Detection and Classification of Pulmonary Nodules Using Convolutional Nerual Network

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—In this work, we present a computer aided diagnosis(CAD) system to help the diagnosis of pulmonary nodules. There are two parts in our CAD system, detection of pulmonary nodules and classification of malignant and benign nodules. We use 3D V-net in detection task and reaches 65% accuracy on LUNA16 DATASET. Based on VGGNet and ResNet, we propose ResVGG in classification task and its accuracy reaches 98.73%. Extensive experimental results demonstrate reliability of our CAD system.**

***Index Terms*—pulmonary nodule detection and classification , deep convolutional neural network**

I. Introduction

Lung cancer is one of the cancers with the highest incidence in the world. Due to the deterioration of the environment and the increase in the number of smokers, the incidence of cancer in various countries has risen rapidly in recent years. Early treatment of lung cancer can effectively reduce the occurrence of late complications and recurrence, and can greatly improve the survival rate of patients. One of the early phenomena of lung cancer is the appearance of Pulmonary nodules.

Pulmonary nodules are an imaging feature of lung disease. For the diagnosis of Pulmonary nodules, it is recognized that the best detection method is Computer Tomography (CT). CT is the most important means for diagnosis, follow-up and efficacy evaluation of pulmonary nodules, but there is radiation during scanning, which has certain effects on human body. In general, low-dose CT can be used, at one-fifth or less of the conventional dose, with minimal impact on the human body. The normal population is recommended to have a yearly scan. Thin-section CT, enhanced CT, MAGNETIC resonance imaging, positron emission computed tomography (PET) and other imaging examinations can be used to analyze the properties of pulmonary nodules from the aspects of morphology and metabolic function, which is of great reference value for the diagnosis of pulmonary nodules.

The pulmonary nodules are approximately spherical in physiological structure and appear as circular shadows in CT images. In the medical field, pulmonary nodules are defined as focal, quasi-circular, dense or asfirm lung shadows with imaging findings of diameter 3 cm.

In the past, the identification of pulmonary nodules required doctors. Without professional training and long-term experience, it was difficult for doctors to visually distinguish pulmonary nodules from other lung tissues.

With the development of deep neural network technology and the increase of medical image data, computer aided diagnosis (CAD) has gradually emerged and achieved great development. CAD can diagnose quickly and accurately, which is of great significance to the development of medicine. But Traditional CAD systems detect candidates based on some simple assumptions (e.g. nodules look like a sphere) [**?**].

Our main contribution are as follows:

- We propose a novel CAD system based on 3D V-Net for detecting pulmonary nodule.
- We verified the CAD can accurately determine whether there are pulmonary nodules in the CT image through the LUNA16 dataset, improving the diagnostic accuracy rate reaches 65%.
- We propose a classifier based on ResVGGNet to classify the nodules, the overall accuracy rate of the classifier is 98.73%.

II. Literature Review

Recent advances in detecting and segmenting pulmonary nodules often depend on the progress of object recognition and classification for nodule screening, which was mainly driven by Deep Neural Networks (DNNs) [**?**] and the improvement of 3D classifiers for false positive reduction that aims at categorizing nodule candidates as "nodule" or "non-nodule" [**?**]. For the object detection networks, Region Proposal Networks (RPN) is a fully-convolutional network that simultaneously

