



Article A Computer-Aided Detection System for the Detection of Lung Nodules Based on 3D-ResNet

Jiaxu Ning ^{1,*}, Haitong Zhao ², Lei Lan ², Peng Sun ^{3,4} and Yunfei Feng ⁵

- ¹ School of Information Science and Engineering, Shenyang Ligong University, Shenyang 110159, China
- ² College of Computer Science & Engineering, Northeastern University, Shenyang 110819, China;
- zhaohaitong214@gmail.com (H.Z.); lanlei1021@gmail.com (L.L.)
- ³ Fidelity Investments Inc., Boston, MA 02210, USA; peng.sun@fmr.com
- ⁴ Department of Computer Science, IOWA State University, Ames, IA 50010, USA
- ⁵ Wal-mart Inc., Bentonville, AR 72712, USA; Yunfei.Feng@walmart.com
- * Correspondence: zhangchangsheng@mail.neu.edu.cn

Received: 12 November 2019; Accepted: 11 December 2019; Published: 16 December 2019



Abstract: In recent years, the research into automatic aided detection systems for pulmonary nodules has been extremely active. Most of the existing studies are based on 2D convolution neural networks, which cannot make full use of computed tomography's (CT) 3D spatial information. To address this problem, a computer-aided detection (CAD) system for lung nodules based on a 3D residual network (3D-ResNet) inspired by cognitive science is proposed in this paper. In this system, we feed the slice information extracted from three different axis planes into the U-NET network set, and make the joint decision to generate a candidate nodule set, which is the input of the proposed 3D residual network after extraction. We extracted 3D samples with 40, 44, 48, 52, and 56 mm sides from each candidate nodule in the candidate set and feed them into the trained residual network to get the probability of positive nodule after re-sampling the 3D sample to $48 \times 48 \times 48$ mm³. Finally, a joint judgment is made based on the probabilities of five 3D samples of different sizes to obtain the final result. Random rotation and translation and data amplification technology are used to prevent overfitting during network training. The detection intensity on the largest public data set (i.e., the Lung Image Database Consortium and Image Database Resource Initiative—LIDC-IDRI) reached 86.5% and 92.3% at 1 frame per second (FPs) and 4 FPs respectively using our algorithm, which is better than most CAD systems using 2D convolutional neural networks. In addition, a 3D residual network and a multi-section 2D convolution neural network were tested on the unrelated Tianchi dataset. The results indicate that 3D-ResNet has better feature extraction ability than multi-section 2D-ConvNet and is more suitable for reduction of false positive nodules.

Keywords: computer-aided detection; deep learning; lung nodule detection; residual network

1. Introduction

Lung cancer is the cancer with the highest mortality in the world. According to statistics, 222,500 new cases of lung cancer and 155,870 deaths are predicted in the United States in 2017 [1]. Lung cancer deaths account for 25% of all cancer deaths, ranking first among all cancers. The seminal National Lung Screening Trial [2] shows that screening for high-risk lung cancer with low-dose computed tomography (CT) is more effective than chest X-ray screening, which can reduce the 7-year mortality rate by 20%.

Since then, screening for lung cancer with low-dose CT has been widespread worldwide. However, a large number of CTs require more radiologists to analyze and diagnose, which greatly increases the burden and the needs of radiologists. To alleviate this, researchers have studied computer-aided

methods for the detection of lung nodules based on CT and have proposed multiple computer-aided detection (CAD) systems for this purpose. These CAD systems can be divided into non-convolution neural network CAD systems, such as the use of a support vector machine (SVM) classifier for lung nodule detection [3] or using image intensity, shape, and texture features for nodule detection [4–6] as proposed by Wiemker and Rafael; and convolution-neural-network-based CAD systems, such as lung nodule detection combined with multi-section based on a 2D convolution neural network.

A typical CAD system consists of two stages: 1) nodule candidate detection and 2) false positive reduction [7]. The aim of the nodule candidate detection is to detect nodule candidates at a very high sensitivity, which usually produces many false positives. In the false positive reduction stage, the system is expected to have a good classification performance to reduce false positive nodules in the candidates at the nodule discovery stage. Finally, the probability of each candidate nodule being a true positive is given. This stage has an especially important impact on the performance of the CAD system.

