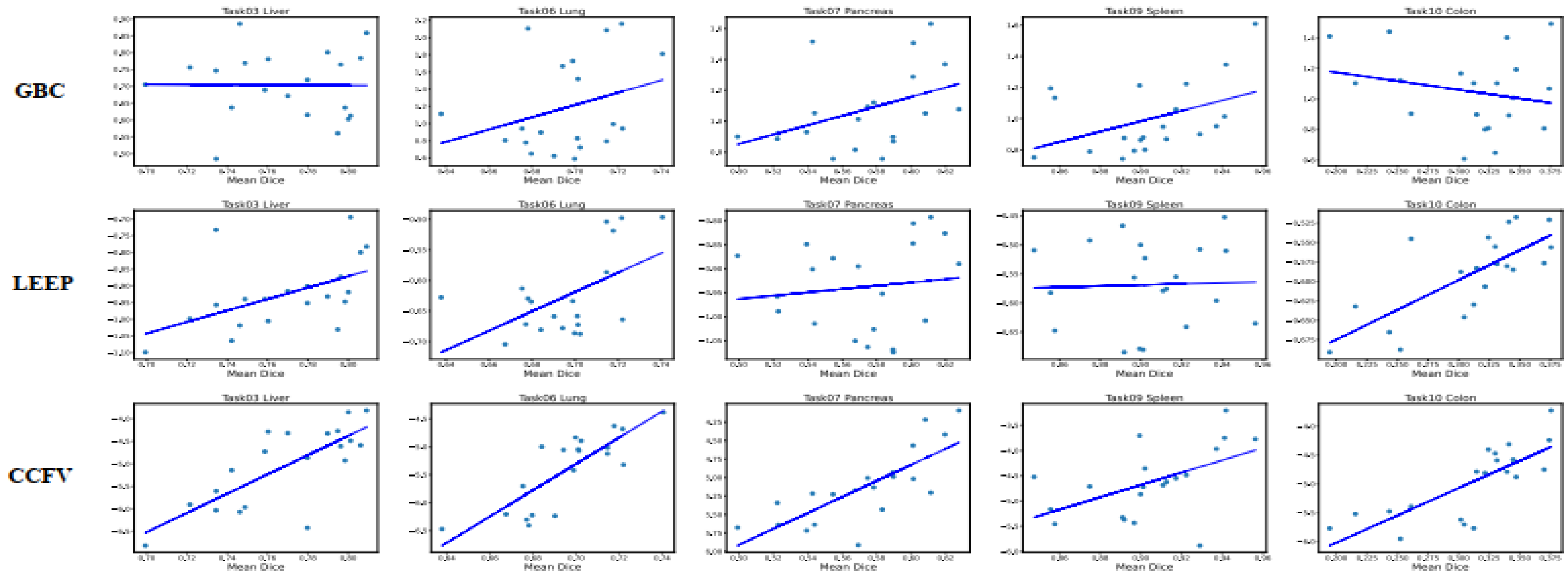
Target: Liver and tumour
Modality: Portal venous phase CT
Size: 420 3D volumes (282 Training +139 Testing)
Source: Memorial Sloan Kettering Cancer Center
Challenge: Label unbalance with large (background), medium (pancreas) and small (tumour) structures.

Yuncheng Yang et.al, Transferability Estimation For Medical Image Segmentation: Method and Observations, ***MICCAI 2023 (Early Accepted)***

