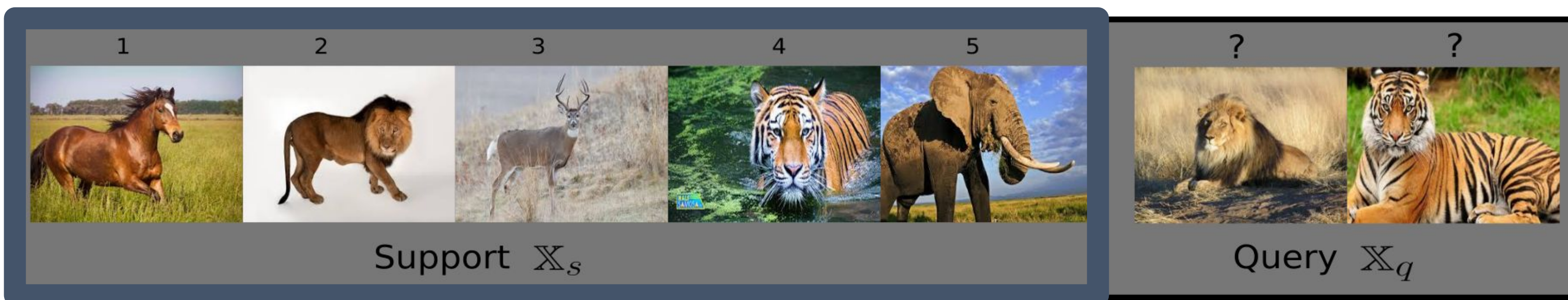
Few-Shot Learning
Apprendre avec peu de données

Ismail Ben Ayed
ETS Montreal
Ismail.benayed@etsmtl.ca

Few-shot learning (FSL)

Training on **base** classes

Few-shot tasks at testing time

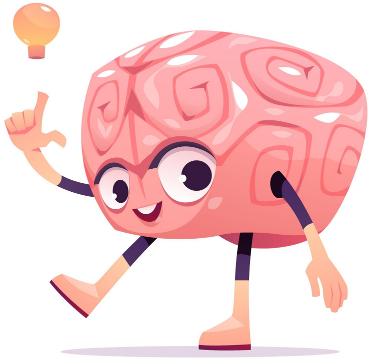


Classify these

Learn from a few examples per **new** class

Few-shot learning

- **Humans recognize easily** with few examples
- Modern ML generalize very poorly



Why it is interesting

Available data sets represent small sub-domains of the world

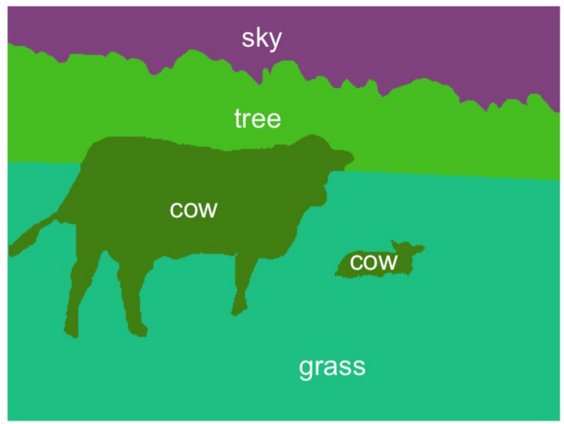
Cityscapes (5k images; 1.5h per image):
Urban scenes, less than 30 classes

A dense prediction task: semantic segmentation

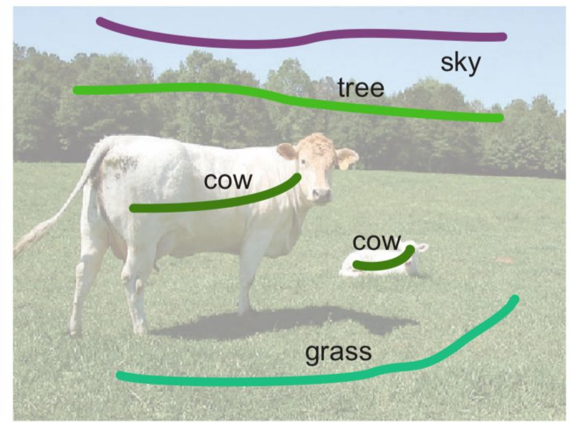
New classes, but with few examples

Semi-supervision

A lot of non-annotated data, and a fraction of points annotated



Full annotations

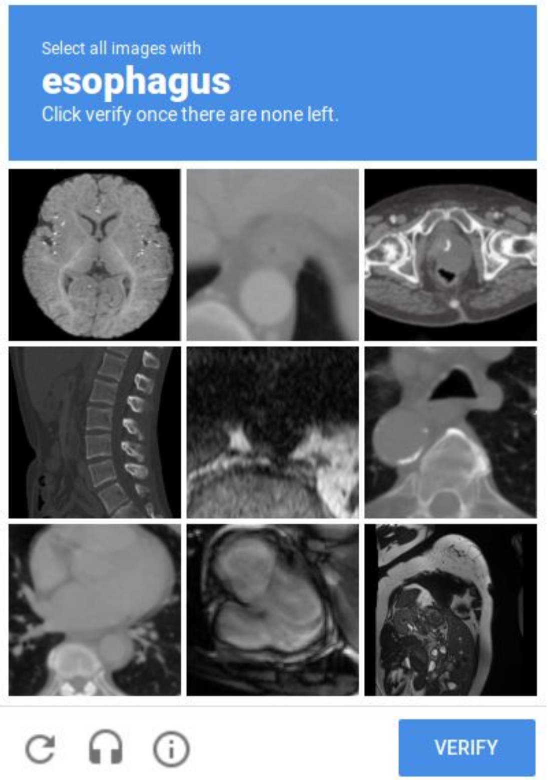


Semi-supervision

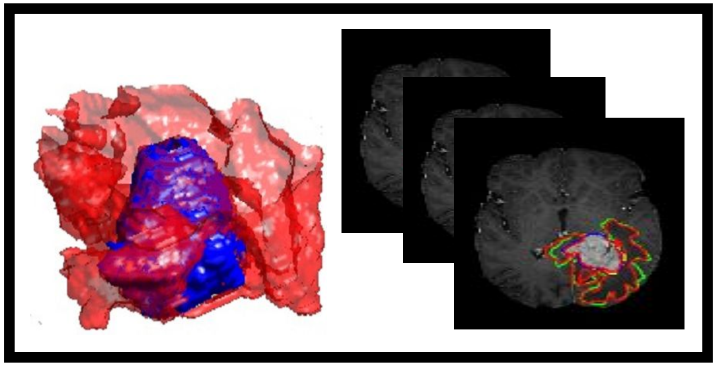
Figures from Lin et al. Scribblesup: Scribble-supervised convolutional networks for semantic segmentation, CVPR 2016

Labels: not only expensive, but might need expert knowledge

Crowdsourcing?

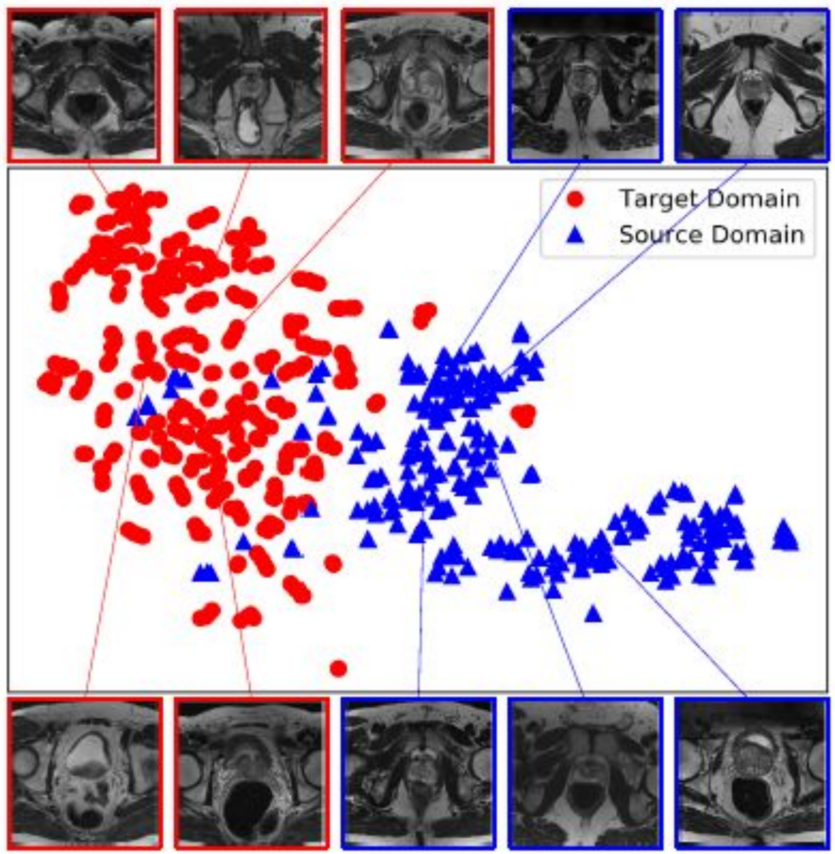


Dense 3D annotations: several hours (of radiologist time)



