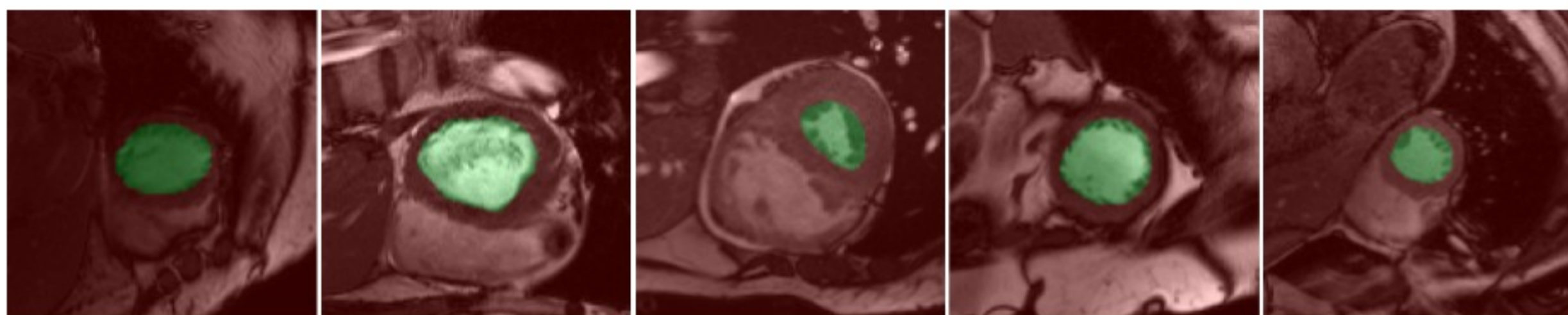
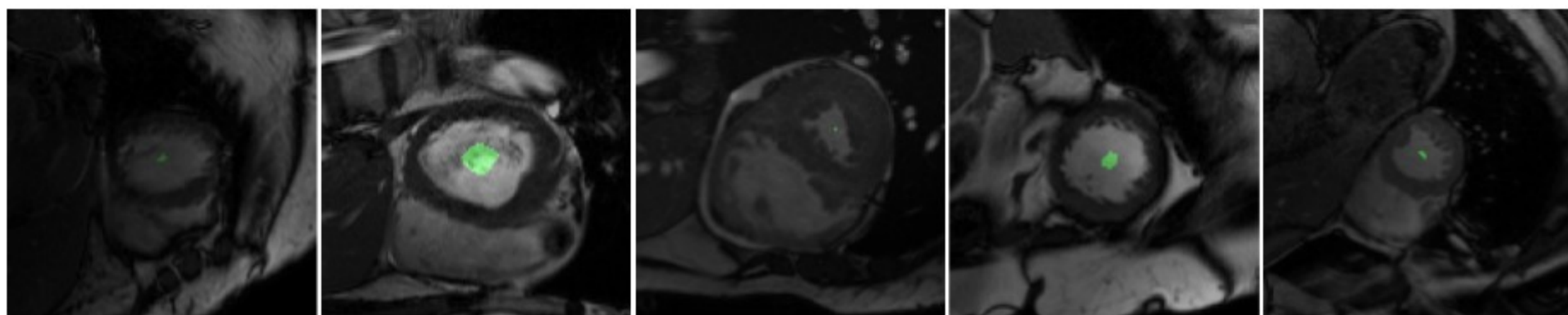


**Abstract:** In weakly supervised segmentation, enforcing global inequality constraints on CNN outputs can leverage unlabeled data, guiding training with domain-specific knowledge, e.g., the target-region size. We introduce a differentiable term, which enforces constraints directly in the loss. Surprisingly, our simple approach yields substantially better results than Lagrangian-dual optimization. We reach 90% of full supervision performances for left ventricle segmentation with only 0.1% of annotations.

