IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 28, 2021, accepted March 17, 2021, date of publication March 24, 2021, date of current version April 7, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3068896

LDNNET: Towards Robust Classification of Lung Nodule and Cancer Using Lung Dense Neural Network

YING CHEN, YERONG WANG, FEI HU, LONGFENG FENG, TAOHUI ZHOU, AND CHENG ZHENG

School of Software, Nanchang Hangkong University, Nanchang 330063, China

Corresponding author: Yerong Wang (wangyerong520@qq.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61762067 and Grant 61867004, and in part by the Natural Science Foundation of Jiangxi Province under Grant 20202BABC202029 and Grant 20202BABC202028.

ABSTRACT Lung nodule classification plays an important role in diagnosis of lung cancer which is essential to patients' survival. However, because the number of lung CT images in current dataset is relatively small and the ratio of nodule samples to non-nodule samples is usually very different, this makes the training of neural networks difficult and poor performance of neural networks. Hence, LDNNET is proposed, which adopts Dense-Block, batch normalization (BN) and dropout to cope with these problems. Meanwhile, LDNNET is an adaptive architecture based on convnets combining softmax classifier which is utilized to alleviate the problems of training deep convnets. Follows are our main work: Firstly, we utilized LDNNET on database LUng Nodule Analysis 2016 (LUNA16) for lung nodule classification and database KAGGLE DATA-SCIENCE-BOWL-2017(Kaggle DSB 2017) for lung cancer classification; Secondly, the comparison experiments are designed to compare the performance of dense connection, pooling layer and the input pixel size of lung CT(Computed Tomography) images; Thirdly, data enhancement, dense connection and dropout layer were utilized in LDNNET to reduce overfitting; Fourthly, pre-processing methods, for instance enhanced contrast, median filtering, Laplacian filtering are compared to the no-processing method to explore the effect of pre-processing on lung CT images classification. Fifthly, accuracy, specificity and sensitivity on LUNA16 are 0.988396, 0.994585 and 0.982072 and these indicators on Kaggle DSB 2017 are 0.999480, 0.999652 and 0.998974. Furthermore, AUC for both two datasets is over 0.98. Consequently, this paper conducts experiments with uniform parameter settings on two publicly available databases and shows that even in challenging situation where lung images are directly utilized as input images without preprocessing, LDNNET is still the more advanced algorithm than other recent algorithms respectively. Moreover, a series of comparative experiments were conducted to further confirm that the proposed algorithm has the higher accuracy and robustness through verification and discussion.

INDEX TERMS Deep dense neural network, classification, lung nodule, lung cancer.

I. INTRODUCTION

Nowadays lung cancer has become a huge death-threat which is threatening human health worldwide. Lung cancer is one of the most malignant tumors in the world and its 5-year-survival rate is only 18% [1] Early and accurate classification of lung nodule and lung cancer is of great significance to the treatment of patients and the improvement of survival. Because, early lung cancer often appears as lung nodules with or without malignant signs on CT, therefore the classification of lung nodules is the first step in the early diagnosis of lung cancer. The second step is to classify lung and non-lung cancer on CT images of lung nodules to assist doctors in medical diagnosis. Various diagnostic procedures are utilized by physicians for early diagnosis of malignant lung nodules, for instance clinical settings, computed tomography (CT) scan analysis (morphological assessment), positron emission tomography (PET) (metabolic assessments) and needle prick biopsy analysis [2]. However, it is difficult to widely realize the detection method of needle prick biopsy analysis. Because, it is difficult for patients to accept needle prick biopsy analysis. Further, the actual use cannot be popularized

The associate editor coordinating the review of this manuscript and approving it for publication was Haruna Chiroma.

in the medical field. Hence, CT images are still the most effective means to detect lung cancer. Hence, it is essential to classify the lung nodule in CT images. Early classification of lung nodules of lung CT images can save the lives of lung cancer patients. Further, a fast and accurate Computer-Aid Diagnosis (CAD) system for lung nodule classification is urgently desired. The classification of lung nodules is important for the diagnosis of lung cancer based on CT images for instance nodule, non-nodule and cancer, non-cancer. Early detection is critical to give patients the best chance of survival and recovery [3]. In the medical process, there are often a large number of CT images for the doctors to diagnose. However, the diagnosis of lung CT images in the past only relied on the experience of the physician. In reality, looking over and judging such numerous images will inevitably increase doctors' workload and even cause error medical accidents [4]. Machine learning methods can be used to classify lung nodules. Gupta *et al.* [5] extracted relevant features from the lung images and used a machine learning classifier to realize feature selection and identification of lung diseases. However, the accuracy of the machine learning method is relatively low. In the paper [5], the classification accuracy of the Improvised Cuttlefish Algorithm (ICFA) method is 97.3%. In our previously published paper [6], the LDDNET network was proposed for robust lung parenchymal segmentation. Paper [6]and this paper are both based on lung CT images, and both utilize convolutional neural network algorithms to realize the intelligentization of medical images. The dataset used in paper [6] is LIDC-IDRI for lung CT image segmentation. Further, the datasets used in this paper are LUNA16 and KAGGLE DSB 2017 for lung CT image classification, which are the subset of the LIDC-IDRI dataset. In this paper, the number of images utilized by the LDNNET network on datasets LUNA16 and KAGGLE DSB 2017 are 50000 and 90000, meanwhile the number of images used by LDDNET in paper [6] on LIDC-IDRI dataset is 4000. The main point of the difference between LDNNET and LDDNET is that the LDDNET network has an up-sampling structure (that is the decoder), and LDDNET finally outputs a binary image of lung parenchymal segmentation. However, this structure is not utilized in LDNNET network in this paper. At last, LDNNET finally outputs the classification results of lung CT images directly. Further, at present, most methods utilized for classifying lung nodules and lung cancer are to use deep neural networks to perform experiments for higher accuracy. However, due to the lung CT data collection, the ratio of nodule and non-nodule samples is often very different, which makes the training of the neural network difficult. In this paper, we have adopted data enhancement, dropout, batch normalization (BN) and dense connections to alleviate the problem of sample imbalance.

The contributions in this paper are summarized as follows:

1) We proposed LDNNET network for lung CT images classification which fully takes the end-to-end advantage of deep learning and directly classify lung CT images without lung parenchymal segmentation and lung nodule segmentation that the current lung nodule classification algorithms require;

2) We conducted the comparison experiments to compare the performance of dense connection, pooling layer and the input pixel size of lung CT (Computed Tomography) images for lung nodule classification and lung cancer classification. The image sizes we set gradually changes to maximize the performance comparison of LDNNET, for instance, 80×80, 64×64, 48×48, 32×32 and 16×16;

3) Data enhancement, dense connection and dropout layer were utilized in LDNNET network to reduce overfitting. At the same time, dense connections increase the robustness of lung features extracted by LDNNET network through feature reuse so that the network has a higher accuracy and robustness than previous algorithms;

4) Pre-processing methods, for instance enhanced contrast, median filtering, Laplacian filtering are compared with the no-processing method to explore the effect of pre-processing on lung CT images for lung nodule classification and lung cancer classification. The experimental results show that even if LDNNET does not perform any preprocessing methods, for instance, enhanced contrast, median filtering, Laplacian filtering, and without parameter adjustment for the databases LIDC-IDRI and LUNA16, the LDNNET is always better than the best-performing algorithm;

(5) This paper conducts experiments on multiple publicly available databases. We conducted experiments on LUNA16 dataset and LDNNET achieved a good classification effect of lung nodules. Meanwhile, on Kaggle DSB 2017 dataset, LDNNET still achieved a good lung classification effect, which is proved that the LDNNET network has good classification performance not only for lung nodules but also for lung cancer. Additionally, a series of experiments are carried out to analyze the advantages and limitations of the LDNNET.

The remainder of this paper is organized as follows: Section 2 outlines deep neural network algorithms based on geometric features and lung feature description algorithms based on transform domain. Section 3 describes the proposed architecture for lung nodule classification and lung cancer classification in detail. The details of two public lung image databases are represented, in which the LUNA16 dataset is for lung nodule classification and the Kaggle DSB 2017 dataset is for lung cancer classification in Section 4. Moreover, environmental configuration of the experiment, network training parameters metrics and measurements for experiments are also described in Section 4. Experimental results evaluation and comparison are represented in Section 5. Section 6 summarizes this paper.

II. RELATED WORKS

A. OVERVIEW OF LUNG NODULE CLASSIFICATION AND LUNG CANCER CLASSIFICATION METHODS

Traditional methods are not using machine learning and do not have higher requirements for equipment performance.

Ahmed *et al.* [7] 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 and clustering-based. Hence the traditional methods are less time consuming in executing. Traditional algorithms for lung nodule classification and lung cancer classification 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) TRADITIONAL LUNG NODULE CLASSIFICATION AND LUNG CANCER CLASSIFICATION ALGORITHMS BASED ON TRANSFORM DOMAIN

These algorithms project normalized lung CT images to a frequency domain and use geometric transformation to extract image features of lung nodules. Wei *et al.* [8] proposed a new two-step CBIR scheme (TSCBIR) for computer-aided diagnosis of lung nodules. Two similarity metrics, semantic relevance and visual similarity were introduced to measure the similarity of different nodules. The first step is to search for K most similar reference ROIs for each queried ROI with the semantic relevance metric. The second step is to weight each retrieved ROI based on its visual similarity to the queried ROI. Gong *et al.* [9] applied a series of preprocessing steps to segment the lung volume and generate the isotropic volumetric CT data. Next, a unique 3D tensor filtering approach and local image feature analysis were used to detect nodule candidates. A 3D level set segmentation method was utilized to correct and refine the boundaries of nodule candidates subsequently. Then, the features of the detected candidates were extracted and the optimal features were selected by using a CFS (Correlation Feature Selection) subset evaluator attribute selection method. Finally, a random forest classifier is trained to classify the detected candidates. Froze *et al.* [10] proposed a methodology to classify lung nodule candidates and non-nodule candidates based on computed tomography (CT) images. The Lung Image Database Consortium (LIDC-IDRI) database was employed for testing. Three techniques are employed to extract texture measurements. The first technique was artificial crawlers (ACs), an artificial life algorithm. The second technique used the rose diagram (RD) to extract directional measurements. The third technique was a hybrid model that combines texture measurements from artificial crawlers and the rose diagram. Chabon *et al.* [11] developed and prospectively validated a machine-learning method termed 'lung cancer likelihood in plasma' (Lung-CLiP), which can robustly discriminate early-stage lung cancer patients from risk-matched controls. This approach achieves performance similar to that of tumour-informed ctDNA detection and enables tuning of assay specificity in order to facilitate distinct clinical applications. Jacobs *et al.* [12] described the different components of the CAD system and presents experiments to optimize the performance of the proposed CAD system. A rich set of 128 features is defined for subsolid nodule candidates. In addition to previously used intensity, shape and texture features, a novel set of contextual features is introduced. Their experiments showed that these features significantly improve the classification performance. Ye *et al.* [13] proposed a new computed tomography (CT) lung nodule computer-aided detection (CAD) method for detecting both solid nodules and ground-glass opacity (GGO) nodules (part solid and nonsolid). Demir and Camurcu [14] developed a computer-aided detection (CAD) system for the detection of lung nodules in computed tomography images. The CAD system consists of four phases, including two-dimensional and three-dimensional preprocessing phases. In the feature extraction phase, four different groups of features were extracted from volume of interests: morphological features, statistical and histogram features, statistical and histogram features of outer surface and texture features of outer surface.

However, traditional methods are less accurate for lung nodule classification and lung cancer classification. These methods are far from meeting the accuracy requirements of clinic physical examination and hospital diagnosis. Further, the traditional methods are summarized by the mature lung CT medical experience in which the selection and setting of lung CT features of lung nodule classification and lung cancer classification are all set manually. Furthermore, in the current context, traditional algorithms based on the lung CT medical experience of algorithm researchers are increasingly unable to adapt to the high-precision data-driven algorithm requirements under the condition of massive and massive medical data. The lung CT medical experience required by traditional algorithms often takes several years or even ten years to form. Therefore, for the aim of adapt to the modern medical system, algorithms based on deep neural networks are widely utilized. This algorithm has a higher accuracy rate in lung nodule classification and lung cancer classification and the generation of the algorithm model always takes only a few hours.

