Seminar (Contrastive learning)

OUYANG YIMING

2021.12.29

(MICCAI2021) Positional Contrastive Learning for Volumetric Medical Image Segmentation

DEWEN ZENG, YAWEN WU, XINRONG HU, XIAOWEI XU, HAIYUN YUAN, MEIPING HUANG, JIAN ZHUANG, JINGTONG HU, YIYU SHI

Motivation and Contribution

1. The scarcity of annotated medical images

2. False negative pairs due to the similarity of different images

3. The doubtful attempt to divide volumes into partitions and consider the slides of corresponding partitions in different volumes as positive pairs and those of different partitions as negative pairs

4. Propose a novel positional contrastive learning (PCL) framework: generating pairs based on the position of a slice in volumetric medical images

Method

1. Pre-train a 2D U-Net (contrastive learning)

2. Fine-tune using the pre-trained encoder as initialization (supervised learning)
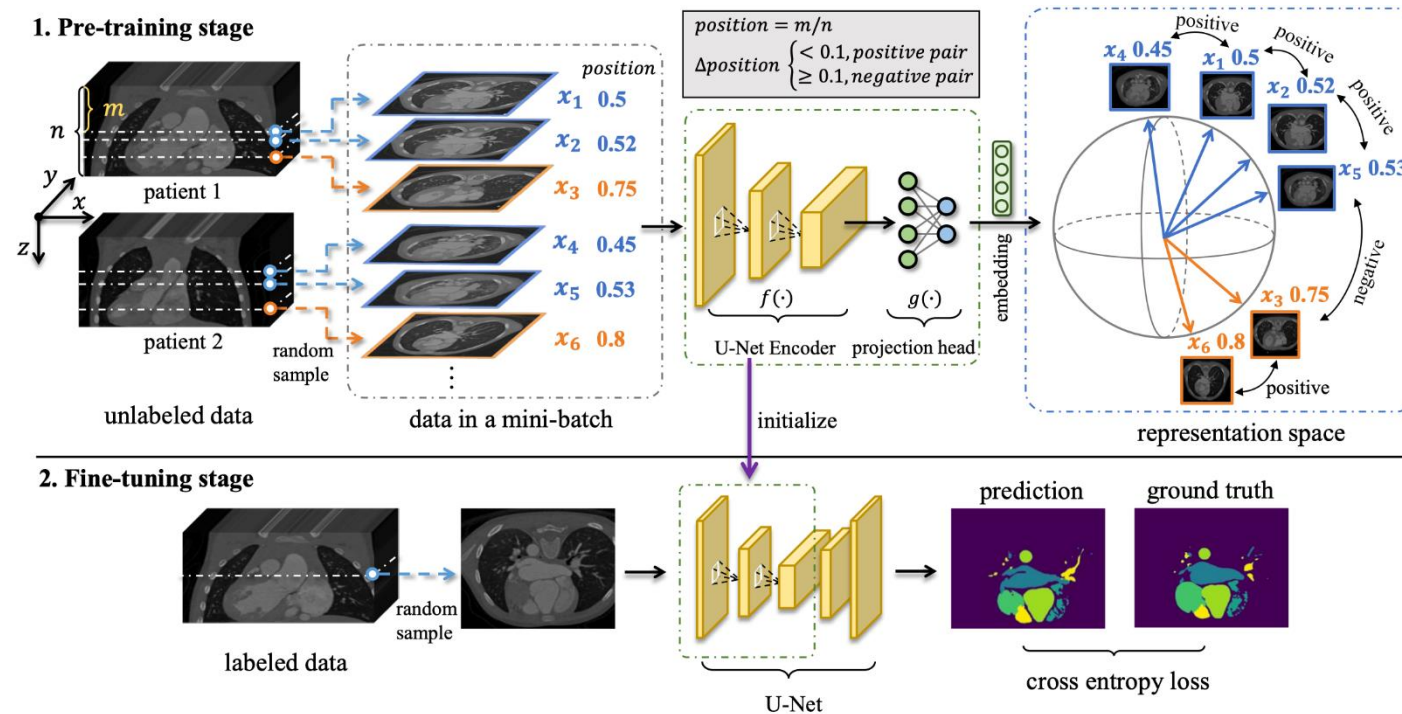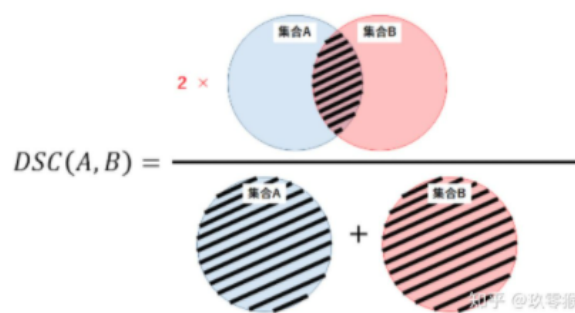
Dataset

