bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

An automatic and efficient pulmonary nodule detection system based on multi-model ensemble

Jian Chen[1☯], Wenlei Wang[2☯], Bin Ju[2☯], Jing Jiang[2], Liying Zhang[2], Huilin He[1], Xuanxuan Zhang[1], Yuqiang Shen[1*]

1 Zhejiang University School of Medicine Fourth Affiliated Hospital, Yiwu, 322000, China
2 The Wowjoy Technology, Hangzhou, 310006, China

☯These authors contributed equally to this work.
* 8016180@zju.edu.cn

Abstract

Accurate pulmonary nodule detection plays an important role in early screening of lung cancer. Although there are many presented CAD systems based on deep learning for pulmonary nodule detection, these methods still have some problems in clinical use. The improvement of false negatives rate of tiny nodules, the reduction of false alarms and the optimization of time consumption are some of them that need to be solved as soon as possible. In view of the above problems, in this paper, we first propose a novel full convolution segmentation framework for lung cavity extraction in preprocessing stage to solve the time consumption problem of the existing pulmonary nodule detection systems. Furthermore, a 2D-NestedUNet segmentation network and a 3D-RPN detection network is stacked to get the high recall and low false positive rate on nodule candidate extraction, especially the recall of tiny nodules. Finally, a false positive reduction method based on multi-model ensemble is proposed for the further classification of nodule candidates. Our methods are evaluated on several public datasets, LUNA16, LNDb and ChestCT2019, which demonstrated the superior performance of our CAD system.

Introduction

Lung cancer is a leading cause of cancer-related death both in men and women. Every year lung cancer results in about 170 million deaths worldwide. The early diagnosis of a lung lesion is recognized as the most important method to reduce the lung cancer mortality rate. The low-dose CT scans which is one diagnostic way can be used to screen for lung cancer in people. Using this screening method can decrease the risk of dying from lung cancer. Now researchers are looking for new attempts to refine CT screening to better predict whether cancer is present [1]. Due to the rich pulmonary vascular structure and the different skill levels of radiologists, the potential malignant lesions are easy to be ignored. In the clinic, an effective way to deal with this problem is to diagnose by two radiologists respectively and then to summarize their answers. However, this may increase the workload of radiologists [2]. With the development of digital medical image processing, using computer-aided diagnosis (CAD) system to assist pulmonologists in clinical diagnosis is a trend. However, the traditional approaches are less accurate and more complicated in detecting pulmonary nodules at early stage of lung cancer so that are hardly applied to help doctors in clinical diagnosis.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

In recent years, with the rapid development of deep learning in the field of medical image analysis, a large number of CNN-based CAD systems have been utilized for the detection of pulmonary nodules [3–5]. Compared with the traditional methods [6, 7], the DNN-based methods have made great progress in accuracy and practicability of pulmonary nodule detection. Nevertheless, many challenges still exist in the lung nodule detection procedure. The detection procedure is usually divided in three stages, i.e. image prepossessing, nodule candidates extraction and false positives reduction. Firstly, in prepossessing step, how to quickly and efficiently extract lung cavity plays an essential role. Due to the various sizes and morphology of pulmonary cavity, the traditional multi-threshold-values based segmentation is not robust enough. Therefore, offering an efficient lung cavity extraction method is very important. Secondly, in clinical practice, radiologists pay more attention to small nodules, because these small nodules are more likely to cause lung cancer in the future. To ensure high sensitivity for them, in other words, CAD is needed to have a high recall rate for small nodules that may better resolve lung cancer early screening. However, the size of small nodules is too small and their radiographic manifestations always appear ground glass attenuation so that it is likely to be neglected. The presented methods, neither using 2D nor 3D system, can find small pulmonary nodules well [3, 4]. That means the developers need invest more research in this area in the future. Thirdly, because abundant tissues exist in the human pulmonary cavity, for instance blood vessel and chest wall, the appearances of these tissues are very similar with the appearance of the pulmonary nodules. This results in producing amounts of false alarms during detection. Making something to accurately distinguish between tissues and nodules thereby reducing the number of false positive signals is very important and necessary [8].

The primary aim of this study is to develop an advanced CAD system that extracts information from medical images efficiently and provides radiologists a precise and timely diagnosis of lung lesion. The key contributions of this paper are summarized below:

- we present a FCNN framework which is based on U-Net [12] for quickly and stably performing pulmonary cavity extraction.
- we stack a 3D-RPN based detection network [3] and a 2D-NestedUNet based segmentation network [13] for providing target candidates to satisfy high recall rate and low false positive rate of pulmonary nodule detection.
- we propose an integrated classification network which consists of ResNet [23], DenseNet [24] and SENet [25], summarizing the results from detection step and classification step to reduce false candidates.

