



Published in final edited form as:

Proc IEEE Int Symp Biomed Imaging. 2018 April ; 2018: 228–231. doi:10.1109/ISBI.2018.8363561.

3D FULLY CONVOLUTIONAL NETWORKS FOR CO-SEGMENTATION OF TUMORS ON PET-CT IMAGES

Zisha Zhong^{*}, Yusung Kim[†], Leixin Zhou^{*}, Kristin Plichta[†], Bryan Allen[†], John Buatti[†], Xiaodong Wu^{*,†}

^{*}Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA

[†]Department of Radiation Oncology, University of Iowa, Iowa City, IA

Abstract

Positron emission tomography and computed tomography (PET-CT) dual-modality imaging provides critical diagnostic information in modern cancer diagnosis and therapy. Automated accurate tumor delineation is essentially important in computer-assisted tumor reading and interpretation based on PET-CT. In this paper, we propose a novel approach for the segmentation of lung tumors that combines the powerful fully convolutional networks (FCN) based semantic segmentation framework (3D-UNet) and the graph cut based co-segmentation model. First, two separate deep UNets are trained on PET and CT, separately, to learn high level discriminative features to generate tumor/non-tumor masks and probability maps for PET and CT images. Then, the two probability maps on PET and CT are further simultaneously employed in a graph cut based co-segmentation model to produce the final tumor segmentation results. Comparative experiments on 32 PET-CT scans of lung cancer patients demonstrate the effectiveness of our method.

Index Terms—

image segmentation; lung tumor segmentation; co-segmentation; fully convolutional networks; deep learning

1. INTRODUCTION

PET-CT imaging has been a highly thriving and successful research field in medical image processing. A lot of research endeavors have been devoted for its clinical use, such as radiation therapy treatment planning [1]. A key and challenging step is the automated and accurate tumor delineation, which plays a vital role for subsequently determining the therapeutic option to achieve improved prognoses [1, 2].

Recently, the co-segmentation technique for tumor delineation using both PET-CT has attracted great attentions [3, 4, 5, 6, 7], where tumor contours on PET and on CT are segmented simultaneously while admitting their possible differences to accommodate the registration inaccuracy and imaging uncertainty. It has demonstrated in those previous works that the design of cost functions in the framework of graph-cut based co-segmentation is critical to achieve good segmentation performance. How to design efficient cost functions is still an open problem and many efforts have been devoted to tackle this challenge.

Representative work on cost function design include using sophisticated image priors (e.g., Gaussian mixture models [3, 4], texture information [6], shape prior [7], etc.) and/or clinical information [4, 5, 6, 7]. More recently, Zhong et al. [8] have introduced the 3D alpha matting technique to compute the region costs for co-segmentation on PET-CT images and validate the efficiency. Although these methods have achieved good performance, they are mostly semi-automatic which need a number of specific user-defined seeds, which may limit their usages in clinical practices.

In this paper, to address this challenge with a more intelligent way, we resort to data-driven deep learning technique to pursue better region costs and attempt to develop a computer-aided automatic processing pipeline for tumor segmentation. Specifically, we propose to integrate the 3D fully convolutional networks (3D-UNet) and the graph cut based co-segmentation model to simultaneously segment lung tumors on PET-CT scans. For the co-registered PET or CT scans, two independent deep 3D-UNets are first trained separately for PET and CT, respectively. Owing to their powerful descriptive capability, both networks can learn high level discriminative features that help generate high quality voxel-level tumor/non-tumor masks and probability maps. Then, the two probability maps, one on PET and the other on CT, are further employed in the graph cut based co-segmentation model [4] with label-consistency constraints to simultaneously produce the final tumor segmentation results on both PET and CT images. Our contribution of this work mainly lies in the following two points:

1. We propose to combine the 3D-UNet and graph cut based co-segmentation for tumors on PET-CT images. The advantage lies in automated localization of tumors, i.e., no need of manually defined object seeds compared to previous semi-automatic methods, which largely facilitates the subsequent clinical processing in diagnosis procedure that may be prone to errors.
2. In our experiments, we evaluate and validate in detail the proposed 3D-UNet based co-segmentation method by comparing with two state-of-the-art semi-automatic methods on 32 PET-CT image datasets.