Pick the Best Pretrained Model via CCFV

https://github.com/EndoluminalSurgicalVision-IMR/CCFV

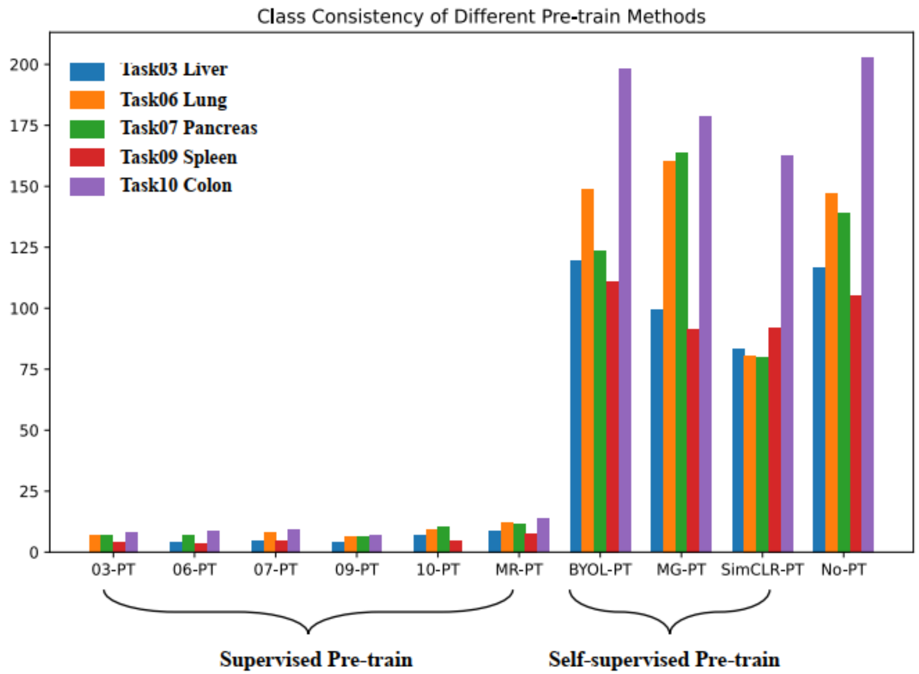
Data/Method	Metrics	Task03	Task06	Task07	Task09	Task10	Avg
LogME	τ	-0.0887	0.2895	-0.1219	-0.0152	-0.0938	-0.0060
	pearson	-0.0927	0.3541	-0.1101	-0.0102	0.0042	0.0290
Transrate	τ	0.3168	0.4635	0.2654	0.6085	0.0961	0.3507
	pearson	0.4440	0.4134	0.3237	0.5857	0.0927	0.3719
Leep	τ	0.5310	0.5231	0.3294	0.0490	0.5499	0.3964
	pearson	0.5134	0.5423	0.1306	0.0378	0.7407	0.3929
GBC	τ	0.1465	0.3320	0.3891	0.6251	0.1503	0.3286
	pearson	-0.0081	0.3064	0.7460	0.5600	0.7139	0.1847
Ours CC-FV	τ	0.6234	0.6508	0.6569	0.5700	0.5550	**0.6112**
	pearson	0.7665	0.8174	0.7703	0.5220	0.7766	**0.7305**

