Computerized Medical Imaging and Graphics 74 (2019) 25–36

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

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

Fast and fully-automated detection and segmentation of pulmonary nodules in thoracic CT scans using deep convolutional neural networks

Xia Huang[a], Wenqing Sun[b], Tzu-Liang (Bill) Tseng[c], Chunqiang Li[d,**,1], Wei Qian[b,*,2]

[a] Department of Metallurgical, Materials, and Biomedical Engineering, University of Texas at El Paso, El Paso, TX, 79968, United States
[b] Department of Electrical and Computer Engineering, University of Texas at El Paso, El Paso, TX, 79968, United States
[c] Department of Industrial, Manufacturing, and Systems Engineering, University of Texas at El Paso, El Paso, TX, 79968, United States
[d] Department of Physics, University of Texas at El Paso, El Paso, TX, 79968, United States

ARTICLE INFO

Article history:
Received 20 September 2018
Received in revised form 9 January 2019
Accepted 18 February 2019

Keywords:
Computer aided diagnosis
Pulmonary nodule detection and segmentation
Convolutional neural networks
Faster regional-CNN
Fully convolutional neural network (FCN)

ABSTRACT

Deep learning techniques have been extensively used in computerized pulmonary nodule analysis in recent years. Many reported studies still utilized hybrid methods for diagnosis, in which convolutional neural networks (CNNs) are used only as one part of the pipeline, and the whole system still needs either traditional image processing modules or human intervention to obtain final results. In this paper, we introduced a fast and fully-automated end-to-end system that can efficiently segment precise lung nodule contours from raw thoracic CT scans. Our proposed system has four major modules: candidate nodule detection with Faster regional-CNN (R-CNN), candidate merging, false positive (FP) reduction with CNN, and nodule segmentation with customized fully convolutional neural network (FCN). The entire system has no human interaction or database specific design. The average runtime is about 16 s per scan on a standard workstation. The nodule detection accuracy is 91.4% and 94.6% with an average of 1 and 4 false positives (FPs) per scan. The average dice coefficient of nodule segmentation compared to the groundtruth is 0.793.

1. Introduction

Worldwide, lung cancer has been having the leading mortality rate of cancer deaths in both males and females for decades with 1.2 million global deaths a year (Fitzmaurice et al., 2017). Due to the inconspicuous symptoms, the majority of lung cancer cases are diagnosed at distant stages with only 4% five-year survival rate (Siegel et al., 2018). Early detection of suspicious pulmonary nodules is crucial to improve the life quality of lung cancer patients. Currently, computed tomography (CT) is considered the best and most widely used imaging modality for early detection and analysis of lung nodules. However, because of the complicated morphological and anatomical appearance of nodules, the nodule identification would be largely dependent on the skill, experience, and vigor of the radiologists (Winkels and Cohen, 2018). After identifying the nodule, precise segmentation is significant for clinical measurements (such as diameter and volume), which objectively provides repeatability of diagnosis and consistency of image interpretation (Liu et al., 2018). Therefore, a fast and fully-automated computer aided detection (CAD) system on nodule detection and segmentation with limited number of false positives (FPs) will dramatically decrease the workload of radiologists as well as the cost of treatment.

For nodule detection many published works proposed a two-stage system, which includes a candidate screening step to rapidly extract nodule regions from pulmonary parenchyma and remove other structures, and a false positive (FP) reduction step to massively eliminate FP candidates from the detected ones until reaching clinically acceptable performance. Ge et al. (2005) used adaptive weighted k-means clustering to segment suspicious candidates, and reduced FPs by combing 3D gradient field and ellipsoid features with a linear discriminant analysis (LDA) classifier. Li et al. (2008) added a multiscale selective filter to enhance nodule and

---

[*] Corresponding author at: Department of Electrical and Computer Engineering, Director: Medical Imaging Informatics Laboratory, University of Texas at El Paso, 500 West University Avenue, El Paso, TX, 79968, United States.

[**] Corresponding author at: Department of Physics, University of Texas at El Paso, 500 West University Avenue, El Paso, TX, 79968, United States.

E-mail addresses: cli@utep.edu (C. Li), wqian@utep.edu (W. Qian).

[1] URL: https://science.utep.edu/photonics/.

[2] URL: http://ee.utep.edu/facultyqian.htm.

https://doi.org/10.1016/j.compmedimag.2019.02.003
0895-6111/

simultaneously suppress normal structures, and then a rule-based classifier was used to reduce FPs based on six shape features and twelve intensity features extracted from the enhanced images. Tan et al. (2011) improved the performance of candidate screening by introducing a maxima of the divergence of the normalized gradient (DNG) to find centers of nodule candidates with a merging stage to remove duplicates and further reduce the number of FPs. Forty-five invariant features, defined on a gauge coordinates system, are used to differentiate nodules from large amount of FPs. Finally, a novel feature-selective classifier based on genetic algorithms and artificial neural networks (FD-NEAT) was first implemented to improve the flexibility and adaptability of a classifier. Its performance was compared with that of two other classifiers based on support vector machines (SVMs) and fixed-topology artificial neural networks (ANNs). Even though these traditional machine learning algorithms achieved remarkable accuracy in nodule detection, some disadvantages including but not limited to arduous human interventions, slow computation time, and mediocre representation capability of hand-crafted features obstruct the further development of traditional CAD system to deal with the large variations of lung nodules from real clinical CT scans. Recently, deep learning techniques in particular convolutional neural networks (CNNs) motivate flurry of researchers to develop powerful and robust algorithms on pulmonary nodule detection, which outperform many traditional machine learning approaches (Anirudh et al., 2016; Tajbakhsh and Suzuki, 2017; Sun et al., 2017a; Ypsilantis and Montana, 2016; Fu et al., 2017; Hamidian et al., 2017; Winkels and Cohen, 2018; Zhu et al., 2018). Setio et al. (2016) delineated multi-view 2D CNNs by taking advantage of nine symmetric planes of nodule cubes without increasing the network complexity. It is fed with nodule candidates obtained by three individual candidate screening algorithms that are exclusively designed for solid, sub-solid, and large nodules respectively. The best detection accuracy is achieved by applying a dedicated mix-fusion method. Inspired by the 3D nature of pulmonary nodules, Huang et al. (2017) exploited a single-scale 3D CNNs to encode much richer and more comprehensive spatial contextual information compared with conventional 2D CNNs. Dou et al. (2017a) significantly boosted the detection accuracy through a multi-level 3D CNNs. In order to cover the large variations of nodules with different sizes of receptive fields, three independent 3D CNNs were involved to learn discriminative features for small, medium, and large size of nodules, respectively. It is a generic 3D CNNs framework that can in principle transfer to other applications to extract targets from variety of complicated mimics. Dou et al. (2017b) proposed a novel two-stage 3D CNN for end-to-end nodule detection with a 3D FCN based nodule candidates screening and a 3D hybrid-loss residual learning based FP reduction. They first tackled the severe imbalance problem of hard and easy samples by employing an online sample filtering scheme. This dynamic scheme naturally splits hard and easy samples based on the loss of each forward propagation of training, so that the training convergence can be fastened. Ding et al. (2017) provided a combination of a 2D Faster R-CNN for the initial nodule candidate detection and a 3D CNN for FP reduction. The Faster R-CNN based nodule detection optimally ensured the sensitivity while maintaining low number of FPs. Jin et al. (2018) constructed a 27-layer 3D residual CNNs, which is much deeper and more effective than the traditional 3D CNNs. A spatial pooling and cropping (SPC) layer ensures the capability of learning multi-level contextual information using a single-scale 3D CNNs architecture. Such design overcomes the restriction of tedious parameter tuning while dealing with model fusion, and drastically accelerates the training and testing process. Moreover, an online hard sample selection (OHSS) unlocks the potential of network to detect extreme nodules with complex morphological characteristics. However, due to the nature of the prevalent two-stage nodule detection framework, some unignorable drawbacks still occur. First, the candidate screening, which should ideally detect all the suspicious nodule candidates, determines the upper-bound sensitivity of the entire CAD system. But morphological difference of nodules makes it impossible to achieve the optimal performance based on single or multiple hand-crafted mathematical models (Murphy et al., 2009; Jacobs et al., 2014; Setio et al., 2015), and the tedious experiment-based parameter adjustment restricts the applications onto real clinical trials. Moreover, the suboptimal segmentation of lung parenchyma negatively impacts the candidate screening especially for juxtapleural nodules (Dai et al., 2015). Second, because of the serial algorithm structure with many subcomponents, the long computation time stands out as another demerit. Therefore, a simpler and more independent nodule detection framework is urgently desired.

Because of the critical clinical value in nodule segmentation, a growing numbers of pulmonary nodule segmentation algorithms have been proposed in literature. They can be roughly categorized into four types: 1) Threshold based methods (Reeves et al., 2006; Magalhães Barros Netto et al., 2012; Tachibana and Kido, 2006 Xia et al., 2016). For instance, Tachibana and Kido (2006) designed a coarse-to-fine scheme that consists of a rough segmentation step using multiple fixed thresholds to roughly identify the nodule regions and a precise segmentation step using a watershed-based algorithm to remove the unnecessary structures attached to them. 2) Morphology based methods (Kubota et al., 2011; Dehmeshki et al., 2008; Vijaya Kishore and Satyanarayana, 2013 Lassen et al., 2015). For example, Dehmeshki et al. (2008) presented an efficient sphericity-oriented region growing algorithm applied on the fuzzy connectivity mask created by a connectivity region growing technique with only one single seed point provided by the user. 3) Statistical model based method (Wang et al., 2009; Dong and Peng, 2014; Tan et al., 2013; Mao et al., 2018). Tan et al. (2013) utilized a hybrid algorithm combining marker-controlled watershed, geometric active contours as well as Markov random field (MRF). Similar to the method Tachibana et al. proposed, they imposed watershed method to generate an initial surface of nodule, followed by the refinement of active contours. And MRF optimally estimates the texture distribution of ground glass opacity, so that it improves the segmentation accuracy for this portion. 4) Clustering methods based on traditional machine learning (Van Ginneken, 2006; Tuinstra, 2008; Messay et al., 2015). Messay et al. (2015) proposed a selective regression neural network (RNN) based algorithm with both fully-automated and semi-automated options. The feature learning process using RNN can automate the parameter setting for each nodule based on the learned features. However, the majority of the aforementioned methods perform well only on specific type of nodules (e.g. solitary pulmonary nodule) or on relatively small size of dataset, which cannot satisfy the variety and complexity of pulmonary nodules. In addition, most of the methods still need human interventions, which largely undermine the purpose of CAD systems. Finally, in order to achieve optimal performance, most of the techniques require massive iterations and parameter tunings, which to a large extent, slow down the overall computation process. Recently, the success of semantic segmentation in computer vision field based on fully convolutional neural networks (FCNs) (Long et al., 2015; Wang et al., 2017; Chen et al., 2018; Lekić and Babić, 2018 Yu et al., 2018) attracts some researchers to concentrate on the application of pulmonary nodule segmentation. Wu et al. (2018) firstly deployed an interpretable and multi-task CNNs model to segment and classify pulmonary nodules by feeding 3D patches and achieved the state-of-the-art performance. However, most of current deep learning based algorithms still rely on several preprocessing steps such as lung parenchyma segmentation, which decrease the level of automation.

