Self-Paced Contrastive Learning for Semi-supervised Medical Image Segmentation with Meta-labels

Jizong Peng[1] Ping Wang[1] Christian Desrosiers[1] Marco Pedersoli[1]

Introduction and Motivation

The scarcity of annotations hinders the use of deep learning in various medical applications.

- Semi-supervised and self-supervised learning (Fig. 1) exploit knowledge from unlabeled images.
- Self-supervised learning often relies on a pretext task to per-train the network.

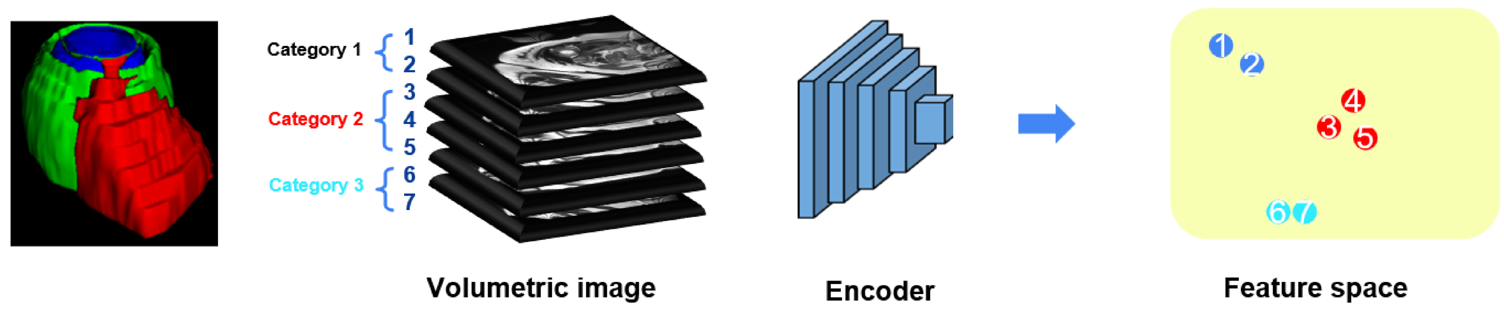


Figure 1: Scheme in [1] to pre-train encoder given slice position as meta-labels.

The pretext task is usually to predict the 2D slice position within a 3D volume. These tags termed as meta-labels are important for the quality of segmentation.

Drawback of current approaches

- Meta labels can be noisy (Fig. 2), leading to unstable performance.
- Limited meta-labels (only slice position) are used given 3D images.
- Represent. learning is only considered in the pre-training stage.

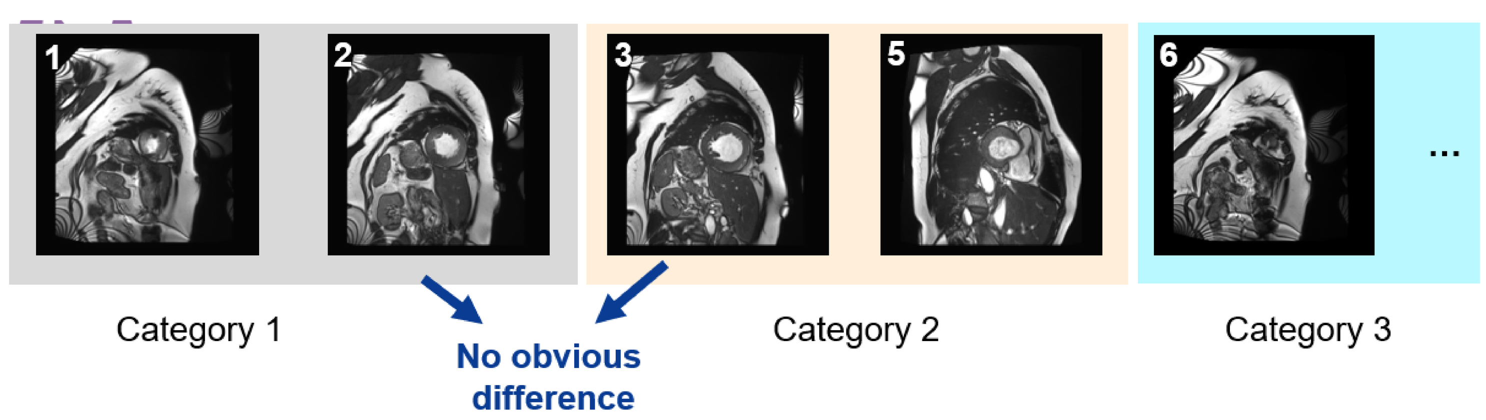


Figure 2: Noisy labels presented in slice position as meta-label [1]

Motivation

We explore a rich set of meta-labels to obtain a better representation during both pre-training and semi-supervised learning steps.

