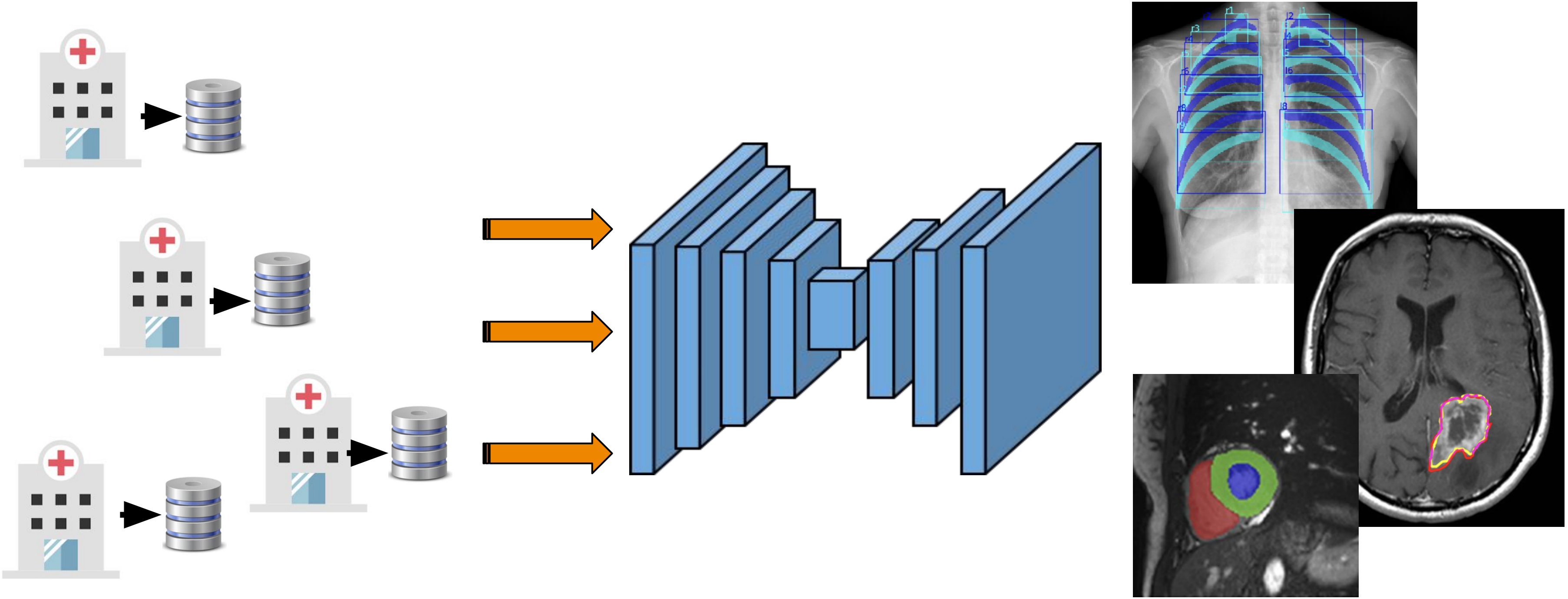
Learning with limited supervision

Jose Dolz

ÉTS, Montreal

Motivation

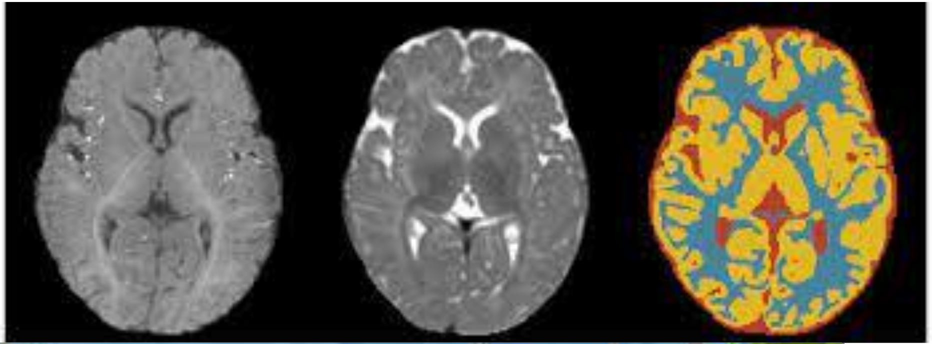


Large labelled (pixel-wise) datasets

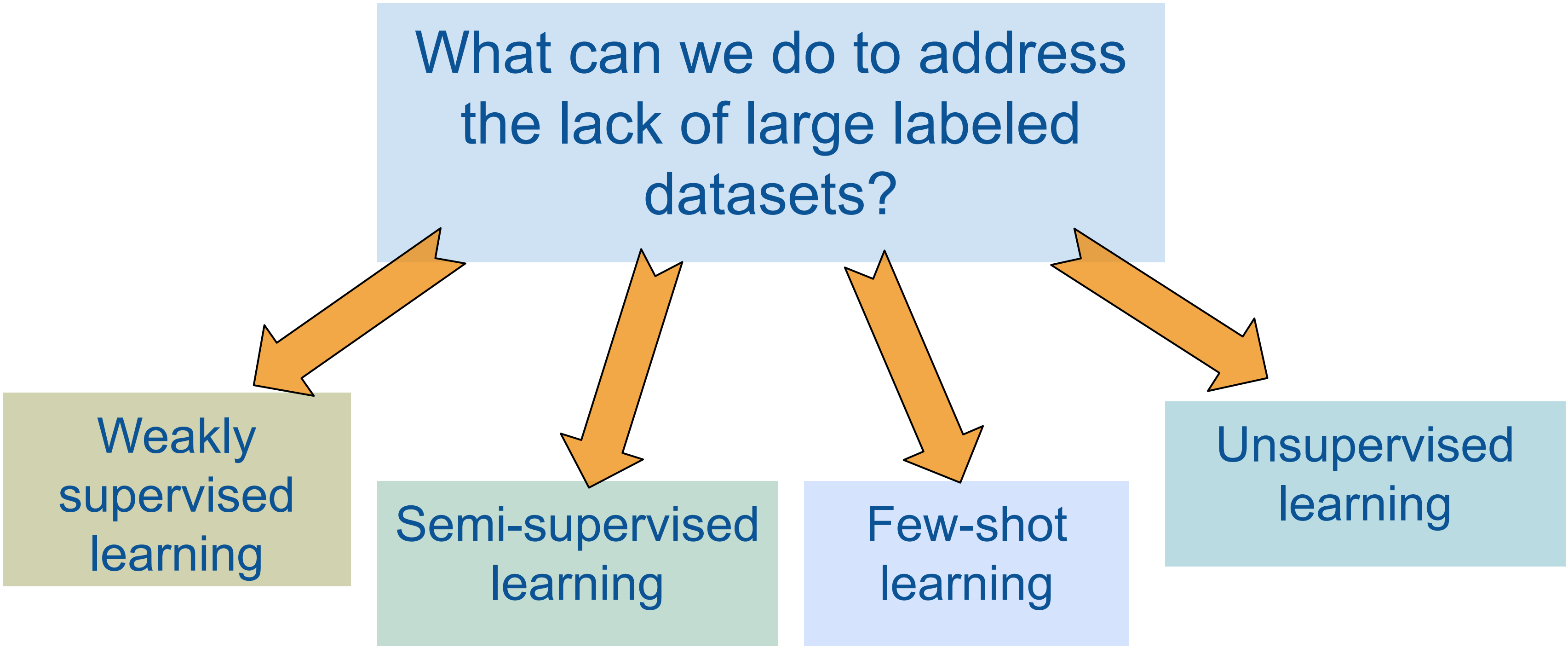
Very good performance in many tasks

Motivation

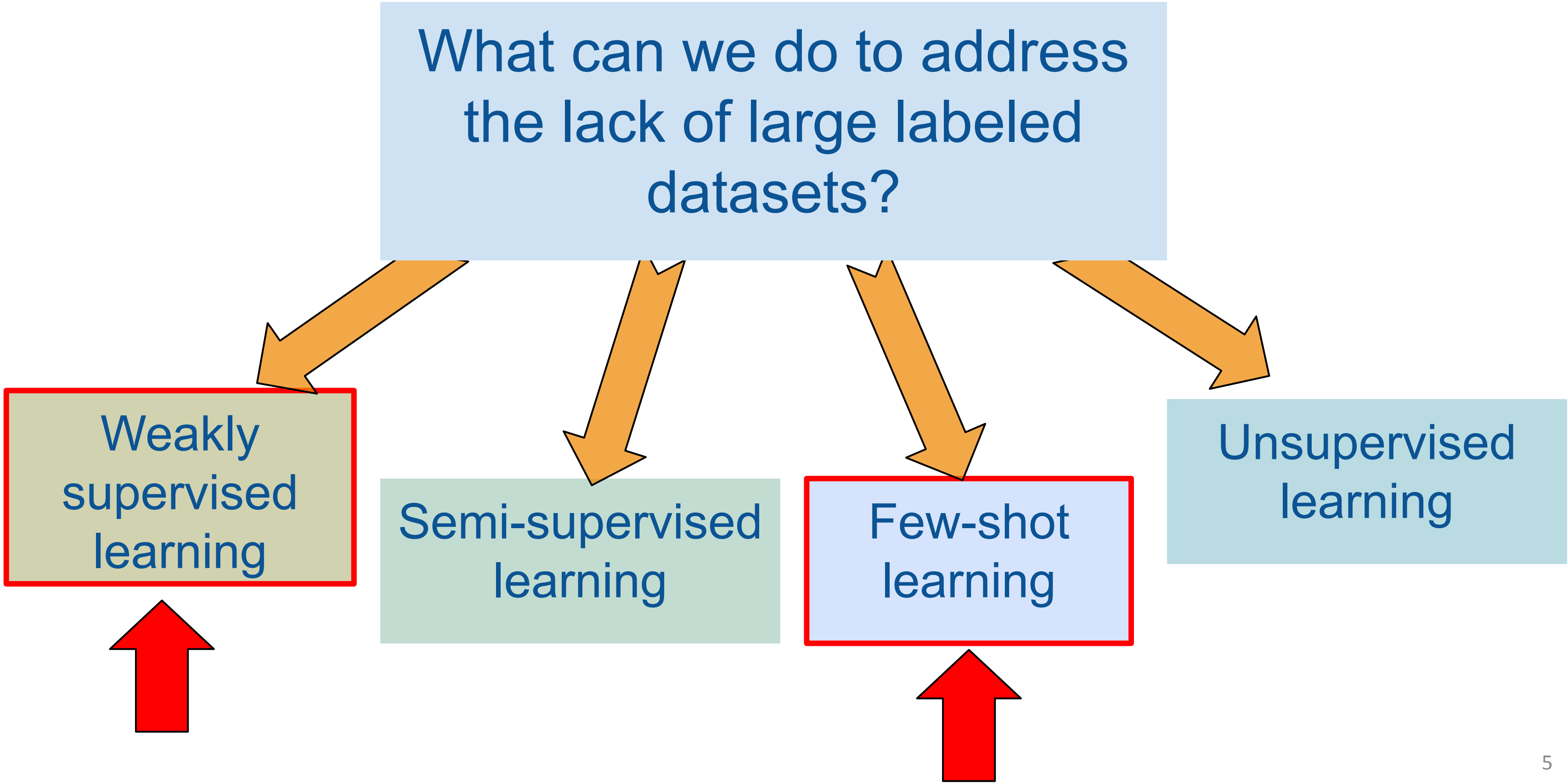
Pixel-wise annotation is a time-consuming task…



Potential solutions



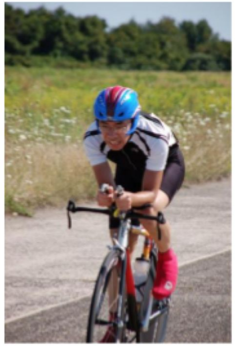
Potential solutions



Data-driven priors (cues)

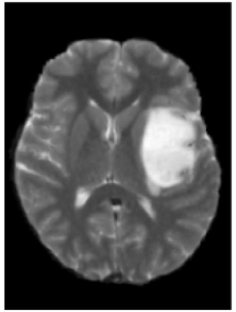
Weakly supervised learning

Image tags



Person

Bike



Tumor

Original Image

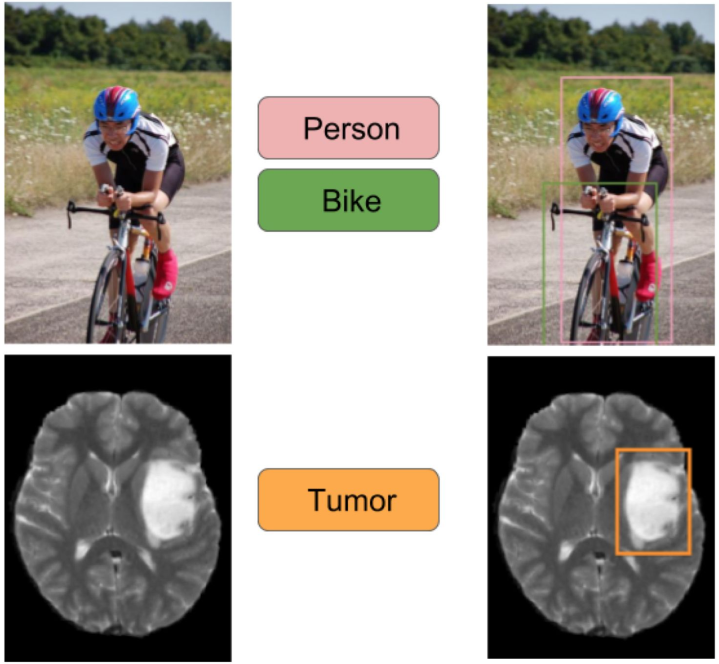
Image tags

- Pathak et al., Constrained convolutional neural networks for weakly supervised segmentation, ICCV 2015
- Kervadec et al., Constrained-CNN losses for weakly supervised segmentation, MedIA 2019.

