Le génie pour l'industrie

MIDL 2020 presentation

Mutual information deep regularization for semi-supervised segmentation

J. Peng, M. Pedersoli, C. Desrosiers

Outline

We proposed a semi-supervised segmentation method for medical image regularized by Mutual Information

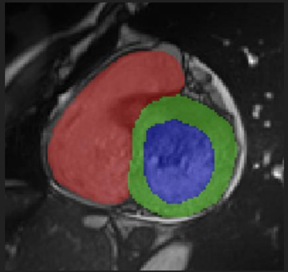
- Introduction on Semi-Supervised Segmentation
- Explanation on Mutual information concept for segmentation
- Our proposed scheme
- Experimental setup and results
- Conclusion

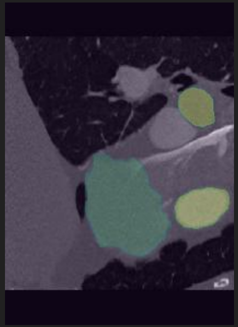
Medical Image segmentation

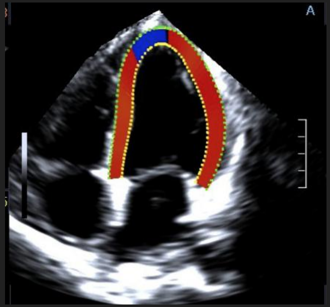
- Important stage for quantification (volume), visualization, intra-operative navigation, radiotherapy and clinical-oriented analysis
- Widely employed with CT MRI, and X-ray and Ultrasound for organs such as brain, lung, spleen, prostate organs.

Modalities: MRI CT Ultrasound

Cardiac Segmentation



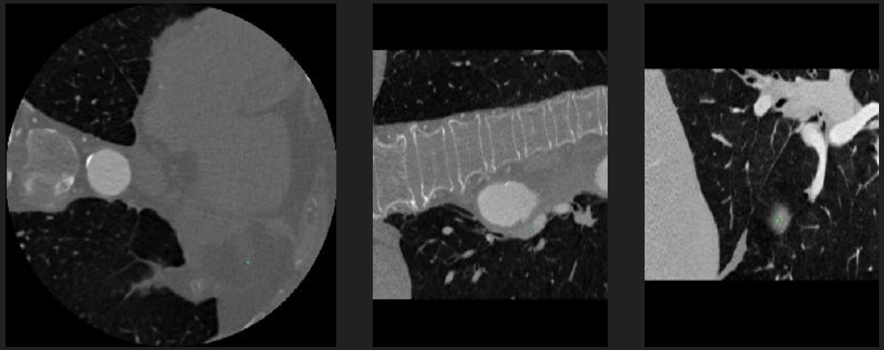




Limited access to annotations

Well know constraint for MIDL:

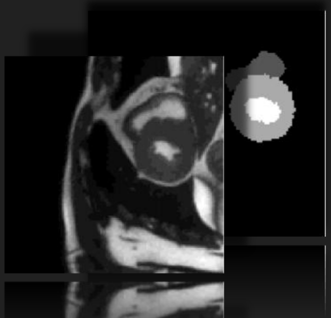
- require experts with long experience
- require annotation slice by slice
- cost hours for a single patient
- privacy requirement



Labeling is hard for 3D volumns
Three different views of a patient from [1]

ref: [1]: Zhuang, Xiahai, et al. "Evaluation of algorithms for Multi-Modality Whole Heart Segmentation: An open-access grand challenge." Medical image analysis 58 (2019): 101537.

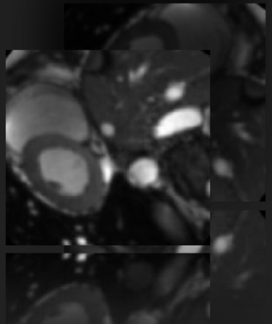
Semi-supervised learning framework



Labeled images

Supervised loss:

- ➢ Dice loss [1]
- ➢ Cross entropy loss [2]
- ➢ Uncertainty based loss [3]



Unlabeled images

Regularization loss:

- ➢ Consistency based reg. [4-7]
- ➢ Prior-enabled based reg.[8]
- ➢ Entropy based reg.[9]
- ➢ Mutual information based reg.

