



Fusion of 3D lung CT and serum biomarkers for diagnosis of multiple pathological types on pulmonary nodules

Yu Fu^{a,1}, Peng Xue^{a,1}, Ning Li^{b,1}, Peng Zhao^b, Zhuodong Xu^b, Huizhong Ji^a, Zhili Zhang^a, Wentao Cui^{a,*}, Enqing Dong^{a,*}

^a School of Mechanical, Electrical & Information Engineering, Shandong University, Weihai 264209, China

^b Department of Radiology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan 250021, China

ARTICLE INFO

Article history:

Received 10 December 2020

Accepted 24 August 2021

Keywords:

Multiple pathological types

Pulmonary nodules

Multi-resolution deep learning

Machine learning

Multimodal information fusion

ABSTRACT

Background and Objective: : Current researches on pulmonary nodules mainly focused on the binary-classification of benign and malignant pulmonary nodules. However, in clinical applications, it is not enough to judge whether pulmonary nodules are benign or malignant. In this paper, we proposed a fusion model based on the Lung Information Dataset Containing 3D CT Images and Serum Biomarkers (LIDC-CISB) we constructed to accurately diagnose the types of pulmonary nodules in squamous cell carcinoma, adenocarcinoma, inflammation and other benign diseases.

Methods: : Using single modal information of lung 3D CT images and single modal information of Lung Tumor Biomarkers (LTBs) in LIDC-CISB, a Multi-resolution 3D Multi-classification deep learning model (Mr-Mc) and a Multi-Layer Perceptron machine learning model (MLP) were constructed for diagnosing multiple pathological types of pulmonary nodules, respectively. To comprehensively use the double modal information of CT images and LTBs, we used transfer learning to fuse Mr-Mc and MLP, and constructed a multimodal information fusion model that could classify multiple pathological types of benign and malignant pulmonary nodules.

Results: : Experiments showed that the constructed Mr-Mc model can achieve an average accuracy of 0.805 and MLP model can achieve an average accuracy of 0.887. The fusion model was verified on a dataset containing 64 samples, and achieved an average accuracy of 0.906.

Conclusions: : This is the first study to simultaneously use CT images and LTBs to diagnose multiple pathological types of benign and malignant pulmonary nodules, and experiments showed that our research was more advanced and more suitable for practical clinical applications.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

Lung cancer is the malignant tumor with the highest morbidity and mortality rate among all cancers in the world, the global cancer statistics in 2018 showed that there were nearly 2.1 million new cases of lung cancer worldwide, accounting for 11.6% of all new cancer cases, and 1.8 million deaths, accounting for 18.4% of all cancer deaths [1]. Researches have indicated that if lung cancer patients can be diagnosed and treated in time at the early stage, the 5-year survival rate can reach 16%–49% [2]. In the early stage of lung cancer, patients are very possible to have tiny nodules in the lungs, therefore, accurate diagnosis of the specific pathological

types of these pulmonary nodules is very important for the detection of lung cancer.

There are many pathological types of pulmonary nodules, benign types include inflammatory pseudotumor, hamartoma, tuberculosis, localized fibrosis, localized chronic inflammation and so on; malignant types include squamous cell carcinoma, adenocarcinoma, small cell carcinoma, etc. The current “gold standard” for the diagnosis of pulmonary nodules is pathological examination of the corresponding tissue obtained by surgery or percutaneous lung biopsy, but this invasive diagnostic technique will also bring significant risks to patients, therefore, how to avoid excessive treatment of benign pulmonary nodules has become an important issue attracting more and more attention from researchers [3].

In clinical, the preferred method of non-invasive examinations is CT imaging diagnosis. Another common non-invasive examination method in clinical practice is to check the content of Lung Tumor Biomarkers (LTBs) in serum. With the continuous development of molecular biology and immunology, the detection of LTBs

* Corresponding authors.

E-mail addresses: wentaocui@sdu.edu.cn (W. Cui), enqdong@sdu.edu.cn, 16601103340@163.com (E. Dong).

¹ Co-first authors

has been widely used in the early diagnosis of lung cancer [4]. National Academy of Clinical Biochemistry (NACB) has published that carcinoembryonic antigen (CEA), cytokeratin 19 fragment (Cyfra21-1), squamous cell carcinoma antigen (SCCA), neurospecific enolase (NSE) and pro-gastrin-releasing peptide (ProGRP) play an important role on the early diagnosis of benign and malignant pulmonary nodules. Nowadays, some researches [5,6] manifested that LTBs such as carbohydrate antigen 125 (CA125), carbohydrate antigen 724 (CA724), alpha-fetoprotein (AFP), carbohydrate antigen 50 (CA50) and carbohydrate antigen 199 (CA199) can also reveal the pathology of lung tumors. In addition, researchers have also discovered many new LTBs which can effectively help medical workers to diagnose pulmonary nodules, such as c-reactive protein (CRP) mentioned in the literature [7,8], human epididymal protein 4 (HE4) mentioned in the literature [9,10], etc.

