Bounding boxes for weakly supervised segmentation: Global constraints get close to full supervision

MIDL 2020, Montréal
Paper O-001

Hoel Kervadec, Jose Dolz, Shanshan Wang, Eric Granger, Ismail Ben Ayed

July 6 2020

ÉTS Montréal
hoel@kervadec.science
https://github.com/LIVIAETS/boxes_tightness_prior

Presentation overview

- On the (un)certainty of weak labels

Presentation overview

- On the (un)certainty of weak labels
- Tightness prior: application to bounding boxes

Presentation overview

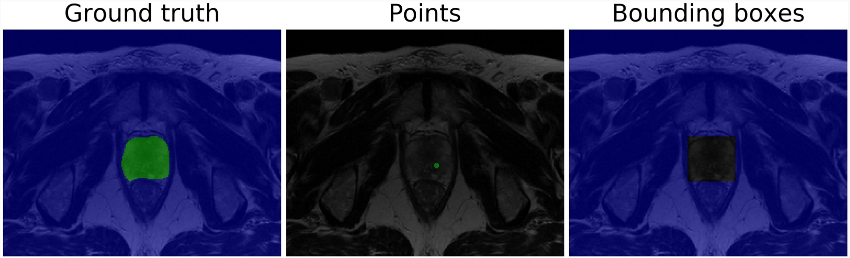
- On the (un)certainty of weak labels
- Tightness prior: application to bounding boxes
- Constraining a deep network during training

Presentation overview

- On the (un)certainty of weak labels
- Tightness prior: application to bounding boxes
- Constraining a deep network during training
- Results and conclusion

On the (un)certainty of weak labels

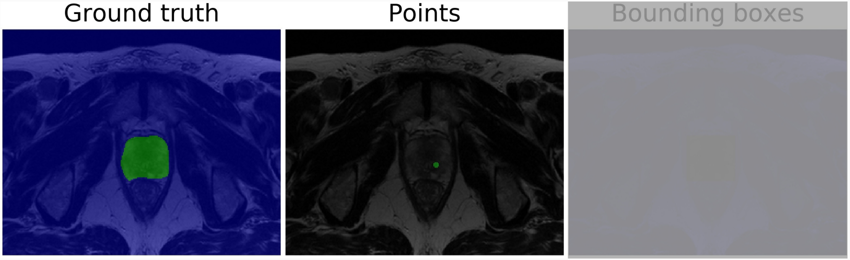
Weak labels



Blue: background, green: foreground, no-color: unknown.

Full labels are expensive, but weak labels are difficult to use

Constrained-CNN losses, with points [Kervadec et al., MedIA'19]



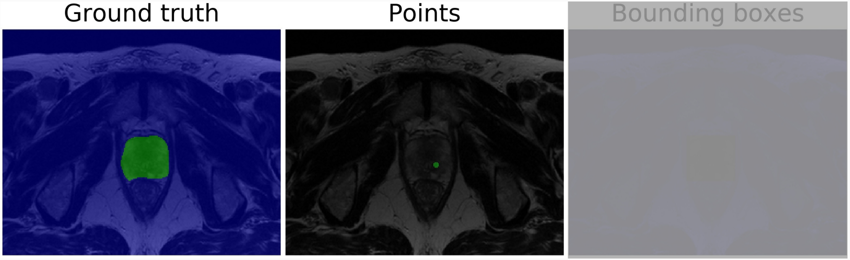
Partial cross-entropy **on the foreground pixels**, with size constraint:

$$\min_{\boldsymbol{\theta}} \sum_{p \in \Omega_L} - \log(s^p_{\boldsymbol{\theta}})$$

$$\text{s.t. } a \leq \sum_{p \in \Omega} s^p_{\boldsymbol{\theta}} \leq b$$

$\boldsymbol{\theta}$ Network parameters

Constrained-CNN losses, with points [Kervadec et al., MedIA'19]



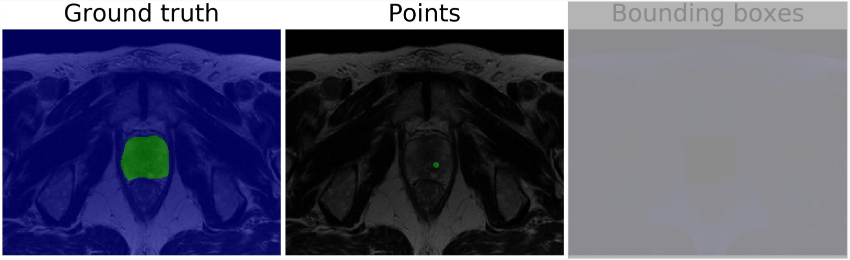
Partial cross-entropy **on the foreground pixels**, with size constraint:

$$\min_{\boldsymbol{\theta}} \sum_{p \in \Omega_L} -\log(s_{\boldsymbol{\theta}}^p)$$

$$\text{s.t. } a \leq \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq b$$

