Learning Anatomically Consistent Embedding for Chest Radiography

Ziyu Zhou*[1,2], Haozhe Luo*[2], Jiaxuan Pang[2], Xiaowei Ding[1†], Michael B. Gotway[3], Jianming Liang[2†]
[1] Shanghai Jiao Tong University [2] Arizona State University [3] Mayo Clinic

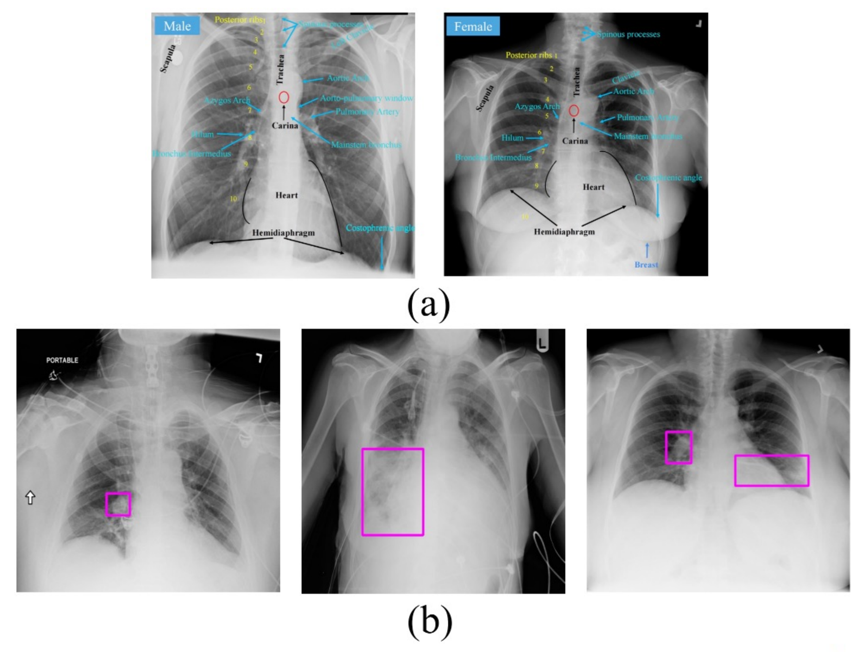
BMVC 2023

Background: Self-supervised learning (SSL) pretrains generic source models without using expert annotation, allowing the pretrained models to be quickly fine-tuned into high-performance application-specific target models and minimizing annotation cost.

Motivation: Compared with photographic images, medical images acquired with the same image protocol **exhibit high consistency** in anatomy. To exploit this anatomical consistency, this paper introduces a novel SSL approach, called **PEAC** (patch embedding of anatomical consistency).