Domain shifts make things worse

Even with full annotations in one domain

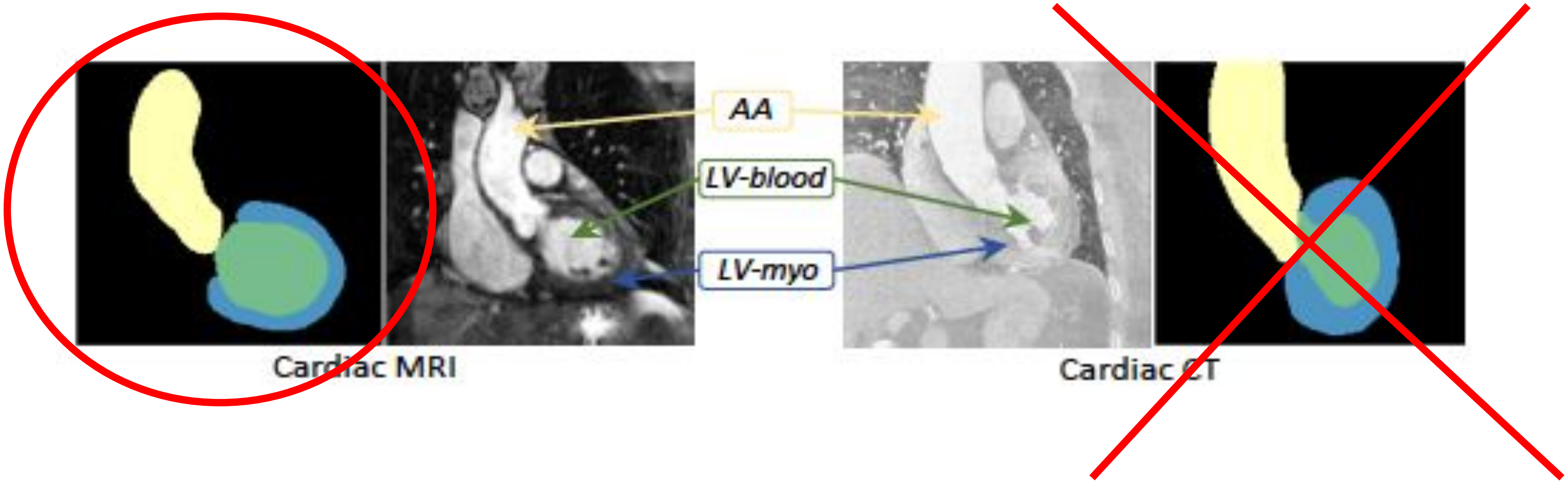


[MRI Prostate segmentation: Figure from Zhu et al., Boundary-weighted Domain Adaptive Neural Network for Prostate MR Image Segmentation ArXiv 2019]

Unsupervised domain adaptation

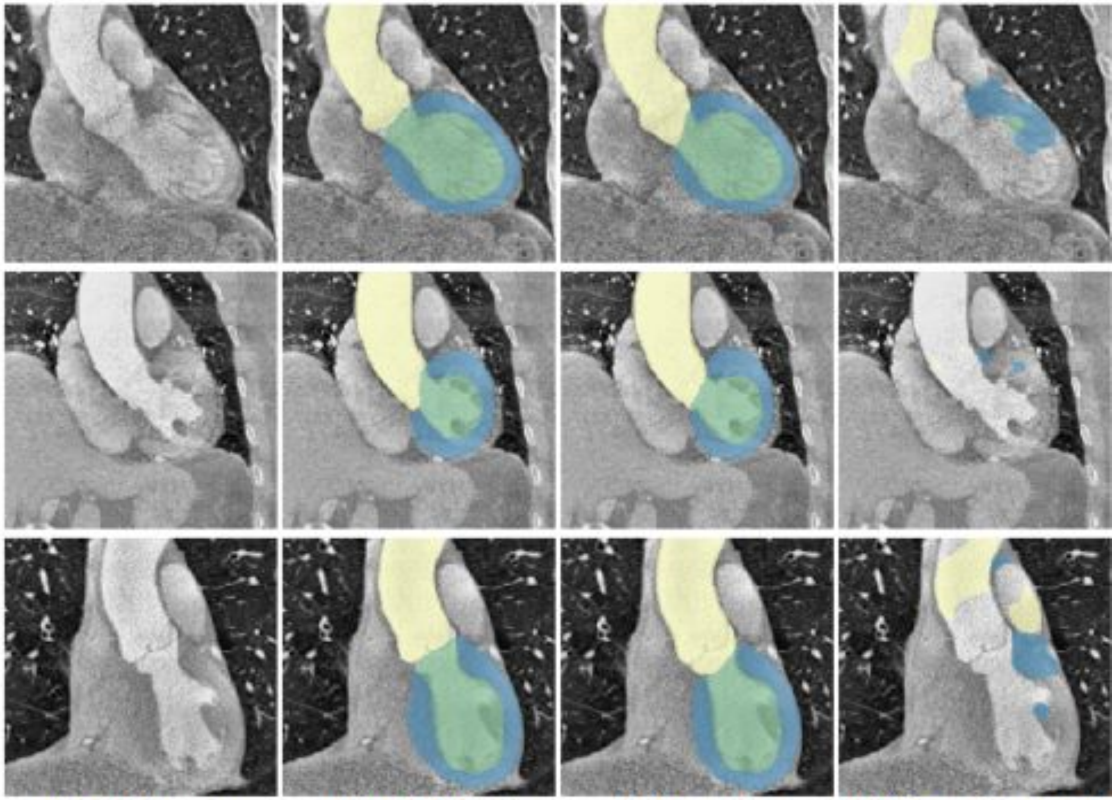
We have labels for
the source domain

No labels for the target



[Images from Dou et al., PnP-AdaNet: Plug-and-play adversarial domain adaptation network with a benchmark at cross-modality cardiac segmentation ArXiv 2018]