$\boldsymbol{\theta}$	Network parameters
Ω	Image space
$\Omega_L \subset \Omega$	Labeled pixels

Constrained-CNN losses, with points [Kervadec et al., MedIA'19]



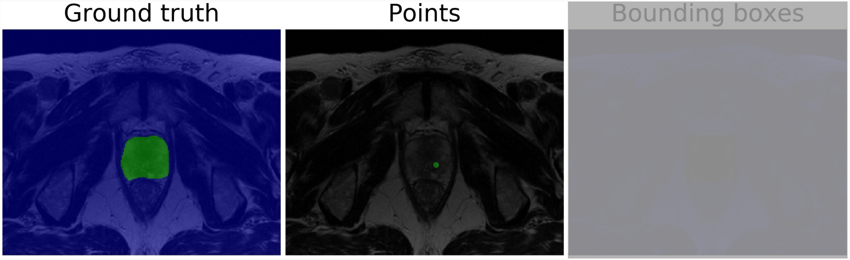
Partial cross-entropy **on the foreground pixels**, with size constraint:

$$\min_{\boldsymbol{\theta}} \sum_{p \in \Omega_L} -\log(s_{\boldsymbol{\theta}}^{p})$$

$$\text{s.t. } a \leq \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^{p} \leq b$$

$\boldsymbol{\theta}$	Network parameters
Ω	Image space
$\Omega_L \subset \Omega$	Labeled pixels
$p \in \Omega$	pixel

Constrained-CNN losses, with points [Kervadec et al., MedIA'19]



Partial cross-entropy **on the foreground pixels**, with size constraint:

$$\min_{\boldsymbol{\theta}} \sum_{p \in \Omega_L} -\log(s_{\boldsymbol{\theta}}^p)$$

$$\text{s.t. } a \leq \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq b$$

$\boldsymbol{\theta}$	Network parameters
Ω	Image space
$\Omega_L \subset \Omega$	Labeled pixels
$p \in \Omega$	pixel
$s_{\boldsymbol{\theta}}^p$	Foreground probability

Constrained-CNN losses, with points [Kervadec et al., MedIA'19]

It works well, but required some precise size information (a, b).

Constrained-CNN losses, with points [Kervadec et al., MedIA'19]

It works well, but required some precise size information (a, b).

How to realistically get it?

Constrained-CNN losses, with points [Kervadec et al., MedIA'19]

It works well, but required some precise size information (a, b).

How to realistically get it?

A bounding box gives a natural upper size.

But cannot do the opposite with a box

But cannot do the opposite with a box

Partial cross-entropy on the background pixels, with size constraint:

Ω_O Outside of the box

$$\min_{\boldsymbol{\theta}} \sum_{p \in \Omega_O} -\log(1 - s_{\boldsymbol{\theta}}^p)$$

$$\text{s.t.} \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq |\Omega_I|$$

But cannot do the opposite with a box

Partial cross-entropy on the background pixels, with size constraint:

$$\min_{\boldsymbol{\theta}} \sum_{p \in \Omega_O} -\log(1 - s_{\boldsymbol{\theta}}^p)$$

$$\text{s.t.} \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq |\Omega_I|$$

Ω_O	Outside of the box
Ω_I	Inside of the box

But cannot do the opposite with a box

Partial cross-entropy on the background pixels, with size constraint:

$$\min_{\boldsymbol{\theta}} \sum_{p \in \Omega_O} -\log(1 - s_{\boldsymbol{\theta}}^p)$$
$$\text{s.t.} \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq |\Omega_I|$$

Ω_O	Outside of the box
Ω_I	Inside of the box
$1 - s_{\boldsymbol{\theta}}^p$	Background probability

Why it does not work?

$$\min_{\boldsymbol{\theta}} \sum_{p \in \Omega_O} - \log(1 - s_{\boldsymbol{\theta}}^p)$$
$$\text{s.t.} \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq |\Omega_I|$$

Why it does not work?

$$
\begin{aligned}
\min_{\boldsymbol{\theta}} \sum_{p \in \Omega_O} & -\log(1 - s_{\boldsymbol{\theta}}^p) \\
\text{s.t. } & \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq |\Omega_I|
\end{aligned}
$$

Introduce *massive* imbalance in training.

