A Lung Dense Deep Convolution Neural Network for Robust Lung Parenchyma Segmentation

YING CHEN, YERONG WANG, FEI HU, AND DING WANG
School of Software, Nanchang Hangkong University, Nanchang 330063, China

ABSTRACT Lung parenchyma segmentation is the prerequisite for an automatic diagnosis system to analyze lung CT (computed tomography) images. However, traditional lung segmentation algorithms have poor adaptability and are not effectively robust regarding lung databases with blood vessels and small voids which can interfere the segmentation. The main work of this paper is as follows: Firstly, a lung dense deep convolutional neural network (LDDNet) is proposed, which adopts some popular optimizer methods, such as dense block, batch normalization (BN) and dropout. The performance of LDDNet is tested on the public lung database LIDC-IDRI which contains many cases of interference for segmentation. Secondly, the labeled with blood vessels and small voids are not contained by the public ground-truth masks of the LIDC-IDRI database, therefore these regions are labeled by us with LabelMe software. Thirdly, for the aim of exploring the effect of image preprocessing on segmenting lung CT images with deep neural network, contrast enhancing, median filtering and Laplacian filtering are used to preprocess the image as comparative experiments. Finally, dataset is classified into four classes by the geometrical shapes to test the performance of LDDNet. The accuracy of the segmentation experiment reaches over 99% and the four classes can all reach over 95%. Additionally, blood vessels and small voids are segmented out from the lung parenchyma which is not achieved by other methods. Experimental results confirm that the proposed LDDNet can segment the lung parenchymal area more accurately and has better robustness in comparison with other neural networks and most of the traditional methods.

INDEX TERMS Deep dense neural network, lung segmentation, robustness.

I. INTRODUCTION Nowadays cancer has become a dangerous disease which is threatening human health to great extent. In the field of cancer, the highest mortality rate is obtained by lung cancer. Accompanied with CT imaging technology developing, CT images have become the most effective and feasible means to detect lung cancer. However, as the CT accuracy has been improved, the number of CT images, which are generated by each scan, is greatly increased too. The diagnostic workload of the radiologist is aggravated and it is possible to miss the diagnosis on account of fatigue. Qi et al. [1] proposed that there are more than 3.12 million new cancer cases per year in China and 2 million deaths from cancer patients. In addition, liver cancer has been replaced by lung cancer as the first cause of cancer death in China. Wu et al. [2] proposed that the 5-year survival rate of patients with advanced lung cancer is not higher than 15% in clinical studies. Furthermore, the 5-year survival rate of early lung cancer patients reaches 70%, indicating that the survival rate of lung cancer patients is closely related to early treatment. Therefore, the most effective and feasible solution to treat lung cancer is early detection.

Zhang et al. [3] proposed the truth that the accurate segmentation of lung parenchyma in lung CT images is an important step in the diagnosis and treatment of lung diseases. Furthermore, it is one of the main bottlenecks restricting the application of computer-aided detection technology in the field of pulmonary disease diagnosis. That traditional segmentation methods are difficult to segment at the boundary of lungs, is proposed by Yuan et al. [4]. On condition that there are blood vessels and small voids in the lung database, a good segmentation effect cannot be obtained by the traditional segmentation method. Meanwhile, most existing lung
segmentation algorithms based on deep neural networks can accurately segment the lung parenchyma regions. However, under nonideal conditions, it is still challenging to design robust lung segmentation algorithms, which can accurately segment the lung regions despite the effects of blood vessels and small voids and user cooperation. “User cooperation” means that the source of the CT images may be different, for instance hospitals and CT equipment. These factors will also affect the final lung parenchyma segmentation effect of the CT images.

The remainder of the paper is organized as follows: Section 2 outlines traditional lung parenchyma segmentation methods and modern neural network-based segmentation methods, for example FCNS, U-Net and so on. Section 3 describes in detail the proposed architecture for lung parenchyma segmentation. The details of public lung image databases and corresponding ground-truth masks are represented in Section 4. Moreover, environmental configuration of the experiment, network training parameters metrics and measurements for experiments are also described in Section 4. Experimental result evaluation and comparison are represented in Section 5. Section 6 summarizes this paper.

II. RELATED WORKS

A. OVERVIEW OF TRADITIONAL LUNG SEGMENTATION METHODS

Traditional method is not to use machine learning, and does not require relatively high equipment requirements. Anter and Hassenian [5] proposed that disadvantages of machine learning is long training time in comparison with the traditional methods, for instance, histogram-based, edge-based, region-based, model-based, watershed, clustering-based. Hence the traditional methods are less time consuming in executing. Traditional lung parenchyma segmentation algorithms do not have the accuracy rate of the neural network-based algorithms. However, less time is taken by the traditional algorithms. Hence traditional algorithms are still of great value.

1) LUNG SEGMENTATION ALGORITHMS BASED ON GRADIENTS OF LUNG CT IMAGES

A computer-aided region segmentation for the plain chest radiographs is proposed by Chondro et al. [6]. Further improvement for region boundaries is performed by utilizing a statistical-based region growing with an adaptive graph-cut technique that increases accuracy within any dubious gradient. Shariaty et al. [7] proposed a new method for lung segmentation in CT scans based on thresholding algorithm, which has a reconstruction operation to detect the attached nodules and add them to the lung mask. Gopalakrishnan and Kandaswamy [8] proposed a histogram utilized Adaptive Multilevel Thresholding (AMT) for estimating the total number of Gaussians and their initial parameters. In addition, the segmented lung parenchyma from the Gaussian Mixture model (GMM) undergoes an Adaptive Morphological Filtering (AMF) to reduce the boundary errors. Zhang et al. [9] proposed an object localization improved GrabCut algorithm for lung parenchyma segmentation that can automatically select the appropriate bounding box that relates to lung parenchyma, then use GrabCut algorithm. The algorithm can adapt to different forms of lung parenchyma and effectively improve the accuracy of segmentation. Traditional regional growth method is used by Tang et al. [10]. This method is used to preliminarily locate the lung boundary contour. Afterwards the lung boundary noise is removed, and the lung boundary is repaired by the adaptive curvature threshold method. Finally, the DRLSE model in the level set method is used to accurately segment the lung region. Hao et al. [11] proposed a novel automatic segmentation method based on an LBF active contour model with information entropy and joint vector. This method extracts the interest area of pulmonary nodules by a standard uptake value (SUV) in Positron Emission Tomography (PET) images, and automatic threshold iteration is used to construct an initial contour roughly. This method has great reference significance for the segmentation of lung parenchyma. Wei et al. [12] proposed in the objective function of image segmentation, based on the Chan-Vese model method. The local boundary statistical characteristic energy term is introduced to improve the accuracy and segmentation speed of lung medical image segmentation. These algorithms need to preprocess lung CT images according to prior knowledge before or during the location of the boundaries of lung parenchyma, which undoubtedly increases the complexity of the algorithms. In addition, the correlated segmentation steps and the various thresholds that need to be determined both make the algorithm very complex.

2) LUNG SEGMENTATION ALGORITHMS BASED ON PIXELS OF LUNG CT IMAGES

Qu et al. [13] proposed a kind of traditional algorithm, in which combined with the fuzzy C-means clustering algorithm. The left and right lung can be distinguished accurately by traditional threshold method to the accuracy of 0.9832 and 0.9807. Furthermore, in paper [13], the final lung parenchymal segmentation results were obtained by clustering, labeling and merging the super-pixel sub-regions of the same type. At last, the average accuracy of segmentation of CT images of the lungs reached 0.9946. Khan [14] proposed a novel approach for segmenting lung parenchyma using the combination of colour features and improved fuzzy-C means clustering in this paper. This method overcomes the disadvantages of existing CT lung parenchyma segmentation techniques since it combines the color features of different pixels present in the entire image. The combination of improved FCM clustering along with the colour features is the major advantage of this method. A method of lung parenchymal segmentation based on improved fuzzy C-means clustering and freeman chain code algorithm is proposed by Zhang et al. [15]. The improved fuzzy C-means clustering algorithm was used to reconstruct the CT image rough segmentation. Then the missing lung parenchyma
margin was repaired by the difference of three-chain code, which is generated by freeman chain code algorithm, to obtain the intact lung parenchyma region. Dharmalingham and Kumar [16] proposed a unique pathological lung segmentation method called reference-model based segmentation that uses shape property of human lung. The proposed segmentation approach constructs a reference lung model from input slices using a novel Sampling Lines Algorithm (SLA) and extracts the shape features. Nithila and Kumar [17] proposed a fast, new algorithm to segment the lung from CT images which places automatically the initial contour to locate the boundary and identify the concave edges. And the result of this algorithm is excellent compared with other active contour models. Peng et al. [18] proposed a hybrid semi-automatic method called Hull-Closed Polygonal Line Method (Hull-CPLM) to detect the boundaries of the lung Region of Interest (ROI). The first step of Hull-CPLM is that an image preprocessing method is constructed to implement the coarse segmentation by using as low as 15% of the manually delineated points as the initial points.