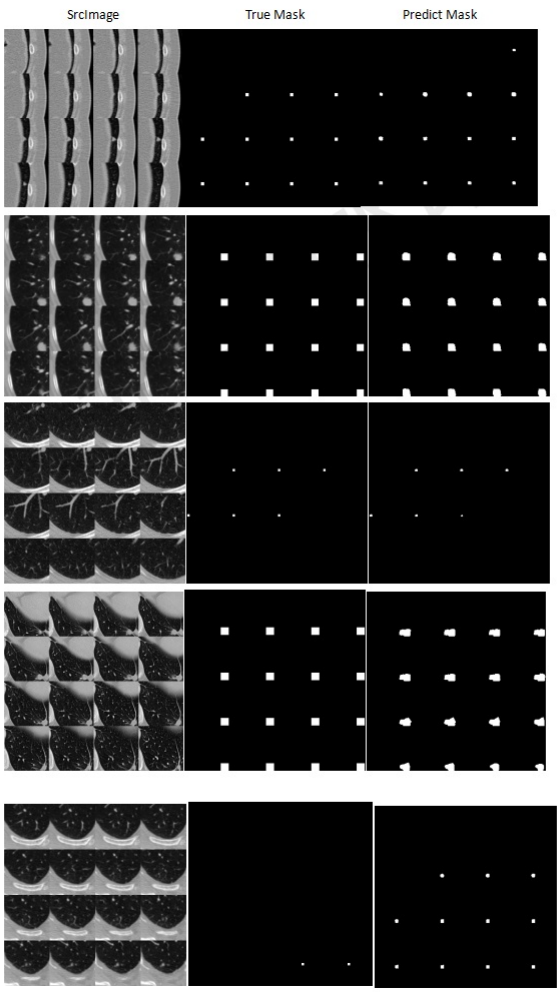

Fig. 1: Our result of detection on the test data. The original image is predicted and compared with the gold standard image. The original image is shown on the left, the gold standard image is shown in the middle, and the predicted image is shown on the right.

predicts object bounds and objectness scores at each position [**?**]. Although a lot of methods which can achieve the same funtion as RPN, RPN itself requires lower cost of GPU in the region proposal step by the virtue of convolutional features shared with down-stream dectetion network. Moreover, overall object detection accuracy can also be improved after RPN learns, these are parts of the reasons why RPN has a better performance in terms of algorithmic and is commonly applied in the different fields of image detection.

For instance, the model in [**?**] utilizes the 3D Faster Region-based Convolutional Neural Network (Faster R-CNN) with an acrchitecture resembling U-net, which achieves the performance of 83.4% in terms of Free Receiving Operating Curve (FROC); the nodule detection system in [**?**] also applies a typical 3D RPN that is also constructed in a U-net like structure. The predicted proposals of the model are used as detection results directly rather than adding an additional classifier because only two types (nodule or non-nodule) are used in the system; Besides the previous two models adopting U-net like structures, models in [**?**] and [**?**] also apply U-net architecture to their segmentation. However, V-net is also accepted in medical image segmentation. [**?**] shows us the sample to use V-net. The left and right parts of V-net represent a compression path and a decompression path respectively. During the progress of the model, the resolution of the data is gradually reduced at the end of each stage by using proper stride, and the crucial features are also extracted from the original data.

After nodule detection, it requires a 3D classifier to reduce false positive cases. Because time and cost are always limited, a two-dimentional DCNN detector of nodule candidates can be a good choice, However, more discriminative features are captured by 3D DCNNs rather than 2D DCNNs and consequently 3D DCNNs are more widely used in classifiers. For example, the model in [**?**] proposed a CAD system using multi-view convolutional networks, which also takes low computation time; Q Dou et al also proposed a method employing 3D CNN that can better tackle nodules of different sizes [**?**]. Besides these multi-scale networks, there is a single-scale nodule detection model [**?**], but the assumption of this model per se leads to some limitations in performance, which was not commonly adopted by the normal criteria. As for the initialization of parameters, the strategy used in the literature [**?**] can be considered in the networks.

Next, we will show two excellent models with high performance and try to find out the structures of these two models. More models can be seen in [**?**]. The first model from J Ding et al is mainly based on Deep Convolutional neural networks (DCNNs) of both 2D and 3D methods [**?**]. To be concrete, this model forgoes the method to deal with the 3D volume of oringinal CT scan directly due to a very high computation cost, but proposes to first uses a deconvolutional structure to Faster R-CNN for candidate detection on axis slices. After this step, two neighbouring slices of each axial slice will be concatenated in an axial direction, and thence be rescaled into $600\times600\times3$ pixels. In the second part, a three-dimensional DCNN is applied to the following process of false positive reduction. In this part the networks have a structure that can be divided into 4 kinds of layers, including a 2-way softmax activation layer put in the final, and more details of the networks can be found in [**?**]. Combining these two parts, this system achieved very high performance in the datasets given by Lung Nodule Analysis (LUNA16).