Why it does not work?

$$\min_{\boldsymbol{\theta}} \sum_{p \in \Omega_O} -\log(1 - s_{\boldsymbol{\theta}}^p)$$
$$\text{s.t.} \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq |\Omega_I|$$

Introduce *massive* imbalance in training.

No *explicit* supervision to predict foreground.

Why it does not work?

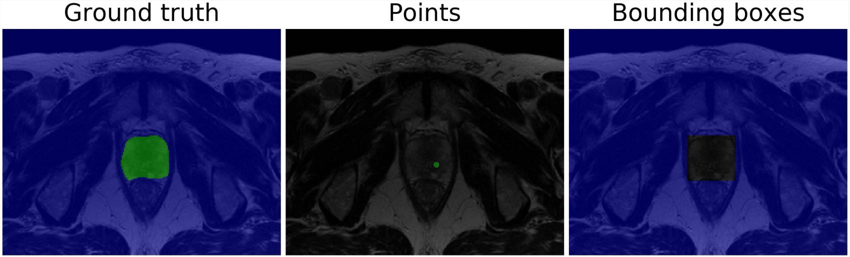
$$\min_{\boldsymbol{\theta}} \sum_{p \in \Omega_O} -\log(1 - s_{\boldsymbol{\theta}}^p)$$
$$\text{s.t.} \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq |\Omega_I|$$

Introduce *massive* imbalance in training.

No *explicit* supervision to predict foreground.

Result: It predicts only background.

Dirty solution – Mixed labels



We could mix the two kind of labels.

But defeat the purpose of having less annotations.

Dirty solution – Ugly heuristic

Or use a heuristic: The center of the box is always foreground.

Dirty solution – Ugly heuristic

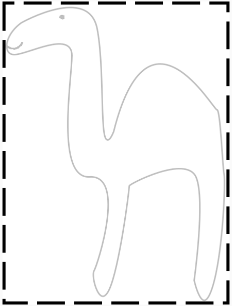
Hypothesis: The same part of the box always belong to the foreground.

Does it hold for more complex, deformable objects?

Dirty solution – Ugly heuristic

Hypothesis: The same part of the box always belong to the foreground.

Does it hold for more complex, deformable objects?

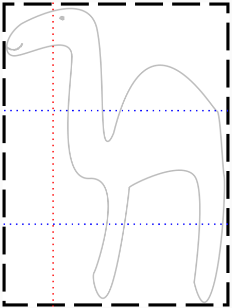


If the camel moves, our heuristic will be wrong.

Tightness prior

Tightness prior

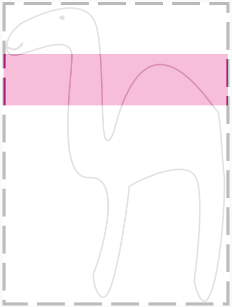
The classical tightness prior [Lempitsky et al., ICCV'09] states that:



Any line parallel to the box will cross the camel, at some point.

Tightness prior

Which can be generalized:



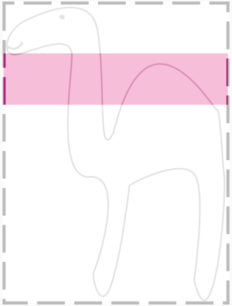
A segment of width w will cross-the camel w times.

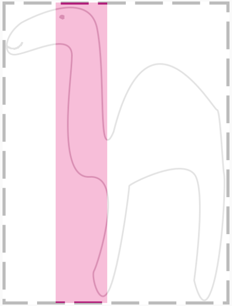
Formal definition

Formal definition

$\mathcal{S}_L := \{s_l\}$	set of segments
w	width of a segment
$y_p \in \{0,1\}$	true label for pixel p

$$\sum_{p \in s_l} y_p \geq w \qquad \forall s_l \in \mathcal{S}_L$$

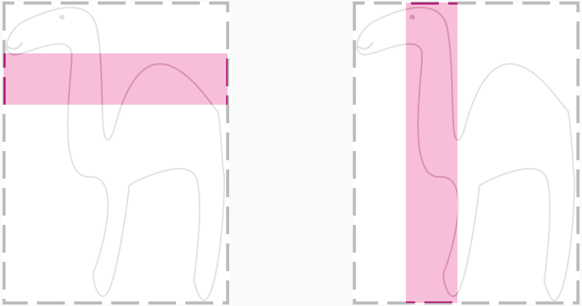




Formal definition

$\mathcal{S}_L := \{s_l\}$	set of segments
w	width of a segment
$y_p \in \{0, 1\}$	true label for pixel p