In this paper, we propose a CNNs based algorithm to automatically detect and segment pulmonary nodules with very limited

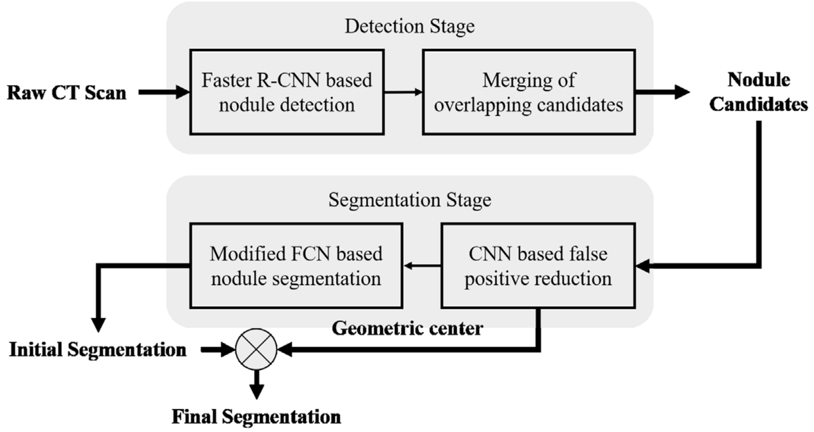


Fig. 1. Top level framework of proposed CNNs based nodule detection and segmentation system.

number of FPs. Comparing with methods using the original nodule patch only, our method improves the area under the curve (AUC) malignancy prediction by 4% with the combination of the original nodule patch, the segmented nodule patch highlighting the shape, and the gradient of nodule patch highlighting the texture with accurate boundary information provided (Sun et al., 2017b). To our best knowledge, this is one of the first works that exploit pure 2D CNNs based algorithm for pulmonary nodule segmentation from the raw CT scan without any manual settings. The output is the corresponding nodule masks with high overlapping ratio compared to radiologists' markings. This study is the follow up study of our previous researches (Qian et al., 1993; DeVore et al., 1995; Sun et al., 2004; Zhang et al., 2007; Ye et al., 2013; Sun et al., 2017a,b, 2016, 2017c, 2017d, 2017e). The reason of using 2D axial slice instead of 3D volume in our system is three-fold. First, a powerful deep learning model relies on large amount of training data. However, due to the stressful workload of manual nodule identification, the available thoracic CT scans with high quality groundtruth annotations are still insufficient to train a robust and discriminative deep learning network. Since a nodule may appear in several neighboring 2D slices of a single CT scan, using 2D instead of 3D slices will naturally augment the size of training set without manual data augmentation process. Second, 3D deep learning models will exponentially add computation complexity compare to 2D models. Therefore, some models require image down sampling or cropping (Jin et al., 2018) to compensate the massive memory consumption. This requirement, to some extent, restricts the feasibility of 3D models on common workstation with limited GPU resources. Third, transfer learning (Weiss et al., 2016; Long et al., 2016) can be performed based on some fine-tuned 2D models such as VGG16 (Simonyan and Zisserman, 2015), ResNet-50, and ResNet-101 (He et al., 2016), significantly accelerates the convergence of training, boosts the performance, and cuts down the computation complexity. The remaining paper is organized as follows. The data and methodologies will be described in Section 2. Section 3 reports the experimental results with qualitative and quantitative performance evaluation, execution performance analysis, and comparison between our proposed method and some state-of-the-art approaches to underline the significance of our study. Conclusions and some potential improvements are stated in Section 4.

2. Materials and methods

2.1. Dataset

Similar to most of aforementioned literatures, we train and evaluate our fully-automated system on a large publicly available dataset, organized by Lung Image Database Consortium image collection (LIDC-IDRI) (Armato et al., 2015, 2011; Clark et al., 2013) with 1018 scans from seven academic centers and eight medical imaging companies. The slice thickness of these scans is ranging from 0.6 mm to 5.0 mm. For each scan, four experienced radiologists performed two-phase nodule assessment and recorded the detailed nodule information such as boundary coordinates, malignancy level etc. into an XML file. In this data set three types of lesions are included: non-nodule, small nodule (<3 mm), and large nodule (≥3 mm). According to some clinical recommendations (Aberle et al., 2011), we only consider large nodules (≥3 mm) in our study. The pre-filtering strategy proposed in the Lung Nodule Analysis 2016 (LUNA16) challenge only selected nodules with the consensus of at least three out of four radiologists with slice thickness no more than 2.5 mm. As a consequence, 888 CT scans with 1186 nodules are involved.

2.2. Methodology

The framework of our proposed nodule segmentation algorithm is shown in Fig. 1. It can be simply divided into four main components: 1) 2D Faster R-CNN based candidate detection to rapidly locate pulmonary nodule patches; 2) merging overlapping candidates by combining 2D patches with close Euclidean distances; 3) traditional three-layer 2D CNN based FP reduction to further eliminate FPs; and 4) modified FCNs based nodule segmentation to precisely segment the initial nodule mask. The geometric centers of detected nodules will guide the system to refine the segmentation and output the final nodule segmentation result. The more detailed explanation of each component is provided below.

2.2.1. Preprocessing

In order to reduce memory consumption, we assigned -1000 HU (air) as lower bound and +3000 HU (bone) as upper bound, then applied a linear mapping to convert the original signed 16-bit CT volume scans into 8-bit intensity values in the range of 0-255. This is the only preprocessing work in our proposed approach.

2.2.2. Faster R-CNN based nodule detection

The Faster R-CNN model evolved from Fast R-CNN (Girshick, 2015), which is mainly composed of a region proposal network (RPN) to propose potential regions of objects and an object detection network (ODN) to classify the region proposals from RPN. The new contribution of Faster R-CNN is that RPN and ODN can share the same convolutional layers, which enabled a more unified system to run at near real-time frame rates on natural images without performance loss (Ren et al., 2017). Based on our experiments, it is

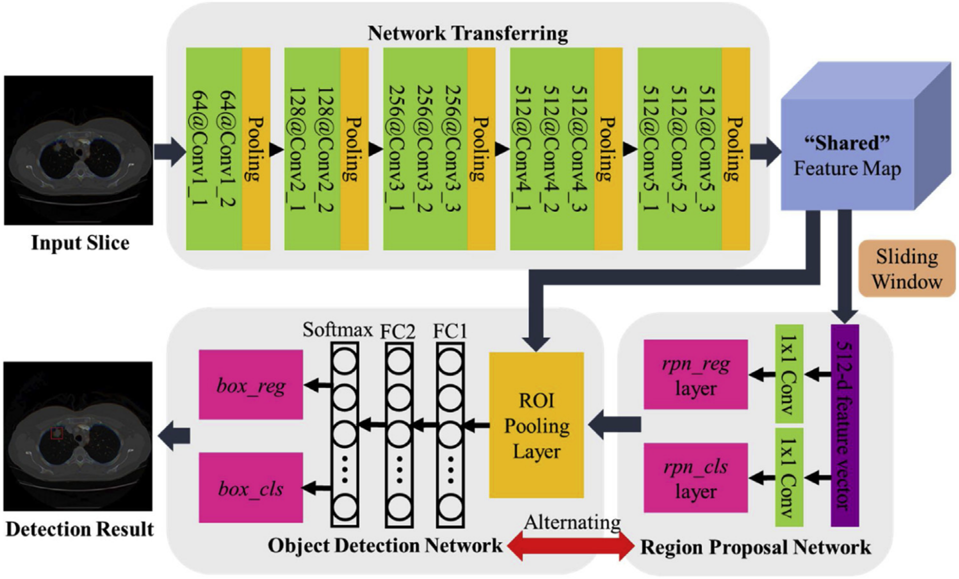


Fig. 2. The architecture of proposed Faster R-CNN based nodule detection.

fully convertible to be applied on nodule detection after conducting several modifications. The network architecture is shown in Fig. 2.

To accelerate training convergence and save computation workload, we reused the weights of first five groups of convolutional and pooling layers with a total of thirteen consecutive convolutional (Conv) and five pooling layers from a pretrained VGG16 model, and the calculated feature maps are shared with both RPN and ODN training. For RPN, we utilized a small network by sliding a 3 × 3 window at a time over the shared feature space to convert this feature map to a 512-dimension feature vector followed by a ReLU layer (Nair and Hinton, 2010). We implement the aforementioned small network by a 3 × 3 convolutional layer. After two sibling 1 × 1 convolutional layers, an RPN regression (*rpn_reg*) layer outputs the bounding box coordinates of each proposal, and an RPN classification layer (*rpn_cls*) estimates the probability of the proposal being a nodule. The design of anchors ensures the capability to parallelly predict multiple nodule proposals at each sliding window location. Different from common objects in natural images with big sizes and elongated shapes, the nodules are with relatively small and square boundaries. Therefore, we used a fixed scale ratio (1:1) and removed the other scale ratios in original RPN design, and implemented seven anchors with ascending common differences: 4 × 4, 6 × 6, 10 × 10, 16 × 16, 24 × 24, 36 × 36, 52 × 52 to fit the size variations of nodules. Because of the sparse distribution of nodules, we also adopt non-maximum suppression (NMS) based on the scores of *rpn_cls* with intersection over union (IoU) threshold of 0.7 between groundtruth and RPN proposals. NMS massively reduces the number of proposals and also potentially improves the detection accuracy, since high-density proposals may cause RPN to identify many surrounding regions that partially overlapped with true nodules (please see Fig. 3 as an example). The learnt proposals from RPN are fed into ODN for further classifications.

