Lung Nodule Detection With Deep Learning In 3D Thoracic MR Images

G V Sowmya[1], Chethan G S[2]
[1, 2] Dept of ISE
[1, 2] JNNCE

Abstract- *The early detection of lung cancer is crucial in reducing mortality. Magnetic resonance imaging (MRI) may be a viable imaging technique for lung cancer detection. Numerous lung nodule detection methods have been studied for computed tomography (CT) images. However, to the best of our knowledge, no detection methods have been carried out for MR images. In this paper, a lung nodule detection method based on deep learning is proposed for thoracic MR images. With parameter optimizing, spatial three channel input construction and transfer learning, a Faster R-CNN network is designed to locate the lung nodule region. Then a false positive (FP) reduction scheme based on anatomical characteristics is designed to reduce FPs and preserve the true nodule. The proposed method is tested on 142 T2-weighted MR scans from the First Affiliated Hospital of Guangzhou Medical University. The sensitivity of the proposed method is 85.2% with 3.47 false positives per scan. Experimental results demonstrate that the designed Faster R-CNN network and the FP reduction scheme are effective in lung nodule detection and FP reduction for MR images.*

Keywords- Deep learning, Convolution Neural Network, lung nodule, Segmentation.

I. INTRODUCTION

Lung cancer has the highest rate of mortality among all type of cancers. It is one of the most widespread diseases and leads to the highest mortality among all the forms of cancer. As is reported in global cancer statistics in 2012, nearly 1.83 million new cases of lung cancer occurred and the estimated deaths are over 1.5 million [1]. The causes of lung cancer include smokes, toxic particles in the air, aging, gene, gender, etc. Unfortunately, people feel incapable of finding useful ways to curb the incidence of cancer and meanwhile it is almost impossible to find a cure when patients are in the terminal stage of cancer under the current condition. The confirmed cases of lung cancer are commonly accompanied by continuing emergence of nodules. Early detection of cancer can greatly improve survival rates of patients. Distinctly, the lung nodule detection in the initial stage of computer-aided diagnosis schemes is remarkable and obligatory.

The early detection of lung cancer is the corner stone to reduce mortality .The 5-year survival rate for all stages of lung cancer is only 16%. When lung cancer is detected at a localized stage, the 5-year survival rate will increase to 52% Computed tomography (CT) is a widely used technique for lung nodule detection. Recently, with the development of magnetic resonance imaging (MRI) technique, MRI has also been used in lung disease diagnosis .MRI is a non-radiation examination. MRI might be a valuable tool in malignant nodule detection and lung cancer screening. MRI can provide not only morphological, but also functional information, such as physiological , pathophysiological , and molecular information. MRI has multiple modalities and each MRI scan has many slices. The computer-aided detection (CAD) for MR images is also widely used. Pulmonary nodule detection is the basis for nodule measurement and classification. The study on automatic pulmonary nodule detection is vital.Most previous nodule detection methods require handcrafted feature extraction. The convolution neural network (CNN) automatically discovers features, shows promising results in many pattern recognition tasks.

Most previous nodule detection methods require handcrafted feature extraction. Recently convolution neural network (CNN) which automatically discovers features, shows promising results in many pattern recognition tasks. This led to the application of CNN in automated pulmonary nodule detection. At first , the nodule candidate was detected through traditional method. Then a patch was cropped for each candidate and CNN was used to classify it as normal or nodule. Later the nodule detection was carried out through 3D CNN method.

II. PROBLEM FORMULATION

Lung cancer has the highest rate of mortality among all type of cancers.Numerous lung nodule detection methods have been studied for CT images. For MR images, few detection methods have been carried out. CT has better image contrast and less air artifacts. Due to the image differences between CT and MR, nodule detection methods for CT images may be not suitable for MR images. In this paper, we propose a lung nodule detection method for thoracic MR images based on deep learning. The development of magnetic resonance imaging (MRI) technique, helps to Reduce mortality. MRI is a non-radiation examination. MRI might be a valuable tool in malignant nodule detection and lung cancer screening. MRI has multiple modalities and each MRI scan has many slices. Computer-aided detection(CAD) for MR images is demanding.

