Lung Nodule Segmentation and Recognition Algorithm Based on Multiposition U-Net

Na Zhang, 1 Jianping Lin, 2 Bengang Hui, 1 Bowei Qiao, 1 Weibo Yang, 1 Rongxin Shang, 1 Xiaoping Wang, 1 and Jie Lei 1

1Department of Thoracic Surgery, Tangdu Hospital, Air Force Military Medical University, Xi’an, China
2Department of Dermatology, Tangdu Hospital, Air Force Military Medical University, Xi’an, China

Correspondence should be addressed to Xiaoping Wang; wxpchest@fmmu.edu.cn and Jie Lei; leijiemd@163.com

Received 28 January 2022; Revised 21 February 2022; Accepted 9 March 2022; Published 23 March 2022

1. Introduction

With the global spread of new crown disease, more and more scholars focus on lung disease [1]. Among them, lung nodules are the main lesions of the lung. Malignant lung nodules can develop into lung cancer, which seriously threatens human health. Due to the small proportion of lung nodules in the lung, doctors might inevitably produce missed detection and false detection in the detection process. Therefore, through the research of artificial intelligence, scholars use intelligent algorithms to assist doctors in making accurate diagnosis. Specifically, it includes the following three aspects:


Methods of computer-aided diagnosis have been proved that they can effectively assist doctors to make accurate diagnosis. However, it is still difficult to study weak lung nodules and complex and diverse signs of lung nodules. Specifically, it includes the following: (1) the model established by the computer does not conform to the doctors’ diagnosis process. (2) In the face of lung nodules with lung wall adhesion, incomplete extraction will occur and result in missed detection. (3) In the face of complex and diverse signs of lung nodules, it is impossible to accurately distinguish from a single body position.

Therefore, a new lung nodule segmentation and recognition algorithm is proposed. (1) According to the doctor’s diagnosis process, a lung nodule segmentation and recognition algorithm process is constructed in line with the doctor’s diagnosis process. (2) The attention mechanism is constructed to focus on the area where the lung parenchyma is located, accurately extract the lung parenchyma, and reduce the missed detection rate of lung nodules. (3) The multiposition feature enhancement network is constructed to recognize the signs of lung nodules.

2. Details of Algorithm

According to the diagnosis process of lung nodules by doctors, the algorithm flow is proposed, as shown in Figure 1. Firstly, the lung parenchyma extraction model of attention mechanism is established, and then, the multiposition bidirectional enhancement feature pyramid network is established to extract lung nodules. Finally, the signs of lung nodules are identified by radiomics and morphological features.

2.1. U-Net Algorithm. U-Net is proposed based on a fully convolutional neural network and has achieved certain results in image segmentation. U-Net consists of an encoder part and a decoder part, as shown in Figure 2. The encoder has four submodules, which are composed of a convolution layer and a pooling layer, so that the image features are gradually reduced and abstracted; the decoder corresponds to it layer by layer, and the function of the deconvolution layer in the decoder is to increase the feature size in turn and use skip connections to connect and merge the deconvolution result of the decoding part and the output of the encoding part correspondingly. Finally, the probability map is output through convolution.

In the lung parenchyma segmentation algorithm, we introduce an attention mechanism to make U-Net focus on the region of interest. And we introduced dense atrous convolution to adjust the receptive field. In the lung nodule extraction algorithm, we introduce a multiangle model into a unified U-Net to achieve multangle detection.

2.2. Lung Parenchyma Segmentation Algorithm. Lung nodules only exist in the lung parenchyma. In order to simulate the diagnosis process of doctors, it is necessary to focus on
the area where the lung nodules are located to realize the segmentation of lung parenchyma. It consists of encoding path and decoding path.

The traditional encoding and decoding method is U-Net, which can obtain spatial information in shallow network, but the learning ability of depth feature is not good. The cascade method leads to high redundancy in the utilization of shallow features in feature fusion, which makes the network huge. Therefore, we introduce the attention mechanism into the network to obtain higher-level features and increase the weight of the target region in order to avoid the interference of background pixels and improve the learning ability of the model.