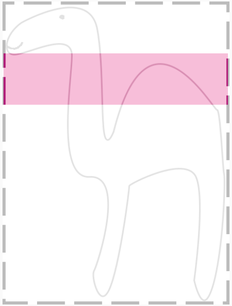
$$\sum_{p \in s_l} y_p \geq w \qquad \forall s_l \in \mathcal{S}_L$$

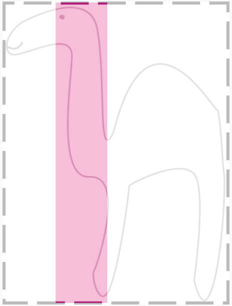


Formal definition

$\mathcal{S}_L := \{s_l\}$	set of segments
w	width of a segment
$y_p \in \{0, 1\}$	true label for pixel p

$$\sum_{p \in s_l} y_p \geq w \qquad \forall s_l \in \mathcal{S}_L$$

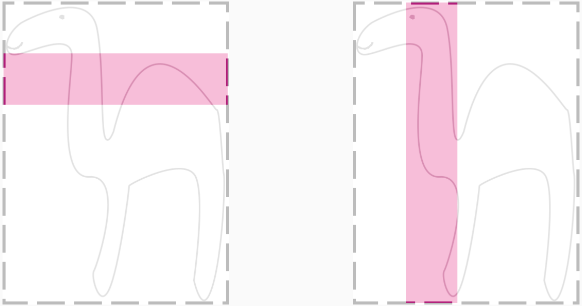




Formal definition

$\mathcal{S}_L := \{s_l\}$	set of segments
w	width of a segment
$y_p \in \{0, 1\}$	true label for pixel p

$$\sum_{p \in s_l} y_p \geq w \qquad \forall s_l \in \mathcal{S}_L$$



Updating the formulation

We can update our bounding box supervision model:

$\mathcal{L}_O$ Loss outside the box

$$\min_{\boldsymbol{\theta}} \mathcal{L}_O(\boldsymbol{\theta})$$
$$\text{s.t.} \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq |\Omega_I|$$

Updating the formulation

We can update our bounding box supervision model:

$\mathcal{L}_O$ Loss outside the box

$$\min_{\boldsymbol{\theta}} \mathcal{L}_O(\boldsymbol{\theta})$$

$$\text{s.t.} \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq |\Omega_I|$$

$$\text{s.t.} \sum_{p \in s_l} s_{\boldsymbol{\theta}}^p \geq w \qquad \forall s_l \in \mathcal{S}_L.$$

Updating the formulation

We can update our bounding box supervision model:

$$\min_{\boldsymbol{\theta}} \mathcal{L}_O(\boldsymbol{\theta})$$

$$\text{s.t.} \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq |\Omega_I|$$

$$\text{s.t.} \sum_{p \in s_l} s_{\boldsymbol{\theta}}^p \geq w \qquad \forall s_l \in \mathcal{S}_L.$$

$\mathcal{L}_O$	Loss outside the box
$\sum_{p \in s_l} s_{\boldsymbol{\theta}}^p$	Sum on continuous values

Updating the formulation

We can update our bounding box supervision model:

$$\begin{aligned} &\min_{\boldsymbol{\theta}} \mathcal{L}_O(\boldsymbol{\theta}) \\ &\qquad \text{s.t. } \sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p \leq |\Omega_I| \\ &\qquad \text{s.t. } \sum_{p \in s_l} s_{\boldsymbol{\theta}}^p \geq w \qquad \forall s_l \in \mathcal{S}_L. \end{aligned}$$

$\mathcal{L}_O$	Loss outside the box
$\sum_{p \in s_l} s_{\boldsymbol{\theta}}^p$	Sum on continuous values

Gives an **optimization problem with dozens of constraints**.

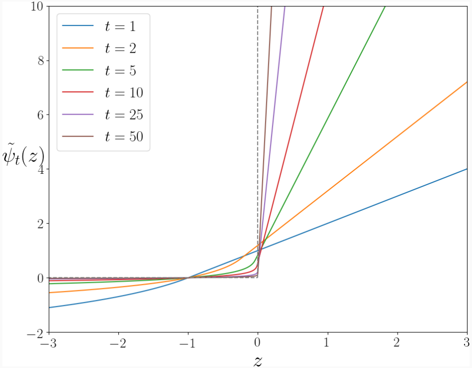
On constrained deep-networks during training

Penalty method such as [Kervadec et al., MedIA'19] or tweaked Lagrangian methods [Nandwani et al., 2019, Pathak et al., 2015] crumble with many competing constraints.

