PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Automated volumetric lung segmentation of thoracic CT images using fully convolutional neural network

Mohammadreza Negahdar, David Beymer, Tanveer Syeda-Mahmood

Mohammadreza Negahdar, David Beymer, Tanveer Syeda-Mahmood, "Automated volumetric lung segmentation of thoracic CT images using fully convolutional neural network," Proc. SPIE 10575, Medical Imaging 2018: Computer-Aided Diagnosis, 105751J (27 February 2018); doi: 10.1117/12.2293723

SPIE.

Event: SPIE Medical Imaging, 2018, Houston, Texas, United States

Automated Volumetric Lung Segmentation of Thoracic CT Images using Fully Convolutional Neural Network

Mohammadreza Negahdar, David Beymer, Tanveer Syeda-Mahmood
IBM Research-Almaden, San Jose, CA, USA

ABSTRACT

Deep Learning models such as Convolutional Neural Networks (CNNs) have achieved state-of-the-art performance in 2D medical image analysis. In clinical practice; however, most analyzed and acquired medical data are formed of 3D volumes. In this paper, we present a fast and efficient 3D lung segmentation method based on V-net: a purely volumetric fully CNN. Our model is trained on chest CT images through volume to volume learning, which palliates overfitting problem on limited number of annotated training data. Adopting a pre-processing step and training an objective function based on Dice coefficient addresses the imbalance between the number of lung voxels against that of background. We have leveraged V-net model by using batch normalization for training which enables us to use higher learning rate and accelerates the training of the model. To address the inadequacy of training data and obtain better robustness, we augment the data applying random linear and non-linear transformations. Experimental results on two challenging medical image data show that our proposed method achieved competitive result with a much faster speed.

Keywords: Deep Learning, Convolutional Neural Network, Volumetric, Lung, Segmentation

1. DESCRIPTION OF PURPOSE

The accurate volumetric segmentation of organs or structures in medical images allows the subsequent quantitative analysis of clinical parameters related to volume and shape. Furthermore, it is an important first step in visual augmentation, treatment planning and computer assisted diagnosis pipelines.

The capabilities of Convolutional neural networks (CNNs) to solve segmentation problem and achieving state-of-the-art performance through their hierarchically learned highly representative features, have revolutionized the medical image analysis domain. As the image is processed through the network layers, shallower layers capture local information and deeper layers capture more global information [2]. The most well-known, in medical image analysis, of novel CNN architectures is U-net which proposed the combination of an equal amount of upsampling and downsampling layers [3]. Upsampling layers are combined with skip connections between opposing convolution and deconvolution layers, where concatenate features from the contracting and expanding paths. From a training perspective this means that entire images can be processed by U-net in one forward pass, resulting in a segmentation map directly. This allows U-net to take into account the full context of the image. However, depth information is ignored by 2D CNNs.

Although CNNs have attained impressive performance in 2D medical image segmentation [3, 4], 3D segmentation still remains a challenge. In fact, more complicated anatomical environments of 3D medical images translates to CNNs with more parameters to capture representative features. In 3D U-net, it is shown that a full 3D segmentation can be achieved by feeding U-net with a few 2D annotated slices from the same volume [5]. However, solving a 3D segmentation problem through slice-by-slice fashion and applying 2D CNN segmentation will ignore the sequential information between consecutive slices and cannot take full advantage of the special information encoded in the volumetric data [6].

A 3D-variant of U-net architecture, called V-net, proposed performing 3D image segmentation using 3D convolutional layers with an objective function directly based on the Dice coefficient [1]. Although 3D convolution exploits different views of a volume data, it require a larger number of parameters and higher computational load. However, this network is primarily trained to process MRI volumes. Furthermore; as a known caveat for training a 3D CNN, it comes with various optimization challenges such as over-fitting, gradient vanishing and slow convergence speed.