Bad generalization to the target

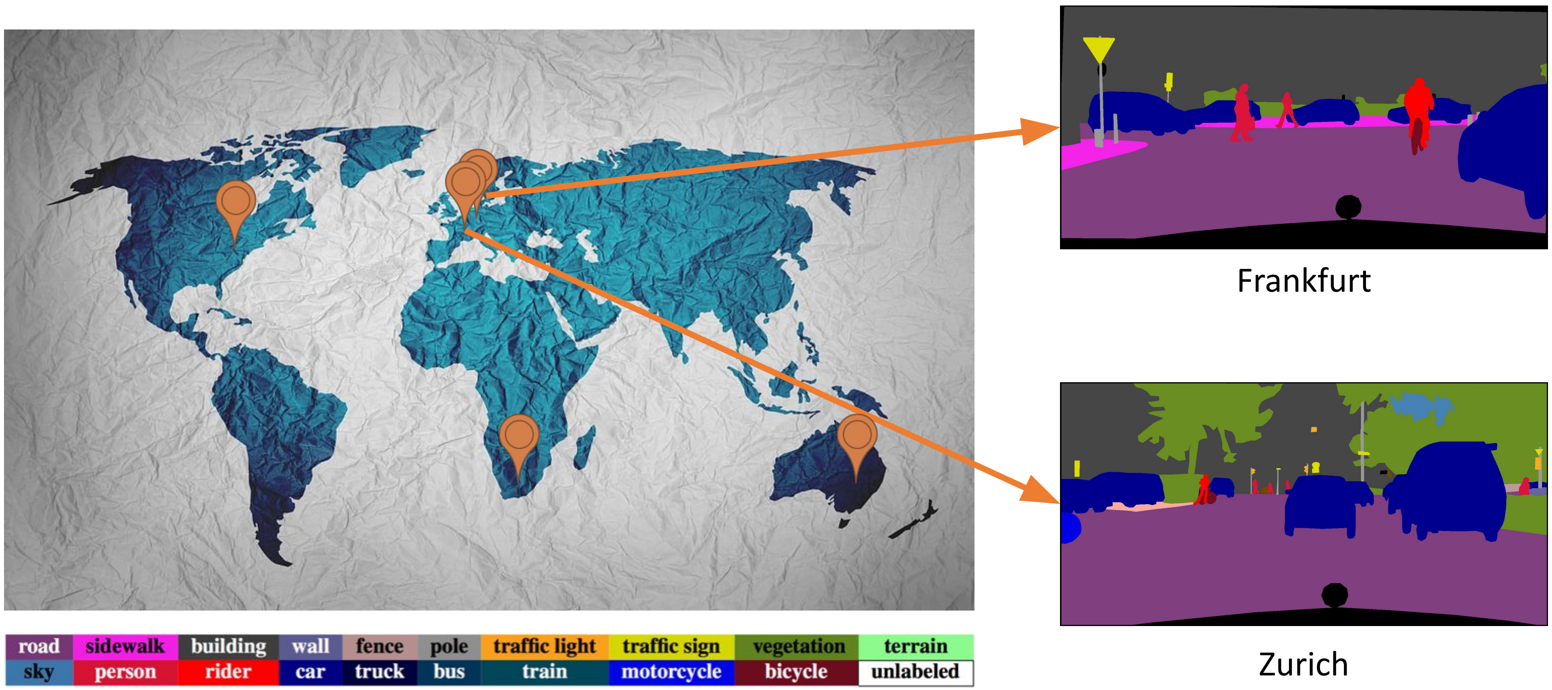


(a) CT image (b) CT label (c) Seg-CT (d) Seg-CT-noDA

[Images from Dou et al., PnP-AdaNet: Plug-and-play adversarial domain adaptation network with a benchmark at cross-modality cardiac segmentation ArXiv 2018]

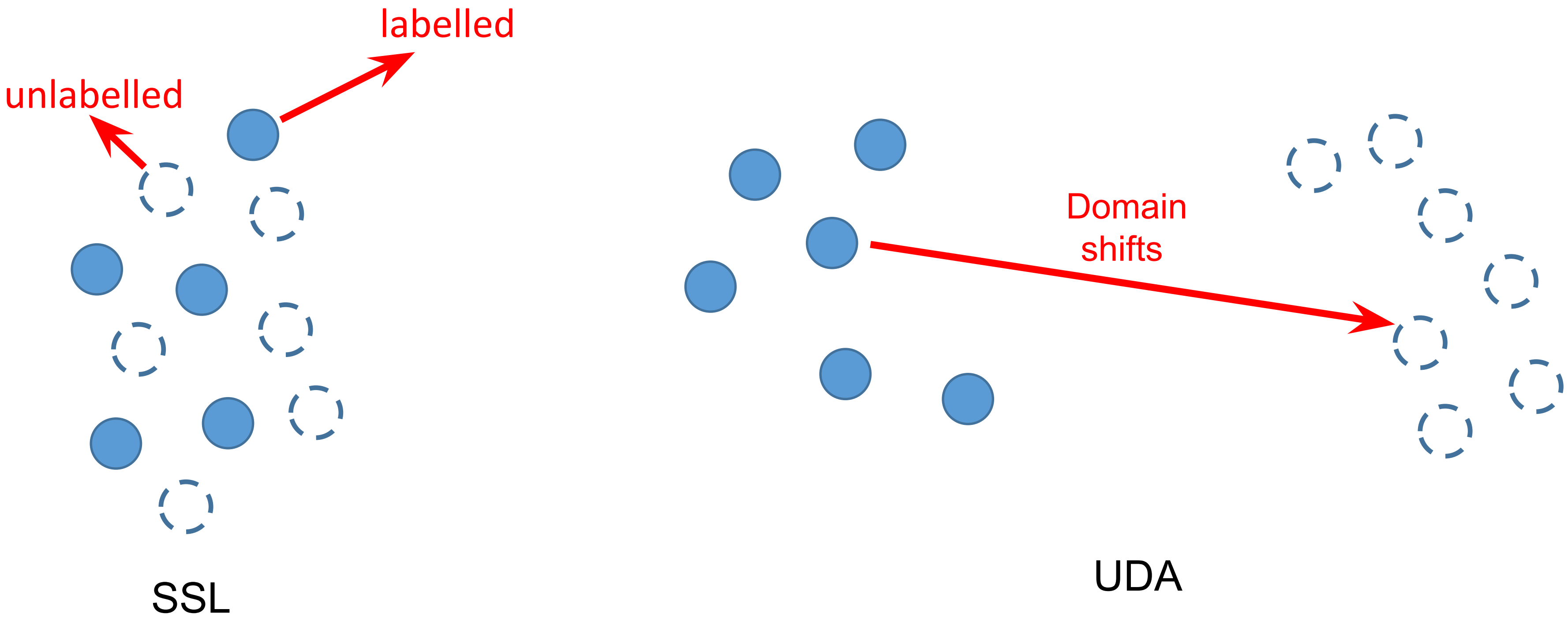
Wide interest in computer vision as well

Domain shifts are everywhere BUT we cannot label everywhere



Cityscapes (5000 images): labeling of 1 image takes 90 min at average [Cordt et al., CVPR 2016]

FSL, SSL and UDA are closely related problems



Few-Shot/SSL/UDA in a nutshell: We are leveraging unlabelled data with priors

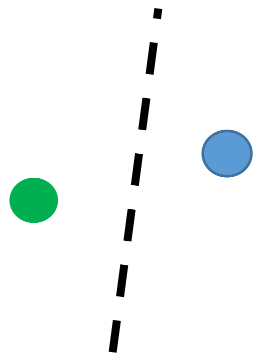
-- Structure-driven priors: Regularization

-- Knowledge-driven priors (e.g., anatomical constraints)

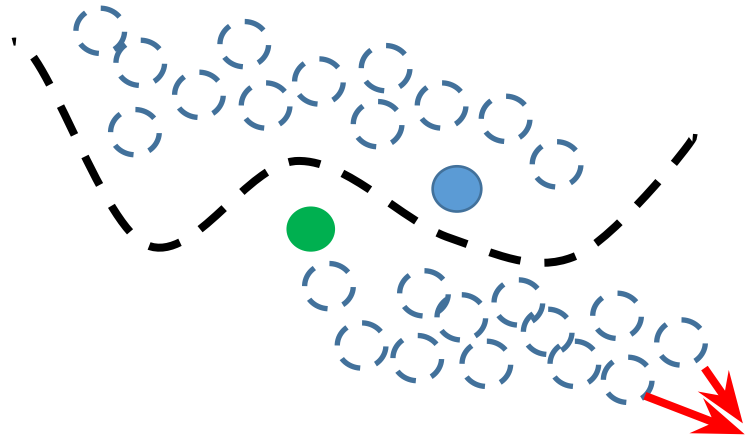
-- Invariance priors (e.g., contrastive learning)

-- Multi-modal priors (e.g., text info associated with the images)

Learning from labeled and unlabeled data



Only labeled points



Unlabeled points

$$\min_{\theta} \sum_{p \in \mathcal{L}} l(\mathbf{y}^p, \mathbf{s}_\theta^p) + \sum_{p,q \in \mathcal{L} \cup \mathcal{U}} w_{p,q} \|\mathbf{s}_\theta^p - \mathbf{s}_\theta^q\|$$

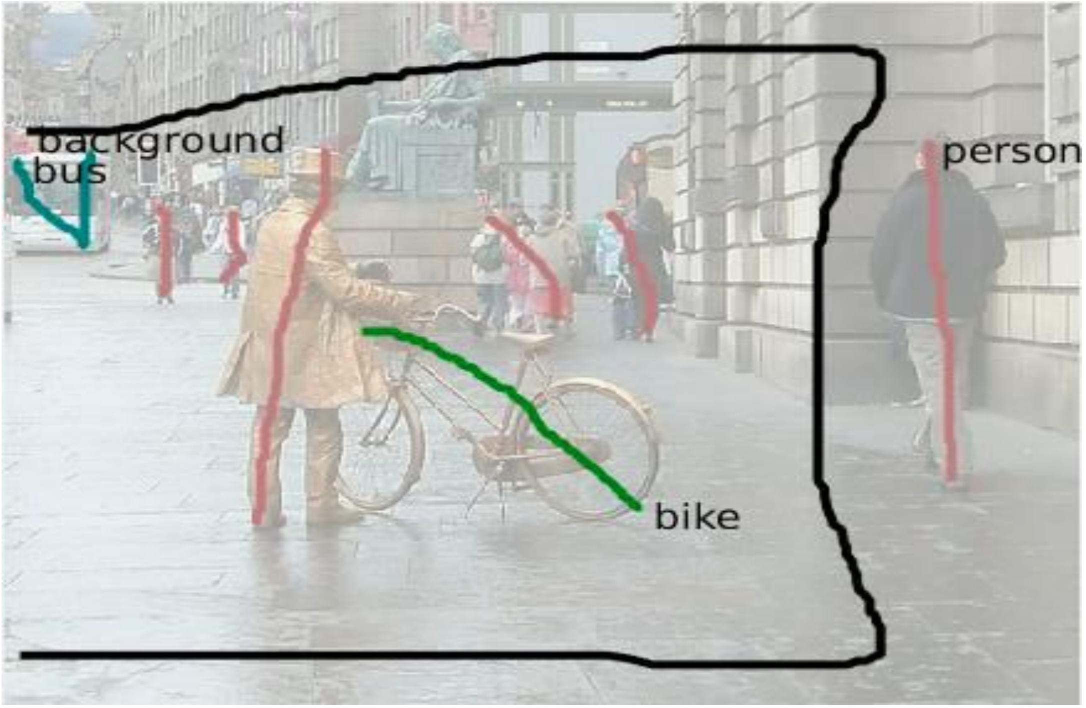
Labeled points

Unlabeled points

Regularized losses for segmentation

[Tang et al., ECCV 2018]