The rapid development of modern medical technology has made more and more data obtained for diagnosing pulmonary nodules. How to use artificial intelligence technology to fuse multimodal information to construct a pulmonary nodules prediction model is a challenging research. Most of existing researches only used single modal information to distinguish benign and malignant nodules, and do not fully fuse LTBs and CT images. Taking this situation into account, we constructed a Lung Information Dataset Containing 3D CT Images and Serum Biomarkers (LIDC-CISB) based on the clear clinical pathological information of pulmonary nodules. On the basis of LIDC-CISB, a fusion model was constructed through deep learning and traditional machine learning methods. In summary, our contributions are as follows:

- 1) To be able to diagnose multiple pathological types of pulmonary nodules, we cooperated with multiple hospitals to construct a new dataset (LIDC-CISB) containing multimodal information of pulmonary nodules marked with clear pathological examination results of pulmonary nodules.
- 2) Based on the single modal 3D CT images and the single modal LTBs in LIDC-CISB, a Multi-resolution 3D Multi-classification deep learning model (Mr-Mc) and a Multi-Layer Perceptron machine learning model (MLP) were constructed to diagnose multiple pathological types of pulmonary nodules, respectively.
- 3) Through transfer learning, we fused Mr-Mc and MLP to construct a fusion model that can use 3D CT images and LTBs to diagnose multiple pathological types of pulmonary nodules. Experiments show that the fusion model can obtain an accuracy (Acc) of 0.906 in the classification of multiple pathological types of pulmonary nodules.

Noted that the frequent abbreviations in this paper and corresponding explanations were shown in [Appendix A](#).

2. Related works

Nowadays, a large number of studies have proposed a variety of methods for diagnosing benign and malignant pulmonary nodules [11–26]. Current researches on pulmonary nodules are mainly have three directions: 1) utilizing CT images to determine benign and malignant pulmonary nodules in binary-classification studies [11–21]. 2) utilizing LTBs to identify benign and malignant pulmonary nodules by machine learning methods [22–28]. 3) integrating simple CT imaging information (such as burr sign, diameter, etc.) and LTBs information to intelligently classify pathological types of pulmonary nodules [29,30].

2.1. Binary-classification methods for diagnosing benign and malignant pulmonary nodules based on CT images

In recent years, machine learning algorithms have been increasingly used in the field of medical image analysis [31–33], and

early studies on the classification of pulmonary nodules based on lung CT images mainly relied on artificially extracted features. Han et al. [11] used Haralick texture features, local binary model features, and Gabor features to build a support vector machine (SVM) for the binary-classification of benign and malignant pulmonary nodules, and obtained an AUC of 0.927 on the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) dataset [34]. Jacobs et al. [12] used the texture, shape and background characteristics of pulmonary nodules to classify subsolid and solid pulmonary nodules, and obtained 80% sensitivity. Sergeeva et al. [20] used random forest and SVM methods to construct a binary-classification model for diagnosing benign and malignant pulmonary nodules, and obtained an Acc of 81.30% on LIDC-IDRI dataset.

The above researches used traditional machine learning methods and limited artificially extracted features to construct the model, but the hidden features and 3D information of pulmonary nodules were missing in feature extraction process. Deep learning methods break the limitation of manual feature extraction, and use multi-layer structure to automatically extract features for classification, therefore, it is widely used in the field of medical image processing [35–37]. Shen et al. [14] used the multi-scale convolutional neural networks with an alternately stacked structure to extract discriminative features from CT data, and obtained an Acc of 86.84 on the subset of LIDC-IDRI. After that, Shen et al. [15] replaced max-pooling with multi-crop pooling operation in the original network to obtain multi-scale features to diagnose benign and malignant pulmonary nodules, and achieved an Acc of 87.14%. Zhu et al. [17] developed an end-to-end pulmonary nodule detection and classification system, which used 3D Faster R-CNN as detection network and 3D dual path network (3DPN) as classification network, they finally obtained an Acc of 90.44% on the subset of LIDC-IDRI dataset. However, these studies have always been limited to the binary-classification of benign and malignant pulmonary nodules, and did not discuss the classification of multiple pathological types.

2.2. Binary-classification methods for diagnosing benign and malignant pulmonary nodules based on LTBs

In order to meet the clinical need for accurate diagnosis of benign and malignant pulmonary nodules, many researchers have begun to use LTBs to construct models for the classification of benign and malignant pulmonary nodules. Jiang et al. [24] used thymidine kinase 1 (TK1), CEA, Cyfra21-1 and NSE to diagnose benign and malignant pulmonary nodules through logistic regression method and obtained an AUC of 0.946. Wang et al. [25] analyzed the effects of four LTBs (CEA, NSE, Cyfra21-1 and haptoglobin) in identifying benign and malignant pulmonary nodules, and obtained a sensitivity of 0.850 and a specificity of 0.857.

In addition, Wang et al. [25] also analyzed the classification sensitivity of different markers to different pathological types of pulmonary nodules, and found that Cyfra21-1 was more sensitive to the detection of squamous cell carcinoma, CEA was more sensitive to the detection of adenocarcinoma, and NSE was more sensitive to the detection of small cell carcinoma. However, most classification methods of benign and malignant pulmonary nodules utilizing LTBs information used statistical analysis, and the types of LTBs used in their models are not uniform. Researchers analyzed the diagnostic value of certain types of LTBs according to their own research needs. At the same time, these studies mostly focused on the classification of benign and malignant pulmonary nodules, and did not discuss the classification of multiple pathological types of pulmonary nodules.