Data-driven priors (cues)

Image tags

Bounding boxes



Original Image

Image tags

Bounding boxes

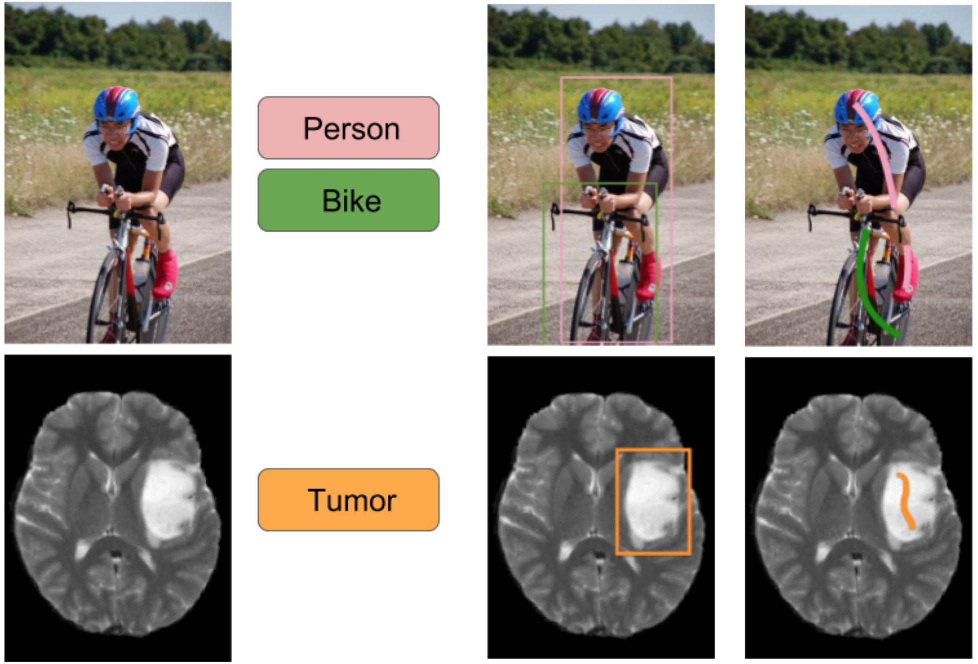
- Pathak et al., Constrained convolutional neural networks for weakly supervised segmentation, ICCV 2015
- Kervadec et al., Constrained-CNN losses for weakly supervised segmentation, MedIA 2019.

Data-driven priors (cues)

Image tags

Bounding boxes

Scribbles



Original Image | Image tags | Bounding boxes | Scribbles

- Pathak et al., Constrained convolutional neural networks for weakly supervised segmentation, ICCV 2015
- Kervadec et al., Constrained-CNN losses for weakly supervised segmentation, MedIA 2019.

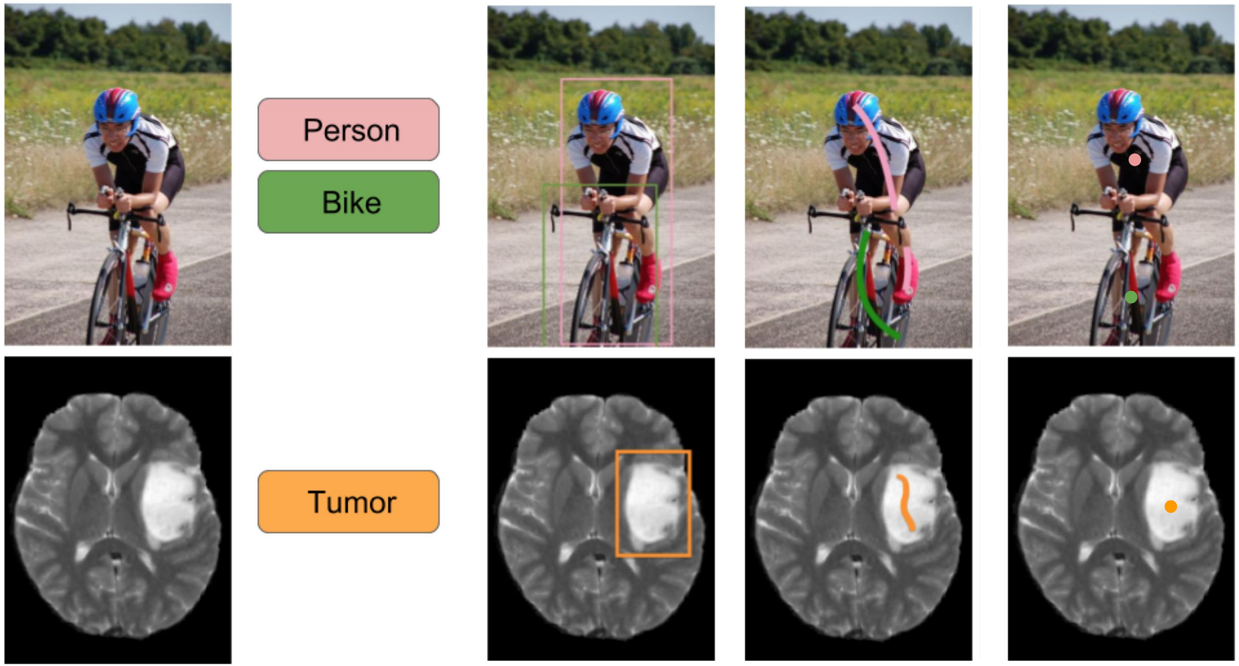
Data-driven priors (cues)

Image tags

Bounding boxes

Scribbles

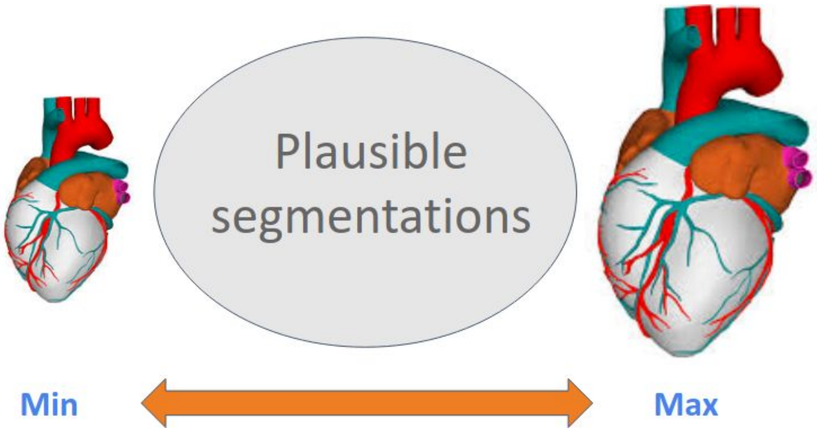
Points



Original Image | Image tags | Bounding boxes | Scribbles | Points

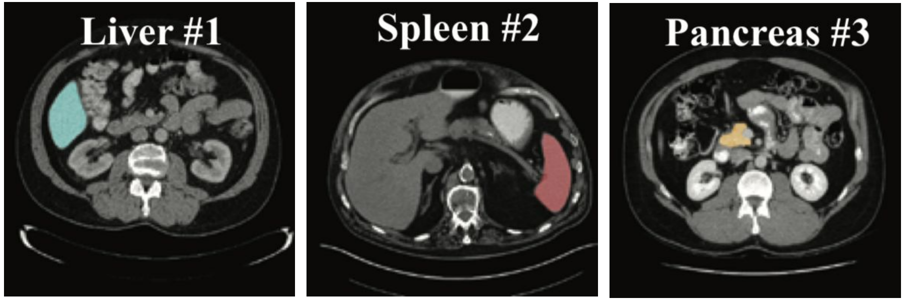
- Pathak et al., Constrained convolutional neural networks for weakly supervised segmentation, ICCV 2015
- Kervadec et al., Constrained-CNN losses for weakly supervised segmentation, MedIA 2019.

Knowledge-driven priors



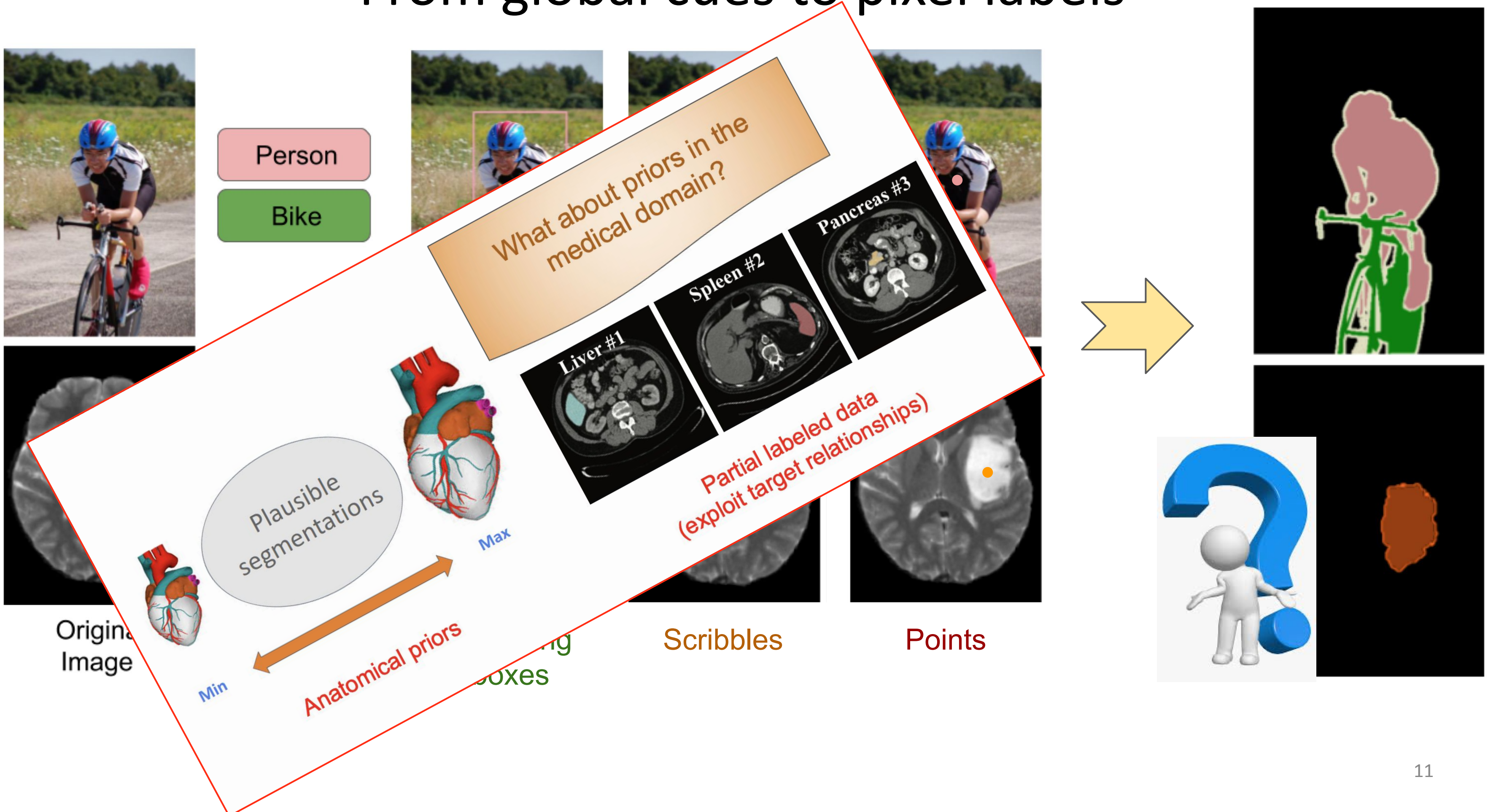
Anatomical priors

What about priors in the medical domain?