For CT nodule detection, most deep learning based methods consist of two steps. First nodule candidate was extracted and cropped. Then a deep learning network was designed to classify the candidate as nodule or normal. As the size of lung nodule differs a lot, cropping each region with a fixed size may be not reasonable. To solve this problem, Faster R-CNN is designed for lung nodule detection in this paper.Faster R-CNN takes the whole image as input and no candidate extraction is needed. Pulmonary nodule detection is basis for nodule measurement and classification. Nodule detection methods consist of four steps: pre processing, lung parenchyma segmentation, nodule detection and false positive (FP) reduction. Thorax and lung parenchyma extracted first based on region growing .Dense structures inside the lung parenchyma were selected. A support Vector machine (SVM) classifier was trained to classify the dense structure as nodule or FP. Most deep learning based methods consists of two steps, first nodule candidate was extracted and then classify the candidate as nodule or normal. Faster R-CNN is designed for lung nodule detection.

III. LITERATURE SURVEY

[1] **Y. Li, H. Chen, and X. Wei, "An Automated Lung Parenchyma Segmentation Method for 3D Thoracic MR Images"**In this paper, an automated lung segmentation method was proposed for thoracic magnetic resonance (MR) images.For each MR slice, testing lung parenchyma exist or not is enough. Thus a rough lung parenchyma segmentation method is implemented. For the first step, the thorax was segmented combining thresholding and shape analysis.The region outside the thorax usually has very low intensity, which makes thresholding effective to segment the thorax. 3D Otsu threshold and shape analysis are combined to get the thorax region. First, 3D Otsu is used to find the optimal threshold for an MR scan. Then, the 2D hole-filling operator is applied slice by slice. Finally, slice with incorrect segmentation is found based on Hausdorff distance and re-segmented using iterative thresholding. For the second step ,the air-filled regions were detected based on k-means clustering. For the third step, the anatomical characteristics were employed to separate the initial lung parenchyma from the air-filled regions. As the surrounding tissues may connect to the lung, a contour refinement step was finally designed.

[2] C. Li, G. Zhu, X. Wu, et al., "False-Positive Reduction on Lung Nodules Detection in Chest Radiographs by Ensemble of Convolutional Neural Networks"They proposed Ensemble of convolutional neural networks (E-CNNs) used to significantly reduce the number of false-positive on lung nodules detection in chest radiographs (CXRs). With the rapid development of deep learning, convolutional neural network (CNN) has made big progress, especially in image classification and object detection .The CNN is used for detecting the lung nodules. Firstly the nodule signals are enhanced thenthe patches are cut from the enhanced CXRs with the slide - window method, and fed into the CNN modules to detect the lung nodules.The CNN is a highly nonlinear filter that can be trained by using input images and the corresponding teaching labels. It always consists of multi convolutional layers, pooling layers and fully connected layers. One single CNN has limited learning capacity, and may not learn all the essential features to distinguish a nodule from various types of non-nodule structure, but multi different CNNs can deal with much more non-nodules. So the Ensemble of CNNs (E-CNNs) is used to reduce the false-positives in this paper .With different scales of subsample of the patches, the proposed E-CNNs model deal with the non -nodule patch differently. When all the patches can be marked as the true positive, the false positives can be completely reduced.