2.3. Methods of fusing multi-modal information to diagnose pulmonary nodules

2.3.1. Binary-classification methods for diagnosing benign and malignant pulmonary nodules

With the continuous research on the classification of benign and malignant pulmonary nodules, studies have shown that if the characteristics of LTBDs and clinical imaging features are better combined to establish a model, the accuracy of diagnosing benign and malignant pulmonary nodules will be higher. Gong et al. [29] used 78 quantitative imaging features extracted from CT images and 5 LTBDs features (SCCA, CEA, Cyfra21-1, CA153 and CA199) to construct two SVM models to predict benign and malignant pulmonary nodules. After that, a fusion model was constructed by weighting these two SVM classifiers and results showed that the AUC of the fusion model was 0.85, which was higher than that of the two single SVM models.

2.3.2. Multi-classification methods for diagnosing multiple pathological types of pulmonary nodules

Zhao et al. [30] used 8 kinds of LTBDs information (CEA, NSE, Cyfra21-1, CA125, CA199, CA724, SCC, CA153) and 9 kinds of CT image features (size, location, boundary, spicule sign, lobulation, pleural indentation, vessel convergence sign, lymphadenectasis and cavity sign) to establish a BP neural network model and an SVM model to identify inflammation, tuberculosis and tuberculoma, inflammatory pseudotumor, tumor differentiation or cancer, respectively. All these 5 pathological types of pulmonary nodules are confirmed by pathological results. A 40-case test set was used to estimate the effect of BP and SVM model, as results, BP model achieved an Acc of 0.6 and the SVM model achieved an Acc of 0.8. Although literature [30] used double modal information to study the classification of multiple pathological types of pulmonary nodules, it did not merge double modal information, and did not make full use of the 3D CT image information. How to fuse the complete 3D CT image information and standard LTBDs information together to diagnose multiple pathological types of pulmonary nodules is the focus of the next research.

3. LIDC-CISB dataset

Current researches on CT imaging of pulmonary nodules mainly rely on the LIDC-IDRI dataset. Limited by this dataset, scholars can only conduct binary-classification studies to determine benign and malignant pulmonary nodules, moreover, this dataset does not have corresponding LTBDs. With regard to this, we cooperated with Shandong Provincial Hospital Affiliated to Shandong First Medical University to construct a Lung Information Dataset Containing 3D CT Images and Serum Biomarkers (LIDC-CISB) with biopsy-based cytological analysis as annotation. Based on LIDC-CISB, we can conduct multi-classification studies of multiple pathological types of pulmonary nodules. In addition, there are CT images and LTBDs data of the same patient in the same period in LIDC-CISB, using these data, we can construct a model that can combine CT images and LTBDs multimodal information to determine multiple pathological types of pulmonary nodules.

The data in LIDC-CISB were collected through clinical studies from January 2017 to December 2019. LIDC-CISB contains a Lung CT Image Dataset (LCID) and a Lung Tumor Biomarkers Dataset (LTBD). The pathological types of pulmonary nodules in LIDC-CISB include inflammation (Inf), Hamartoma (Ham), squamous cell carcinoma (SCC), adenocarcinoma (Ade), small cell carcinoma (SCC*) and other benign diseases (OBD), the detailed data distribution is shown in Table 1.

Table 1

Data distribution of LIDC-CISB.

Datasets	Inf	Ham	SCC	Ade	SCC*	OBD	Total
LCID	88	18	46	765	21	77	1015
LTBD	40	8	26	381	13	35	503

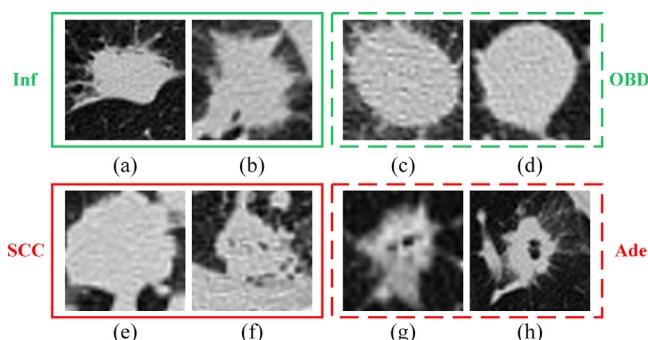


Fig. 1. CT images of pulmonary nodules. The samples surrounded by green bounding boxes are benign, and those surrounded by red bounding boxes are malignant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.1. Introduction of LCID

In order to construct the high-quality LCID, the following principles should be followed in the process of data collection: first, the locations and sizes of pulmonary nodules must be confirmed by 3 independent radiologists for three rounds, and the pathological results are obtained by examining the frozen section of the patient's lesion after surgery or percutaneous lung biopsy under CT guidance; second, whether the data is used depends on the image quality and pathological results; finally, the pulmonary nodules that meet above requirements must belong to primary lung disease, and the patient has not received radiotherapy or chemotherapy before surgery.