Due to the improvement of modern computer computing power and the development of deep learning [8], convolution neural networks are increasingly being applied in the image analysis field [9]. Because convolution neural networks have a strong ability to extract image features, CAD systems based on them have shown better performance than other methods and achieved a higher competition performance metric (CPM) score. The approach proposed by Nasrullah et al. uses two deep 3D faster R-CNN and U-Net encoder–decoders with MixNet to detect and learn the features of the lung nodules, respectively [10]. For the DeepLung model, a 3D faster region-based convolutional neural network (R-CNN) was designed for nodule detection with 3D dual path blocks and a U-NET-like encoder–decoder structure to effectively learn nodule features, and deep 3D dual path networks (DPNs) were designed for nodule detection and classification, respectively [11].

Although the existing CAD systems based on convolution neural networks have achieved better performance than CAD based on other approaches [12–14], there are still two problems in these studies: (1) Too many false positive nodules are introduced in the nodule discovery stage, and the convolution neural network is prone to false positive nodules in the subsequent false positive reduction stage. (2) These studies mostly use 2D convolution neural networks for false positive reduction. It is well-known that a CT image has three dimensions. A 2D convolution neural network cannot make full use of the 3D feature and spatial information of the image data. So, there is still much room for the precision of the network to improve.

Aiming at the above problems, we use the spatial information of data as much as possible in the phases of both candidate nodule detection and false positive nodule reduction. The main contributions are as follows: (1) A nodule discovery method based on the U-NET framework [15] is designed, which is based on a 2D convolution neural network and has good feature extraction ability. Through extracting information from three axial planes for detection, it makes good use of the spatial information of the data. This candidate nodule discovery method has high discovery intensity and greatly reduces the number of false positive nodules. (2) A 3D residual network structure is proposed to reduce false positives. For 3D images such as CT, the 3D convolution neural network has a stronger ability to extract features than a 2D network. In order to make the network easy to train, we add residual blocks to alleviate gradient dispersion.

2. Data

The dataset in this paper was from the largest available open dataset (i.e., the Lung Image Database Consortium (LIDC-IDRI)) [16] and the Tianchi racing data set [17]. The LIDC-IDRI database was initiated by the American National Cancer Institute, aiming to study early cancer detection in high-risk groups. The LIDC-IDRI database contains a total of 1018 CT images that come from seven institutions and were labeled by four experienced radiologists in two stages. In the first stage, each radiologist independently diagnosed and labeled the nodule position. The following are the three kinds of labeled objects: (1) R (diameter) \geq 3 mm nodules, (2) R < 3 mm nodules, and (3) R \geq 3 mm non-nodules. In the second stage, annotations from all four radiologists were reviewed in an unblended fashion and each

radiologist decided to either accept or reject each annotation. The database is very uneven, and the thickness of CT images varies from 0.6 to 5.0 mm. According to the recommendations of the American Radiology Association [18], we discarded the scans with a thickness greater than 3 mm, yielding 888 scans. According to the current screening rules for lung cancer, we treated only nodules that were classified as $R \ge 3$ mm as significant, while nodules that were $R \ge 3$ mm and non-junctions were classified as meaningless. Since the nodule labels were labeled by different radiologists, the nodule labels whose distances were closer than the sum of their radii were merged. The merger regulation is to average their diameter and coordinates separately. We selected the nodules labeled by 3 or 4 of the 4 radiologists as standard positive nodules. These damage tissues were the detection target of our algorithm. Other nodules labeled by less than 3 radiologists or nodules nor true positive nodules.

3. Method

This paper proposes a two-stage CAD system framework. The implementation process of the proposed CAD system is shown in Figure 1. The two stages are: 1) candidate detection; and 2) false positive reduction. The network structure of the first stage is described in Figure 1a. It detects and extracts the candidate nodules from the preprocessed CT images. For this task, we designed a modified U-NET [15]. As shown in the bottom-left of Figure 1, the CT sections of X–Y, X–Z, and Y–Z are fed into the modified U-NET to detect candidate nodules. Finally, the candidate nodules in different dimensions of the sections are merged into 3D candidate nodules, which is the input of the next stage. The task of the second stage is to reduce false positives from the candidates generated by the first stage. A ResNet structure is employed at this stage, as shown in Figure 1b. As shown in the picture at the bottom-middle of Figure 1, we extracted five kinds of 3D samples with different sizes $(40^3 \text{ means } 40 \times 40 \times 40 \text{ mm}^3, 44^3 \text{ means } 44 \times 44 \times 44 \text{ mm}^3, 48^3 \text{ means } 48 \times 48 \times 48 \text{ mm}^3, 52^3 \text{ means } 40 \times 40 \times 40 \text{ mm}^3, 50^3 \text{ means } 40 \times 40 \times 40^3 \text{ mm}^3, 50^3 \text{ mm}^3, 50^3 \text{ mm}^3, 50^3 \text{ mm}^3, 50^3 \text{$ $52 \times 52 \times 52$ mm³, and 56^3 means $56 \times 56 \times 56$ mm³) from each candidate and adjusted the size to $48 \times 48 \times 48$ mm³ uniformly. The setting of the sizes of 3D scans was determined by experimental trials and the experience of the researchers. Next, we fed them into a 34-level 3D residual network to reduce false positives and get the positive probability nodules. When training the models, we randomly divided the 888 scans into 5 parts: 3 for training, 1 for cross-validation, and 1 for testing. The details of the implementation of the two stages are introduced in the following subsections.