Majda and Abdelhamid [19] proposed an improved version of the standard graph cuts algorithm based on the Patch-Based similarity metric. The weights between each pixel and its neighboring pixels are based on the obtained new term. The graph is then created using these weights between its nodes. Finally, the segmentation is completed with the minimum cut/Max-Flow algorithm. Similar to lung parenchyma segmentation algorithm based on image gradients, the lung parenchyma segmentation algorithms based on pixels are also easily affected by blood vessels and small voids, which leads to a decrease in segmentation accuracy. These algorithms are also not adequately robust.

These traditional methods are less time consuming and less demanding on the equipment, but the accuracy of the segmentation is not high. Comparative analyses between our proposed method and other lung segmentation methods are summarized in TABLE 1.

### B. OVERVIEW OF SEGMENTATION METHODS OF LUNG PARENCHYMA BASED ON DEEP NEURAL NETWORK

In recent years, great success has been achieved by Convolutional Neural Network (CNN) in image classification [33]–[35], target recognition [36]–[38], detection in addition other fields. CNN is used as a feature extraction method by many networks, furthermore the network has been improved to achieve better results. At present, great results in the segmentation of lung parenchymal images is achieved by CNN. Long et al. [22] proposed Fully convolutional networks (FCNs) which adopts contemporary classification networks (AlexNet, the VGG net, and GoogLeNet) into fully convolutional networks and transfer their learned

### TABLE 1. Comparisons of previous studies on lung segmentation.

| Methods | Strengths | Weaknesses |
|---------|-----------|------------|
| Traditional lung segmentation algorithms based on gradients [6-12] | These methods of segmenting the lung parenchyma are less time consuming. The large number of data samples for training and high-quality laboratory equipment are not needed. | Finding the right threshold takes a considerable amount of time. These methods can not be sufficiently robust on unconstrained capture conditions and on low-quality lung databases. |
| Traditional lung segmentation algorithms based on pixels [13-18] | These methods are fast and do not need a large number of samples. | These are also not robust enough on low-quality lung databases. In addition, the result of the segmentation is easily affected by the blood vessels and small voids. |
| Lung segmentation algorithms based on the deep neural network methods [6][14] [20-32] | This type of methods has greatly improved the accuracy rate than traditional segmentation methods. For non-ideal lung images, the performance of segmentation is also good. | On the basis of the improved accuracy, these methods have not separated blood vessels and small voids in lung CT images, which will greatly affect the accuracy of lung cancer detection. And deep neural networks do not separate the glued parts well between the left and right lungs. |
| Our proposed method (LDDNet) | With the existence of blood vessels and small voids around the lung parenchyma, the good segmentation result is obtained by LDDNet. The sticky left and right lungs were segmented out. Furthermore blood vessels and small voids are segmented out from the lung CT images too. The accuracy of the LDDNet method is higher than that of existing deep neural network methods. | Dense connections are more time consuming. Moreover, comparing to some traditional lung segmentation methods and other deep neural networks, our proposed method requires higher laboratory equipment requirements. Furthermore, due to the memory capacity of the device, the lung images inputted into the LDDNet network can not be very large. |
representations by fine-tuning to the segmentation task. An encoder and decoder are contained in FCNs. Only the encoder is included in CNNs. Moreover, $1 \times 1$ sized spatial structure is used to replace $1 \times 1$ sized structure to acquire the prediction masks by FCNs. Yuan et al. [4] proposed an automatic segmentation algorithm for lung CT images based on U-Net, in which the accuracy of four classes reached 0.991, 0.978, 0.983, 0.997 respectively. Gaussian and Laplacian filtering are performed on the original CT images. Then the preprocessed images and the original images are taken as inputs respectively by using U-Net to execute segmentation. All the segmented lung regions are fused by linear regression to extract the parenchymal region of the lung. U-Net not only can be used to segment lung parenchyma. Wang et al. [39] proposed the structure of Dense U-Net for retinal vessel segmentation, which is similar to that of LDDNet. However, on LDDNet network, the Dense Block structure is only used in the Encoder part, not in Decoder part. On contrast the Dense Block structure is utilized in both encoder and decoder part by Dense U-Net. In addition, the method adopted in this paper is to divide the image into small pieces and then perform separate segmentation. After the segmentation is completed, the segmented small images are combined to obtain the final segmentation result. Recently, Neural Architecture Search (NAS) has successfully identified neural network architectures for semantic image segmentation, which exceed human designed ones on large-scale image classification. Liu et al. [31] proposed to search the network level structure in addition to the cell level structure, which forms a hierarchical architecture search space. Auto-DeepLab, the architecture searched specifically for semantic image segmentation, attains state-of-the-art performance without any ImageNet pretraining. Isensee et al. [32] presented nnU-Net (‘nonew-Net’), a framework that automatically adapts itself to any given new dataset. The nnU-Net strips away the architectural bells and whistles that are typically proposed in the literature and relies on just a simple U-Net architecture embedded in a robust training scheme. Ronneberger et al. [40] used U-Net to segment biomedical images. The U-Net architecture is shown in the FIGURE 1.

Harrison et al. [23] proposed progressive holistically-nested networks (P-HNNs) for pathological lung segmentation, which is a bottom-up deep-learning based approach. This method is expressive enough to handle variations in appearance, while remaining unaffected by any variations in shape. P-HNNs incorporate the deeply supervised learning framework. It is enhanced with a simple, effective, progressive multi-path scheme, which more reliably merges outputs from different network stages by P-HNNs. Jin et al. [24] developed a 3D generative adversarial network (GAN), which effectively learns lung nodule property distributions in 3D space. They used GAN to generate simulated training images where nodules lie on the lung border, which are cases where the published P-HNN (progressive holistically-nested network) model struggles. By this method, the P-HNN model learns to better segment lung regions under challenging situations. Gordienko et al. [25] proved the efficiency of 2D CXR analysis by lung segmentation and skeletal
exclusion techniques by using deep learning methods. A lung CT image segmentation using the U-net architecture is proposed by Skourt et al. [26], which is one of the most used architectures in deep learning for image segmentation. This network can be trained end-to-end from very few images and outperforms many methods. A multi-stage training strategy, network-wise training, is proposed by Hwang and Park [27], in which the current stage network is fed with both input images and the outputs from pre-stage network. It is shown that this strategy has an ability to reduce falsely predicted labels and produce smooth boundaries of lung fields. Deep convolutional neural networks (NNS) architectures are used for automated multiclass segmentation of anatomical organs in chest radiographs (CXRs), namely for lungs, clavicles, and heart by Novikov et al. [28]. Delayed subsampling, exponential linear units, highly restrictive regularization, and a large number of high-resolution low-level abstract features are used by them. Anthimopoulos et al. [29] introduced a deep purely convolutional neural network for the semantic segmentation of interstitial lung diseases. The training was performed in an end-to-end and semisupervised fashion, utilizing both labeled and nonlabeled image regions. Dai et al. [30] proposed Structure Correcting Adversarial Network (SCAN) to segment lung fields and the heart in CXR images. A critic network is incorporated to impose on the convolutional segmentation network the structural regularities inherent in human physiology by SCAN. This method generalizes well to CXR images from different patient populations and disease profiles. Moreover, by using only very limited training data available, this model reaches human-level performance without relying on any pre-trained model. The segmentation accuracy of these deep neural network methods has been greatly improved, but it is more time-consuming than traditional methods. Further, it does not segment small voids and blood vessels in the lung CT images, nor does it solve the problem of left and right lung adhesion.

III. THE PROPOSED ARCHITECTURE FOR LUNG PARENCHYMA SEGMENTATION

A. FLOWCHART OF THE LUNG PARENCHYMA SEGMENTATION SYSTEM

An encoder-decoder structure is used by our lung parenchyma segmentation system. The input lung CT images are encoded into a fixed-length vector representation by the encoder. Further, this fixed-length vector is decoded into the final segmentation result by the decoder. The original lung images serve as the input images. Furthermore, the output is binary images, in which the lung parenchyma is represented by pixel value of 255 and the background is represented by the pixel value of 0. Flowchart of the lung parenchyma segmentation system is demonstrated by FIGURE 2.