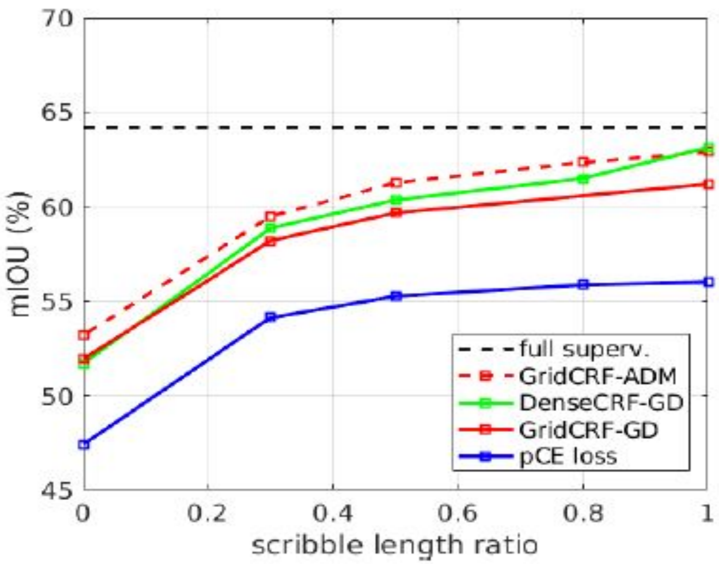
$$\min_{\theta} \sum_{p \in \mathcal{L}} l(\mathbf{y}^p, \mathbf{s}_\theta^p) + \sum_{p,q \in \mathcal{L} \cup \mathcal{U}} w_{p,q} ||\mathbf{s}_\theta^p - \mathbf{s}_\theta^q||$$

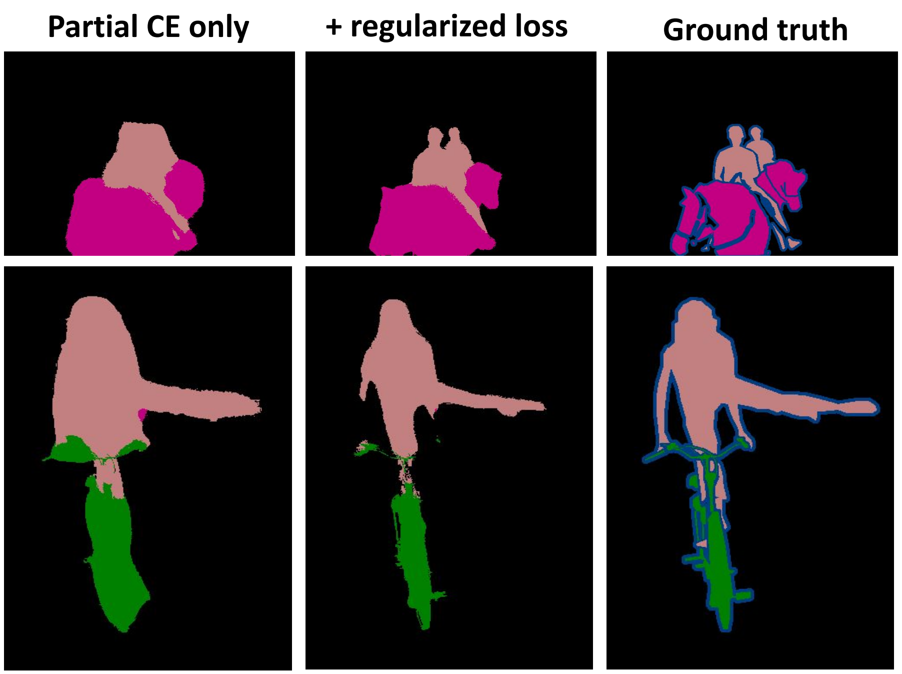


Regularized losses for segmentation

[Tang et al., ECCV 2018]

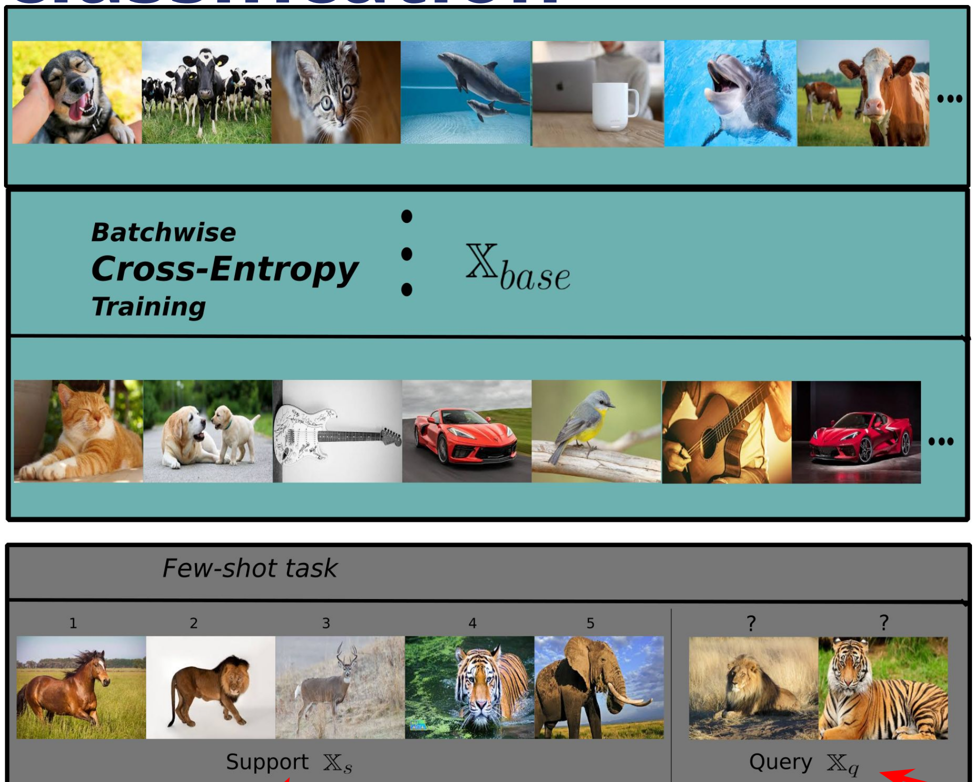
97.6% of full supervision performance with **3%** of the labels!





Laplacian regularized few-shot image classification

[Ziko et al., ICML 2020]



Basic training
(no meta-learning)

Supervision term with a few labeled samples

Laplacian regularization (unlabeled points)