Related Work

Before the rapid development of CNN in medical image processing, the pulmonary nodules detection was mostly based on hardcrafted features. In recent years, with the successful application of deep learning in medical images, the intelligent screening system of pulmonary nodules has been greatly developed. Based on the exploited deep architecture, these approaches can be divided into proprecessing, candidate nodules extraction and false positive reduction. In [3, 4, 14], the pulmonary cavity of the CT scans was obtained by traditional image processing methods such as region growth and morphological operations. [15] proposed the spot detection for extracting the nodule candidates. Using this method, nodule and non-nodule samples could be distinguished according to their different shape, size, specific texture, density and other features. However, due to the choices of thresholds, the detection performance of this method for

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

small nodules and tiny nodules was very poor and not robust. In recent years, many excellent CNN-based CAD systems for automatic pulmonary nodules detection are presented, which can roughly be divided into 2D image slices and 3D volume images. In [3], 2D image slices were fed into Faster-RCNN [16] to locate suspected nodules and output their sizes in CT images. First from the perspective of 3D volume images, [4] established a pulmonary nodules detection system with the aid of 3D-Faster-RCNN framework. This method could make full use of the spatial information of 3D images. In [17], the attention mechanism is brought into the approach presented in [4] to further acquire 3D spatial semantic information. In [18], a new self-supervised pulmonary nodule detection framework is proposed based on 3D-RPN to improve the sensitivity of nodule detection by adopting multi-scale features to increase the resolution of nodule cubes. Compared with the 2D models, these 3D methods employ spatial semantic information to better analyze the morphological characteristics of nodules and locate them much accurately. However, their recalls for small nodules are not high so that would be easy to increase the false negatives rate for small nodules and tiny nodules. [19] presented a 3D full convolutional network which is based on V-Net [20] to extract nodule candidates. A high recall is the advantage of this solution. For reducing more false positives, [4, 6] proposed a simple 3D convolutional network adding the back of detection module to filter false candidates. [21] introduced the attention mechanism after false positive reduction module to facilitae the performance of classification for pulmonary nodules (benign and malignant).

Methods

Our proposed CAD system can be roughly divided into three stages: 1) FCNN based lung segmentation, 2) multi-model ensemble based pulmonary nodule extraction, and 3) false positive reduction (the whole pipeline shown in Fig 1).

Fig 1. The Overview of Our CAD system. A whole CT scan is fed into our system to predict nodule candidates. The process consists of three stages, i.e. preprocessing, pulmonary nodules extraction and false positive reduction. In first stage, raw data will be processed to extract lung cavity and then fed into second stage to find pulmonary nodule candidates. Finally, the candidates will be filtered in stage 3 to reduce false alarms.

Novel FCNN for Lung Segmentation

Accurate lung segmentation is the basics for rapidly constructing automated pulmonary nodule detection system with high-accuracy and sensitivity. By using the traditional method, the 2D single slice is first processed with a Gaussian filter to remove the fat, water and kidney background and then followed with a 3D connection region extraction module to remove irrelevant regions. However, this method is time consuming and unstable. In order to accelerate this processing for experiment and deployment, this paper propose a lung segmentation network, which is based on U-Net. The encoder path includes five consecutive convolution modules, the number of feature maps is doubled at each convolution module. Each convolution module consists of two convolution filters with kernel_size = 3, followed by a max pooling function with kernel_size = 2 and stride = 1. Each convolution filter is followed with a BatchNormalization and a nonlinear activation function ReLU. Finally, the final layer in the encoder path produces a high dimensional image representation with high semantic information.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Assume that the input of the segmentation network is a single-channel grayscale image with a size of 1×H×W, denoted as I_{in}, and the output of each convolution module i^{th} in the encoder path is denoted as X_e^i, which represents the number of output channels, each convolution module i^{th} takes as input to X_e^{i-1}. At the end of the encoder, a feature map F_e will be obtained whose size is reduced by 4 times compared to I_{in}.

Different from the traditional U-Net structure, we make some modification for our FCNN in the convolution mode of the decoder path, which is to use the same convolution structure as the encoder path to implement feature decoding, and at the end of the decoder, a convolution filter with kernel_size = 1×1 used as the network output layer. The entire network will eventually output a segmentation prediction result with the same size as the input image. The predicted value of each pixel represents the likelihood that this pixel belongs to a lung cavity. The overall lung cavity segmentation network is shown in Fig 2, using the Dice coefficient as loss function which is same to original U-Net [12].

Fig 2. The Schematic of Our FCNN framework for Lung Segmentation.

Nodule Candidate Extraction based on Multi-model Ensemble

In this section, we design a multi-model ensemble network, which is combined with a 2D segmentation network and a 3D detection network to sensitively screen the nodule candidates.

2D NestedUNet for Pulmonary Nodule Extraction

NestedUNet [22] can accelerate network optimization by using dense convolutional blocks which bridge the semantic gap between the feature maps of the encoder and decoder. However, we found that NestedUNet with depth supervision module is difficult to be optimized and easy to cause gradient explosion while training for pulmonary cavity segmentation in this work. Thus we use the fast version of NestedUNet with the same loss function from the original NestedUNet as our segmentation module, the specific formulas are summarized bellow:

$$Loss = \lambda Loss_{BCE} + Loss_{Dice} \tag{1}$$

$$Loss_{BCE} = \frac{1}{N}\sum_{n=1}^{N} r_n \log(p_n) + (1 - r_n)\log(1 - p_n) \tag{2}$$

$$Loss_{Dice} = 1 - \frac{\sum_{n=1}^{N} p_n r_n + \epsilon}{\sum_{n=1}^{N} p_n + \epsilon} - \frac{\sum_{n=1}^{N}(1 - p_n)(1 - r_n) + \epsilon}{\sum_{n=1}^{N}(2 - p_n - r_n) + \epsilon} \tag{3}$$

where p_n is the probability that pixel n is predicted as lung cavity, r_n is the true category of pixel n, $r_n = 1$ for the lung cavity and, $r_n = 0$ is regarded as backgound. $n \in \{1, 2, 3 \ldots N\}$, $N = H \times W$, ϵ is a minimal number.