1. Congenital Heart Disease (CHD) dataset, CT

2. MMWHS dataset, CT (same annotation types as CHD)

3. ACDC dataset, MRI

4. HVSMR dataset, MRI (a bit different annotation types from ACDC)

Evaluation

Dice

对于分割过程中的评价标准主要采用Dice相似系数(Dice Similariy Coefficient,DSC),Dice系数是一种集合相似度度量指标,通常用于计算两个样本的相似度,值的范围 $0-1$,分割结果最好时值为 1 ,最差时值为 0

$$Dice(P,T) = \frac{|P_1 \wedge T_1|}{(|P_1| + |T_2|)/2} \Leftrightarrow Dice = \frac{2TP}{FP + 2TP + FN}$$

Experiment(Semi-supervised Learning)

Table 1: Comparison of the proposed PCL method with baseline methods on CHD and ACDC. M is the number of patients used in the fine-tuning process. Results are reported in the form of mean(standard deviation) on 5-fold cross-validation. PCL provides better results than the baselines for all values of M.

CHD (68 patients in total)							
Method	$M=2$	$M=6$	$M=10$	$M=15$	$M=20$	$M=30$	$M=51$
Random	0.184(.06)	0.508(.06)	0.584(.05)	0.627(.05)	0.658(.04)	0.693(.04)	0.754(.02)
Rotation [6]	0.171(.06)	0.488(.07)	0.575(.04)	0.625(.04)	0.651(.04)	0.691(.04)	0.749(.03)
PIRL [14]	0.196(.07)	0.504(.08)	0.617(.05)	0.658(.03)	0.674(.04)	0.714(.04)	0.761(.03)
SimCLR [3]	0.192(.06)	0.515(.06)	0.599(.06)	0.631(.05)	0.666(.05)	0.699(.05)	0.756(.03)
GCL [2]	0.255(.10)	0.564(.04)	0.646(.03)	0.669(.04)	0.697(.04)	0.725(.04)	0.766(.03)
PCL	**0.356(.08)**	**0.600(.06)**	**0.661(.05)**	**0.686(.05)**	**0.716(.04)**	**0.735(.05)**	**0.774(.03)**

ACDC (100 patients in total)							
Method	$M=2$	$M=6$	$M=10$	$M=15$	$M=20$	$M=30$	$M=80$
Random	0.588(.07)	0.782(.03)	0.840(.03)	0.876(.01)	0.894(.01)	0.909(.01)	0.928(.00)
Rotation [6]	0.572(.08)	0.809(.03)	0.868(.02)	0.886(.01)	0.898(.01)	0.910(.01)	0.925(.00)
PIRL [14]	0.492(.03)	0.823(.04)	0.865(.01)	0.880(.02)	0.896(.02)	0.912(.01)	0.927(.00)
SimCLR [3]	0.352(.06)	0.725(.08)	0.824(.04)	0.869(.02)	0.894(.01)	0.913(.01)	0.927(.00)
GCL [2]	0.636(.05)	0.803(.04)	0.872(.01)	0.891(.01)	0.902(.01)	0.913(.01)	0.927(.01)
PCL	**0.671(.06)**	**0.850(.01)**	**0.885(.01)**	**0.904(.01)**	**0.909(.01)**	**0.919(.00)**	**0.929(.00)**

Experiment(Transfer Learning)

Table 2: Transfer learning comparison of the proposed PCL method with the baselines. Except for Random, all the methods are pre-trained on CHD and ACDC without labels and fine-tuned on MMWHS and HVSMR respectively.