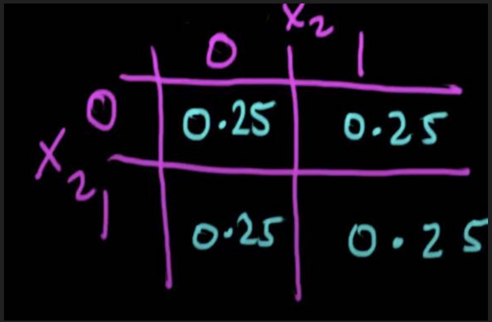
Ref:
[1]: Sudre, Carole H., et al. [2]:Zhang, Zhilu, and Mert Sabuncu, 2018. [3]: Kendall, Alex, et al.,CVPR 2018. [4]: Perone, et al., 2018. [5]: Li, Xiaomeng, et al., 2018. [6]: Wang, Fan, et al.,.ICCV 2013, [7]: Cui, Wenhui, et al. 2019. [8]: Zheng, Han, et al. MICCAI 2019. [9]: Vu, Tuan-Hung, et al. CVPR 2019.

Mutual information

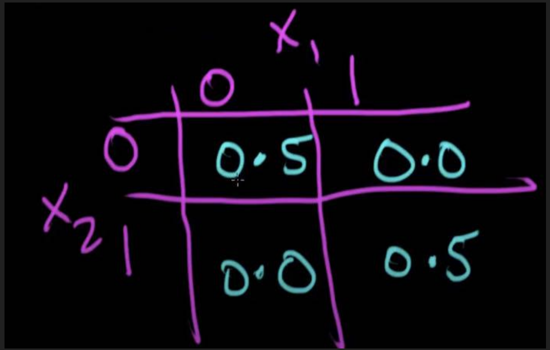
measures the amount of information that two variables *X*, *Y* share:.

$$I(X;Y) = D_{\mathrm{KL}}\big(p(X,Y)\,||\,p(X)\,p(Y)\big)$$

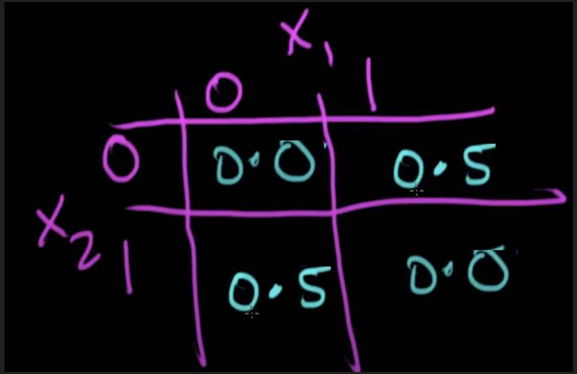
if *X*, *Y* are independent, then $p(X,Y) = p(X)p(Y)$, $I(X;Y) = 0$



MI = 0



MI Maximized



MI Maximized

Mutual information for different transformation

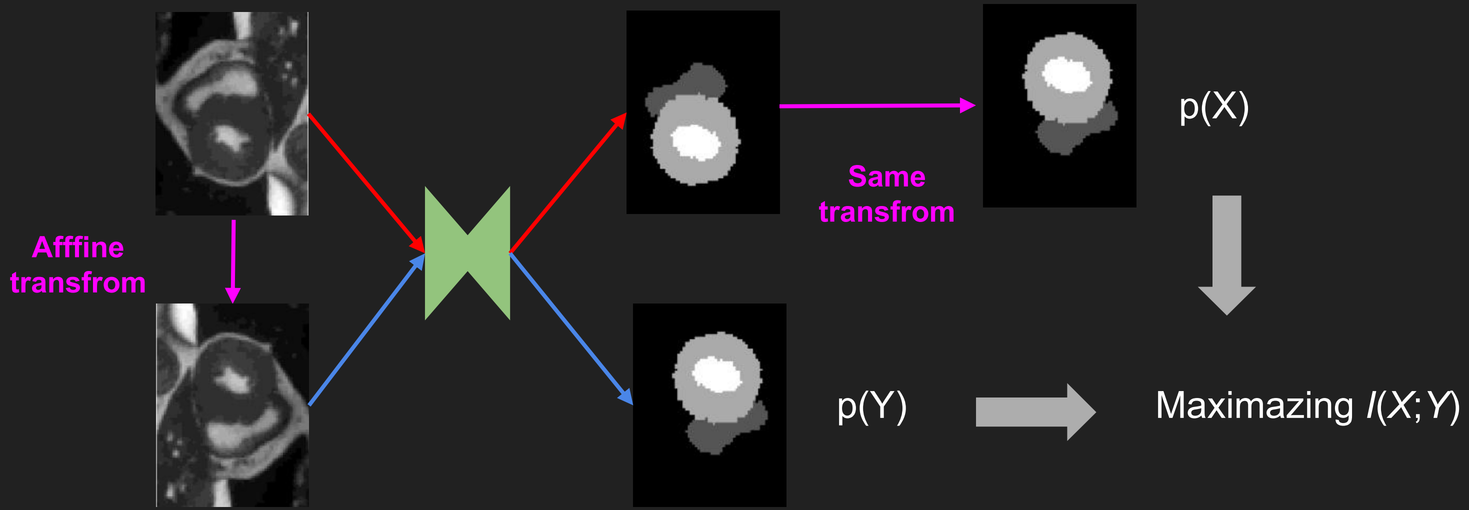
measures the amount of information that two variables *X*, *Y* share:.