Medical Imaging 2018: Computer-Aided Diagnosis, edited by Nicholas Petrick, Kensaku Mori, Proc. of SPIE Vol. 10575, 105751J · © 2018 SPIE · CCC code: 1605-7422/18/$18 · doi: 10.1117/12.2293723

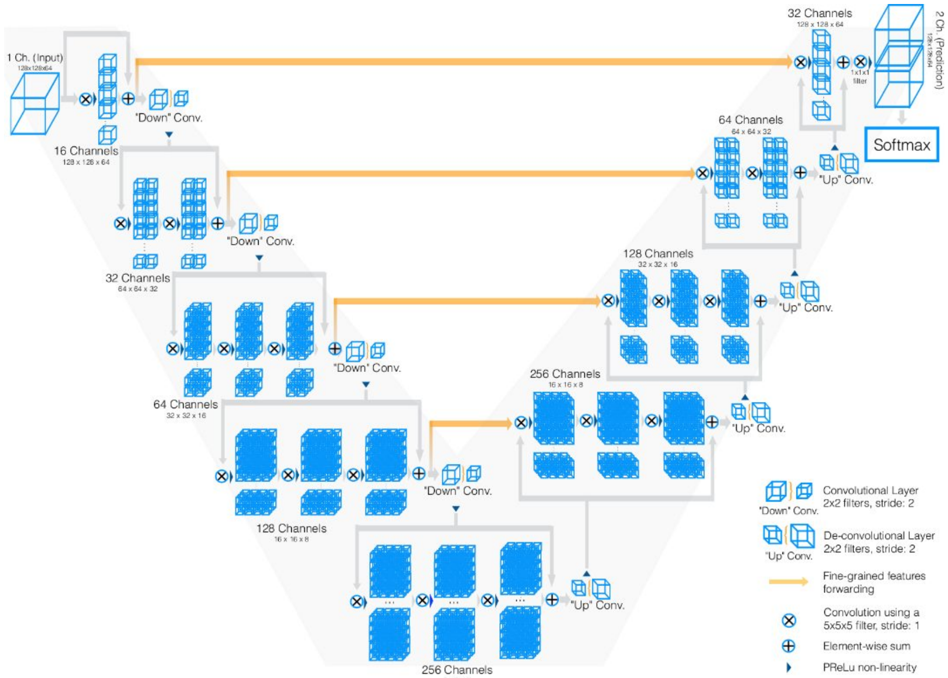

Figure 1. V-net architecture as proposed in [1].

We aim to adopt V-net model, leverage it with the information of standard protocol of lung image acquisition, and train end-to-end to segment 3D Lung in thoracic CT images. Our method employs batch normalization for training which enables us to use higher learning rate and accelerates the training of the model.

2. METHODS

Dataset: For training and testing the proposed method, we used dataset of EMPIRE10 [7], a thoracic CT registration challenge in MICCAI 2010. The dataset consists of 30 pairs of thoracic CT scans. Each pair of scans is taken from a single subject either in inhale and exhale respiratory phases or two phases from 4D CT. Subjects may suffer from lung diseases or appear healthy. Data from a variety of scanners is included and a variety of voxel sizes occur. Binary lung masks are provided for each scan. The trained model has also been tested on dataset of VESSEL12, a lung blood vessels segmentation challenge in ISBI 2012. The dataset consists of 23 thoracic CT scans from a variety of sources and represent a variety of clinically common scanners and protocols. About half of the scans contain abnormalities such as emphysema, nodules or pulmonary embolisms.

Thoracic CT images are acquired using a relatively static imaging protocol, where lung is approximately located in the same position and same scale. Lung usually occupies less than a quarter of the acquired CT images. Hence, we apply an intensity-based segmentation on our CT images using simple-ITK in order to locate the lung bounding box. Then, CT images and their associated labels are cropped such that regions outside the lung volume are excluded where possible. Hereby, we take advantage of the imbalance between lung voxels and background, which in turn alleviates memory insufficiency problem.

2.1 Architecture

We base our algorithm on V-net architecture [1] which originally presented to perform 3D segmentation of prostate in MR images. V-net uses 3D convolutional layers for efficient volume to volume learning through their deep and large receptive field (Figure 1). Using 3D CNN lets encode representations from volumetric receptive fields, and therefore extracting