CHD transferring to MMWHS (20 patients in total)						
Method	$M=2$	$M=4$	$M=6$	$M=8$	$M=10$	$M=16$
Random	0.232(.14)	0.661(.10)	0.732(.07)	0.769(.06)	0.808(.05)	0.834(.05)
Rotation [6]	0.247(.16)	0.659(.13)	0.751(.07)	0.768(.07)	0.803(.06)	0.850(.04)
PIRL [14]	0.251(.10)	0.670(.11)	0.755(.07)	0.774(.06)	0.821(.05)	0.851(.04)
SimCLR [3]	0.269(.17)	0.683(.10)	0.751(.07)	0.783(.06)	0.818(.05)	0.850(.04)
GCL [2]	0.262(.11)	0.703(.07)	0.768(.05)	0.805(.04)	0.820(.04)	0.851(.03)
PCL	**0.339(.15)**	**0.748(.08)**	**0.792(.05)**	**0.820(.05)**	**0.840(.04)**	**0.869(.03)**

| ACDC transferring to HVSMR (10 patients in total)				
Method	$M=2$	$M=4$	$M=6$	$M=8$
Random	0.742(.06)	0.813(.05)	0.842(.03)	0.842(.04)
Rotation [6]	0.737(.07)	0.816(.06)	0.845(.03)	0.844(.03)
PIRL [14]	0.740(.05)	0.826(.04)	0.849(.03)	0.846(.03)