2) LUNG NODULE CLASSIFICATION AND LUNG CANCER CLASSIFICATION ALGORITHMS BASED ON DEEP NEURAL NETWORK

Traditional lung nodule classification and lung cancer classification algorithms require manually extracted features. Meanwhile, features of lung CT images can be automatically extracted by the algorithms based on deep neural network, which can be more accurate and effective than traditional algorithms.

Zuo *et al.* [4] put forward a multi-resolution convolutional neural network (CNN) to extract features of various levels and resolutions from different depth layers in the network for classification of lung nodule candidates. Inspired by the AlphaGo system, Ali *et al.* [3] proposed a kind of deep reinforcement learning network. Their deep learning algorithm takes a raw CT image as input, views it as a collection of states and outputs a classification of whether a nodule is present or not. The dataset used to train the model is the LIDC/IDRI database, which is hosted by the LUng Nodule Analysis (LUNA)

challenge. The training results yielded an overall accuracy of 0.991. Mobiny *et al.* [15] proposed a memory-augmented capsule network for the rapid adaptation of CAD models to new domains. It consists of a capsule network that was meant to extract feature embeddings from some high-dimensional inputs, and a memory-augmented task network meant to exploit its stored knowledge from the target domains. When trained on the LUNA dataset, the memory-augmented capsule network requires only 30 additional samples from their collected lung nodule and incidental lung nodule datasets to achieve clinically relevant performance. In paper [16], Heuvelmans retrospectively validated their Lung Cancer Prediction Convolutional Neural Network (LCP-CNN), which was trained on US screening data, on an independent dataset of indeterminate nodules in an European multicentre trial, to rule out benign nodules maintaining a high lung cancer sensitivity. William *et al.* [17] proposed a novel lung nodule detection and classification model using one stage detector called as ''I3DR-Net''. The model was formed by combining pre-trained natural images weight of Inflated 3D ConvNet (I3D) backbone with feature pyramid network to multi-scale 3D Thorax Computed tomography scan (CT-scan) dataset. Nasrullah *et al.* [2] proposed a novel deep learning-based model with multiple strategies for the precise diagnosis of the malignant nodules. Due to the recent achievements of deep convolutional neural networks (CNN) in image analysis, two deep three-dimensional (3D) customized mixed link network (CMixNet) architectures were utilized for lung nodule detection and classification, respectively. Nodule detections were performed through faster R-CNN on efficiently-learned features from CMixNet and U-Net like encoder–decoder architectures. Classification of the nodules was performed through a gradient boosting machine (GBM) on the learned features from the designed 3D CMixNet structure. Xie *et al.* [18] proposed a multi-view knowledge-based collaborative (MV-KBC) deep model to separate malignant from benign nodules using limited chest CT data. The MV-KBC model learns 3D lung nodule characteristics by decomposing a 3D nodule into nine fixed views. Ali *et al.* [19] proposed transferable texture Convolutional Neural Networks (CNN) to improve the classification performance of pulmonary nodules in CT scans. An Energy Layer (EL) was incorporated in the proposed scheme, which extracts texture features from the convolutional layer. The inclusion of EL reduces the number of learnable parameters of the network, which further reduces the memory requirements and computational complexity. Naik and Edla [20] surveyed around 108 research papers to focus on the contribution of deep learning methodologies in detection of malignant tumor in Lung CT scans. Naik and Edla [20] also presented challenges and opportunities in classifying lung nodule by using advanced deep learning strategies. Lin and Li [21] proposed a Taguchi-based convolutional neural network (CNN) for classifying nodules into malignant or benign. For setting parameters in a CNN, most users adopt trial and error to determine structural parameters. Dey *et al.* [22] proposed four two-pathway Convolutional Neural Networks (CNN), including a basic 3D CNN, a novel multi-output network, a 3D DenseNet, and an augmented 3D DenseNet with multi-outputs. These four networks were evaluated on the public LIDC-IDRI dataset and outperformed most existing methods. However, Dey *et al.* [22] work directly on 3D volumes yields instead of using 2D slices or approximating 3D im-age with multi-views better results for the lung nodule classification problem. These 3D images are better than 2D images in the classification of lung nodules. However, the server performance requirements are higher here, which is not conducive to the promotion of the algorithm. Hence, our proposed LDNNET network is still experimenting on 2D images. Al-Shabi *et al.* [23] proposed a novel CNN architecture called Gated-Dilated (GD) network to classify nodules as malignant or benign. Unlike previous studies, the GD network uses multiple dilated convolutions instead of max-poolings to capture the scale variations. Moreover, the GD network has a Context-Aware sub-network that analyzes the input features and guides the features to a suitable dilated convolution. Al-Shabi *et al.* [23], lung CT images were divided into two categories in the experiment. Meanwhile, our network is used for this classification of nodule and non-nodule. Furthermore, the number of lung CT images for LDNNET network is 100000 and 90000 for LUNA16 and Kaggle DSB 2017 dataset, which is greater than the number of 1000 lung CT images from LIDC-IDRI used in paper [23]. Nobrega *et al.* [24] proposed approach which aims to explore the performance of deep transfer learning for lung nodules malignancy classification. The deep features returned were classified using Naive Bayes, MultiLayer Perceptron (MLP), Support Vector Machine (SVM), Near Neighbors (KNN) and Random Forest (RF) classifiers. Qin *et al.* [25] proposed a computer-aided diagnosis (CAD) system for simultaneous accurate pulmonary nodule detection and false positive reduction. They build a full 3D CNN model that employs 3D U-Net architecture and a 3D DenseNet-based model is presented to reduce false positive. Ardila *et al.* [26] proposed a deep learning algorithm that uses a patient's current and prior computed tomography volumes to predict the risk of lung cancer. The model achieves a state-of-the-art performance (94.4% area under the curve). The main work in this paper is similar to the paper [27], Detection and classification of pulmonary nodules using convolutional neural networks: A survey. Meanwhile, the network LDNNET proposed in this paper actually have the process of lung nodule detection. Because the title and subject are both classification of lung CT images, this part of lung nodule detection does not mention too much. Furthermore, classification experiment is also used in the Kaggle DSB 2017 dataset to classify lung cancer and non-lung cancer. In paper [27], the author introduced the fundamental knowledge of CNN as well as the reasons for their suitability to medical images analysis and proved by experimental results that CNNs have transformed greatly the early diagnosis and management of lung cancer. However, in paper [27], the author used experiments to illustrate that

TABLE 1. Comparisons of previous studies on lung nodule classification.

Methods	Strengths	Weaknesses
Traditional lung nodule classification and lung cancer classification algorithms based on transform domain	These methods are fast and do not require a large number of samples for training.	On th basis of the improved accuracy of lung nodule classification, these methods are not robust enough for the reason the input lung image are separate part of lung CT images rather than the whole original lung images. The lung CT medical experience required by traditional algorithms often takes several years or even ten years to form.
Lung nodule classification and lung cancer classification algorithms based on deep neural network	This type of algorithm has a higher accuracy rate than the algorithm based on transform domain.	Training the lung nodule classification model takes a a considerable amount time, meanwhile the traditional algorithms do not need the process of training.
Our proposed method (LDNNET network)	The LDNNET network has the higher accuracy and robustness for lung nodule classification and lung cancer classification.	In LDNNET network, dense connections are utilized to fully learn the pixel information of the lung CT images. However, it also takes a longer time to train the network model than other network. And for there are many connections in LDNNET network, this also causes the batch size setting value during network training to be lower to match the capacity of the network training equipment.

CNN is very practical in the direction of medical image analysis, meanwhile no new neural network model is proposed for the classification of lung nodules. In this paper, LDNNET network is proposed to cope with the classification of lung nodules and good experimental results are achieved. Zhao *et al.* [28] proposed three schemes of designing CNN models and resulting eleven deep CNN models. In paper [28], classification criteria were benign lung nodules and malignant lung nodules. The dataset used is LIDC-IDRI. Meanwhile in this paper, classification criteria are lung nodules and non-nodules and the dataset is LUNA16.

Comparative analysises between our proposed method and other lung nodule classification and lung cancer classification methods are summarized in TABLE 1.

In recent years, great success has been achieved by Convolutional Neural Network (CNN) in image classification [29]–[31], target recognition [32]–[34], 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.

III. THE PROPOSED ARCHITECTURE FOR LUNG NODULE CLASSIFICATION

A. FLOWCHART OF THE LUNG NODULE CLASSIFICATION SYSTEM

The method proposed in this paper is an end-to-end structure which fully takes the end-to-end advantages of deep learning and directly classify lung CT images without lung parenchymal segmentation and lung nodule segmentation that the current lung nodule classification algorithms require. The proposed structure, LDNNET, adopts many structures of modern neural network, for instance, Dense-Block, batch normalization (BN) and dropout. Meanwhile, LDNNET is an adaptive architecture based on convnets combining softmax classifier which are utilized to alleviate the problems of training deep convnets. We input the original lung CT images into the network and finally get the classification results of lung CT images. The structure is shown in the FIGURE 1 below.

B. LDNNET FOR LUNG NODULE AND LUNG CANCER CLASSIFICATION

We proposed a LDNNET network. LDNNET 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. Network structure of LDNNET is shown as follow FIGURE 2.

Huang *et al.* [35] proposed that as CNNs become increasingly deep, a new research problem emerges: as information about the input or gradient passes through many layers, it can vanish and ''wash out'' by the time it reaches the end (or beginning) of the network. To solve this problem, we applied the Dense-Block structure in our LDNNET network. Dense-Block has good effects for image classification and has been adopted by the latest papers. Fielding and Zhang [36] proposed a Swarm Optimised DenseBlock Architecture Ensemble (SODBAE) method, a joint optimisation and training process that explores a constrained search space over a skeleton Dense Block Convolutional Neural Network (CNN) architecture. Being evaluated using the CIFAR-10 dataset, the proposed model shows great superiority in classification performance over other state-of-the-art methods. Dense-Block can also extract image features

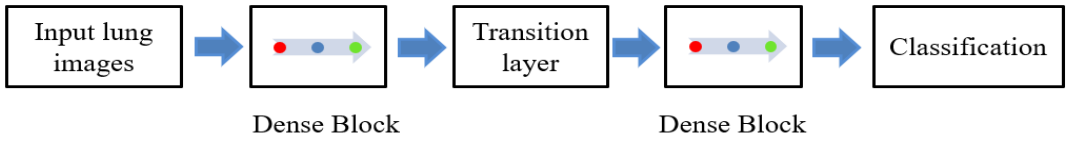


FIGURE 1. Flowchart of the lung nodule classification system.