2. METHODOLOGY

In recent years, deep learning has been proven to significantly outperform conventional statistical learning based approaches combined with manually-designed features from image, text or multi-modality inputs [9]. The medical imaging community has been rapidly entering the arena, and deep learning is quickly demonstrated to be the state-of-the-art tool for a wide variety of medical tasks, including segmentation [10, 11]. In this paper, we focus on investigating the FCN based semantic segmentation for tumor delineation on PET-CT scans. In literature, since the original FCN is proposed in [12], a number of variants have been developed in medical community, including 2D-UNet [13], 3D-UNet [14], 3D-DSN [15], etc. We attempt to utilize the 3D-UNet with an encoder-decoder architecture on our task owing to its powerful performance.

Our segmentation pipeline mainly consists of three steps (Fig. 1): data preprocessing, 3D UNets based FCN for probability maps generation, and graph-cut based co-segmentation.

Step (1): Data preprocessing is a basic yet vital step in medical image segmentation. Common operations include image registration, spatial resampling, image intensity value thresholding, etc. We will present details of this step in Section 3.1. Step (2): Two 3D-UNets are separately trained on preprocessed inputs, one on CT and one on PET. The 3D-UNets allows for capturing implicit and informative high-level features of tumors/non-tumors, which, however, usually produce coarse segmentation results that may not precisely localize tumor boundaries. Consequently, we perform Step (3) to further adopt the graph based co-segmentation model to refine the segmentation by considering potential label consistency between the dual-modality PET and CT images. Due to the space limit, we mainly present Step (2) in detail.

In our 3D-UNet framework, the encoder module contains 4 convolutional and max-pooling layers with 32, 64, 128, 256 feature maps, respectively, the decoder module contains 4 deconvolutional and convolutional layers with 256, 128, 64, 32 feature maps, respectively. In convolutional layers, the size of all convolutional kernels is $3 \times 3 \times 3$. For all max-pooling layers, the pooling size is $2 \times 2 \times 2$ with stride 2. For all deconvolutional layers, we upsample the input features maps with factor 2. Similar to [14], we concatenate the feature maps after the deconvolutional with those corresponding features in encoder module. After the decoder, a softmax classifier is used to generate voxel-level probability maps and predictions.

Finally, with the high-quality probability maps obtained from the 3D-UNets on PET and CT, a graph cut based co-segmentation model [4] is constructed and optimized to obtain the final segmentation results. The detail procedure and description are referred to [4].

3. EXPERIMENTS

3.1. Experimental Settings

32 co-registered PET-CT scan pairs from different patients with primary non-small cell lung cancer were obtained. The image spacing varies from $0.78 \times 0.78 \times 2 \text{ mm}^3$ to $1.27 \times 1.27 \times 3.4 \text{ mm}^3$. The image slice image size is 512×512 . The number of slices varies from 112 to 293. The tumor contour on each of the PET and CT scans are labeled by two physicians and we adopted the Simultaneous Truth And Performance Level Estimation (STAPLE) algorithm [16] to generate the reference standard for each PET and CT scan.

In our experiments, we first resampled all scans with an isotropic spacing of $1 \times 1 \times 1$ in voxels and then cropped fixed size of 3D volumes ($128 \times 128 \times 64$) centered on the mass gravity of each tumor. Additionally, in order to remove uncorrelated image details, we took similar image intensity value thresholding strategy as that in [8]. All 32 PET-CT scan pairs were split into two sets: 20 for training and 12 for testing. Several simple translation, rotation and flip operations were adopted for data augmentation and the final training set contains 3000 3D PET-CT scan pairs.