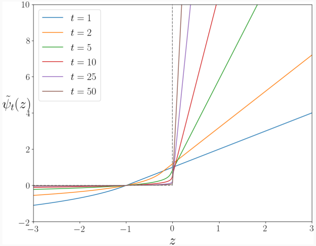
On constrained deep-networks during training

Penalty method such as [Kervadec et al., MedIA'19] or tweaked Lagrangian methods [Nandwani et al., 2019, Pathak et al., 2015] crumble with many competing constraints.

Recent work on extended log-barrier [Kervadec et al., 2019b] is much more robust:



Extended log-barrier



The ext. log-barrier is integrated directly into the loss function.

Model to optimize:

$$\min_x \mathcal{L}(x)$$

$$\text{s.t. } z \leq 0$$

Model w/ extended log-barrier:

$$\min_x \mathcal{L}(x) + \tilde{\psi}_t(z)$$

Final model

$$\min_{\boldsymbol{\theta}} \; \mathcal{L}_O(\boldsymbol{\theta}) + \lambda \left[\sum_{s_l \in \mathcal{S}_L} \tilde{\psi}_t \left(w - \sum_{p \in s_l} s_{\boldsymbol{\theta}}(p) \right) \right] + \tilde{\psi}_t \left(\sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p - |\Omega_I| \right)$$

Two simple hyper-parameters: weight λ for the tightness prior, t common to all constraints.

Final model

$$\min_{\boldsymbol{\theta}} \ \mathcal{L}_{O}(\boldsymbol{\theta}) + \lambda \left[\sum_{s_l \in \mathcal{S}_L} \tilde{\psi}_t \left(w - \sum_{p \in s_l} s_{\boldsymbol{\theta}}(p) \right) \right] + \tilde{\psi}_t \left(\sum_{p \in \Omega} s_{\boldsymbol{\theta}}^{p} - |\Omega_I| \right)$$

Two simple hyper-parameters: weight λ for the tightness prior, t common to all constraints.

Final model

$$\min_{\boldsymbol{\theta}} \; \mathcal{L}_O(\boldsymbol{\theta}) + \lambda \left[\sum_{s_l \in \mathcal{S}_L} \tilde{\psi}_t \left(w - \sum_{p \in s_l} s_{\boldsymbol{\theta}}(p) \right) \right] + \tilde{\psi}_t \left(\sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p - |\Omega_I| \right)$$

Two simple hyper-parameters: weight λ for the tightness prior, t common to all constraints.

Final model

$$\min_{\boldsymbol{\theta}} \ \mathcal{L}_O(\boldsymbol{\theta}) + \lambda \left[\sum_{s_l \in \mathcal{S}_L} \tilde{\psi}_t \left(w - \sum_{p \in s_l} s_{\boldsymbol{\theta}}(p) \right) \right] + \tilde{\psi}_t \left(\sum_{p \in \Omega} s_{\boldsymbol{\theta}}^p - |\Omega_I| \right)$$

Two simple hyper-parameters: weight λ for the tightness prior, t common to all constraints.

Evaluation and results

Datasets and baseline

Evaluate on two dataset:

- PROMISE12: prostate segmentation [Litjens et al., 2014]
- ATLAS: Ischemic stroke lesions [Liew et al., 2018]

Datasets and baseline

Evaluate on two dataset:

- PROMISE12: prostate segmentation [Litjens et al., 2014]
- ATLAS: Ischemic stroke lesions [Liew et al., 2018]