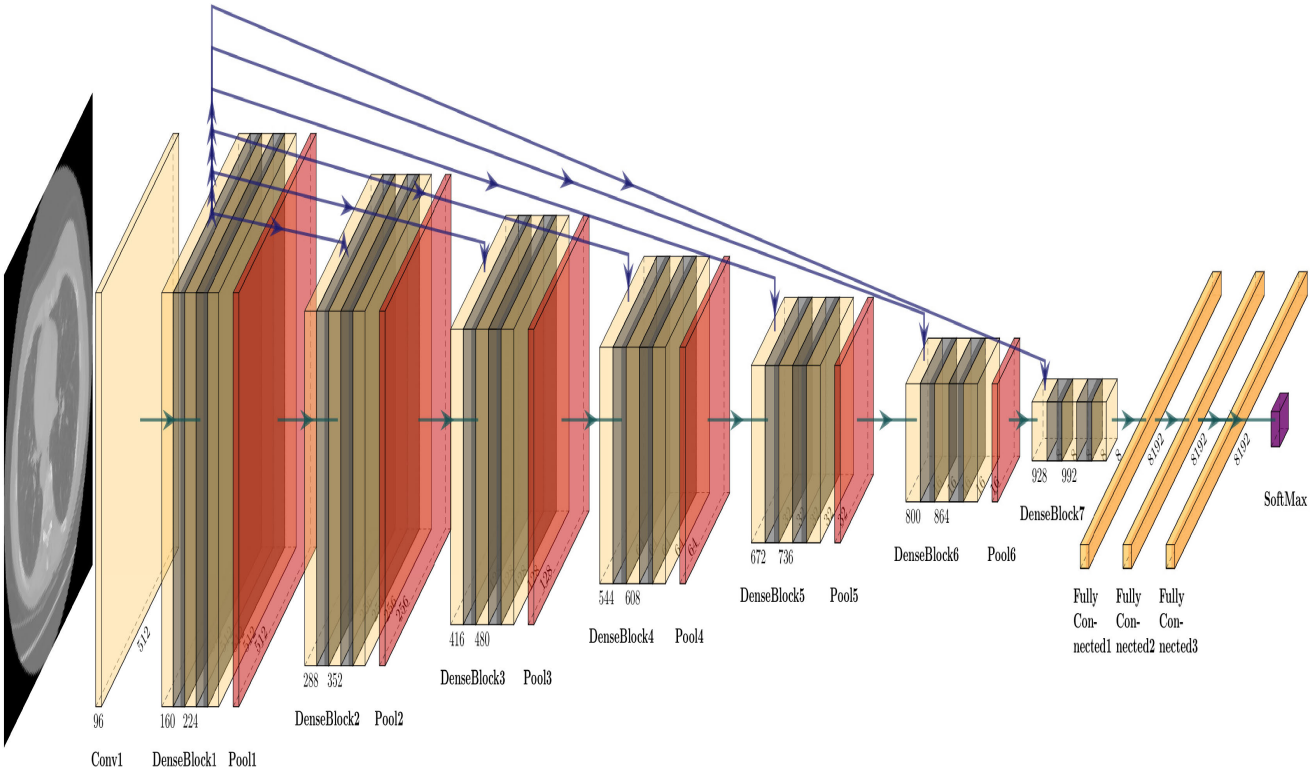


FIGURE 2. Network structure of LDNNET.

better recently. For instance, in the paper [37], zhao *et al.* proposed a novel medical image fusion algorithm based on deep convolutional generative adversarial network and dense block models, which is used to generate fusion images with rich information. Chen *et al.* [38] proposed a new dense block which uses complex connections between each layer to build a more powerful generator. Next, to improve the perceptual quality, they found a new set of feature maps to compute the perceptual loss, which would make the output image look more real and natural. Dense-Block, Batch Normalization (BN), Relu, and Convolution which size is 3×3 and a layer of Concatenation were included in the structure of Dense-Block. For Batch Normalization, in paper [39], Sergey *et al.* proved that Batch Normalization: accelerating deep network training by reducing internal covariate shift which is essential for LDNNET network training. For ReLu, in the paper [40], Xavier *et al.* proposed that while logistic sigmoid neurons are more biologically plausible than hyperbolic tangent neurons, the latter work better for training multi-layer neural networks. And rectifying neurons are an even better model of biological neurons and yield equal or better performance than hyperbolic tangent networks in spite of the hard non-linearity and non-differentiability at zero, creating sparse representations with true zeros, which seem remarkably suitable for naturally sparse data. There is a direct connection between every two layers in the Dense-Block, which makes the most of features of lung CT images for the aim of ensuring the largest information flow in the network. Assuming there are L convolutional layers, the traditional convolution network has L connections. However, there are Lx(L+1) connections in the LDNNET network. Through this structure, the number of parameters in the network can be reduced and the image information can be fully utilized to train the LDNNET network. Besides the disappearance of the gradient during the training process will seriously restrict the improvement of the accuracy of the deep neural network. The structure of Dense-Block can alleviate this problem for some extent for the dense connections.

The LDNNET network is composed of 7 Dense-Blocks. Huang *et al.* [35] introduced the Dense Convolutional Network (DenseNet), which connects each layer to every other layer in a feed-forward fashion. Whereas traditional convolutional networks with L layers have L connections—one between each layer and its subsequent layer their network has L(L+1)/2 direct connections. In paper [35], the structure

of Dense-Block consist of the proposed network DenseNet. In this paper, each Dense-Block consists of two BN-RELU-Convolution structures. The LDNNET network begins with a convolution with 5×5 kernels, stride 1, padding 0, number 96. In the convolution of Dense-Blocks, the size is kernel 3×3, stride 1 and padding 0. AlexNet [41] used the convolution with $11 \times 11 \times 3$ kernels, stride 4. The larger the convolution kernel, the larger the receptive field, the more image information can be seen, the better the features can be got by the deep neural network. However, a large convolution kernel will cause a sudden increase in the amount of calculation, which is not conducive to the increase of the depth of the model and the calculation performance will also be reduced. VGG [30] used evaluation of networks to increase depth using an architecture with very small (3×3) convolution filters, which shows that a significant improvement on the prior-art configurations can be achieved by pushing the depth to 16–19 weight layers. Hence, we get the point of that using a combination of two 3×3 convolution kernels work better than using one 5×5 convolution kernel and the total amount of parameters ($3\times3\times2$ VS $5\times5\times1$) is reduced. Furthermore, the convolutions with 3×3 kernels are utilized to increase the number of layers of the neural network. And the amount of network calculations is reduced to obtain better experimental performance in the field of lung nodule classification and lung cancer classification. The number of convolution kernels in Dense-Block is 64. The effect of the convolution with kernel size 1×1 is cross-channel characteristic and reduce the convolution kernel parameters to achieve the purpose of simplifying the model to speed up calculation. For this reason, the convolution with kernel size 1×1 is also utilized by us. Every Dense-Block (except the last Dense-Block) is followed by an average pooling layer with size 2×2 and stride 2, namely for a transition layer. At the end of LDNNET network, there are four FC-Dropout structures. For Dropout, in paper [42], Stefan *et al.* proposed that Dropout training is designed for deep neural networks and has been successful on high-dimensional single-layer natural language tasks. Dropout layer preserves the Bayes decision boundary and should therefore induce minimal bias in high dimensions and Dropout training improves the exponent in the generalization bound for empirical risk minimization. Each FC (Fully Connected Layer) contains 8192 neurons which is followed by a dropout layer with the dropout rate 0.5, except for the fourth FC. At last, the probabilities of each input data are obtained by the classifier.

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. Kingma and Ba [43] introduced Adam, an algorithm for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments. The method is straightforward to implement, is computationally efficient, has little memory requirements, is invariant to diagonal rescaling of the gradients and is well suited for problems that are large in terms of data and/or parameters. 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.

On the classification layers of the LDNNET network, the convnets transform the feature maps extracted from the original image into final class scores. If the output scores are regarded as class probabilities, a softmax function is used as the output:

$$\mathrm{f}_j(\mathrm{y}) = \frac{e^{y_j}}{\sum_{k=1}^{k=K} e^{y_k}} \tag{1}$$

where $\mathrm{f_j}$ (y) denotes the probabilities of the *jth* class, $\mathrm{y_x}$ is the output value of the *xth* neuron.

Cross entropy function is utilized as cost function on our convnets. J is cost function, $\tilde{\mathrm{f}}_j$ (y) denotes the desired label probabilities of the training images and $\mathrm{f_j}$ (y) denotes the actual network output:

$$J(\theta) = \frac{1}{\mathrm{s}} \sum_{i=1}^{\mathrm{s}} \sum_{j=1}^{n} \tilde{f}_j(z) \log \left[f_j(z) \right] \tag{2}$$

where θ are parameters, **s** is the value of batch sizes, n spans the classes numbers of input data sets. In our experiment, the classes we set are all 2, that is, n is equal to 2.

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) [53]. The dataset LIDC-IDRI has a total of 1018 CT scans, and each CT image has a label file in xml format. The dataset LUng Nodule Analysis 2016(LUNA16) [55] is the subset of LIDC-IDRI. LUNA16 consists of the CT scans where all nodules $\geq$ 3 mm are accepted by at least 3 out of 4 radiologists. Meanwhile these type of CT scans are not included. Hence LUNA16 has a total of 888 CT scans for lung nodule analysis. This database also contains annotations collected in two phases with four experienced radiologists. Each radiologist marked lesions they identified as non-nodule (<3 mm) and nodule ($\geq$3 mm). We labeled the lung CT images of the LUNA16 according to the annotations provided by the LUNA16. The classification of lung nodules comes from the annotation file in xml format in the dataset LUNA16. Number 1 is in the name of nodule lung CT images and 0 is on behalf of non-nodule lung CT images. This dataset is used to classify nodule and non-nodule of lung CT images. The difference between LIDC-IDRI and LUNA16 can be seen as follows: Zhao *et al.* [44] utilized the LIDC-IDRI dataset for lung nodule classification, in which the lung CT images are divided into three categories "nodule $>$ or $=$ 3 mm", "nodule $<$ 3 mm" and "non-nodule $>$ or $=$ 3 mm"). Meanwhile in this paper, the dataset used for classification is LUNA16, in which the CT scans can

TABLE 2. The architecture details of LDNNET.

Name	Type/Kernel/Stride		Output size
Input			512x512x1
Conv1	Convolution/5x5/1		512x512x96
Dense block 1	Batch Normalization, Relu, Convolution/1x1/1 Batch Normalization, Relu, Convolution/3x3/1 concatenation	X 2	512x512x160 512x512x224
Transition layer 1	Average-pooling/2x2/2		256x256x224
Dense block 2	Batch Normalization, Relu, Convolution/1x1/1 Batch Normalization, Relu, Convolution/3x3/1 concatenation	X 2	256x256x288 256x256x352
Transition layer 2	Average-pooling/2x2/2		128x128x352
Dense block 3	Batch Normalization, Relu, Convolution/1x1/1 Batch Normalization, Relu, Convolution/3x3/1 concatenation	X 2	128x128x416 128x128x480
Transition layer 3	Average-pooling/2x2/2		64x64x480
Dense block 4	Batch Normalization, Relu, Convolution/1x1/1 Batch Normalization, Relu, Convolution/3x3/1 concatenation	X 2	64x64x544 64x64x608
Transition layer 4	Average-pooling/2x2/2		32x32x608
Dense block 5	Batch Normalization, Relu, Convolution/1x1/1 Batch Normalization, Relu, Convolution/3x3/1 concatenation	X 2	32x32x672 32x32x736
Transition layer 5	Average-pooling/2x2/2		16x16x736
Dense block 6	Batch Normalization, Relu, Convolution/1x1/1 Batch Normalization, Relu, Convolution/3x3/1 concatenation	X 2	16x16x800 16x16x864
Transition layer 6	Average-pooling/2x2/2		8x8x864
Dense block 7	Batch Normalization, Relu, Convolution/1x1/1 Batch Normalization, Relu, Convolution/3x3/1 concatenation	X 2	8x8x928 8x8x992
FC1	Fully Connected		8192
Dropout1	Dropout-(50%)		———
FC2	Fully Connected		8192
Dropout2	Dropout-(50%)		———
FC3	Fully Connected		8192
Dropout3	Dropout-(50%)		———
FC4	Fully Connected		Class numbers
Cost	Softmax		———

be divided into nodule and non-nodule. Further, the nodules are all $\geq$ 3 mm and there are no lung nodules smaller than 3mm in the LUNA16 dataset. In addition, in this paper, the input image size is 48 $\times$ 48 and in paper [44], the size is 53 $\times$ 53, which are both not the one whole CT slice. Monkam *et al.* [45] also utilized LIDC-IDRI dataset for lung nodule classification. The lung CT images are divided into two categories ''micro-nodule'' and ''non-nodule''). And the image sizes are set to 20 $\times$ 20, 16 $\times$ 16, 12 $\times$ 12, 8 $\times$ 8 and 4 $\times$ 4 which are all not the original sizes of CT scans and are relatively small in size, hence, ''micro-nodule'' is called.

In 2017, Kaggle had a competition for lung cancer classification. The dataset is called Kaggle DSB 2017 [56]. This dataset provided the lung CT images in DICOM format. CT images with lung cancer are marked as 1, otherwise they are marked as 0. There are nine labeled attributes for each nodule, i.e., subtlety, internal structure, calcification, sphericity, margin, lobulation, spiculation, radiographic solidity and malignancy. For each nodule, its malignancy rating (attribute rating) is evaluated by radiologists, the score of which ranges from 1 to 5 where 1 denotes highly unlike malignant (highly without the given attribute) and 5 denotes highly malignant. In this paper, lung CT images are identified as non-cancer

TABLE 3. Detailed parameters of the dataset LUNA16 AND Kaggle DSB 2017.

