Medical Image Analysis 40 (2017) 172–183

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

# Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation

Shuo Wang[a,c,1], Mu Zhou[b,1], Zaiyi Liu[d], Zhenyu Liu[a], Dongsheng Gu[a,c], Yali Zang[a,c], Di Dong[a,c,1], Olivier Gevaert[b,1], Jie Tian[a,c,e,*]

[a] CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] Stanford Center for Biomedical Informatics Research (BMIR), Department of Medicine, Stanford University, CA 94305, USA
[c] University of Chinese Academy of Sciences, Beijing 100049, China
[d] Guangdong General Hospital, Guangzhou, Guangdong 510080, China
[e] Beijing Key Laboratory of Molecular Imaging, Beijing 100190, China

ARTICLE INFO

*Article history:*
Received 9 January 2017
Revised 29 June 2017
Accepted 29 June 2017
Available online 30 June 2017

*Keywords:*
Lung nodule segmentation
Convolutional neural networks
Deep learning
Computer-aided diagnosis

ABSTRACT

Accurate lung nodule segmentation from computed tomography (CT) images is of great importance for image-driven lung cancer analysis. However, the heterogeneity of lung nodules and the presence of similar visual characteristics between nodules and their surroundings make it difficult for robust nodule segmentation. In this study, we propose a data-driven model, termed the Central Focused Convolutional Neural Networks (CF-CNN), to segment lung nodules from heterogeneous CT images. Our approach combines two key insights: 1) the proposed model captures a diverse set of nodule-sensitive features from both 3-D and 2-D CT images simultaneously; 2) when classifying an image voxel, the effects of its neighbor voxels can vary according to their spatial locations. We describe this phenomenon by proposing a novel central pooling layer retaining much information on voxel patch center, followed by a multi-scale patch learning strategy. Moreover, we design a weighted sampling to facilitate the model training, where training samples are selected according to their degree of segmentation difficulty. The proposed method has been extensively evaluated on the public LIDC dataset including 893 nodules and an independent dataset with 74 nodules from Guangdong General Hospital (GDGH). We showed that CF-CNN achieved superior segmentation performance with average dice scores of 82.15% and 80.02% for the two datasets respectively. Moreover, we compared our results with the inter-radiologists consistency on LIDC dataset, showing a difference in average dice score of only 1.98%.



## 1. Introduction

Lung cancer is the leading cause for cancer related deaths and carrying a dismal prognosis with a 5-year survival rate at only 18% (Siegel et al., 2016). Treatment therapy monitoring and lung nodule analysis (Aerts et al., 2014) using computed tomography (CT) images are important strategies for early lung cancer diagnosis and survival time improvement. In these approaches, accurate lung nodule segmentation is necessary that can directly affect the subsequent analysis results. Specifically, given the fact of growing volumes of clinical imaging data, developing a data-driven segmentation model is of great clinical importance to avoid tedious manual processing and reduce inter-observer variability (Kubota et al., 2011).

Despite development of approaches for lung nodule segmentation in recent years (Farag et al., 2013; Kubota et al., 2011; Lassen et al., 2015), achieving accurate segmentation performance continues to require attention because of the heterogeneity of lung nodules as shown on CT images (Fig. 1). The presence of similar visual characteristics between nodules and their surroundings poses a technical challenge for developing robust segmentation models. For example, juxtapleural nodules (Fig. 1(b)) have an intensity similar to that of lung wall; thus, they are difficult to distinguish using intensity values only. In addition, cavitary nodules with black hole inside (Fig. 1(c)) and calcific nodules (Fig. 1(d)) are challenging cases because of the intensity dissimilarity within different part of nodules. Similarly, non-solid nodules such as ground-glass opacity (GGO, Fig. 1(e)) are also problematic because a simple morpholog-

* Corresponding author.
*E-mail addresses:* di.dong@ia.ac.cn (D. Dong), olivier.gevaert@stanford.edu (O. Gevaert), jie.tian@ia.ac.cn, tian@ieee.org (J. Tian).
[1] Contributed equally.

http://dx.doi.org/10.1016/j.media.2017.06.014

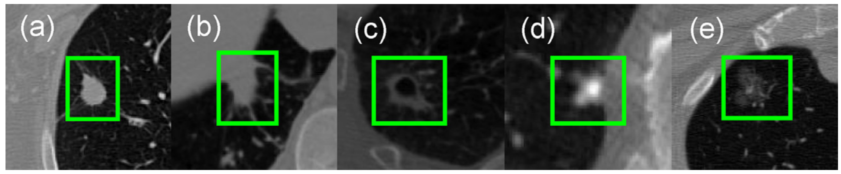


**Fig. 1.** Example images of lung nodules with different locations and shapes in CT: (a) common isolated nodule. (b) juxtapleural nodule. (c) cavitary nodule. (d) calcific nodule. (e) ground-glass opacity (GGO) nodule.

ical operation is not suitable for these cases due to the fact of low intensity contrast in CT data (Dehmeshki et al., 2008).