The 3D-UNet was implemented using open source TensorFlow package and ran on NVIDIA GeForce GTX 1080 Ti GPU with 11GB of memory. 3D-UNet was trained by Adam optimization method with a mini-batch size of 1 and for 20 epochs. To prevent from

overfitting, the weight decay and early-stop techniques were adopted to obtain the best performance on test set where the Dice coefficient (DSC) was evaluated and reported as adopted in [8]. We conducted quantitative comparisons to Song *et al.*'s [4] and our previous matting-based co-segmentation method [8]. And for the co-segmentation method, the parameters were selected based on the similar strategy as that in [8].

3.2. Results and Analysis

Table 1 reports the mean DSCs and standard deviations for the evaluated methods on the test scans. In addition to co-segmentation, we also report those results from the traditional graph-cut based segmentation methods using their respective manually-crafted region costs [4, 8]. From these quantitative results, we have the following observations. First, compared to the two previous semi-automatic methods [4, 8], the proposed 3D-UNet based co-segmentation approach has achieved significantly better DSCs either with or without using the co-segmentation model on PET and CT scans. This demonstrates that the trained 3D-UNets can learn more descriptive yet discriminative features over the other methods to distinguish between tumor and non-tumor voxels. Second, when incorporated with the co-segmentation model, the performance of all methods can be consistently improved, which also proves the efficiency of co-segmentation by considering the potential label-consistency constraint on dual-modality PET-CT scans. Third, we observe that the segmentation performance on PET is inferior to those on CT. The main reason could be due to the different imaging principles between PET and CT. In general, the tumor boundaries on PET are gradually varied cross a wide image intensity range, and are not obvious as those on CT. This could become a main challenge on PET based medical information extraction. Finally, the proposed method does not need any specific manually-defined object seeds, this advantage over the previous semi-automatic methods can largely help alleviate doctors' burdens and consequently facilitate clinical usages.

Fig. 2 shows the segmentation results of three evaluated methods on two PET-CT scan pairs. Due to the space limit, we only show the contours obtained from co-segmentation. Those example results demonstrated that our 3D-UNets based method can locate tumor boundaries more accurately than the compared methods.

4. CONCLUSION

In this paper, we have developed an integrated image segmentation system that combines the powerful fully convolutional networks (3D UNets) and the graph-cut based co-segmentation method, in which the UNet based FCN can generate high quality voxel-level tumor confidences that were further used to locate the tumor boundary with the powerful co-segmentation model. Experiments on 32 datasets demonstrated the effectiveness of the proposed method with higher performance compared to the state-of-the-art methods.

Acknowledgments

This work was supported in part by the National Science Foundation (NSF) under Grant CCF-1733742, and in part by the the National Institutes of Health (NIH) under Grants R01-EB004640 and 1R21CA209874.