Taking the proposals predicted by RPN, the ODN is involved to serve as a binary classifier to determine nodule regions using the architecture of traditional CNNs. A ROI pooling layer is imposed to map each proposal to a smaller feature map by implementing max pooling operation of the values in a fixed 7 × 7 sub-window. Then two 4096-way fully-connected layers (FC1 and FC2) are conducted to produce a lower dimension feature vector, followed by two independent Softmax layers to output bounding boxes and probability scores of predicted nodules (*box_reg* and *box_cls*). In Faster R-CNN, RPN and ODN are mutually finetuned by adopting a pragmatic four-step alternating training procedure (Ren et al., 2017). As such, we achieved a unified network with sharing convolutional layers for RPN and ODN and the loss function for a single batch of *N* images is defined as follows.

$$\mathcal{L}_t = \frac{1}{N}\left[\frac{1}{\mathrm{N}_{rr}}\sum_i \mathcal{L}_r\left(t_i,\, t_i^*\right) + \frac{1}{\mathrm{N}_{rc}}\sum_i \mathcal{L}_c\left(p_i,\, p_i^*\right) + \frac{1}{\mathrm{N}_{br}}\sum_j \mathcal{L}_r\left(t_j,\, t_j^*\right) + \frac{1}{\mathrm{N}_{bc}}\sum_j \mathcal{L}_c\left(p_j,\, p_j^*\right)\right] \tag{1}$$

$$\mathcal{L}_r\left(t,\, t^*\right) = \mathrm{R}(t - t^*) \tag{2}$$

$$\mathcal{L}_c\left(p,\, p^*\right) = -log[pp^* + (1-p)(1-p^*)] \tag{3}$$

where N_{rr}, N_{rc}, N_{br}, and N_{bc} are numbers of inputs in *rpn_reg*, *rpn_cls*, *box_reg*, *box_cls* layers respectively, $\mathcal{L}_r$ and $\mathcal{L}_c$ represent loss associated the regression and classification layers, t_i represents the four coordinates of predicted nodule proposal, t_i^* is the coordinates of the corresponding groundtruth nodule, p_i and p_i^* denote the predicted and true probability of current anchor to be a nodule in RPN. Similarly, t_j, t_j^*, p_j, p_j^* implicit the same concepts in ODN. Besides, R is a robust loss function (smooth L_1 loss) explained in (Girshick, 2015).

2.2.3. *Merging of overlapping nodule candidates*

After Faster R-CNN, most nodule candidates can be detected. However, a single nodule may appear in several slices, therefore it may have more than one candidate representing the same nodule especially for some nodules with blurry edges. Based on some intuitive observations, these candidates will lie in close proximity to each other. Thus, a simple and computation-efficient merging operation is implemented by recursively combining candidates within five voxels of each other until no further merge is needed. This merging procedure ensures that a single nodule is identified within a single 2D slice rather than multiple slices alongside each other, dramatically decreases the number of unnecessary detections, and fastens the processing speed of the following FP reduction since the base number of candidates are lowered.

2.2.4. *CNN based false positive reduction*

With the nodule candidates extracted by Faster R-CNN, true nodule regions are successfully identified with small amount of

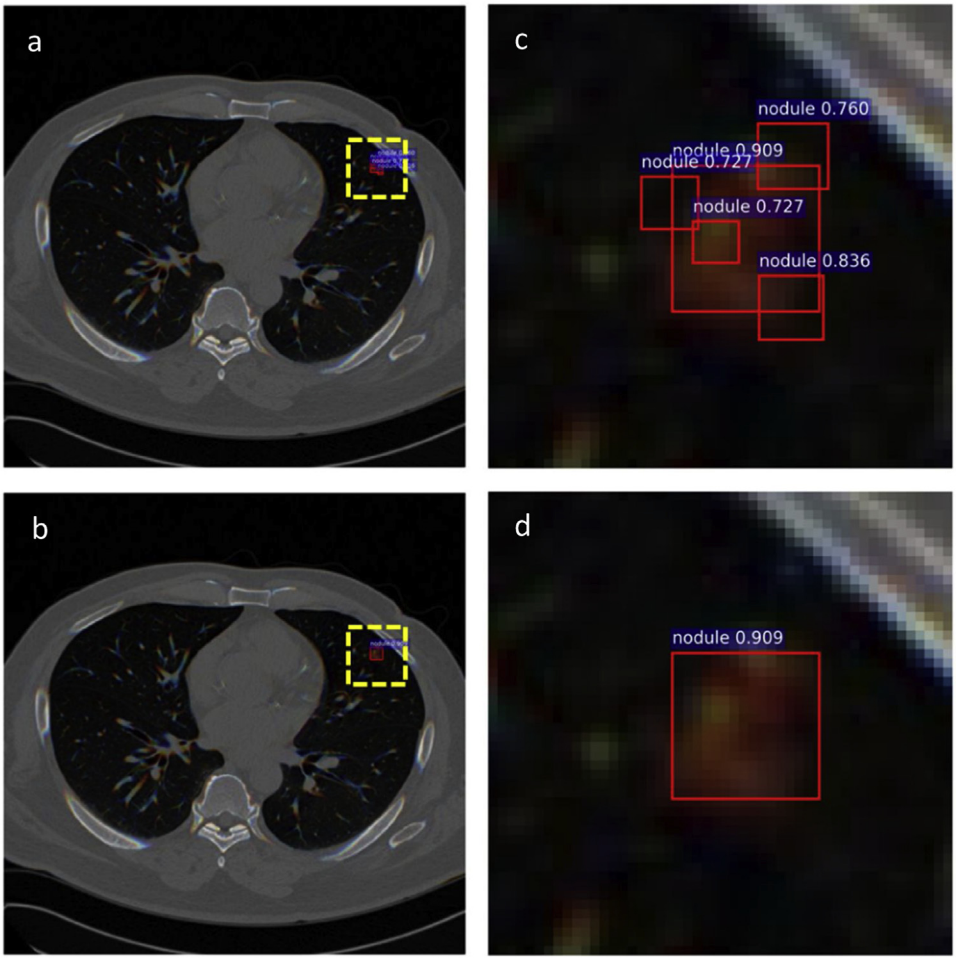


Fig. 3. Comparison of nodule candidate detection with and without non-maximum suppression (NMS) operation. (a) Nodule candidates without NMS operation. (b) Nodule candidates with NMS operation. (c) and (d) are zoom-in view of the region marked by yellow dashed square in (a) and (b). The predicted nodule candidates by Faster R-CNN are marked by red solid squares with the classification probability on top of them. Note that only the candidates with classification probability of nodules larger than 0.7 are displayed in this example. With NMS operation, our system successfully detects a ground glass opacity (GGO) nodule without partially overlapped duplicates (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

FPs. However, the existence of FPs still prohibits the use of CAD system in clinical practice. Considering the requirement of computation time and the advantages of Faster R-CNN component, a simple 2D CNN based classifier is sufficient to handle FP reduction task. Based on our previous work (Sun et al., 2017a) with the addition of multi-view nodule patches obtained from nine symmetrical planes presented in a previous work (Setio et al., 2016), an individual CNN with three pairs of convolution-pooling layers and one fully-connected layer is implemented. The loss is calculated by using the cross-entropy error, and weights are updated using mini-batches of 128 images. Then the testing is incorporated based on the Faster R-CNN initial candidate detection results. The kernel size for each convolutional layer is 5 × 5, 3 × 3, and 3 × 3 and the numbers of filters are 24, 48, 64, respectively. The input image size is 64 × 64 for both training and testing sets. After FP reduction step, the occurrence of FPs is largely decreased while maintaining a high detection sensitivity. Since the FP reduction network will be executed after Faster R-CNN initial candidate detection, different initial candidate FP levels with corresponding sensitivity may cause variations on the overall performance. In order to achieve the best performance, we empirically set three initial candidate FP levels (small, medium, large) and individually employ FP reduction on each FP level. We quantify the results in Section 3.

2.2.5. *Modified FCN based nodule segmentation*

FP reduction eliminated the most unlikely nodule candidates detected by Faster R-CNN. Among the remaining candidates, the 100 by 100 patches containing the nodule detection bounding boxes are created. Compared to FP reduction network that conducts a single imagewise two-class classification, more precise pixelwise classification is performed for segmentation purpose. Therefore, larger receptive field with richer background texture is needed. The modified deconvolutional neural network (Long et al., 2015) is used to generate the detailed nodule segmentation contour. In this experiment, the VGG16 is imposed as the backbone, and the weights of all the convolutional layers are initiated by ImageNet VGG16 pretrained model, while the weights of the later deconvolutional layers are randomized. In VGG16, there are five groups of convolutional layers altogether (Conv1 to Conv5), and each group contains a few consecutive convolutional layers. The original FCN used three convolutional layers (Conv3, Conv4 and Conv7) to generate the segmentation results. Since the lower convolutional layers have higher resolution, incorporating these layers should help with segmentation precision. In this study, two extra rounds of training are added to FCN training procedure to incorporate the first two convolutional layers (Conv1 and Conv2) to deconvolutional networks as well. The architecture of our modified FCN is shown in Fig. 4.