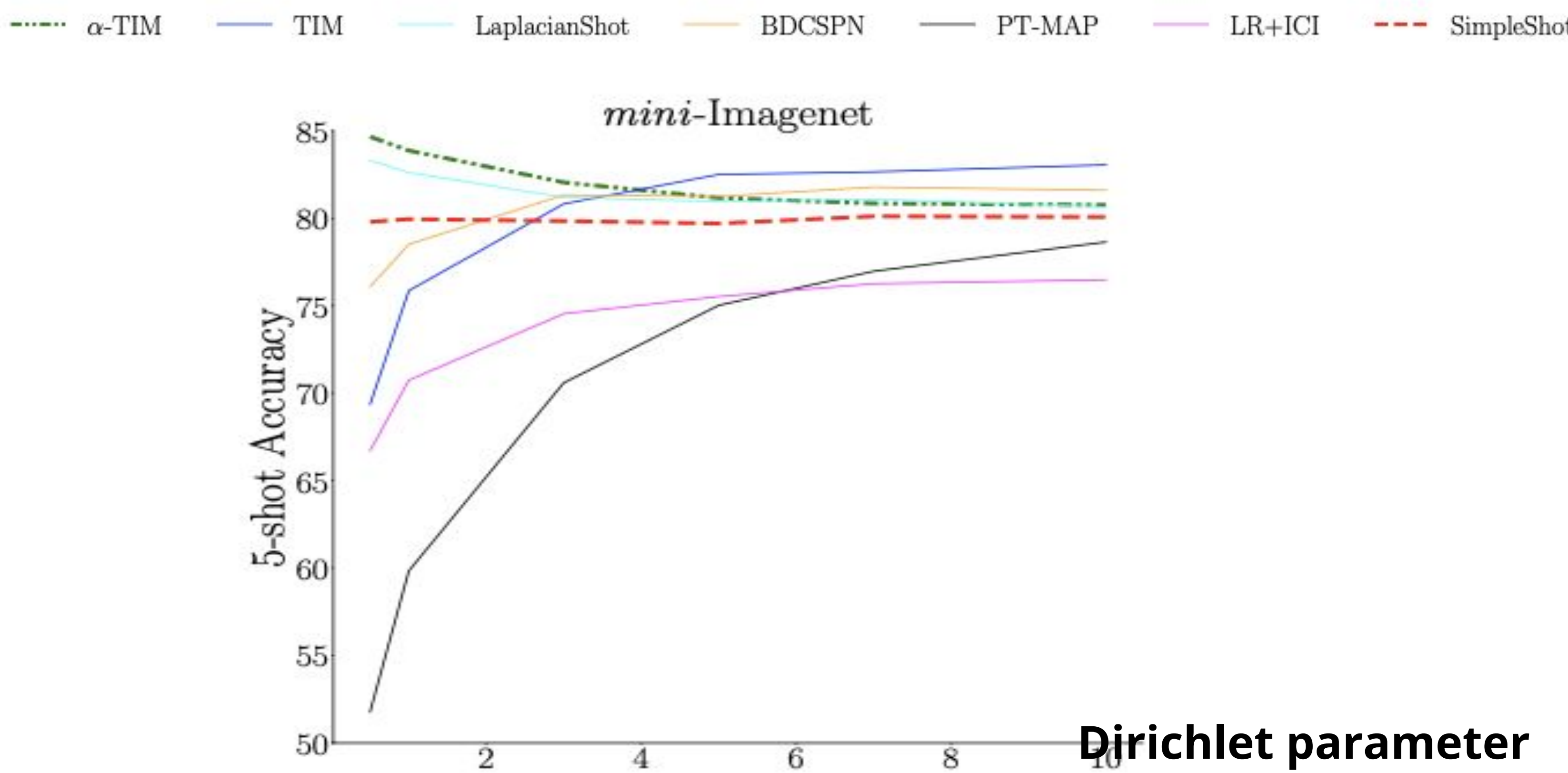
The results question and abundant meta-learning literature

Methods	Network	1-shot	5-shot
MAML [Finn et al., 2017]	ResNet-18	49.61 ± 0.92	65.72 ± 0.77
Chen [Chen et al., 2019]	ResNet-18	51.87 ± 0.77	75.68 ± 0.63
RelationNet [Sung et al., 2018]	ResNet-18	52.48 ± 0.86	69.83 ± 0.68
MatchingNet [Vinyals et al., 2016]	ResNet-18	52.91 ± 0.88	68.88 ± 0.69
ProtoNet [Snell et al., 2017]	ResNet-18	54.16 ± 0.82	73.68 ± 0.65
Gidaris [Gidaris and Komodakis, 2018]	ResNet-15	55.45 ± 0.89	70.13 ± 0.68
SNAIL[Mishra et al., 2018]	ResNet-15	55.71 ± 0.99	68.88 ± 0.92
AdaCNN [Munkhdalai et al., 2018]	ResNet-15	56.88 ± 0.62	71.94 ± 0.57
TADAM [Oreshkin et al., 2018]	ResNet-15	58.50 ± 0.30	76.70 ± 0.30
CAML [Jiang et al., 2019]	ResNet-12	59.23 ± 0.99	72.35 ± 0.71
TPN [Yanbin et al., 2019]	ResNet-12	59.46	75.64
TEAM [Qiao et al., 2019]	ResNet-18	60.07	75.90
MTL [Sun et al., 2019]	ResNet-18	61.20 ± 1.80	75.50 ± 0.80
VariationalFSL [Zhang et al., 2019]	ResNet-18	61.23 ± 0.26	77.69 ± 0.17
Transductive tuning [Dhillon et al., 2020]	ResNet-12	62.35 ± 0.66	74.53 ± 0.54
MetaoptNet[Lee et al., 2019]	ResNet-18	62.64 ± 0.61	78.63 ± 0.46
SimpleShot [Wang et al., 2019]	ResNet-18	63.10 ± 0.20	79.92 ± 0.14
CAN+T [Hou et al., 2019]	ResNet-12	67.19 ± 0.55	80.64 ± 0.35
LaplacianShot (ours)	ResNet-18	**72.11** ± 0.19	**82.31** ± 0.14

Several recent baselines: [Chen et al., ICLR'19]; [Tian et al., ECCV'20] [Dhillon et al., ICLR'20]; [Ziko et al., ICML'20]; [Boudiaf et al., NeurIPS'20]

More realistic benchmark: Further surprises

[Veilleux et al., NeurIPS 2021]



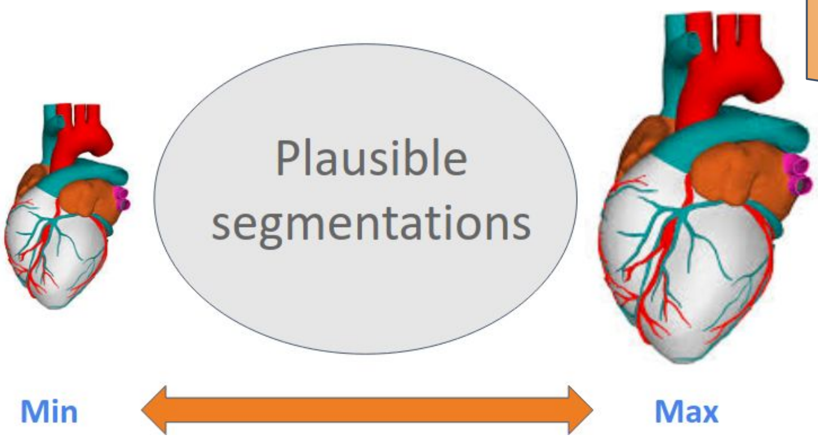
Dirichlet parameter

More realistic few-shot tasks (Dirichlet-sampled)

Constrained optimization (in deep networks)

Data meet domain knowledge

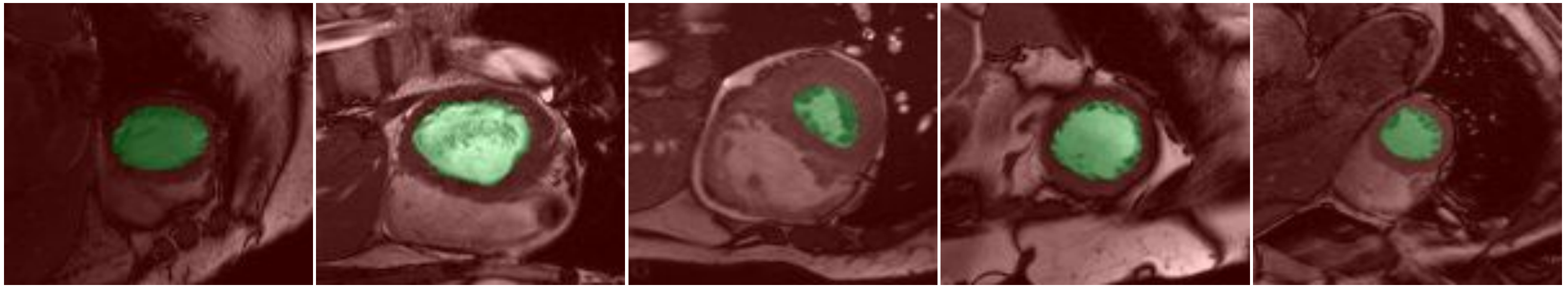
A lot of priors in medical imaging



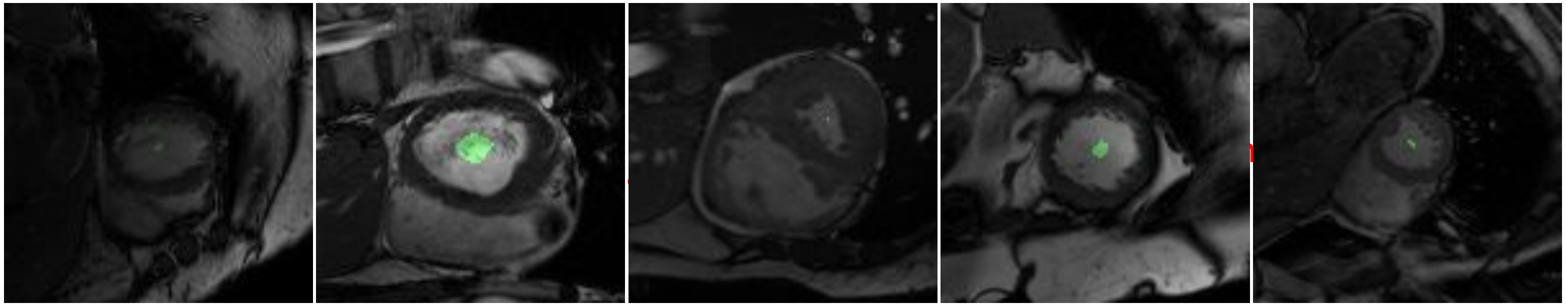
Anatomical priors (e.g., shapes)

