Mask uncertainty regularization to improve machine learning-based medical image segmentation

Vladimir Groza* [1], *Benoit Huet* [1], *Nozha Boujemaa* [1]
[1] *Median Technologies, France*
**corresponding author e-mail: vladimir.groza@mediantechnologies.com*

Background

Segmentation of the different body structures on CT and MRI scans remains a challenging problem that requires accurate ground truth (GT) segmentation. One of the important aspects is the lack of reliability caused by radiologists annotation disagreement coupled with insufficient quality of the medical images.

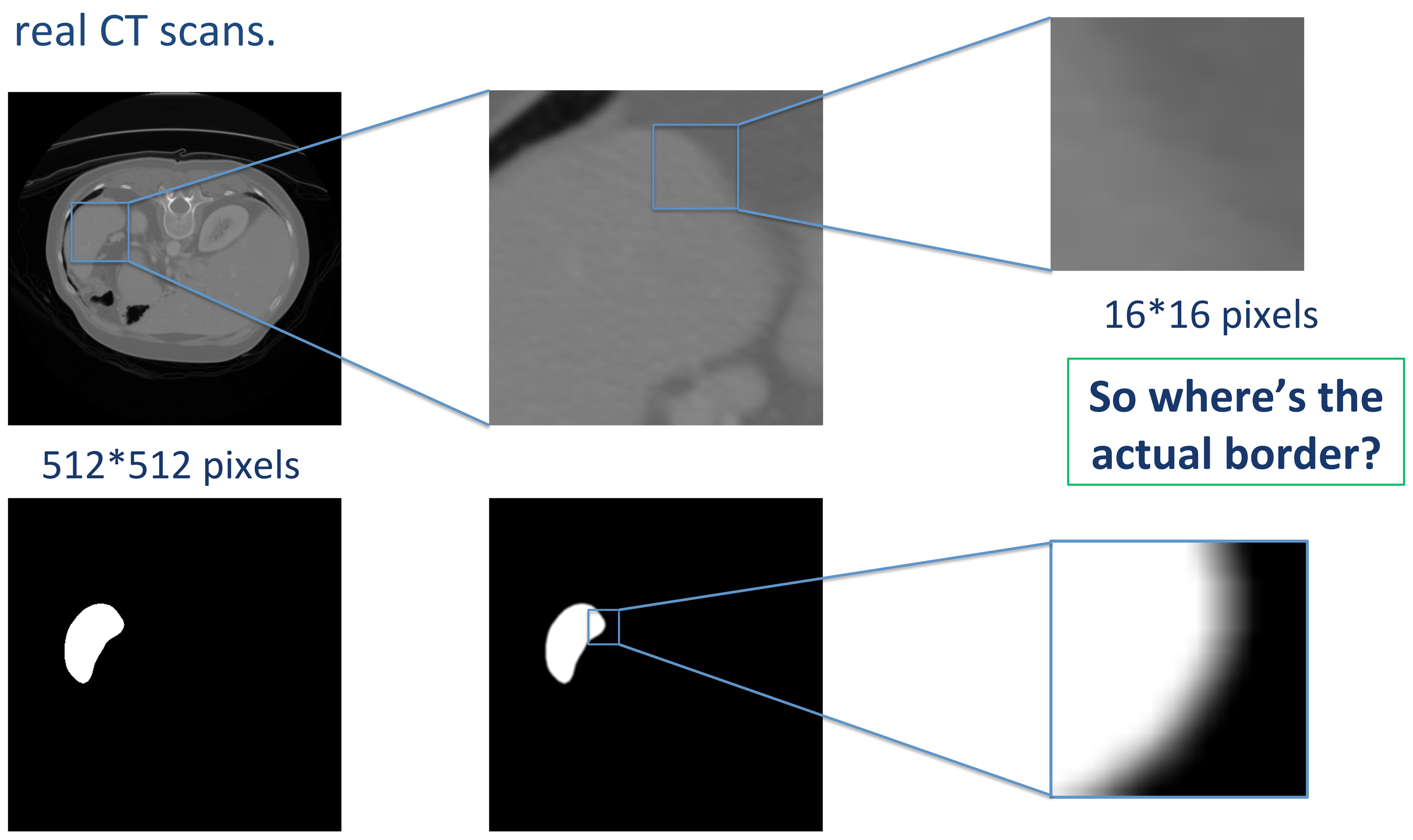
An independent multiple annotation is needed to overcome frequent disagreements between radiologists decisions mostly on the organs border, which is always blurred and affects on the providing an accurate segmentation even on contrasted CT/MRI images.