Intensity-based methods using morphological operation (Diciotti et al., 2011; Messay et al., 2010) and region growing (Dehmeshki et al., 2008; Kubota et al., 2011) have been studied. Energy optimization methods including level set (Farag et al., 2013) and graph cut (Ye et al., 2010) were also researched for lung nodule segmentation. However, the robustness is still problematic especially for segmenting juxtapleural nodules. For example, in morphology-based methods, the morphological template size is difficult to generalize with nodules of various diameters (Kubota et al., 2011). Sophisticated methods can process juxtapleural nodules by applying a shape constraint (Farag et al., 2013; Keshani et al., 2013) or relying on user interactive parameter settings (Messay et al., 2015). However, it may not be applicable for irregular shaped nodules where the shape hypothesis can be violated. In addition, user interactive parameters such as well centralized seed point (Messay et al., 2015) or stroke (Lassen et al., 2015) are difficult to tune for different types of nodules. The limitations of directly applying raw intensity value for segmentation suggest the need of novel solutions for capturing high-level, nodule-sensitive features from CT volumes.

Recently, convolutional neural networks (CNN) have been emerged as powerful tools for learning discriminative feature hierarchies adapted to different vision tasks (Havaei et al., 2016; Shen et al., 2016). Benefiting from the unique feature learning ability from hierarchical network layers, CNN models have shown encouraging results in medical image segmentation tasks (Moeskops et al., 2016; Valverde et al., 2017; Zhang et al., 2015), indicating the usefulness of CNN-based models for medical object segmentation. However, the applicability of developing CNN-based approaches to model heterogeneous lung nodule CT volumes (as seen in Fig. 1) has remained uncertain. In particular, the design of network hierarchy that is capable of capturing both 2-D and 3-D lung nodule features has not been explicitly addressed.

In this study, we investigate the problem of developing a deep hierarchy of convolutional neural networks in the context of lung nodule segmentation. We follow a voxel classification scheme that aims to distinguish nodule voxels from healthy voxels in CT images. In addressing the challenge of analyzing heterogeneous CT data, we propose a central focused convolutional neural networks (CF-CNN) that is adaptive to lung nodule segmentation for various types of nodules. Overall, our technical contributions in this work are four-fold:

1. The proposed CF-CNN model can achieve appealing segmentation performance for a variety of lung nodules especially for juxtapleural nodules without nodule shape hypothesis or user-interactive parameter setting (Fig. 1).
2. We present a two-branch CNN structure to leverage both 3-D features and multi-scale 2-D features. The 3-D-patch branch learns multi-view features from multiple CT slices and the 2-D-patch branch learns multi-scale features through multiple 2-D patches. The multi-scale patch strategy enables the model to learn multi-scale features without involving multiple networks (Shen et al., 2015) (Section 2.1.2).
3. We design a novel central pooling layer to retain much patch-center features rather than patch edge features. This strategy reserves much target-voxel-focused information and thereby achieved improved performance as opposed to uniformly distributed max pooling (Section 2.1.3).
4. During model training, we propose a sampling method to process imbalanced training labels and extract challenging patches to allow efficient model training. In this strategy, voxels are sampled where each voxel is assigned a weight score denoting its difficulty for segmentation (Section 2.3).

### 1.1. Related work

Approaches for lung nodule segmentation involved the detection of a Volume of Interest (VOI) covering the nodule area and segmentation inside this VOI. These methods can be generally classified into morphology methods (Diciotti et al., 2011; Messay et al., 2010), region growing methods, (Kubota et al., 2011; Song et al., 2016), energy optimization methods (Farag et al., 2013; Lassen et al., 2015), and machine-learning methods (Lu et al., 2013; Wu et al., 2010).

In morphology methods, morphological operations such as logic opening operation were applied for nodule-attached vessels removal (Kostis et al., 2003), then the connected component selection can separate lung nodules. However, the fixed-size morphological template is difficult to separate nodules that usually have wide contact surfaces with other anatomical objects. Consequently, more complex morphological operations that combine shape hypothesis were introduced. For instance, Kuhnigk et al. (2006) showed that the radius of vessels decreases while the vessels evolve along the periphery of the lungs. In addition, rolling ball filters (Messay et al., 2010) combined with rule-based analysis was also proposed for juxtapleural nodules. One notable difficulty for morphology methods is the morphological template size selection (Kubota et al., 2011), because it is difficult to find a suitable morphology template for various size of nodules. Non-solid nodules in particular are challenging for morphology operation (Diciotti et al., 2011).

In region growing methods, segmentation starts with a user-specified seed point, and voxels are included into nodule set iteratively until the pre-defined converge criterion is satisfied. These methods work well for isolated nodules. However, when analyzing juxtapleural nodules, region growing algorithm is known to be difficult to converge. Therefore, Dehmeshki et al. (2008) introduced a shape hypothesis and proposed sphericity contrast based region growing method to detach nodule from lung wall. Instead of using the current voxel intensity only, Kubota et al. (2011) constructed a probability map to denote the likelihood of each voxel belonging to nodule according to the local intensity value, then a region growing method was used to separate the nodule from background area. The common challenge for region growing methods is the converge criteria. Although shape constraint can be considered, irregular-shaped nodules remain difficult to process because the shape hypothesis can be violated.

In energy optimization methods, nodule segmentation is converted into an energy minimization task. The level-set-based meth-