Partially labeled data
(e.g., exploiting organ relationships)

Example: Left ventricle segmentation in cardiac MRI with volumetric constraints



Full annotations

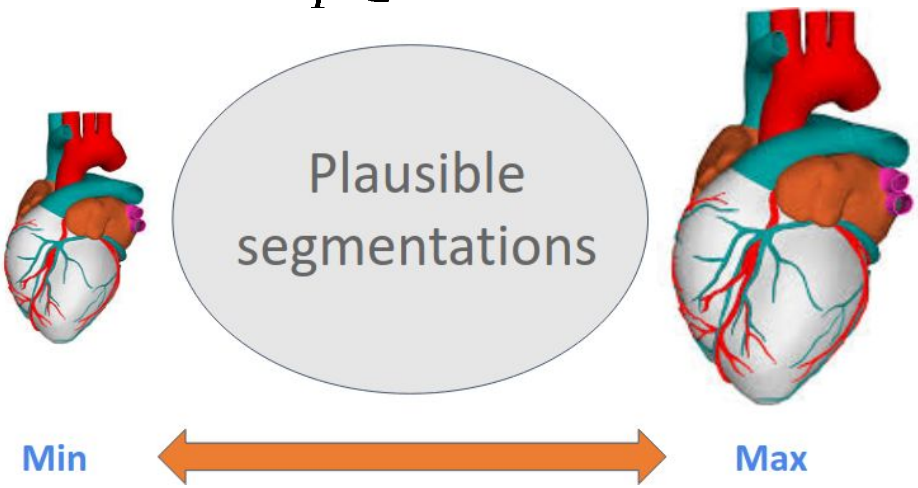


Partial annotations for cross-entropy

Example: Left ventricle segmentation in cardiac MRI with volumetric constraints

[Kervadek et al., MedIA 2019]

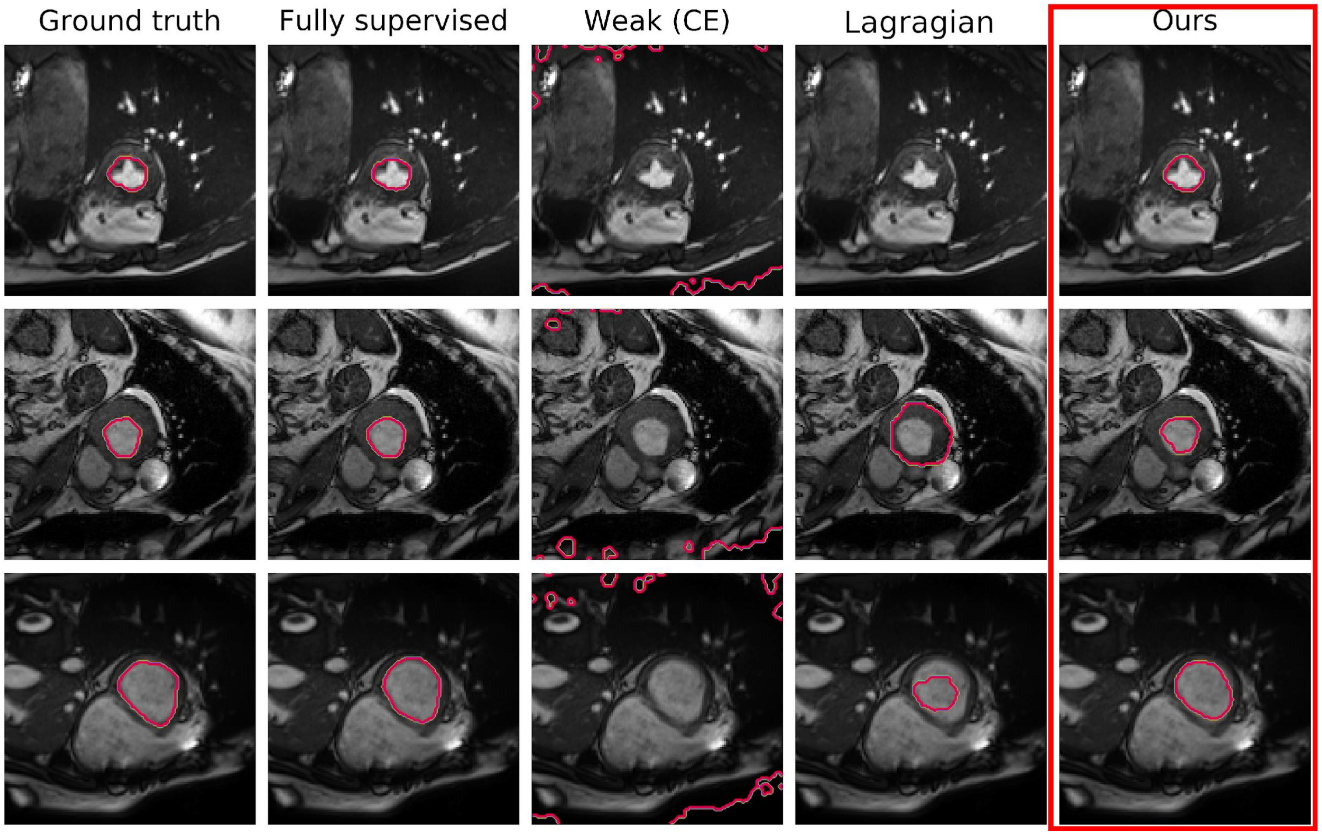
$$\min \leq \sum_{p \in \Omega} s_\theta^{p,l} \leq \max$$



Size information

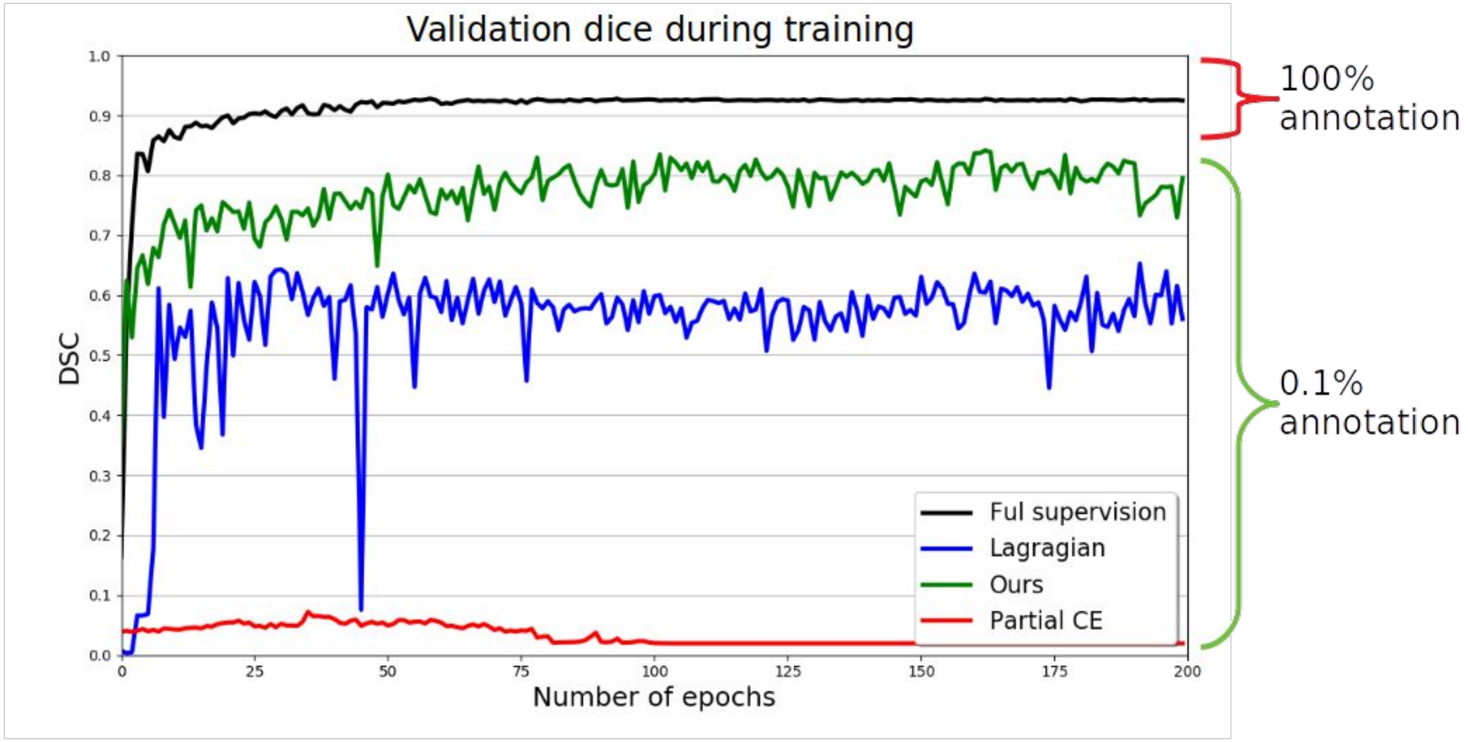
Example: Left ventricle segmentation in cardiac MRI with volumetric constraints