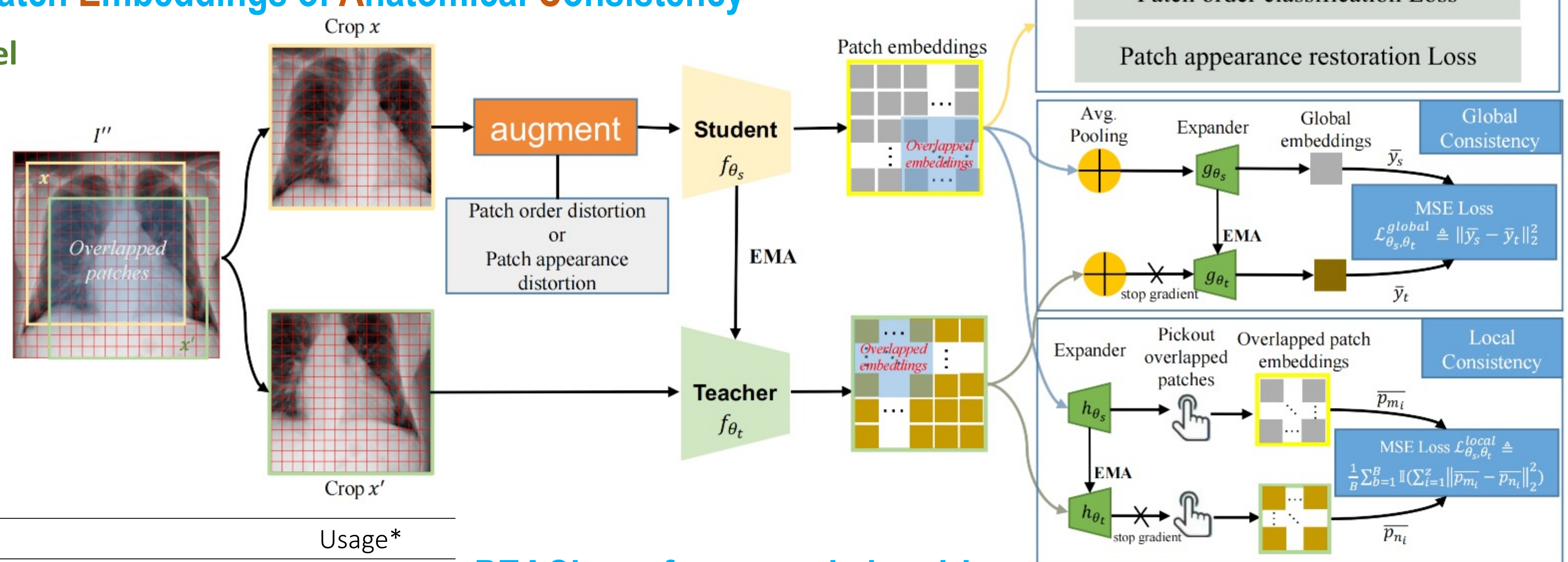
Contributions: We propose to learn **global and local consistencies** via stable grid-based matching, transfer pre-trained PEAC models to diverse downstream tasks, and demonstrate that (1) PEAC achieves significantly better performance than the existing state-of-the-art fully/self-supervised methods; (2) PEAC captures the anatomical structure consistency across views of the same patient and across patients of different genders, weights, and healthy statuses.

Motivations: (a) there are large (global) and small (local) anatomical structures; (b) lung diseases can be local or global.



PEAC: Patch Embeddings of Anatomical Consistency

1. built on **Student-Teacher Model**
2. autodidactically gleans fine-grained image details by **patch order classification and patch appearance restoration**;
3. learns coarse-grained global features by **global consistency**;
4. learns fine-grained local features from overlapped patches by **local consistency**.



Dataset	Task	Usage*
ChestX-ray14	14 thoracic diagnoses classification	P | F
CheXpert	14 thoracic diseases classification	F
ShenZhen	Tuberculosis classification	F
RSNA Pneumonia	Lung opacity, abnormality classification	F
JSRT	Lungs, heart, clavicles segmentation	F

* The usage of each dataset in our experiments is denoted with P for pretraining, F for finetuning.