Dataset	Collection Statistics	Details
LUNA16	Image format	MHD(Descriptive information of lung CT) RAW(CT information)
	Number of Patients	888
	Number of series	888 CT
	Number of nodules	1186
	Number of non-nodules	551065
	Number of images	754976
	Ratio of nodule and non-nodule	464:1
	Diameter of nodules	Greater than 3mm less than 50mm
	Image Size (GB)	110
	Classification standard/Label	Non-nodule/0 and Nodule/1
Kaggle DSB 2017	Image format	DICOM
	Number of Patients	1397
	Number of series	1397CT
	Number of images	96209
	Image Size (GB)	150
	Classification standard/Label	Non-cancer/0 and Cancer/1

(score$\leq$3) and nodule (score$\geq$4). This dataset is used to classify cancer and non-cancer of lung CT images. The detailed parameters of LUNA16 and Kaggle DSB 2017 are as shown in TABLE 3.

The dataset LUNA16 is saved as the formats of MHD and RAW. The files of MHD format contain the descriptive information of lung CT for instance object type and element type. The files of RAW format contain the original lung CT image information. There is a one-to-one correspondence between a RAW file and an MHD file to get the PNG format lung CT images for training. The dataset Kaggle DSB 2017 is saved as the format of DICOM. DICOM contains not only lung CT image information, but also patient information, hospital information and equipment information. However, only the part of image information is needed to be inputted into deep neural network training. Here PNG format was selected as the image format for LDNNET. In addition, the network training will be burdened by the extra information in the DICOM format other than images. The class packet PYDICOM is utilized to extract the image information inside the DICOM format. Furthermore, the image is saved in the PNG format by the class packet CV2 (OPENCV).

2) ENHANCEMENT AND LABEL FOR CT IMAGES

From the above table 3, it can be seen that the ratio of nodule and non-nodule are not balanced, so data enhancement measures are adopted to solve this problem. And overfitting is a result of network parameters greatly outnumbering the number of features in the input images [3] The number of lung CT images for experiment, which are chosen from datasets LUNA16 and Kaggle DSB 2017, is relatively small. Hence the method of enhancing data is executed by us for avoiding overfitting. The images are expanded by a factor of 20 in this paper. The methods like rotations, horizontal, flip, clip, blurry and so on, are utilized in this paper. The examples of the enhancement of lung CT images can be seen from FIGURE 4. The image in the column a is the original lung CT image. Furthermore, the images of column b, column c and column d are obtained by some random cropping and rotation *et al.* operations in the first column.

3) DATASET SETTINGS USED IN THE EXPERIMENT

Firstly, in dataset LUNA16, 100000 of the 754,976 original lung CT images are selected. There is a problem in dataset LUNA16 that the unbalanced ratio of non-nodule (551065) and nodule (1186) sample images is too large, so in order to balance the samples during LDNNET network training. Hence, we select all the lung CT images of nodule. Further, these images are data-enhanced, according to the way in FIGURE 4 above. Then in the rest of LUNA16 dataset, we select 50,000 lung CT images of non-nodule. Hence, we select the same number of samples for nodule and non-nodule images for LDNNET the network training. This approach can greatly guarantee the authenticity of the experiment.

In dataset Kaggle DSB 2017, 90000 original images are selected from the total of 96209, which are almost all the images in Kaggle DSB 2017 dataset. Secondly, among dataset used for our experiment, the number of training set images, validation set images and test set images account for 54000, 18000, 18000 for Kaggle DSB 2017 and 60000, 20000, 20000 for LUNA16. The ratio is 0.6, 0.2, 0.2. These images are divided into training set images, validation set images and test set images according to random, not manual intervention. This dataset obtained in this way is more objective, so that the network LDNNET can learn the real classification capability to the greatest extent. Thirdly, our classification experiment uses a lot of data on the two datasets, especially in Kaggle DSB 2017, almost all the images are used for the experiment. In addition, the test set has reached 20000 in LUNA116 dataset. The detailed parameters of the dataset are displayed in TABLE 4.

The above settings of our experimental datasets ensure the authenticity of the experiment to the greatest extent. Hence, we don't take 10-fold cross-validation. Due to the relatively large number of datasets in our experiment, we adopt

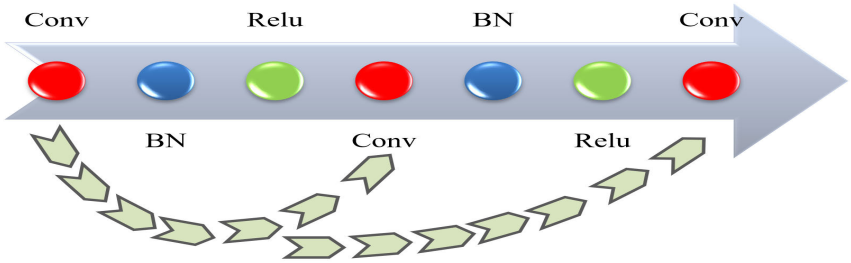


FIGURE 3. Structure of the Dense-Block.

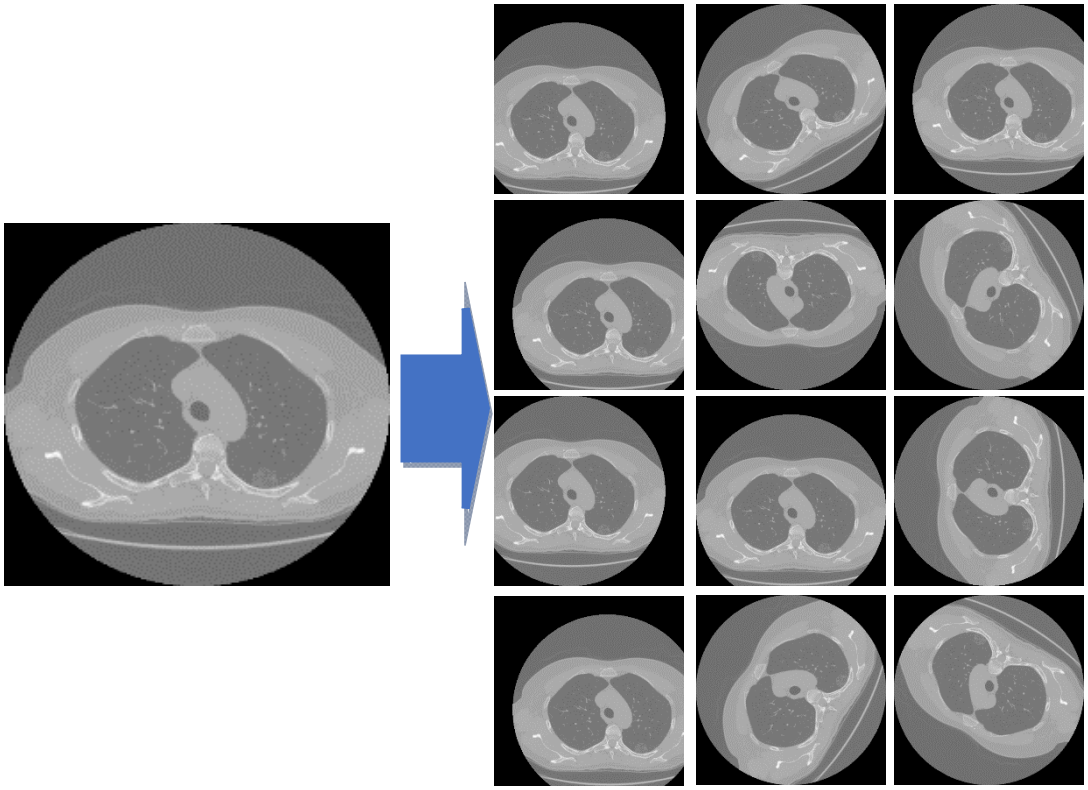


FIGURE 4. Examples of the enhancement of datasets of LUNA16 and Kaggle DSB 2017.

3-fold cross-validation which is enough for the experiment in this paper. This process is iteratively executed 3 times using the 3 different combinations of subgroups. The area under the receiver operating characteristic (ROC) curve (AUC) measurements can be seen from FIGURE 5. And AUC for LUNA16 is over 0.98 and AUC for Kaggle DSB 2017 is over 0.99.

The nodule classification on the LUNA16 dataset is shown in FIGURE 6. In FIGURE 6, the lung CT image of part b is the original image on the LUNA16 dataset whose size is 512×512. The lung candidate nodule images to be inputted into LDNNET network are like part a and part c of FIGURE 6, in which the nodule part and the non-nodule part are marked with red boxes in the part b of FIGURE 6. The pixel size of the lung CT images of the nodule candidates is 48×48. The lung nodule image samples gotten from LUNA16 dataset are lung CT images of nodule candidates which are produced from this kind of image like part a and part c of FIGURE 6. Part a and part c of FIGURE 6 are not taken from part b. Furthermore, part b is just a sample of the original lung CT image. LUNA16 dataset divides the data into two classes: nodule and non-nodule. The classification of lung CT images on the LUNA16 dataset is completed by the configuration files on the LUNA16 dataset. The size of the lung CT images sent to LDNNET network for training is 48×48 which are the lung CT images of the nodule candidates obtained from the LUNA16 dataset.

The classification of the lung CT images on the Kaggle DSB 2017 dataset is shown in FIGURE 7. The sample images obtained from the Kaggle DSB 2017 dataset are the original

TABLE 4. Detailed parameters of the dataset used in this paper.

Properties	LUNA16	Kaggle DSB 2017
Percentage of training	60%	60%
Percentage of validation	20%	20%
Percentage of testing	20%	20%
Image Size	48x48x1 pixels	512x512x1 pixels
Image Format	PNG	PNG
Original CT format	MHD RAW	DICOM
Number of the images	100000	90000
Number of training epochs	200	200

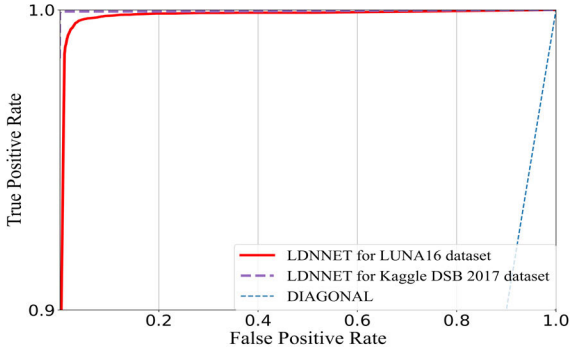


FIGURE 5. ROC curves for LDNNET networks on LUNA16 dataset and Kaggle DSB 2017 dataset.

lung CT images whose pixel size is 512 × 512. The lung CT images on the Kaggle DSB 2017 dataset are divided into two parts of cancer and non-cancer which can be seen from part a, part b and part c of FIGURE 7. In part a and part b of FIGURE 7, each cancer position is marked with a red box. The discrimination between cancer and non-cancer is also based on the configuration file on the Kaggle DSB 2017 dataset. The size of the lung CT images sent to LDNNET network for training is 512×512 which is the original size of images obtained from Kaggle DSB 2017 dataset.

B. EXPERIMENT ENVIRONMENTAL CONFIGURATION AND NETWORK TRAINING PARAMETERS

In this paper, TABLE 4 shows the experiment environment details. Furthermore, experiment environmental configuration and network training parameters are shown as follows. Because the memory of our experimental equipment is relatively small. The size of memory is 32G. Hence, the size of batch is set to 32. The learning rate is set to 0.0001. The dropout rate is set to 0.5. The number of the epoch is set to 200. From FIGURE 8, it can be seen that the model accuracy gets steady after 150 of the iteration steps. Hence, we set 200 as the iteration number to ensure that the training convergence is enough, in addition to make the most use of the lung CT images. The experimental accuracy by epochs of LDNNET network with no preprocessing, with median filtering, with enhanced contrast and with Laplacian filter for lung nodule classification on LUNA16 dataset and lung cancer classification on Kaggle DSB 2017 dataset are shown in part a and part b of FIGURE 8.

For the performance comparison parameters of the proposed LDNNET network, accuracy, precision, recall, dice similarity coefficient (DSC), specificity, sensitivity and FPs/Scan are chosen by us. The details of these metrics are shown as TABLE 6. In TABLE 6, TP, FP, FN, and TN are represented by the number of True Positives, False Positives, False Negatives and True Negatives, respectively. The parameter scan represents the number of scan series of the dataset which is executed in this paper.