Creating full mask for segmentation is **redundant and time-consuming**. Each slice in a MRI needs to be manually segmented. We want to use **weak labels** in the form of points ; which are much faster to produce.



a) Fully supervised labels



b) Weakly supervised labels

We propose [1] to **train the network under constraints**: the size of the predicted segmentation should be within reasonable bounds.

$$\min_{\theta} E(\hat{\mathbf{x}}_{\theta} | \mathbf{y})$$

s.t.  $a < S < b$

Size of the prediction (pixels)

"Sane" size (pixels)

Set of pixels having a label

Where:

$$S = \sum_{p \in \Omega} x_p$$

$$(a, b) = \begin{cases} (97, 1722) & \text{if } \Omega_L \neq \{\emptyset\} \\ (0, 0) & \text{otherwise} \end{cases}$$

Set of pixels in the image

For slices containing no left-ventricle, the set of labeled pixels will be empty, and the predicted size should be 0.

Otherwise, we use **bounds inferred from a single patient**. We compute the size of left-ventricle for all slices, and pick a and b as the minimum and maximum of those sizes.

Using a standard **Lagrangian method** to solve this is **difficult**, as we would need to alternate between complete training of the network and the update of the Lagrangian multipliers.

This was **partially solved by [2]**, where they use a proposal method and impose the constraints on the "fake" ground truth generated. But this model tends to **reinforce early mistakes** over time and proved to be **unstable** during training.