PEAC's performance is inspiring

Result 1: PEAC outperforms fully-supervised pretrained models

Backbone	Pretraining data	Pretraining method	ChestX-ray14	CheXpert	ShenZhen	RSNA Pneumonia
ResNet-50	No pretraining (i.e., training from scratch)		80.40 ± 0.05	86.60 ± 0.17	90.49 ± 1.16	70.00 ± 0.50
	ImageNet-1K	Fully-supervised	81.70 ± 0.15	87.17 ± 0.22	94.96 ± 1.19	73.04 ± 0.35
	ChestX-ray14	Fully-supervised	–	87.40 ± 0.26	96.32 ± 0.65	71.64 ± 0.37
ViT-B	No pretraining (i.e., training from scratch)		70.84 ± 0.19	80.78 ± 0.13	84.46 ± 1.65	66.59 ± 0.39
	ImageNet-21K	Fully-supervised	77.55 ± 1.82	83.32 ± 0.69	91.85 ± 3.40	71.50 ± 0.52
	ChestX-ray14	Fully-supervised	–	84.37 ± 0.42	91.23 ± 0.81	66.96 ± 0.24
	ChestX-ray14	$PEAC^{-3}$ (self-supervised)	80.04 ± 0.20	88.10±0.29	96.69 ± 0.30	<u>73.77 ± 0.39</u>
Swin-B	No pretraining (i.e., training from scratch)		74.29 ± 0.41	85.78 ± 0.01	85.83 ± 3.68	70.02 ± 0.42
	ImageNet-21K	Fully-supervised	81.32 ± 0.19	87.94 ± 0.36	94.23 ± 0.81	73.15 ± 0.61
	ChestX-ray14	Fully-supervised	–	87.22 ± 0.22	91.35 ± 0.93	70.67 ± 0.18
	ChestX-ray14	$PEAC^{-1}$ (self-supervised)	<u>81.90 ± 0.15</u>	<u>88.64 ± 0.19</u>	<u>97.17 ± 0.42</u>	73.70 ± 0.48
	ChestX-ray14	PEAC (self-supervised)	**82.78 ± 0.21**	**88.81 ± 0.57**	**97.39 ± 0.19**	**74.39 ± 0.66**

Result 2: PEAC outperforms self-supervised models pretrained on ImageNet

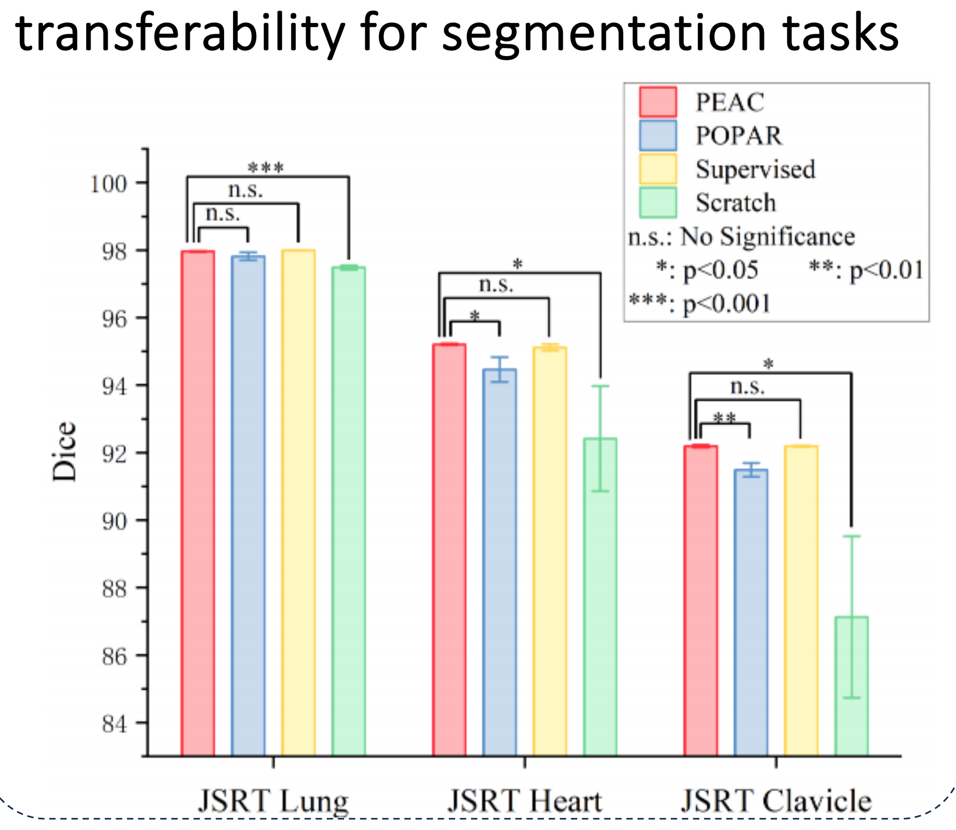
Backbone	Pretrained dataset	Method	ChestX-ray14	CheXpert	ShenZhen	RSNA Pneumonia
ViT-B	ImageNet	MoCo V3	79.20 ± 0.29	86.91 ± 0.77	85.71 ± 1.41	72.79 ± 0.52
		SimMIM	79.55 ± 0.56	87.83 ± 0.46	92.74 ± 0.92	72.08 ± 0.47
		DINO	78.37 ± 0.47	86.91 ± 0.44	87.83 ± 7.20	71.27 ± 0.45
		BEiT	74.69 ± 0.29	85.81 ± 1.00	92.95 ± 1.25	72.78 ± 0.37
		MAE	78.97 ± 0.65	87.12 ± 0.54	93.58 ± 1.18	72.85 ± 0.50
	ChestX-ray14	$PEAC^{-3}$	80.04 ± 0.20	88.10±0.29	96.69 ± 0.30	<u>73.77 ± 0.39</u>
Swin-B	ImageNet	SimMIM	81.39 ± 0.18	87.50 ± 0.23	87.86 ± 4.92	73.15 ± 0.73
	ChestX-ray14	$PEAC^{-1}$	<u>81.90 ± 0.15</u>	<u>88.64 ± 0.19</u>	<u>97.17 ± 0.42</u>	73.70 ± 0.48
		PEAC	**82.78 ± 0.21**	**88.81 ± 0.57**	**97.39 ± 0.19**	**74.39 ± 0.66**