B. LDDNet FOR LUNG CT SEGMENTATION

We proposed a lung dense deep convolutional neural network (LDDNet). LDDNet is a deep neural network with direct connections between any two layers. The input of each layer of the network is the fusion of all the previous layer outputs. Besides the feature map learned by this layer is directly passed to all subsequent layers as input. In addition, the depth of the network can be reduced without reducing the performance of segmentation. Additionally, the same segmentation effect can be achieved as the deeper network structure by this network. Moreover, this kind of structure can well alleviate the disappearance of gradients, save parameters, ease underfitting, and reuse features. Furthermore, many convolutions, are used, which size are $1 \times 1$. Through this method, the dimension can be reduced. For instance, the amount of calculation can be reduced from $4 \times 4 \times 5$ to $4 \times 4 \times 3$ through $1 \times 1$ convolutions. As a result, the accuracy rate is slightly improved. Network structure of LDDNet can be seen in FIGURE 3.

The Dense-Block structure is used in the LDDNet. Every Dense-Block is composed of layers of Batch Normalization, Relu, and Convolution which size is $3 \times 3$, in addition to a layer of Concatenation. We can see from FIGURE 3, FIGURE 4 that the Dense-Block structure is adopted by the LDDNet network. As demonstrated in FIGURE 4, the top convolutional layers connect the later convolutional layers. Moreover, each convolutional layer is followed by the rectified linear unit (ReLu) layers and batch normalization (BN) layers. Assuming there are L convolutional layers, the traditional convolution network has L connections, however there are $L \times (L+1)$ connections in the LDDNet based on the dense block. There is a direct connection between every two layers, which makes the most of features of lung CT images. The effect is that features are reused to the greater extent which is the goal that deep neural network has been pursuing. The overfitting and gradient vanishing effectively are alleviated by dense-block, which adds the larger feature value of the bottom layers to the small feature value of the top layers.
In addition, since there is currently no lung CT image dataset with a large data size. Hence this will cause the deep neural network to underfit in training. Besides the disappearance of the gradient during the training process will seriously restrict the improvement of the accuracy of the neural network. The structure of Dense-Block can alleviate this problem for some extent for the dense connections.

The network structure of LDDNet is shown in TABLE 2. This neural network can be feature-multiplexed in the case of a small dataset size, reducing the speed at which the gradient disappears. The network used in this paper is divided into two parts. The first half of the network is a convolution process, namely for LDDNet-encoder. The second half of the network is LDDNet-decoder in which deconvolution is used besides fusing the output of the average pooled layer. The convolutional layer is abbreviated as ‘Conv’. The deconvolution layer is abbreviated as ‘Conv_d’. The final output of LDDNet-encoder is defined as ‘Final’. The fusion of the output of the pooling layer and the deconvolution layer is abbreviated as ‘ADD’. The dropout layer is abbreviated as ‘Dt’.

In the part of LDDNet-encoder, the size of input image of this experiment is $512 \times 512 \times 1$. Initially through Conv1, in which the number of convolution kernels is 64, the size is $3 \times 3$ and the stride is 1. Furthermore, five Dense-Blocks follow behind Conv1. Each Dense-Block contains two convolutional layers, in which the number of convolution kernels is 64, the size is $3 \times 3$ and the stride is 1. The number of channels is increased in the graphics to 128 by each Dense-Block. Then five pooling layers are connected after each Dense-Block. The pooling layer we take is average pooling, in which the specification is $2 \times 2$, and the stride is 2. The pooling layer narrows the image size. Following the above structure are two convolutional layers for instance, Conv2 and Conv3. Conv2 is composed of 8192 kernels with $7 \times 7$ size and step 1. Conv3 is composed of 8192 kernels with $1 \times 1$ size and step 1. The Final layer is composed of n kernels with $1 \times 1$ size and step 1, where n is 2 in this paper, namely for the lung pixels and non-lung pixels. Besides two converged concatenations are also contained. In addition, followed by Conv2 and Conv3 respectively, there is a dropout layer, in which the probability of setting the discard is 0.5.

In the part of LDDNet-decoder, Conv_d1, Conv_d2, Conv_d3, Conv_d4, Conv_d5 are performed to obtain the segmentation result. Meanwhile, the size of Conv_d1 is the same as the output size of Pool4. Besides Conv_d1 is merged with Pool5 to get Add1, Conv_d2 is merged with Pool4 to get Add2, Conv_d3 is merged with Pool3 to get Add3, Conv_d4 is merged with Pool2 to get Add4, Conv_d5 is merged with Pool1 to get Add5. Then through the structure BN-Relu-Conv8, Conv8 is composed of 192 kernels with $3 \times 3$ size and step 1. Through the structure BN-Relu-Conv9, Conv9 is composed of 192 kernels with $1 \times 1$ size and step 1.
### TABLE 2. The structure of the LDDNet.

| Layer Name     | Type/Kernel/Stride                                      | Output Size   |
|----------------|--------------------------------------------------------|---------------|
| Input          |                                                        |               |
| Conv 1         | Convolution/3x3/1/64, Batch Normalization, ReLU,       | 512x512x1     |
|                | Convolution/3x3/1/64, Batch Normalization, ReLU,       | 512x512x64    |
|                | Concatenation                                          |               |
| Dense-Block 1  |                                                        | 512x512x192   |
| Average Pooling layer 1 | Average-pooling/2x2/2, Batch Normalization, ReLU,       | 256x256x192   |
| Dense-Block 2  | Convolution/3x3/1/64, Batch Normalization, ReLU,       |               |
|                | Concatenation                                          |               |
| Average Pooling layer 2 | Average-pooling/2x2/2, Batch Normalization, ReLU,       | 128x128x320   |
| Dense-Block 3  | Convolution/3x3/1/64, Batch Normalization, ReLU,       |               |
|                | Concatenation                                          |               |
| Average Pooling layer 3 | Average-pooling/2x2/2, Batch Normalization, ReLU,       | 64x64x448     |
| Dense-Block 4  | Convolution/3x3/1/64, Batch Normalization, ReLU,       |               |
|                | Concatenation                                          |               |
| Average Pooling layer 4 | Average-pooling/2x2/2, Batch Normalization, ReLU,       | 32x32x576     |
| Dense-Block 5  | Convolution/3x3/1/64, Batch Normalization, ReLU,       |               |
|                | Concatenation                                          |               |
| Average Pooling layer 5 | Average-pooling/2x2/2, Convolution/7x7/1/8192, Dropout (50%) | 16x16x704     |
| Conv 2         |                                                        | 16x16x8192    |
| Dt 1           | Dropout (50%)                                          |               |
| Conv 3         | Convolution/1x1/1/8192, Dropout(-50%)                  | 16x16x8192    |
| Final          | Deconvolution/1x1/1/2, Addition(+Pool5)                | 16x16x2(Class number) |
| Conv_d1-Add1   |                                                        |               |
| BN-Relu-Conv4  | Batch Normalization, ReLU, Convolution/3x3/1/576      | 16x16x704     |
| Conv_d2-Add2   | Deconvolution/3x3/2/576, Addition(+Pool4)             | 16x16x576     |
| BN-Relu-Conv5  | Batch Normalization, ReLU, Convolution/3x3/1/448      | 32x32x576     |
| Conv_d3-Add3   | Deconvolution/3x3/2/448, Addition(+Pool3)             | 32x32x448     |
| BN-Relu-Conv6  | Batch Normalization, ReLU, Convolution/3x3/1/320      | 64x64x448     |
| Conv_d4-Add4   | Deconvolution/3x3/2/320, Addition(+Pool2)             | 64x64x320     |
| BN-Relu-Conv7  | Batch Normalization, ReLU, Convolution/3x3/1/192      | 128x128x320   |
| Conv_d5-Add5   | Deconvolution/3x3/2/192, Addition(+Pool1)             | 128x128x192   |
| BN-Relu-Conv8  | Batch Normalization, ReLU, Convolution/3x3/1/192      | 256x256x192   |
| BN-Relu-Conv9  | Batch Normalization, ReLU, Convolution/1x1/1/192      | 256x256x192   |
| Conv_d6        | Deconvolution/3x3/2/192, Addition(+Pool1)             | 512x512x2(Class number) |
| prediction     |                                                        | 512x512x1     |
As described above, the result of deconvolution is as follows: Conv_d1, Conv_d2, Conv_d3, Conv_d4, Conv_d5, Conv_d6; fusion result: Add1, Add2, Add3, Add4, Add5. The final output segmentation result is a $512 \times 512$ graph, which is a binarized segmentation result. The number of classes in the experiment is 2, namely for the lung parenchyma area and the background area. In the segmentation results, the pixel value of 255 represents the part of the lung parenchyma and the pixel of 0 represents the background.