The thickness of CT images in LCID is 1 mm, and the diameter of pulmonary nodules is between 3mm 3 cm. All CT images were obtained by dual source multi-slice spiral CT scanner (SOMATOM Definition AS, SOMATOM Definition Flash CT etc.), the scanning parameters were as follows: 120KVP, pitch was 1 mm, slice thickness was 1 mm, pixel spacing was 0.77 mm, resolution was 512×512 , stored in DICOM format. In our experiment, the slice thickness of all CT images was adjusted to 0.77 mm. Fig. 1 showed some examples in LCID.

Noted that the CT images in the LCID were provided by 936 patients whose ages ranged from 12 to 78, with an average age of 57. We could get more than one pulmonary nodule from CT of some patients, so, the number of pulmonary nodules is bigger than the number of patients. Among these patients, 48.50% are from male patients, and 51.50% are from female patients. All patients with malignant nodules are in the early stage of the disease and have no obvious clinical features.

3.2. Introduction of LTBD

The LTBDs commonly used to diagnose pathological types of pulmonary nodules include CEA, CA50, CA125, carbohydrate antigen 242 (CA242), CA724, Cyfra21-1, NSE, ProGRP, and SCCA. Based on clear clinical pathological information of pulmonary nodules, we constructed LTBD by collecting the data of above 9 types of LTBDs, all examinations of which were performed by clinical pathologists. There are currently 503 samples in LTBD, in which 216 samples (available LTBDs data) contain all 9 LTBDs information, including 18

cases of Inf, 10 cases of SCC, 172 cases of Ade, and 16 cases of OBD.

3.3. Multi-resolution method for extracting 3D volume data of pulmonary nodules

Before extracting the 3D volume data of pulmonary nodules, lung CT images needs to be normalized according to Eq. (1):

$$N(v) = \begin{cases} 0, & v < -1000 \text{ (HU)} \\ \frac{v+1000}{1400}, & -1000 \text{ (HU)} \leq v \leq 400 \text{ (HU)} \\ 1, & v > 400 \text{ (HU)} \end{cases} \quad (1)$$

where, v is the original value of CT images, and $N(v)$ is the normalized value, $N(v) \in [0,1]$.

In the process of extracting pulmonary nodules from lung CT images, the size of the pixel windows used to extract pulmonary nodules has a greater impact on the processing of the target [16,38,39]. At present, researchers mostly used multi-scale methods to extract pulmonary nodules from lung CT images [13–19]. However, the multi-scale method is less effective in some cases, here, a novel multi-resolution method was proposed to extract pulmonary nodules. The multi-resolution method is to specify several pixel windows of different sizes based on the statistics of the pulmonary nodules' diameter, and then apply the corresponding pixel windows to extract 3D volume data according to the actual size of the pulmonary nodule. Because the diameter of pulmonary nodules in the constructed LCID is between 3 mm – 3 cm and the thickness of reconstruction slice is 0.77 mm, we chose three sizes as low resolution ($20 \times 20 \times 20$), medium resolution ($30 \times 30 \times 30$), and high resolution ($45 \times 45 \times 45$) in the selection of pixel window. So, the nodules extracted through these three types of pixel windows are all cubic volumes.

During extracting pulmonary nodules, which type of pixel window was selected depends on the diameter D of the pulmonary nodule, and the specific correspondence was as follows: when $3 \text{ mm} \leq D < 14 \text{ mm}$, selecting a low-resolution pixel window; when $14 \text{ mm} \leq D < 21 \text{ mm}$, selecting a medium resolution pixel window; when $21 \text{ mm} \leq D \leq 30 \text{ mm}$, selecting a high-resolution pixel window. For larger pulmonary nodules, high-resolution pixel windows are selected to extract pulmonary nodules, and then low- and medium-resolution pulmonary nodules are obtained by down-sampling; for medium nodules, medium-resolution pixel windows are selected to extract pulmonary nodules, and then high-resolution data are obtained by B-spline interpolation algorithm, and low-resolution data are obtained by down-sampling; for small nodules, low-resolution pixel windows are selected to extract pulmonary nodules, and then high- and medium-resolution pulmonary nodules are obtained by B-spline interpolation algorithm. Fig. 2 is a schematic diagram of the multi-resolution method for extracting pulmonary nodules.

This multi-resolution method can not only ensure the integrity of pulmonary nodule extraction, but also avoid extracting too much background information.

4. Methods

4.1. The overall workflow

In this paper, the overall research framework mainly has the following three parts: the construction of Mr-Mc model based on LCID dataset, the construction of MLP model based on LTBD dataset, the deep fusion of Mr-Mc model and MLP model. The purpose of our research is to build a model that can incorporate multi-modal information to accurately diagnose multiple pathological types of lung nodules. Fig. 3 showed the overall workflow of our work.