V. EXPERIMENTAL RESULTS EVALUATION AND COMPARISON

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

Image preprocessing methods are combined with the LDNNET neural network to obtain the classification result of the lung CT images. For the aim of testing the accuracy and robustness of our proposed method, no preprocessing and preprocessing methods are obtained before inputting the lung CT images into the LDNNET in TABLE 7 and TABLE 8. Image preprocessing will reduce or improve the performance of deep neural networks. One of our work is to compare the impacts of some common preprocessing methods on LDNNET. The preprocessing methods of the lung CT images are composed of enhanced contrast, median filtering, Laplacian filtering. The median filtering method is a nonlinear 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, a mask is utilized to enhance the 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 LDNNET, meanwhile LDNNET with no preprocessing is executed as the comparison. The results of experiments on LUNA16 dataset and Kaggle DSB 2017 dataset are shown in TABLE 7 and TABLE 8.

It can be seen from the above TABLE 7 and TABLE 8 that through preprocessing methods which consist of enhanced contrast, median filtering and Laplacian filtering, the accuracy of the algorithm is generally not improved or even the accuracy rate has decreased, whether it is on the LUNA16 dataset or Kaggle DSB 2017 dataset. The reason

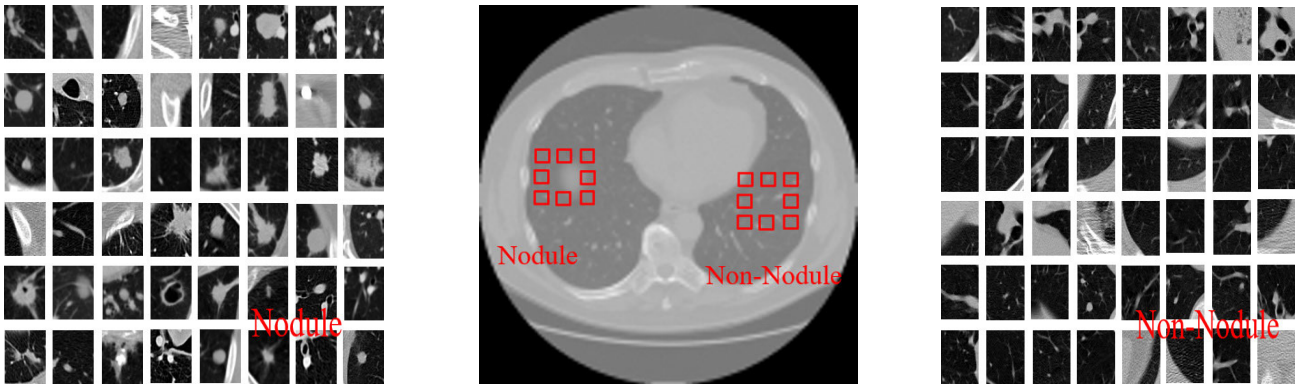


FIGURE 6. Samples of lung nodule and lung non-nodule on LUNA16 dataset.

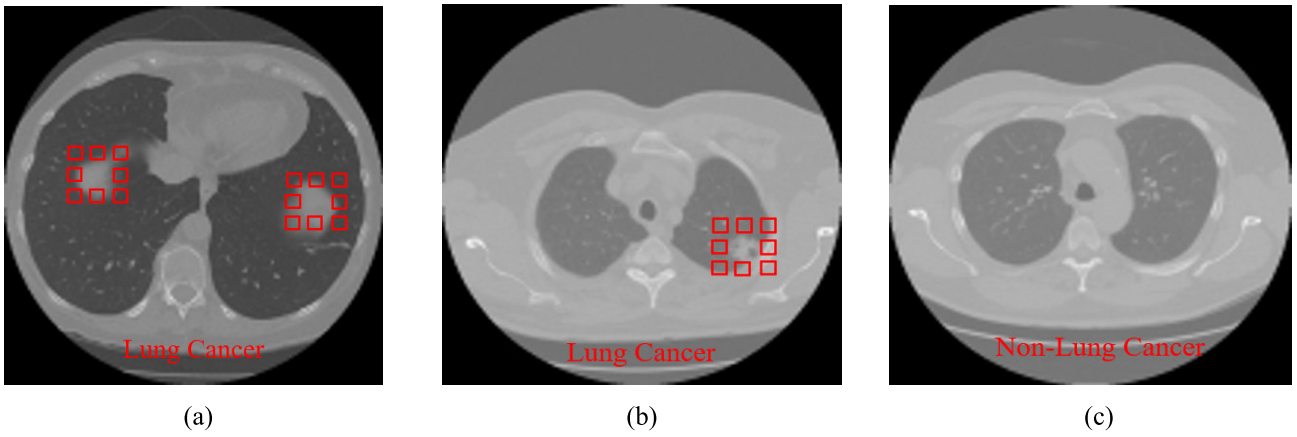


FIGURE 7. Samples of lung cancer and non-lung cancer on Kaggle DSB 2017 dataset.

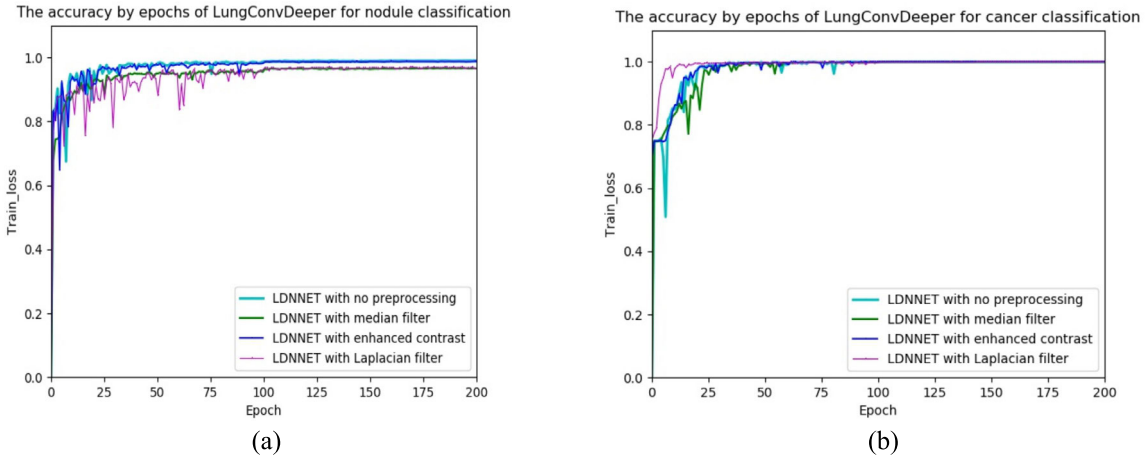


FIGURE 8. The accuracy by epochs of LDNNET for lung nodule classification on LUNA16 dataset (a) and lung cancer classification on Kaggle DSB 2017 dataset (b).

for the decrease in accuracy is that the image preprocessing method on lung CT images will cause the original pixel information of the image to be lost. Furthermore, the network LDNNET proposed in this paper is insufficient in learning the classification information based on the lung CT image in the pixel, which leads to the accuracy decline. Further, the accuracy of network LDNNET on both LUNA16 dataset and Kaggle DSB 2017 dataset are above 0.964726 and 0.999117,

TABLE 5. Detailed parameters of the experimental environment.

Parameter	Details
CPU	Intel (R) Core (TM) i9-7900X CPU @ 3.30GHz 3.31GHz
GPU	NVIDIA GeForce GTX 1080 Ti
Memory	32.0 GB
Operating System	WINDOWS 10
Programming Software	PYCHARM
Programming Language	PYTHON
Neural Network Framework	TENSORFLOW[54]

TABLE 6. Detailed parameters of the experimental environment.

Metric	Detail
Accuracy	(TP + TN)/(TP + FP + TN + FN)
Precision	TP/(TP+FP)
Recall	TP/(TP+FN)
Dice coefficient	2TP/(2TP + FP + FN)
Specificity	TN/(TN + FP)
Sensitivity	TP/(TP+FN)
FPs/Scan	FP/Scan

TABLE 7. The comparison of accuracy by LDNNET on the LUNA16 dataset.

Method	Accuracy	Precision	Recall	DSC	Sensitivity	Specificity
LDNNET with median filter	0.964765	0.973607	0.954634	0.964027	0.954634	0.974678
LDNNET with Laplacian filter	0.964726	0.967289	0.961182	0.964226	0.961182	0.968195
LDNNET with enhanced contrast	0.986160	0.990713	0.981214	0.985941	0.981214	0.991000
LDNNET with no preprocessing	**0.988396**	**0.994396**	**0.982071**	**0.988196**	**0.982072**	**0.994585**

TABLE 8. The comparison of accuracy by LDNNET on the Kaggle DSB 2017 dataset.

Method	Accuracy	Precision	Recall	DSC	Sensitivity	Specificity
LDNNET with median filter	**0.999636**	0.999636	**0.999179**	**0.999282**	**0.999179**	0.999791
LDNNET with Laplacian filter	0.999117	**0.999794**	0.996716	0.998253	0.996716	**0.999930**
LDNNET with enhanced contrast	0.999532	0.999384	0.998768	0.999076	0.998768	0.999791
LDNNET with no preprocessing	0.999480	0.998974	0.998974	0.998974	0.998974	0.999652

which shows that LDNNET network has good robustness to classify lung nodule and lung cancer. However, in some special cases, preprocessing methods can slightly improve the accuracy, for example, the image processing method of median filter on the Kaggle DSB 2017 dataset in TABLE 7. The accuracy can reach 0.999636. Median filtering is a non-linear smoothing algorithm, which sets the gray value of each pixel to the median of the gray values of all pixels in a certain neighborhood window of that point. In the field of image processing, median filtering can protect the pixel edges of the image from being blurred and remove noise. Therefore, in the TABLE 8, the accuracy of preprocessing with median filtering is higher than that without preprocessing by 0.015%. This preprocessing of median filter does not destroy the pixel information here, but strengthens the classification information between pixels and then the network LDNNET can learn more of this information to improve the accuracy. In conclusion, LDNNET network both achieve high accuracy in lung nodule classification on LUNA16 dataset and lung cancer classification on Kaggle DSB 2017 dataset with the effect preprocessing methods of lung CT images. Therefore, LDNNET network has both good robustness and good performance in the field of lung nodule classification and lung cancer classification.

B. EFFECT OF INPUT IMAGE SIZE

Medical CT images are generally huge. However, it is impossible to enter the original lung CT image into LDNNET during the process of classification. The pixel size of the original lung CT image acquired in the Kaggle DSB 2017 dataset is 512 × 512. Hence, in this paper, the image of the LUNA16 dataset is resized to 512 × 512 and other pixel sizes for comparison. The pixel size of the original lung CT image acquired in the LUNA16 dataset is 48 × 48. Because the image pixel size of 48 × 48 is relatively small, in this paper, the lung CT images of the comparison experiment will not be performed according to the multiple reduction. Otherwise, there will be input images, for instance, 3 × 3 and 6 × 6. The pixel sizes of these types of images are too small, which will make the judgment standard of network accuracy lose academic fairness. Therefore, in this paper, the pixel sizes of the lung CT image to be inputted into LDNNET network are set as 80 × 80, 64 × 64, 48 × 48, 32 × 32 and 16 × 16. In this set of comparative experiments, this series of pixel

TABLE 9. Statistical analysis of different input image size by LDNNET on the LUNA16 dataset.

Size of lung CT images	Accuracy	Precision	Recall	DSC	Sensitivity	Specificity
80x80	0.983192	0.991435	0.974433	0.98286	0.974433	0.991763
64x64	0.987394	0.994150	0.980279	0.987166	0.980279	0.994356
48x48	**0.988396**	**0.994396**	**0.982071**	**0.988196**	**0.982072**	**0.994585**
32x32	0.981534	0.989303	0.973186	0.981178	0.973186	0.989703
16x16	0.944102	0.958202	0.92743	0.942565	0.927430	0.960415

TABLE 10. Statistical analysis of different input image size by LDNNET on the Kaggle DSB 2017 dataset.

Size of lung CT images	Accuracy	Precision	Recall	DSC	Sensitivity	Specificity
512x512	0.999480	0.998974	0.998974	0.998974	0.998974	0.999652
256x256	**0.999896**	**1.000000**	**0.999589**	**0.999795**	**0.999589**	**1.000000**
128x128	0.999480	0.999589	0.998358	0.998973	0.998358	0.998973
64x64	0.999324	0.998973	0.998973	0.998665	0.998358	0.999652
32x32	0.998909	0.997947	0.997742	0.997845	0.997742	0.999304

sizes are utilized to fully demonstrate the impact of pixel size on network performance which can be seen in TABLE 9 and TABLE 10.

