Semi-supervised SPECT/CT Segmentation Using Convolutional Neural Networks

Junyu Chen, Ye Li, Yong Du, Eric C. Frey

The Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins Medical Institutes

Department of Electrical and Computer Engineering, Johns Hopkins University

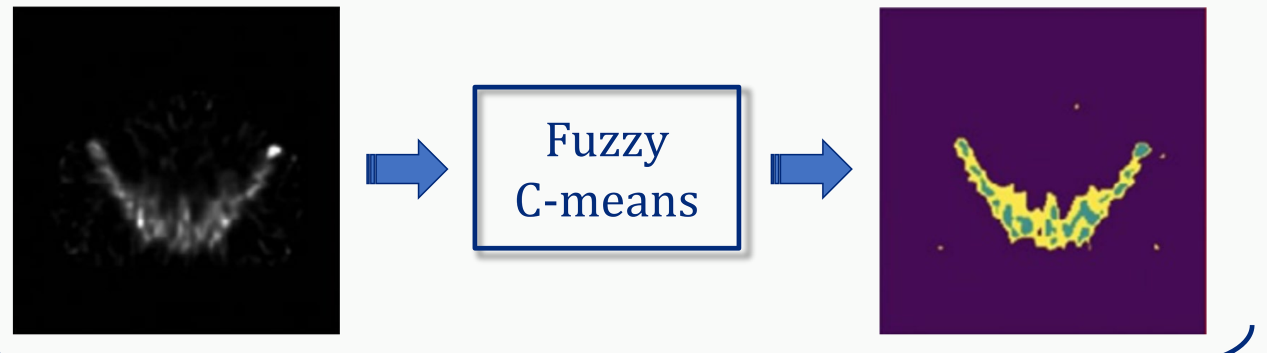
SNMMI 2021 Annual Meeting

Introduction

Quantitative bone single-photon emission computed tomography (QBSPECT) has the potential to provide a better quantitative assessment of bone metastasis than planar bone scintigraphy due to its ability to better quantify activity in overlapping structures. An important element of assessing the response of bone metastasis is accurate image segmentation. However, limited by the properties of QBSPECT images, the segmentation of anatomical regions-of-interests (ROIs) still relies heavily on the manual delineation by experts. This work proposes a fast and robust automated segmentation method for partitioning a QBSPECT image into lesion, bone, and background. Specifically, we present a new unsupervised segmentation loss function and its semi- and supervised variants for training a convolutional neural network. The method can operate in unsupervised, semi-supervised, or fully-supervised modes depending on the availability of annotated training data.

Objective

We aim to develop a fully automated method for partitioning voxels into back-ground, lesion, and bone in QBSPECT images. To this end, we propose a set of loss functions for training segmentation convolutional neural networks (ConvNets). The loss functions are based on the classical Fuzzy C-means (FCM) algorithm [1]. An important property of the proposed loss functions is that they incorporate the fundamental idea of fuzzy clustering, where the amount (fuzziness) of the overlap between segmentation classes is controlled via a user-defined hyperparameter.



Can a ConvNet learn this unsupervised segmentation scheme? Yes!!!