Partial labeled data
(exploit target relationships)

From global cues to pixel labels



Constrained optimization (CNN)

Optimize (A) such that (B)

Task Set of constraints

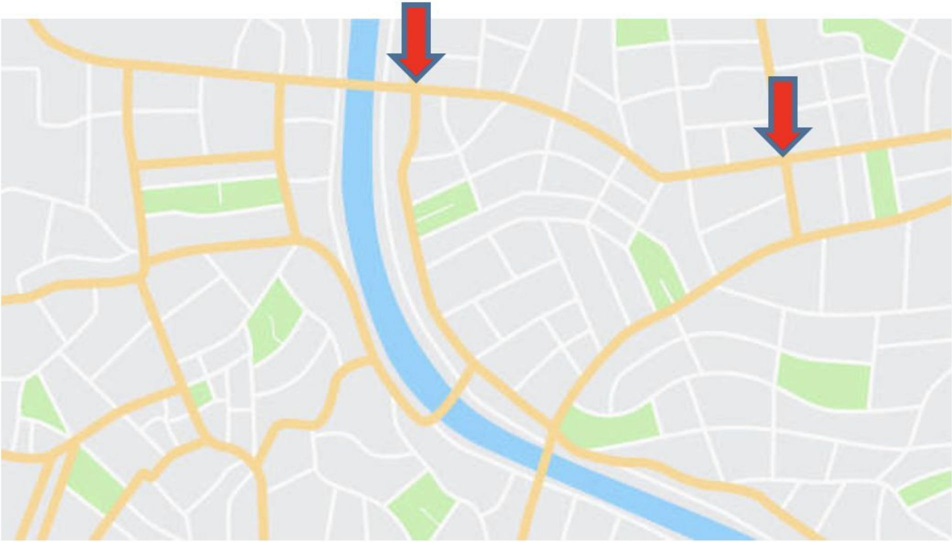
Constrained optimization (CNN)

Optimize (A)

Task

such that (B)

Set of constraints



How we can go from point A to B?

Constrained optimization (CNN)

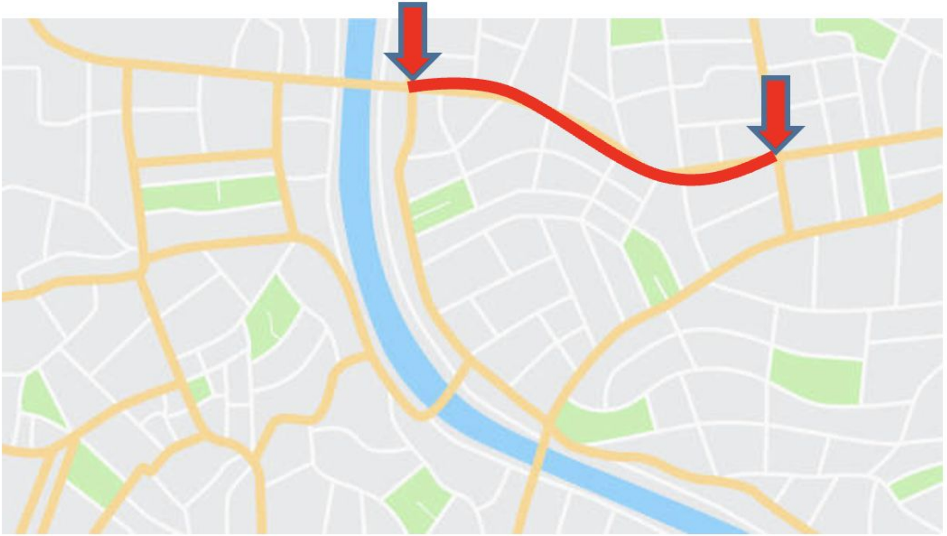
Optimize (A)

Task

such that (B)

Set of constraints

How we can go from point A to B?



Constrained optimization (CNN)

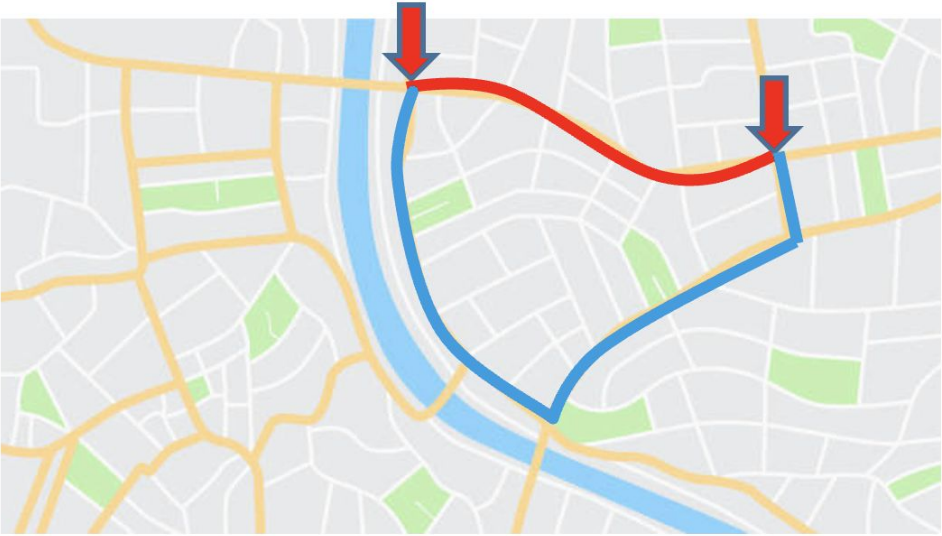
Optimize (A)

Task

such that (B)

Set of constraints

How we can go from point A to B?



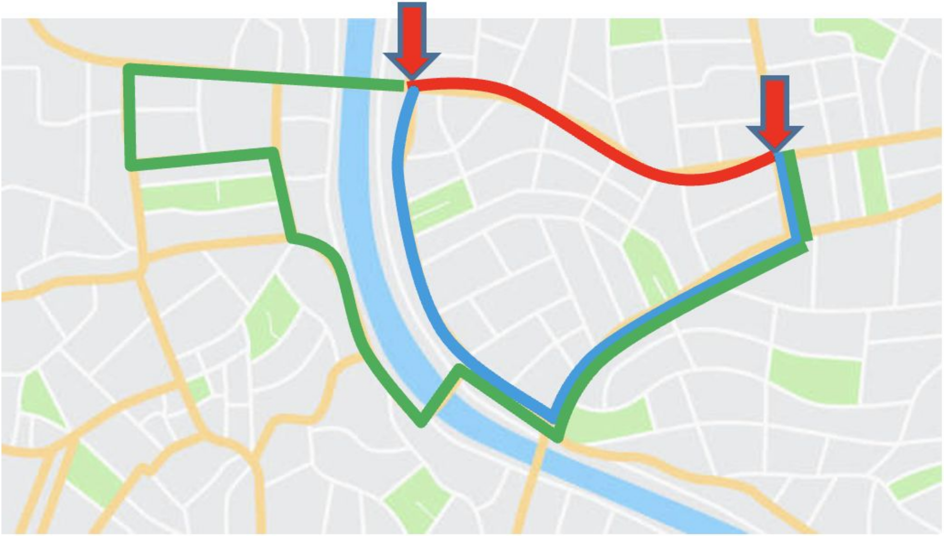
Constrained optimization (CNN)

Optimize (A)

Task

such that (B)

Set of constraints



How we can go from point A to B?

Which is the best route?

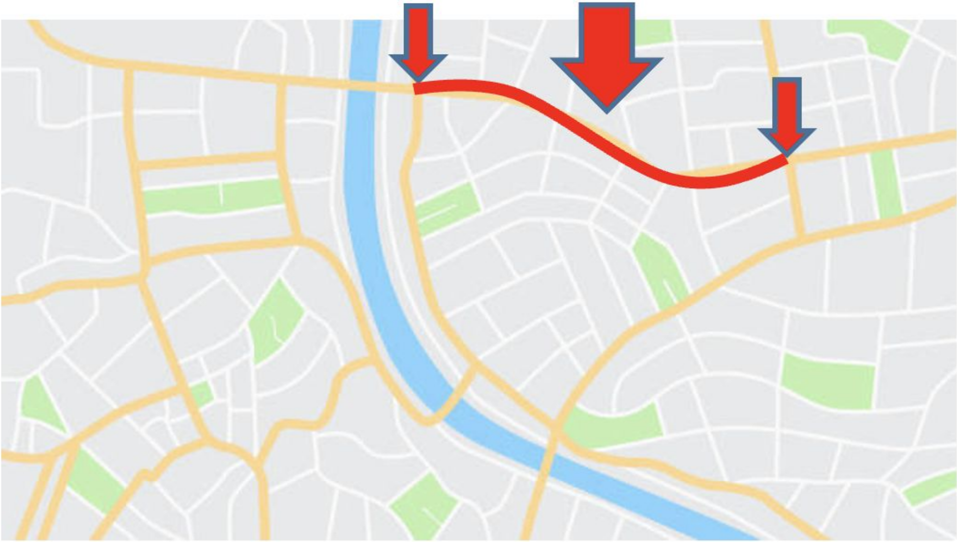
Constrained optimization (CNN)

Optimize (A)

Task

such that (B)

Set of constraints



How we can go from point A to B?

Which is the best route?

Constraint: shortest

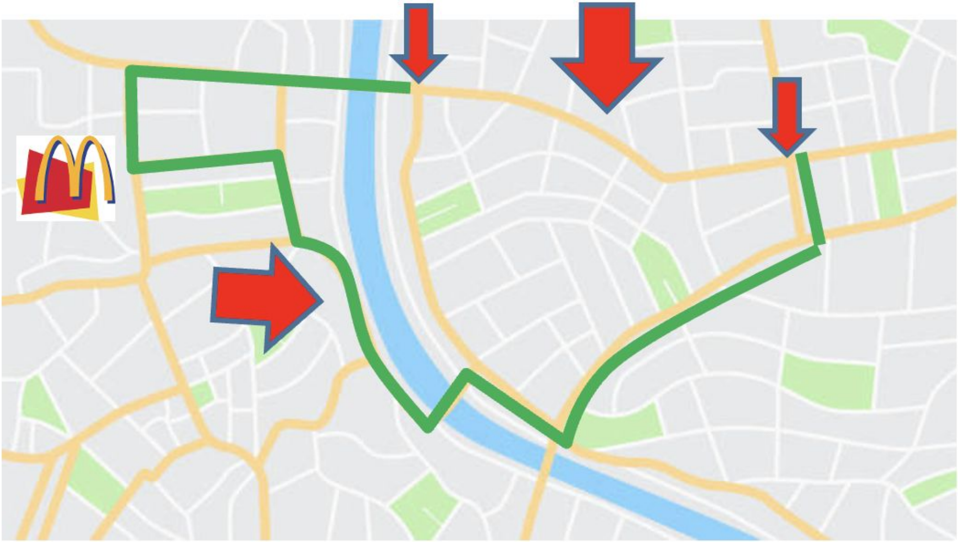
Constrained optimization (CNN)

Optimize (A)

Task

such that (B)

Set of constraints



How we can go from point A to B?