From TABLE 9, the size sequence of lung CT images is 80×80, 64×64, 48×48, 32×32 and 16×16. The results show that when the pixel size of the image is 48×48, the network LDNNET has the highest accuracy in classifying lung nodules which is 0.988396. However, the data in TABLE 9 here is not the same as in TABLE 10. The accuracy of the LDNNET network will decrease as the image pixels decrease always. The pixel size of the original lung CT image obtained from Kaggle DSB 2017 dataset in TABLE 10 is 512×512, which is more suitable for reduction in multiples, but the pixel size of the original lung CT image obtained from LUNA16 dataset in TABLE 9 is 48×48. The size of these images can no longer be reduced by multiples. In TABLE 9 the pixel size of the lung CT images of 64×64 and 80×80 are both enlarged from 48×48. This will cause 64×64 and 80×80 lung CT images with large pixel size, but they lose the lung nodule classification information in the original images. Therefore, the LDNNET network cannot learn lung nodule classification skills well which resulted in a decrease in accuracy. For lung CT images of different pixel sizes inputted into the LDNNET network, LDNNET can give very good accuracy for each pixel size.

Meanwhile as demonstrated in TABLE 10, LDNNET is trained by different sizes of the input images for instance, 512×512 and so on. Additionally, the performance of the bigger image size is better relatively. Because the reduction of the image will lead to the loss of semantic information, eventually the LDNNET network cannot learn enough features for lung cancer classification. LDNNET 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 512×512 image is sixteen times the pixel information of the 128×128 image. Hence, when the pixels obtained by LDNNET are insufficient, the deep neural network cannot learn sufficient lung cancer classification performance by itself. At last, images with larger pixels will have a higher accuracy than images with smaller pixels by LDNNET. Although in TABLE 10, the classification accuracy of the 256×256 pixel image size is 0.998996, which is slightly higher than the classification accuracy of the 512×512 pixel image size by 0.041%. The reason for the violation of the overall trend here should be that 256×256 pixels are smaller than 512×512 pixels. As the number of training pixels is reduced, overfitting is more likely to occur. But this still hasn't changed the overall trend that as the pixels decrease, the accuracy decreases. Hence, it can be concluded from the TABLE 10 that as the image pixels become smaller, the effect of LDNNET classification of lung nodule is gradually getting worse.

Hence the conclusion can be made that the network still has high robustness in lung nodule classification and lung cancer classification for input lung CT images of different pixel sizes.

C. 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 structure of average pooling and maximum pooling can be seen in FIGURE 9. There are also many other types of pooling layers, for instance, Yu *et al.* [46] proposed the mixed pooling layer to regularize CNNs, which replaces the deterministic pooling operations with a stochastic procedure by randomly using the conventional max pooling and average pooling methods. Inspired by Dropout (that randomly sets half the activations to zeros), the proposed mixed pooling method replaces the conventional deterministic pooling operations with a stochastic procedure, randomly employing the max

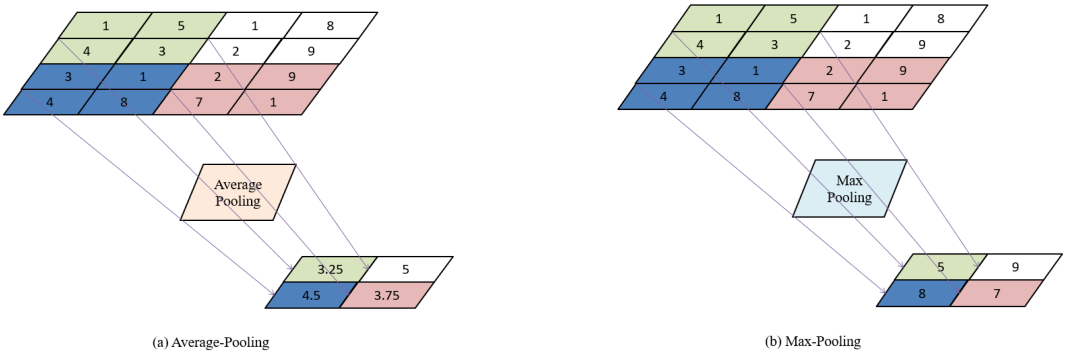


FIGURE 9. The structures of the two types of pooling layers.

TABLE 11. Statistical analysis of type and number of pooling by LDNNET on the LUNA16 dataset.

Type of pooling	Accuracy	Precision	Recall	DSC	Sensitivity	Specificity
LDNNET (five-layer average pooling)	0.983732	0.989351	0.977629	0.983455	0.977629	0.989703
LDNNET (five-layer maximum pooling)	0.980532	0.989592	0.970847	0.98013	0.970847	0.990008
LDNNET (six-layer average pooling)	0.986160	**0.992107**	0.979811	0.985921	0.979811	**0.992373**
LDNNET (six-layer maximum pooling)	0.981689	0.988687	**0.988687**	0.98135	**0.989093**	0.974121
LDNNET (seven-layer average pooling)	**0.988396**	0.994396	0.982071	**0.988196**	0.982072	0.994585
LDNNET (seven-layer maximum pooling)	0.980879	0.987895	0.973264	0.980525	0.973264	0.988330
LDNNET (eight-layer average pooling)	0.984541	0.990682	0.977941	0.984270	0.977941	0.991000
LDNNET (eight-layer maximum pooling)	0.981920	0.990788	0.972484	0.981551	0.972484	0.991152
LDNNET (nine-layer average pooling)	0.986083	0.990017	0.981760	0.985871	0.981760	0.990313
LDNNET (nine-layer maximum pooling)	0.981226	0.989606	0.972250	0.980852	0.972250	0.990008

TABLE 12. Statistical analysis of type and number of pooling by LDNNET on the Kaggle DSB 2017.

Type of pooling	Accuracy	Precision	Recall	DSC	Sensitivity	Specificity
LDNNET (five-layer average pooling)	0.999636	0.999589	0.998974	0.999281	0.998974	0.999860
LDNNET (five-layer maximum pooling)	0.999896	0.999794	**0.999795**	0.999795	**0.999795**	0.999930
LDNNET (six-layer average pooling)	0.999480	0.999589	0.998358	0.998973	0.998358	0.999861
LDNNET (six-layer maximum pooling)	0.999792	0.999795	0.999384	0.999589	0.999384	0.999930
LDNNET (seven-layer average pooling)	0.999480	0.998974	0.998974	0.998974	0.998974	0.999652
LDNNET (seven-layer maximum pooling)	0.999896	**1.000000**	0.999589	0.999795	0.999589	**1.000000**
LDNNET (eight-layer average pooling)	0.999740	0.999589	0.999384	0.999487	0.999384	0.999861
LDNNET (eight-layer maximum pooling)	**0.999948**	**1.000000**	**0.999795**	**0.999897**	**0.999795**	**1.000000**
LDNNET (nine-layer average pooling)	0.999584	0.999589	0.998768	0.999179	0.998768	0.999861
LDNNET (nine-layer maximum pooling)	0.999740	0.999589	0.999384	0.999487	0.999384	0.999861

pooling and average pooling methods during the training of CNNs.

To further explore the effect of pooling layer function on LDNNET classification of lung nodules and classification of lung cancer, comparative experiments are designed in this paper. There are five-layer average pooling, five-layer maximum pooling, six-layer average pooling, six-layer maximum pooling, seven-layer average pooling, seven-layer maximum pooling, eight-layer average pooling, eight-layer maximum pooling, nine-layer average pooling and nine-layer maximum pooling. The results of the comparative experiments can be seen in TABLE 11 and TABLE 12.

As can be seen from the TABLE 11, both average pooling and maximum pooling have relatively little influence on the accuracy of the LDNNET network on lung nodule classification, in which no matter which pooling layer structure is used, the accuracy rate is between 0.980000 and 0.987000. Furthermore, the accuracy rate error is between 0.005000. It can also be seen from the comparison in the TABLE 11 that under the premise of the same number of pooling layers, the average pooling layer has a higher accuracy than the maximum pooling layer, in which the difference of average accuracy is about 0.0048.

In TABLE 12, since the accuracy of LDNNET network on Kaggle DSB 2017 dataset is above 0.999480, there is no obvious difference between the average pooling layer and the maximum pooling layer on LDNNET network for lung cancer classification, in which the accuracy of average pooling is 0.00027 lower than the maximum accuracy.

From above, we can get the conclusion that the average pooling layer has better performance than the maximum pooling layer on LDNNET network for lung nodule classification

TABLE 13. Statistical analysis of the effect of dense connection by LDNNET on the LUNA16 dataset.

Existence of Dense Connection	Accuracy	Precision	Recall	DSC	Sensitivity	Specificity
LDNNET with Dense Connection	**0.988396**	**0.994396**	**0.982071**	**0.988196**	**0.982072**	**0.994585**
LDNNET without Dense Connection	0.973477	0.975558	0.970691	0.973119	0.970691	0.976203

TABLE 14. Statistical analysis of the effect of dense connection by LDNNET on the Kaggle DSB 2017 dataset.

Existence of Dense Connection	Accuracy	Precision	Recall	DSC	Sensitivity	Specificity
LDNNET with Dense Connection	**0.999480**	**0.998974**	**0.998974**	**0.998974**	**0.998974**	**0.999652**
LDNNET without Dense Connection	0.998181	0.995899	0.996921	0.996410	0.996921	0.998608

TABLE 15. Statistical analysis of different numbers of DENDE blocks by LDNNET on the LUNA16 dataset.

Number of Dense-Blocks	Accuracy	Precision	Recall	DSC	Sensitivity	Specificity
LDNNET with five Dense Blocks	0.984850	0.990688	0.978564	0.984589	0.978564	0.991000
LDNNET with six Dense Blocks	0.986122	0.992029	0.979811	0.985882	0.979811	0.992297
LDNNET with seven Dense Blocks	**0.988396**	**0.994396**	**0.982071**	**0.988196**	**0.982072**	**0.994585**
LDNNET with eight Dense Blocks	0.985929	0.992259	0.979188	0.985680	0.979188	0.992525
LDNNET with nine Dense Blocks	0.984464	0.990448	0.978019	0.984194	0.978019	0.990771

TABLE 16. Statistical analysis of different numbers of DENDE blocks by LDNNET on the Kaggle DSB 2017 dataset.

Number of Dense-Blocks	Accuracy	Precision	Recall	DSC	Sensitivity	Specificity
LDNNET with five Dense Blocks	**0.999896**	**0.999795**	**0.999795**	**0.999795**	**0.999795**	**0.999930**
LDNNET with six Dense Blocks	0.999376	0.998973	0.998563	0.998768	0.998563	0.999652
LDNNET with seven Dense Blocks	0.999480	0.998974	0.998974	0.998974	0.998974	0.999652
LDNNET with eight Dense Blocks	0.999584	0.999589	0.998768	0.999179	0.998768	0.999861
LDNNET with nine Dense Blocks	0.999688	**0.999795**	0.998974	0.999384	0.998974	**0.999930**

but worse for lung cancer classification. Further, both two types of pooling layer structures can make the LDNNET network have a high accuracy rate, which also shows that the LDNNET network has good robustness in this aspect of the type of pooling layer.

D. EFFECT OF DENSE CONNECTION

LDNNET utilized the structure of 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-Block alleviates 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 nodule classification and lung cancer classification by LDNNET.

From the above TABLE 13, the application of dense connection has a significant influence on the accuracy of the network, in which LDNNET without dense connection has the higher accuracy than LDNNET with dense connection by 1.53254% on LUNA16 dataset.

In the TABLE 14, the effect of dense connection on Kaggle DSB 2017 on accuracy is not as great as that of dense connection on LUNA16 dataset. But in the case of utilizing dense connection, it will still increase the accuracy by 0.13014% compared with the case of not using dense connection. From the expression of the experimental results in this paragraph and the previous paragraph, we can conclude that dense connection can indeed improve the accuracy of the LDNNET network in terms of lung nodule classification and lung cancer classification.

Because the dense connection in LDNNET is realized by the Dense-Block structure. To further explore the effect of dense connection on LDNNET for classification of lung nodules and classification of lung cancer, further comparative experiments are designed. There are five Dense-Blocks, six Dense-Blocks, seven Dense-Blocks, eight Dense-Blocks and nine Dense-Blocks within LDNNET. The results of the comparative experiment can be seen in TABLE 15 and TABLE 16.