Materials & Methods

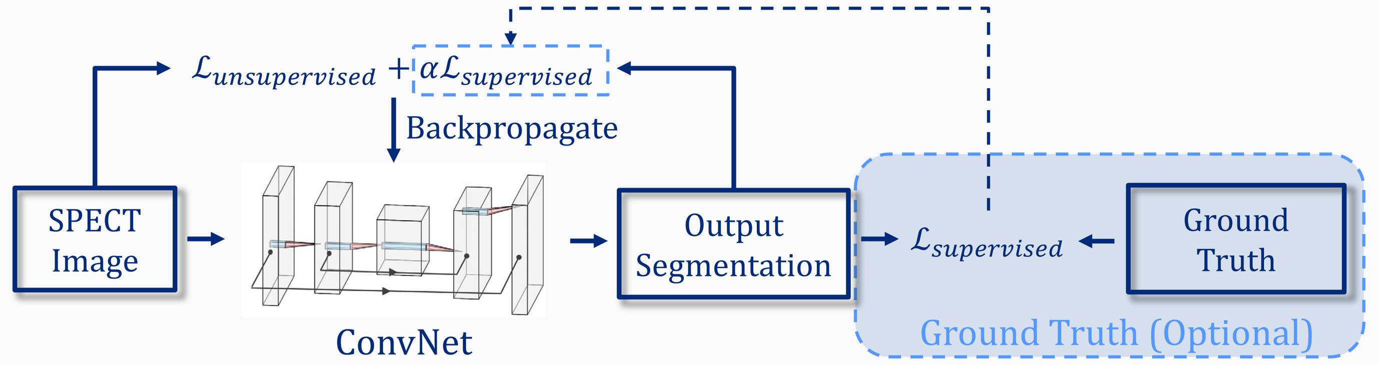


Fig. 1. Schematic of the proposed method.

We propose to model the membership functions, u, in the objective function of Robust FCM [2] by using the softmax output of the last layer of the ConvNet model, $f(\boldsymbol{y};\theta)$, using:

$$\mathcal{L}_{RFCM}(\boldsymbol{y};\theta)=\sum_{j\in\Omega}\sum_{k=1}^{C} f_{jk}^{q}(\boldsymbol{y};\theta)\|y_j-v_k\|^2+\beta\mathcal{R}(f_{jk}^{q}(\boldsymbol{y};\theta)),$$

where $f_{jk}(\boldsymbol{y};\theta)$ is the k^{th} channel softmax output from the ConvNet at location j, q controls the fuzzy overlap between classes, v_k is the class mean, and $\mathcal{R}(f_{jk}^{q}(\boldsymbol{y};\theta))$ is a noise regularizing term:

$$\mathcal{R}(f_{jk}^{q}(\boldsymbol{y};\theta))=\sum_{j\in\Omega}\sum_{k=1}^{C} f_{jk}^{q}(\boldsymbol{y};\theta)\sum_{l\in N_j}\sum_{m\in M_k} f_{lm}^{q}(\boldsymbol{y};\theta)$$

Where term N_j represents the neighboring voxels of voxel j, M_k is a set containing $\{1,\dots,C\}\backslash\{k\}$ (i.e., class numbers other than k).

While networks trained using unsupervised loss functions can be more robust to test images from an "unseen" domain, their performances are generally limited due to the sole dependence on intensity information. We propose to train a ConvNet using both supervised and unsupervised loss functions so that the network can consider the intensity statistics in an individual image while embracing the supervised information provided by the ground truth g. The proposed FCM-based supervised loss is defined as:

$$\mathcal{L}_{FCM_{label}}(\boldsymbol{y};\theta)=\sum_{j\in\Omega}\sum_{k=1}^{C} f_{jk}^{q}(\boldsymbol{y};\theta)\|g_{jk}-\mathbf{1}\|^2$$

The final semi-supervised loss can be written as:

$$\underbrace{\mathcal{L}_{semi-RFCM}^{\alpha}(\boldsymbol{y};\theta)}_{\text{Semi-supervised loss}}=\underbrace{\mathcal{L}_{RFCM}(\boldsymbol{y};\theta)}_{\text{Unsupervised loss}}+\underbrace{\alpha\mathcal{L}_{FCM_{label}}(\boldsymbol{y};\theta)}_{\text{Supervised loss}}.$$

Results

The proposed algorithm was assessed on three datasets, where the training and a testing set consist of the attenuation maps generated from the XCAT [3] phantom and their corresponding simulated SPECT images; another testing set is a set of clinical SPECT/CT bone scans obtained from an institutional-review-board-approved protocol. It is important to mention that the networks were ***trained using purely simulated images***. We compared the results obtained using the proposed loss functions with those from two widely used supervised loss functions, Dice loss ($\mathcal{L}_{DSC}$) and Cross-entropy loss ($\mathcal{L}_{CE}$). We also compared the results to a newly proposed unsupervised loss function, Mumford-Shah loss ($\mathcal{L}_{MS}$) [4], and a semi-supervised variant of it.