SimCLR [3]	0.700(.07)	0.779(.05)	0.808(.04)	0.815(.04)
GCL [2]	0.770(.05)	0.818(.05)	0.842(.03)	0.843(.03)
PCL	**0.781(.05)**	**0.832(.05)**	**0.857(.03)**	**0.857(.03)**

(MICCAI2021) Semi-supervised Contrastive Learning for Label-efficient Medical Image Segmentation

XINRONG HU, DEWEN ZENG, XIAOWEI XU, YIYU SHI

Motivation and Contribution

1. Most existing contrastive learning methods target image classification tasks

2. Semantic segmentation, on the other hand, requires pixel-wise classification

3. Propose a semi-supervised framework consisting of self-supervised global contrast and supervised local contrast to take advantage of the available labels

Method

(a) global contrastive

(b) self-supervised local contrastive

(c) supervised local contrastive
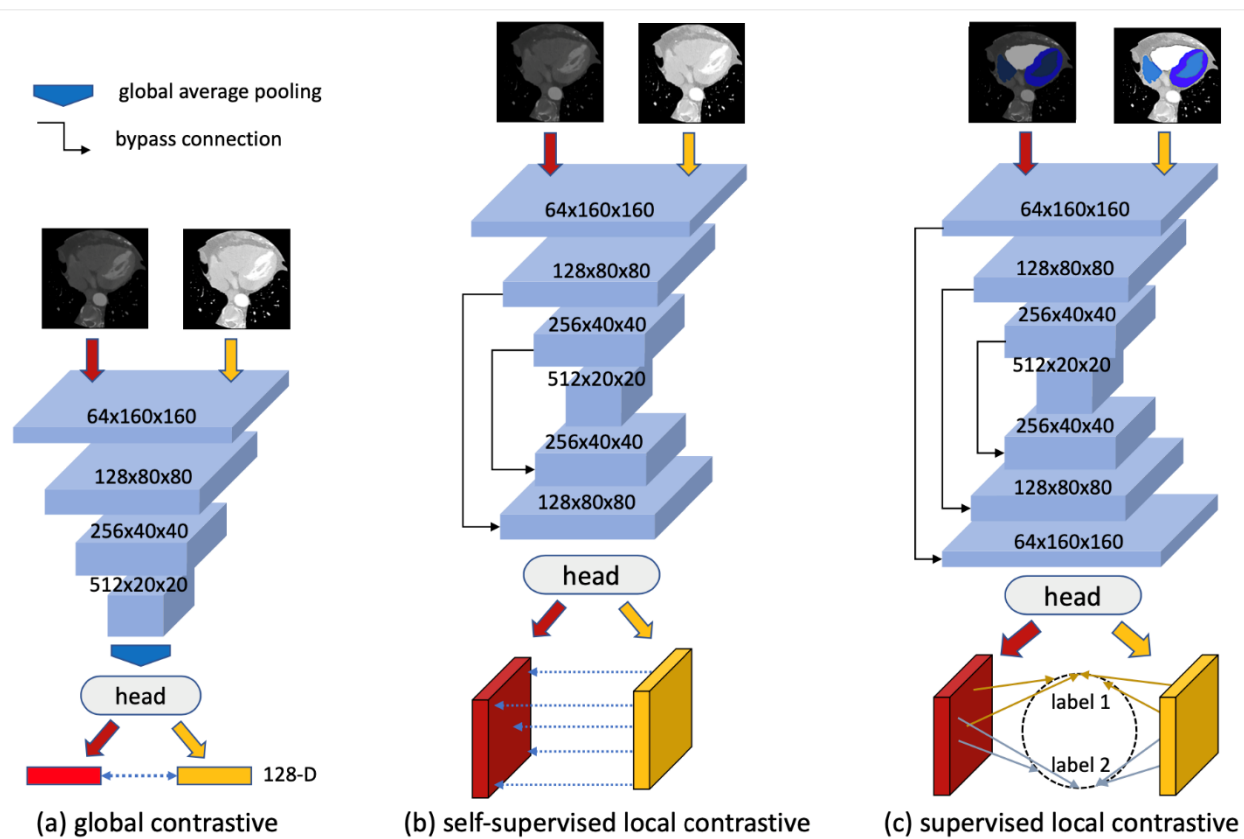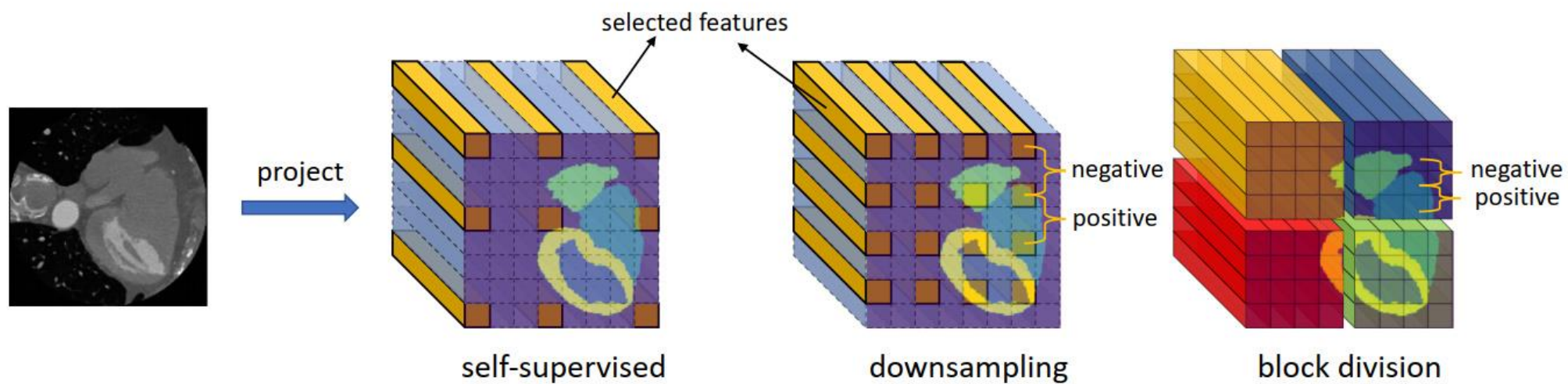
Method