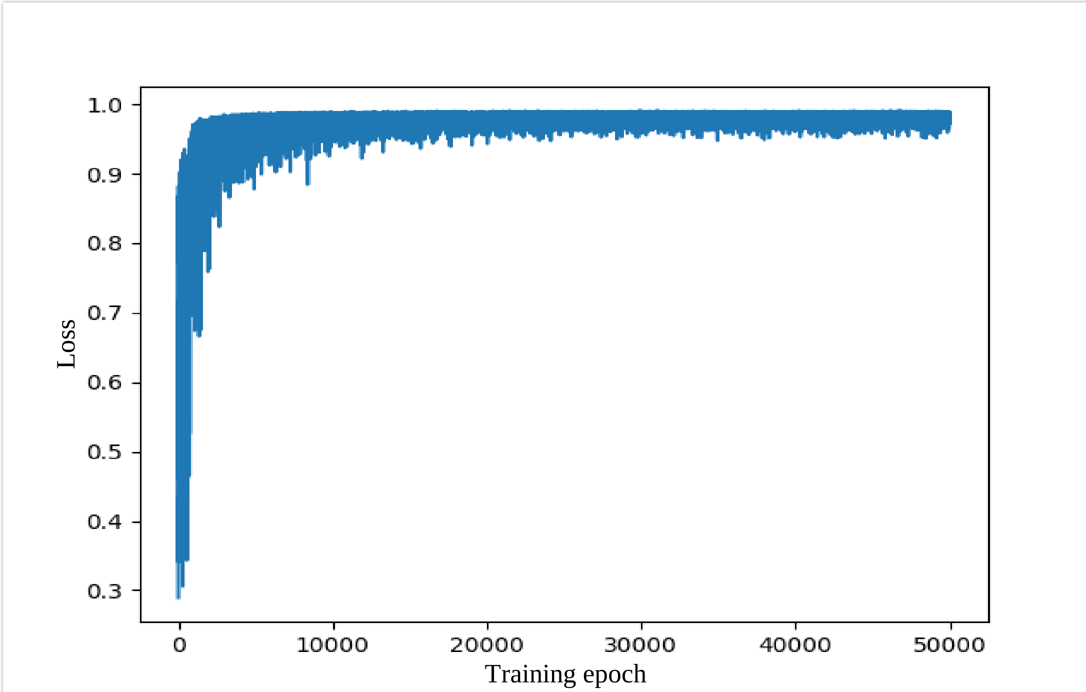

Figure 2. Learning curve of the proposed model for training on volumetric thoracic CT images.

more discriminative features via richer 3D spatial information. Nonetheless adopting a model for a specific problem is not trivial, since a large number of mutually dependent parameter values and algorithmic choices have to be chosen.

We leveraged the V-net model by using batch normalization for training which enables us to use higher learning rate and accelerates the training of the model. Furthermore, the proposed end-to-end network is economical regarding to storage consumption, because of the cropping step.

2.2 Training and Testing

We randomly picked %80 of EMPIRE10 scans (48 CT scans) for training and used the rest (12 CT scans) for testing our algorithm. We reserve whole VESSEL12 dataset for testing the trained model. To address the limited number of training dataset and obtain better robustness, the dataset is augmented applying a random affine transform and a randomly deformed version of each scan by using a deformation field obtained through a 2x2x2 grid of control-points and B-spline interpolation. We used a fixed learning rate 0.0001 and a batch size of 2 to avoid memory problem.

Segmented lung is compared against the provided mask for all test datasets using Dice coefficient, where 1 represents perfect agreement and 0 represents no overlap:

$$D = \frac{2\,|X \cap Y|}{|X|+|Y|} \tag{1}$$

where X and Y are the ground truth and test regions, respectively.