$$I(X;Y) = D_{\mathrm{KL}}\left(p(X,Y)\,||\,p(X)\,p(Y)\right)$$

How about having X and Y as a segmentation distribution?

Case1:

We explore the consistency with MI



Mutual information on nearby patches

measures the amount of information that two variables *X*, *Y* share:.

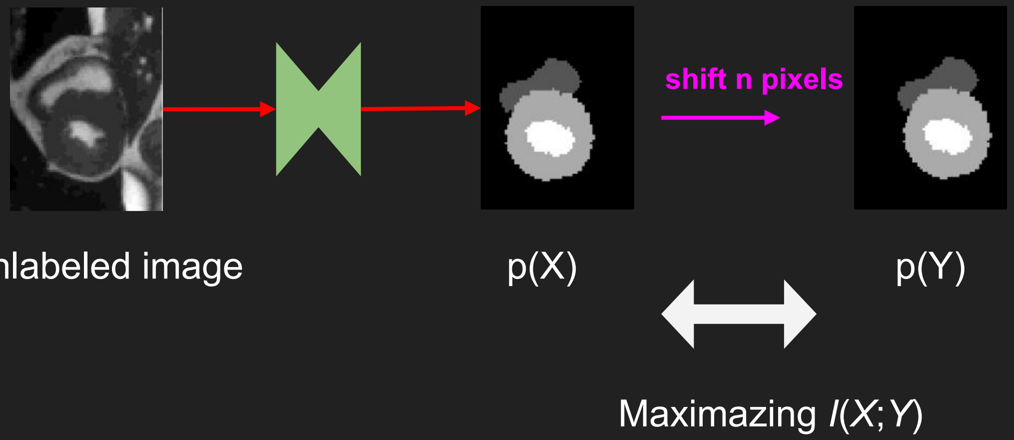
$$I(X;Y) = D_{\mathrm{KL}}\big(p(X,Y)\,||\,p(X)\,p(Y)\big)$$

How about having X and Y as a segmentation distribution?

Case2:

We explore the structural information of nearby patches by maximazing the MI.

MI does not require a strict assignment mapping.



How to compute I(X; Y) for 2D?

We compute the joint distribution by using product of the two marginal distribution (conditionally independent given the same input image)

$$\mathbf{P}_{pq} = \frac{1}{|\mathcal{D}_u|\,|\mathcal{T}|\,|\Omega|} \sum_{\mathbf{x}\in\mathcal{D}_u} \sum_{T\in\mathcal{T}} \sum_{(i,j)\in\Omega} \mathbf{f}_{ij} \cdot (\mathbf{f}^{T}_{i+p,j+q})^{\top}$$

2D convolution

We compute the MI from the joint probability matrix.

$$I(\mathbf{P}) = \sum_{k=1}^{C} \sum_{k'=1}^{C} \mathbf{P}(k,k') \cdot \log \frac{\mathbf{P}(k,k')}{\left(\sum_{k'} \mathbf{P}(k,k')\right) \cdot \left(\sum_{k} \mathbf{P}(k,k')\right)}.$$