3D RPN for Nodule Extraction

Pulmonary nodule detection, as a crucial step in the CAD systems, aims to accurately locate the nodule, meanwhile catching more true nodule candidates and reducing the number of non-nodule candidates as much as possible. We present a 3D RPN model for detecting nodule candidates from CT images, where a modified U-Net is the backbone network [4]. The structure of pulmonary nodule detection is shown in Fig 3. The

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

network consists of an encoder and a decoder. Due to the limitation of GPU memory, the input of the 3D RPN network is cropped from the 3D CT image. The size of the input image is 128×128×128. The encoder includes 3D convolutional layers, 3D residual blocks, and 3D maximum pooling layers. Each 3D residual block contains three residual units. The decoder has deconvolutional layers, combining units and residual blocks. Each combining unit concatenates the output of the convolutional layer and the corresponding deconvolutional layer, which likes the long connection in U-Net. In the left combining unit, the location information is introduced as an extra input, the feature map of this combining unit has the size of $32 \times 32 \times 32 \times 131$. Finally, the output feature is resized to $32 \times 32 \times 32 \times 3 \times 5$ after two convolutional layers. The last two dimensions correspond to the anchors and regression box respectively. There are three anchors with different scales, corresponding to three bounding boxes with the length of 3, 10 and 20 mm. The loss function is defined as below:

$$L = 0.5L_{cls} + pL_{reg} \tag{4}$$

where $p \in \{0, 1\}$ (0 for negative and 1 for positive). The classification loss L_{cls} is the cross-entropy loss and the regression loss L_{reg} is the smooth L1 loss function.

Fig 3. The Structure of 3D RPN for Nodule Detection. Each cube represents a 4D tensor. The size is $Length \times Width \times Height \times Channels$. The last two dimensions of the output tensor correspond to the three anchors and five regression box indicators (nodule coordinates x, y, z, nodule size r and probability).

Combination of Nodule Candidates

After a CT scan volume is fed to the aforesaid 2D and 3D pulmonary nodule extraction methods respectively, two lists of pulmonary nodule candidates will be provided, here denoted as L_1 and L_2. Then we can combine these candidates by using algorithm 1. Firstly, based on NMS algorithm, we remove the parts of candidates whose detection boxes have significant overlaps with others. After that, we delete the parts of candidates from 2D segmentation that have high overlaps with above 3D handled results also based on NMS. The 3D detection method has high accurate location as benefit point however its recall rate is relatively lower than 2D segmentation method. Meanwhile, using 2D segmentation network will provide a large number of false positive candidates. Combing the results of two different methods can help us to get accurate position of pulmonary nodules and avoid excessive false alarms. That is the reason why we use model ensemble strategy. Besides, based on CUDA platform, we use multi-thread technology to fast perform the calculation during the comparison process so that can dramatically reduce the time cost of fusion phase.

False Positives Reduction

As previously mentioned, nodule candidates from detection phase may contain many false alarms, although we have taken some measures to suppress this issue. To reduce the non-nodule candidates, we propose a 3D DCNN-based false positive reduction module which contains three network models. Fig 4 shows our architectural diagram for false positives reduction, the module is based on the 3D network of ResNet [23], DenseNet [24] and SENet [25]. For the center of a given nodule candidate, we extract a 3D data cube of size $32 \times 32 \times 32$ that possibly includes pulmonary nodule as input to the three network models, and then get the probabilities p_i (i $\in \{1, 2, 3\}$) of the three classifiers. To further eliminate false positives, we design the strategy described as

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Algorithm 1 Merging two nodules results extraction

```
L_1 and L_2 are two results
L_R is the result of the merge
L_2 = NMS(L_2)
for l_2 in L_2 do
    for l_1 in L_1 do
        if IOU(l_2, l_1) > thresh then
            remove(l_2)
        end if
    end for
end for
L_R = L_1 + L_2
```

follows: 1) for predicted probability p_i, if $p_i \geq T_1$, the final classify probability of the input cube is equal to the average value of all $p_i \geq T_1$; otherwise, the final result is the average value of all p_i. 2) multiplying the classify probability from 1) with the detection probability from detection step as the final predicted result of the CAD system, if the value $\geq T_2$, input will be denoted as true alarm; conversely, as false alarm. During the experiments, T_1 and T_2 are set as 0.5.

Fig 4. The Architectural Diagram for False Positive Reduction. 1) the input size of the false positive reduction module is $32 \times 32 \times 32$. 2) p1, p2 and p3 are the output probabilities of ResNet, DenseNet and SENet respectively. 3) the final probability is the average value of p1, p2 and p3.

Experiments

Dataset

In this paper we employed three lung nodule datasets in all.

- **LUNA16** Dataset [9] comes from LIDC-IDRI [27] which consists of 1018 cases from several institutions. The dataset excluded scans with a slice thickness greater than 2.5 mm. In total, 888 CT scans are included. For each CT scan, nodule candidates are provided and the corresponding class labels (0 for non-nodule and 1 for nodule) are also annotated. 1186 nodules are annotated across 601 patients.
- **LNDb** Dataset [10] contains 294 CT scans collected at the hospital CHUSJ in Porto, Portugal between 2016 and 2018. The dataset excluded scans with a slice thickness greater than 1 mm. All CTs were annotated by at least one radiologist. Finally, the dataset includes 1897 annotations by the 5 radiologists with 1429 corresponding unique findings.
- **ChestCT2019** Dataset [11] is proved by TIANCHI medical AI competition. We get 1397 samples with nodule from the CT scans of 1000 patients. The slice thickness of CTs is less than 2 mm. The biggest difference between ChestCT2019 and the two above datasets is that the slice spacings of CTs in ChestCT are almost all 5 mm.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Data Preprocessing

The HU values commonly observed in pulmonary nodules in CT scans are used. The Hounsfield scale of tissue density is based on two values: air as -1000 HU and water as 0 HU. Density of other tissues is related to this range, usually from -1000 to +400 HU for pulmonary CT image analysis. Because in pulmonary CT scan exist abundant tissues, in order to decrease the impact of other tissues and avoid losting more information details after we compress raw images into 256-bit grayscale images, we selected [-1350, 150] as the observation range. As the images were collated from different institutions using different CT scanners which results in a wide variation of image parameters within our training dataset, we rescaled the slice thickness of all used CT scans to 1.0 mm.