There are also CAD models specially designed for avoiding the friction between different tasks. For instance, the system invented by H Tang et al [**?**] managed to incorporate decoupled feature maps and a segmentation refinement subnet, which gives a full solution to the diagnosis of pulmonary nodules. Some parts of this system are also readily to be applied to other CAD models. This system can be divided into 3 parts: nodule candidate screening (NCS), decoupled false positive reduction (DFPR) and segmentation refinement (SR). In the first part, the system uses a $3\times3\times3$ 3D layer and two parallel $1\times1\times1$

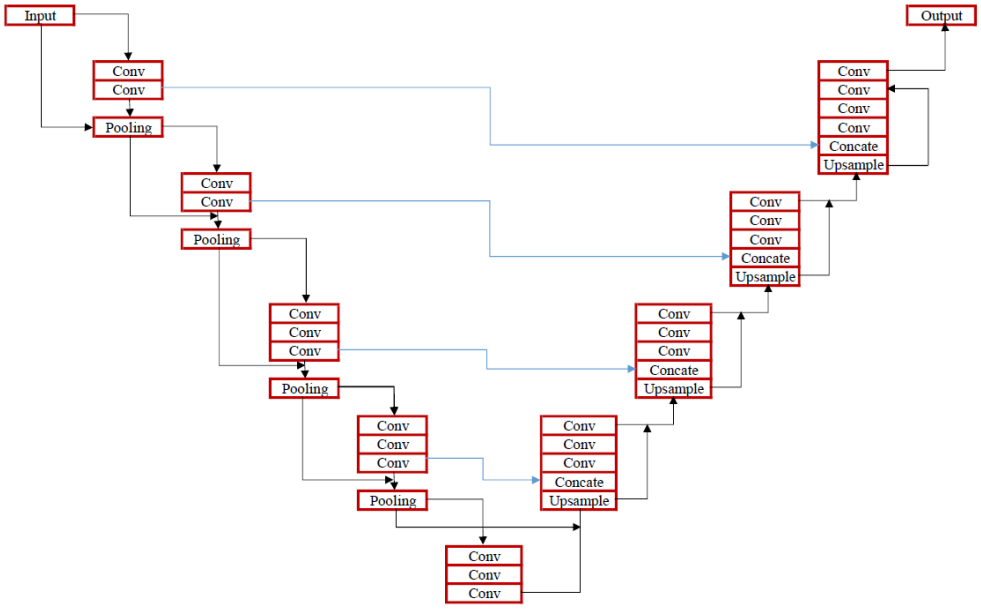

Fig. 2: Network structure of 3D V-Net

layers in order to generate nodule candidates, which reduce the same multi-task loss function to a large extent. In the second part, the system will apply a 3D pooling layer to get features not from the map which has the same features as RPN, but from the early feature map which has a miniature receptive field. This part has the same function as the NCS. In the third part, the system will gradually upsample the cropped map that is processed in the previous steps and connect the results with low-level semantically strongly features. The aim of this part is to minimize both the soft dice loss of the predicted mask sets and the soft dice loss of the ground truth mask sets.

After these examples, we are going to show the preprocess of our model on the data at first, then explain the methods we used in our nodule detection and finally illustrate how we classify the nodule candidates.

III. Model

To detect pulmonary nodules in the primitive CT images of the lungs, three steps are commonly adopted. First, an image segmentation algorithm is used to generate a mask image of the lung region, and then the lung region image was generated based on the mask image. Sencond, the image of lung region generated by lung segmentation and the mask image generated by the nodule labeling information were used to train the pulmonary nodule divider based on the convolutional neural network. Third, after the suspected pulmonary nodules are found, common image classification algorithms (such as CNN, etc.) can be used to classify the suspected pulmonary nodules, and the probability of whether the suspected pulmonary nodules are real pulmonary nodules can be obtained. Nevertheless, LUng Nodule Analysis 2016 (LUNA16) challenge [**?**] has helped us to directly locate the pulmonary nodules, so what we do in the aritcle is only the second and the third step.