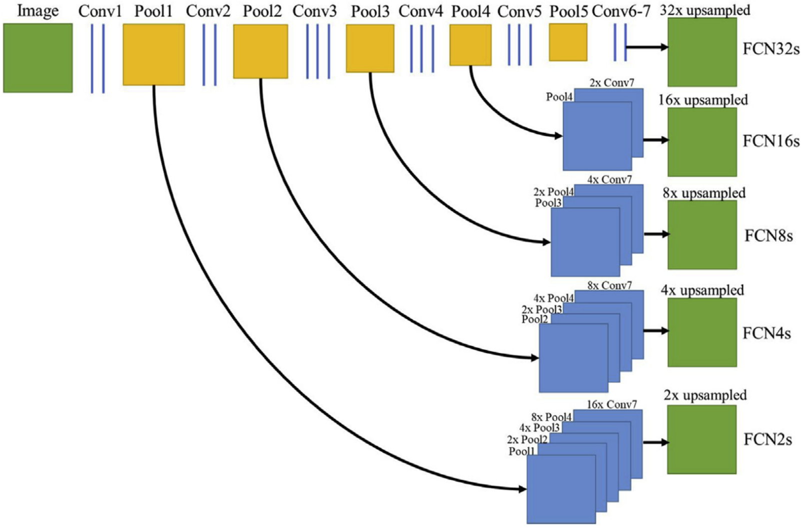


Fig. 4. The architecture of our modified FCN.

To initiate the FCN, the ImageNet VGG16 weights are downloaded and a classification FCN is trained based on it. Each convolutional layer provides the local neighborhood information of the image, and each pooling layer down-samples the image by the stride of 2. To fine tune the VGG16, the third fully connected layer (fc8) is removed, and the output node is set to 2 since our dataset has two subsets: nodule and non-nodule. Then all the fully connected layers are converted to convolutional layers, and all the weights are preserved in the transformed model.

Then the fine-tuned VGG16 weights are used to initiate FCN32 s, and an extra deconvolutional layer connected to the last convolutional layer is initialized with random noise attached to the end. Then we up sample the deconvolutional layer result by 2 to generate the FCN32 s segmentation mask. The added deconvolutional layer and up sampling layer can be considered as a block, after FCN32 s being trained, another block is attached to the end to train FCN16 s. Because there is another up sample layer in FCN16 s compared to FCN32 s, the original deconvolutional results are twice as large. Different deconvolutional results are up sampled to the identical size as the original image, and concatenated together to generate the segmentation output. The whole process is repeated 5 times so the FCN2s output has the same size and resolution of the original image. All the deconvolutional layers are up sampled at the same size and concatenated together. Compared to original FCN8s, our model utilizes the lower level convolution results thus yielding higher resolutions and preserved the details of the nodule. Since segmentation is indeed a pixel wise classification task, we still use cross entropy loss as the loss function during the FCN training.

2.2.6. *Postprocessing*

The final segmentation results are obtained after the fusion of detected nodule centers and initial segmentation masks to remove non-nodule segmentations. This fusion operation only selects the corresponding segmented object with shortest Euclidean distances between detected nodule center and object centers in the initial segmentation mask.

3. Results and discussion

3.1. *Experimental design*

For the Faster R-CNN based nodule detection, we take the 888 filtered CT scans with 1186 high confidence pulmonary nodules from LUNA16 dataset and retrieve their original annotations with precise nodule boundary coordinates recorded in XML format from LIDC-IDRI. To create the bounding square for each nodule, we calculate the smallest bounding rectangle and treat the longer side as the side of the result bounding square and record the four edge coordinates in groundtruth file. Because of the 2D network design, a total of 7909 axial nodule slices are generated. Leveraging from the ten patient-level subsets by LUNA16, we include nodule slices of subset0 to subset5 as training, subset6 to subset7 as validation, and subset8 to subset9 as testing. As a consequence, 534 scans with 5040 nodule slices, 178 scans with 1411 nodule slices, and 176 scans with 1458 nodules slices are assigned to training, validation, and testing set, respectively. For the training step, we adapt backpropagation and stochastic gradient descent (SGD) (LeCun et al., 1989) to train our end-to-end network. The minibatch size of images per iteration is set as 2 and the minibatch size of anchors per image is set as 10. For each image, based on the number of positive anchors (nodule), we randomly crop equal number of negative anchors (background) so that the sampled positive and negative anchors have the ratio of 1:1. Except for the first five groups of convolutional layers from VGG16, we randomly initialize all the other layers by drawing weights from a zero-mean Gaussian distribution with standard deviation 0.01. The learning rate is 0.001 for the first 10 epochs, and 0.0001 for the remaining with the momentum of 0.9 and a weight decay of 0.0005. Based on the aforementioned settings, the loss is stable after 20 epochs training. The total training time for Faster R-CNN is about 2.5 h. For the modified FCN based nodule segmentation, the learning rates for the five deconvolutional training are set to 10–12, 10–14, 10–16, 10–14, 10–15, respectively. The momentum is set to 0.99, weight decay is 0.0005, the iteration is set to 50,000 for each round of training. Besides the initialization round, we train FCN with five rounds manner, adding one fully connected layer each time. Each round of FCN training is about 5 h, and total training time for FCN is about 30 h.

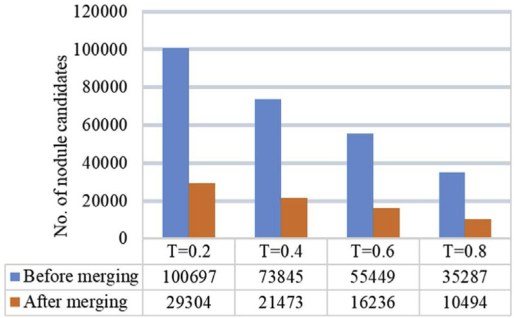
All the training and testing of CNNs are implemented using Tensorflow (Abadi et al., 2015) on a standard PC with a single 6GB memory GPU GeForce GTX 980Ti.

3.2. *Evaluation metrics*

To evaluate nodule detection accuracy we calculate the detection sensitivity and the corresponding average FPs per scan for every prediction probability level, and the Free Receiver Operating Characteristic (FROC) is imposed to visualize the sensitivity versus

Table 1
The system performance and CPM score comparison of the proposed method and other state-of-the-art approaches. Note that "online" means models with online descriptions available on LUNA16 competition website: https://luna16.grand-challenge.org/Results/, "*" represents models with limited details provided.

Team	GPU	CNN Dimension	CPM
PAtech (PA_tech)^{online}	Tesla K80 × 4	3D	0.951
JianPeiCAD (weiyixie)^{online}	NVIDIA Titan X × 2	3D	0.950
LUNA16FONOVACAD (zxp774747)^{online}	NVIDIA M40 × 2	3D	0.947
iFLYTEK-MIG (yinbaocai)*	–	–	0.941
zhongliu_xie (zhongliu.xie)^{online}	NVIDIA K80 × 2	3D	0.922
iDST-VC (chenjx1005)*	–	–	0.897
qfpxfd (qfpxfd) (Ding et al., 2017)	–	3D	0.891
CASED (CASED) (Jesson et al., 2017)	–	3D	0.887
3DCNN_NDET (lishaxue3)*	–	3D	0.882
Aidence (mjharte)*	–	–	0.871
junxuan20170516 (chenjx1005)*	–	3D	0.865
MEDICAI (bharadwaj)*	–	3D	0.862
Ethan20161221 (ethanhwang2012)*	–	–	0.856
resnet (QiDou) (Dou et al., 2017b)	NVIDIA Titan X × 1	3D	0.839
CCELargeCubeCnn (Intel_wuhui)^{online}	CPU	3D	0.833
Ours	NVIDIA 980 Ti × 1	2D	**0.880**



	T=0.2	T=0.4	T=0.6	T=0.8
🟦 Before merging	100697	73845	55449	35287
🟧 After merging	29304	21473	16236	10494

Fig. 5. Number of detected nodule candidates before and after candidate merging process with different cut-off thresholds of prediction probability for all 888 LUNA16 scans.

FPs relationship. A predicted nodule candidate will be counted as a true positive (TP) if it stays within five voxels to the real nodule center. The competition performance metric (CPM) score (Niemeijer et al., 2011) is calculated by taking the average detection sensitivity of seven predefined FPs per scan: 0.125, 0.25, 0.5, 1, 2, 4, and 8.

With regards to the nodule segmentation, we evaluate the segmentation similarity between automated segmentation and groundtruth mask through the calculation of the dice coefficients (Sørensen, 1948), which is defined by the following equation.

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{4}$$

3.3. Nodule detection

After nodule candidate detection by Faster R-CNN, we apply a merging step to combine the candidates that are geometrically close to each other. The number of candidates before and after the merging process is shown in Fig. 5. The number of candidates decrease by 70.9%, 70.9%, 70.7%, and 70.3% at cut-off prediction probability threshold of 0.2, 0.4, 0.6, and 0.8, respectively, which verifies the superiority of using the merging operation as a simple postprocessing step in 2D nodule candidate detection. The FROC curve of our Faster R-CNN based nodule candidate detection is shown in Fig. 6a. When the threshold of prediction probability is set as 0.5, we achieve the sensitivity of 95.2% with an average of 19.8 FPs per scan, which outperforms the traditional nodule detection algorithm in LUNA16 with an overall sensitivity of 94.4% with an average of 620.6 FPs per scan. Our proposed merging method massively reduces the burden of FP reduction task. The FROC curves after 2D CNN based FP reduction are shown in Fig. 6b-d by taking initial candidate FP level of 10 (L10), 15 (L15), 20 (L20) from Faster R-CNN results. We list the sensitivity at each initial candidate FP level using our nodule detection method in Table 1 with CPM score of 0.866, 0.875, and 0.880 for L10, L15, and L20, respectively. The best CPM is achieved in L20 with sensitivity of 91.4% and 94.6% at an average of 1 and 4 FPs per scan, respectively. The number of trainable parameters of Faster R-CNN, FP reduction, and modified FCN networks are 2.4×10^6, 7.1×10^4, and 1.3×10^8. The 3D U-net-based architecture in CASED (Jensen et al., 2017), which has identical CPM value compared to our approach, has approximated 6.6×10^6 learning parameters. Our Faster R-CNN + CNN model for detection has about 36% parameters of CASED model. By investigating different initial candidate FP levels from detecting network, these results show that a less conservative initial candidate selection threshold yields stronger performance (better CPM value), when FP reduction is employed. Moreover, a more progressive initial candidate selection threshold only performs better sensitivity at smaller FP levels (less than 0.25 FPs per scan in our case) in FROC curve.