The proposed scheme

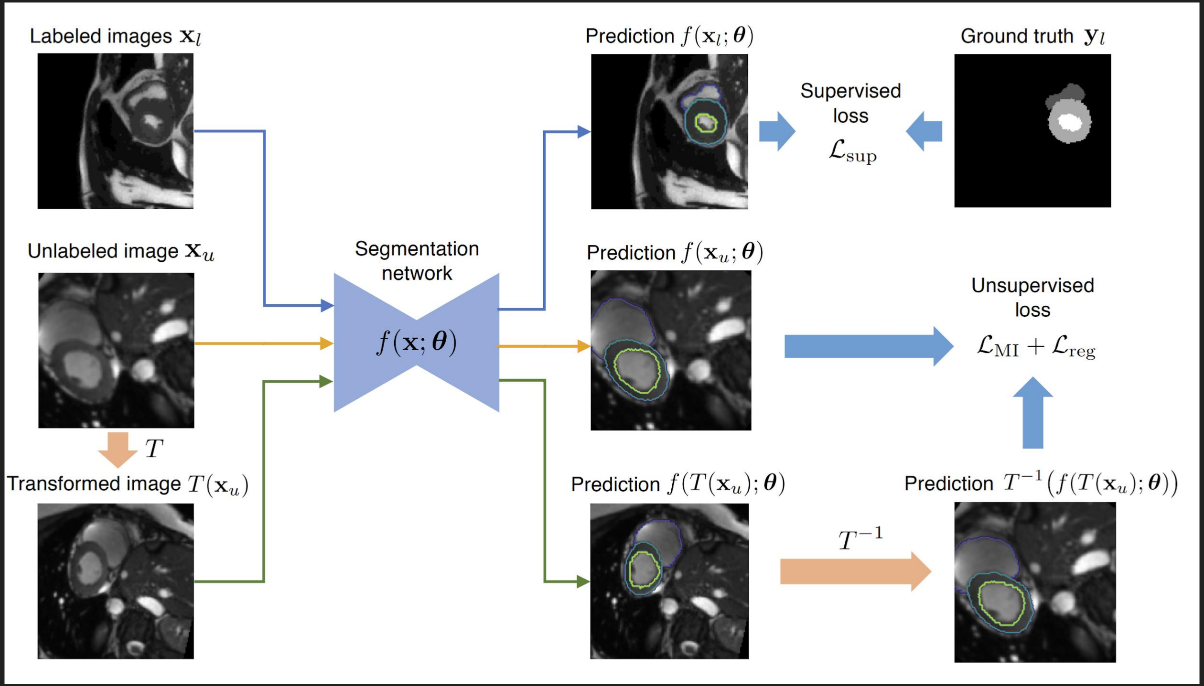
We examine the proposed idea in a semi supervised learning setting on three benchmark datasets: ACDC, prostate and spleen.

ACDC: 4% data as labeled, 83.5% as unlabeled, 12.5% as validation

Prostate: 14% data as labeled, 66% as unlabeled, 20% as validation

Spleen: 10% as labeled data, 78% as unlabeled, 12% as validation.

We compared our method against mean teacher [1] and entropy minimization [2]