A. Preprocess the Data

The file annotations.csv is indispensable to generate the mask of pulmonary nodules. It annotates the coordinates of X, Y, Z as well as the diameter of pulmonary nodules. Using these data, we could generate a cubic area which the center point is the coordinate and length is the diameter. The cubic area is the mask of pulmonary nodules that we need. Then we try denoising the image in order to find a reasonable gray value interval and restore the image. We set the window width and window level (-1000,600) to wipe off the image noise of background, such as highlight of bones, the metal lines of the CT bed etc.. It is imperative to make the input picture at the same size for fully connected layer. So we crop the picture in the same size. The CT image with a layer thickness greater than 1mm and the corresponding Mask images were sampled with interpolation (linear interpolation method was used for CT image and nearest-neighbour interpolation method for Mask image), and the layer thickness after interpolation sampling was 1mm. The size of the Patch area (96,96,16) in CT image and Mask image was taken according to a certain step length, and the valid Patch Mask image and corresponding Patch image were judged and retained. The final step is to prepare benign and malignant pulmonary nodules classification data. Coordinates are read from the candidates.csv file, and images of the size of (48,48,48) are taken as candidate pulmonary nodules images centring on this coordinate, and the images are

divided into two categories according to the label value (0 or 1). After doing these steps, we have data that we can train with. And we found that there were 1,351 pulmonary nodules and 549,714 non-pulmonary nodules. There's a huge difference between positive and negative samples. Therefore, we performed data enhancement processing on the data and randomly sampled 20% of the data from 549714 non-pulmonary nodules images, which were enlarged by 40 times (rotation, translation, inversion, etc.) for 1,351 pulmonary nodules images. A total of 601 of the 888 cases had pulmonary nodules in the CT data, and a total of 16,475 patches were taken out from the 601 cases. We chose 80 percent of the data for training and 20 percent for testing. Now we have prepared the data.

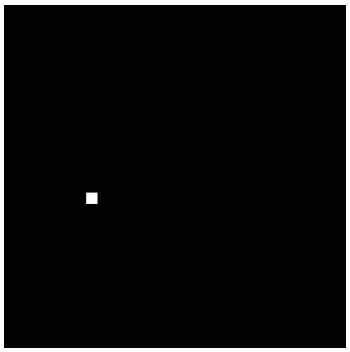

Fig. 3: Accorind to candidates.csv, we generate the mask of pulmonary nodule. This cubic area's center point is the coordinate and length is the diameter which candidates.csv annotates. It points out the position of the pulmonary nodule.

B. Nodule Detection

We use 3D V-Net network [**?**] model to achieve detection. Schematic model could be seen in 2.

The whole network is divided into compressed path and uncompressed path, just similar to [**?**], shrink and expand feature maps. In each stage it reduces the feature by half. And residual connection is used in each stage to accelerate the convergence. PRelu activation [**?**] is used as activation function in the whole net work to improve the peformance of fitting data. At the end of network, 1x1x1 kernel is used to process data of the same size as the input.

The following points should be noted:

- The convolutional layer connecting the output layer adopts the convolution kernel of 1x1x1 size, while the remaining Conv layer adopts the convolution kernel of 3x3x3 size.
- 3dMaxPooling layer is adopted in the pooling layer.
- The Upsample layer is implemented by the deconvolution layer.
- Concate layer splices the results of convolutional layer and Upsample layer in the decoding network.

It could be seen in 3 that the mask of pulmonary nodule only take a small place in the whole image. The imbalance between foreground and background often causes in the process of gradient descent of loss function, the network falls into the local optimal solution. And this local optimal solution is strongly affected by the background. In the area of object detection, focal loss is often adopted to alleviate the imbalance between positive and negative samples. Here we adpot Dice Loss as our loss function.

$$D = \frac{2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \tag{1}$$

In the equation, p_i means predicted foreground, g_i means groundtruth(real foreground) and N means the number of pixel.

It measures the overlap of the two samples and cope with the problem when the foreground only take a little propotion in the whole image.