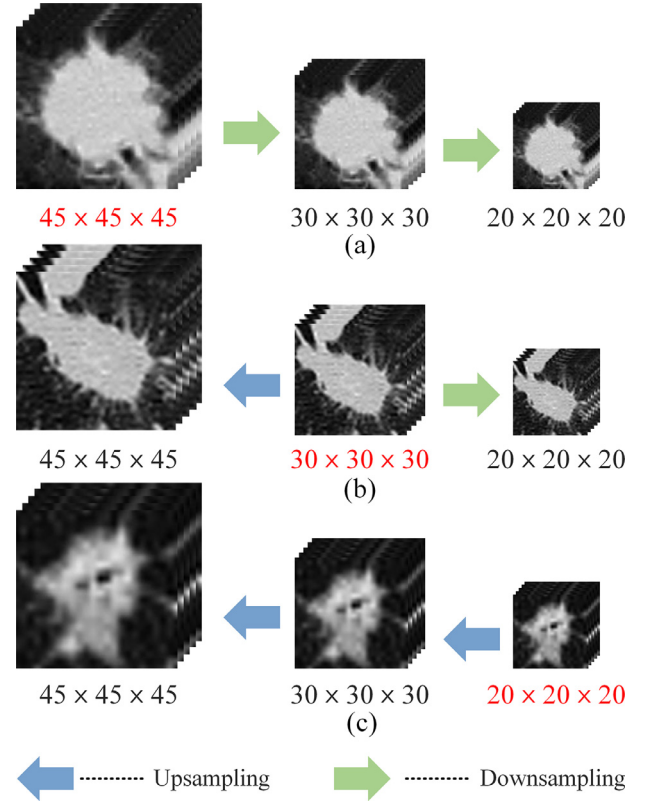


Fig. 2. Multi-resolution method to extract pulmonary nodules. The red numbers represent the actual sizes of pulmonary nodules, the green arrows represent down-sampling, and the blue arrows represent up-sampling. Fig. 2(a) shows the process of extracting large pulmonary nodules, Fig. 2(b) shows the process of extracting medium pulmonary nodules, Fig. 2(c) shows the process of extracting small pulmonary nodules. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.2. Multi-resolution 3D multi-classification deep learning model

Hinton et al. [40] proposed deep learning in 2006, which combines low-level features of the network to form high-level abstract features to learn the relationships existing within data. The emergence of convolutional neural networks (CNN) has promoted the application of deep learning in the field of image processing. A convolutional layer can be described as:

$$\mathbf{C} = g(\omega \otimes \mathbf{u} + b) \quad (2)$$

where \mathbf{C} denotes the output of the convolution layer, \mathbf{u} denotes the input, ω denotes the convolutional kernel, " \otimes " represents convolution operation, b denotes the offset and $g(\bullet)$ denotes the activation function.

Since dual path network (DPN) [41] has the advantages of both residual network (ResNet) [42] and densely connected network (DenseNet) [43], it has achieved excellent results in many classification tasks. So, we used 3D DPN (3DPN) units to constructed a Multi-resolution 3D Multi-classification deep learning model (Mr-Mc), which was shown in Fig. 4.

The Mr-Mc model is mainly composed of a low-resolution network (Archi-Low resolution with input size of $20 \times 20 \times 20$), a medium-resolution network (Archi-Med resolution with input size of $30 \times 30 \times 30$) and a high-resolution network (Archi-High resolution with input size of $45 \times 45 \times 45$). Each single resolution network is formed by stacking 3DPN units, the detailed structure of Mr-Mc was displayed in Table 2.

Taking the high-resolution network (Archi-High resolution) in Fig. 4 as an example to illustrate the process of the calculation,

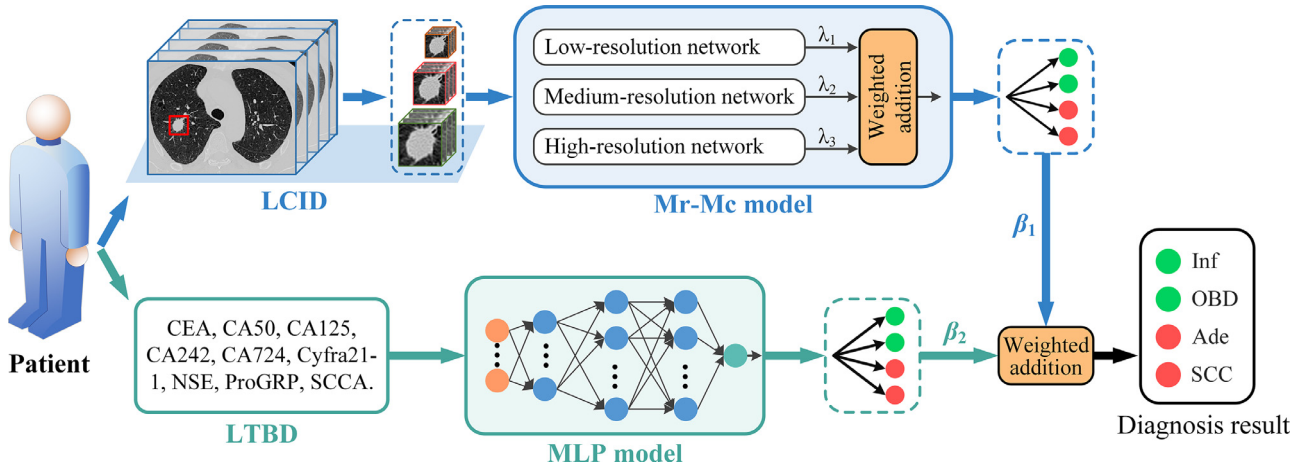


Fig. 3. The overall workflow of our work.