Implementation Details

Lung Segmentation

The training data of lung cavity segmentation comes from LUNA16. The original label contains five types, background, left lung, right lung, and nodules etc. For lung cavity segmentation, we first merged the label = 3, label = 4 and label = 2 as the lung cavity, label = 1 and label = 5 as the non-lung cavity, and the samples with false label were eliminated. Then, a single CT slice image was selected along the axial direction as input for training. To effectively reducing the over-segmentation rate, when selecting 2D slice images, we set the proportion of slice including and excluding the lung cavity to 7:3. Prevent model overfitting and enhancing its generalization ability, in data augmentation stage we taken some measures like 1) random flipping horizontally or vertically, 2) adding random Gaussian noise, and 3) standardization with the mean of 0 and the variance of 1. During the training, SGD with momentum of 0.9 and initiative learning rate 0.001 is used as optimizer to minimize the our loss function.

Pulmonary Nodule Extraction

a) *2D NestedUNet for Pulmonary Nodule Extraction*: The dataset that we use in this task is from LIDC-IDIR [27], because it contains the particular and relative precise description of pulmonary nodules compared with other datasets, which is helpful to build an accurate segmentation model. For making suitable data as input to network, a large number of patches are randomly cropped from 2D slices according to the center coordinates and sizes of the nodules, whose size is 256×256 and must contain complete nodules within the region. To reduce the false positives, we also select the patches which have no nodules for training. In addition, the approaches such as random horizontal and vertical flip, random shift from the crop center, random angle rotation, random color dithering and Gaussian smoothing are used to augment training samples. During the training, SGD with momentum of 0.9 and initiative learning rate 0.001 is used as optimizer to minimize the our BCEDiceLoss (binary cross entropy loss and dice coefficient loss combined with weight).

b) *3D RPN for Nodule Extraction*: As the 3D RPN network is deep, and has more parameters than 2D RPN, the model is easy to overfit on small dataset. To solve this problem, we use similar methods with above task to enlarge the number of training data. In Section we have discussed the significance of the location of tiny nodule in early pulmonary nodule detection. Hence we increase the sampling frequencies of tiny nodules in the training set. Specifically, the sampling frequencies of nodules larger than 10 mm and 20 mm are expanded 2 and 4 times higher than other nodules, respectively. The sampling frequency of tiny nodules smaller than 3mm also increases 2 times. Some of negative samples have similar appearances with pulmonary nodules, making them

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 1. Comparison with Different Pulmonary Nodule Detection Methods on Lung Database [28]

Methods	F1 score	Sensitivity	Specificity	Accuracy	Dice	Avg. per scan(s)
F. Liao [4]	0.9408	0.8753	**0.9994**	0.9841	0.9313	63
U-Net [12]	0.9658	0.9696	0.9872	0.9872	/	/
R2U-Net [29]	0.9832	**0.9944**	0.9832	0.9918	/	/
LGAN [30]	/	/	/	/	0.9850	/
Our FCNN	**0.9836**	0.9834	0.9985	**0.9965**	**0.9864**	7.7

difficult to be classified correctly. Hard example mining is a common technique to solve this problem.

False Positive Reduction

As for training data of the proposed networks, we firstly rescale each CT in [0, 255] using the specified windowing values, and then normalize it in [0, 1]. After that, for each selected candidate, we use the center of candidate as the centroid and then crop a $64 \times 64 \times 64$ cube. In the training data, there is an imbalance problem on the number of positive and negative samples. To solve the problem, we use the following strategies to augment data:

- Random crop: for each $64 \times 64 \times 64$ cube, we crop smaller cubes in the size of $32 \times 32 \times 32$ from it.
- Rotation: for each $32 \times 32 \times 32$ cube, we rotate it by 90°, 180°, 270° from three orthogonal dimensions (coronal, sagittal and axial position), thus finally augmenting 9 times for each chosen candidate.
- Multiple windowings: for each candidate, we use different windowing values (-1350 to 150 HU, 1200 to 600 HU and -1000 to 400 HU), thus augmenting 3 times.

The training procedure of three classification networks are almost identical. We use SGD with the momentum of 0.9 as optimizer to minimize the loss function and update the model parameters. The initial learning rate is 0.001 and decay is 0.9. The batch size is set to be 128. Due to the imbalance problem of positive and negative samples, we finally choose focal loss [26] as loss function. And the experiments proved that the effect of focal loss is better than cross entropy loss.

Results

Lung segmentation is very important for analyzing lung related diseases. In this experiment, we compare the performance of our FCNN model with the current methods on dataset [28], including the traditional benchmark method [4], U-Net model [12], R2U-Net [29], and LGAN [30]. 2D Dice score, F1 score and sensitivity were calculated with same settings for each method. The performance comparison is shown in Table 1. Compared with the state of the art methods, our FCNN model has the highest score, with an average F1 score of 0.9836, an accuracy of 0.9965, and an average Dice score of 0.9864. Although the traditional method had the highest specificity, it requires a series of thresholds, morphological manipulations, and composition analysis. Compared with this method, our method provides an end-to-end solution, which takes an average of 3.4 seconds per scan.