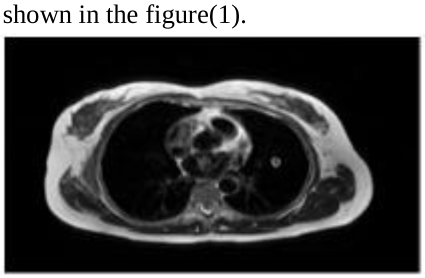
[3] **[4] Q. Dou, H. Chen, L. Yu, et al., "Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection"**False positive reduction is one of the most crucial components in an automated pulmonary nodule detection system, which plays an important role in lung cancer diagnosis and early treatment. The objective of this paper is to effectively address the challenges in this task and therefore to accurately discriminate the true nodules from a large number of candidates. A novel method employing 3D convolutional neural networks (CNNs) for false positive reduction in automated pulmonary nodule detection. In general, each 3D convolutional network consists of 3D convolutional, 3D max-pooling and fully-connected layers to hierarchically extract representations (also called features) and a softmax layer for the final regression to probabilities. Each layer contains a number of channels, and every channel encodes a different pattern. For 3D CNN, each channel in the convolutional /max-pooling layer is actually a 3D feature volume, rather than a 2D feature map in conventional CNNs. The 3D feature volume includes a group of neurons structured in a cubic manner. The 3D CNNs can encode richer spatial

information and extract more representative features via their hierarchical architecture trained with 3D samples.

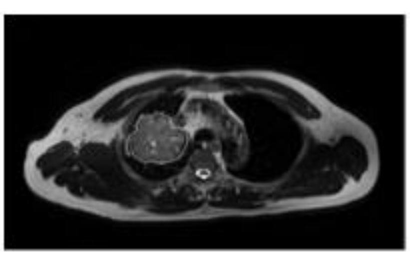
[5] H. Jiang, H. Ma, W. Qian, et al., "An automatic detection system of lung nodule based on multi-group patch-based deep learning network"In this paper, a volume of detailed tissues was contained in lung parenchyma including nodules, vessels, or other impurities. Notably vascular morphology was so conspicuous inside the lung that they may affect the detection performance of lungnodules. In lung CT images, all the pixels of vessels and nodules emerged as bright tissues among the darker adjacent pixels. The morphology and structure of vessels were distinct from those of lung nodules. In general, vessels looked like tubular structures, whereas lung nodules appeared as ellipse, irregular sphere or cotton-like structures. The aim of this step was to eliminate the vascular structures in the lung, thus they can be beneficially analyzed as nodule-like structures. Some mature vessel enhancement algorithms such as Sato filter, vessel enhancing diffusion (VED) is proposed here. An effective lung nodule detection scheme based on multi-group patches cut out from the lung images, which are enhanced by the Frangi filter.

IV. SYSTEM DESIGN

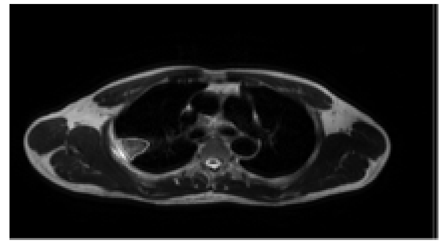
The data used here is T2-weighted MR scans, all MR scans were performed on 3T MR imager . Dataset includes 142 T2W-MR scans. Each scan consists of 13 to 33 slices with a 7.7mm spatial resolution along the axial direction. The total number of slices is 3403,in that 800 slices are with lung nodule Some slices contain more than one lung nodule region, resulting in a total number of 862 nodule regions. According to the position, nodules can be divided are shown in the figure(1).



Isolated lung nodule



juxta vascular nodule



juxta pleural nodule

Figure.1.Examples of lung nodules

Faster R-CNN : Faster R-CNN is composed of two modules. The first module is a Region Proposal Network (RPN), which can generate proposed region for each image. The second module is a Faster R-CNN detector, which classifies region proposals. RPN and fast R-CNN share convolution layer. TheRPN and Fast R-CNN share convolutional layers. In RPN, the convolution layers of a pre-trained network are first followed by a 3×3 convolutional layer. Then two 1×1 convolutional layers are added for classification and regression, respectively. To deal with different scales and aspect ratios of objects, anchors are introduced in the RPN .To deal with different scales and aspects , anchors are introduced in the RPN.Each anchor is associated with a scale and an aspect ratio. The default setting of anchors are 3 scales (128×128, 256×256 and 512×512 pixels) and 3 aspect ratios (1: 1, 1: 2,and 2: 1), leading to k = 9 anchors at each location. A nodule generally occupies a small portion in the image. Thus the anchor scale should be adjusted.