[Kervadek et al., MedIA 2019]



Example: Left ventricle segmentation in cardiac MRI with volumetric constraints

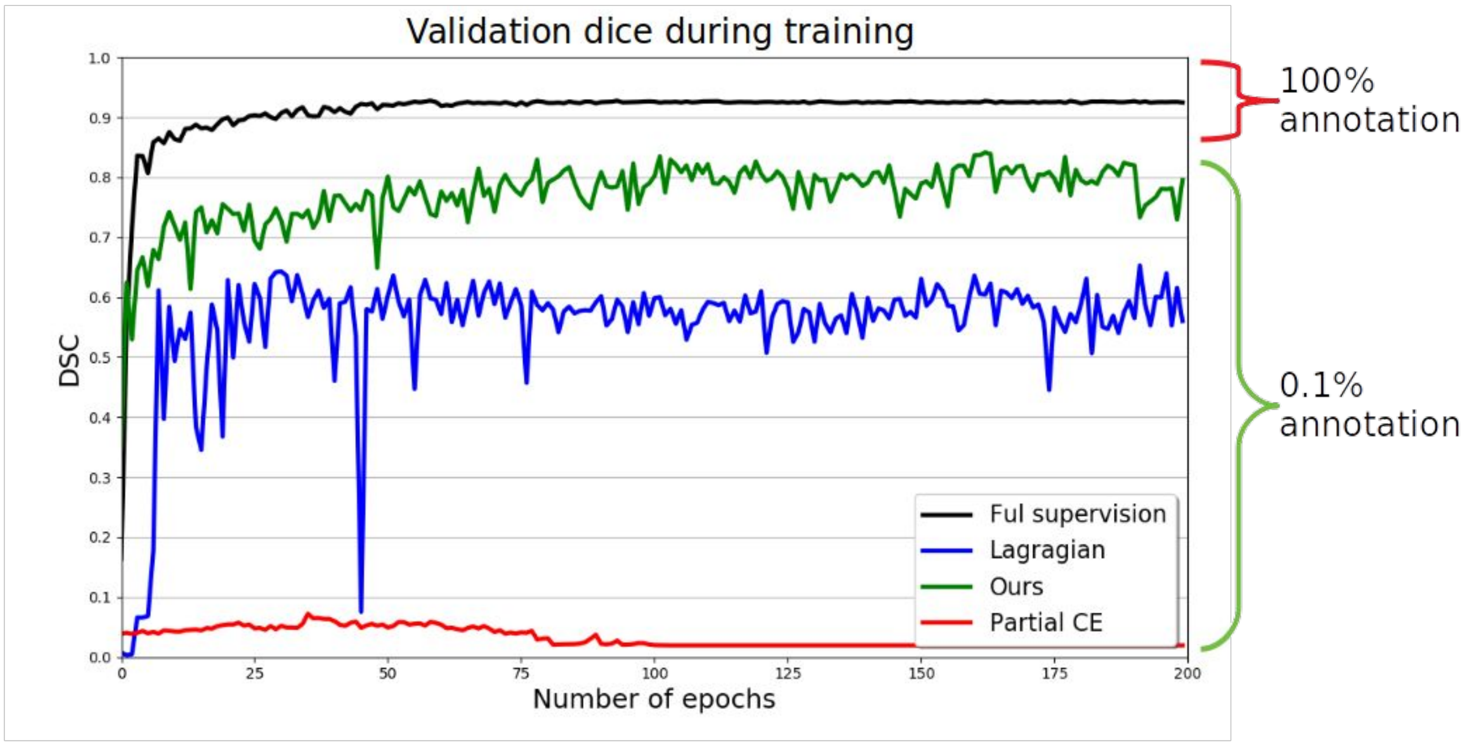
[Kervadek et al., MedIA 2019]



Example: Left ventricle segmentation in cardiac MRI with volumetric constraints

[Kervadek et al., MedIA 2019]

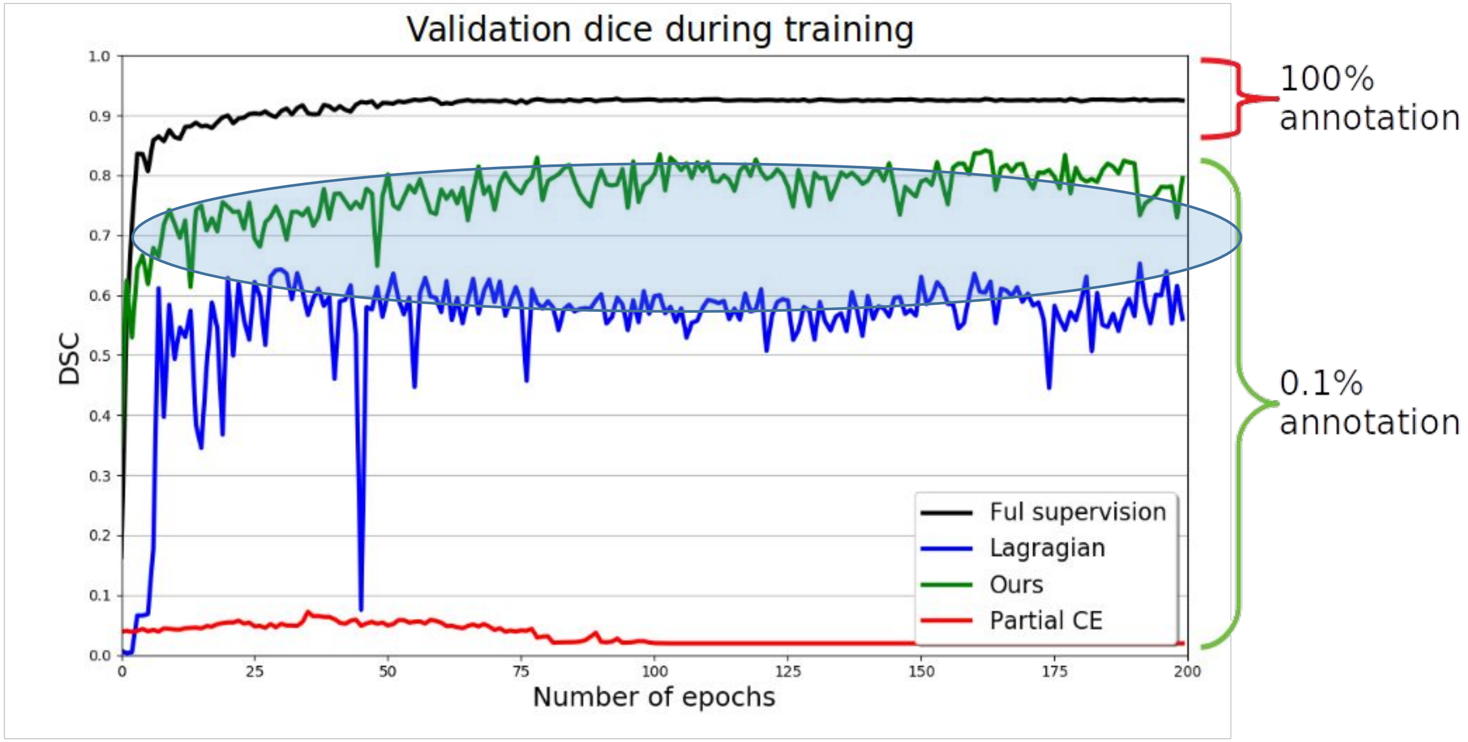
The exciting part: 90% of full supervision Dice with 0.1% of labels