5. REFERENCES

- [1]. Hatt Mathieu, Tixier Florent, Pierce Larry, Kinahan Paul, Le Rest Catherine, and Visvikis Dimitris, "Characterization of PET/CT images using texture analysis: the past, the present... any future?," *European Journal of Nuclear Medicine & Molecular Imaging*, vol. 44, no. 1, pp. 151–166. [PubMed: 27271051]
- [2]. Bagci Ulas, Udupa Jayaram K., Mendhiratta Neil, Foster Brent, Xu Ziyue, Yao Jianhua, Chen Xinjian, and Mollura Daniel J., "Joint segmentation of anatomical and functional images: Applications in quantification of lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT images," *Medical Image Analysis*, vol. 17, no. 8, pp. 929–945, Dec. 2013. [PubMed: 23837967]
- [3]. Han Dongfeng, Bayouth John, Song Qi, Taurani Aakant, Sonka Milan, Buatti John, and Wu Xiaodong, "Globally Optimal Tumor Segmentation in PET-CT Images: A Graph-Based Co-Segmentation Method," *Information processing in medical imaging : proceedings of the ... conference*, vol. 22, pp. 245–256, 2011. [PubMed: 21761661]
- [4]. Song Q, Bai J, Han D, Bhatia S, Sun W, Rockey W, Bayouth JE, Buatti JM, and Wu X, "Optimal Co-Segmentation of Tumor in PET-CT Images With Context Information," *IEEE Transactions on Medical Imaging*, vol. 32, no. 9, pp. 1685–1697, Sept. 2013. [PubMed: 23693127]
- [5]. Li H, Bai J, Abu Hejle T, Wu X, Bhatia S, and Kim Y, "Automated Cosegmentation of Tumor Volume and Metabolic Activity Using PET-CT in Non-Small Cell Lung Cancer (NSCLC)," *International Journal of Radiation Oncology * Biology * Physics*, vol. 87, no. 2, pp. S528, Oct. 2013.
- [6]. Lartizien C, Rogez M, Niaf E, and Ricard F, "Computer-Aided Staging of Lymphoma Patients With FDG PET/CT Imaging Based on Textural Information," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 3, pp. 946–955, May 2014. [PubMed: 24081876]
- [7]. Ju W, Xiang D, Zhang B, Wang L, Kopriva I, and Chen X, "Random Walk and Graph Cut for Co-Segmentation of Lung Tumor on PET-CT Images," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5854–5867, Dec. 2015. [PubMed: 26462198]
- [8]. Zhong Zisha, Kim Yusung, Buatti John, and Wu Xiaodong, "3d Alpha Matting Based Co-segmentation of Tumors on PET-CT Images," in *Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment*, Lecture Notes in Computer Science, pp. 31–42. Springer, Cham, Sept. 2017.
- [9]. Goodfellow Ian, Bengio Yoshua, and Courville Aaron, *Deep Learning*, The MIT Press, Cambridge, Massachusetts, Nov. 2016.
- [10]. Shen Dinggang, Wu Guorong, and Suk Heung-Il, "Deep Learning in Medical Image Analysis," *Annual review of biomedical engineering*, vol. 19, pp. 221–248, June 2017.
- [11]. Litjens Geert, Kooi Thijs, Bejnordi Babak Ehteshami, Setio Arnaud Arindra Adiyoso, Ciompi Francesco, Ghafoorian Mohsen, van der Laak Jeroen A. W. M., van Ginneken Bram, and Sánchez Clara I., "A Survey on Deep Learning in Medical Image Analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017, arXiv: 1702.05747. [PubMed: 28778026]
- [12]. Shelhamer Evan, Long Jonathan, and Darrell Trevor, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017. [PubMed: 27244717]
- [13]. Ronneberger Olaf, Fischer Philipp, and Brox Thomas, "U-Net: Convolutional Networks for Biomedical Image Segmentation," arXiv:1505.04597 [cs], May 2015, arXiv: 1505.04597.
- [14]. Özgün Çiçek Ahmed Abdulkadir, Lienkamp Soeren S., Brox Thomas, and Ronneberger Olaf, "3d U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," in *MICCAI 2016* Oct. 2016, Lecture Notes in Computer Science, pp. 424–432, Springer, Cham.
- [15]. Dou Qi, Chen Hao, Jin Yueming, Yu Lequan, Qin Jing, and Heng Pheng-Ann, "3d Deeply Supervised Network for Automatic Liver Segmentation from CT Volumes," in *MICCAI 2016* Oct. 2016, Lecture Notes in Computer Science, pp. 149–157, Springer, Cham.
- [16]. Warfield SK, Zou KH, and Wells WM, "Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation," *IEEE Transactions on Medical Imaging*, vol. 23, no. 7, pp. 903–921, July 2004. [PubMed: 15250643]