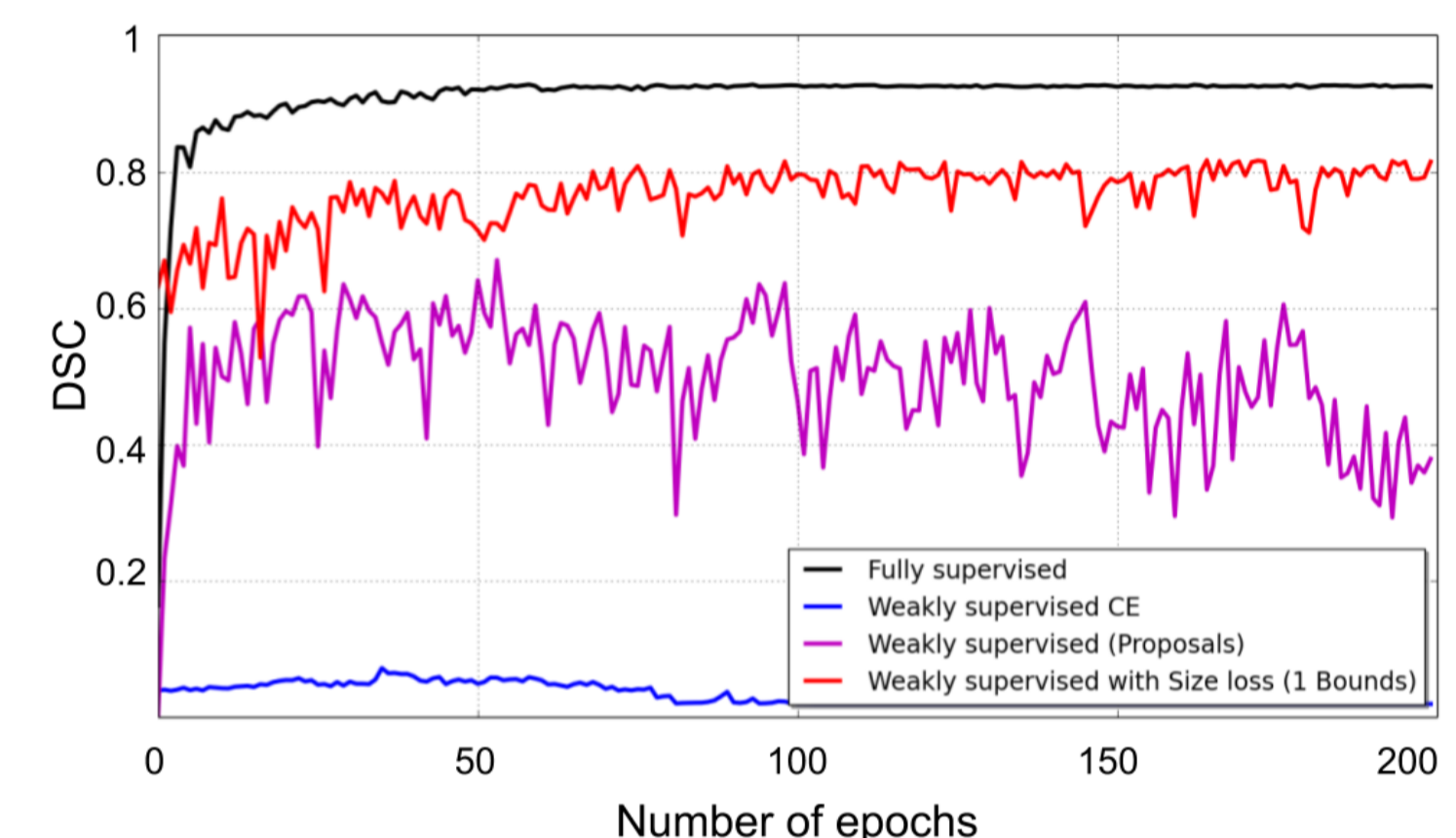
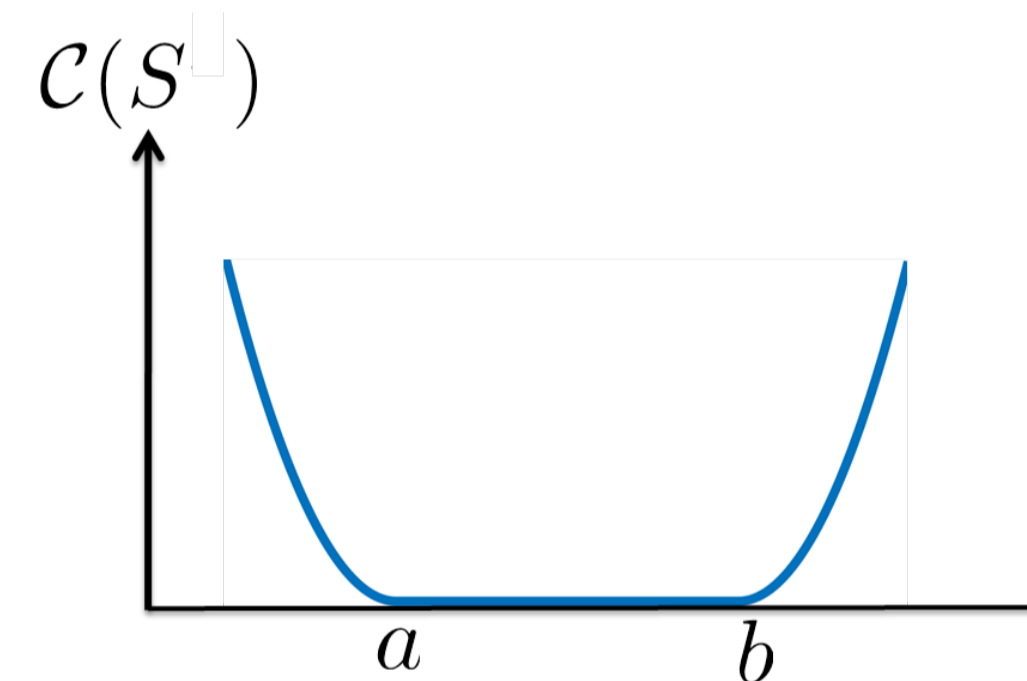
Instead, we **incorporate a penalty** for violated constraints directly **into the loss function**:

$$E(\mathbf{x} | \mathbf{y}) = \sum_{p \in \Omega_L} y_p \log \frac{1}{x_p} + \lambda \mathcal{C}(S)$$

$$\mathcal{C}(S) = 1_{[S < a]} (S - a)^2 + 1_{[S > b]} (S - b)^2$$

In all images, we use a **partial cross-entropy**: we apply it only on pixels having a label.

In parallel, if the predicted size is out of bounds, we **penalize all pixels equally**.