3.1. Candidate Detection

For a CAD system, the candidate detection plays a decisive role in the performance of the next phase. The candidate detection algorithm should improve the detection intensity as much as possible while reducing the number of false positives on the basis of guaranteeing a high detection intensity. To improve the detection intensity and reduce the number of false positives, we designed a convolution neural network based on U-NET to detect candidates. In contrast to Z-NET [19] which only used vertical sections for candidate detection, we cut three axis planes of each CT to detect candidate nodules through the same U-NET. We obtained the position information p(x, y, z) and diameter of each candidate as well as the probability of the given candidate. We only focused on candidates with $R \ge 3$ mm since nodules R < 3 mm and non-nodules were considered as meaningless. We merged candidate sets of three sections to improve the detection intensity. The merger rule was to merge the candidates whose distances of (x, y, z) coordinates did not exceed 5 mm and the new coordinates took the average.

The images were preprocessed before passing through U-NET. There were great differences between the scans because they came from different facilities and different devices, which would have a great impact on our CAD system. We re-sampled the segmented images so that each voxel represented 1 mm³. Then we adjusted the pixel intensity from (-1000, 400) to (0, 1), and the portion outside this range was discarded.



Figure 1. Computer-assisted detection (CAD) system overview.

The original U-NET is very prone to over-fitting because this model contains a large number of parameters [20]. To solve this problem, several adjustments were made to the structure of U-NET. (1) We adjusted the U-NET input size from 572×572 to 320×320 ; the input size of 320×320 can contain all the lung tissue areas of all existing image data. (2) The number of convolutional cores was appropriately reduced on the premise of securing the number of original network layers. The modified U-NET did not show a significant drop in performance and this avoids over-fitting [21]. The adjusted network structure is shown by Figure 1. It consists of down-sampling steps and up-sampling steps. Each down-sampling step consists of two consecutive convolutional layers and one pooling layer. The convolution kernel size of a convolution layer was 3×3 , and the initial number of convolution kernels was 32. We doubled the number of convolution cores in every down-sampling step. The rectified linear unit (ReLU) [22] was applied to every convolution layer. We used the filling mode of the same padding to keep the size. The pooling layer used a 2×2 maximum pool in steps of 2 for down-sampling. Each up-sampling step consisted of two consecutive convolution layers—a connection layer and an up-sampling layer. The convolution kernel size of the convolution layer was also 3×3 , and the number of convolution kernels was halved in each stage. ReLU and the same padding were applied to each convolutional layer. The connection layer duplicated the corresponding same-size features in the down-sampling step and linked them with the corresponding features in the upper sampling line. The up-sampling layer used 2×2 steps for up-sampling to achieve size amplification. The last layer of the network adopted a 1×1 convolution operation. The sigmoid activation function was used in this layer. A probability map with the size of $320 \times 320 \times 1$ —the same as the input size—was obtained, and the candidate nodules were extracted based on the U-NET output probability mapping.

In order to prevent the network from over-fitting the false positive reduction caused by the imbalance of the positive nodules and the false positive nodules in the candidate set, we discarded some low-probability candidates and only focused on the higher-probability candidates. We selected a probability threshold based on experience and the performance on the cross-validation set. We reduced the probability threshold from 1.0, and the amplitude of each reduction was 0.1 until the detection intensity on the cross-validation set was constant after three consecutive reductions. This value was chosen as the threshold. This allowed us to ensure high detection intensity while minimizing the number of false positive nodules.

3.2. 3D Sample Extraction

Since the nodule's R (diameter) \geq 30 mm, in order to include all the nodule information and enough background information in the dicing blocks, we extracted $48 \times 48 \times 48 \text{ mm}^3$ 3D samples centered on the candidate p(x, y, z). Because the size distribution of the nodules is very uneven, in order to improve the performance of the CAD system, we also extracted dicing blocks of $40 \times 40 \times 40 \text{ mm}^3$, $44 \times 44 \times 44 \text{ mm}^3$, $52 \times 52 \times 52 \text{ mm}^3$, and $56 \times 56 \times 56 \text{ mm}^3$ and resized them to $48 \times 48 \times 48 \text{ mm}^3$. These 3D samples got their respective nodule probabilities via the same false positive reduction network and the average was used as the final probability.