In order to visualize the well performance of our proposed architecture, we compare the predicted results of our FCNN on two different CT slices with the traditional

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig 5. Comparison with Traditional Method from F.Liao [4] and Our FCNN. The first column lists two different scans and their corresponding ground truth in second column. In the third column are the segmentation results by F. Liao and in last column are the results by our method.

Table 2. The Average Recall Comparison of Detection Models

Method	LUNA16	LNDb	ChestCT2019	Ave. per scan(sec)
DSB	74.40	20.05	6.94	83
DeepLung	83.40	20.83	13.31	91
Ours(only 3D-RPN)	84.54	34.77	14.05	**70**
Ours	**90.40**	**45.57**	**37.01**	147

method. Fig 5 shows the segmentation results. The first and second column shows the original input CT slice and the corresponding ground truth, and the other columns display the segmentation results from traditional method and our FCNN respectively. The significant improvement in lung segmentation by using our FCNN model can be observed.

We trained the detection model on LUNA16 trainset and evaluated the performance on the validation set of LUNA16. LNDb and ChestCT2019 are also evaluated. In addition, we compared the performance with the DSB and DeepLung models on the three dataset [4, 5]. DSB model from Data Science Bowl 2017 pulmonary nodule detection competition and DeepLung from LUNA16 competition. The average recall comparison of the detection models is shown in Table 2.

From the above table, we can see that our model ensemble method performs better on each data set than the others. The average recall of our model ensemble method on the LUNA16 is 90.4%, 45.57% on LNDB and 37.01% on ChestCT2019. The performance of the only 3D-RPN used method is poor than our complete method, but it takes the shortest time. In [4], the average recall of the DSB method model trained on the Data Science Bowl 2017 competition data set is 0.8562, but it is designed to neglect the very small nodules during training, the LUNA16 dataset is not suitable.

Conclusion

In this paper, we propose a high-performance pulmonary nodule detection system based on stacking of deep convolution networks. We propose a novel full convolution segmentation network based on U-Net for lung segmentation to solve the efficiency problem of the existing pulmonary nodule detection systems. Furthermore, a 3D-RPN detection network and a 2D-NestedUNet segmentation network is stacked to get the high recall and low false positive rate on pulmonary nodule extraction. Finally, a false positive reduction method based on multi-model ensemble is proposed for the further classification of nodule candidates. Extensive experimental results on public available datasets, LUNA16, LNDb and ChestCT2019, demonstrate the superior performance of our CAD system. We believe that our CAD system will be a very powerful tool for early diagnosis of pulmonary cancer.

Acknowledgments

The authors thank for the help of Basic Public Research Project of Zhejiang LY20H180013.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

1. Litjens G J, Kooi T, Bejnordi B E, et al. A survey on deep learning in medical image analysis. Medical Image Analysis. 2017; 60–88.

2. Lauritzen PM, Andersen JG, Stokke MV, et al. Radiologist-initiated double reading of abdominal CT: retrospective analysis of the clinical importance of changes to radiology reports. BMJ quality & safety. 2016; 25:595–603.

3. Ding J, Li A, Hu Z, et al. Accurate Pulmonary Nodule Detection in Computed Tomography Images Using Deep Convolutional Neural Networks. Medical image computing and computer assisted intervention. 2017; 559–567.

4. F. Liao, M. Liang, Z. Li, X. Hu and S. Song. Evaluate the Malignancy of Pulmonary Nodules Using the 3-D Deep Leaky Noisy-OR Network. IEEE Transactions on Neural Networks and Learning Systems. 2019; 3484–3495.

5. W. Zhu, C. Liu, W. Fan and X. Xie. DeepLung: Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). 2018; 673–681

6. Bin C, Takayuki K, Hirotoshi H, et al. Automatic segmentation of pulmonary blood vessels and nodules based on local intensity structure analysis and surface propagation in 3D chest CT images. International Journal of Computer Assisted Radiology and Surgery. 2012; 465–482.

7. Jibi John, M.G. Mini. Multilevel Thresholding Based Segmentation and Feature Extraction for Pulmonary Nodule Detection. Procedia Technology. 2016; 957–963.

8. S S Siegelman, N F Khouri, F P Leo, E K Fishman, R M Braverman, E A Zerhouni. Solitary pulmonary nodules: CT assessment. Radiology. 1986; 160(2):307–312.

9. Setio A A, Traverso A, De Bel T, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. Medical Image Analysis. 2017; 1–13.

10. Joao Pedrosa, Guiherme Aresta, Carlos Ferreira, et al. LNDb: A Lung Nodule Database on Computed Tomography.

11. TIANCHI competition 2019. Available from: `https://tianchi.aliyun.com/competition/entrance/231724/information`.

12. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. International Conference on Medical image computing and computer-assisted intervention. 2015; 234–241.

13. Zongwei Zhou, Md Mahfuzur R S, Nima T, Jianming L. UNET++: A Nested U-Net Architecture for Medical Image Segmentation. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. 2018; 3–11.

14. Kiraly A P, Higgins W E, Mclennan G, et al. Three-Dimensional Human Airway Segmentation Methods for Clinical Virtual Bronchoscopy. Academic Radiology, 2002; 9(10): 1153–1168.

15. Setio A A, Jacobs C, Gelderblom J, et al. Automatic detection of large pulmonary solid nodules in thoracic CT images[J]. Medical Physics. 2015; 42(10): 5642–5653.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

16. Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks. Neural information processing systems. 2015; 91–99.