Comparing with the top 15 models under the nodule detection track in LUNA16 challenge that employed 3D CNNs, our 2D nodule detection method successfully outperforms some of the 3D methods (Row 10 to 15 in Table 1) while keeping the computation and memory usage efficiency. Different from the methods (Row 1, 2, 3, and 5) using multiple powerful GPUs, our 2D model shows the capability of using cost-efficient GPUs. For instance, compared with Zhu et al., 2018 that employed a 3D Faster R-CNN network with 3D dual-path blocks and U-net-like encoder-decoder structure to compactly and effectively exploit features, our light and straightforward approach has much smaller number of learning parameters, thus the whole 2D slices that contain the entire natural contextual information can be fed into the training without considering the GPU memory limitation. They tried to save memory consumption by cropping 3D patches with pixel size 96 × 96 × 96 from original scans, which may also add computation burden at the testing phase.

Nodules that are successfully detected by candidate detection network but eliminated by FP reduction network are shown in Fig. 7. About 4% of correctly detected nodules in our testing set are falsely removed. These nodules present either irregular shapes or ambiguous boundaries. Bringing more data targeting these nodule representatives might potentially boost the sensitivity performance.

3.4. Nodule segmentation

Not considering FPs, the mean and standard deviation of dice coefficients (*DSC*) regarding the nodule segmentation using FCN2s,

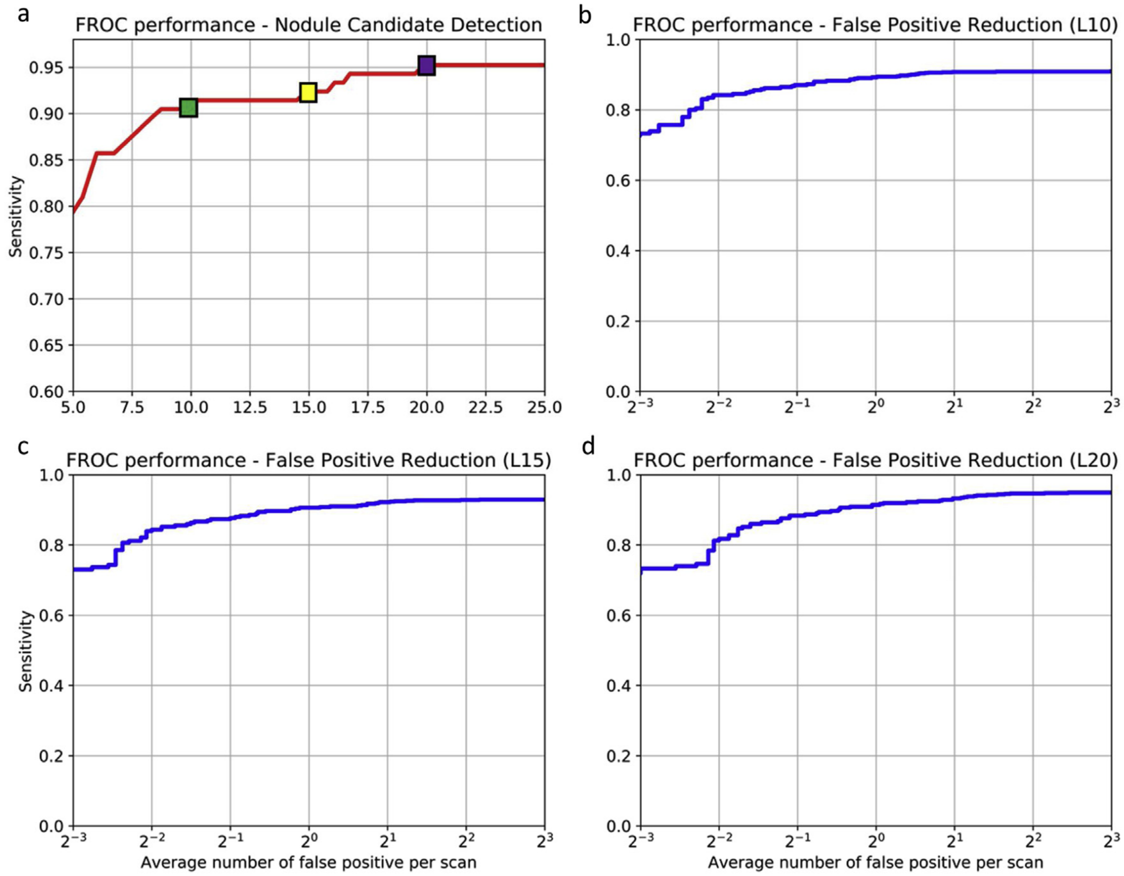


Fig. 6. The FROC curve of our nodule detection results. (a) FROC curve of Faster R-CNN based nodule candidate detection without FP reduction. (b) FROC curve of FP reduction by taking initial candidate FP level of 10 in.(a) as initial candidates marked by green square. (c) FROC curve of FP reduction by taking initial candidate FP level of 15 in.(a) as initial candidates marked by blue square. (d) FROC curve of FP reduction by taking initial candidate FP level of 20 in.(a) as initial candidates marked by purple square (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

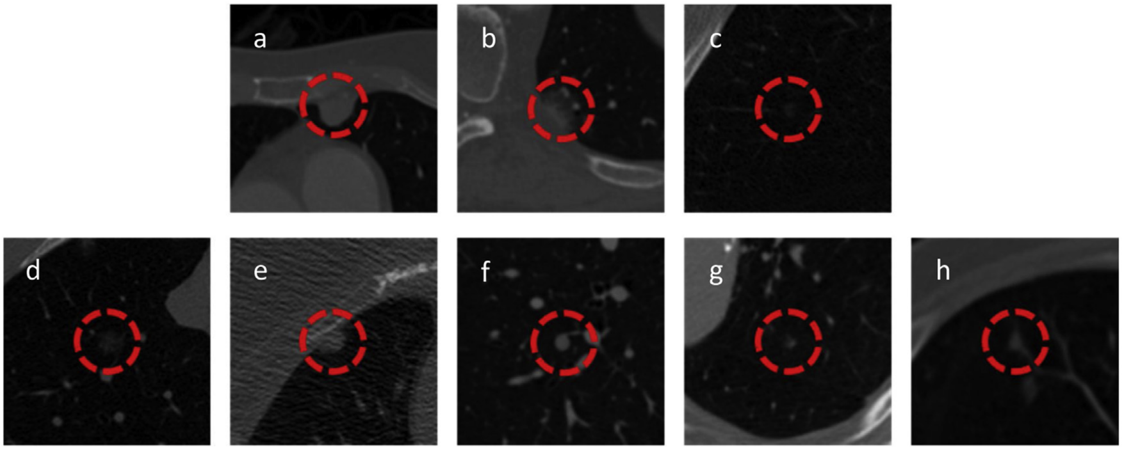


Fig. 7. Nodules in the testing set that are eliminated by FP reduction network. Red dashed circles mark the nodule locations. The top row (a to c) are nodules in subset8, and the bottom row (d to h) are nodules in subset9 (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

FCN4s, and FCN8s are presented in Table 2, which quantitatively proves that FCN2s with lower level convolution will potentially yield higher resolution and preserve the boundary information of the nodule. Therefore, higher *DSC* is obtained with the comparison to groundtruth mask. This proves our assumption that using lower level convolutional layers can help improve the segmentation accuracy. The higher level deconvolutional results provide the rough nodule locations and shape, while the lower level deconvolutional results provide the nodule boundary details.

Table 2
The mean and standard deviation of segmentation dice coefficients among FCN2s, FCN4s, and FCN8s.

	Mean	Std
FCN2s	**0.793**	**0.082**
FCN4s	0.723	0.149
FCN8s	0.674	0.172

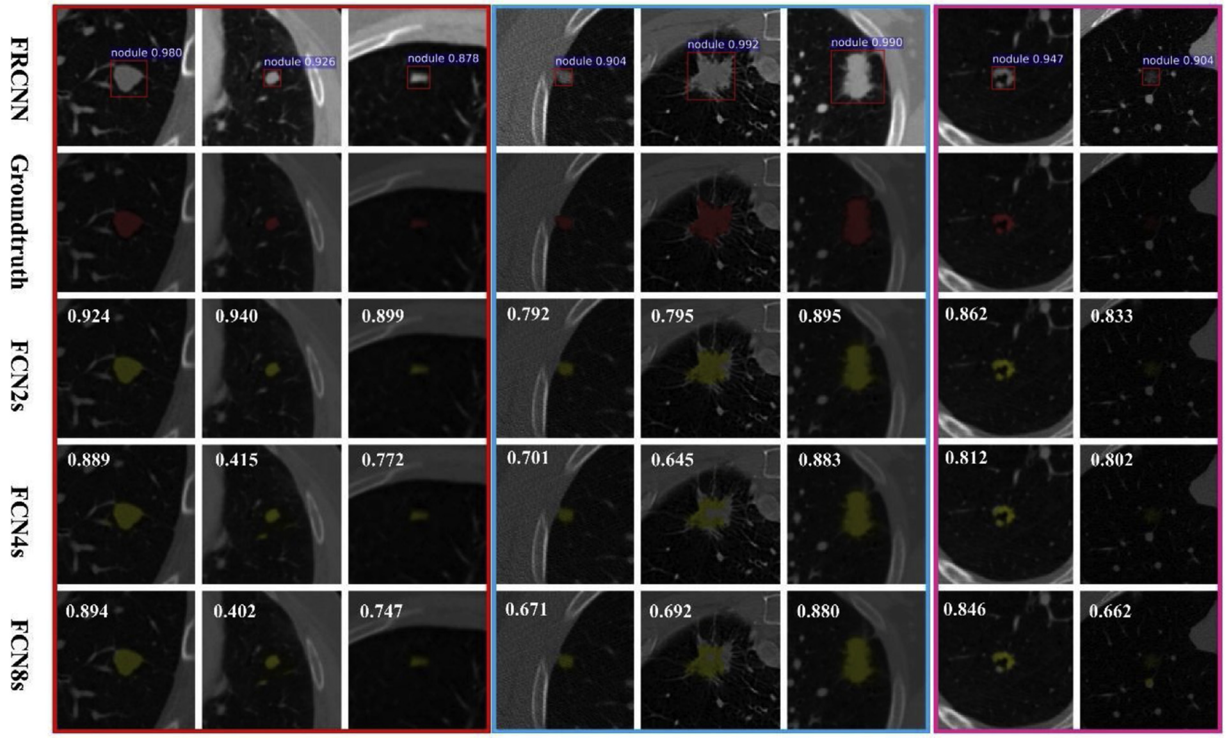


Fig. 8. Visualization results of our proposed nodule segmentation system with various anatomical characteristics. The columns 1–3 marked by red rectangle represent three isolated nodules, the columns 4–6 marked by light blue rectangle denote one juxta-pleural (column 4) and two juxta-vascular (column 5 and 6) nodules, and the columns 7 and 8 marked by light purple rectangle show one subsolid nodule with center excavation and one ground glass opacity (GGO) nodule. The first row represents original nodule patches after Faster R-CNN detection and FP reduction with predicted bounding boxes marked by solid red squares as well as the classification probabilities in light blue background. The second row represents the corresponding annotations by radiologists. The manual segmentations are emphasized by red masks. The third to fifth rows denote nodule segmentation results by FCN2s, FCN4s, and FCN8s, respectively. The fully-automated segmentations are emphasized by yellow masks. The white decimals implicit the dice coefficients for each segmentation compared to groundtruth markings (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 3
Performance of the proposed method and other state-of-the-art approaches.