3.3. False-Positive Reduction: 3D Depth Residual Network

In existing CAD systems [12], the 2D convolutional neural network is mostly adopted in the false positive reduction stage. However, 2D convolutional neural networks cannot make full use of 3D image information. Studies have shown that the convolutional neural network depth is an extremely important factor in determining its performance [23]. The performance of a convolutional neural network will increase as its depth increases. At the same time, deeper networks become more difficult to train and degenerate. In order to solve these problems, a deep residual network [24] was proposed by He et al. Based on the above discussion, we designed a 3D deep residual network in order to improve the performance of reducing false positives.

The structure of the 3D deep residuals network is shown in Figure 1b. The main network structure refers to the 34-layer residual network in [24]. The original residual network was a 2D residual network and the CT involved in this paper provides 3D image information. The original 2D network cannot make good use of spatial information, so we improved the original 2D residual network and propose a 3D residual network. At the same time, the number of original network convolution kernels is large. We properly reduced the number of convolution kernels under the existing data set to prevent over-fitting and to prevent difficulty in training. The improved 3D residual network was composed of 33 convolution layers and a fully-connected layer. The convolution layer used $3 \times 3 \times 3$ convolution kernels and the initial convolution kernel was 16. For the same size output, the number of convolution kernels does not change. We doubled the number of convolution kernels when the size of feature map was reduced by half. We achieved down-sampling by setting the step size to 2. We implemented residual operations by inserting shortcut connections between convolutional layers. While the input and output of the residual block were the same, we added the input characteristics and output characteristics of the residual block directly before passing through the activation function. When the output size of the residual block increased, we let the input feature map pass through a $1 \times 1 \times 1$ convolutional layer of step size 2 to uniform size and then added it to the output. At the end of the network was a global average pooling and a fully-connected layer using the sigmoid's activation function.

In the training process, we used an efficient first-order stochastic gradient optimization algorithm (Adam [25]) to optimize the model. The cross-entropy error was used as a measure of the loss. The number of samples per batch (mini-batches) was 40, and batch normalization [26] was used to accelerate training before each convolutional layer's activation function. The weight initialization strategy is the standard initialization proposed by Glorot et al. [27]. Training was stopped when the accuracy in the crossing validation dataset did not improve after three epochs.

3.4. Data Augmentation

Data augmentation is an important way to improve the performance of convolutional neural networks because it prevents over-fitting. We adopted a series of data augmentation methods during the candidate detection stage and false positive reduction stage.

(1) Candidate detection phase: In the candidate detection phase, we adopted the following data augmentation methods: (-20, 20) degree random rotation, image mirror, $R \ge 30$ mm random translation. Through data amplification, the detection intensity increased slightly.

(2) False-positive-reduction stage: Because the candidate set is unbalanced, the false positives are far greater in number than the true positives. Deep residual networks trained with unbalanced data can result in predictions trending toward false positives, degrading the performance of the network. Therefore, the amplification process at this stage was only for the positive nodules of the training set and the cross-validation set. For each positive nodule, we took the following amplification methods: (-10, 10) degrees random rotation, 1 mm translation in each axis, and extraction of 3D samples with side lengths of 40, 44, 48, 52, and 56 mm and adjusting them to $48 \times 48 \times 48$ mm³ uniformly.

3.5. Evaluation

The output of the CAD system is a CSV file containing all the label information of candidates. Each label contains (x, y, z) coordinate information and a score. The score range is (0, 1). The higher the score, the higher the probability that the corresponding nodule is a true positive. When the label coordinates were located at a circular space centered on the standard reference coordinate with the radius of the nodule as radius, we identified it as a correctly detected positive nodule. Otherwise it was treated as a false positive nodule. The free response operating characteristic (FROC) curve of the result set and the CPM [28] were used as evaluation criteria to measure the CAD system's performance. The CPM calculated the average intensity of the FROC curves at seven points (1/8, 1/4, 1/2, 1, 2, 4, and 8 frames per second per scan (FPs/scan)). It shows the performance of CAD systems at important points and can be a good measure of the performance of a CAD system.