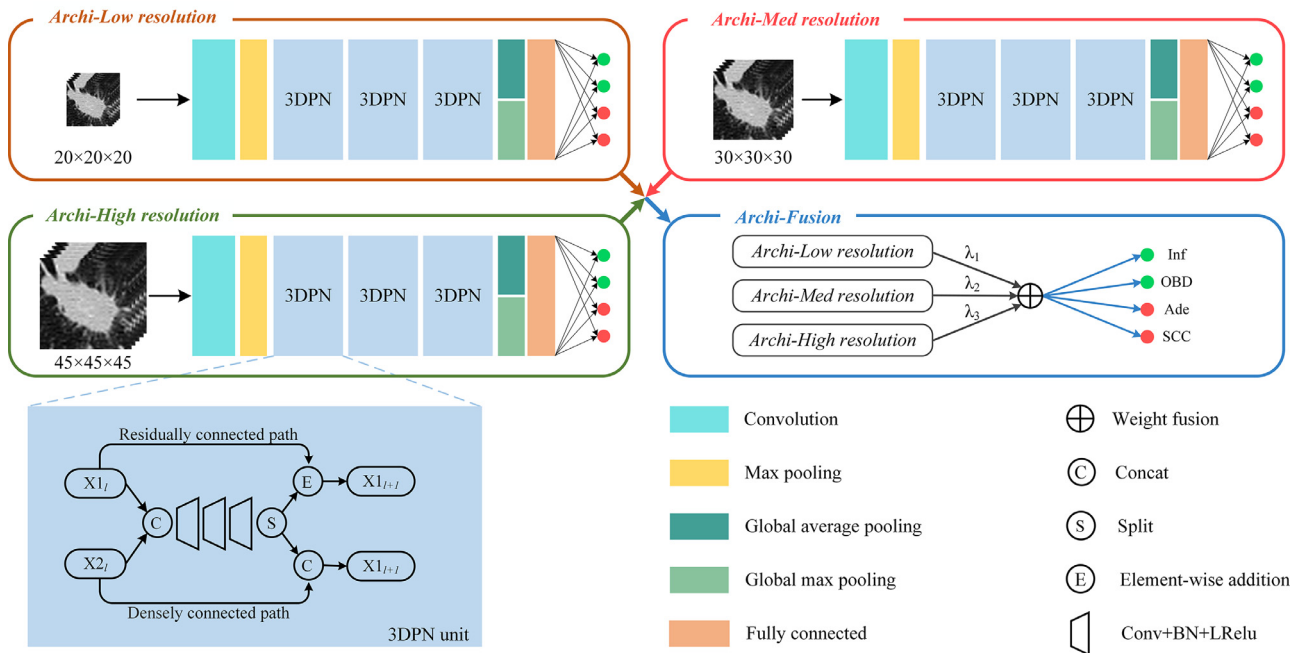


Fig. 4. The structure of the Mr-Mc model.

where $45 \times 45 \times 45$ represents the size of the input. The pulmonary nodules are extracted features by three 3DPN units after a convolution layer and a max pooling layer, and then after performing global average pooling (GAP) and global max pooling (GMP) on the features, we take average of these two outputs and pass it through a softmax classifier to get the output of single-resolution model. At last, the outputs of three single-resolution networks are fused by weight coefficients $\lambda_1, \lambda_2, \lambda_3$ to obtain predicted probability of Mr-Mc. In Mr-Mc, except for last layer of the fully connected network that uses softmax activation function, all other layers of Mr-Mc use LReLU activation function. These two activation functions are defined as follows:

$$f_{LReLU}(\xi) = \max(\xi, \alpha\xi) \quad (3)$$

$$f_{softmax}(\xi_i) = \frac{\exp(\xi_i)}{\sum_{j=1}^n \exp(\xi_j)} \quad (4)$$

where α is a positive number less than 1, when $\alpha = 0$, LReLU becomes to ReLU. n represents the number of categories. Mr-Mc is

trained by minimizing the cross-entropy loss defined as:

$$L = -\frac{1}{m} \sum_{j=1}^m \sum_{i=1}^n \mathbf{l}_{ji} \cdot \log \mathbf{p}_{ji}(\mathbf{l}_{ji}; \mathbf{x}_j; \theta) \quad (5)$$

where m is the number of samples in per mini-batch, \mathbf{l}_j denotes the class label (one-hot) of image \mathbf{x}_j , \mathbf{p}_j denotes the class prediction of image \mathbf{x}_j , θ denotes all parameters of Mr-Mc.