C. Nodule Classify

[**?**] is a classical network in the image classification area. It use two kernels in small size in place of one big kernel. But it occurs gradient disapear as the layer of neural network. And [**?**] alleviate this problem through adding connection between layers not adjacent to each other. This structre, which could be seen in 4, benefits a lot when doing backwark propogation. The equation of backward propogation is as bellow:

$$\frac{\partial \text{loss}}{\partial x_l} = \frac{\partial \text{loss}}{\partial x_L} \cdot \frac{\partial x_L}{\partial x_l} = \frac{\partial \text{loss}}{\partial x_L} \cdot \left(1 + \frac{\partial}{\partial x_L} \sum_{i=l}^{L-1} F\left(x_i, W_i\right)\right) \tag{2}$$

$\frac{\partial \text{loss}}{\partial x_L}$ means when the loss function reaches L's gratitude, 1 in brackets indicates that the short-circuit mechanism can propagate the gradient nondestructively, and this alleviates the situation of gradient disappearance. We use an improved

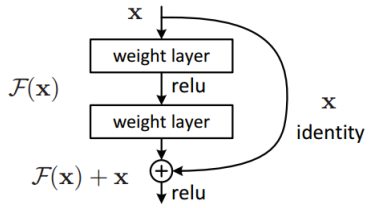

Fig. 4: skip-connection in the ResNet

version of the VGG network to achieve benign and malignant classification, we called it ResVGG. The structure could be seen in 5. The residual connection is used to prevent the gradient from disappearing during the training. And to the following points should be noted:

- All Conv layers adopt 3×3×3 convolution kernel, which in VGG is 3×3, and we expand the dimension for the 3D image.
- 3dMaxPooling layer is adopted in the pooling layer.
- The FC layer is the full connection layer. The first FC USES Relu activation function, while the second FC USES Softmax activation function.

We adopt cross-entropy function as our loss function.

$$L = -\left[y \log \hat{y} + (1-y) \log\left(1-\hat{y}\right)\right] \tag{3}$$

where $\hat{y}$ means the predicted value, and y means the true value.

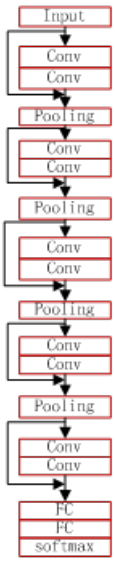

Fig. 5: The structre of ResVGGNet

TABLE I: The performance of the classifier under different indexes

	Precision	Recall	f1-score
0	0.99795	0.98282	0.99033
1	0.96736	0.99606	0.98150

IV. Results

In this work, we propose a novel CAD system based on 3D V-Net for detecting pulmonary nodule and a classifier based on ResVGGNet to classify the nodules. We peform experiments to checkout whether the structure of the model is effective. The result of detection could be seen in 1, I and II. And the result of classification could be seen in 6. We believe that our computer aided diagnosis system will make sense and assist docotors in the diagnosis of pulmonary nodules.

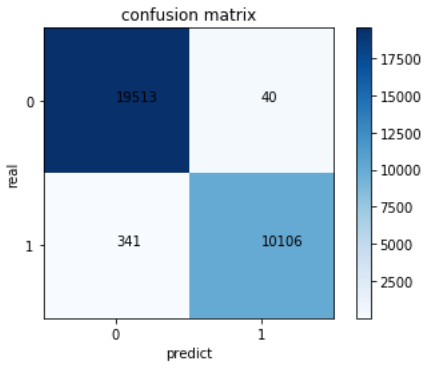

Fig. 6: By analyzing the confounding matrix, the overall accuracy rate of the classifier is 98.73% , the false positive rate is 1.718% and the omission rate is 0.394%.

TABLE II: Accuracy of different CAD models

System	accuracy
3D CNN	0.46
3D M-Net	0.60
3D U-Net	0.58
3D W-Net	0.63
Ours	**0.65**

V. Future Studies

Due to the time and device restrictions, we use few models to achieve the detection and classification of pulmonary nodules. With the research of convolutional networks, more models are put award. And in the future, we will use these models to improve the accuracy of classification and detection.

Acknowledgment

Thanks very much for my group members in this internship. I would also like to express my gratitude to Dr Teoh Teik Toe for his discussion of topic selection and guidance on some issues in the early stage of my project. I hope we can meet at Nanyang Technological University in the future.

References

[1] E. Lopez Torres, E. Fiorina, F. Pennazio, C. Peroni, M. Saletta, N. Camarlinghi, M. E. Fantacci, and P. Cerello, "Large scale validation of the m5l lung cad on heterogeneous ct datasets," *Medical Physics*, vol. 42, no. 4, pp. 1477–1489, 2015.