Result 3: PEAC outperforms recent self-supervised models pretrained on medical images

Backbone	Pretrained dataset	Method	ChestX-ray14	CheXpert	ShenZhen	RSNA Pneumonia
ResNet-50	ChestX-ray14	SimSiam	79.62 ± 0.34	83.82 ± 0.94	93.13 ± 1.36	71.20 ± 0.60
		MoCoV2	80.36 ± 0.26	86.42 ± 0.42	92.59 ± 1.79	71.98 ± 0.82
		Barlow Twins	80.45 ± 0.29	86.90 ± 0.62	92.17 ± 1.54	71.45 ± 0.82
ViT-B	ChestX-ray14	SimMIM	79.20 ± 0.19	83.48 ± 2.43	93.77 ± 1.01	71.66 ± 0.75
		$PEAC^{-3}$	80.04 ± 0.20	88.10±0.29	96.69 ± 0.30	**73.77 ± 0.39**
Swin-B	ChestX-ray14	SimMIM	79.09 ± 0.57	86.75 ± 0.96	93.03 ± 0.48	71.99 ± 0.55
		$POPAR^{-1}$	<u>80.51 ± 0.15</u>	<u>88.16 ± 0.66</u>	<u>96.81 ± 0.40</u>	73.58 ± 0.18
		$PEAC^{-1}$	**81.90 ± 0.15**	**88.64 ± 0.19**	**97.17 ± 0.42**	<u>73.70 ± 0.48</u>

With the best bolded and the second best underlined, a statistical analysis is conducted between the best vs. others, where blue-highlighted boxes indicate no statistically significant difference at level p = 0.05.