Feature extraction models: Different feature extraction models can be used in Faster RCNN. In this paper, three of the most impressive CNN models are tested. One is VGG16. The other two are residual network (ResNet) with different architectures. ResNet uses a residual learning framework to ease the training of deep networks. It takes a standard deep CNN and add shortcut connections which bypass few convolutional layers at a time.

VGG16: VGG16 is composed of 13 convolutional layers and 4 pooling layers. For each convolutional layer, the kernel size is 3×3 and the stride is 1. For each pooling layer, the kernel size is 2×2 and the stride is 2.

ResNet50: ResNet50 is one of the residual models. The architecture of ResNet50 is shown in Table 1. When ResNet50 is used as the feature extraction model in Faster R-CNN, the output of Conv4_x is the feature map.

ResNet101: ResNet101 is another architecture of ResNet, shown in Table 1. It is the same with ResNet50, except the

Conv4_x layer, use the output of Conv4_x as the feature map in Faster R-CNN

Detection using faster R-CNN: Directly using MR images to train the parameters of feature extraction model will get overfitting. To solve this problem, transfer learning is employed. First a pre-trained model from natural images is employed as the initial parameters. Then fine-tuning is performed to get the final parameters value.Natural image is a three-channel image, including R, G and B channels. A T2W-MR slice is one channel gray level image. To fine-tune the existing trained models, we should construct a three-channel input for each MR slice. One construction manner is using the gray level image in each channel and the three channels are the same. In this construction, each slice in one MR scan is separately treated and the spatial information is discarded. Spatial information is discarded, air artifacts may be easily detected as lung nodule, shown in Figure(2).Compared with lung nodule, air artifacts show different characteristics.In one situation, air artifact can only be seen in one slice.In the other situation, air artifacts in consecutive slices show different shapes at the same position. Based on the difference between air artifact and nodule, a three-channel input adding spatial information is designed. In the spatial three-channel input, one channel is the gray level image. The other two channels are its two consecutive slices. For the top or the bottom slice, the two consecutive slices are the following two slices or the upper two slices. For the other slices, the two consecutive slices are one following slice and one upper slice. Thus spatial information is added in the input image.

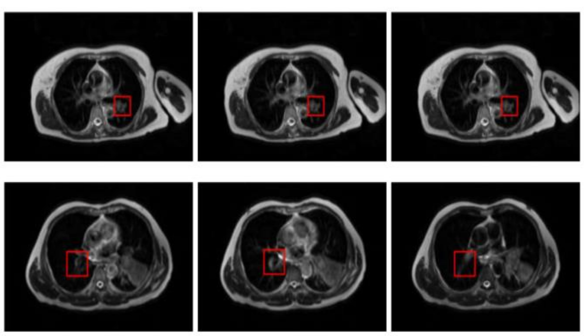


Figure 2. Examples of air artifacts

FP Reduction: Though spatial three-channel input is constructed for Faster R-CNN, many false positives (FPs) still exist. One type of FP is the tissue outside the lung parenchyma, indicated by the rectangle regions ,and the other type of FP is the air artifact region inside the lung parenchyma, indicated by the rectangle region as shown in the Fig. 3(b). An FP reduction scheme based on the anatomical characters of lung nodule is proposed. Lung nodules has two common characteristics:

1. Larger lung nodules can be seen in consecutive slices.
2. Lung nodules are inside the lung parenchyma.

For the first characteristic, two elements are important. The first element is deciding each region Rj belong to large or small. The second element is how to judge one large region is consecutive or not. For the first element, a fix threshold Ts is used to justify Rj belong to large or small. For the second element, the area overlap ratios between Rj and its corresponding regions in the two consecutive slices are used. If these two area overlap ratios are both less than Tr, then Rj is regarded as large isolated region and should be removed. Thus the first criterion can be defined below.