4. Experimental Results

4.1. Candidates Detection

We extracted the candidate set using three sections via U-NET and then merged the three candidates. The detection intensity of the three sections merged into the final detection intensity. In order to compare the detection performance of candidate nodules with other CAD systems, we compared the solid [29], sub-solid [5], and large solid [30] algorithms in the proposed system [12] with their merged detection sensitivity. As shown in Table 1, the detection sensitivity of the three axial sections, Z-order, Y-order, and X-order, were 91.5%, 80.3%, and 84.4%, respectively. The detection sensitivity after merging the three candidate sets and removing the candidate nodules with lower probability was 94.2%. Compared to the solid, sub-solid, and large solid algorithms, the false positives produced by our algorithm for nodule discovery based on three-section U-NET guaranteeing the discovery intensity of similar nodules were only 30% (81.6/269.2) of those obtained by the original algorithm.

Table 1. Detection sensitivity of candidate detection algorithms.

Candidate Dection	Sensitivity (%)	FPs per Scan		
Z-order	91.5	46.9		
Y-order	80.3	21.9		
X-order	84.4	46.4		
Combined set	95.2	81.6		
Solid	85.7	329.3		
Sub-solid	36.1	287.2		
Large solid	31.8	47.1		
Reduced set	93.3	269.2		

For a given candidate nodule set, we classified positive and false positive nodules by using 3D-ResNet. The results of false positive reduction are shown in Figure 2. With the help of the test set amplification, we obtained an average detection sensitivity of 78% at seven operating points. We found that 3D-ResNet test data augmentation (TDA) significantly improved the network performance. The average detection intensity increased from 72.1% to 78% through the test set amplification. The detection sensitivity of the nodule detection phase was 0.952, and our CAD system achieved 86.5% and 93.3% detection sensitivity at 1 FPs/scan and 4 FPs/scan, respectively.



Figure 2. (a) The line chart of 3D-ResNet test data augmentation (TDA) and free-response operating characteristic (FROC) of noTDA in the false positive reduction stage; (b) 3D-ResNet and late-fusion 9 views. The space between the dotted red/blue lines is the 0.95 confidence interval of the data of each method, respectively.

As a contrast, we implemented the multifaceted 2D ConvNets of DIAG CONVNET [12] in the false positives detection stage. For each candidate, nine $50 \times 50 \times 50$ mm³ sections were extracted from $64 \times 64 \times 64$ mm³ 3D samples. Each section corresponded to different planes of 3D samples and via 2D ConvNets. Finally, we used a late-fusion method to merge them. The FROC curves of our 3D-ResNet and the late-fusion 9 views network are shown in Figure 2b. 3D-ResNet's CPM = 0.78 was higher than late-fusion 9 views' (CPM = 0.729). At the same time, the detection intensity of 3D-ResNet was also higher than late-fusion 9 views at 1 FPs/scan and 4 FPs/scan.

Figure 3 gives the candidate 3D samples with positive probability from 0.1 to 0.9 after removal of the false positive nodes. It is obvious from the figure that the more pulmonary nodules the candidate node had, the greater the positive probability.



Figure 3. Candidate 3D samples of different probability.

4.3. Comparison with Existing CAD Systems

4.3.1. Comparison with 2DConvNet-CAD in CUMedVis

In order to better evaluate the performance of 3D-ResNet on different data sets, we compared the CAD system in this paper with Conv2D in DIAG CONVNET. The test data used for comparison was the totally unrelated Tianchi Race [17] data, and the corresponding candidate set of the data was generated by the U-NET of the candidate nodule generation phase above. In order to compare the two networks to determine which had better feature extraction and false positive reduction capability, we performed test set augmentation on both networks based on the same candidate nodule set. The performance of 2DConvNet-CAD in CUMedVis [31] and 3D-ResNet CAD on the unrelated Tianchi dataset is shown in Figure 4. The CPM of 3D-ResNet CAD was 0.503, which is better than the 2DConvNet-CAD CPM score (0.429). It is obvious that the 3D convolution neural network had a better spatial feature extraction capability for 3D image data with spatial features such as CT and had a better performance than the 2D convolution neural network. At the same time, when the network structure is complex and the network is deep, the residual network can alleviate the phenomenon of gradient dispersion greatly and accelerate training.



Figure 4. Comparison with 2DConvNet-CAD on the Tianchi dataset.