Result 4: PEAC exhibits prominent transferability for segmentation tasks

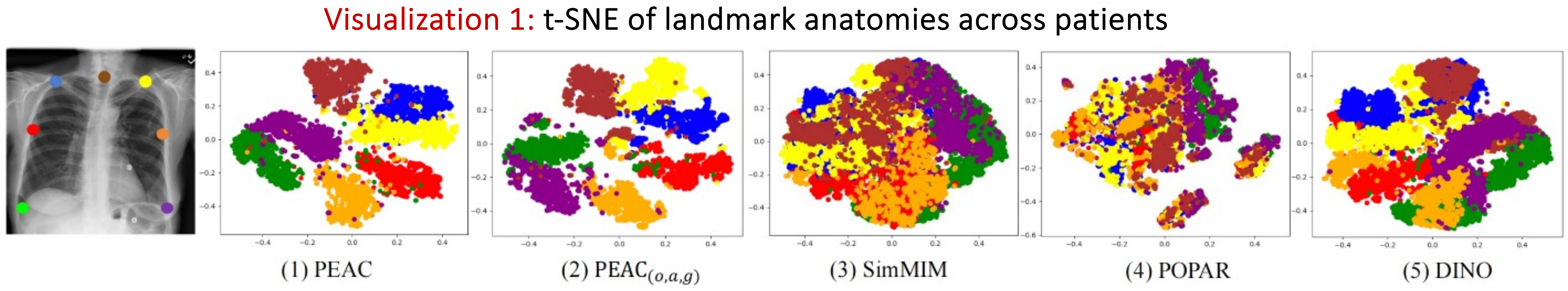


Result 5: PEAC is outstanding in small data regime

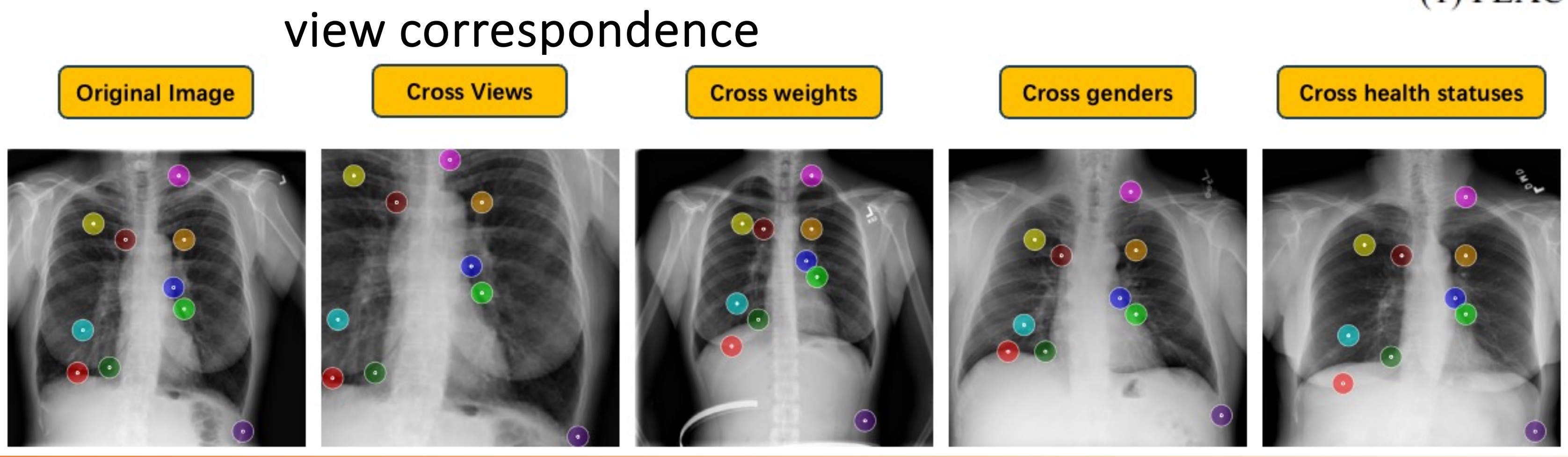
Method	25%	50%	100%
MoCo-v2	74.71	76.89	80.36
Barlow Twins	76.23	77.59	80.45
SimSiam	73.05	75.20	79.62
$DiRA_{MoCo\text{-}v2}$	<u>77.55</u>	<u>78.74</u>	<u>81.12</u>
PEAC	77.78	79.29	82.78

Visualization of PEAC model

Visualization 1: t-SNE of landmark anatomies across patients


(1) PEAC (2) $PEAC_{(o,a,g)}$ (3) SimMIM (4) POPAR (5) DINO

Visualization 2: Cross-patient and cross-view correspondence



Visualization 3: Zero-shot co-segmentation