	SPECT Lesion		SPECT Bone		CT Bone	
Method	DSC	Surface DSC	DSC	Surface DSC	DSC	Surface DSC
$\mathcal{L}_{DSC}$	0.356	0.177	0.677	0.928	0.604	0.614
$\mathcal{L}_{CE}$	0.400	0.197	0.664	0.931	0.591	0.707
$\mathcal{L}_{semi-MS}^{0.1}$	0.677	0.590	0.701	0.939	0.652	0.760
$\mathcal{L}_{FCM_{label}}$	0.363	0.213	0.628	0.896	0.778	0.770
$\mathcal{L}_{semi-RFCM}^{0.1}$	0.747	0.788	0.742	0.950	0.794	0.841

Table 1. Quantitative evaluations on clinical SPECT/CT dataset.

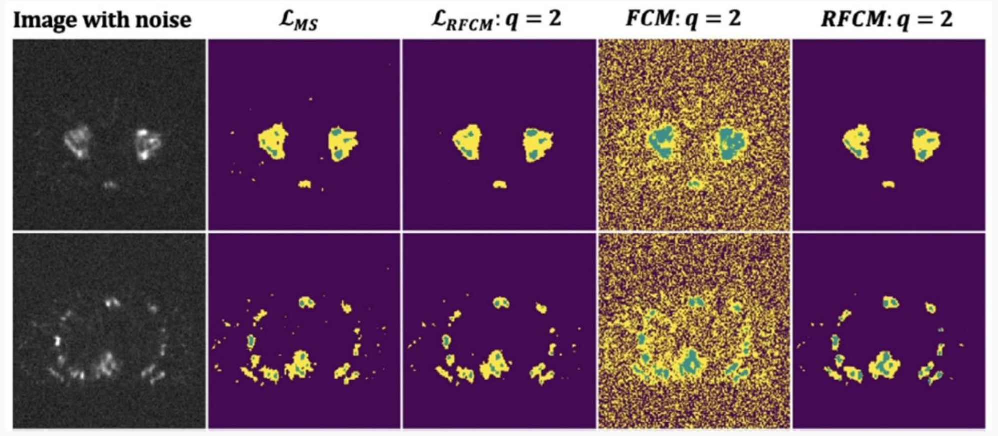


Fig. 2. Qualitative comparisons of different unsupervised segmentation methods on simulated SPECT images with additive Gaussian noise.

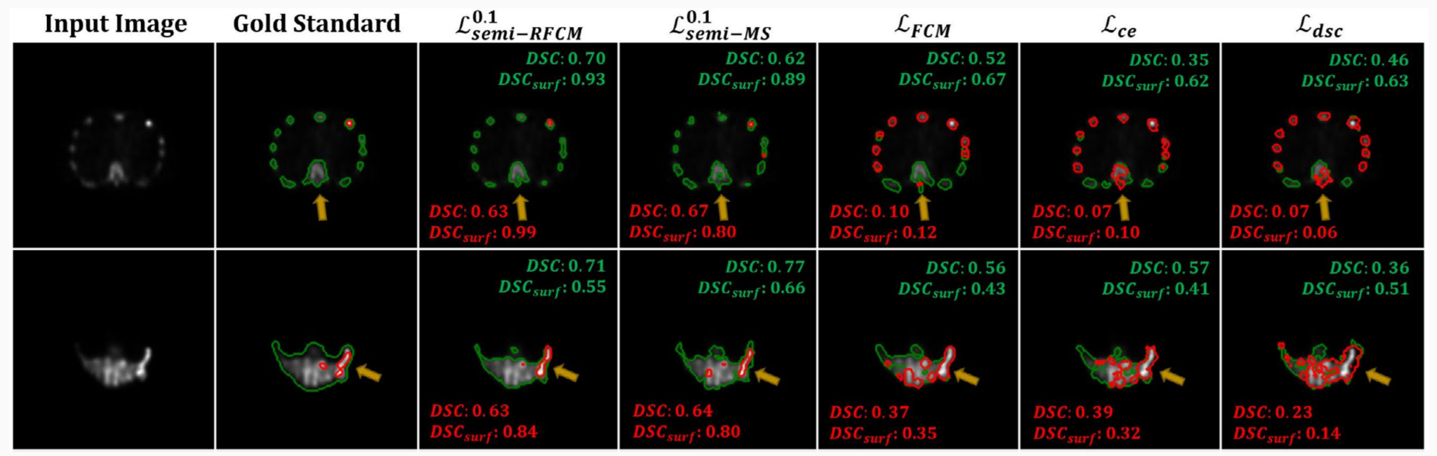


Fig. 3. Qualitative comparisons of different loss functions on clinical SPECT images.