Which is the best route?

Constraint: shortest but going through a McDonnalds

Constrained optimization (CNN)

Optimize (A) such that (B)

$$\min_{\theta} \mathcal{H}(S, Y) \quad s.t. \quad \sum_{n=0}^{N} s_n = A$$

Let's assume we know the target size (A)

Constrained problem

Constrained optimization (CNN)

Let's assume we know the target size (A)

Optimize (A) such that (B)

$$\min_{\theta} \mathcal{H}(S, Y) \quad s.t. \quad \sum_{n=0}^{N} s_n = A$$

Constrained problem

$$\min_{\theta} \mathcal{H}(S, Y) + \lambda(\sum_{n=0}^{N} s_n - A)$$

Unconstrained problem (penalties to the rescue!)

Equality constraints

A=B

General definition

$$\min_{\theta} \mathcal{H}(S) \quad s.t. \quad g(\mathbf{s}) = C$$

Constraint

Equality constraints

A=B

General definition

$$\min_{\theta} \mathcal{H}(S) \quad s.t. \quad g(\mathbf{s}) = C$$

Constraint

$$\Downarrow$$

$$\min_{\theta} \mathcal{H}(S) + \lambda(g(\mathbf{s}) - C)$$

Penalty

Equality constraints

A=B

General definition

$$\min_{\theta} \mathcal{H}(S) \quad s.t. \quad g(\mathbf{s}) = C$$

Constraint

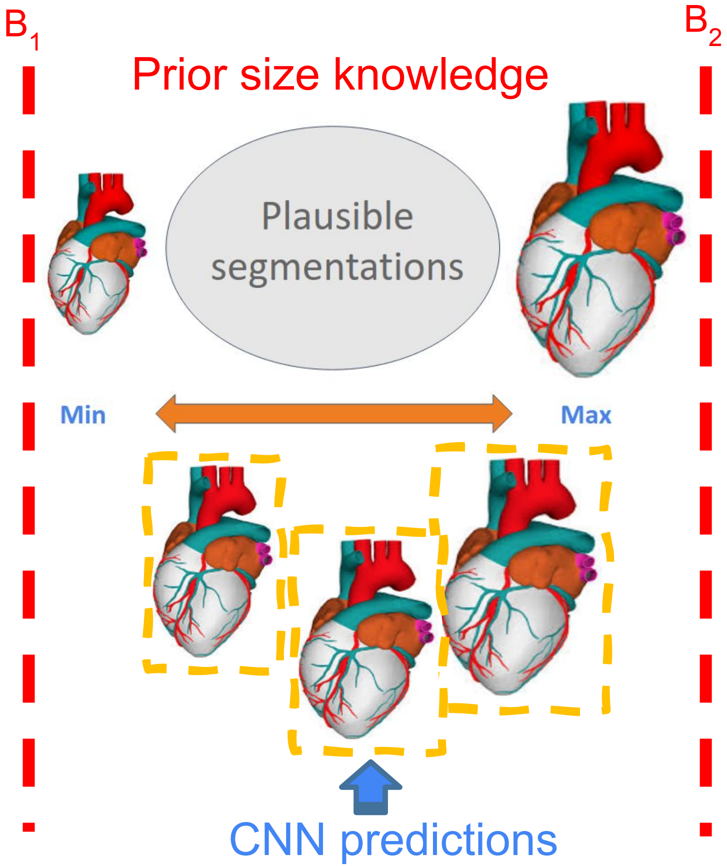
$$\min_{\theta} \mathcal{H}(S) + \lambda(g(\mathbf{s}) - C)$$