Example: Left ventricle segmentation in cardiac MRI with volumetric constraints

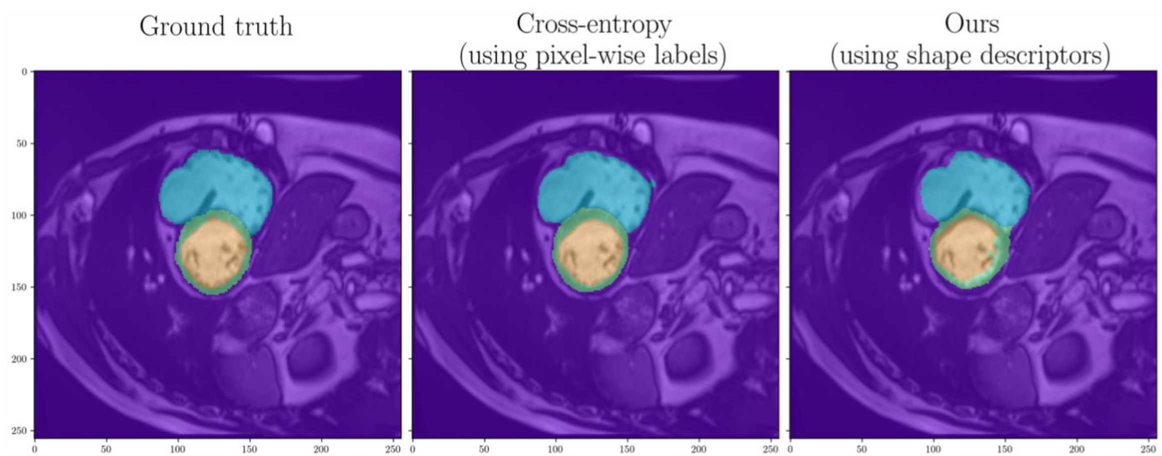
[Kervadek et al., MedIA 2019]

The surprising part: Lagrangian optimization is much worse than a simple penalty



Beyond size: Exploring shape priors as functions of network outputs

[Kervadek et al., MIDL 2021 (Best-paper award)]



(a) A visual comparison of the different supervision methods on the ACDC dataset.

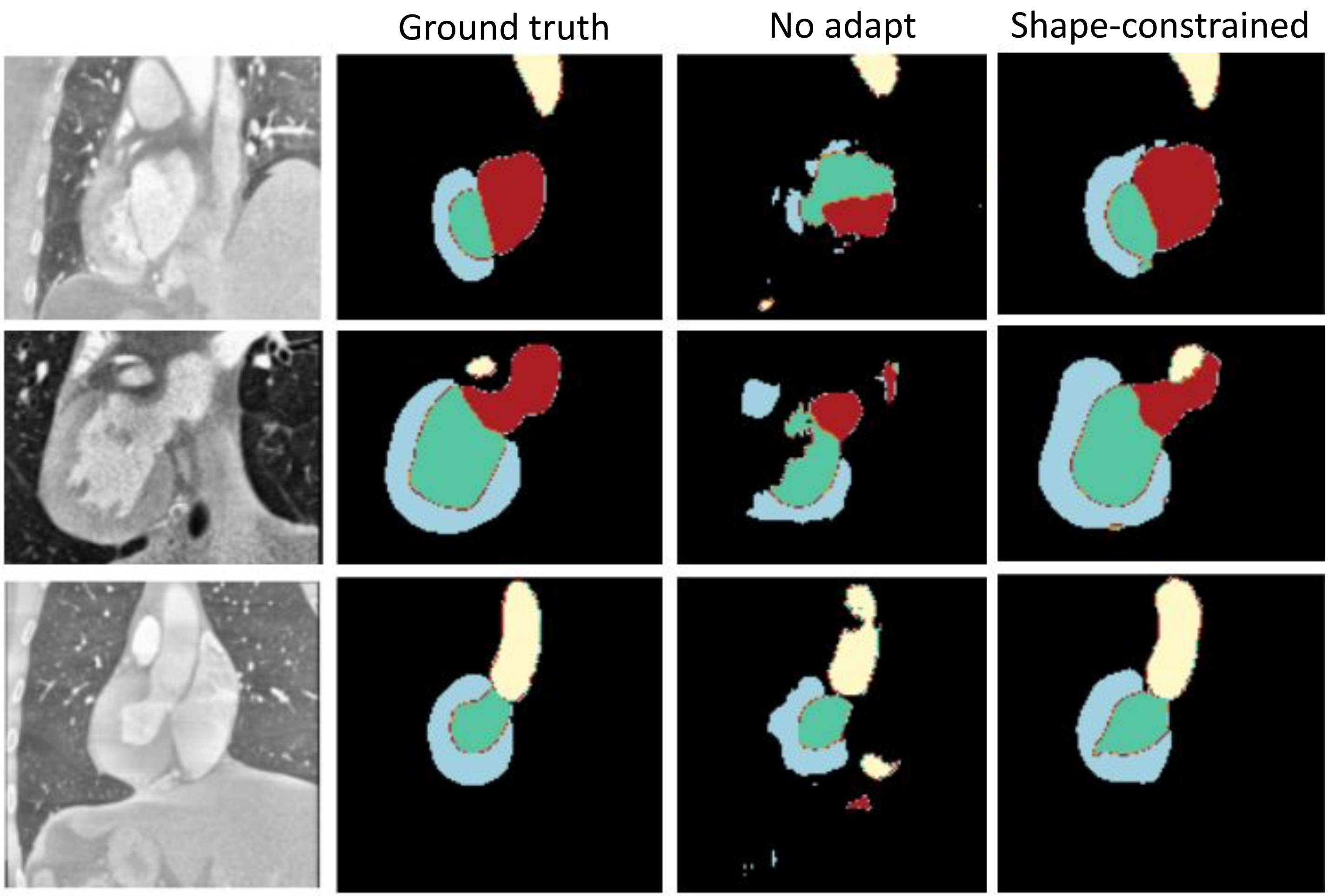
Pixel	Label
0	RV
1	BACKGROUND
2	LV
⋮	
65536	BACKGROUND

(b) Pixel-wise labels (**65k** discrete values)

Shape descriptor (in pixels)	Class: RV	Class: MYO	Class: LV
Object volume $\mathfrak{V}$	3100	800	1600
Centroid location $\mathfrak{C}$	(125, 80)	(125, 125)	
Avg. dist. to centroid $\mathfrak{D}$	(20, 15)	(15, 20)	(10, 10)
Object length $\mathfrak{L}$	750	1000	500

(c) Shape descriptors (**16** continuous values)

A few shape descriptors are surprisingly powerful in unsupervised domain adaptation [Bateson et al., MedIA 2022 (under revision)]



References & acknowledgments

I. M. Ziko, J. Dolz, E. Granger and I. Ben Ayed, Laplacian regularized few-shot learning, *International Conference on Machine Learning (ICML), 2020*
https://github.com/imtiazziko/LaplacianShot

O. Veilleux, M. Boudiaf, P. Piantanida and I. Ben Ayed, Realistic Evaluation of Transductive Few-Shot Learning, *Neural Information Processing Systems (NeurIPS), 2021*
https://github.com/oveilleux/realistic_transductive_few_shot

H. Kervadec, J. Dolz, M. Tang, E. Granger, Y. Boykov, I. Ben Ayed, Constrained-CNN losses for weakly supervised segmentation, *Medical Image Analysis (MedIA), 2019*
https://github.com/LIVIAETS/SizeLoss_WSS

M. Tang, F. Perazzi, A. Djelouah, I. Ben Ayed, C. Schroers, Y. Boykov, On regularized losses for weakly supervised CNN segmentation, *European Conference on Computer Vision (ECCV), 2018*
https://github.com/meng-tang/rloss

M. Bateson, J. Dolz, H. Kervadec, H. Lombaert, I. Ben Ayed, Souce-free domain adaptation for image segmentation, *Medical Image Analysis (MedIA), 2022 (under revision)*
https://github.com/mathilde-b/SFDA

H. Kervadec, H. Bahig, L. Letourneau-Guillon, J. Dolz, I. Ben Ayed, Pixel-wise supervision for segmentation: A few global shape descriptors might be surprisingly good!, *Medical Imaging with Deep Learning (MIDL), 2021*
https://github.com/hkervadec/shape_descriptors

Merci!