Methods	Nodule amount	IoU
Tachibana and Kido (2006)	23	50.7 ± 21.9%
Wang et al. (2009)	64	58%
Messay et al. (2015)	68	63 ± 16%
Kubota et al. (2011)	23	69 ± 18%
	82	59 ± 19%
Tan et al. (2013)	23	65%
Lassen et al. (2015)	19	52 ± 7%
	40	50 ± 14%
Messay et al. (2015)	66	71.70 ± 19.89%
	77	69.23 ± 13.82%
Wang et al. (2017)	493	71.16 ± 12.22%
Ours	**223**	**70.24 ± 12.04%**

Table 3 shows the comparisons between our proposal segmentation result and other researchers' results based on the LIDC dataset. Since these methods use intersection over union (IoU) as the measurement metric, we report our results based on the same measurement in this table. Among all the listed approaches, our method achieves high performance on a relatively large amount of testing dataset.

In clinical practice, nodules have various types of clinical characteristics. The capability of segmenting large variety of nodules is necessary. Nine nodule characteristics are given in LIDC groundtruth markings, reflecting nodules' calcification, malignancy level and so on. We chose four representative characteristics and separate our testing set into different groups according to the corresponding characteristic scores. Based on the quantitative results in Table 4, our segmentation algorithm possesses the robustness of processing various types of nodules with similar performance.

Table 4
Dice coefficients on different nodule groups based on clinical characteristics. Note that nodules in the testing set are grouped based on their clinical characteristic scores. The numbers in square brackets represent the number of nodules in the corresponding group. We average the characteristic scores from four radiologists.

Characteristics	Scores					
	1	2	3	4	5	6
Calcification	– –	– –	0.794 [10]	0.817 [4]	0.788 [1]	0.795 [208]
Sphericity	0.784 [1]	0.740 [19]	0.823 [72]	0.781 [76]	0.798 [55]	– –
Spiculation	0.783 [157]	0.824 [36]	0.779 [14]	0.874 [6]	0.858 [10]	– –
Malignancy	0.737 [14]	0.780 [62]	0.807 [71]	0.812 [43]	0.800 [33]	– –

Visual comparisons between groundtruth, FCN2s, FCN4s, and FCN8s are shown in Fig. 8 by displaying several representative nodules. Our nodule segmentation algorithm can precisely segment large variety of nodules with different anatomical characteristics. Due to the complicated boundary pattern, the *DSC* for poor-circumscribed (juxta-pleural and juxta-vascular) and fuzzy-boundary (ground glass opacity) nodules is a little worse than well-circumscribed nodules. But since the main components of such nodules are successfully segmented, the quantitative performance decay will not impact the robustness in terms of visualization.

Because we are conducting an end-to-end nodule segmentation directly from raw CT scans, the combined dice coefficient including FPs can reflect the overall performance. The false positives (FPs) can be involved in both detection and segmentation steps. For the

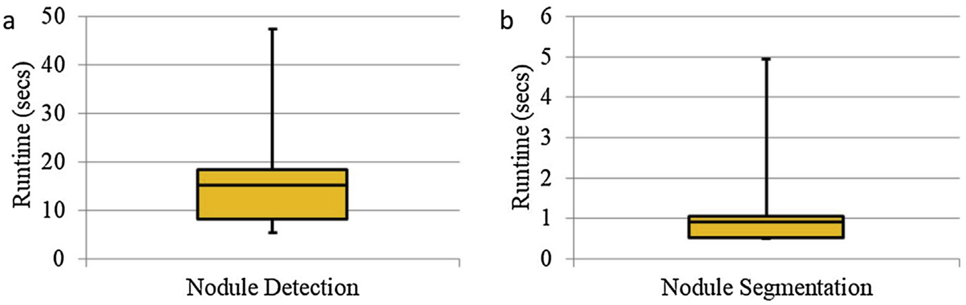


Fig. 9. Box plot of execution time for our proposed (a) nodule detection algorithm and (b) nodule segmentation algorithm.

Table 5
The mean and standard deviation of execution time for our proposed nodule detection and segmentation algorithm.

	Mean (secs)	Std
Detection (Faster RCNN + CNN)	15.319	8.126
Segmentation (FCN2s)	0.917	0.835

segmentation algorithm, the FPs can occur when the algorithm falsely detects other structures (vessels, airway walls, etc.) while avoiding the true nodules, which brings the dice coefficient to zero on FPs (Type I). For the detection algorithm, the FPs are apparently added when the algorithm fails to separate FPs from detected candidates (Type II). Leveraging from our full CAD system (detection + segmentation), the Type I FPs are eliminated by detection network since the center coordinates far from detected candidates will be removed after the postprocessing step illustrated in Section 2.2.6. Therefore, only Type II FPs are remained. We can simply calculate the combined dice score by considering all positive detected regions (TPs and FPs). For the given testing set, we choose sensitivity of 91.4% with 1 FPs per scan in the detection step and feed all the candidates in the segmentation network, the end-to-end dice score will be decreased to 0.426 (0.793 if we exclude the Type II FPs). Therefore, the combined dice coefficient provides the convention to be compared to other end-to-end nodule segmentation algorithms.

3.5. *Execution performance*

As for execution perspective, we demonstrate the efficiency of the proposed method by running our automated nodule segmentation system for all 888 LUNA16 scans. Fig. 9 shows statistical box plots regarding nodule detection (Faster R-CNN + CNN) and nodule segmentation (FCN2s). Note that the execution time is scan wised, so there are some variations due to the variations of slice numbers. The mean values and standard deviations are indicated in Table 5. Our end-to-end fully-automated nodule segmentation system achieves an average of 16 s per scan without any human interventions.

4. Conclusions

In this paper, we developed a four-step pure CNNs based pulmonary CAD algorithm on thoracic CT scans, which can automatically and efficiently segment lung nodules with reasonable amount of FPs. Even though there is still a performance gap compared to some state-of-the-art methods that involved much denser 3D convolutional layers with much more complex designs, our results and evaluations demonstrated the capability of using pure 2D CNNs on a standard workstation to detect and segment pulmonary nodules with high performance. Such automation and efficiency significantly facilitate the translation from scientific researches to real applications on the computerized lung nodule segmentation trend. For the future outlook, further evaluation based on independent testing dataset is desired to perform a more comprehensive comparison. Besides, for any 2D detection scheme implemented on volumetric imaging, the model performance is also critical when being applied across an entire volume. We will assess it in the future. Finally, it is attractive to implement more advanced 2D CNNs architectures (such as generative adversarial network in Goodfellow et al., 2014) to further boost the performance.

Conflict of interest statement

No conflict of interest.

Acknowledgement

This work was supported by US National Science Foundation (award number 1429708 and 1827745).

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Kaiser, L., Kudlur, M., Levenberg, J., Man, D., Monga, R., Moore, S., Murray, D., Shlens, J., Steiner, B., Sutskever, I., Tucker, P., Vanhoucke, V., Vasudevan, V., Vinyals, O., Warden, P., Wicke, M., Yu, Y., Zheng, X., 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. None., http://dx.doi.org/10.1038/nn.3331.

Aberle, D.R., Adams, A.M., Berg, C.D., Black, W.C., Clapp, J.D., Fagerstrom, R.M., Gareen, I.F., Gatsonis, C., Marcus, P.M., Sicks, J.D., Lung, National, Screening Trial Research Team, 2011. Reduced lung-cancer mortality with low-dose computed tomographic screening. N. Engl. J. Med. 365 (5), 395–409, http://dx.doi.org/10.1056/NEJMoa1102873.

Anirudh, R., Thiagarajan, J.J., Bremer, T., Kim, H., 2016. Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data. SPIE Med. Imaging 9785, http://dx.doi.org/10.1117/12.2214876.

Armato, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., Kazerooni, E.A., MacMahon, H., Van Beek, E.J.R., Yankelevitz, D., Biancardi, A.M., Bland, P.H., Brown, M.S., Engelmann, R.M., Laderach, G.E., Max, D., Pais, R.C., Qing, D.P.Y., Roberts, R.Y., Smith, A.R., Starkey, A., Batra, P., Caligiuri, P., Farooqi, A., Gladish, G.W., Jude, C.M., Munden, R.F., Petkovska, I., Quint, L.E., Schwartz, L.H., Sundaram, B., Dodd, L.E., Fenimore, C., Gur, D., Petrick, N., Freymann, J., Kirby, J., Hughes, B., Vande Casteele, A., Gupte, S., Sallam, M., Heath, M.D., Kuhn, M.H., Dharaiya, E., Burns, R., Fryd, D.S., Salganicoff, M., Anand, V., Shreter, U., Vastagh, S., Croft, B.Y., Clarke, L.P., 2011. 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), 915–931, http://dx.doi.org/10.1118/1.3528204.