Dataset

1. Hippocampus dataset, MRI

2. MMWHS dataset, CT

Experiment

Table 1: Comparison between state-of-the-art methods and the proposed methods w.r.t. subsequent segmentation dice scores on two datasets. Ablation studies of our methods (w/o global contrast) are also included.

| Methods	Hippocampus			MMWHS		
	% of data labeled in X_{tr}			% of data labeled in X_{tr}		
	5	10	20	10	20	40
random	0.788	0.833	0.852	0.328	0.440	0.715
global [7]	0.817	0.834	0.861	0.359	0.487	0.724
global+local(self) [4]	0.808	0.843	0.858	0.367	0.490	0.730
Mixup [18]	0.818	0.847	0.861	0.365	0.541	0.755
TCSM [14]	0.796	0.838	0.855	0.347	0.489	0.733
local(stride)	0.818	0.845	0.860	0.354	0.485	0.743
local(block)	0.817	0.843	0.862	0.366	0.475	0.736
global+local(stride)	0.822	**0.851**	0.863	**0.384**	0.525	0.758
global+local(block)	**0.824**	0.849	**0.866**	0.382	**0.553**	**0.764**

Experiment
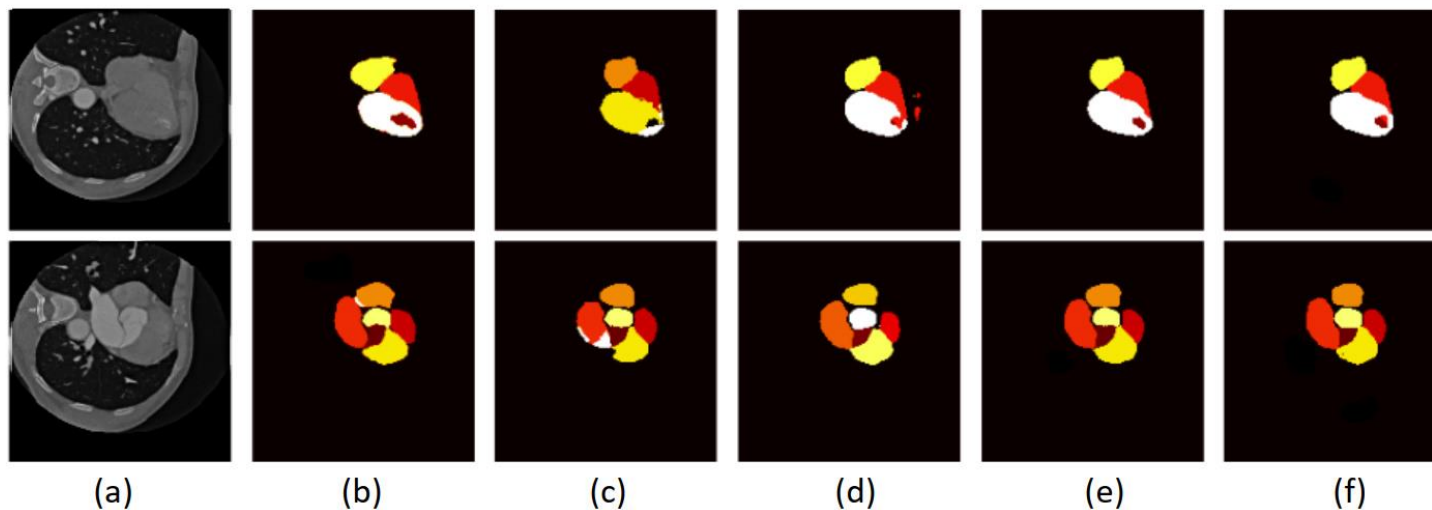
Fig. 3: Visualization of segmentation results on MMWHS with 40% labeled data. (a) and (b) represent input image and ground truth. (c) ~ (f) are the predictions of random, global+local(self), global+local(stride), and global+local(block). Different color represents different structure categories. (Better to view in colors)
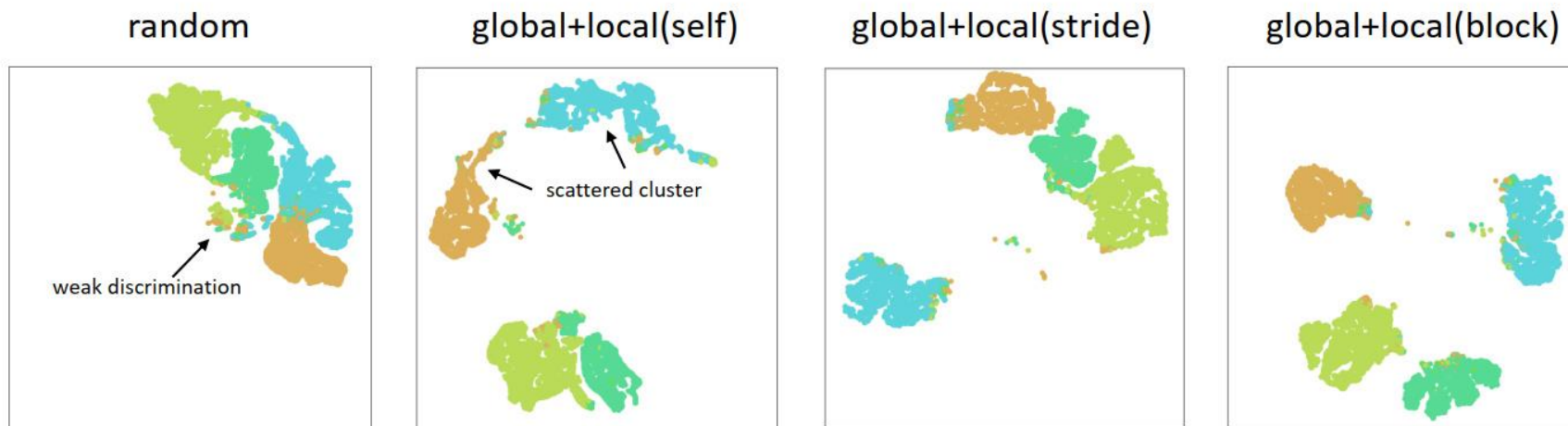
Experiment

Fig. 4: Visualization of embedding features after applying t-SNE on MMWHS with 40% labeled data. Different color represents different classes. The features are all from the feature map before the last point-wise convolution. (Better to view in colors.

Thanks