4.3.2. Comparison with Ohter CAD Systems that Use LIDC-IDRI Data Sets

To compare with other existing CAD systems more widely, we list the performance of other CAD systems that use the LIDC-IDRI data set in Table 2. In contrast to the proposed two-stage nodule detection approach, the compared methods use one step to process the input CT images, regardless

of what kind of model is employed. The feature merging method is employed in the proposed CAD system to merge nodules from different dimensions of sections. As shown in Table 2, our 3D-ResNet CAD had better performance than other CAD systems based on selecting the same layer thickness and the same size data. Especially at 4 FPs/scan, the detection intensity of 3D-ResNet CAD reached 92.3%. Although the CAD system proposed by Cascio et al. obtained a sensitivity of 97.0%, it found 111 nodules among the same samples. In contrast to this, the proposed CAD system found more true positive nodules. Thus, the proposed CAD system had a stronger ability to detect lung tumor nodules, which is essential for diagnosis in the medical field.

Table 2. CAD system using the Lung Image Database Consortium (LIDC-IDRI) dataset published in recent years.

CAD Systems	Year	#Scans	Slice Thickness	Nodule Size (mm)	Agreement Levels	Nodules	Sensitivity (%) /FPs/Scan
3D-ResNet	-	888	≥2.5	≥3	at least 3	1186	92.3/4.0 86.5/1.0
CUMedVis [31]	2016	888	≥ 2.5	≥ 3	at least 3	1186	90.7/4.0 84.8/1.0
DIAG_CONVNET [12]	2016	888	≥ 2.5	≥ 3	at least 3	1186	90.1/4.0 85.4/1.0
LungCAM [32]	2015	949	-	≥ 3	at least 3	1749	80.8/8.0
Van Ginneken et al. [13]	2015	865	≥ 2.5	≥ 3	at least 3	1147	76.0/4.0 73.0/1.0
Brown et al. [4]	2014	108	0.5–3	≥ 4	at least 3	68	75.0/2.0
Choi and Choi [6]	2013	58	0.5–3	3–30	at least 1	-	95.3/2.3
Tan et al. [33]	2013	360	-	≥ 3	at least 4	-	83.0/4.0
Teramoto and Fujita [34]	2013	84	0.5–3	5-20	at least 1	103	80.0/4.2
Cascio et al. [35]	2012	84	1.25-3	≥ 3	at least 1	148	97.0/6.1 88.0/2.5
Guo and Li [36]	2012	85	1.25-3	≥ 3	at least 3	111	80.0/7.4 75.0/2.8
3D MixNet [10]	2019	89	1.25-3	≥ 3	at least 3	127	88.83/3.5
Li et al. [37]	2018	85	1.25–3	≥ 3	at least 3	119	86.0/2.4

5. Conclusions

In this paper we proposed a nodule discovery algorithm based on a three-section U-NET and a CAD system false positive reduction algorithm based on a 3D residual network. The system presented two successes. First, in terms of candidate detection, compared to the single-section U-NET used by Z-NET, we used the optimized U-NET for detection in three sections. Then, we used a high threshold to extract the candidate nodule set and finally merged the three candidate sets. In this way, the spatial information of the data can be better used to improve the detection intensity at the nodule detection stage, while the high threshold reduces the number of false positive nodules to a large extent. As shown by Table 1, at a similar detection intensity of 95.2%, the number of false positive nodules generated by this method was only 30% that of the other detection methods. Secondly, in the false positives reduction stage, to make better use of the data's spatial information, we abandoned the traditional 2D convolution neural network and used a 3D convolution neural network. Because spatial information is essential for spatial data like CT scans, using spatial information effectively can greatly improve the ability of the network to identify nodules. At the same time, we increased the depth of the network as much as possible to achieve better performance. We adopted a 3D residual network to make the deep 3D network easy to train. For each nodule, we extracted five 3D samples of different sizes into a joint judge to eliminate the effect of uneven nodule size distribution. The results show that this improved performance to some extent. Finally, we implemented a multi-section 2D convolution neural network (CUMedVis) [31] and compared it with our algorithm on the same data. The results show that our 3D residual network had better performance and a higher CPM score (0.780). Based on the experimental results, we conclude that 3D convolution neural networks have a stronger spatial feature extraction capability for 3D data with spatial information (e.g., CT scans) than 2D networks, and therefore can obtain better performance and CPM scores. In this paper, we focused on the optimization of both nodule candidate detection and false positives reduction, and achieved good results.

Author Contributions: Conceptualization, J.N.; methodology, J.N. and H.Z.; software, L.L.; validation, J.N., H.Z., L.L., and P.S.; formal analysis, J.N. and Y.F.; investigation, J.N.; resources, H.Z.; data curation, L.L., P.S., and Y.F.; writing—original draft preparation, J.N.; writing—review and editing, J.N. and H.Z.; visualization, J.N., P.S., and Y.F.; supervision, J.N.; project administration, H.Z.