### C. THE DIFFERENCE BETWEEN LDDNet AND U-Net

U-Net is improved based on FCN. The difference is that U-Net does not simply encode and decode pictures like FCN. U-Net accurately locates the high-pixel features extracted from the shrink path. The new feature map is combined during encoder to preserve some important feature information from the previous decoder process.

As can be seen from FIGURE 3 and FIGURE 1, only L connections are contained by the traditional L-layer convolutional neural network, while $L \times (L+1)/2$ connections are contained by LDDNet. The difference between U-Net and LDDNet is that in the encoder part of the neural network, the network layer in the U-Net does not have a direct connection with all the previous layers. In contrast LDDNet has a direct connection between each layer and the previous network layer. The fusion of all the previous layers is the input of the next layer.

As can be seen from FIGURE 3 and FIGURE 1, the U-Net network here is very similar to the dense network. The obvious difference between U-Net and LDDNet is the following Equation 1 and Equation 2.

\[ C_n = W_n(C_{n-1}) + C_{n-1} \]
\[ C_n = W_n(C_0, C_1, C_2, \ldots, C_{n-1}) \]

The Equation 1 is U-Net. Here, n represents the n layers, $C_n$ represents the output of the n layers, $W_n$ represents a nonlinear transformation. For U-Net, the output of the n-layer is the output of the n-1 layer plus the nonlinear transformation of the output of the n-1 layer. The Equation 2 is LDDNet. $[C_1, C_2, \ldots, C_{n-1}]$ indicates that the output feature map of the 0 to n-1 layer is concatenation. Concatenation is the merging of channels. U-Net is the sum of values, and the number of channels is constant. $W_n$ includes convolutions of BN, ReLU and $3 \times 3$. The output of the n layer is the concatenation of the output of the 0 to n-1 layer. In the end of the Dense-decoder, cross entropy function is adopted as the cost function.

### IV. EXPERIMENTAL CONFIGURATION

#### A. DESCRIPTION OF THE LUNG CT IMAGE DATABASE

1) PROPERTIES FOR CT IMAGES

Diagnostic, lung cancer screening chest tomography (CT) scans and indicates annotated lesions are included in Lung Image Database Consortium’s image collection (LIDC-IDRI) [41]. It is an internet-accessible international resource for developing, training, and evaluating computer-aided diagnostic (CAD) methods for lung cancer detection and diagnosis. LIDC-IDRI is sponsored by the National Cancer Institute (NCI), further promoted by the National Institutes of Health (FNIH) Foundation, and participated by the Food and Drug Administration (FDA). Seven academic centers and eight medical imaging companies collaborated to create this dataset which contains 1018 cases and 244527 images. Each subject includes images from a clinical thoracic CT scan and an associated XML file that records the results of a two-phase image annotation process performed by four experienced thoracic radiologists. Each radiologist independently reviewed each CT scan and marked lesions belonging to one of three categories (“nodule > or =3 mm,” “nodule <3 mm,” and “non-nodule > or =3 mm”). The detailed parameters of LIDC-IDRI are as shown in TABLE 3.

In fact, lung CT images are acquired in DICOM format by CT machine in hospitals. However, DICOM format must be converted to image format to be suitable for deep neural network training. Here PNG format was selected as the image format for LDDNet. In addition, the network training will be burdened by the extra information in the DICOM format other than images. The class packet PYDICOM is used to extract the image information inside the DICOM format. Furthermore, the image is saved in the PNG format by the class packet CV2 (OPENCV).

TABLE 3. Detailed parameters of the dataset LIDC-IDRI.

| Collection Statistics | Details |
|-----------------------|---------|
| Modalities            | CT (computed tomography) |
|                       | DX (digital radiography) |
|                       | CR (computed radiography) |
| Number of Patients    | 1010 |
| Number of Studies     | 1038 |
| Number of scans       | 1018 CT |
|                       | 290 CR/DX |
| Number of Images      | 244527 |
| Image format          | DICOM |
| Image Size (GB)       | 124 |

2) ENHANCEMENT AND LABEL FOR CT IMAGES

The number of lung CT images for experiment, which are chosen from dataset LIDC-IDRI, is relatively small. Hence the method of enhancing data is executed by us. The images are expanded by a factor of 20 in this paper. The methods like rotation, translation, clip, blurry and so on, are used in this paper. The instances of the enhancement of lung images can be seen from FIGURE 5. The images in the first line are the lung that is panning. Furthermore, the images in the bottom line are rotated.

The dataset LIDC-IDRI has been published and there is no segmentation of the image. In this paper, the database is manually labeled by us with the software LabelMe. To evaluate robustness of LDDNet, we adopt software LabelMe for labeling the lung regions and the noise regions occluding the lung regions. First, the outer ring of lung is labeled as ‘_lung_’. Second, the external region of lung is labeled as ‘_background_’. Finally, the blood vessels and spaces occluding the lung area are labeled as ‘_background_’. From the
third row of column c in FIGURE 9, the second row of column c in FIGURE 15, we can see that our LDDNet has good robustness. The blood vessels and spaces are segmented out from the lung parenchyma region. The process of labeling is manifested in FIGURE 6.

3) CLASSIFICATION OF CT IMAGES
This experiment was conducted on the LIDC-IDRI database. The images for experiment were derived from LIDC-IDRI and the selected images were manually labeled. Two teachers and six graduate students participated in the work of labeling. These people are engaged in lung CT image segmentation and detection research in the school of software. Furthermore, our work of labeling has been validated by the radiologist of Jiangxi Cancer Hospital in China in Jiangxi Province. Our database was marked by eight people, and each person completed five groups. The number of pictures in each group is 100. The number of pictures we selected is 4000. Among them, the number of training set images, verification set images and test set images account for 2400, 800, 800. The ratio is 0.6, 0.2, 0.2. The detailed parameters of the dataset are displayed in TABLE 4.

Database LIDC-IDRI consists of complete lung regions and lung regions occluded by blood vessels and small voids. LIDC-IDRI comes from different collection equipment and different people. Furthermore, patients’ lungs also have different health conditions. For example, canceration of the lungs or stagnant water in the lungs, result in different geometrical shapes of the patient’s lung CT images. In order to fully verify the robustness of the network, the images are divided into four classes according to the geometry. Furthermore, each class is trained with LDDNet to test the segmentation performance. The classification criteria are whether the left and right lungs are clear, whether the left and right lungs

| Properties                    | LIDC-IDRI |
|-------------------------------|-----------|
| Number of groups              | 40        |
| Number of Image per group     | 100       |
| Percentage of training        | 60%       |
| Percentage of validation      | 20%       |
| Percentage of testing         | 20%       |
| Image Size                    | 512x512x1 pixels |
| Image Format                  | PNG       |
| Original CT format            | DICOM     |
| Number of the images          | 4000      |
| Number of training epoch      | 200       |
| Number of class one images    | 1234      |
| Number of class two images    | 617       |
| Number of class three images  | 1210      |
| Number of class four images   | 939       |
are adhered, whether they are upper slices of the CT image of the lung, and whether the left and right lungs are symmetrical. The specific characteristics of each class are described below.

As can be seen from FIGURE 7, all of lung CT images are classified into four classes:

1. **Class 1**: Class 1 is lung images whose left lung images and right lung images are clear, symmetrical. Such images occupy the largest proportion in the dataset in TABLE 4.

2. **Class 2**: Class 2 is lung images whose right lung images and left lung images are bonded. This kind of sticky image is relatively difficult to segment. In addition, the proportion of this kind of images in the dataset is the smallest. But the accuracy of the segmentation effect is still above 0.95. It shows that LDDNET has good robustness. As is shown in the red circles in FIGURE 7 class 2.

3. **Class 3**: Class 3 is lung images whose images are top CT slices. This class of image belong to the upper slice of the lung CT images. The left and right lungs are relatively small, and there are bronchial tubes among the imaged, as shown in the red circles in FIGURE 7 of class 3.

4. **Class 4**: Class 4 is lung images whose left lung and right lung are not symmetrical. This class of image has a smaller proportion of the right lung than the left lung as shown in the red circles in FIGURE 7.