Criterion one: If a detected region with an area larger than Ts does not show larger area overlap ratios with its two consecutive slices, this region should be removed. For the second characteristic, the criterion to remove FP is to justify if the detected region is inside the lung parenchyma or not. However for lung with a juxta heart nodule or a juxta pleural nodule, the lung parenchyma may not be segmented as a whole. In this situation, the nodule is not inside the lung parenchyma and will be removed.

Criterion two: If a detected region displays in a slice without lung parenchyma, this region should be removed.

Lung parenchyma segmentation: Lung parenchyma segmentation method is implemented here .Firstly the thorax is segmented. Then the lung parenchyma candidate is extracted in each slice. The region outside the thorax usually has very low intensity, which makes thresholding effective to segment the thorax.For the MR slices , the thorax segmentation results are indicated as red boundary shown in the below figure. To get the lung parenchyma candidate ,each MR slice is segmented by low threshold T.

$$BW_{i,j} = \begin{cases} 1, & \text{if } I_{i,j} < T \\ 0, & \text{otherwise} \end{cases}$$

where i,j is the gray value of pixel(i,j). T is the threshold.

In parenchyma candidate ,some non lung tissues still remain ,hence 3D connected component analysis is made to remove these tissues. Treating the MR scan as a whole,3D connected component labelling is applied. The lung volume V, is selected from labelled volumes depending on the volume size .It is expressed by,

$$V = \begin{cases} V1 \cup V2, & \text{if } V2 > 0.5\,V1; \\ V1, & \text{otherwise} \end{cases}$$

where V1,V2 are the largest and second largest volumes among the labeled volumes, V is used as final segmentation result shown in Figure(3)

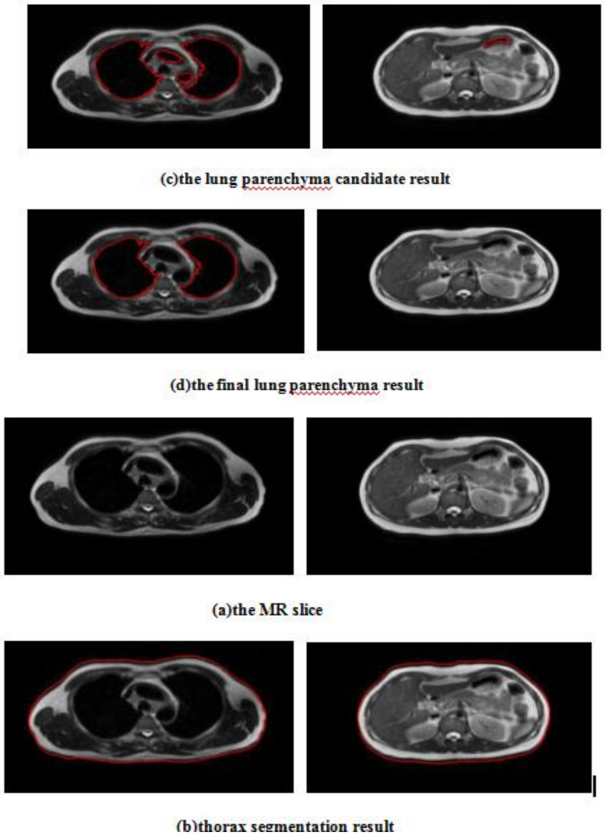


Figure 3.Lung parenchyma Segmentation

FP Reduction Procedure: The whole FP reduction scheme is detailed as follows and depicted in Figure(4).

Step 1: If no lung parenchyma exist in slice i, then the detected regions in slice i are regarded as FPs and removed. If slice i contains lung parenchyma, then go to Step 2.

Step 2: For each region Rj in slice i, if its area is less than Ts, Rj will be reserved. If its area is larger than Ts, then go to Step 3. In this experiment, Ts=500.