Penalty

This can be modeled with linear/quadratic penalties, KL divergence, etc

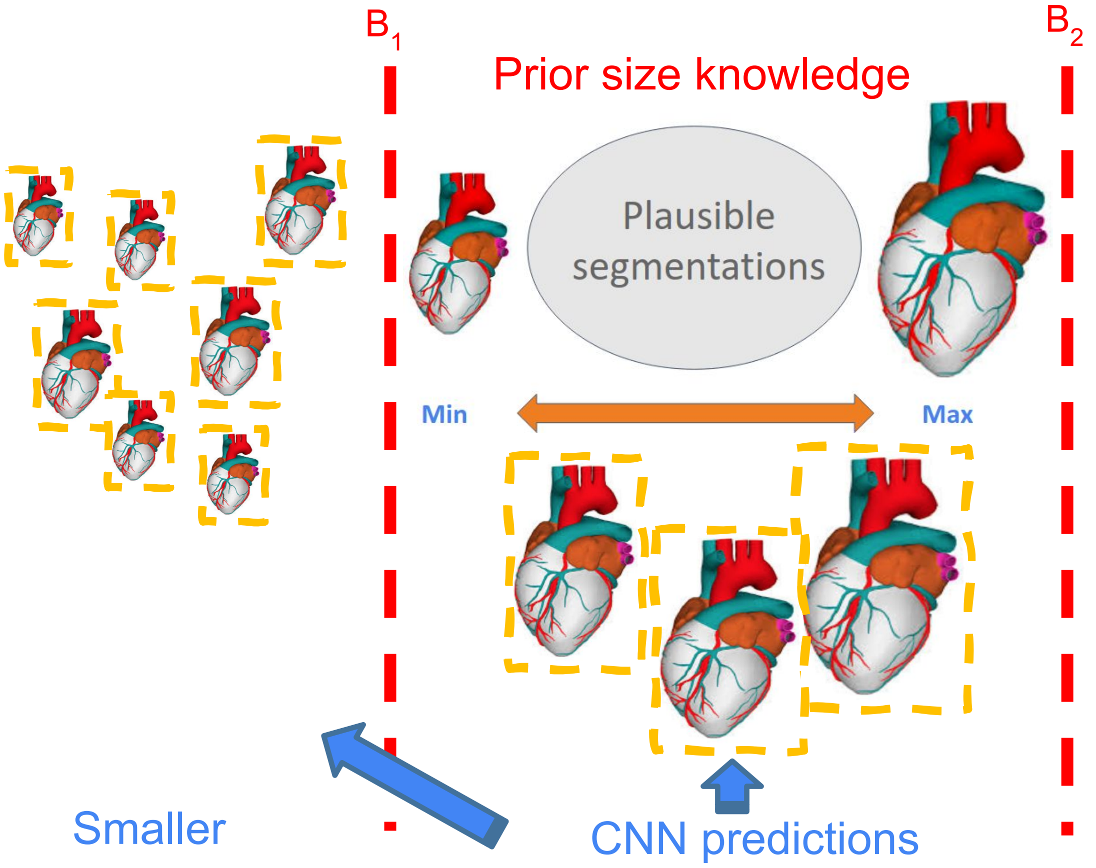
Inequality constraints

A<B,
A>B,
B1 < A < B2

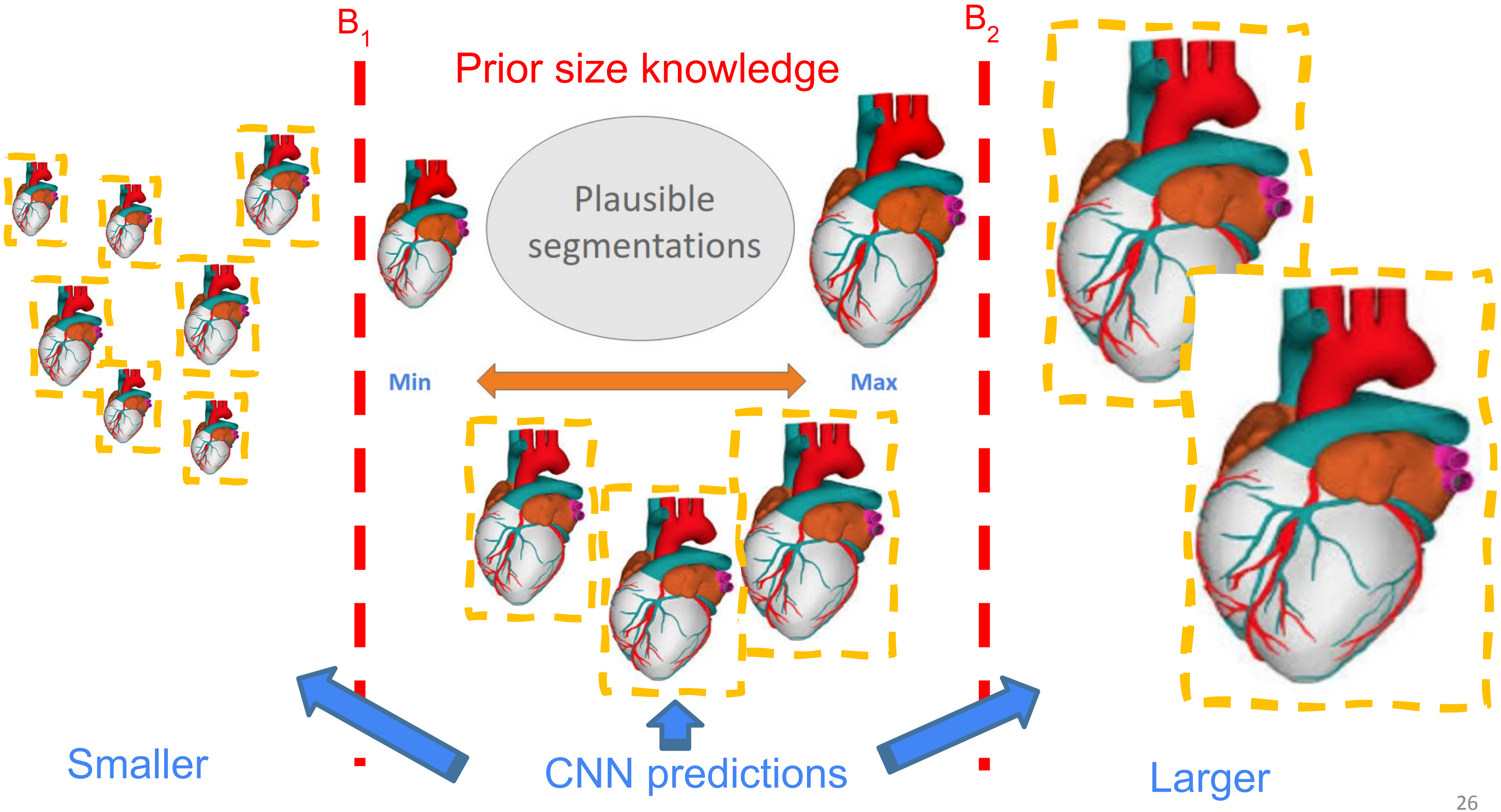


Inequality constraints

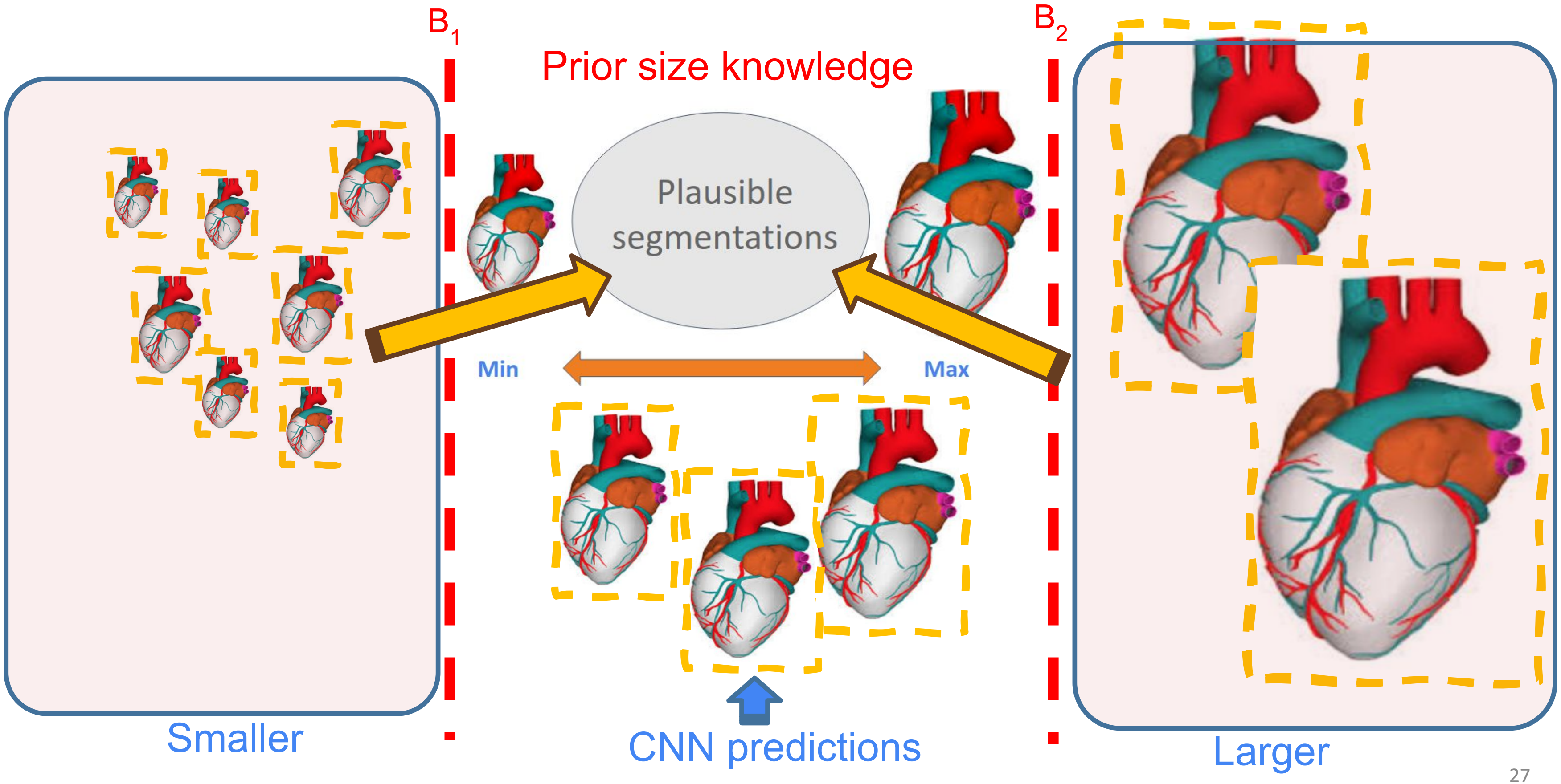
$A<B$,
$A>B$,
$B1 < A < B2$



Inequality constraints



Inequality constraints



Inequality constraints

Size loss
[Kervadec et al, MedIA'19]

No cavity Cavity

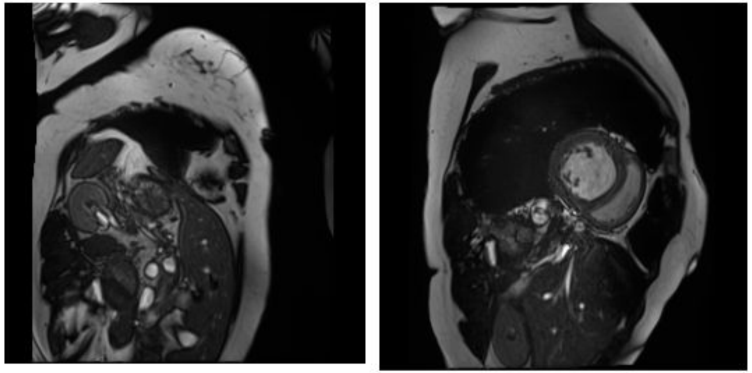


Image-tag information

$$\sum_{p \in \Omega} s_\theta^{p,c} \leq 0$$

For negative image tags

Inequality constraints

Size loss
[Kervadec et al, MedIA'19]

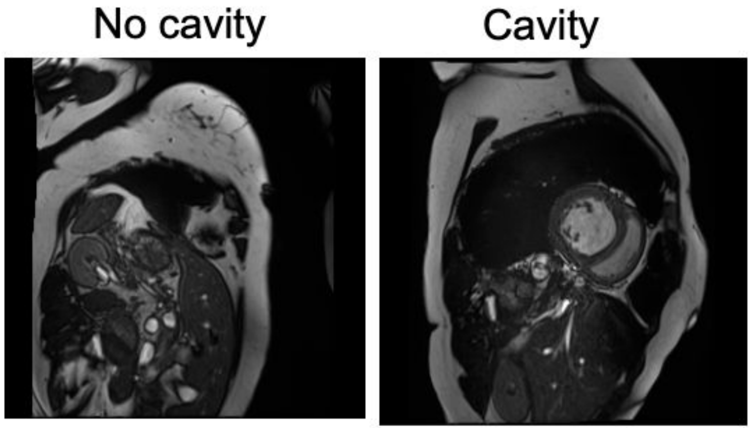
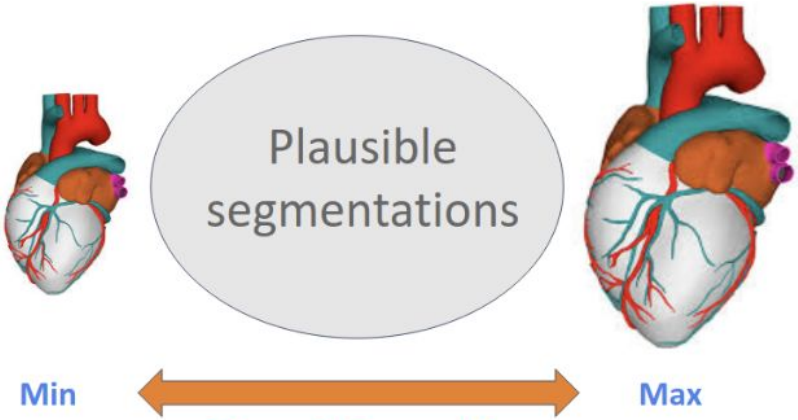


Image-tag information

$$\sum_{p \in \Omega} s_\theta^{p,c} \leq 0$$

For negative image tags



Size information

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

For positive image tags

Inequality constraints

Formal definition

Size loss
[Kervadec et al, MedIA'19]

$$\min_{\boldsymbol{\theta}} \; \mathcal{H}(S) \quad \text{s.t} \quad a \leqslant \sum_{p \in \Omega} S_p \leqslant b$$

Inequality constraint

[Kervadec et al., Constrained-CNN losses for weakly supervised segmentation. MedIA 2019]

Inequality constraints

Formal definition

Size loss
[Kervadec et al, MedIA'19]

$$\min_{\boldsymbol{\theta}} \mathcal{H}(S) \quad \text{s.t} \quad a \leqslant \sum_{p\in\Omega} S_p \leqslant b$$

$$\mathcal{H}(S) + \lambda \mathcal{C}\left(V_S\right)$$

$$\boxed{V_S = \sum_{p\in\Omega} \mathbf{s}_\theta^{p,c}}$$

[Kervadec et al., Constrained-CNN losses for weakly supervised segmentation. MedIA 2019]

Inequality constraints

Size loss
[Kervadec et al, MedIA'19]

Formal definition

$$\min_{\boldsymbol{\theta}} \mathcal{H}(S) \quad \text{s.t} \quad a \leqslant \sum_{p \in \Omega} S_p \leqslant b$$

$$\mathcal{H}(S) + \lambda \mathcal{C}(V_S)$$

$$\boxed{V_S = \sum_{p \in \Omega} \mathbf{s}_\theta^{p,c}}$$

$$\mathcal{H}(S) = -\sum_{p \in \mathcal{L}} log(s_\theta^p)$$

CE on the labeled pixels (if any)

[Kervadec et al., Constrained-CNN losses for weakly supervised segmentation. MedIA 2019]

Inequality constraints

Formal definition

Size loss
[Kervadec et al, MedIA'19]

$$\min_{\boldsymbol{\theta}} \ \mathcal{H}(S) \quad \text{s.t} \quad a \leqslant \sum_{p \in \Omega} S_p \leqslant b$$

$$\mathcal{H}(S) + \lambda \mathcal{C}(V_S)$$

$$V_S = \sum_{p \in \Omega} \mathbf{s}_\theta^{p,c}$$

$$\mathcal{H}(S) = -\sum_{p \in \mathcal{L}} log(s_\theta^p)$$

CE on the labeled pixels (if any)

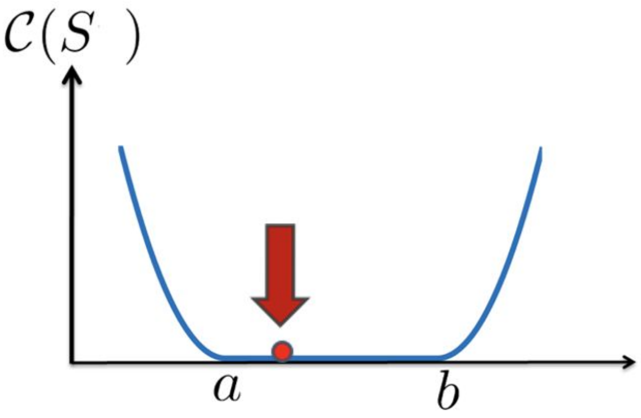
$$\mathcal{C}(V_S) = \begin{cases} (V_S - a)^2, & \text{if } V_S < a \\ (V_S - b)^2, & \text{if } V_S > b \\ 0, & \text{otherwise} \end{cases}$$

[Kervadec et al., Constrained-CNN losses for weakly supervised segmentation. MedIA 2019]

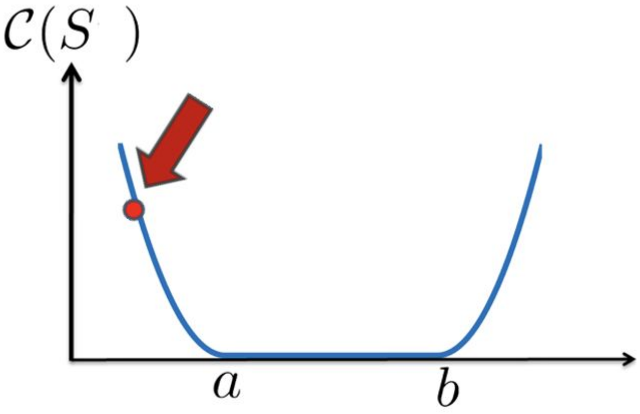
Inequality constraints

Visual intuition

Size loss
[Kervadec et al, MedIA'19]



Constraint A satisfied



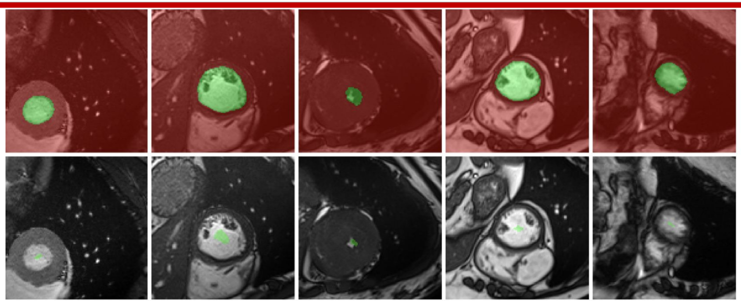
Constraint B violated

[Kervadec et al., Constrained-CNN losses for weakly supervised segmentation. MedIA 2019]

Inequality constraints

Size loss
[Kervadec et al, MedIA'19]

	Model	Method	DSC (Val)
Weakly supervised	Partial CE		0.1497
	CE + Tags	Lagrangian Proposals (Pathak et al., 2015a)	0.7707
	Partial CE + Tags	Direct loss (Ours)	0.7924
	CE + Tags + Size*	Lagrangian Proposals (Pathak et al., 2015a)	0.7854
	Partial CE + Tags + Size*	Direct loss (Ours)	0.8004
	CE + Tags + Size**	Lagrangian Proposals (Pathak et al., 2015a)	0.7900
	Partial CE + Tags + Size**	Direct loss (Ours)	**0.8708**
	CE + 3D Size**	Lagrangian Proposals (Pathak et al., 2015a)	N/A
	Partial CE + 3D Size**	Direct loss (Ours)	0.8580
Fully supervised	Cross-entropy		0.8872