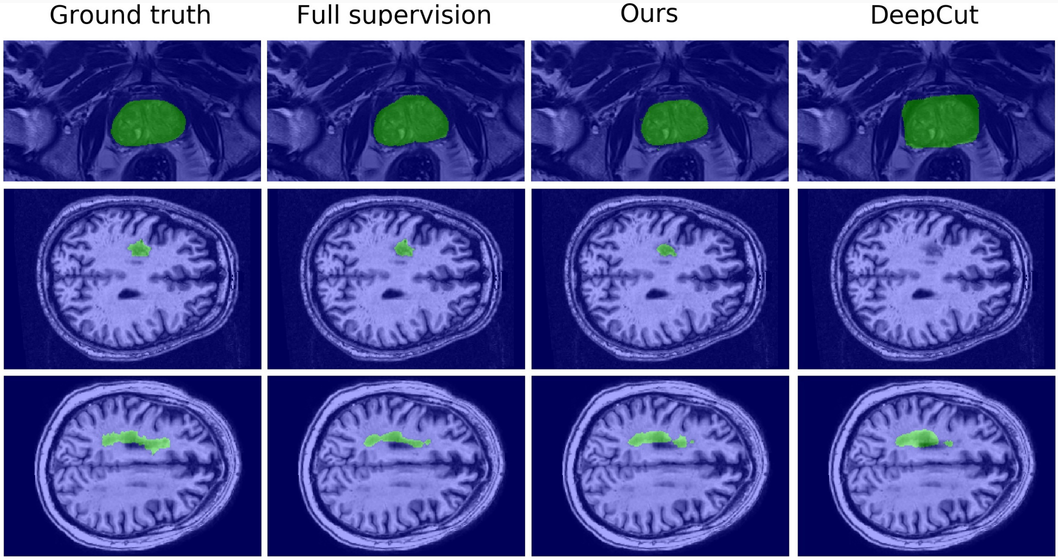
Use DeepCut [Rajchl et al., 2016] as baseline and comparison.

Results

Method	PROMISE12	ATLAS
	DSC	DSC
Deep cut [Rajchl et al., 2016]	0.827 (0.085)	0.375 (0.246)
$\mathcal{L}_{\mathrm{O}}$		
s.t. tightness prior	NA	0.161 (0.145)
s.t. tightness prior + box upper bound	**0.835 (0.032)**	**0.474 (0.245)**
Full supervision (Cross-entropy)	0.901 (0.025)	0.489 (0.294)

Results on both PROMISE12 and ATLAS datasets.

Results



Conclusion

Tightness prior, as a series of constraints, enables direct use of bounding boxes.

Compatible with other losses.

Conclusion

Tightness prior, as a series of constraints, enables direct use of bounding boxes.

Compatible with other losses.

More details in the paper (inner working of $\mathcal{L}_O$, computational cost, tightness sensitivity).

Code is publicly available:
`https://github.com/LIVIAETS/boxes_tightness_prior`

References i

Kervadec, H., Dolz, J., Tang, M., Granger, E., Boykov, Y., and Ben Ayed, I. (2019a).
Constrained-cnn losses for weakly supervised segmentation.
Medical Image Analysis.

Kervadec, H., Dolz, J., Yuan, J., Desrosiers, C., Granger, E., and Ben Ayed, I. (2019b).
Constrained deep networks: Lagrangian optimization via log-barrier extensions.
arXiv preprint arXiv:1904.04205.

References ii

Lempitsky, V., Kohli, P., Rother, C., and Sharp, T. (2009).
Image segmentation with a bounding box prior.
In *2009 IEEE 12th international conference on computer vision*, pages 277–284. IEEE.

Liew, S.-L., Anglin, J. M., Banks, N. W., Sondag, M., Ito, K. L., Kim, H., Chan, J., Ito, J., Jung, C., Khoshab, N., et al. (2018).
A large, open source dataset of stroke anatomical brain images and manual lesion segmentations.
Scientific data, 5:180011.

Litjens, G., Toth, R., van de Ven, W., Hoeks, C., Kerkstra, S., van Ginneken, B., Vincent, G., Guillard, G., Birbeck, N., Zhang, J., et al. (2014).
Evaluation of prostate segmentation algorithms for mri: the promise12 challenge.
Medical image analysis, 18(2):359–373.

Nandwani, Y., Pathak, A., Singla, P., et al. (2019).
A primal dual formulation for deep learning with constraints.
In *Advances in Neural Information Processing Systems*, pages 12157–12168.

Pathak, D., Krahenbuhl, P., and Darrell, T. (2015).
Constrained convolutional neural networks for weakly supervised segmentation.
In *IEEE International Conference on Computer Vision (ICCV)*, pages 1796–1804.

Rajchl, M., Lee, M. C., Oktay, O., Kamnitsas, K., Passerat-Palmbach, J., Bai, W., Damodaram, M., Rutherford, M. A., Hajnal, J. V., Kainz, B., et al. (2016).
Deepcut: Object segmentation from bounding box annotations using convolutional neural networks.
IEEE transactions on medical imaging, 36(2):674–683.