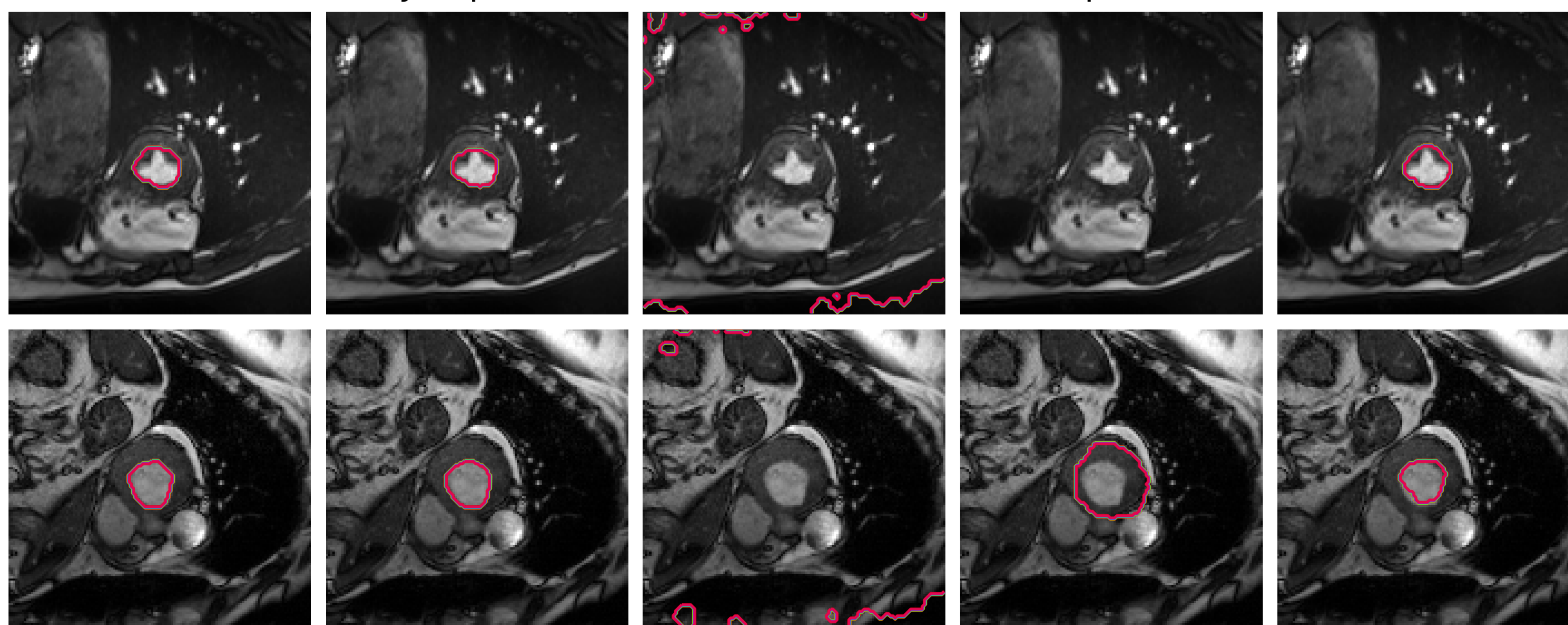


Results reach 90% of the full supervision dice, with a fraction (0.1%) of labels. We tried with one upper bound ( $S < b$ ) and two bounds ( $a < S < b$ ), resulting in a **3% improvement**. This is because we provide more information to the network, and penalize it when it is under-segmenting.

Future work will involve **multi-class problems** and other type of constraints (**length, centroid, compactness** of the segmentation).

We also intent to try different kinds of weak labels, notably **bounding boxes**, and investigate ways to deduce constraints values from the labels directly.

Ground truth    Fully supervised    Weak (CE)    Proposals    Ours



| Method              | Dice        |
|---------------------|-------------|
| Weak cross-entropy  | 7 %         |
| Proposals (1 bound) | 61 %        |
| Ours (1 bound)      | 81 %        |
| Ours (2 bounds)     | <b>84 %</b> |
| Full supervision    | 92 %        |

[1] Kervadec, H., Dolz, J., Tang, M., Granger, E., Boykov, Y., Ben Ayed, I. Size-constraint loss for weakly supervised CNN segmentation. MIDL 2018

[2] Pathak, D., Krahenbuhl, P., Darrell, T. (2015). Constrained convolutional neural networks for weakly supervised segmentation. CVPR 2015