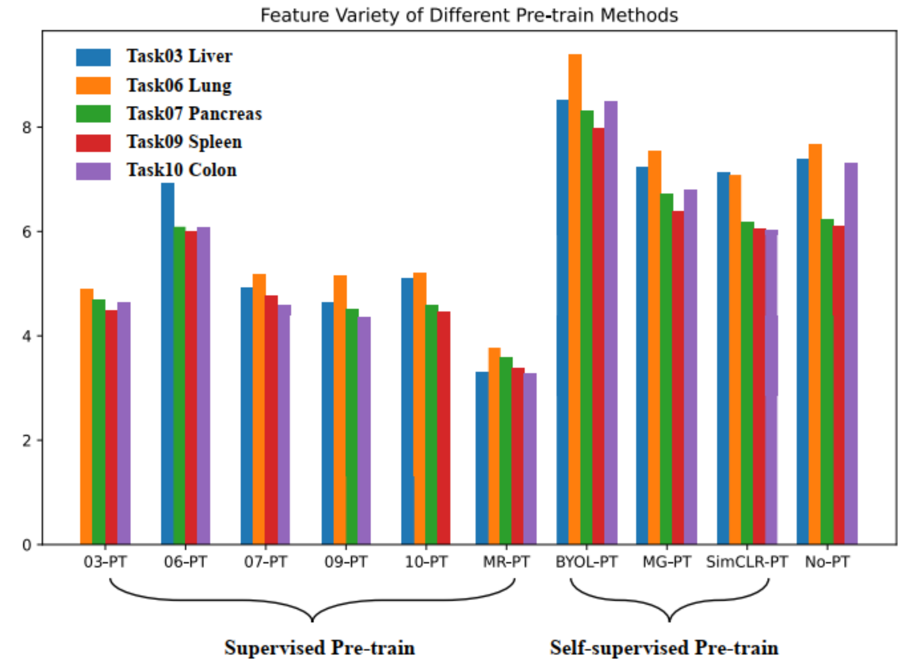


CCFV is a powerful tool

https://github.com/EndoluminalSurgicalVision-IMR/CCFV

- **Supervised Pretraining**
- **Self-supervised Pretraining**
- **…**





Pretrain-then-Finetune

Key problems

- How to pretrain a model?
- How to pick the best pretrained model?
- How to fine-tune the pretrained model?

Guidance

- Dive into the Details of SSL, Medical Image Analysis, 2023

Pretraining

- Partially-Supervised Pretraining via ST-UNet, arxiv, 2023

Transferability

- Class-Consistency and Feature Variety, MICCAI, 2023

Finetuning

- **These are open opportunities to you again!**

Tuning a foundation model is popular!

Google Scholar MEDICAL

Articles

About 375 results (**0.05** sec)

Kirillov, Alexander, et al. "Segment Anything." (2023).

758 Citations by Oct 12th

>**300 "Medical" Citations**

Any time
Since 2023
Since 2022
Since 2019
Custom range...

Sort by relevance
Sort by date

Create alert

Segment anything

Search within citing articles

Segment anything model for **medical** image analysis: an experimental study
MA Mazurowski, H Dong, H Gu, J Yang, N Konz… - **Medical** Image …, 2023 - Elsevier
… , **medical** image domains pose their own set of challenges. Here, we perform an extensive evaluation of SAM's ability to segment **medical** images on a collection of 19 **medical** … **medical** …
Save Cite Cited by 35 Related articles All 2 versions

Segment anything in **medical** images — [PDF] arxiv.org
J Ma, B Wang - arXiv preprint arXiv:2304.12306, 2023 - arxiv.org
… of SAM to **medical** images, with the goal of creating a universal tool for the segmentation of various **medical** targets. Specifically, we first curate a large-scale **medical** image dataset, …
Save Cite Cited by 114 Related articles All 2 versions

Medical sam adapter: Adapting segment anything model for **medical** image segmentation — [PDF] arxiv.org
J Wu, R Fu, H Fang, Y Liu, Z Wang, Y Xu, Y Jin… - arXiv preprint arXiv …, 2023 - arxiv.org
… A **medical** image adapted SAM, which we have dubbed **Medical** SAM Adapter (MSA), shows superior performance on 19 **medical** image segmentation tasks with various image …
Save Cite Cited by 49 Related articles All 2 versions

[HTML] Large Language Models in **Medical** Education: Opportunities, Challenges, and Future Directions — [HTML] jmir.org
A Abd-Alrazaq, R AlSaad, D Alhuwail… - JMIR **Medical** …, 2023 - mededu.jmir.org
… (GPT) series, into **medical** education has the potential to … promise for revolutionizing **medical** curriculum development, … , and copyright concerns in **medical** education. As we navigate the …
Save Cite Cited by 30 Related articles All 5 versions

Segment anything model for **medical** images? — [PDF] arxiv.org
Y Huang, X Yang, L Liu, H Zhou, A Chang… - arXiv preprint arXiv …, 2023 - arxiv.org
… objects), so as to better develop a new generation of general **medical** segmentation models. In this report, we build a large **medical** image dataset named COSMOS 553K including …
Save Cite Cited by 34 Related articles All 3 versions

Pretrain-then-Finetune

Key problems

- How to pretrain a model?
- How to pick the best pretrained model?
- How to fine-tune the pretrained model?

Guidance

- **Dive into the Details of SSL,** ***Medical Image Analysis, 2023***

Pretraining

- **Partially-Supervised Pretraining via ST-UNet,** ***arxiv, 2023***

Transferability

- **Class-Consistency and Feature Variety,** ***MICCAI, 2023***

Finetuning

- **These are open opportunities to you again!**

Many Thanks!