### B. EXPERIMENT ENVIRONMENTAL CONFIGURATION AND NETWORK TRAINING PARAMETERS

In this paper, TABLE 5 shows the experiment environment details. Because the memory of our experimental equipment is relatively small. The size of memory is 32G. Hence, the batch is set to 2. The learning rate is set to 0.00001. The dropout rate is set to 0.5. The number of the iteration is set to 50000. From FIGURE 8, it can be seen that the model accuracy gets steady after 30000 of the iteration steps. Hence, we set 50000 as the iteration number to ensure that the training convergence is enough, in addition to make the most use of the lung CT images.

We assume that $J$ is the cost function. $\tilde{f}_j(z)$ denotes the desired label probabilities of the training image and $f_j(z)$ denotes the actual network output:

$$J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{c} \tilde{f}_j(z) \log [f_j(z)]$$

For optimization, the mini-batch Adam algorithm is adopted to minimize the cost function. The lung CT images are composed of mostly non-lung regions and a few lung regions, which lead to the adoption of the standard stochastic gradient descent (SGD) algorithm. Adam algorithm increases the learning rate for sparse data and decreases it for common data, and updates quickly for sparse features and slowly for common features.

### C. METRICS AND MEASUREMENTS FOR EXPERIMENTS

For the segmentation results of the experimental images, the Jaccard Similarity (JS) is used to evaluate the accuracy of the experimental segmentation. The CT image segmentation results of the lungs were evaluated. Jaccard Similarity (JS) was utilized as the evaluation index. As is shown in the following equation 4:

$$JS(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

where $A$ represents the area segmented by the algorithm, in addition $B$ represents the area marked manually. When the result of neural network segmentation is better, the value of
JS is larger. When the result of neural network segmentation is worse, the value of JS is smaller. The value of JS ranges from 0 to 1. Wherever FP (false positive) and FN (false negative) denote the misclassification number of non-lung pixels and lung pixels in test images respectively. TP (true positive) and TN (true negative) denote the recognition number of lung pixels and non-lung pixels in test images respectively. The accuracy represents the correct segmentation pixels and is calculated as follows.

$$\text{acc} = \frac{TP + TN}{TP + FN + FP + TN} \times 100\% \quad (5)$$

The precision(p), recall(r), Specificity(SP), Sensitivity(SE), Dice similarity coefficient(DSC) are calculated as follows:

$$p = \frac{TP}{TP + FP} \quad (6)$$
$$r = \frac{TP}{TP + FN} \quad (7)$$
$$SP = \frac{TN}{TN + FP} \quad (8)$$
$$SE = \frac{TP}{TP + FN} \quad (9)$$
$$DSC = \frac{2 \times TP}{(TP + FP) + (TP + FN)} \quad (10)$$

V. EXPERIMENTAL RESULTS EVALUATION AND COMPARISON

A. EFFECT OF THE PRE-PROCESS OF LUNG CT IMAGES

Image preprocessing method is combined with the LDDNet neural network to obtain the segmentation result of the lung CT image. For the aim of testing the accuracy and robustness of our proposed method, no preprocessing methods are obtained before inputting the lung CT images into the LDDNet except the contrast experiment in TABLE 6. Image preprocessing will reduce or improve the performance of deep neural networks. One of our work will compare the impact of some common preprocessing methods on LDDNet. The preprocessing methods of the lung CT images are composed of enhanced contrast, median filtering, Laplacian filtering. The median filtering method is a non-linear smoothing technique that sets the gray value of each pixel to the median of the gray values of all pixels in a neighborhood window at that point. Laplacian is a kind of differential operator. Its application can enhance areas with abrupt changes in gray levels and weaken areas with slow changes in gray levels. For contrast-enhancing experiments, we use a mask to enhance image contrast to process the image. Contrast enhancing, median filtering and Laplacian filtering are utilized to preprocess the images before the images are inputted into the LDDNet, meanwhile LDDNet with no preprocessing is executed as the comparison. The experiments results are shown in TABLE 6.

As can be seen from TABLE 6, some special preprocessing methods can increase the accuracy of LDDNet. There are the model accuracy rates of LDDNet for different preprocessing methods. The accuracy of LDDNet with no preprocessing methods can reach 0.994378, meanwhile the accuracy of preprocessing of median filter, Laplacian filter, enhanced contrast can reach 0.994726, 0.994667, 0.985178. The preprocessing of median filter, Laplacian filter can improve accuracy on LDDNet, but enhanced contrast cannot.
TABLE 6. The comparison of accuracy of the proposed methods.

| Method                        | Acc   | Precision | recall | DSC  | Sensitivity | Specificity |
|-------------------------------|-------|-----------|--------|------|-------------|-------------|
| Fuzzy C-Means Clustering [14]| 0.963100 | □         | □      | □    | 0.810900    | 0.968200    |
| Thresholding and Gaussian Mixture Modelling [8] | 0.972200 | □         | □      | □    | 0.972200    | 0.999100    | 0.967000    |
| Patch-Based Similarity [19]  | □     | □         | □      | □    | 0.839800    | □           |
| LDDNet with median filter     | 0.994726 | 0.996882  | 0.996912 | □ | 0.996897    | 0.996912    | 0.982357    |
| LDDNet with Laplacian filter  | 0.994667 | 0.997027  | 0.996698 | □ | 0.996863    | 0.996698    | 0.983145    |
| LDDNet with enhanced contrast | 0.985178 | 0.989843  | 0.992689 | □ | 0.991264    | 0.992689    | 0.943552    |
| LDDNet with no preprocessing  | 0.994378 | 0.996779  | 0.996619 | □ | 0.996699    | 0.996619    | 0.981508    |

Furthermore, the accuracy of LDDNet is higher than the traditional methods [8], [14], [19]. From above, the conclusion can be drawn that LDDNet has better performance than most of the traditional methods. Besides, some special preprocessing methods will improve the performance of LDDNet in the field of lung parenchyma segmentation, for instance median filter, Laplacian filter.

B. EFFECT OF THE CT LUNG IMAGE CLASSIFICATION

The following experimental images are the segmentation results of the four class of lung CT images by LDDNet. In addition, there are segmented images on the bottom in FIGURE 9. Furthermore, the blood vessels and small voids can be segmented out from the lung region in class 1, class 2, class 3, and class 4, from which the robustness of LDDNet can be proved in red circles in figure 9. The specific parameters of the four classes of image segmentation results can be seen in TABLE 13.

As shown in FIGURE 9 (class 2), the CT section of the lung is a relatively tight junction between the left and right lungs. Artificially, it can be well marked by magnifying the image. Nevertheless, good separation of the left and right lungs is not obtained by the traditional method.

The following FIGURE 10, FIGURE 11, FIGURE 12, and FIGURE 13 show the segmentation results of each class of images in detail. In each figure, column a represents original lung CT images, column b represents ground-truth, at last column c represents the segmentation results of the LDDNet network. In images of column d, the false positive and negative errors are presented as green and red.

For relatively good images of class 1, our network segmentation results are good. It can be seen from column a in FIGURE 10 that the images of class 1 are relatively complete of the left and right lungs, and the left and right lungs are not adhered. The image quality of this class is relatively high. In addition, the segmentation effect of this class of image is also good for LDDNet. It can be seen from the first row and third row of column d in FIGURE 10 that not only the lung parenchyma can be segmented, but also the blood vessels and small voids can be well separated here.

For the relatively good images of the CT lungs with respect to the left and right lungs, it can be seen from the first row of column d in FIGURE 11 that the left and right lungs have been segmented. Furthermore, the segmentation effect is very good for the edge regions of the lungs, for example, the red circles marked column d. As can be seen from column a in FIGURE 11, the images of this class are all images in which the left and right lungs are bonded. The image quality of this class of image is worse than that of class one, and the segmentation effect of this class of images on LDDNet is also relatively poor. It can be seen from column d in FIGURE 11 that the segmentation of the lung parenchyma is realized, moreover the bonded parts of the left and right lungs are also divided.
As for the good images of the upper part of the CT lung slice, as can be seen from FIGURE 12, image size of the upper part is smaller than that of class 1 and class 2, which makes segmentation difficult. As can be seen from the second row of column d, the effect of segmentation is good. As can be seen from column a in FIGURE 12, the images of class 3 are all images of the upper slice of the lung CT. The image quality of this class of image is worse than that of class 2, and the size of the left and right lungs in the image is smaller. The segmentation performance of this class of images for LDDNet is also relatively poor. It can be seen from column d in FIGURE 12 that the lung parenchyma is realized, furthermore the edges of the lung parenchyma are relatively well divided, for example, the red circled parts of column d.