17. Yuemeng Li, Yong Fan. DeepSEED: 3D Squeeze-and-Excitation Encoder-Decoder Convolutional Neural Networks for Pulmonary Nodule Detection. https://arxiv.org/abs/1904.03501

18. Jingya Liu, Liangliang Cao, Oguz Akin, Yingli Tian. Accurate and Robust Pulmonary Nodule Detection by 3D Feature Pyramid Network with Self-supervised Feature Learning. https://arxiv.org/abs/1907.11704

19. O. Ozdemir, R. L. Russell and A. A. Berlin. A 3D Probabilistic Deep Learning System for Detection and Diagnosis of Lung Cancer Using Low-Dose CT Scans. https://arxiv.org/abs/1902.03233

20. Milletari F, Navab N, Ahmadi S, et al. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. International Conference on 3D Vision. 2016; 565–571.

21. Mundher Al-Shabi, Lee Hwee-Kuan, Maxine Tan. Gated-Dilated Networks for Lung Nodule Classification in CT scans. IEEE Access. 2019; 178827–178838.

22. Eunwoo Kim, Chanho Ahn, Songhwai Oh. NestedNet: Learning Nested Sparse Structures in Deep Neural Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018; 8669–8678.

23. Kaiming He, Xiangyu Zhang, Shaoqing Ren, et al. Deep Residual Learning for Image Recognition. CVPR 2016; 770–778.

24. G. Huang, Z. Liu, L. v. d. Maaten and K. Q. Densely Connected Convolutional Networks. CVPR 2017; 4700–4708.

25. Jie Hu, Li Shen, Gang Sun. Squeeze-and-Excitation Networks. CVPR 2018; 7132–7141.

26. Tsung-Yi L, Priya G, Ross G, Kaiming H, Piotr D. Focal Loss for Dense Object Detection. ICCV 2017; 2980–2988.

27. Armato S G, Mclennan G, Bidaut L, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans. Medical Physics. 2011; 38(2):915–931.

28. Kaggle competition 2019. Available from: https://www.kaggle.com/kmader/finding-lungs-in-ct-data.

29. Alom M Z, Hasan M, Yakopcic C, et al. Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation. https://arxiv.org/abs/1802.06955

30. Tan J, Jing L, Huo Y, et al. LGAN: Lung Segmentation in CT Scans Using Generative Adversarial Network. https://arxiv.org/abs/1901.03473

bioRxiv preprint https://doi.org/10.1101/2020.04.14.040931; this version posted April 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure1