Funding: This work was funded by the National Natural Science Foundation Program of China (61572116 and 61572117).

Acknowledgments: Thanks to the China Scholarship Council and reviewers.

Conflicts of Interest: We declare that we have no financial or personal relationships with other people or organizations that could inappropriately influence our work. There is no professional or other personal interest of any nature or kind in any product, service, and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled "An Aided Detection System for Lung Tumor Node Based on 3D Residual Network". This article does not contain any studies with human participants performed by any of the authors. Informed consent was obtained from all individual participants included in the study. All authors declare that they have no conflicts of interest.

References

- Siegel, R.L.; Miller, K.D.; Fedewa, S.A.; Ahnen, D.J.; Meester, R.G.; Barzi, A.; Jemal, A. Colorectal cancer statistics, 2017. *CA Cancer J. Clin.* 2017, 67, 177–193. [CrossRef] [PubMed]
- National Lung Screening Trial Research Team; Aberle, D.R.; Adams, A.M.; Berg, C.D.; Black, W.C.; Clapp, J.D.; Fagerstrom, R.M.; Gareen, I.F.; Gatsonis, C.; Marcus, P.M.; et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N. Engl. J. Med.* 2011, 365, 395–409. [CrossRef] [PubMed]
- 3. Bergtholdt, M.; Wiemker, R.; Klinder, T. Pulmonary nodule detection using a cascaded SVM classifier. In Proceedings of the Medical Imaging 2016: Computer-Aided Diagnosis: International Society for Optics and Photonics, San Diego, CA, USA, 28 February–2 March 2016.
- Brown, M.S.; Lo, P.; Goldin, J.G.; Barnoy, E.; Kim, G.H.J.; McNitt-Gray, M.F.; Aberle, D.R. Toward clinically usable CAD for lung cancer screening with computed tomography. *Eur. Radiol.* 2014, 24, 2719–2728. [CrossRef] [PubMed]
- Jacobs, C.; van Rikxoort, E.M.; Twellmann, T.; Scholten, E.T.; de Jong, P.A.; Kuhnigk, J.M.; Oudkerk, M.; de Koning, H.J.; Prokop, M.; Schaefer-Prokop, C.; et al. Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images. *Med. Image Anal.* 2014, *18*, 374–384. [CrossRef]
- 6. Choi, W.J.; Choi, T.S. Automated pulmonary nodule detection system in computed tomography images: A hierarchical block classification approach. *Entropy* **2013**, *15*, 507–523. [CrossRef]
- Pehrson, L.M.; Nielsen, M.B.; Ammitzbøl Lauridsen, C. Automatic pulmonary nodule detection applying deep learning or machine learning algorithms to the LIDC-IDRI database: A systematic review. *Diagnostics* 2019, 9, 29. [CrossRef]
- 8. Li, R.; Wang, S.; Gu, D. Ongoing Evolution of Visual SLAM from Geometry to Deep Learning: Challenges and Opportunities. *Cogn. Comput.* **2018**, *10*, 875–889. [CrossRef]
- 9. Liu, H.; Fang, J.; Xu, X.; Sun, F. Surface material recognition using active multi-modal extreme learning machine. *Cogn. Comput.* **2018**, *10*, 937–950. [CrossRef]
- 10. Nasrullah, N.; Sang, J.; Alam, M.S.; Xiang, H. Automated detection and classification for early stage lung cancer on CT images using deep learning. *Proc SPIE* **2019**, *10995*, 109950S.
- 11. Zhu, W.; Liu, C.; Fan, W.; Xie, X. Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification. In Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018; pp. 673–681.
- 12. Setio, A.A.A.; Ciompi, F.; Litjens, G.; Gerke, P.; Jacobs, C.; Van Riel, S.J.; Wille, M.M.W.; Naqibullah, M.; Sánchez, C.I.; van Ginneken, B. Pulmonary nodule detection in CT images: False positive reduction using multi-view convolutional networks. *IEEE Trans. Med. Imaging* **2016**, *35*, 1160–1169. [CrossRef]
- Van Ginneken, B.; Setio, A.A.; Jacobs, C.; Ciompi, F. Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans. In Proceedings of the 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), Brooklyn, NY, USA, 16–19 April 2015; pp. 286–289.