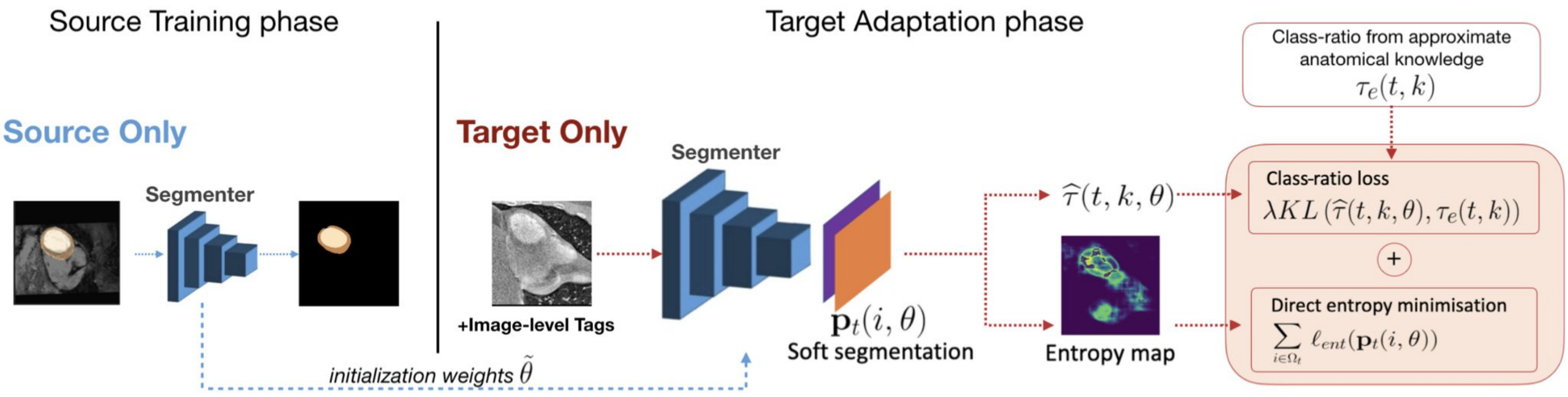


[Kervadec et al., Constrained-CNN losses for weakly supervised segmentation. MedIA 2019]

Inequality constraints

Source-free Domain Adaptation
[Bateson et al., MICCAI'20]

KL divergence

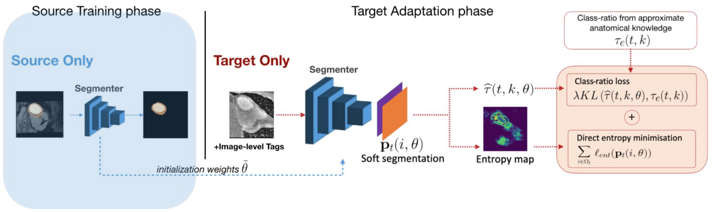


Bateson et al., Source-Relaxed Domain Adaptation for Image Segmentation. MICCAI'20

Inequality constraints

KL divergence

Source-free: no access to source data when adapting



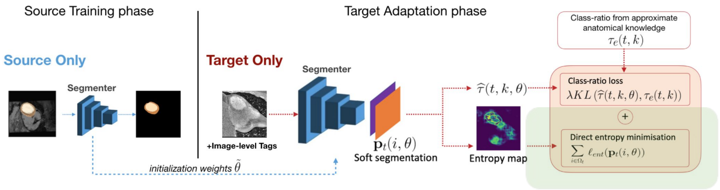
1-Train the network on the source domain

$$\min_{\theta}\sum_{p\in\mathcal{L}} l(\mathbf{y}^p,\mathbf{s}_\theta^p)$$

Set of labeled SOURCE pixels

Bateson et al., Source-Relaxed Domain Adaptation for Image Segmentation. MICCAI'20

Inequality constraints



KL divergence

Source-free: no access to source data when adapting

2-Adapt the model without accessing the source data

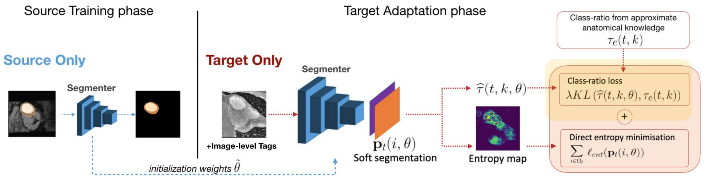
$$\mathcal{L}_{\mathcal{H}} = -\sum_{p\in\mathcal{T}}\sum_{k} \mathbf{s}_\theta^{p,k}\log \mathbf{s}_\theta^{p,k} + \mathcal{D}_{KL}(\hat{\tau},\tau_e)$$

Minimize entropy on predicted TARGET pixels

Bateson et al., Source-Relaxed Domain Adaptation for Image Segmentation. MICCAI'20

Inequality constraints

KL divergence



Source-free: no access to source data when adapting

2-Adapt the model without accessing the source data

$$\mathcal{L}_{\mathcal{H}} = -\sum_{p \in \mathcal{T}} \sum_{k} \mathbf{s}_\theta^{p,k} \log \mathbf{s}_\theta^{p,k} + \mathcal{D}_{KL}(\hat{\tau}, \tau_e)$$

$\tau_e(t,k)$ Estimated size by an auxiliary network trained on the source

$$\hat{\tau}(t,k,\theta) = \frac{1}{|\Omega_t|} \sum_{i \in \Omega_t} \mathbf{s}_\theta^{i,k}$$

Minimize entropy on predicted TARGET pixels

Size regularizer

Computed size from the segmentation of the target image

Bateson et al., Source-Relaxed Domain Adaptation for Image Segmentation. MICCAI'20

Inequality constraints

L2 Penalty

But we can do more than simply the size

Shape descriptors
[Kervadec et al, MIDL'21]

Shape moment $$\mu_{p,q}^{(k)}(s_{\boldsymbol{\theta}}) := \sum_{i \in \Omega} s_{\boldsymbol{\theta}}^{(i,k)} x_{(i)}^p y_{(i)}^q,$$

Central moment $$\bar{\mu}_{p,q}^{(k)} := \sum_{i \in \Omega} s_{\boldsymbol{\theta}}^{(i,k)} \left(x_{(i)} - \frac{\mu_{1,0}^{(k)}}{\mu_{0,0}^{(k)}} \right)^p \left(y_{(i)} - \frac{\mu_{0,1}^{(k)}}{\mu_{0,0}^{(k)}} \right)^q.$$

[Kervadec et al., Beyond pixel-wise supervision for segmentation: A few global shape descriptors might be surprisingly good!. MIDL'21]

Inequality constraints

L2 Penalty

But we can do more than simply the size

From shape and central moment

Shape descriptors
[Kervadec et al, MIDL'21]

Volume

$$\mathfrak{V}^{(k)}(s_{\boldsymbol{\theta}}) := \mu_{0,0}^{(k)}(s_{\boldsymbol{\theta}}).$$

Centroid

$$\mathfrak{C}^{(k)}(s_{\boldsymbol{\theta}}) := \left(\frac{\mu_{1,0}^{(k)}(s_{\boldsymbol{\theta}})}{\mu_{0,0}^{(k)}(s_{\boldsymbol{\theta}})}, \frac{\mu_{0,1}^{(k)}(s_{\boldsymbol{\theta}})}{\mu_{0,0}^{(k)}(s_{\boldsymbol{\theta}})}\right).$$

Length

Laplacian

$$\mathfrak{L}^{(k)}(s_{\boldsymbol{\theta}}) := \sum_{i,j \in \mathcal{G}_\Omega} |s_{\boldsymbol{\theta}}^{(i,k)} - s_{\boldsymbol{\theta}}^{(j,k)}| L_{\Omega,i,j}.$$

[Kervadec et al., Beyond pixel-wise supervision for segmentation: A few global shape descriptors might be surprisingly good!. MIDL'21]

Few-shot learning

Setting

Training on **base** classes

Few-shot learning

Setting

Training on **base** classes

Few-shot tasks at testing time

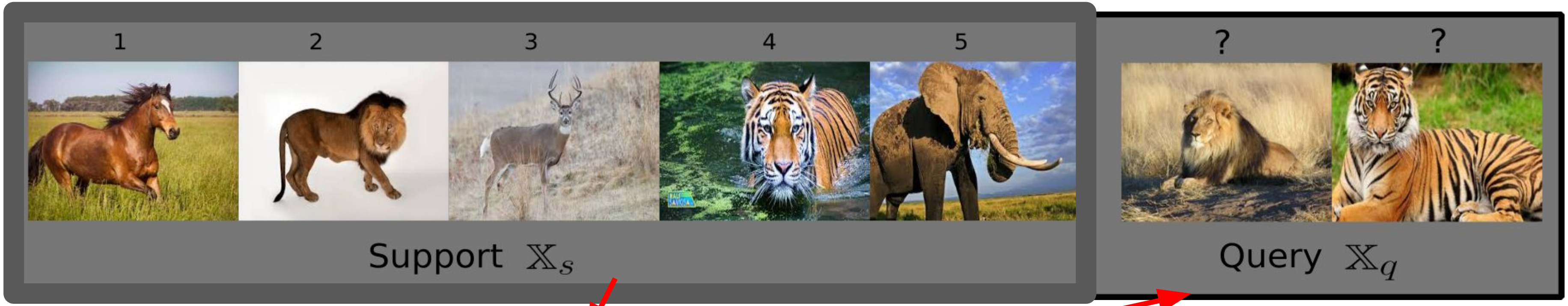
Learn from a few examples per **new** class