Approach

Augmentation techniques aim at increasing model generalization by manipulating inputs in order to enrich set of the variability of images in regards to their GT [1].

We introduce a method to regularize the models learning process by augmenting the information only in the associated GT masks.

On most of the studied CT scans the border pixels degradation which we observed is commonly in the band of 3-5 pixels. The range of 5 pixels was taken as a baseline in order to better mimic the borders in real CT scans.



Dataset

The Medical Decathlon public data set [2] used in this work consists of 41 training and 20 testing volumes.

To perform correct evaluation we used only 41 volumes from the training dataset with the available ground truth (GT) for spleen.

The total number of 3650 2D CT images was used to recreate train and test subsets with the ratio of 9:1, while the train set was also split into 5 folds to perform the 5-fold cross validation scheme.

Pipeline and Methods

To perform segmentation of the spleen in abdominal CT scans, we used the *LinkNet-base* pipeline with the *se-resnext50* backbone and corresponding decoder.

Decoder: each block consists of the 2D Convolution + Upsampling + 2D Convolution layer, each followed by the Batch Normalization and ReLU.

The training process includes two consecutive phases.
First : 100 epochs using the *Adam* optimizer with initial LR of 1e-4 and the *ReduceLROnPlateau* scheduling.
Second: 20 epochs using the SGD optimizer and Cyclic LR scheduling with the *base LR* of 5e-7 and *max LR* of 1e-3.

Valdation & Test:
- **No Augmentations**
- **No Smoothing**

Random Train Augmentations [3]:
- Random 90 degree rotations
- Horizontal / vertical flips
- Transpose
- Brightness / Contrast / Gamma

Results

Results for the binary spleen segmentation on each particular fold and for the mean averaged cross-fold ensemble are given in Table 1, where the results in brackets correspond to mask smoothing.

The proposed method significantly improves every single fold almost up to the level of the ensemble.

	DICE	Precision	Recall
Fold 0	0.9034 (**0.9280**)	**0.9504** (0.9434)	0.9018 (**0.9370**)
Fold 1	0.9098 (**0.9141**)	**0.9405** (0.9312)	0.9272 (**0.9352**)
Fold 2	0.9072 (**0.9328**)	0.9375 (**0.9464**)	0.9279 (**0.9405**)
Fold 3	0.9324 (**0.9418**)	0.9458 (**0.9601**)	**0.9466** (0.9389)
Fold 4	0.9169 (**0.9334**)	0.9267 (**0.9449**)	0.9370 (**0.9481**)
Ensemble	0.9432 (**0.9486**)	**0.9584** (0.9542)	0.9445 (**0.9535**)