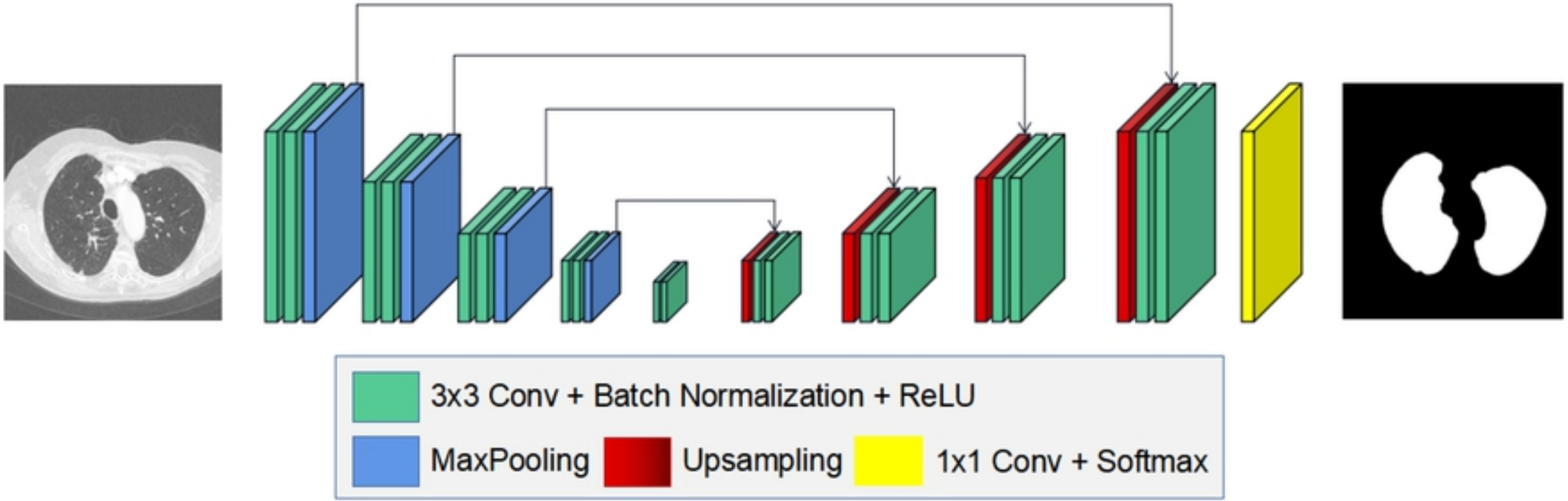

3x3 Conv + Batch Normalization + ReLU
MaxPooling
Upsampling
1x1 Conv + Softmax

Figure2

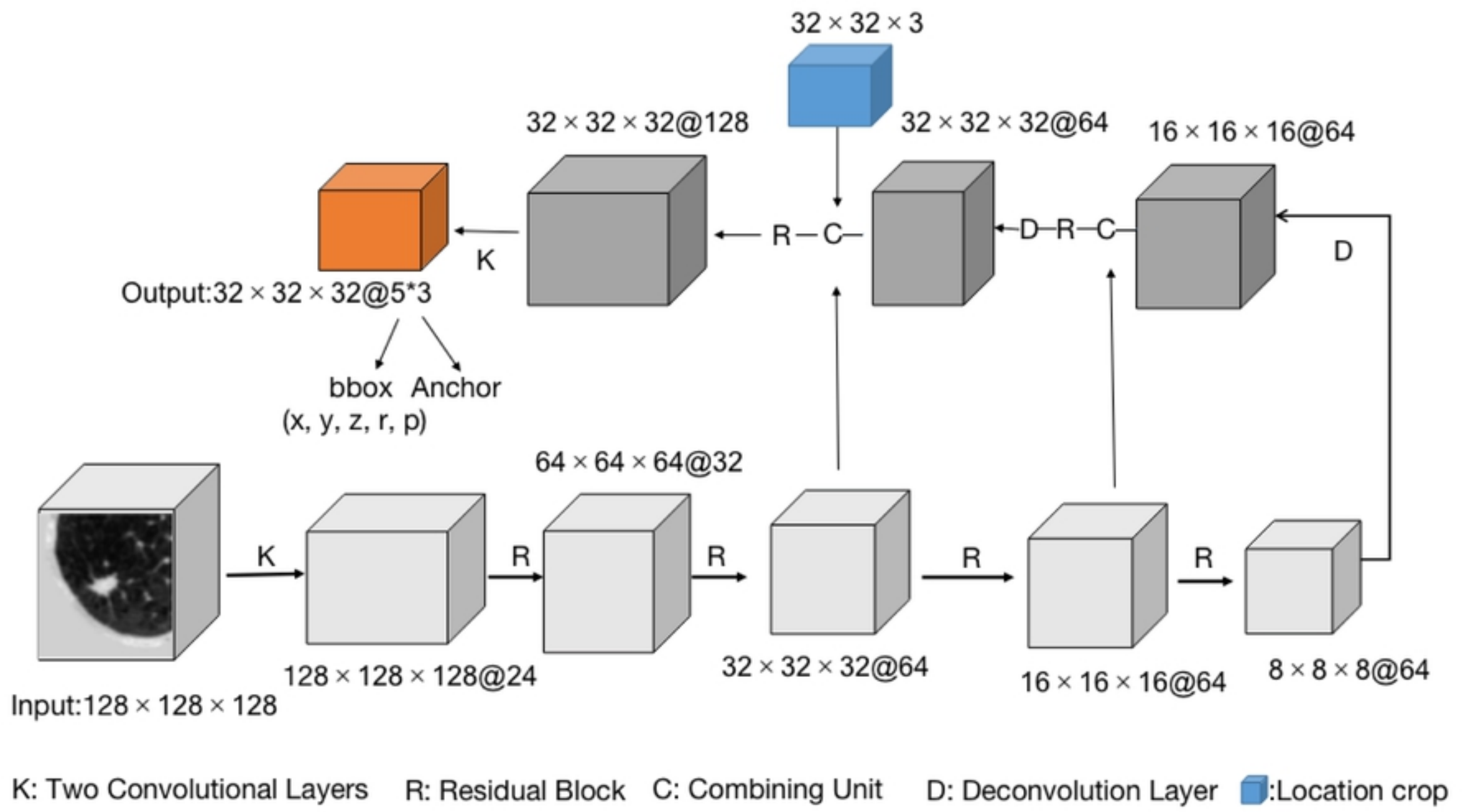

32 × 32 × 3
32 × 32 × 32@128
32 × 32 × 32@64
16 × 16 × 16@64
R
C
D
R
C
D
K
Output:32 × 32 × 32@5*3
bbox
Anchor
(x, y, z, r, p)
64 × 64 × 64@32
K
R
R
R
R
Input:128 × 128 × 128
128 × 128 × 128@24
32 × 32 × 32@64
16 × 16 × 16@64
8 × 8 × 8@64
K: Two Convolutional Layers
R: Residual Block
C: Combining Unit
D: Deconvolution Layer
:Location crop