The corresponding attention model is shown in Figure 3. The module contains two inputs, upsampling feature $g$ and coding feature $x_i$. Through the convolution operation of $(1,1,1)$, $W_g^T g_i$ and $W_x^T x_i^l$ are obtained. On this basis, an attention model is built as follows:

$$J = \sigma_2 \left( \sigma_1 \left( W_x^T x_i^l + W_g^T g_i + b_g \right) \right) + \frac{1}{1 + \exp(-x_i)},$$  

where $\sigma$ represents the convolution kernel.

In view of the limitations of U-Net network, the continuous pooling kernel convolution step operation reduces the resolution, which results in the loss of detail information, and increases the receptive field by expanding the convolution kernel, which results in the increase of parameters and the difficulty of training. Thus, we add a dense connected block (DAC) to improve the network and integrate inception, residual network, and hole convolution. Its structure is shown in Figure 4. Four
cascade branches are used, and the corresponding receptive fields are 3, 7, 9, and 19. It is activated by ReLu. Based on the idea of residual network, the original features and other features are fused to enhance feature mining. Finally, combined with the cavity convolution of different scale expansion rates, the multi-scale feature extraction is realized.

In the standard U-Net framework, the sampling function of the maximum pooling kernel is used to reduce and...
mediate the resolution of the feature image, respectively, but
the training process will lead to feature loss and accuracy
reduction. Therefore, we construct the following structure,
as shown in Figure 5. The transposed convolution is to
transpose the convolution kernel in the ordinary convolu-
tion operation which we usually use and then take the out-
put of the ordinary convolution as the input of the transposed
convolution, and the output of the transposed

![Network Structure](image)

**Figure 6: Network structure.**

![Database Display](image)

(a) Lung nodule  
(b) Lung cancer  
(c) Grid-like increased density

**Figure 7: Database display.**
convolution is the ordinary convolution input. This structure is composed of multiscale convolution cores in parallel, which can catch up with the characteristics of different dimensions and provide the learning ability of the network.

Since lung parenchyma segmentation belongs to pixel segmentation and can be regarded as a binary classification problem, we introduce the Dice loss function to define the loss function:

\[
\text{Loss} = 1 - \frac{2 \sum_{i} p_i g_i}{\sum_{i} p_i^2 + \sum_{i} g_i^2},
\]

where \(N\) represents the number of pixels and \(g_i \in \{0, 1\}\) is used to distinguish between foreground and background. \(p_i \in (0, 1)\) represents the prediction result of the \(i\)-th pixel.

2.3. Lung Nodule Extraction and Sign Recognition Algorithm. Lung nodules have 3D features, so it is difficult to extract lung nodules from axial position alone. Therefore, we construct sagittal and coronal images according to axial images, improve U-Net network, and combine with bidirectional enhanced feature fusion network to enhance the extraction of lung nodules at different scales. The Mish activation function is introduced to shorten the network transmission time and improve the efficiency, as shown Figure 6.

The depth of the network is 5 layers, and the edge padding operation is used to replace the traditional U-Net crop operation, so that the output image size of the network is consistent with the input to realize feature fusion.

Because the high-level features mainly contain rich semantic information of the target, the low-level features mainly contain accurate location information of the target. Therefore, we build a two-way enhanced feature pyramid network to fuse the accurate low-level information with the high-level information through bottom-up path enhancement, so as to shorten the distance of information transmission.

Through two-way cross-scale connection, the low-level features of lung nodules are made full use and extracted, and the low-level fine-grained features with the high-level semantic features are better integrated. Feature vectors are enriched, the whole feature level is enhanced, and the utilization of features at all levels by the backbone network is improved. The network can also effectively extract the features of small nodules, so the problem of small target nodule loss is solved in the process of lung nodule segmentation.

Activation function is a way to introduce “nonlinearity” into neural network, which plays an important role in network training and evaluation. Liu et al. [19] proposed a new deep learning activation function, Mish activation function, which is a nonmonotonic, smooth, and continuous neural network activation function, and its function expression is as follows:

\[
f(x) = x \tanh \left( \ln \left( 1 + e^x \right) \right).
\]

Mish function retains a small amount of negative information, which can allow a small negative gradient to flow in, so the information flow is ensured and the gradient disappearance problem of ReLU function in the process of back propagation is eliminated.