As for the asymmetrical images of the left and right lungs in the CT image, it can be seen from FIGURE 13 that the segmentation results are also very good. It can be seen from column a in FIGURE 13 that the images of class 1 are asymmetrical images of the left and right lungs, and most of these images are missing in the upper part of the right lung. The image quality of this class of image is worse than that of class 1 but better than that of class 2 and class 3. The segmentation effect of this class of images for LDDNet is also relatively poor. It can be seen from column d in FIGURE 13 that the lung parenchyma is realized, in addition the upper edge of the right lung here is relatively smooth. For example, column d is marked with red circles.

From FIGURE 10, FIGURE 11, FIGURE 12, and FIGURE 13 above we can know that the good segmentation results for the four classes of images is obtained by LDDNet. From this, it can be proved that LDDNet is robust on lung parenchyma segmentation.

C. EFFECT OF THE BLOOD VESSELS AND SMALL VOIDS
Lung parenchyma segmentation algorithm can be interfered by the blood vessels and small voids. These regions cannot be segmented out by traditional method. Furthermore, these regions are not obtained by their ground-truth masks too. Here is a total of three CT slices of processing results in FIGURE 14 and FIGURE 15. Original lung CT images are revealed by column a. Ground-truth masks are shown through column b. Network segment results are represented by column c. Visual segmentation results are exhibited in column d. In images of column d, the false positive and negative errors are presented as green and red. LDDNet is used for segmentation as the contrast experiment for effect of the blood vessels and small voids. Effect of the blood vessels and small voids for LDDNet for segmentation of lung parenchyma on lung CT images can be seen from FIGURE 14 and FIGURE 15.
From the red circles marked in FIGURE 14 of column c, the blood vessels are segmented out from the lung parenchyma by LDDNet. In the third row of column c in FIGURE 14, although the blood vessel segmentation in the red circle is not very clear, but the basic outline still can be segmented well. In FIGURE 14, it can be indicated from the first row of column d that the left and right lung were successfully separated.

It can be seen from the red circles marked in FIGURE 15 that LDDNet also has a good effect on the segmentation of lung parenchyma. From column c in FIGURE 15 that small voids in the lung parenchyma have been segmented out. However, in order to further test the robustness of LDDNet, we labeled the small voids in the lung parenchyma, but some small voids were not labeled. In the second row and third row of column b of FIGURE 15, we did not label the small voids, but let LDDNet directly perform the segmentation experiment. The results, manifested from the second row and third row of column c in FIGURE 15, is that LDDNet still segments out the small voids from the lung parenchyma. From the above two sets of comparative experiments, a conclusion can be made that LDDNet network is robust in segmenting the lung parenchyma which has interference terms for instance small voids and blood vessels.

**D. EFFECT OF INPUT IMAGE SIZE**

Medical CT images are generally huge. However, it is impossible to enter the original lung CT image into LDDNet during segmentation. Hence, in this paper, the image is resized to $512 \times 512$ and other pixel sizes for comparison.

We make follow experiments to compare the effect of image size on LDDNet in TABLE 7. As demonstrated in TABLE 7, LDDNet is trained by different sizes of the input images, $512 \times 512$ and so on. Additionally, the performance of the bigger image size is better. Because the reduction of the image will lead to the loss of semantic information, eventually the LDDNet network cannot learn enough features. LDDNet needs to learn the pixel level information of the images during training. For instance, after reducing the size of the input image, the pixel information of the $512 \times 512$ image is sixteen times the pixel information of the $128 \times 128$ image. Hence, when the pixels obtained by LDDNet are insufficient, the deep neural network cannot learn sufficient segmentation performance by itself. At last, images with larger pixels will have a higher accuracy than images with smaller pixels by LDDNet. It can be concluded from the TABLE 7 that as the image pixels become smaller, the effect of LDDNet segmentation of lung parenchyma is gradually getting worse.
E. EFFECT OF TYPE OF POOLING

The most commonly used pooling layers are maximum pooling and average pooling. Average pooling is to average only the feature points in the neighborhood. This pooling can reduce the error caused by the increase in the variance of the estimated value in the limited size of the neighborhood. Maximum pooling is to take the maximum of the feature points in the neighborhood. This kind of pooling can reduce the convolution layer parameter error which is caused by the deviation of the estimated mean value.

Two different types of the pooling layers are used for comparative experiments. The result indicated in TABLE 8, the max pooling has a little better accuracy than average pooling for LDDNet. Precision, recall, DSC, sensitivity, specificity, Jaccard Similarity in two types of pooling are very similar. Even the DSC is the same. So, a conclusion can be made that the average pooling and maximum pooling have very little effect on LDDNet for lung parenchyma segmentation. The structures of the two types of pooling layers is revealed in FIGURE 16 below.

To further explore the effect of pooling layer function on LDDNet segmentation of lung parenchyma, we designed comparative experiments. We designed comparative experiments. There are four-layer average pooling, four-layer maximum pooling, five-layer average pooling, five-layer maximum pooling, six-layer average pooling, and six-layer maximum pooling. The results of the comparative experiment can be seen in Table 9.

As can be seen from the TABLE 9, the effect of the pooling layer on the experimental results is very subtle. The pooling layer has not greatly improved the effect of LDDNet on segmenting lung parenchyma. When the type and number of

---

**TABLE 7.** Statistical analysis of different input image size by LDDNet.

| Size of lung CT images | Acc     | Precision | recall   | DSC     | Sensitivity | Specificity  |
|-----------------------|---------|-----------|----------|---------|-------------|--------------|
| 512x512               | 0.994378| 0.996779  | 0.996619 | 0.996699| 0.9966197  | 0.981508     |
| 256x256               | 0.991906| 0.995519  | 0.994958 | 0.995239| 0.994958   | 0.974587     |
| 128x128               | 0.986386| 0.991697  | 0.992268 | 0.991982| 0.992268   | 0.953389     |
| 64x64                 | 0.961851| 0.992565  | 0.963724 | **0.977936** | 0.963724  | 0.948317     |
| 32x32                 | 0.945139| **0.968828** | 0.966572 | 0.967699 | 0.966572   | 0.823492     |
pooling layers are different, the accuracy of LDDNet experiments will be slightly improved, for instance, four-layer average pooling, four-layer maximum pooling. Through the above TABLE 8 and TABLE 9, LDDNet can still achieve an accuracy rate of more than 0.99 in the case of different pooling layers, which also illustrates the robustness of LDDNet.

F. EFFECT OF DENSE CONNECTION
LDDNet utilized the dense-connection which can get the better result than the deep neural network with same depth. Secondly, we utilized the Dense-Block to reduce overfitting. The dense blocks alleviate the overfitting and gradient vanishing effectively, which adds the larger feature value of the bottom layers to the small feature value of the top layers. Hence, we have designed two comparative experiments, one with dense connections and one without dense connections, for the aim of exploring effect of type of dense connection on lung parenchyma segmentation by LDDNet.

It can be seen from the TABLE 10 that removing the dense connection from LDDNet will result in much worse results. The accuracy of the experiment dropped from 0.9943 to 0.9289. Dense connection can make full use of the image to train the segmentation model. And LDDNet will naturally drop a lot if there is not enough feature accuracy. Therefore, dense connection is also one of the methods used by LDDNet to reduce overfitting and improve accuracy.
Because the dense connection in LDDNet is realized by the Dense Block structure. To further explore the effect of dense connection on LDDNet segmentation of lung parenchyma, we designed comparative experiments. There are four Dense Blocks, five Dense Blocks, six Dense Blocks within LDDNet. The results of the comparative experiment can be seen in TABLE 11.

As can be seen from the TABLE 11, Dense Block is very important for LDDNet to make full use of feature. The accuracy of LDDNet using different numbers of dense blocks can reach more than 0.994.

**G. COMPARISON WITH OTHER METHODS AND ROBUSTNESS EVALUATION**

Agnes et al. [20] proposed a convolutional deep and wide network (CDWN) to segment lung region from the chest CT scan for further medical diagnosis. CDWN learns the required filters to extract hierarchical feature representations at convolutional layers. Geng et al. [21] used the first three parts of VGG-16 network structure for convolution and fused the multi-scale convolution features. Furthermore, each pixel is predicted using MLP to segment the parenchymal region. Ronneberger et al. [40] presented a network and training
strategy that relies on the strong use of data augmentation to use the available annotated samples more efficiently. Hee-won et al. [43] proposed a novel lung segmentation method to minimize the juxta-pleural nodule issue and added the final nodule candidates to the area of the Chan–Vese model (CV) results to modify the lung contour. Chen et al. [44] designed modules which employ atrous convolution in cascade or in parallel to capture multi-scale context by adopting multiple atrous rates. CRF [44] models are composed of unary potentials on individual pixels or image patches and pairwise potentials on neighboring pixels or patches. The resulting adjacency crfs structure is limited in its ability to model long-range connections within the image and generally results in excessive smoothing of object boundaries. In order to improve segmentation and labeling accuracy, the crfs framework has been expanded to incorporate hierarchical connectivity and higher-order potentials defined on image regions. DenseCRF means densely interconnected conditional random field, in which all pairs of variables are allowed to be connected by pairwise potentials from training data for learning parameters.