Step 3: Get the detection result in slice i-1 and slice i+1,Si-1 and Si+1. Compute the area overlap ratios between Rj and Si-1, and Rj and Si-1. If these two area overlap ratios are both less than Tr, then Rj is regarded as FP and removed. In this experiment, Tr=0.2. This procedure can remove FPs.

V. RESULT ANALYSIS

PERFORMANCE OF FEATURE EXTRACTION MODEL In order to test which model is more suitable, VGG16 , ResNet50 and ResNet101 are compared. The training set is divided into two parts. One part is used for training Faster R-CNN and the scan size is 68. The other part is used to get the validation result and the scan size is 29. The FROC curves for different feature extraction models under the validation dataset are shown in figure 5.

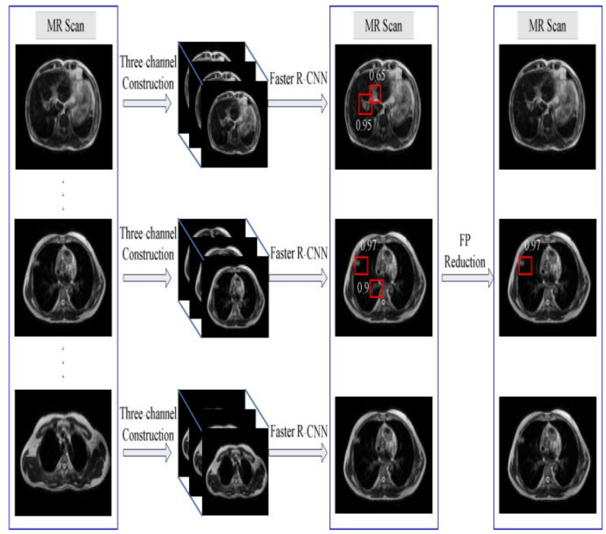


Figure 4.The whole nodule detection procedure for a MR scan

It can be seen that ResNet101 module gets the best performance compared with the other two modules. ResNet101 is the deepest model among the three feature extraction models. A deeper feature extraction model can get deeper features and has larger receptive field. Larger receptive field may lead to better detection performance.

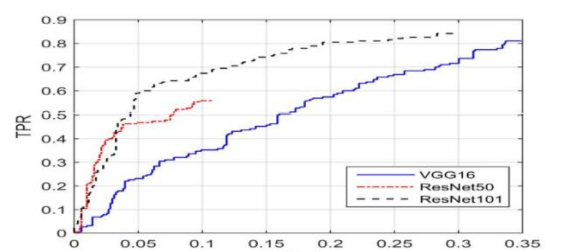


Figure 5. FROC curves for different feature extraction models.

Adding spatial information, to eliminate the influence of air artifacts, the spatial three channel input is designed for Faster R-CNN. To validate the effectiveness of spatial three-channel, another input using gray level image in each of the three channels is compared. This three-channel input is called gray three channel. For these two inputs, separate Faster R-CNN networks under ResNet101 are trained using the whole training set. The image-based FROC curves are drawn shown in figure 6. It can be seen that spatial three-channel designed in this paper gets better lung nodule detection performance compared with gray three-channel. At 75%, 80% and 85% sensitivity, FPs/I for spatial three-channel are 0.17, 0.19 and 0.28. For gray three-channel, the corresponding FPs/I are 0.19, 0.23 and 0.31. The performance difference between the two inputs is statistically significant(p=0.017<0.05) by t-test analysis.