4.3. Multilayer perceptron model for multiple classification of pulmonary nodules

Common machine learning algorithms include logistic regression (LR), decision trees, support vector machines (SVM), and multilayer perceptron (MLP), etc. When constructing a multi-classification model for diagnosing multiple pathological types of pulmonary nodules based on LTBS, the multi-classification model constructed through LR and SVM will lead to an excessive number of classifiers, so we utilized the MLP model to complete the task. The specific structure of MLP was shown in Fig. 5, the first layer is input layer; the middle three layers are hidden layers with ReLU

Table 2
The detailed structure of Mr-Mc.

	Layers	Input	Conv	MP	3DPN-C	3DPN-C	3DPN-C	GAP&GMP	FC	Output
Low resolution	Kernel	—	$5 \times 5 \times 5$	—	$1 \times 1 \times 1$	$1 \times 1 \times 1$	$1 \times 1 \times 1$	—	—	—
					$[3 \times 3 \times 3] \times 2$	$[3 \times 3 \times 3] \times 1$	$[3 \times 3 \times 3] \times 1$			
					$1 \times 1 \times 1$	$1 \times 1 \times 1$	$1 \times 1 \times 1$			
Med resolution	Stride Shape of features	$20 \times 20 \times 20$ @1	$2 \times 2 \times 2$ $10 \times 10 \times 10$ @64	$2 \times 2 \times 2$ $5 \times 5 \times 5$ @64	$1 \times 1 \times 1$	$2 \times 2 \times 2$	$2 \times 2 \times 2$	560	256	4
					$5 \times 5 \times 5$	$3 \times 3 \times 3$	$2 \times 2 \times 2$			
					@192	@304	@560			
High resolution	Kernel	—	$5 \times 5 \times 5$	—	$1 \times 1 \times 1$	$1 \times 1 \times 1$	$1 \times 1 \times 1$	—	—	—
					$[3 \times 3 \times 3] \times 2$	$[3 \times 3 \times 3] \times 1$	$[3 \times 3 \times 3] \times 1$			
					$1 \times 1 \times 1$	$1 \times 1 \times 1$	$1 \times 1 \times 1$			
High resolution	Stride Shape of features	$30 \times 30 \times 30$ @1	$2 \times 2 \times 2$ $15 \times 15 \times 15$ @64	$2 \times 2 \times 2$ $8 \times 8 \times 8$ @64	$1 \times 1 \times 1$	$2 \times 2 \times 2$	$2 \times 2 \times 2$	560	256	4
					$8 \times 8 \times 8$	$4 \times 4 \times 4$	$2 \times 2 \times 2$			
					@192	@304	@560			
High resolution	Kernel	—	$5 \times 5 \times 5$	—	$1 \times 1 \times 1$	$1 \times 1 \times 1$	$1 \times 1 \times 1$	—	—	—
					$[3 \times 3 \times 3] \times 2$	$[3 \times 3 \times 3] \times 1$	$[3 \times 3 \times 3] \times 1$			
					$1 \times 1 \times 1$	$1 \times 1 \times 1$	$1 \times 1 \times 1$			
High resolution	Stride Shape of features	$45 \times 45 \times 45$ @1	$2 \times 2 \times 2$ $23 \times 23 \times 23$ @64	$2 \times 2 \times 2$ $12 \times 12 \times 12$ @64	$1 \times 1 \times 1$	$2 \times 2 \times 2$	$2 \times 2 \times 2$	560	256	4
					$12 \times 12 \times 12$	$6 \times 6 \times 6$	$3 \times 3 \times 3$			
					@192	@304	@560			

Conv represents the convolution operation, MP represents the maximum pooling, 3DPN-C represents the convolution operation in the 3DPN unit, and FC represents the fully connected layer.

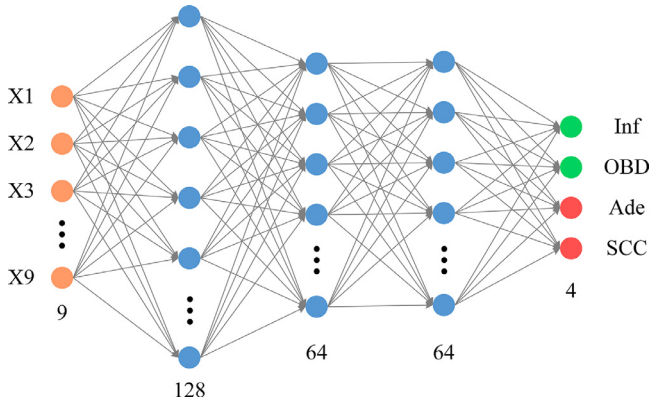


Fig. 5. The structure of MLP. X1-X9 represent 9 kinds of LTBs information, and the numbers below each layer represent the number of neurons in that layer. Repeated experiments for many times, we finally determined that the number of neurons in the first layer of the hidden layer was 128, and the number of neurons in the second and third layers was 64.

activation function and Dropout (drop rate = 0.75); the last layer is output layer with softmax activation function.