The results declared in TABLE 12 confirm that LDDNet achieves substantially better results than the other methods in terms of average accuracy, precision, recall, DSC,
TABLE 12. Statistical analysis (mean) of different lung segment methods.

| Network                                      | Acc    | precision | recall | DSC   | Sensitivity | Specificity |
|----------------------------------------------|--------|-----------|--------|-------|-------------|-------------|
| Fuzzy C-Means Clustering [14]                | 0.963100 | —         | —      | —     | 0.810900    | 0.968200    |
| Thresholding and Gaussian Mixture [8]        | 0.972200 | —         | —      | 0.972200 | 0.999100    | 0.967000    |
| Patch-Based Similarity [19]                  | —      | —         | —      | 0.839800 | —           | —           |
| Thresholding [20]                            | 0.770000 | 0.460000  | 0.810000 | 0.580000 | 0.810000    | 0.760000    |
| CDWN with dropout [20]                       | 0.910000 | 0.970000  | 0.760000 | 0.850000 | 0.760000    | 0.990000    |
| CDWN without dropout [20]                    | 0.980000 | 0.950000  | 0.950000 | 0.950000 | 0.950000    | 0.990000    |
| VGG-16[21]                                   | 0.986700 | —         | —      | 0.986700 | 0.975000    | 0.999100    |
| Heewon al [43]                               | 0.965000 | —         | —      | 0.957900 | 0.923500    | 0.990800    |
| Deeplab v3 [44]                              | 0.981600 | —         | —      | 0.973100 | 0.940000    | 0.983200    |
| U-Net[26]                                    | —      | —         | —      | —     | —           | —           |
| LDDNet (Proposed)                            | 0.994378 | 0.996779  | 0.996619 | 0.996699 | 0.996619    | 0.981508    |
| LDDNet (with median filter)                  | 0.994726 | 0.996882  | 0.996912 | 0.996897 | 0.996912    | 0.982357    |
| LDDNet (four-layer average pooling)          | 0.995109 | 0.996916  | 0.997336 | 0.997126 | 0.997336    | 0.982434    |

Sensitivity, Specificity. The average accuracy for the segmentation of lung regions in the lung dataset is 0.99. This is significantly better than the thresholding method result [20], whose accuracy is 0.77 without applying post-processing techniques and other methods too. The accuracy of LDDNET can reach 0.994378. The performance of LDDNet provides an improvement over thresholding by 29.12%, CDWN with dropout by 9.26%, CDWN without dropout by 1.45%, VGG-16 by 0.77%, Heewon by 3.51%, Deeplab by 1.29%. Further, we adjusted the related parameters of LDDNet. Through the adjustment of parameters such as dense connection, pooling layer, input image size, etc., we can get the accuracy rate 0.994726, 0.995109 through the setting of median filtering and four-layer average pooling. Therefore, it can be concluded that the performance of LDDNet in segmenting lung parenchyma is better than other methods. LDDNet’s segmentation accuracy is higher than other methods. Therefore, the proposed LDDNet is meaningful in the field of lung CT segmentation of lung parenchyma.

Comparative experiments of four classes of images were also conducted in this paper. The comparison of LDDNet’s segmentation effect parameters for the four classes of lung CT images segmentation in the experiment can be seen in TABLE 13.

For the LDDNet network, the training results of our images classified into four classes are compared with the existing methods. As can be declared that in class 1, class 2, class 3 and class 4, high accuracy is obtained by LDDNet. Furthermore, the average segmentation accuracy can reach 0.98 or higher in TABLE 13. Furthermore, LDDNet has the ability to segment images on different conditions of lung CT images. It can be demonstrated that Class 1 and Class 4 have higher accuracy than other methods. On account of that class 1 and class 4 are all the relatively complete images. However, Class 2 and Class 3 cannot have better accuracy than other methods. Such results are acceptable for the reason that class 2 and 3 are not very high quality. The Class 2 image is an image in which the left and right lungs are bonded, and the edges are not smooth. Therefore, LDDNet has a poor ability to learn pixels in the edge region. However, it can be seen from FIGURE 11, FIGURE 14 and FIGURE 15 that although the accuracy of segmentation is not high, the internal blood vessels and small voids are also segmented out by LDDNet. Class 3 is the image of the upper slice of the lung CT, in which the left and right lungs occupy smaller pixels. The segmentation performance of LDDNet to learn must be insufficient, so it will cause the accuracy of segmentation to decrease. However, from FIGURE 12 the LDDNet network still segmented out the small voids and blood vessels. So, no matter when testing LDDNet for image classification or training on the dataset as a whole, LDDNet can achieve a good accuracy. In summary, LDDNet has good robustness while ensuring good segmentation performance.

VI. CONCLUSION

Lung segmentation algorithms play an important role in lung cancer detection system and directly affect the accuracy of lung cancer verification and recognition. Lung dense deep
convolutional neural network (LDDNet) for lung segmentation is introduced in this paper. The grayscale and texture features of the CT image of the lung are utilized by LDDNet. The high accuracy and robustness on lung CT images can be achieved by LDDNet. Through using image preprocessing and no image preprocessing as contrast proves that some special image preprocessing methods like median filter, Laplacian filter and so on, will promote the CT lung parenchyma segmentation results of LDDNet. In contrast, some will reduce the accuracy of segmentation for instance enhanced contrast etc. Besides ground-truth for LIDC-IDRI is marked manually, blood vessels and small voids are also labeled by us for the aim of analyzing the effect of these interference factors on lung segmentation. Finally, the better segmentation result is achieved in term of accuracy and robustness. With LDDNet, the final result is improved compared to most of the traditional lung segmentation methods and deep convolutional networks for instances VGG, U-Net, CDWN and so on. Therefore, the method of segmentation of lung CT images will continue to be studied for clinically specific applications. Too many blood vessels and small voids in the lung CT images can seriously interfere with the accuracy of segmentation. In the future, designing a more robust algorithm for automatic segmentation of lung parenchyma and a more accurate method for labeling lung parenchyma is one of the future research directions. Besides more attention should be attached to the segmentation of blood vessels and small voids. Our goal is to improve the segmentation accuracy of lung CT images, furthermore to help doctors’ clinical diagnosis and promote medical progress better.

ACKNOWLEDGMENT

The authors would like to thank the National Cancer Institute (NCI) for providing the LIDC-IDRI database by the Cancer Image Archive (TCIA) Public Access.