Tab. 1 Results with and without GT border smoothing

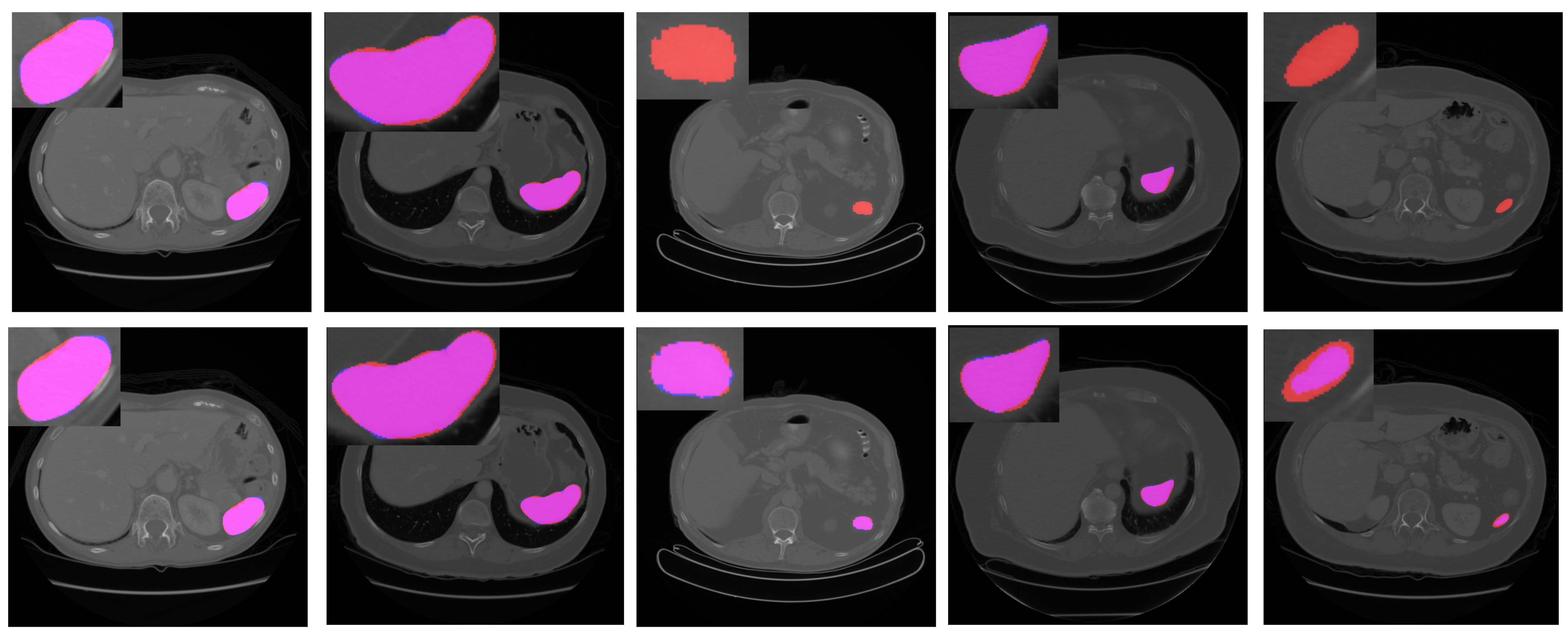


Fig. 1 Segmentation prediction without (first row) and with (second row) border smoothing

Conclusions

This study presented a novel regularization approach considering the GT masks used to improve the organs segmentation quality. We compare the impact of this method particularly on the abdominal CT scans and spleen as an example.

One of the important situation, where this improvement is valuable, is the further studying the segmented organ's surface and it's variation. Here each missed or extra pixel or "pixel-layer" completely changes the structure and properties of the surface.

- More accurate segmentation of the region along the border of the organ of interest.
- Usually leads to the reduced number of false negative pixels.
- Significant improvement on the top and bottom of the organ of interest

References

1. Muller R., Kornblith S., Hinton G.: When Does Label Smoothing Help?, arXiv: 1906.02629 (2019).
2. Amber L. Simpson et al.: A large annotated medical image dataset for the development and evaluation of segmentation algorithms, arXiv:1902.09063 (2019).
3. Buslaev A., Iglovikov V., Khvedchenya E., Parinov A., Druzhinin M., Kalinin A.: Albumentations: Fast and Flexible Image Augmentations. Information vol. 11-2 (2020).