Armato III, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Clarke, L.P., 2015. Data from LIDC-IDRI [WWW document]. Cancer Imaging Arch., http://dx.doi.org/10.1111/j.1558-5646.2010.01177.x.

Chen, K., Fu, K., Yan, M., Gao, X., Sun, X., Wei, X., 2018. Semantic segmentation of aerial images with shuffling convolutional neural networks. Geosci. Remote Sens. Lett. IEEE 15 (2), 173–177, http://dx.doi.org/10.1109/LGRS.2017.2778181.

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., Prior, F., 2013. The cancer imaging archive (TCIA): maintaining and operating a public information repository. J. Digit. Imaging 26 (6), 1045–1057, http://dx.doi.org/10.1007/s10278-013-9622-7.

Dai, S., Lu, K., Dong, J., Zhang, Y., Chen, Y., 2015. A novel approach of lung segmentation on chest CT images using graph cuts. Neurocomputing 168, 799–807, http://dx.doi.org/10.1016/j.neucom.2015.05.044.

Dehmeshki, J., Amin, H., Valdivieso, M., Ye, X., 2008. Segmentation of pulmonary nodules in thoracic CT scans: a region growing approach. IEEE Trans. Med. Imaging 27 (4), 467–480, http://dx.doi.org/10.1109/TMI.2007.907555.

DeVore, R.A., Yang, Z., Kallergi, M., Qian, W., Clark, R.A., Clarke, L.P., Lucier, B.J., 1995. Effect of wavelet bases on compressing digital mammograms. IEEE Eng. Med. Biol. Mag. 14 (5), 570–577, http://dx.doi.org/10.1109/51.464773.

Ding, J., Li, A., Hu, Z., Wang, L., 2017. Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)., http://dx.doi.org/10.1007/978-3-319-66179-7_64.

Dong, F., Peng, J., 2014. Brain MR image segmentation based on local Gaussian mixture model and nonlocal spatial regularization. J. Vis. Commun. Image Represent. 25 (5), 827–839, http://dx.doi.org/10.1016/j.jvcir.2014.01.014.

Dou, Q., Chen, H., Yu, L., Qin, J., Heng, P.A., 2017a. Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection. IEEE Trans. Biomed. Eng. 64 (7), 1558–1567, http://dx.doi.org/10.1109/TBME.2016.2613502.

Dou, Q., Chen, H., Jin, Y., Lin, H., Qin, J., Heng, P.A., 2017b. Automated pulmonary nodule detection via 3D convnets with online sample filtering and hybrid-loss residual learning. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10435., pp. 630–638, http://dx.doi.org/10.1007/978-3-319-66179-7_72.

Fitzmaurice, C., Allen, C., Barber, R.M., Barregard, L., Bhutta, Z.A., Brenner, H., Dicker, D.J., Chimed-Orchir, O., Dandona, R., Dandona, L., Fleming, T., Forouzanfar, M.H., Hancock, J., Hay, R.J., Hunter-Merrill, R., Huynh, C., Hosgood, H.D., Johnson, C.O., Jonas, J.B., Khubchandani, J., Kumar, G.A., Kutz, M., Lan, Q., Larson, H.J., Liang, X., Lim, S.S., Lopez, A.D., MacIntyre, M.F., Marczak, L., Marquez, N., Mokdad, A.H., Pinho, C., Pourmalek, F., Salomon, J.A., Sanabria, J.R., Sandar, L., Sartorius, B., Schwartz, S.M., Shackelford, K.A., Shibuya, K., Stanaway, J., Steiner, C., Sun, J., Takahashi, K., Vollset, S.E., Vos, T., Wagner, J.A., Wang, H., Westerman, R., Zeeb, H., Zoeckler, L., Abd-Allah, F., Ahmed, M.B., Alabed, S., Alam, N.K., Aldhahri, S.F., Alem, G., Alemayohu, M.A., Ali, R., Al-Raddadi, R., Amare, A., Amoako, Y., Artaman, A., Asayesh, H., Atnafu, N., Awasthi, A., Saleem, H.B., Barac, A., Bedi, N., Bensenor, I., Berhane, A., Bernabé, E., Betsu, B., Binagwaho, A., Boneya, D., Campos-Nonato, I., Castañeda-Orjuela, C., Catalá-López, F., Chiang, P., Chibueze, C., Chitheer, A., Choi, J.-Y., Cowie, B., Damtew, S., das Neves, J., Dey, S., Dharmaratne, S., Dhillon, P., Ding, E., Driscoll, T., Ekwueme, D., Endries, A.Y., Farvid, M., Farzadfar, F., Fernandes, J., Fischer, F., G/hiwot, T.T., Gebru, A., Gopalani, S., Hailu, A., Horino, M., Horita, N., Husseini, A., Huybrechts, I., Inoue, M., Islami, F., Jakovljevic, M., James, S., Javanbakht, M., Jee, S.H., Kasaeian, A., Kedir, M.S., Khader, Y.S., Khang, Y.-H., Kim, D., Leigh, J., Linn, S., Lunevicius, R., El Razek, H.M.A., Malekzadeh, R., Malta, D.C., Marcenes, W., Markos, D., Melaku, Y.A., Meles, K.G., Mendoza, W., Mengiste, D.T., Meretoja, T.J., Miller, T.R., Mohammad, K.A., Mohammadi, A., Mohammed, S., Moradi-Lakeh, M., Nagel, G., Nand, D., Le Nguyen, Q., Nolte, S., Ogbo, F.A., Oladimeji, K.E., Oren, E., Pa, M., Park, E.-K., Pereira, D.M., Plass, D., Qorbani, M., Radfar, A., Rafay, A., Rahman, M., Rana, S.M., Søreide, K., Satpathy, M., Sawhney, M., Sepanlou, S.G., Shaikh, M.A., She, J., Shiue, I., Shore, H.R., Shrime, M.G., So, S., Soneji, S., Stathopoulou, V., Stroumpoulis, K., Sufiyan, M.B., Sykes, B.L., Tabarés-Seisdedos, R., Tadese, F., Tedla, B.A., Tessema, G.A., Thakur, J.S., Tran, B.X., Ukwaja, K.N., Uzochukwu, B.S.C., Vlassov, V.V., Weiderpass, E., Wubshet Terefe, M., Yebyo, H.G., Yimam, H.H., Yonemoto, N., Younis, M.Z., Yu, C., Zaidi, Z., Zaki, M.E.S., Zenebe, Z.M., Murray, C.J.L., Naghavi, M., 2017. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015. JAMA Oncol. 3 (4), 524–548, http://dx.doi.org/10.1001/jamaoncol.2016.5688.

Fu, L., Ma, J., Ren, Y., Han, Y.S., Zhao, J., 2017. Automatic detection of lung nodules: false positive reduction using convolution neural networks and handcrafted features. SPIE Med. Imaging 10134, http://dx.doi.org/10.1117/12.2253995.

Ge, Z., Sahiner, B., Chan, H.-P., Hadjiiski, L.M., Cascade, P.N., Bogot, N., Kazerooni, E.A., Wei, J., Zhou, C., 2005. Computer-aided detection of lung nodules: false positive reduction using a 3D gradient field method and 3D ellipsoid fitting. Med. Phys. 32 (8), 2443–2454, http://dx.doi.org/10.1118/1.1944667.

Girshick, R., 2015. Fast R-CNN. Proceedings of the IEEE International Conference on Computer Vision, http://dx.doi.org/10.1109/ICCV.2015.169.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., arXiv Prepr. arXiv 2014. Generative Adversarial Networks., http://dx.doi.org/10.1017/CBO9781139058452.

Hamidian, S., Sahiner, B., Petrick, N., Pezeshk, A., 2017. 3D convolutional neural network for automatic detection of lung nodules in chest CT. Proc. SPIE Int. Soc. Opt. Eng. 10134, http://dx.doi.org/10.1117/12.2255795.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), http://dx.doi.org/10.1109/CVPR.2016.90.

Huang, X., Shan, J., Vaidya, V., 2017. Lung nodule detection in CT using 3D convolutional neural networks. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), http://dx.doi.org/10.1109/ISBI.2017.7950542.

Jacobs, C., van Rikxoort, E.M., Twellmann, T., Scholten, E.T., de Jong, P.A., Kuhnigk, J.M., Oudkerk, M., de Koning, H.J., Prokop, M., Schaefer-Prokop, C., van Ginneken, B., 2014. Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images. Med. Image Anal. 18 (2), 374–384, http://dx.doi.org/10.1016/j.media.2013.12.001.

Jesson, A., Guizard, N., Ghalehjegh, S.H., Goblot, D., Soudan, F., Chapados, N., 2017. CASED: curriculum adaptive sampling for extreme data imbalance. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)., http://dx.doi.org/10.1007/978-3-319-66179-7_73.

Jin, H., Li, Z., Tong, R., Lin, L., 2018. A deep 3D residual CNN for false-positive reduction in pulmonary nodule detection. Med. Phys. 45 (5), 2097–2107, http://dx.doi.org/10.1002/mp.12846.

Kubota, T., Jerebko, A.K., Dewan, M., Salganicoff, M., Krishnan, A., 2011. Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models. Med. Image Anal. 15 (1), 133–154, http://dx.doi.org/10.1016/j.media.2010.08.005.

Lassen, B.C., Jacobs, C., Kuhnigk, J.M., Van Ginneken, B., Van Rikxoort, E.M., 2015. Robust semi-automatic segmentation of pulmonary subsolid nodules in chest computed tomography scans. Phys. Med. Biol. 60 (3), 1307–1323, http://dx.doi.org/10.1088/0031-9155/60/3/1307.

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., 1989. Backpropagation applied to handwritten zip code recognition. Neural Comput. 1 (4), 541–551, http://dx.doi.org/10.1162/neco.1989.1.4.541.

Lekić, V., Babić, Z., 2018. Using GANs to enable semantic segmentation of ranging sensor data. Zooming Innovation in Consumer Technologies Conference (ZINC), http://dx.doi.org/10.1109/ZINC.2018.8448963.

Li, Q., Li, F., Doi, K., 2008. Computerized detection of lung nodules in thin-section CT images by use of selective enhancement filters and an automated rule-based classifier. Acad. Radiol. 15 (2), 165–175, http://dx.doi.org/10.1016/j.acra.2007.09.018.