REFERENCES

[1] Z. W. Qi, Detection and Diagnosis of Lung Tumor Based on CT Image, Harbin, China: Harbin Institute of Technology, 2014.
[2] C. R. Wu, J. Zhu, and M. Q. Ye, “Review of computer aided detection and diagnosis techniques for CT image of pulmonary nodules,” in Proc. Data Acquisition Process., May 2016, pp. 868–881.
[3] W. L. Zhang, X. Q. Lu, and Y. Gu, “Research progress of lung parenchyma segmentation based on CT images of the lungs,” Chin. J. Med. Phys., vol. 39, no. 9, pp. 902–907, 2017.
[4] T. Yuan, H. Y. Cheng, and Y. H. Chen, “CT image segmentation algorithm based on U-Net network,” J. Autom. Instrum., vol. 6, pp. 59–61, Oct. 2017.
[5] A. M. Anter and A. E. Hassenian, “CT liver tumor segmentation hybrid approach using neurosopic sets, fast fuzzy c-means and adaptive watershould algorithm,” Artif. Intell. Med., vol. 97, pp. 105–117, Jun. 2019.
[6] P. Chondro, C.-Y. Yao, S.-J. Ruan, and L.-C. Chien, “Low order adaptive region growing for lung segmentation on plain chest radiographs,” Neuro-computing, vol. 275, pp. 1002–1011, Jan. 2018.
[7] F. Shariaty, S. Hosseiniou, and V. Y. Rud, “Automatic lung segmentation method in computed tomography scans,” J. Phys., Conf. Ser., vol. 1236, Jun. 2019, Art. no. 012028.
[8] S. Gopalakrishnan and A. Kandaswamy, “Automatic delineation of lung parenchyma based on multilevel thresholding and Gaussian mixture modelling,” Comput. Model. Eng. Sci., vol. 114, no. 2, pp. 141–152, 2018.
[9] S. Zhang, Y. Zhao, and P. Bai, “Object localization improved GrabCut for lung parenchyma segmentation,” Procedia Comput. Sci., vol. 131, pp. 1311–1317, Oct. 2018.
[10] S. Y. Tang, M. Yang, Y. Miao, and J. N. Bai, “CT image segmentation of lungs with regional growth and horizontal integration,” Electron. Technol. Appl., vol. 44, no. 5, pp. 133–139, 2018.
[11] R. Hao, Y. Qiang, and X. Yan, “Juxta-vascular pulmonary nodule segmentation in PET-CT imaging based on an LBF active contour model with information entropy and joint vector,” Comput. Math. Methods Med., vol. 2018, Jan. 2018, Art. no. 2183847.
[12] Y. Wei, X. H. Xu, T. Jia, and D. Z. Zhao, “CT image lung nodule detection algorithm based on optimized level set method,” J. Syst. Simul., vol. 18, no. S2, pp. 909–911, 2006.
[13] Y. Qu, “Study on lung parenchymal segmentation algorithm of lung CT image,” M.S. thesis, Shandong Univ. Traditional Chin. Med., Jinan, China, 2014.
[14] Z. F. Khan, “Automated segmentation of lung parenchyma using colour based fuzzy C-Means clustering,” J. Electr. Eng. Technol., vol. 14, no. 5, pp. 2163–2169, Sep. 2019.
[15] W. L. Zhang, X. Q. Lu, and Y. Gu, “Pathological lung segmentation for images based on Freeman chain code,” Comput. Eng. Des., vol. 39, no. 10, pp. 195–198, 2018.
[16] V. Dharmalingham and D. Kumar, “A model based segmentation approach for lung segmentation from chest computer tomography images,” Multimedia Tools Appl., vol. 79, nos. 15–16, pp. 10003–10028, Apr. 2020.
[17] E. E. Nithila and S. S. Kumar, “Segmentation of lung from CT using various active contour models,” Biomed. Signal Process. Control, vol. 47, pp. 57–62, Jan. 2019.
[18] T. Peng, Y. Wang, T. C. Xu, and X. Chen, “Segmentation of lung in chest radiographs using hull and closed polygonal line method,” IEEE Access, vol. 7, pp. 137794–137810, 2019.
[19] A. Majda and E. H. Abdelhamid, “Graph cuts segmentation approach using a patch-based similarity measure applied for interactive CT lung image segmentation,” Int. J. Comput. Inf. Eng., vol. 12, no. 7, pp. 520–524, 2018.
[20] S. Akila Agnes, J. Anitha, and J. Dinesh Peter, “Automatic lung segmentation in low-dose chest CT scans using convolutional deep and wide network (CDWN),” IEEE Access, vol. 11, pp. 1–11, Nov. 2018.
[21] L. Geng, S. Zhang, J. Tong, and Z. Xiao, “Lung segmentation method with dilated convolution based on VGG-16 network,” Comput. Assist. Surgery, vol. 24, no. sup2, pp. 27–33, Oct. 2019.
[22] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 640–651.
[23] P. A. Harrison, Z. Xu, K. George, L. Le, M. S. Ronald, and J. M. Daniel, “Progressive and multi-path holistically nested neural networks for pathological lung segmentation from CT images,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2017, pp. 621–629.
[24] D. Jin, Z. Xu, Y. Tang, A. P. Harrison, and D. J. Mollura, “CT-realistic lung nodule simulation from 3D conditional generative adversarial networks for robust lung segmentation,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2018, pp. 732–740.
[25] Y. Giodenko, P. Gang, J. Hui, Z. Wei, Y. Kochura, O. Alienin, O. Rokovyi, and S. Stenroen, “Deep learning with lung segmentation and bone shadow exclusion techniques for chest X-ray analysis of lung cancer,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2018, Art. no. 2183847.
[26] S. Mougiakakou, “Semantic segmentation of pathological lung tissue with information entropy and joint vector,” in Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham, Switzerland: Springer, 2017, pp. 92–99.
[27] S. Hwang and S. Park, “Accurate lung segmentation network via network-wise training of convolutional networks,” in Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham, Switzerland: Springer, 2017, pp. 133–140.
[28] A. A. Novikov, D. Lenis, D. Major, J. Hladuvka, M. Wimmer, and K. Buhler, “Fully convolutional architectures for multiclass segmentation in chest radiographs,” IEEE Trans. Med. Imag., vol. 37, no. 8, pp. 1865–1876, Aug. 2018.
[29] M. Anthimopoulos, S. Christodoulidis, L. Ebner, T. Geiser, A. Christe, and S. Mougiakakou, “Semantic segmentation of pathological lung tissue with dilated fully convolutional networks,” IEEE J. Biomed. Health Inform., vol. 23, no. 2, pp. 714–722, Mar. 2019.
[30] W. Dai, N. Dong, and Z. Wang, “SCAN: Structure correcting adversarial network for organ segmentation in chest X-rays,” in Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham, Switzerland: Springer, 2018, pp. 263–273.

[31] C. Liu, L.-C. Chen, F. Schroff, H. Adam, W. Hua, A. L. Yuille, and L. Fei-Fei, “Auto-DeepLab: Hierarchical neural architecture search for semantic image segmentation,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 82–92.

[32] F. Isensee, J. Petersen, S. A. Kohl, P. F. Jager, and K. H. Maier-Hein, “Automated design of deep learning methods for biomedical image segmentation,” 2019, arXiv:1904.08128. [Online]. Available: https://arxiv.org/abs/1904.08128

[33] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. Comput. Vis. Pattern Recognit., 2016, pp. 770–778.

[34] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014, arXiv:1409.1556. [Online]. Available: http://arxiv.org/abs/1409.1556

[35] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 2818–2826.

[36] R. Girshick, “Fast R-CNN,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 1440–1448.

[37] J. Dai, Y. Li, K. He, and J. Sun, “R-FCN: Object detection via region-based fully convolutional networks,” (2016), arXiv:1605.06409. [Online]. Available: https://arxiv.org/abs/1605.06409

[38] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, “SSD: Single shot multibox detector,” in Proc. Eur. Conf. Comput. Vis., 2016, pp. 21–37.

[39] C. Wang, Z. Zhao, Q. Ren, Y. Xu, and Y. Yu, “Dense U-net based on patch-based learning for retinal vessel segmentation,” Entropy, vol. 21, no. 2, p. 168, 2019.

[40] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, Oct. 2015, pp. 234–241.

[41] LIDC-IDRI. National Cancer Institute—The Leading Agency for Cancer Research and Funding in the United States, Lung Image Database Consortium’s Image Collection (LIDC-IDRI). Accessed: Jan. 18, 2019. [Online]. Available: https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI

[42] Tensorflow. Accessed: Jan. 18, 2019. [Online]. Available: https://www.tensorflow.org/

[43] H. Chung, H. Ko, S. J. Jeon, K.-H. Yoon, and J. Lee, “Automatic lung segmentation with juxta-pleural nodule identification using active contour model and Bayesian approach,” IEEE J. Transl. Eng. Health Med., vol. 6, 2018, Art. no. 1800513.

[44] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, “Rethinking atrous convolution for semantic image segmentation,” 2017, arXiv:1706.05587. [Online]. Available: http://arxiv.org/abs/1706.05587

YING CHEN was born in 1981. He received the B.S., M.S., and Ph.D. degrees from Jilin University, Changchun, China. He is currently an Associate Professor at Nanchang Hangkong University, Nanchang, China. His main research interests include pattern recognition, machine learning, biometrics, image processing, and deep learning.

YERONG WANG was born in 1993. He received the B.S. degree in software engineering from the Changchun Institute of Technology, Changchun, China. He is currently pursuing the M.S. degree with Nanchang Hangkong University, under the supervision of Prof. Y. Chen. His main research interests include software engineering, image processing, and machine learning.

FEI HU was born in 1997. She received the B.S. degree in computer science and technology from Anhui Normal University, Wuhu, China. She is currently pursuing the M.S. degree with Nanchang Hangkong University, under the supervision of Prof. Y. Chen. Her main research interests include machine learning and image processing.

DING WANG was born in 1992. He received the B.S. degree in applied electronic technology from the College of Business and Technology, Hubei University of Technology, Changchun, China. He is currently pursuing the M.S. degree with Nanchang Hangkong University, under the supervision of Prof. Y. Chen. His main research interests include software engineering, image processing, and cyber security.