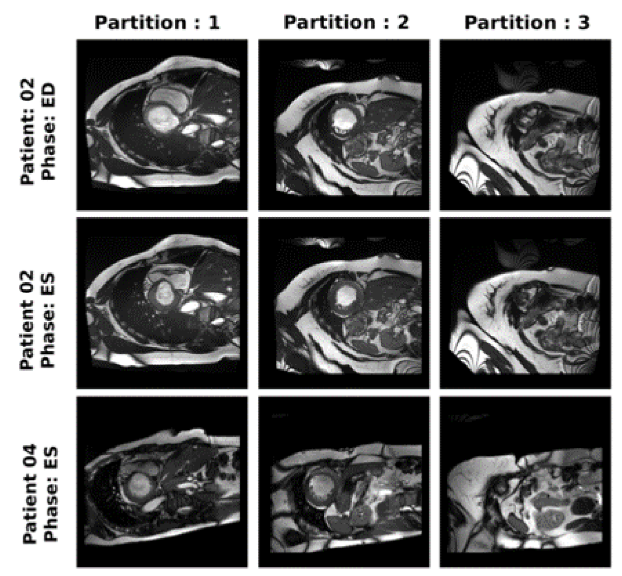
[1] K. Chaitanya, *et al.*, Contrastive learning of global and local features for medical image segmentation with limited annotations. NeurIPS, 2020

Method

Multiple Meta-Labels

Given a set of volumetric images, we extract multiple easy-to-obtain meta-labels from 2D slices:

- Slice position
- Scan ID (Patient ID)
- Cardiac phase



Self-paced Mechanism:

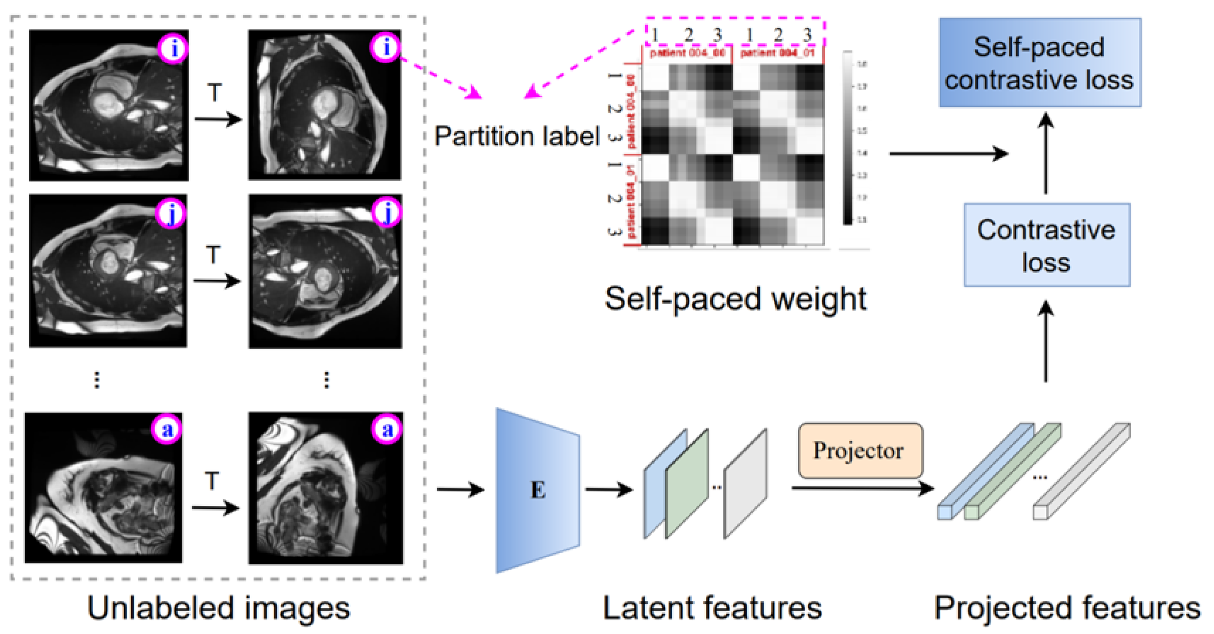


Figure 3: Our proposed self-paced scheme with meta-labels

$$\mathcal{L}^k_{\text{con}} = \frac{1}{2N}\sum_{i=1}^{2N}\frac{1}{|\mathcal{P}^k(i)|}\sum_{j\in\mathcal{P}^k(i)} \underbrace{-\log\frac{\exp\left(\mathbf{z}_i^{\intercal}\mathbf{z}_j/\tau\right)}{\sum_{a\in\mathcal{A}(i)}\exp\left(\mathbf{z}_i^{\intercal}\mathbf{z}_a/\tau\right)}}_{\ell_{ij}} \tag{1}$$

We introduce a self-paced coefficient w_{ij} measuring the reliability of contrastive pairs

$$\mathcal{L}^k_{\text{SP-con}} = \frac{1}{2N}\sum_{i=1}^{2N}\frac{1}{|\mathcal{P}^k(i)|}\sum_{j\in\mathcal{P}^k(i)} w_{ij}\,\ell_{ij} + R_\gamma(w_{ij}) \tag{2}$$

Two regularization strategies for w_{ij}

$$R^{\text{hard}}_\gamma(w_{ij}) = -\gamma\, w_{ij}\,; \qquad R^{\text{linear}}_\gamma(w_{ij}) = \gamma\big(\tfrac{1}{2}w_{ij}^2 - w_{ij}\big). \tag{3}$$

Final self-paced loss combining three meta labels:

$$\mathcal{L}_{\text{TOTAL}} = \sum_{k=1}^{K}\lambda_k\,\mathcal{L}^k_{\text{SP-con}} + \mathcal{L}_{\text{sup}} \tag{4}$$