Figure3

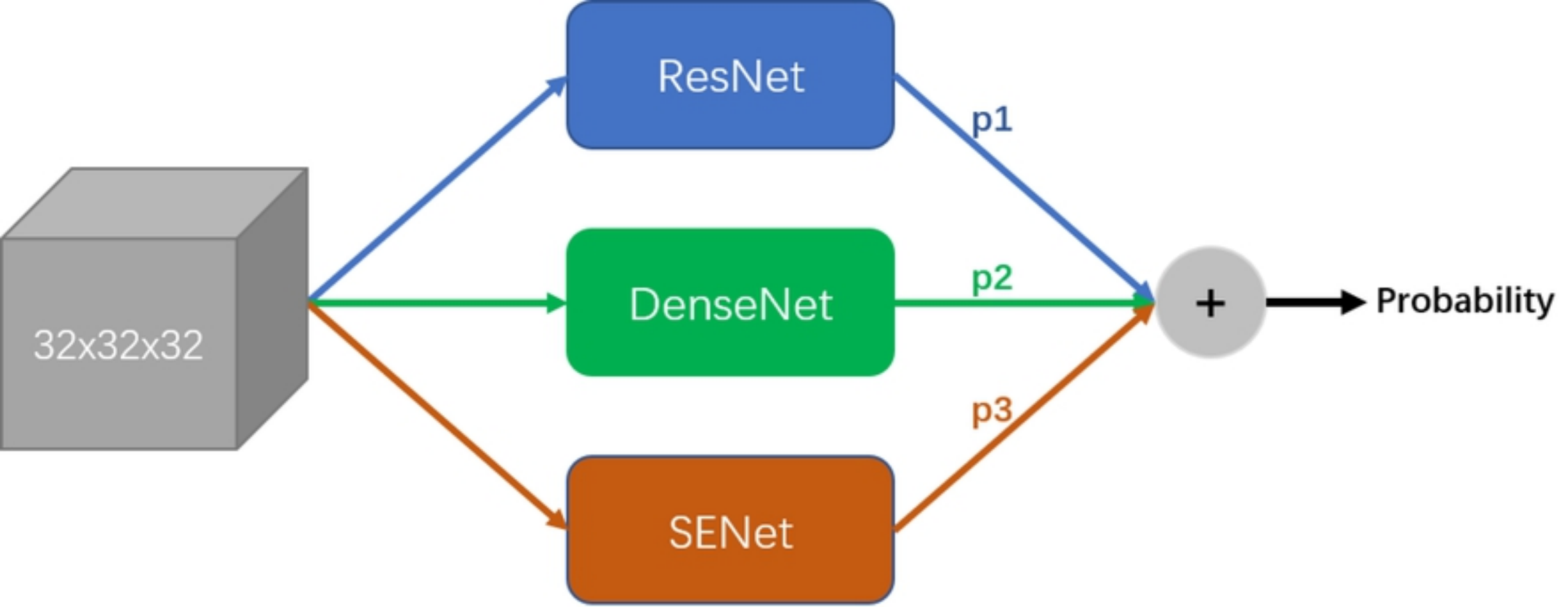


Figure4

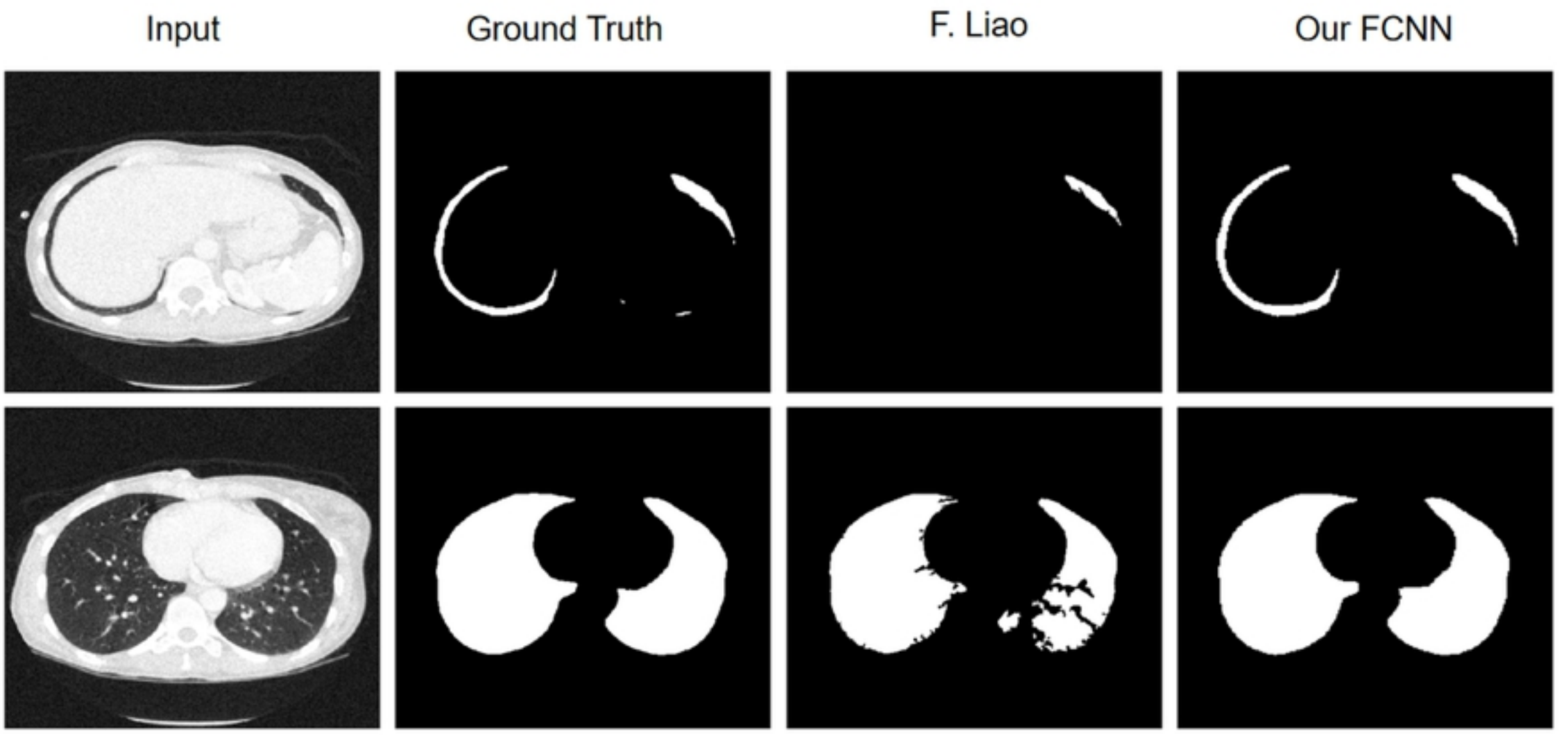


Figure5