Few-shot learning

Setting

Training on **base** classes

Few-shot tasks at testing time

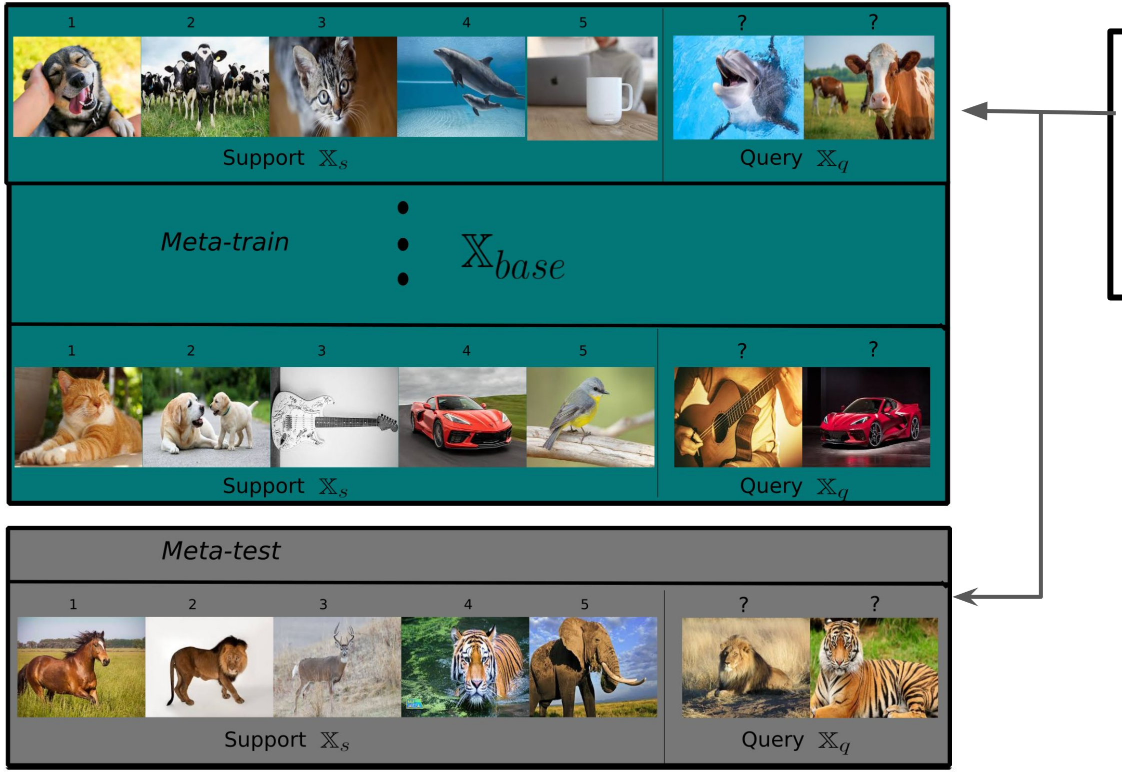


Learn from a <u>few examples</u> per **new** class

Classify these

Few-shot learning

Literature

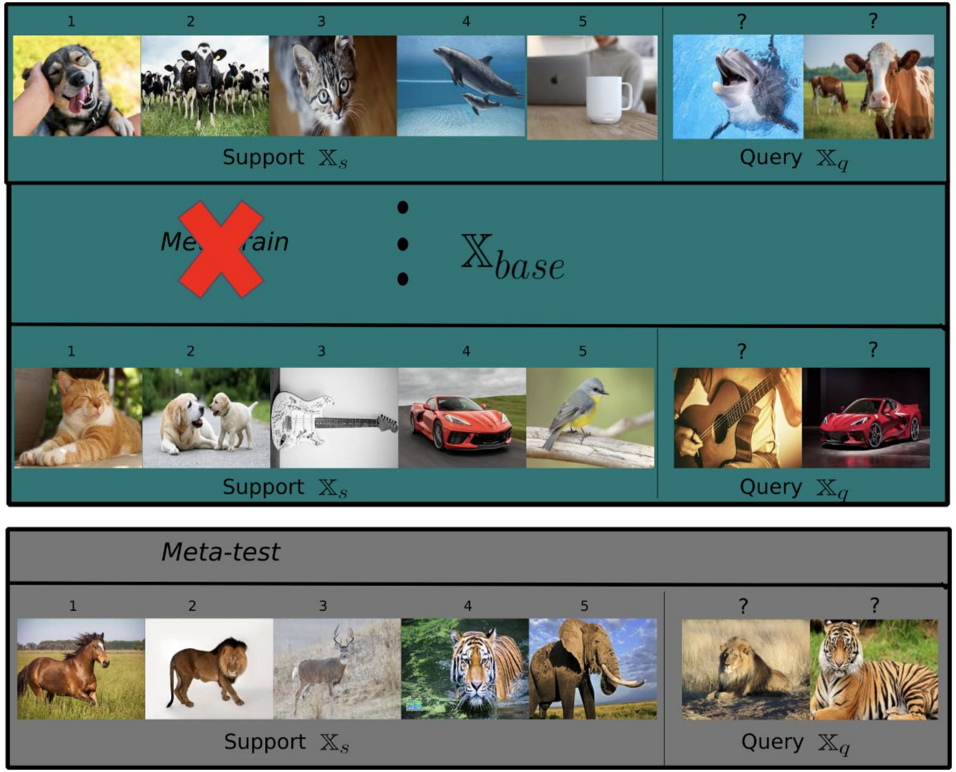


Create artificial episodes for ***episodic training***
(Learn initial model)

Vinyal et al, (Neurips '16),
Snell et al, (Neurips '17),
Sung et al, (CVPR ' 18),
Finn et al, (ICML' 17),
Ravi et al, (ICLR' 17),
Lee et al, (CVPR' 19),
Hu et al, (ICLR '20),
Ye et al, (CVPR '20), . . .

Few-shot learning

Literature



No need to ***meta-train***

[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],
[Veilleux et al., NeurIPS'21],...

Few-shot learning

Classification

TIM
[Boudiaf et al, NeurIPS'20]

1) Standard training on the base classes to get an initial model

Few-shot learning

Classification

TIM
[Boudiaf et al, NeurIPS'20]

1) Standard training on the base classes to get an initial model

2) During adaptation:

$$\min_{\mathbf{W}} \; \lambda \cdot \mathrm{CE} - \widehat{\mathcal{I}}_{\alpha}(X_{\mathcal{Q}}; Y_{\mathcal{Q}})$$

Few-shot learning

Classification

TIM
[Boudiaf et al, NeurIPS'20]

1) Standard training on the base classes to get an initial model

2) During adaptation:

$$\min_{\mathbf{W}} \lambda \cdot \mathrm{CE} - \widehat{\mathcal{I}}_\alpha(X_\mathcal{Q}; Y_\mathcal{Q})$$

CE on supervised images (i.e., support)

Few-shot learning

Classification

TIM
[Boudiaf et al, NeurIPS'20]

1) Standard training on the base classes to get an initial model

2) During adaptation:

$$\min_{\mathbf{W}} \ \lambda \cdot \mathrm{CE} - \widehat{\mathcal{I}}_\alpha(X_\mathcal{Q}; Y_\mathcal{Q})$$

On test (query) samples

$$\underbrace{-\sum_{k=1}^{K} \widehat{p}_k \log \widehat{p}_k}_{\widehat{\mathcal{H}}(Y_\mathcal{Q}):\ \text{marginal entropy}} + \alpha \underbrace{\frac{1}{|\mathcal{Q}|} \sum_{i \in \mathcal{Q}} \sum_{k=1}^{K} p_{ik} \log(p_{ik})}_{-\widehat{\mathcal{H}}(Y_\mathcal{Q}|X_\mathcal{Q}):\ \text{conditional entropy}}$$

Few-shot learning

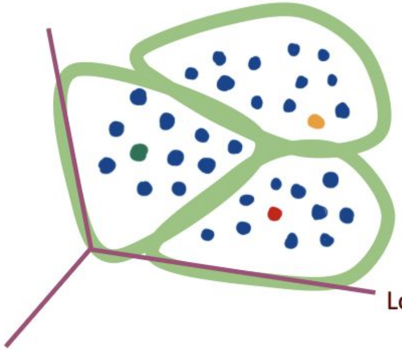
Classification

TIM
[Boudiaf et al, NeurIPS'20]

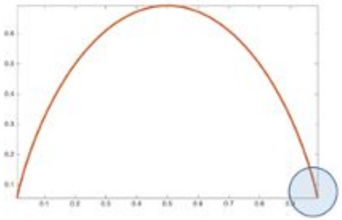
1) Standard training on the base classes to get an initial model

2) During adaptation:

$$\min_{\mathbf{W}} \; \lambda \cdot \mathrm{CE} - \widehat{\mathcal{I}}_{\alpha}(X_{\mathcal{Q}}; Y_{\mathcal{Q}})$$



Low-entropy classifier



$$\underbrace{-\sum_{k=1}^{K} \widehat{p}_k \log \widehat{p}_k}_{\widehat{\mathcal{H}}(Y_{\mathcal{Q}}):\ \text{marginal entropy}} \quad + \alpha \underbrace{\frac{1}{|\mathcal{Q}|} \sum_{i \in \mathcal{Q}} \sum_{k=1}^{K} p_{ik} \log(p_{ik})}_{-\widehat{\mathcal{H}}(Y_{\mathcal{Q}}|X_{\mathcal{Q}}):\ \text{conditional entropy}}$$

Problem if only the entropy is minimized!

Few-shot learning

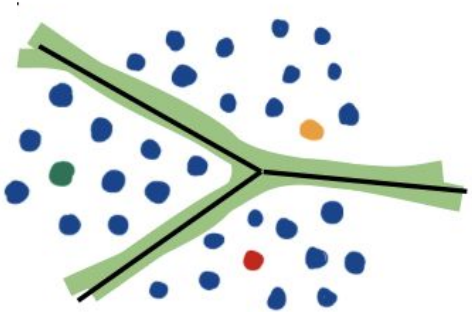
Classification

TIM
[Boudiaf et al, NeurIPS’20]

1) Standard training on the base classes to get an initial model

2) During adaptation:

$$\min_{\mathbf{W}} \ \lambda \cdot \mathrm{CE} - \widehat{\mathcal{I}}_\alpha(X_\mathcal{Q}; Y_\mathcal{Q})$$



$$\underbrace{-\sum_{k=1}^{K} \widehat{p}_k \log \widehat{p}_k}_{\widehat{\mathcal{H}}(Y_\mathcal{Q}):\ \text{marginal entropy}} + \alpha \underbrace{\frac{1}{|\mathcal{Q}|} \sum_{i \in \mathcal{Q}} \sum_{k=1}^{K} p_{ik} \log(p_{ik})}_{-\widehat{\mathcal{H}}(Y_\mathcal{Q}|X_\mathcal{Q}):\ \text{conditional entropy}}$$

Few-shot learning

Classification

TIM
[Boudiaf et al, NeurIPS'20]

Method	Transd.	Backbone	*mini*-ImageNet 1-shot	*mini*-ImageNet 5-shot	*tiered*-ImageNet 1-shot	*tiered*-ImageNet 5-shot	CUB 1-shot	CUB 5-shot
MAML [9]		ResNet-18	49.6	65.7	-	-	68.4	83.5
RelatNet [40]		ResNet-18	52.5	69.8	-	-	68.6	84.0
MatchNet [45]		ResNet-18	52.9	68.9	-	-	73.5	84.5
ProtoNet [38]		ResNet-18	54.2	73.4	-	-	73.0	86.6
MTL [39]	✗	ResNet-12	61.2	75.5	-	-	-	-
vFSL [50]		ResNet-12	61.2	77.7	-	-	-	-
Neg-cosine [26]		ResNet-18	62.3	80.9	-	-	72.7	89.4
MetaOpt [22]		ResNet-12	62.6	78.6	66.0	81.6	-	-
SimpleShot [46]		ResNet-18	62.9	80.0	68.9	84.6	68.9	84.0
Distill [41]		ResNet-12	64.8	82.1	71.5	86.0	-	-
RelatNet + T [14]		ResNet-12	52.4	65.4	-	-	-	-
ProtoNet + T [14]		ResNet-12	55.2	71.1	-	-	-	-
MatchNet+T [14]		ResNet-12	56.3	69.8	-	-	-	-
TPN [28]		ResNet-12	59.5	75.7	-	-	-	-
TEAM [34]		ResNet-18	60.1	75.9	-	-	-	-
Ent-min [7]	✓	ResNet-12	62.4	74.5	68.4	83.4	-	-
CAN+T [14]		ResNet-12	67.2	80.6	73.2	84.9	-	-
LaplacianShot [51]		ResNet-18	72.1	82.3	79.0	86.4	81.0	88.7
TIM-ADM		ResNet-18	73.6	**85.0**	**80.0**	**88.5**	81.9	90.7
TIM-GD		ResNet-18	**73.9**	**85.0**	79.9	**88.5**	**82.2**	**90.8**

No meta-learning

Few-shot learning

Segmentation

RePRI
[Boudiaf et al, CVPR'21]

* The initial model is trained over the base classes following standard segmentation training (i.e., CE)

Few-shot learning

Segmentation

RePRI
[Boudiaf et al, CVPR'21]

* The initial model is trained over the base classes following standard segmentation training (i.e., CE)

$$\min \quad -\frac{1}{|\mathcal{L}|}\sum_{p\in\mathcal{L}} l(\mathbf{y}^p, \mathbf{s}_\theta^p) - \lambda_{\mathcal{H}}\frac{1}{|\mathcal{Q}|}\sum_{j\in\mathcal{Q}} \mathbf{s}_\theta^j \log(\mathbf{s}_\theta^j) + \lambda_{KL}(\hat{\mathbf{s}}_\theta^{\mathcal{Q}} \log(\frac{\hat{\mathbf{s}}_\theta^{\mathcal{Q}}}{\tau}))$$

CE on supervised images (i.e., support)

Few-shot learning

Segmentation

RePRI
[Boudiaf et al, CVPR'21]

* The initial model is trained over the base classes following standard segmentation training (i.e., CE)

$$\min \quad -\frac{1}{|\mathcal{L}|}\sum_{p\in\mathcal{L}} l(\mathbf{y}^p, \mathbf{s}_\theta^p) - \lambda_{\mathcal{H}}\frac{1}{|\mathcal{Q}|}\sum_{j\in\mathcal{Q}} \mathbf{s}_\theta^j \log(\mathbf{s}_\theta^j) + \lambda_{KL}(\hat{\mathbf{s}}_\theta^{\mathcal{Q}} \log(\frac{\hat{\mathbf{s}}_\theta^{\mathcal{Q}}}{\tau}))$$

CE on supervised images (i.e., support)

Entropy on unsupervised images (i.e., queries)

Few-shot learning

Segmentation

RePRI
[Boudiaf et al, CVPR'21]