The size of lung nodules is not regularly the same, but the lung nodules are spherical in space and are round-like in axial, coronal, and sagittal positions. In this case, we normalized the image to \(200 \times 200\) and extract radiomics and morphological features.

Spiculation sign is the main malignant sign of the lung. Therefore, we divided the data into training set and test.
t-test and LASSO algorithm are used to screen the discriminative features in the training set of radiomics label. Radiomics feature as first-order statistics, 3D shape, 2D shape, etc. which reflect the boundary and grayscale of lung nodules used for analysis. The method of 10 times cross-validation is used for fitting, and the omics logistic regression model is established. Conventional image features and clinical data are obtained by chi-square and t-test, and morphological logistic regression model is established. Integrate the two and make comprehensive statistics.

Through the above algorithm, it is applied to the images of lung nodules in axial, coronal, and sagittal positions, respectively. If there is at least one azimuth image that meets the spiculation sign, it is determined that the lung nodule is the spiculation sign.

### 3. Experiments and Results Analysis

The International Early Lung Cancer Action Project database is used in the experiment [35], and 500 sets of lung CT data are collected clinically. To verify the effectiveness of the algorithm, a database was constructed, including images with lung nodules, images of lung cancer, and images with grid-like increased density, as shown in Figure 7. Professional doctors are invited to outline the lung parenchyma and lung nodule areas and mark the signs of lung nodules.

The lung CT data is 16-bit data, and the image we see is an image displayed by dynamically adjusting the window width and window level, as shown in Figure 8. In terms of algorithm design, a container of 16 should be used.

Lung nodules smaller than 5 mm are called micronodules, which are generally benign and do not require special treatment. Lung nodules of 6-8 mm are partially solid nodules, so as to detect the possibility of malignant transformation as early as possible. Therefore, we focused on 6-8 mm lung nodules. Conduct studies and expand their lung nodules to a uniform size. Since spiculation has great threat to

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Md (%)</th>
<th>Vd (%)</th>
<th>Ud (%)</th>
<th>CM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level sets</td>
<td>79</td>
<td>21</td>
<td>20</td>
<td>79</td>
</tr>
<tr>
<td>Active contour model</td>
<td>83</td>
<td>20</td>
<td>18</td>
<td>81</td>
</tr>
<tr>
<td>Dual-branch residual network</td>
<td>86</td>
<td>17</td>
<td>15</td>
<td>84</td>
</tr>
<tr>
<td>Ours</td>
<td>91</td>
<td>15</td>
<td>11</td>
<td>88</td>
</tr>
</tbody>
</table>

![Figure 9: Effect graphs of lung parenchyma segmentation.](image-url)
human beings and high probability of malignancy, it becomes the focus of current research. The algorithm builds the model from the axial, coronal, and sagittal views and uses the circularity to preliminarily screen the region of interest through the segmentation algorithm to obtain the lung nodule region. Md, Vd, Ud, and CM are introduced to measure the segmentation performance of the algorithm.

\[
Md = \frac{E_g \cap E_s}{E_g \cup E_s} \times 100\% ,
\]

\[
Vd = \frac{E_s \cap E_s}{E_s} \times 100\% ,
\]

\[
Ud = \frac{E_g \cap E_s}{E_g} \times 100\% ,
\]

\[
CM = \frac{1}{3}\{Md + (1 - Nd) + (1 - Ud)\} \times 100\% ,
\]

where \(E_g\) is the area outlined by the doctor, \(E_s\) is the area extracted by the algorithm, and \(\cap\) represents XOR operation. Md and CM are directly proportional to the algorithm performance, and Vd and Ud are inversely proportional to the algorithm performance.

Table 3: Lung nodule recognition rate.