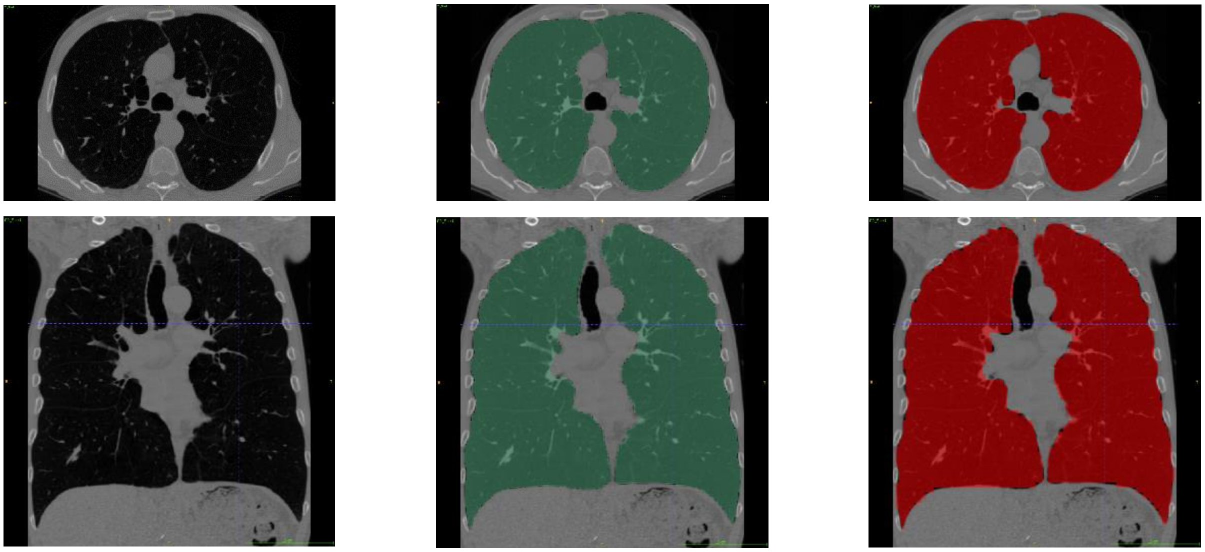

Figure 3. Qualitative comparison: Mid-axial (top) and mid-coronal (bottom) planes of the original CT image (left), ground truth (middle), and segmented lung with the proposed deep learning model (right).

3. RESULTS

First, we analyze the learning process of the proposed model which is based on Dice loss layer that directly minimizes this commonly used segmentation error measure. As shown in Figure 2, the validation loss consistently increases and converges notably fast. Our training scheme and data augmentation allows us to train the large V-net on a relatively small dataset without considerable overfitting. Our model is trained in 50k iterations.

We evaluated the trained model on the EMPIRE10 testing dataset and the average dice coefficient was 0.983±0.002. Figure 3 shows qualitative comparison of lung segmentation result against the provided ground truth. Segmentation of the shown exemplary 3D volume in Figure 3, (420×312×537 voxels) took less than 8s on a NVIDIA GeForce GTX Titan X GPU with 12 GB of memory.

Next, we used 23 thoracic CT scans of VESSEL12 for testing the trained model and achieved dice coefficient was 0.990±0.002 as segmentation accuracy.

Our trained model outperforms classic intensity-based algorithms in lung segmentation problem in low dose and low quality CT images. Figure 4 shows the performance of the trained model in segmentation of a low quality thoracic CT image of the Lung image Database Consortium (LIDC) [8].

4. CONCLUSIONS

In this paper, we proposed a fast and accurate end-to-end 3D lung segmentation method using deep learning which exploits both local features and more global contextual features simultaneously. Our proposed model is based on a modified version of V-net, where 3D CNN and using lose function based on dice coefficient yields more accurate results for volume to volume learning. Adopting batch normalization for training enables us to choose higher learning rate and accelerates the training of the proposed model.