- Roth, H.R.; Lu, L.; Seff, A.; Cherry, K.M.; Hoffman, J.; Wang, S.; Liu, J.; Turkbey, E.; Summers, R.M. A new 2.5 D representation for lymph node detection using random sets of deep convolutional neural network observations. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Boston, MA, USA, 14–18 September 2014; pp. 520–527.
- 15. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical image computing and computer-assisted intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
- Armato, S.G., III; McLennan, G.; Bidaut, L.; McNitt-Gray, M.F.; Meyer, C.R.; Reeves, A.P.; Zhao, B.; Aberle, D.R.; Henschke, C.I.; Hoffman, E.A.; et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Med. Phys.* 2011, *38*, 915–931. [PubMed]
- 17. Tianchi Medical AI Competition. Available online: https://tianchi.aliyun.com/competition/entrance/231601/information (accessed on 1 February 2019).
- Kazerooni, E.A.; Austin, J.H.; Black, W.C.; Dyer, D.S.; Hazelton, T.R.; Leung, A.N.; McNitt-Gray, M.F.; Munden, R.F.; Pipavath, S. ACR–STR practice parameter for the performance and reporting of lung cancer screening thoracic computed tomography (CT): 2014 (Resolution 4). *J. Thorac. Imaging* 2014, 29, 310–316. [CrossRef] [PubMed]
- 19. Moira, B.; Robbert van der, G.; Michael de, K.; Jeroen, M.; Guido, Z. ZNET-Lung Nodule Detection. Available online: https://www.github.com/gzuidhof/LUNA16/ (accessed on 1 November 2018).
- 20. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
- 21. Cogswell, M.; Ahmed, F.; Girshick, R.; Zitnick, L.; Batra, D. Reducing overfitting in deep networks by decorrelating representations. *arXiv* **2015**, arXiv:1511.06068.
- 22. Glorot, X.; Bordes, A.; Bengio, Y. Deep sparse rectifier neural networks. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, Ft. Lauderdale, FL, USA, 11–13 April 2011; pp. 315–323.
- 23. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
- 24. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
- 25. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* 2014, arXiv:1412.6980.
- 26. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv* **2015**, arXiv:1502.03167.
- Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010; pp. 249–256.
- 28. Niemeijer, M.; Loog, M.; Abramoff, M.D.; Viergever, M.A.; Prokop, M.; van Ginneken, B. On combining computer-aided detection systems. *IEEE Trans. Med. Imaging* **2010**, *30*, 215–223. [CrossRef]
- 29. Murphy, K.; van Ginneken, B.; Schilham, A.M.; De Hoop, B.; Gietema, H.; Prokop, M. A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification. *Med. Image Anal.* **2009**, *13*, 757–770. [CrossRef]
- 30. Setio, A.A.; Jacobs, C.; Gelderblom, J.; van Ginneken, B. Automatic detection of large pulmonary solid nodules in thoracic CT images. *Med. Phys.* **2015**, *42*, 5642–5653. [CrossRef]
- 31. Dou, Q.; Chen, H.; Yu, L.; Qin, J.; Heng, P.A. Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection. *IEEE Trans. Biomed. Eng.* **2016**, *64*, 1558–1567. [CrossRef] [PubMed]
- Lopez Torres, E.; Fiorina, E.; Pennazio, F.; Peroni, C.; Saletta, M.; Camarlinghi, N.; Fantacci, M.E.; Cerello, P. Large scale validation of the M5L lung CAD on heterogeneous CT datasets. *Med. Phys.* 2015, 42, 1477–1489. [CrossRef] [PubMed]
- Tan, M.; Deklerck, R.; Cornelis, J.; Jansen, B. Phased searching with NEAT in a time-scaled framework: Experiments on a computer-aided detection system for lung nodules. *Artif. Intell. Med.* 2013, 59, 157–167. [CrossRef] [PubMed]

- 34. Teramoto, A.; Fujita, H. Fast lung nodule detection in chest CT images using cylindrical nodule-enhancement filter. *Int. J. Comput. Assist. Radiol. Surg.* **2013**, *8*, 193–205. [CrossRef] [PubMed]
- 35. Cascio, D.; Magro, R.; Fauci, F.; Iacomi, M.; Raso, G. Automatic detection of lung nodules in CT datasets based on stable 3D mass–spring models. *Comput. Biol. Med.* **2012**, *42*, 1098–1109. [CrossRef] [PubMed]
- 36. Guo, W.; Li, Q. High performance lung nodule detection schemes in CT using local and global information. *Med. Phys.* **2012**, *39*, 5157–5168. [CrossRef]
- 37. Li, C.; Zhu, G.; Wu, X.; Wang, Y. False-positive reduction on lung nodules detection in chest radiographs by ensemble of convolutional neural networks. *IEEE Access* **2018**, *6*, 16060–16067. [CrossRef]



 \odot 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).