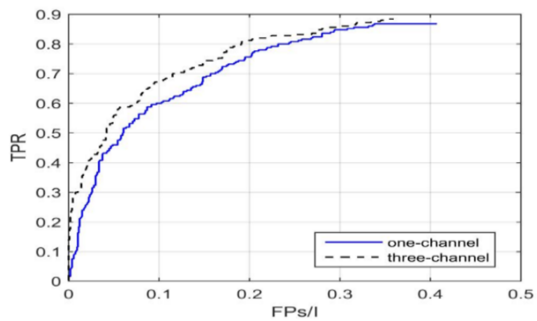


Figure 6. FROC curves for different input constructions

The performance of the proposed FP reduction scheme is tested. Two criteria are proposed to remove FPs. Let C1 represent Criterion one and C2 represent Criterion two. The image-based FROC curves for Faster R-CNN, Faster R-CNN with C1 and Faster R-CNN with C1+C2 are drawn in the below figure. It can be seen that the designed FP reduction scheme gets better lung nodule detection performance compared with Faster R-CNN. Using t-test analysis, the performance difference between Faster RCNN with C1+C2 and Faster R-CNN is statistically significant (p<0.05). The performance difference between C1 andC1+C2 is also tested. The figure shows that C1+C2 shows better detection result and the improvement is statistically significant (p<0.05). It may be concluded that different FP regions are separately removed by C1 and C2 criterion as shown in figure 7.

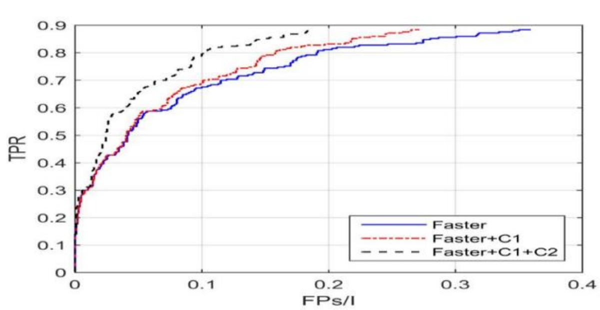


Figure 7. FROC curves for FP reduction schemes

VI. CONCLUSION

In this article, a lung nodule detection method for thoracic T2W-MR images is proposed based on deep learning. The proposed method takes the whole image as input and can detect nodules with different sizes and shapes. With optimized parameter, spatial three-channel input and transfer learning, Faster R-CNN is designed for lung nodule detection. This detection scheme can avoid candidate extraction and be less dependent on scale. As Faster RCNN does not consider anatomical characteristics, many FP regions exist in the detection results. To reduce FPs and preserve true nodule, an FP reduction scheme based on the anatomical characteristics of lung nodule is designed. Experimental results show that the designed Faster R-CNN can detect most of the nodules and the proposed FP reduction scheme can obviously reduce FP regions. It can be concluded that FP reduction scheme based on anatomical characteristics may be better than image features in Faster R-CNN detection.

REFERENCES

[1] Yanfeng Li, Linlin Zhang, Houjin Chen1, and Na Yang, " Lung Nodule Detection with Deep Learning in 3D Thoracic MR Images",2019.

[2] Y. Li, H. Chen, and X. Wei, "An Automated Lung Parenchyma Segmentation Method for 3D Thoracic MR Images ",2017.

[3] S. Ren, K. He, R. Girshick, et al., "Faster R-CNN: towards real-time object detection with region proposal networks," IEEE Trans.Pattern Anal. MachIntell., vol. 39, no. 6, pp. 1137-1149, 2017.

[4] C. Li, G. Zhu, X. Wu, et al., "False-Positive Reduction on Lung Nodule Detection in Chest Radiographs by Ensemble of Convolutional Neural Networks," IEEE Access, vol. 6, pp. 16060-7,2018.

[5] Q. Dou, H. Chen, L. Yu, et al., "Multilevel contextual 3-d cnns for false positive reduction in pulmonary nodule detection," IEEE Trans. Biomed. Eng.vol. 64, no. 7, pp. 1558-1567, 2017.

[6] H. Jiang, H. Ma, W. Qian, et al., "An automatic detection system of lung nodule based on multi-group patch-based deep learning network," IEEE TransInf. Technol. B., vol. 22, no. 4, pp. 1227-37,2017.