In the TABLE 15, with the increase in the number of Dense-Blocks, the accuracy of the LDNNET network is constantly improving for lung nodule classification which means that Dense-Block enables the LDNNET network to learn more fully the classification information of lung nodules, thereby improving accuracy.

In the TABLE 16, on Kaggle DSB 2017 dataset, the effect of Dense-Blocks is not evident like TABLE 15. The accuracy of utilizing different numbers of Dense-Block can reach more than 0.999376. Therefore, it does make little sense to modify the number of Dense-Block to improve the accuracy of the LDNNET network on the Kaggle DSB 2017 dataset.

In conclusion, from above TABLE 13, TABLE 14, TABLE 15 and TABLE 16, Dense-Block can improve the accuracy of the LDNNET network in lung nodule classification and lung cancer classification, therefore in this paper, this structure Dense-Block is adopted to increase accuracy.

E. COMPARISON WITH OTHER METHODS AND ROBUSTNESS EVALUATION

Masood *et al.* [47] proposed a Computer-Assisted Decision Support System in Pulmonary Cancer by using the novel model based on deep learning and metastasis information obtained from MBAN (Medical Body Area Network). The proposed model, DFCNet is based on the deep fully convolutional neural network (FCNN) which is used for classification of each detected pulmonary nodule into four lung cancer stages. Lan *et al.* [48] proposed a new network named RUN to complete nodule detection in a single step by bypassing the candidate selection. The system introduced the shortcut of the residual network to improve the traditional U-Net, thereby solving the disadvantage of poor results due to its lack of depth. Furthermore, the experimental results were compared with the traditional U-Net. Jung *et al.* [49] proposed a three dimensional deep convolutional neural network (3D DCNN) with shortcut connections and a 3D DCNN with dense connections for lung nodule classification. The shortcut connections and dense connections successfully alleviate the gradient vanishing problem by allowing the gradient to pass quickly and directly. Connections help deep structured networks to obtain general as well as distinctive features of lung nodules. Moreover, they increased the dimension of DCNNs from two to three to capture 3D features. Nasrullah *et al.* [2] proposed a novel deep learning-based model with multiple strategies for the precise diagnosis of the malignant nodules. Due to the recent achievements of deep convolutional neural networks (CNN) in image analysis, two deep three-dimensional (3D) customized mixed link network (CMixNet) architectures were utilized for lung nodule detection and classification, respectively. Nodule detections were performed through faster R-CNN on efficiently-learned features from CMixNet and U-Net like encoder–decoder architecture. Dou *et al.* [50] proposed a novel framework with 3D convolutional networks (ConvNets) for automated detection of pulmonary nodules from low-dose CT scans, which is a challenging yet crucial task for lung cancer early diagnosis and treatment. El-Bana *et al.* [51] proposed a two-stage framework that exploits the ever-growing advances in deep neural network model and that is comprised of a semantic segmentation stage followed by localization and classification. The recently published DeepLab model was employed for semantic segmentation and that it significantly improves the accuracy of nodule detection compared to the classical U-Net model and its most recent variants was shown. Polat and Danaei Mehr [52] proposed two Convolutional Neural Network (CNN)-based models as deep learning methods to diagnose lung cancer on lung CT images. To investigate the performance of the two proposed models (Straight 3D-CNN with conventional softmax and hybrid 3D-CNN with Radial Basis Function (RBF)-based SVM), the altered models of two-well known CNN architectures (3D-AlexNet and 3D-GoogleNet) were considered. It is needed to be emphasized that the accuracy of paper [51] is relatively high, with an accuracy of 0.964000 on LUNA16 dataset and an accuracy of 0.912000 on Kaggle DSB 2017 dataset. However, the algorithm in paper [51] is divided into two steps. The first step is semantic segmentation and positioning of lung CT images. The second step is to classify lung CT images. On the contrary, LDNNET network directly uses original lung CT images, and does not use semantic segmentation and positioning. In the end, LDNNET network can obtain better results which indicated that our network has better performance and higher robustness. The above methods are current methods which are widely applied in lung nodule classification and lung cancer classification. Further, these methods are utilized to compare the performance of the LDNNET network. Performance comparison for LDNNET and other methods on the LUNA16 dataset and Kaggle DSB 2017 dataset can be seen in TABLE 17 and TABLE 18.

In TABLE 17, LDNNET network obtained the highest accuracy rate among other methods on lung nodule classification, which is 0.988396. In contrast, the accuracies of lung nodules classification by CNN [47] and DFCNet [47] are lower, only 0.740100 and 0.801200. The accuracies of Deeplab-V3 plus(ex_65) + FRCNN-Inception-V2 [51] and Multi-Resolution CNN + Knowledge Transfer [4] are slightly lower than that of our proposed network LDNNET, and the accuracy values are 0.970000 and 0.973300. The sensitivity of LDNNET network is higher than other methods in TABLE 17. Meanwhile, the LDNNET network achieved the lowest value of FPs/Scan, which is 0.07995. Therefore, LDNNET has better performance on lung nodule classification than the current mainstream algorithms. The LDNNET network obtains lower FPs/Scan while obtaining the higher accuracy rate. Hence the LDNNET will have lower false positives, which is more suitable for application in actual medical institutions to assist the doctor in diagnosing the patients' lung CT images, for instance hospitals and clinics.

In TABLE 18, the LDNNET network achieved the highest accuracy, which is 0.999480. The performance of LDNNET provides an improvement over straight 3D-CNN + softmax classifier [52] by 9.7%, Hybrid 3D-CNN + RBF-based SVM [52] by 8.2%, DeeplabV3plus(ex_65) + Mobilenet-V1_1.0_224 [51] by 6.9%, DeeplabV3plus(ex_65) + Inception-V3 [51] by 4.2%. The index FPs/Scan obtained by the network is 0.00357 which is enough for commercialization and medical application. Hence, LDNNET surpassed other algorithms on accuracy of lung cancer classification on Kaggle DSB 2017 dataset. Combining the contents of TABLE 17 and TABLE 18, we can conclude that here the LDNNET network has a high accuracy rate on both the lung nodule classification of the LUNA16 dataset and the lung cancer classification of the Kaggle DSB 2017 dataset, which

TABLE 17. Performance comparison for LDNNET and other models on the LUNA16 dataset.

Method	Sensitivity	FPs/Scan	Accuracy	Specificity
CNN[47]	0.702300	**4.7**	0.740100	0.794700
DFCNet[47]	0.731400	4.2	0.801200	0.819500
RUN:Resdiual-U-Net[48]	0.909000	2	——	——
D48[49]	0.913000	1.6	0.948000	0.984000
3D CMixNet Faster-RCNN[2]	0.900000	1.1	——	——
DeepNet[50]	0.848000	1	——	——
ResNet[50]	0.867000	1	——	——
ResNet+HL[50]	0.905000	1	——	——
ESB-ALL[49]	0.933000	0.7	0.963000	0.993000
Deeplab-V3 plus(ex_65) + FRCNN-Inception-V2[51]	0.964000	0.6	0.970000	0.994000
Multi-Resolution CNN + Knowledge Transfer[4]	0.972600	——	0.973300	0.973800
Our proposed LDNNET network	**0.982072**	0.07995	**0.988396**	**0.994584**

TABLE 18. Performance comparison for LDNNET and other models on the Kaggle DSB 2017 dataset.

Method	Sensitivity	FPs/Scan	Accuracy	Specificity
straight 3D-CNN + softmax classifier[52]	0.864000	——	0.902300	0.930900
Hybrid 3D-CNN + RBF-based SVM[52]	0.885300	——	0.918000	0.942300
DeeplabV3plus(ex_65) + Mobilenet-V1_1.0_224[51]	0.900100	——	0.930000	0.943000
DeeplabV3plus(ex_65) + Inception-V3[51]	0.912000	——	0.956600	0.972400
Our proposed LDNNET network	**0.998973**	**0.00357**	**0.999480**	**0.999652**

proves that the LDNNET network is also robust on these two datasets.

VI. CONCLUSION

In this paper, a new network structure of the LDNNET was proposed. LDNNET network is utilized on two datasets, LUNA16 and Kaggle DSB 2017. LDNNET network is robust for the reason that both in the classification of lung nodules and the classification of lung cancer have both achieved good performance. Further, the accuracy rate for the classification of lung nodules has been reached 0.982072 on dataset LUNA16 by LDNNET network, which surpasses other existing methods. Meanwhile, LDNNET network achieves the accuracy of 0.999480 on lung cancer classification on dataset Kaggle DSB 2017. This accuracy is much higher than the comparison method of other papers. Actually, for this accuracy of 0.999480, LDNNET can be deployed on the field of medical examination on commercial clinics and hospital to assist doctors in diagnosing patients. Furthermore, the comparative experiments were designed in preprocessing, dense connection, input image size, pooling layer and depth of neural network on LUNA16 and Kaggle DSB 2017 datasets for a sufficient comparison experiment. In addition, the results of comparison experiments proved that these parameters have an impact on the performance of the network. Meanwhile, by changing these parameters, the accuracy of lung nodules on LUNA16 still can reach over 0.988396 and the accuracy of lung cancer classification on Kaggle DSB 2017 can reach over 0.999480 too. Hence, that LDNNET network is robust, which can be proved by above comparative experiments. It should be emphasized that the lung CT images sent into LDNNET network of LUNA16 dataset are 48×48 pixel size for lung nodule classification, meanwhile, the lung CT images sent into LDNNET network of Kaggle DSB 2017 dataset are 512×512 pixel size for lung cancer classification. The LDNNET network can achieve higher accuracy than other existing algorithms for different pixel sizes of input lung CT images and different experimental targets. LDNNET does not use semantic segmentation and positioning, hence this network is easier to be deployed and utilized under actual circumstances. Finally, the conclusion can be made that the proposed network not only has good performance in lung nodule classification and lung cancer classification, but also has good robustness on the field of input image pixel size, network structure and both two datasets, LUNA16 and Kaggle DSB 2017. Focusing on the field of medical image classification, lung CT image classification scenarios by using more effective network structures to extract common features from images of different types and performing cross-type matching to improve the performance, which is the first work we have planned for the future. In addition, merging the features of different types of medical images, for instance, brain CT images, liver CT images, bone CT images and so on, to achieve more accurate and more robust medical image classification will be considered in future work. Furthermore, we will apply the LDNNET network to other medical diseases fields, for instance classification and detection of diseases namely breast cancer, brain tumor, colon cancer and diabetic retinopathy. Our pursuit is always to apply deep convolution neural network deployment to commercial organizations and medical units to promote auxiliary medical diagnosis systems.

ACKNOWLEDGMENT

The research in this paper uses the LUng Nodule Analysis 2016 (LUNA16) database which is acquired from the Lung Image Database Consortium, Image Database Resource Initiative (LIDC-IDRI) database which is provided by

the Cancer Image Archive (TCIA) Public Access and Kaggle DATA-SCIENCE-BOWL-2017 (Kaggle DSB 2017) database which is obtained from Kaggle's Data Science Bowl (DSB) 2017.

REFERENCES

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2016," *CA A, Cancer J. Clinicians*, vol. 66, pp. 7–30, Jan. 2015.

[2] N. Nasrullah, J. Sang, M. S. Alam, M. Mateen, B. Cai, and H. Hu, "Automated lung nodule detection and classification using deep learning combined with multiple strategies," *Sensors*, vol. 19, no. 17, p. 3722, Aug. 2019.

[3] I. Ali, G. R. Hart, G. Gunabushanam, Y. Liang, W. Muhammad, B. Nartowt, M. Kane, X. Ma, and J. Deng, "Lung nodule detection via deep reinforcement learning," *Frontiers Oncol.*, vol. 8, p. 108, Apr. 2018.

[4] W. Zuo, F. Zhou, Z. Li, and L. Wang, "Multi-resolution CNN and knowledge transfer for candidate classification in lung nodule detection," *IEEE Access*, vol. 7, pp. 32510–32521, 2019.

[5] N. Gupta, D. Gupta, A. Khanna, P. P. R. Filho, and V. H. C. de Albuquerque, "Evolutionary algorithms for automatic lung disease detection," *Measurement*, vol. 140, pp. 590–608, Jul. 2019.