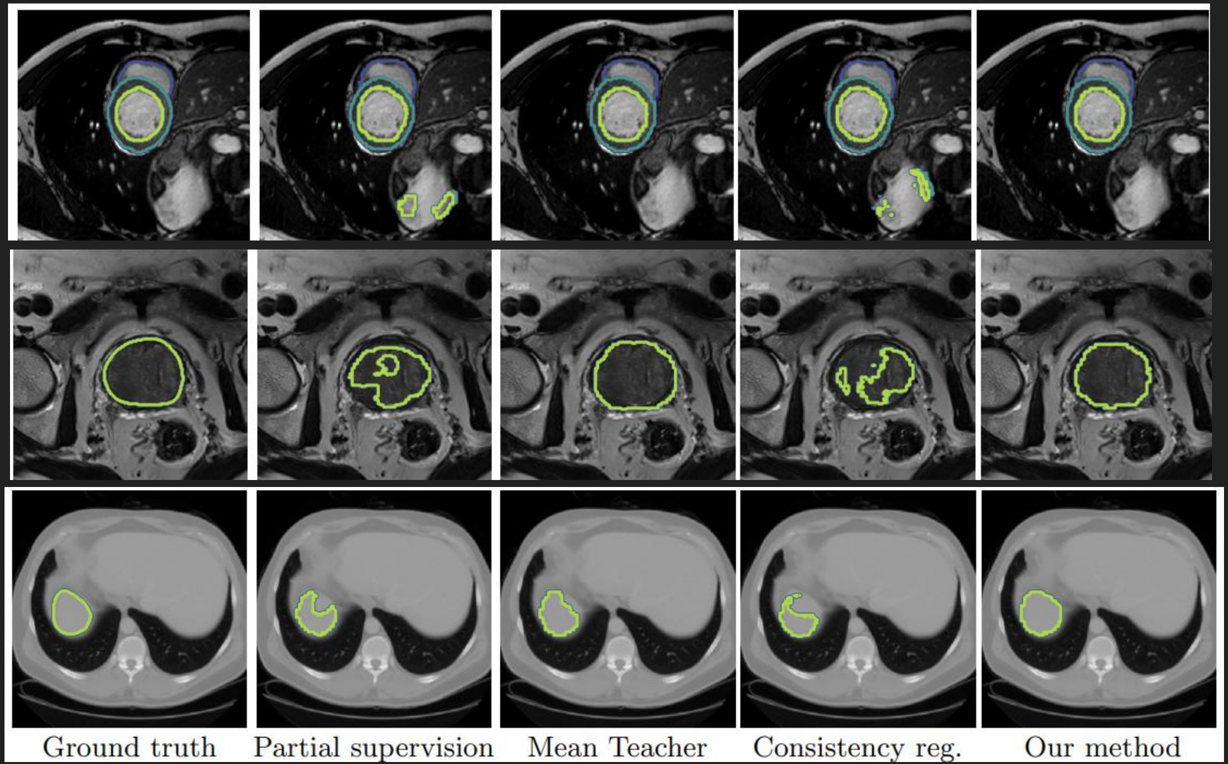
ref:
[1]: Perone, Christian S., and Julien Cohen-Adad. "Deep semi-supervised segmentation with weight-averaged consistency targets." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2018. 12-19.
[2]: Vu, Tuan-Hung, et al. "Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019.

Experimental results

Table 1: Mean 3D DSC of tested methods on the ACDC, Prostate and Spleen datasets. RV, Myo and LV refer to the right ventricle, myocardium and right ventricle classes, respectively. We test our method using KL and MSE for $\mathcal{L}_{\text{reg}}$. Mutual information corresponds to our method without loss term $\mathcal{L}_{\text{reg}}$ and Consistency regularization to our KL-based method without $\mathcal{L}_{\text{MI}}$. Reported values are averages (standard deviation in parentheses) for 3 runs with different random seeds.

	ACDC					
	RV	**Myo**	**LV**	**Mean**	**Prostate**	**Spleen**
Full supervision	88.98 (0.09)	84.95 (0.15)	92.44 (0.33)	88.79 (0.13)	87.33 (0.40)	93.52 (0.48)
Partial supervision	73.25 (0.36)	75.54 (1.27)	86.89 (0.26)	78.56 (0.42)	84.20 (0.73)	87.38 (1.05)
Entropy min.	73.85 (1.29)	74.92 (0.85)	86.12 (0.53)	78.30 (0.87)	83.04 (0.51)	90.21 (0.31)
Mean Teacher	82.99 (0.49)	80.43 (1.02)	89.33 (0.33)	84.25 (0.56)	86.15 (0.19)	93.22 (0.34)
Mutual information	81.98 (0.62)	75.75 (0.47)	87.89 (0.11)	81.87 (0.32)	83.75 (1.21)	90.35 (0.36)
Consistency reg.	82.30 (0.60)	79.43 (0.81)	88.55 (0.37)	83.42 (0.48)	84.88 (0.54)	91.50 (0.61)
Ours (MSE)	82.82 (0.35)	79.91 (0.72)	88.84 (0.77)	83.85 (0.39)	85.77 (0.46)	93.12 (0.19)
Ours (KL)	85.08 (0.10)	81.08 (0.42)	90.72 (0.44)	85.63 (0.20)	86.63 (0.07)	93.37 (0.13)

Experimental results



ref: [1]: Zhuang, Xiahai, et al. "Evaluation of algorithms for Multi-Modality Whole Heart Segmentation: An open-access grand challenge." Medical image analysis 58 (2019): 101537.

Conclusion

- We proposed a semi-supervised learning algorithm employing mutual information as a deep regularization

- Mutual information can capture structural/schematic information of organs which are usually under regular shape.

- We examined the proposed idea on three benchmark dataset and have proven that the proposed method achieved significant improvement compared with baseline method and comparable performance compared with SOTA methods.