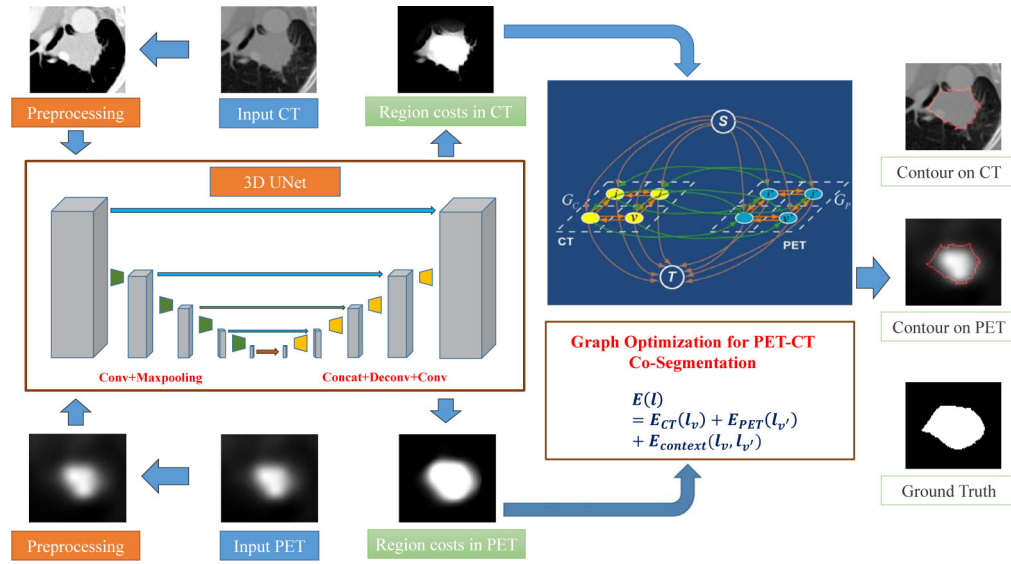


Fig. 1.

The flowchart of proposed 3D-UNet based PET-CT co-segmentation framework. Note that we designed two independent 3D-UNets for PET and CT, respectively, to produce high-quality regions costs for subsequent graph-based co-segmentation.

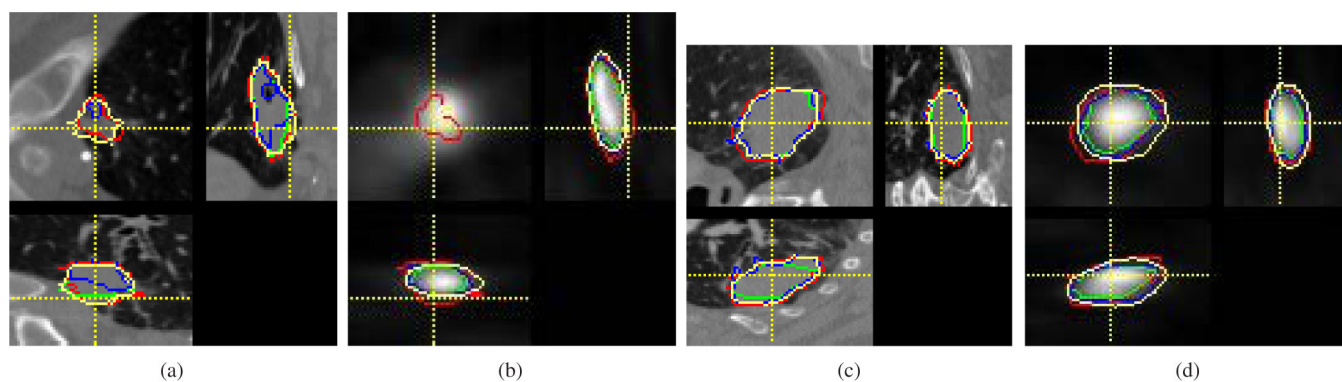


Fig. 2. Segmentation results of compared methods on two PET-CT scans: No. 0002122 and No. 002110. **Red:** ground truth, **Green:** Song's method [4], **Blue:** Zhong's method [8], **Yellow:** Proposed method.

Table 1.

Average DSC's and standard deviations of three compared methods on test set. Bolded numbers mean statistically significant (under $p < 0.5$) over the compared methods.

Methods	Modalities	No-CoSeg	CoSeg
Song et al. [4]	CT	0.577 ± 0.349	0.624 ± 0.240
	PET	0.607 ± 0.151	0.642 ± 0.148
Zhong et al. [8]	CT	0.767 ± 0.108	0.781 ± 0.099
	PET	0.697 ± 0.146	0.722 ± 0.120
Proposed	CT	0.856 ± 0.074	0.869 ± 0.049
	PET	0.757 ± 0.088	0.760 ± 0.088