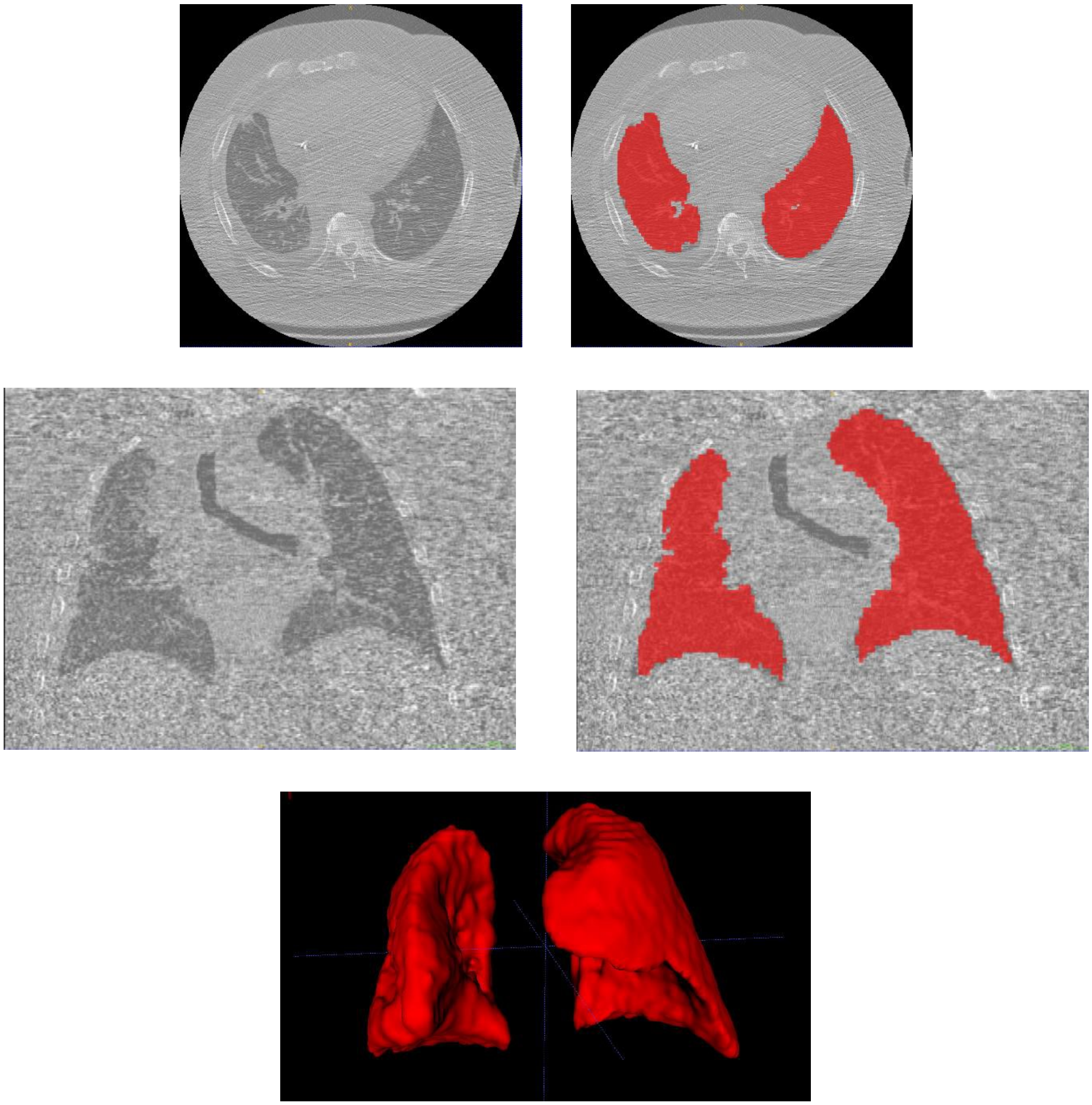

Figure 4. Performance of the proposed model in segmentation of LIDC CT image: axial (top) and coronal (middle) planes of the CT image (left) and segmented lung with the proposed deep learning model (right). Volumetric rendering of the segmented lung (bottom).

REFERENCES

[1] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation."3D Vision (3DV), 2016 Fourth International Conference on, 565-571 (2016).

[2] M. D. Zeiler, and R. Fergus, "Visualizing and understanding convolutional networks."European conference on computer vision, 818-833 (2014).

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation."International Conference on Medical Image Computing and Computer-Assisted Intervention, 9351, 234-241 (2015).

[4] F. Xing, Y. Xie, and L. Yang, "An Automatic Learning-Based Framework for Robust Nucleus Segmentation," IEEE Transactions on Medical Imaging, 35(2), 550-566 (2016).

[5] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp *et al.*, "3D U-Net: learning dense volumetric segmentation from sparse annotation."International Conference on Medical Image Computing and Computer-Assisted Intervention, 9901, 424-432 (2016).

[6] P. Moeskops, J. M. Wolterink, B. H. M. van der Velden *et al.*, [Deep Learning for Multi-task Medical Image Segmentation in Multiple Modalities] Springer International Publishing, Cham(2016).

[7] K. Murphy, B. van Ginneken, J. M. Reinhardt *et al.*, "Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge," IEEE Trans Med Imaging, 30(11), 1901-20 (2011).

[8] S. G. Armato, 3rd, G. McLennan, L. Bidaut *et al.*, "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans," Med Phys, 38(2), 915-31 (2011).