Liu, H., Geng, F., Guo, Q., Zhang, C., Zhang, C., 2018. A fast weak-supervised pulmonary nodule segmentation method based on modified self-adaptive FCM algorithm. Soft Comput. 22 (12), 3983–3995, http://dx.doi.org/10.1007/s00500-017-2608-5.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, http://dx.doi.org/10.1109/CVPR.2015.7298965.

Long, M., Zhu, H., Wang, J., Jordan, M.I., arXiv Prepr. arXiv 2016. Deep Transfer Learning with Joint Adaptation Networks.

Magalhães Barros Netto, S., Corrĉa Silva, A., Acatauassú Nunes, R., Gattass, M., 2012. Automatic segmentation of lung nodules with growing neural gas and support vector machine. Comput. Biol. Med. 42 (11), 1110–1121, http://dx.doi.org/10.1016/j.compbiomed.2012.09.003.

Mao, Q., Zhao, S., Gong, T., Zheng, Q., 2018. An effective hybrid windowed fourier filtering and fuzzy C-mean for pulmonary nodule segmentation. J. Med. Imaging Health Inform. 8 (1), 72–77, http://dx.doi.org/10.1166/jmihi.2018.2235.

Messay, T., Hardie, R.C., Tuinstra, T.R., 2015. Segmentation of pulmonary nodules in computed tomography using a regression neural network approach and its application to the Lung Image Database Consortium and Image Database Resource Initiative dataset. Med. Image Anal. 22 (1), 48–62, http://dx.doi.org/10.1016/j.media.2015.02.002.

Murphy, K., van Ginneken, B., Schilham, A.M.R., de Hoop, B.J., Gietema, H.A., Prokop, M., 2009. A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification. Med. Image Anal. 13 (5), 757–770, http://dx.doi.org/10.1016/j.media.2009.07.001.

Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted Boltzmann machines. In: Proc. 27th Int. Conf. Mach. Learn.

Niemeijer, M., Loog, M., Abramoff, M.D., Viergever, M.A., Prokop, M., van Ginneken, B., 2011. On combining computer-aided detection systems. IEEE Trans. Med. Imaging 30 (2), 215–223, http://dx.doi.org/10.1109/TMI.2010.2072789.

Qian, W., Kallergi, M., Clarke, L.P., 1993. Order statistic-neural network hybrid filters for gamma camera-bremsstrahlung image restoration. IEEE Trans. Med. Imaging 12 (1), 58–64, http://dx.doi.org/10.1109/42.222667.

Reeves, A.P., Chan, A.B., Yankelevitz, D.F., Henschke, C.I., Kressler, B., Kostis, W.J., 2006. On measuring the change in size of pulmonary nodules. IEEE Trans. Med. Imaging 25 (4), 435–450, http://dx.doi.org/10.1109/TMI.2006.871548.

Ren, S., He, K., Girshick, R., Sun, J., 2017. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. 39 (6), 1137–1149, http://dx.doi.org/10.1109/TPAMI.2016.2577031.

Setio, A.A.A., Jacobs, C., Gelderblom, J., Van Ginneken, B., 2015. Automatic detection of large pulmonary solid nodules in thoracic CT images. Med. Phys. 42 (10), 5642–5653, http://dx.doi.org/10.1118/1.4929562.

Setio, A.A.A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., Van Riel, S.J., Wille, M.M.W., Naqibullah, M., Sanchez, C.I., Van Ginneken, B., 2016. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans. Med. Imaging 35 (5), 1160–1169, http://dx.doi.org/10.1109/TMI.2016.2536809.

Siegel, R.L., Miller, K.D., Jemal, A., 2018. Cancer statistics, 2018. CA Cancer J. Clin. 68 (1), 7–30, http://dx.doi.org/10.3322/caac.21442.

Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. Int. Conf. Learn. Represent, http://dx.doi.org/10.1016/j.infsof.2008.09.005.

Sørensen, T.J., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. Kongelige Danske Videnskabernes Selskab.

Sun, X., Qian, W., Song, D., 2004. Ipsilateral-mammogram computer-aided detection of breast cancer. Comput. Med. Imaging Graph. 28 (3), 151–158, http://dx.doi.org/10.1016/j.compmedimag.2003.11.004.

Sun, W., Zheng, B., Qian, W., 2016. Computer aided lung cancer diagnosis with deep learning algorithms. SPIE Med. Imaging 9785, http://dx.doi.org/10.1117/12.2216307.

Sun, W., Huang, X., Tseng, T.-L. (Bill), Qian, W., 2017a. Automatic lung nodule graph cuts segmentation with deep learning false positive reduction. SPIE Med. Imaging 10134, http://dx.doi.org/10.1117/12.2251302.

Sun, W., Zheng, B., Huang, X., Qian, W., 2017b. Balance the nodule shape and surroundings: a new multichannel image based convolutional neural network scheme on lung nodule diagnosis. SPIE Med. Imaging 10134, http://dx.doi.org/10.1117/12.2251297.

Sun, W., Tseng, T.-L. (Bill), Zhang, J., Qian, W., 2017c. Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data. Comput. Med. Imaging Graph. 57, 4–9, http://dx.doi.org/10.1016/j.compmedimag.2016.07.004.

Sun, W., Zheng, B., Qian, W., 2017d. Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis. Comput. Biol. Med. 89, 530–539, http://dx.doi.org/10.1016/j.compbiomed.2017.04.006.

Sun, W., Ph.D. Thesis 2017e. Deep Learning Method vs. Hand-Crafted Features for Lung Cancer Diagnosis and Breast Cancer Risk Analysis. The University of Texas at El Paso.

Tachibana, R., Kido, S., 2006. Automatic segmentation of pulmonary nodules on CT images by use of NCI lung image database consortium. SPIE Med. Imaging 6144, http://dx.doi.org/10.1117/12.653366.

Tajbakhsh, N., Suzuki, K., 2017. Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs Pattern Recognit. 63, 476–486, http://dx.doi.org/10.1016/j.patcog.2016.09.029.

Tan, M., Deklerck, R., Jansen, B., Bister, M., Cornelis, J., 2011. A novel computer-aided lung nodule detection system for CT images. Med. Phys. 38 (10), 5630–5645, http://dx.doi.org/10.1118/1.3633941.

Tan, Y., Schwartz, L.H., Zhao, B., 2013. Segmentation of lung lesions on CT scans using watershed, active contours, and Markov random field. Med. Phys. 40 (4), 043502, http://dx.doi.org/10.1118/1.4793409.

Tuinstra, T.R., Ph.D. Thesis 2008. Automatic Segmentation of Small Pulmonary Nodules in Computed Tomography Data Using a Radial Basis Function Neural Network with Application to Volume Estimation. University of Dayton.

Van Ginneken, B., 2006. Supervised probabilistic segmentation of pulmonary nodules in CT scans. In: Med. Image Comput. Comput. Interv. 2006., http://dx.doi.org/10.1007/11866763_112.

Vijaya Kishore, V., Satyanarayana, R.V.S., 2013. Performance evaluation of edge detectors – morphology based ROI segmentation and nodule detection from DICOM lung images in the noisy environment. Proceedings of the 2013 3rd IEEE International Advance Computing Conference, IACC 2013, http://dx.doi.org/10.1109/IAdCC.2013.6514386.

Wang, Q., Song, E., Jin, R., Han, P., Wang, X., Zhou, Y., Zeng, J., 2009. Segmentation of lung nodules in computed tomography images using dynamic programming and multidirection fusion Techniques1. Acad. Radiol. 16 (6), 678–688, http://dx.doi.org/10.1016/j.acra.2008.12.019.

Wang, S., Zhou, M., Liu, Z., Liu, Z., Gu, D., Zang, Y., Dong, D., Gevaert, O., Tian, J., 2017. Central focused convolutional neural networks: developing a data-driven model for lung nodule segmentation. Med. Image Anal. 40, 172–183, http://dx.doi.org/10.1016/j.media.2017.06.014.

Weiss, K., Khoshgoftaar, T.M., Wang, D.D., 2016. A survey of transfer learning. J. Big Data, 3–9, http://dx.doi.org/10.1186/s40537-016-0043-6.

Winkels, M., Cohen, T.S., arXiv 2018. 3D G-CNNs for Pulmonary Nodule Detection.

Wu, B., Zhou, Z., Wang, J., Wang, Y., 2018. Joint learning for pulmonary nodule segmentation, attributes and malignancy prediction. Proceedings - International Symposium on Biomedical Imaging, http://dx.doi.org/10.1109/ISBI.2018.8363765.

Xia, Z., Wang, X., Sun, X., Liu, Q., Xiong, N., 2016. Steganalysis of LSB matching using differences between nonadjacent pixels. Multimed. Tools Appl. 75 (4), 1947–1962, http://dx.doi.org/10.1007/s11042-014-2381-8.

Ye, H., Sun, C., Ren, P., Dai, L., Peng, B., Wang, K., Qian, W., Zhang, J., 2013. Mini-array of multiple tumor-associated antigens (TAAs) in the immunodiagnosis of breast cancer. Oncol. Lett. 5 (2), 663–668, http://dx.doi.org/10.3892/ol.2012.1062.

Ypsilantis, P.-P., Montana, G., arXiv 2016. Recurrent Convolutional Networks for Pulmonary Nodule Detection in CT Imaging. https://arxiv.org/abs/1609.09143.

Yu, B., Yang, L., Chen, F., 2018. Semantic segmentation for high spatial resolution remote sensing images based on convolution neural network and pyramid pooling module. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11 (9), 3252–3261, http://dx.doi.org/10.1109/JSTARS.2018.2860989.

Zhang, Y., Sankar, R., Qian, W., 2007. Boundary delineation in transrectal ultrasound image for prostate cancer. Comput. Biol. Med. 37 (11), 1591–1599, http://dx.doi.org/10.1016/j.compbiomed.2007.02.008.

Zhu, W., Liu, C., Fan, W., Xie, X., 2018. DeepLung: deep 3D dual path nets for automated pulmonary nodule detection and classification. Proceedings – 2018 IEEE Winter Conference on Applications of Computer Vision, http://dx.doi.org/10.1109/WACV.2018.00079.