[2] C. Szegedy, A. Toshev, and D. Erhan, "Deep neural networks for object detection," in *Advances in Neural Information Processing Systems 26* (C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, eds.), pp. 2553–2561, Curran Associates, Inc., 2013.

[3] S. G. Armato III, M. L. Giger, and H. MacMahon, "Automated detection of lung nodules in ct scans: Preliminary results," *Medical Physics*, vol. 28, no. 8, pp. 1552–1561, 2001.

[4] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems 28* (C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, eds.), pp. 91–99, Curran Associates, Inc., 2015.

[5] N. Sun, D. Yang, S. Fang, and H. Xie, "Deep convolutional nets for pulmonary nodule detection and classification," in *Knowledge Science, Engineering and Management* (W. Liu, F. Giunchiglia, and B. Yang, eds.), (Cham), pp. 197–208, Springer International Publishing, 2018.

[6] 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*, vol. 30, no. 11, pp. 3484–3495, 2019.

[7] H. Tang, D. R. Kim, and X. Xie, "Automated pulmonary nodule detection using 3d deep convolutional neural networks," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 523–526, 2018.

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds.), (Cham), pp. 234–241, Springer International Publishing, 2015.

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

[10] A. A. A. Setio, F. Ciompi, G. Litjens, P. Gerke, C. Jacobs, S. J. van Riel, M. M. W. Wille, M. Naqibullah, C. I. Sánchez, and B. van Ginneken, "Pulmonary nodule detection in ct images: False positive reduction using multi-view convolutional networks," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1160–1169, 2016.

[11] Q. Dou, H. Chen, L. Yu, J. Qin, and P. Heng, "Multilevel contextual 3-d cnns for false positive reduction in pulmonary nodule detection," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1558–1567, 2017.

[12] N. Khosravan and U. Bagci, "S4nd: Single-shot single-scale lung nodule detection," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018* (A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger, eds.), (Cham), pp. 794–802, Springer International Publishing, 2018.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, December 2015.

[14] A. A. A. Setio, A. Traverso, T. de Bel, M. S. Berens, C. van den Bogaard, P. Cerello, H. Chen, Q. Dou, M. E. Fantacci, B. Geurts, R. van der Gugten, P. A. Heng, B. Jansen, M. M. de Kaste, V. Kotov, J. Y.-H. Lin, J. T. Manders, A. Sóñora-Mengana, J. C. García-Naranjo, E. Papavasileiou, M. Prokop, M. Saletta, C. M. Schaefer-Prokop, E. T. Scholten, L. Scholten, M. M. Snoeren, E. L. Torres, J. Vandemeulebroucke, N. Walasek, G. C. Zuidhof, B. van Ginneken, and C. Jacobs, "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The luna16 challenge," *Medical Image Analysis*, vol. 42, pp. 1 – 13, 2017.

[15] J. Ding, A. Li, Z. Hu, and L. Wang, "Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks," in *Medical Image Computing and Computer Assisted Intervention MICCAI 2017* (M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. L. Collins, and S. Duchesne, eds.), (Cham), pp. 559–567, Springer International Publishing, 2017.

[16] H. Tang, C. Zhang, and X. Xie, "Nodulenet: Decoupled false positive reduction for pulmonary nodule detection and segmentation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019* (D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, and A. Khan, eds.), (Cham), pp. 266–274, Springer International Publishing, 2019.

[17] A. A. A. Setio, A. Traverso, T. de Bel, M. S. N. Berens, C. van den Bogaard, P. Cerello, H. Chen, Q. Dou, M. E. Fantacci, B. Geurts, R. van der Gugten, P. A. Heng, B. Jansen, M. M. J. de Kaste, V. Kotov, J. Y.-H. Lin, J. T. M. C. Manders, A. Sónora-Mengana, J. C. García-Naranjo, E. Papavasileiou, M. Prokop, M. Saletta, C. M. Schaefer-Prokop, E. T. Scholten, L. Scholten, M. M. Snoeren, E. L. Torres, J. Vandemeulebroucke, N. Walasek, G. C. A. Zuidhof, B. van Ginneken, and C. Jacobs, "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge," 2016.

[18] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," 2016.

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," 2015.

[21] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015.