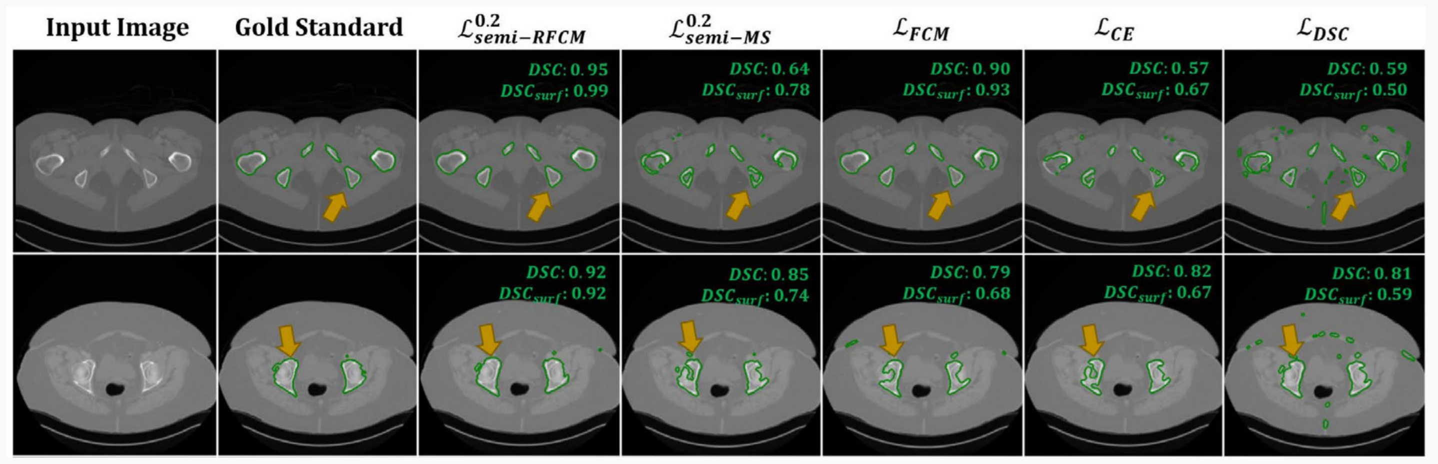


Fig. 3. Qualitative comparisons of different loss functions on clinical CT images..

Conclusions

In this work, we proposed a set of novel FCM-based loss functions for semi-, unsupervised, and supervised SPECT/CT segmentation using deep neural networks. An advantage of the proposed loss functions is that they enable the ConvNets to consider both voxel intensity and semantic information in the image during the training stage. The proposed loss functions also retain the fundamental property of the conventional fuzzy clustering, where the fuzzy overlap between the channels of softmax outputs can be adjusted by a hyperparameter in the loss function. Various experiments demonstrated that the model trained using a dataset of simulated images generalized well and led to fast and robust segmentation on both simulated and clinical SPECT/CT images.

This work has been published by Medical Physics: [5]

References

1. Bezdek JC, Ehrlich R, Full W. FCM: the fuzzy c-means clustering algorithm. *Comput Geosci.* 1984;10:191–203.
2. Pham DL. Spatial models for fuzzy clustering. *Comput Vis Image Underst.* 2001;84:285–297.
3. Segars WP, Mahesh M, Beck TJ, Frey EC, Tsui BM. Realistic CT simulation using the 4D XCAT phantom. *Med Phys.* 2008;35:3800–3808.
4. Kim B, Ye JC. Mumford-shah loss functional for image segmentation with deep learning. *IEEE Trans Image Process.* 2020;29:1856–1866.
5. Chen J, Li Y, Luna LP, Chung HW, Rowe SP, Du Y, Solnes LB, Frey EC. Learning Fuzzy Clustering for SPECT/CT Segmentation via Convolutional Neural Networks. *Medical Physics.* 2021 Apr 17.

Acknowledgements:
This work was supported by a grant from the National Cancer Institute, U01-CA140204. The views expressed in written materials or publications and by speakers and moderators do not necessarily reflect the official policies of the NIH; nor does mention by trade names, commercial practices, or organizations imply endorsement by the U.S. Government.