<table>
<thead>
<tr>
<th>FPF (%)</th>
<th>Axial</th>
<th>Axial + coronal</th>
<th>Axial + coronal + sagittal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isolated nodules</td>
<td>89</td>
<td>93</td>
<td>95</td>
</tr>
<tr>
<td>Vascular adhesive nodule</td>
<td>78</td>
<td>85</td>
<td>91</td>
</tr>
</tbody>
</table>

SEN, SPE, FPF, and ROC are introduced to measure the recognition performance:

\[
SEN = \frac{TP}{TP + FN} ,
\]

\[
SPE = \frac{TN}{TN + FP} ,
\]

\[
FPF = \frac{FP + FN}{TP + FP + TN + FN} .
\]

3.1. Performance of Lung Parenchyma Segmentation Algorithm. The performance of the algorithm is shown in Table 1. Reference [4] builds a 3D model to realize extraction of lung parenchyma. Reference [5] used graph cuts to build an energy model to guide lung parenchyma segmentation. Reference [10] constructed U-Net from the perspective of deep learning to extract lung parenchyma pairs and achieved certain positive results. The above algorithm is not effective for the extraction of lung nodules and lung parenchyma with lung wall adhesions that are relatively close. However, this paper proposes an algorithm to build an attention mechanism, reduces the amount of parameters, focuses on the lung parenchyma area, and realizes the extraction of lung parenchyma. The effect is shown in Figure 9, and good results are achieved.

3.2. Performance of Lung Nodule Segmentation Algorithm. The algorithm performance is shown in Table 2. Level set algorithm [14] focuses on the boundary of lung nodules and extracts lung nodules. Active contour model [16] extracts lung nodules. Dual-branch residual network [21] realizes lung nodule segmentation. The extraction effect of the algorithm proposed is shown in Figure 10. It can be seen that the algorithm constructs a bidirectional enhanced
feature pyramid network from three body positions to realize lung nodule segmentation, and the effect is better than other algorithms.

3.3. Performance of Lung Nodule Sign Recognition Algorithm.
We constructed axial, coronal, and sagittal models to judge the characteristics of lung nodules from three positions. Lung nodule recognition result is show in Table 3. It has high detectability for isolated lung nodules, and the detection rate of lung nodules with vascular adhesion is slightly low, but the overall trend is upward.

The lung nodules are spherical, and the lung nodules cannot be accurately targeted from only one position, so they are reflected from multiple positions. In the nodular region as shown in Figure 11(a), it will appear as a circle in a single position, and it may appear as a long strip from other positions as show in Figure 11(b), which can be judged as a nonlung nodule region. In turn, then the advantages of the multiposition algorithm are reflected.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>SEN (%)</th>
<th>SPE (%)</th>
<th>FPF (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texture</td>
<td>79</td>
<td>74</td>
<td>20</td>
</tr>
<tr>
<td>SVM</td>
<td>81</td>
<td>79</td>
<td>19</td>
</tr>
<tr>
<td>MSCN</td>
<td>84</td>
<td>83</td>
<td>18</td>
</tr>
<tr>
<td>Ours</td>
<td>91</td>
<td>88</td>
<td>15</td>
</tr>
</tbody>
</table>

The performance of the algorithm is shown in Table 4, and its ROC curve is shown in Figure 12. Ref [29] extracted texture features to classify lung nodules. Ref [27] constructed multiscale revolutionary neural networks to recognize the signs of lung nodules. The proposed algorithm realizes the signs of lung nodules by texture information and image radiomics features and has a good effect.

![Figure 11: Nodule and vessel images.](image)

![Figure 12: ROC.](image)
4. Conclusion

Aiming at the difficult problem of lung nodule segmentation and sign recognition, we constructed a lung nodule segmentation and recognition algorithm based on multiposition U-Net from the doctor’s diagnosis process. From the perspective of lung parenchyma segmentation, lung nodule segmentation, and lung nodule recognition, the attention mechanism model, multiposition feature enhancement model, morphology, and radiomics model are constructed, respectively, and finally, the automatic recognition of lung nodule signs can be realized, which can assist doctors to make accurate diagnosis.

Data Availability

The International Early Lung Cancer Action Project can be accessed through the link (https://veet.via.cornell.edu/lungdb.html). For the clinical data used to support the findings of this study, they are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors’ Contributions

Na Zhang and Jianping Lin contributed equally to this work.

Acknowledgments

This work is supported the Health Scientific Research Project of Shaanxi Province (2021B004).

References

their applications (ISSPA), pp. 870–875, Montreal, QC, Canada, 2012.