* The initial model is trained over the base classes following standard segmentation training (i.e., CE)

$$\min \quad -\frac{1}{|\mathcal{L}|}\sum_{p\in\mathcal{L}} l(\mathbf{y}^p, \mathbf{s}_\theta^p) - \lambda_{\mathcal{H}} \frac{1}{|\mathcal{Q}|}\sum_{j\in\mathcal{Q}} \mathbf{s}_\theta^j \log(\mathbf{s}_\theta^j) + \lambda_{KL}(\hat{\mathbf{s}}_\theta^{\mathcal{Q}} \log(\frac{\hat{\mathbf{s}}_\theta^{\mathcal{Q}}}{\tau}))$$

CE on supervised images (i.e., support)

Entropy on unsupervised images (i.e., queries)

KL (to impose target proportions)

$$\hat{\mathbf{s}}_\theta^{\mathcal{Q}} = \frac{1}{|\mathcal{Q}|}\sum_{j\in\mathcal{Q}} \mathbf{s}_\theta^j$$

$$\tau \in [0,1]$$

Prior proportion

Unified view
(Based on InfoMax)

$$I(X,Y) = H(Y) - H(Y|X)$$

Label marginal Posteriors

Unified view
(Based on InfoMax)

$$I(X, Y) = H(Y) - H(Y|X)$$

Label marginal

uniform

$$Cte - KL(\tau||\mathbf{u})$$

TIM [Boudiaf et al, NeurIPS'20]

Unified view
(Based on InfoMax)

$$I(X,Y) = H(Y) - H(Y|X)$$

Label marginal

prior

$$Cte - KL(\tau||\mathbf{q})$$

RePRI [Boudiaf et al, CVPR'21]
AdaEnt [Bateson et al., MICCAI'20]

Unified view
(Based on InfoMax)

$$I(X,Y) = H(Y) - H(Y|X)$$

Label marginal

prior

$$Cte - KL(\tau||\mathbf{q})$$

Relax to inequality constraints

Size loss [Kervadec et al, MedIA'19]
Shape descriptors [Kervadec et al., MIDL'21]

Unified view
(Based on InfoMax)

$$I(X,Y) = H(Y) - H(Y|X)$$

But there are more!!

Weakly supervised segmentation

Size loss [Kervadec et al, MedIA’19]

Shape descriptors [Kervadec et al., MIDL’21]

Few-shot learning

TIM [Boudiaf et al, NeurIPS’20]

RePRI [Boudiaf et al., CVPR’21]

Unsupervised Domain adaptation

AdaEnt [Bateson et al, MICCAI’19]

Unified view
(Based on InfoMax)

$$I(X,Y) = H(Y) - H(Y|X)$$

But there are more!!

Weakly supervised segmentation
Size loss [Kervadec et al, MedIA'19]
Shape descriptors [Kervadec et al., MIDL'21]

Few-shot learning
TIM [Boudiaf et al, NeurIPS'20]
RePRI [Boudiaf et al., CVPR'21]

Unsupervised Domain adaptation
AdaEnt [Bateson et al, MICCAI'19]

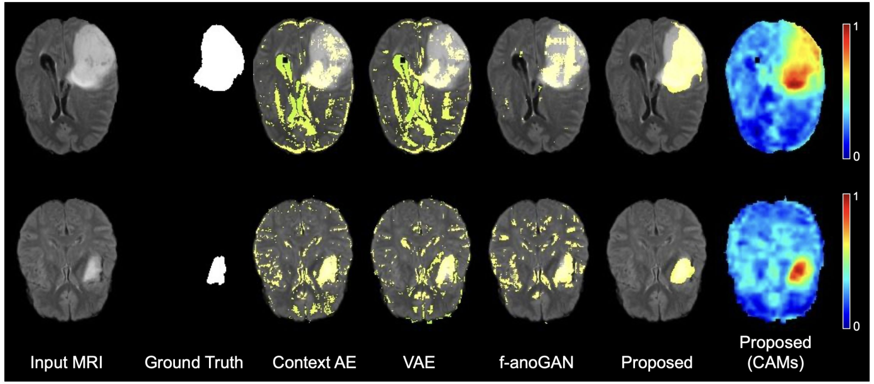
Generalized Few-shot segmentation
DiAM [Hajimiri et al, CVPR'23]

Mixed-supervised segmentation
[Dolz et al, IPMI'21]
[Liu et al, MedIA'22]

Generalized class discovery
MiB [Chiaroni et al, Arxiv'22]

Other research topics

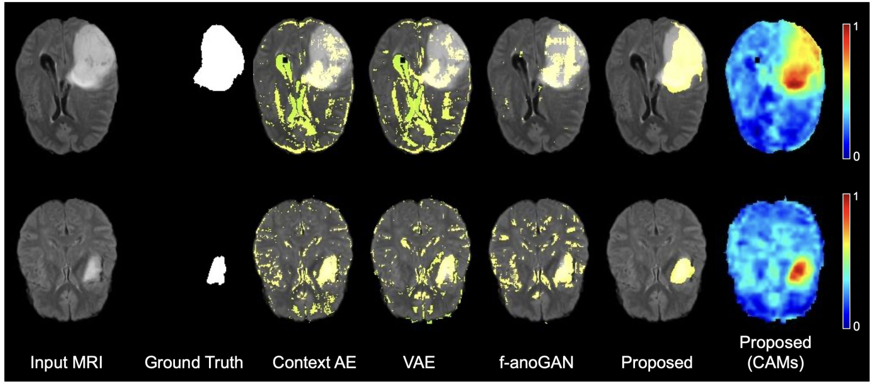
Unsupervised anomaly segmentation



Silva-Rodriguez et al, BMVC'21
Silva-Rodriguez et al, MedIA'22

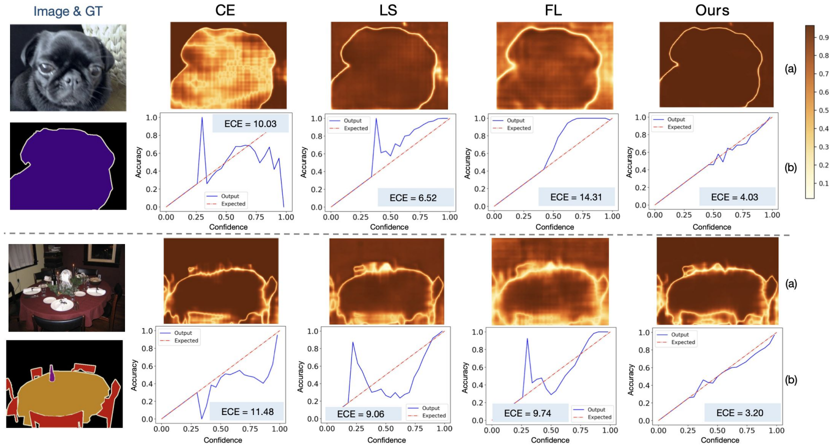
Other research topics

Unsupervised anomaly segmentation



Silva-Rodriguez et al, BMVC'21
Silva-Rodriguez et al, MedIA'22

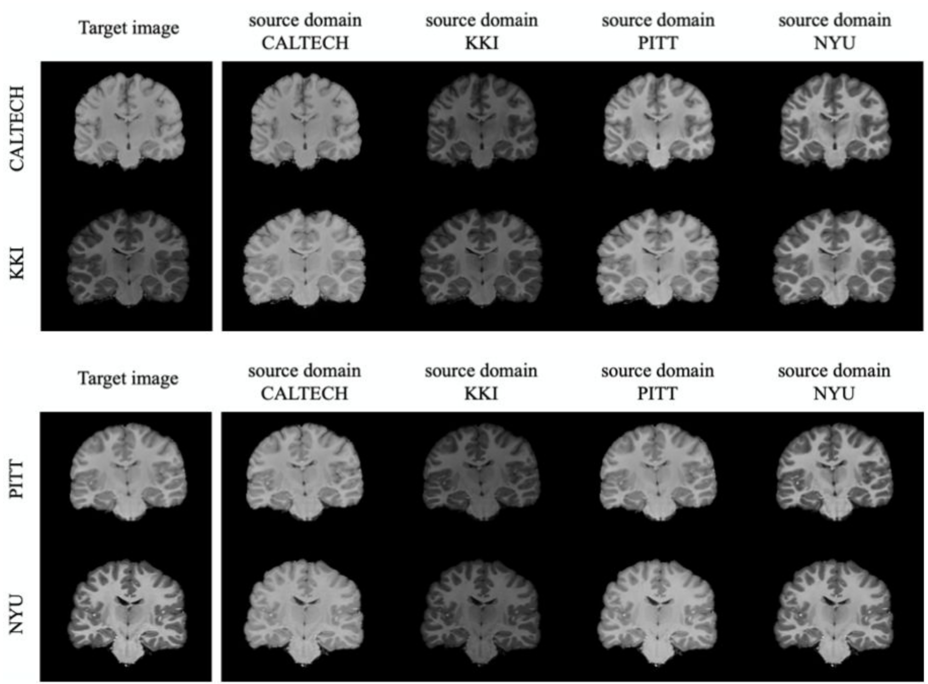
Calibrating neural networks



MbLS [Liu et al, CVPR'22]
CALS [Liu et al, CVPR'23]
MEEP [Larrazabal et al, MICCAI'23]
[Murugesan et al, MedIA'23]
[Murugesan et al, Arxiv'23]

Other research topics

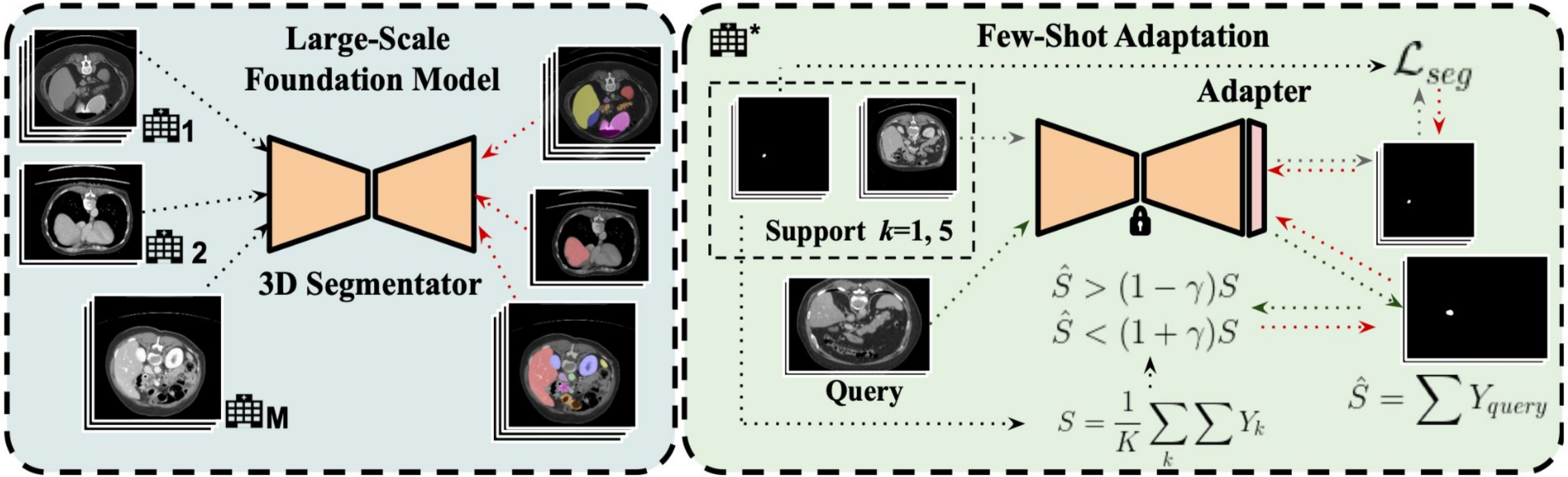
Data harmonization



Harmonizing-Flows [Beizaee et al, IPMI'23]

Other research topics

Foundation models



Silva-Rodriguez et al, Arxiv'23

Take-home message

- The lack of labeled data open the door for many interesting challenges
- Leveraging prior domain knowledge (in the form of constraints, or others methods) can further improve the discriminative performance.
- Few constraints have been explored under low-labeled data regime
- Room for improvement (many opportunities beyond weakly and few-shot supervised segmentation)

A big thank to my collaborators!








I am hiring!

Prompted text in Dall-E: ‘an image with the text 'I am hiring', with a realistic bear holding a beer