[6] Y. Chen, Y. Wang, F. Hu, and D. Wang, "A lung dense deep convolution neural network for robust lung parenchyma segmentation," *IEEE Access*, vol. 8, pp. 93527–93547, 2020.

[7] A. M. Anter and A. E. Hassenian, "CT liver tumor segmentation hybrid approach using neutrosophic sets, fast fuzzy C-means and adaptive watershed algorithm," *Artif. Intell. Med.*, vol. 97, pp. 105–117, Jun. 2019.

[8] G. Wei, H. Cao, H. Ma, S. Qi, W. Qian, and Z. Ma, "Content-based image retrieval for lung nodule classification using texture features and learned distance metric," *J. Med. Syst.*, vol. 42, no. 1, p. 13, Jan. 2018.

[9] J. Gong, J.-Y. Liu, L.-J. Wang, X.-W. Sun, B. Zheng, and S.-D. Nie, "Automatic detection of pulmonary nodules in CT images by incorporating 3D tensor filtering with local image feature analysis," *Phys. Medica*, vol. 46, pp. 124–133, Feb. 2018.

[10] B. R. Froz, A. O. de Carvalho Filho, A. C. Silva, A. C. de Paiva, R. A. Nunes, and M. Gattass, "Lung nodule classification using artificial crawlers, directional texture and support vector machine," *Expert Syst. Appl.*, vol. 69, pp. 176–188, Mar. 2017.

[11] J. J. Chabon, E. G. Hamilton, D. M. Kurtz, M. S. Esfahani, E. J. Moding, H. Stehr, J. Schroers-Martin, B. Y. Nabet, B. Chen, A. A. Chaudhuri, and C. L. Liu, "Integrating genomic features for non-invasive early lung cancer detection," *Nature*, vol. 580, pp. 245–251, Apr. 2020.

[12] C. Jacobs, E. M. van Rikxoort, T. Twellmann, E. T. Scholten, P. A. de Jong, J.-M. Kuhnigk, M. Oudkerk, H. J. de Koning, M. Prokop, C. Schaefer-Prokop, and B. van Ginneken, "Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images," *Med. Image Anal.*, vol. 18, no. 2, pp. 374–384, Feb. 2014.

[13] X. Ye, X. Lin, J. Dehmeshki, G. Slabaugh, and G. Beddoe, "Shape-based computer-aided detection of lung nodules in thoracic CT images," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 7, pp. 1810–1820, Jul. 2009.

[14] Ö. Demir and A. Yilmaz Çamurcu, "Computer-aided detection of lung nodules using outer surface features," *Bio-Med. Mater. Eng.*, vol. 26, pp. S1213–S1222, Aug. 2015.

[15] A. Mobiny, P. Yuan, P. A. Cicalese, S. K. Moulik, N. Garg, C. C. Wu, K. Wong, S. T. Wong, T. C. He, and H. V. Nguyen, "Memory-augmented capsule network for adaptable lung nodule classification," *IEEE Trans. Med. Imag.*, early access, Jan. 12, 2021, doi: 10.1109/TMI.2021.3051089.

[16] M. A. Heuvelmans, P. M. A. van Ooijen, S. Ather, C. F. Silva, D. Han, C. P. Heussel, W. Hickes, H.-U. Kauczor, P. Novotny, H. Peschl, M. Rook, R. Rubtsov, O. von Stackelberg, M. T. Tsakok, C. Arteta, J. Declerck, T. Kadir, L. Pickup, F. Gleeson, and M. Oudkerk, "Lung cancer prediction by deep learning to identify benign lung nodules," *Lung Cancer*, vol. 154, pp. 1–4, Apr. 2021.

[17] I. W. Harsono, S. Liawatimena, and T. W. Cenggoro, "Lung nodule detection and classification from Thorax CT-scan using RetinaNet with transfer learning," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 1319, pp. 1–8, Apr. 2020.

[18] Y. Xie, Y. Xia, J. Zhang, Y. Song, D. Feng, M. Fulham, and W. Cai, "Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT," *IEEE Trans. Med. Imag.*, vol. 38, no. 4, pp. 991–1004, Apr. 2019.

[19] I. Ali, M. Muzammil, I. U. Haq, A. A. Khaliq, and S. Abdullah, "Efficient lung nodule classification using transferable texture convolutional neural network," *IEEE Access*, vol. 8, pp. 175859–175870, 2020.

[20] A. Naik and D. R. Edla, "Lung nodule classification on computed tomography images using deep learning," *Wireless Pers. Commun.*, vol. 116, pp. 655–690, Jan. 2021.

[21] C.-J. Lin and Y.-C. Li, "Lung nodule classification using Taguchi-based convolutional neural networks for computer tomography images," *Electronics*, vol. 9, no. 7, p. 1066, Jun. 2020.

[22] R. Dey, Z. Lu, and Y. Hong, "Diagnostic classification of lung nodules using 3D neural networks," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 774–778.

[23] M. Al-Shabi, H. K. Lee, and M. Tan, "Gated-dilated networks for lung nodule classification in CT scans," *IEEE Access*, vol. 7, pp. 178827–178838, 2019.

[24] R. V. M. D. Nobrega, S. A. Peixoto, S. P. P. D. Silva, and P. P. R. Filho, "Lung nodule classification via deep transfer learning in CT lung images," in *Proc. IEEE 31st Int. Symp. Comput. Based Med. Syst. (CBMS)*, Jun. 2018, pp. 244–249.

[25] Y. Qin, H. Zheng, Y. M. Zhu, and J. Yang, "Simultaneous accurate detection of pulmonary nodules and false positive reduction using 3D CNNs," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 1005–1009.

[26] D. Ardila, A. P. Kiraly, S. Bharadwaj, B. Choi, J. J. Reicher, L. Peng, D. Tse, M. Etemadi, W. Ye, G. Corrado, D. P. Naidich, and S. Shetty, "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography," *Nature Med.*, vol. 25, no. 6, pp. 954–961, Jun. 2019.

[27] P. Monkam, S. Qi, H. Ma, W. Gao, Y. Yao, and W. Qian, "Detection and classification of pulmonary nodules using convolutional neural networks: A survey," *IEEE Access*, vol. 7, pp. 78075–78091, 2019.

[28] X. Zhao, S. Qi, B. Zhang, H. Ma, W. Qian, Y. Yao, and J. Sun, "Deep CNN models for pulmonary nodule classification: Model modification, model integration, and transfer learning," *J. X-Ray Sci. Technol.*, vol. 27, no. 4, pp. 615–629, Sep. 2019.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[30] 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

[31] 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.

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

[33] K. 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

[34] 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.

[35] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.

[36] B. Fielding and L. Zhang, "Evolving deep DenseBlock architecture ensembles for image classification," *Electronics*, vol. 9, no. 11, p. 1880, Nov. 2020.

[37] C. Zhao, T. F. Wang, and B. Y. Lei, "Medical image fusion method based on dense block and deep convolutional generative adversarial network," *Neural Comput. Appl.*, vol. 11600, pp. 1–16, Oct. 2020.

[38] B. X. Chen, T. J. Liu, K. H. Liu, H. H. Liu, and S. C. Pei, "Image super-resolution using complex dense block on generative adversarial networks," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 2866–2870.

[39] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.

[40] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. 14th Int. Conf. Artif. Intell. Statist. (AISTATS)*, 2011, pp. 315–323.

[41] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2012, pp. 1097–1105.

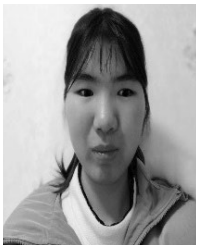
[42] S. Wager, W. Fithian, S. Wang, and P. S. Liang, "Altitude training: Strong bounds for single-layer dropout," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 100–108.
[43] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *Comput. Sci.*, vol. 1412, pp. 1–15, Dec. 2014.
[44] X. Zhao, L. Liu, S. Qi, Y. Teng, J. Li, and W. Qian, "Agile convolutional neural network for pulmonary nodule classification using CT images," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 13, no. 4, pp. 585–595, Apr. 2018.
[45] P. Monkam, S. Qi, M. Xu, H. Li, F. Han, Y. Teng, and W. Qian, "Ensemble learning of multiple-view 3D-CNNs model for micro-nodules identification in CT images," *IEEE Access*, vol. 7, pp. 5564–5576, 2019.
[46] D. Yu, H. Wang, P. Chen, and Z. Wei, "Mixed pooling for convolutional neural networks," in *Proc. Int. Conf. Rough Sets Knowl. Technol.*, 2014, pp. 364–375.
[47] A. Masood, B. Sheng, P. Li, X. Hou, X. Wei, J. Qin, and D. Feng, "Computer-assisted decision support system in pulmonary cancer detection and stage classification on CT images," *J. Biomed. Informat.*, vol. 79, pp. 117–128, Mar. 2018.
[48] T. Lan, Y. Li, J. Kimani Murugi, Y. Ding, and Z. Qin, "RUN: Residual U-Net for computer-aided detection of pulmonary nodules without candidate selection," 2018, *arXiv:1805.11856*. [Online]. Available: http://arxiv.org/abs/1805.11856
[49] H. Jung, B. Kim, I. Lee, J. Lee, and J. Kang, "Classification of lung nodules in CT scans using three-dimensional deep convolutional neural networks with a checkpoint ensemble method," *BMC Med. Imag.*, vol. 18, no. 1, p. 48, Dec. 2018.
[50] Q. Dou, H. Chen, Y. Jin, H. Lin, J. Qin, and P.-A. Heng, "Automated pulmonary nodule detection via 3D convnets with online sample filtering and hybrid-loss residual learning," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2017, pp. 630–638.
[51] S. EL-Bana, A. Al-Kabbany, and M. Sharkas, "A two-stage framework for automated malignant pulmonary nodule detection in CT scans," *Diagnostics*, vol. 10, no. 3, p. 131, Feb. 2020.
[52] H. Polat and H. D. Mehr, "Classification of pulmonary CT images by using hybrid 3D-deep convolutional neural network architecture," *Appl. Sci.*, vol. 9, no. 5, p. 940, Mar. 2019.
[53] 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: Apr. 20, 2012. [Online]. Available: https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI
[54] *Tensorflow*. Accessed: Nov. 9, 2015. [Online]. Available: https://www.tensorflow.org/
[55] *Lung Nodule Analysis 2016 (LUNA16)*. Accessed: 2016. [Online]. Available: https://luna16.grand-challenge.org/Home/
[56] *KAGGLE DATA-SCIENCE-BOWL-2017(Kaggle DSB 2017)*. Accessed: 2017. [Online]. Available: https://www.kaggle.com/c/data-science-bowl-2017/

YING CHEN was born in 1981. He received the B.S., M.S., and Ph.D. degrees from Jilin University, Changchun, Jilin, China.

He is currently an Associate Professor with Nanchang Hangkong University, Nanchang, Jiangxi, China. His main research interests include pattern recognition, machine learning, biometrics, image processing, and deep learning.

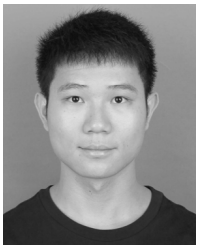
YERONG WANG was born in 1993. He received the B.S. degree in software engineering from the Changchun Institute of Technology, Changchun, Jilin, China. He is currently pursuing the M.S. degree with Nanchang Hangkong University under the supervision of Prof. Ying Chen.

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, Anhui, China. She is currently pursuing the M.S. degree with Nanchang Hangkong University under the supervision of Prof. Ying Chen.

Her main research interests include machine learning and image processing.



LONGFENG FENG was born in 1998. He received the B.S. degree in computer science and technology from the Dongguan University of Technology, Dongguan, Guangdong, China. He is currently pursuing the M.S. degree with Nanchang Hangkong University under the supervision of Prof. Ying Chen.

His main research interests include machine learning and image processing.

TAOHUI ZHOU was born in 1998. He received the B.S. degree in computer science and technology from Nanchang Hangkong University, Nanchang, Jiangxi, China, where he is currently pursuing the M.S. degree under the supervision of Prof. Ying Chen.

His main research interests include machine learning and image processing.

CHENG ZHENG was born in 1998. He received the B.S. degree in computer science and technology from Nanchang Hangkong University, Nanchang, Jiangxi, China, where he is currently pursuing the M.S. degree under the supervision of Prof. Ying Chen.

His main research interests include machine learning and image processing.