Results

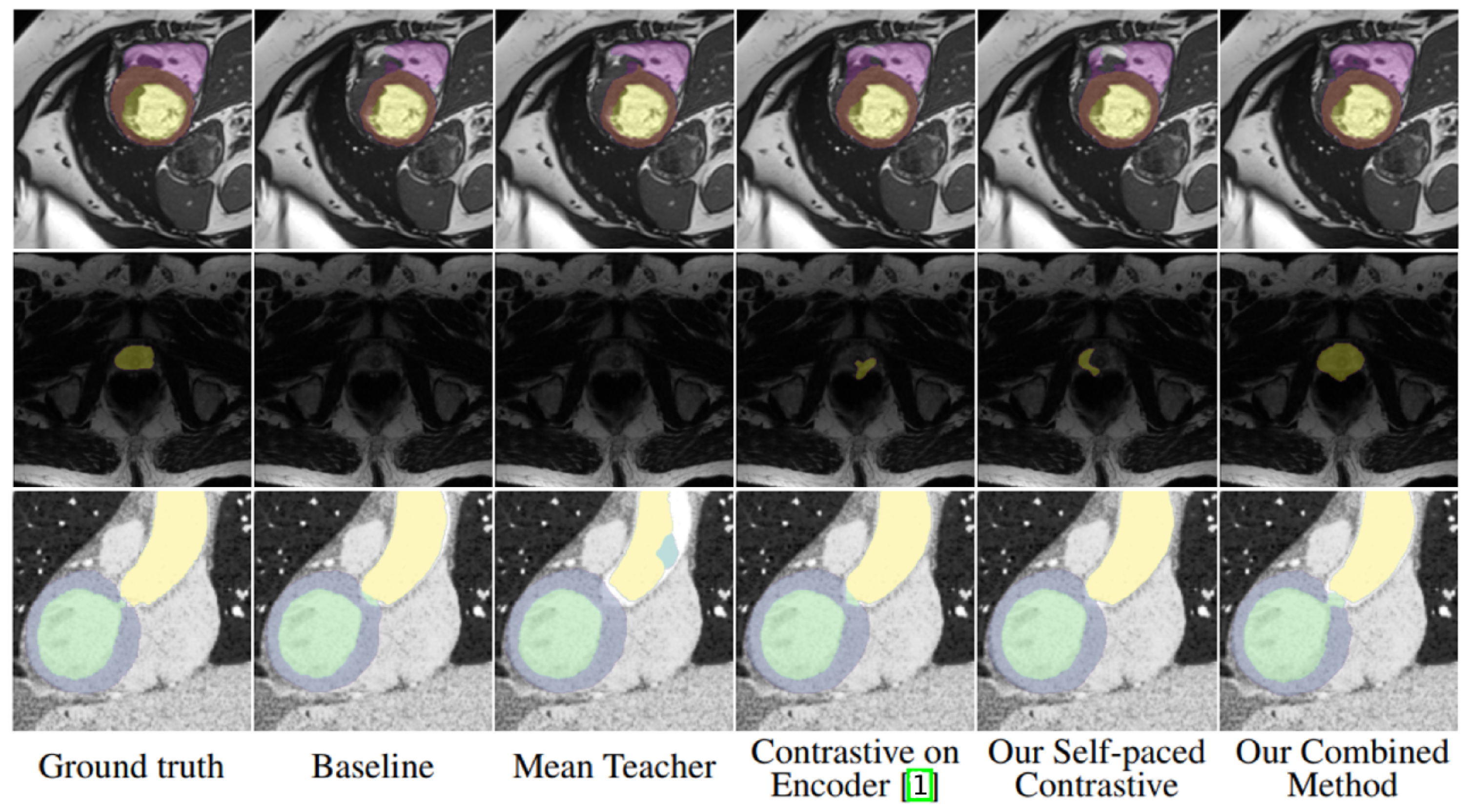


Table 3: 3D DSC performance (and standard deviation) of our method and other approaches on three medical image datasets with few labelled scans. Bold red-colored values are the best performing methods, underlined blue-colored ones correspond to the second best performing method.

Method	ACDC			PROMISE12			MMWHS	
	1 scan	2 scans	4 scans	3 scans	5 scans	7 scans	1 scan	2 scans
Entropy Min.	60.47 (1.03)	69.81 (0.99)	76.19 (1.21)	53.47 (5.70)	65.66 (0.42)	73.52 (2.71)	72.28 (0.58)	78.39 (1.54)
Mix-up	60.87 (1.28)	67.45 (1.04)	76.18 (0.49)	41.38 (2.80)	64.55 (1.93)	73.56 (0.61)	71.50 (0.54)	80.12 (0.84)
Adv. Training	63.05 (0.80)	70.68 (0.27)	75.89 (0.94)	61.58 (2.10)	71.00 (1.20)	81.05 (1.34)	73.47 (1.42)	80.40 (0.93)
Mean Teacher	62.85 (0.67)	72.84 (0.22)	79.12 (0.08)	52.96 (1.97)	68.38 (2.04)	77.37 (0.87)	72.36 (1.35)	81.01 (0.57)
Discrete MI	69.27 (1.41)	77.74 (0.42)	80.06 (0.24)	47.77 (3.58)	68.29 (2.35)	77.63 (1.13)	72.38 (1.04)	82.45 (1.36)
Contrastive	70.05 (2.66)	79.11 (2.02)	81.25 (2.15)	61.15 (2.95)	74.62 (1.69)	80.08 (1.39)	76.45 (0.62)	82.93 (0.42)
Our Method	**79.80** (0.33)	**83.20** (0.25)	**84.84** (0.15)	**74.47** (0.36)	**83.78** (0.30)	**84.52** (0.17)	**78.97** (0.52)	**84.87** (0.11)

Performance comparison with 6 state-of-art approaches for semi-supervised segmentation: Our proposed method outperforms them by a large margin. More results on pretrain-fine-tune /pretrain-semi-supervised setting are in paper / suppl. material.

Conclusions

- We propose a self-paced strategy for contrastive learning with noisy weak labels.
- We demonstrate the usefulness of a contrastive loss on meta-data in both pre-training and semi-supervised settings.
- We show that combining multiple meta-labels improves performance on the final task.

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE Université du Québec ÉTS Le génie pour l'industrie

[1] Department of Software and IT Engineering, ETS Montreal, Canada

Neurips 2021, Paper ID: 26698