4.4. Fusion model for diagnosing multiple pathological types of pulmonary nodules

LIDC-CISB has cases that contain both CT image data and LTBs data, using the information of these cases, we can fuse lung CT images and LTBs to construct a fusion model for the diagnosis of multiple pathological types of benign and malignant pulmonary nodules. Due to the small number of cases containing both CT image data and LTBs in LIDC-CISB, it is difficult to meet the requirements of sample number for retraining the fusion model. Therefore, we used transfer learning [44] to build a fusion model by fusing the trained Mr-Mc and MLP models. First, loading the weights of the Mr-Mc model trained on LCID and the MLP model trained on LTBD into the fusion model, and then the weight coefficients at the tail end of the fusion model were determined by grid search. The structure of the fusion model was shown in Fig. 6.

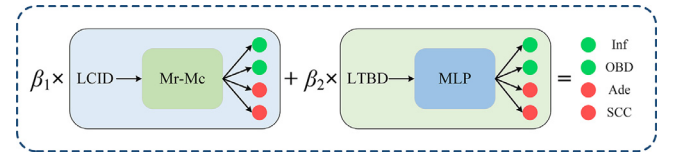


Fig. 6. The structure of the fusion model.

The output \mathbf{y} of the fusion model is obtained by weighting the outputs of Mr-Mc and MLP according to Eq. (6).

$$\mathbf{y} = P(\mathbf{y}_i = k\mathbf{d}_i) = \sum_{n=1}^2 \beta_n \mathbf{p}_n \quad (6)$$

where \mathbf{d}_i represents the input of the i th sample, \mathbf{y} represents the predicted result of the fusion model, k represents the number of categories, β_n represents the weight coefficient, and \mathbf{p}_n represents the output of Mr-Mc and MLP. In the case of a small amount of data, a fusion model with considerable performance can be constructed through transfer learning.

5. Experiments and results

In this paper, we constructed three models for diagnosing multiple pathological types of pulmonary nodules based on the LIDC-CISB dataset. Due to the limited amount of data in LIDC-CISB, we only studied four pathological types (Inf, SCC, Ade and OBD) of pulmonary nodules.

5.1. Evaluation metrics

To evaluate these constructed models, we used the accuracy (Acc), area under the receiver operating characteristic curve (AUC), sensitivity (Se) and precision (Pr) as evaluation metrics, which were defined as:

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$AUC = \int_0^1 t_{pr}(f_{pr})df_{pr} = P(\sigma_p > \sigma_n) \quad (8)$$

Table 3
Distribution of data in LCID.

Dataset Types	Inf	SCC	Ade	OBD	Total
Num. of patients	69	44	429	59	601
Male/female	39/30	35/9	181/248	38/21	293/308
Original nodules	88	46	461	76	671
Expanded training set	3465(55)	2610(29)	2232(248)	3150(50)	11,457(382)
Expanded validation set	98(14)	80(8)	62(62)	91(13)	331(97)
Test set	19	9	151	13	192

$$Se = \frac{TP}{TP + FN} \quad (9)$$

$$Pr = \frac{TP}{TP + FP} \quad (10)$$

where TP , TN , FP , FN represent the number of true positive, true negative, false positive, false negative, respectively, t_{pr} is the true positive rate, f_{pr} is the false positive rate, σ_p and σ_n are the confidence scores for a positive and negative instance, respectively.

5.2. Experimental analysis of Mr-Mc model

5.2.1. Data preprocessing of LCID

From Table 1 we can see that the LCID dataset was a long-tailed dataset, and Ade accounts for the majority. Unless otherwise specified, the following LCID dataset only contains four types of Inf, Ade, SCC and OBD. In the experiment, in order to keep the data distribution in the experiment consistent with the actual distribution of pulmonary nodule types, the test set was constructed by randomly selecting 20% of each type samples in the original LCID dataset. In the remaining samples after extracting the test set samples in the original LCID dataset, we first under-sampled the Ade samples, then the under-sampled Ade samples and other types of pulmonary nodule samples constituted the training set and the validation set (the ratio of each type of samples in the training set and validation set was 3:1). Next, we will introduce in detail the reason and specific method for under-sampling of Ade samples. Since Ade occupies the majority in the LCID dataset, if all the Ade samples are used for training, the trained model will be biased to predict the input data as Ade, which may greatly reduce the practicality of the model. Under-sampling the categories with a large number of samples is an effective method to solve the problem of data imbalance. In order to be able to select representative Ade samples, the Boosting algorithm was used to under-sample Ade samples while ensuring that Ade can be effectively distinguished from other types of pulmonary nodules. Finally, a total of 310 Ade samples with high information entropy were selected from 614 (765–151) samples, where 765 represents the total number of Ade samples we collected and 151 represents the number of Ade samples in the test set.

In the expansion of the data set, we expanded the data in the categories with fewer samples (Inf, SCC, OBD) in the training set and validation set offline. First, we expanded the data belonging to Inf and OBD in the training set and validation set by rotating 3D volume data [90°, 270°] on X-axis, Y-axis and Z-axis; second, we expanded the data belonging to SCC in the training set and validation set by rotating 3D volume data [90°, 180°, 270°] on X-axis, Y-axis and Z-axis. During model training, we also expanded the data in the training set online. In the process of online data expansion, the data are first randomly shifted by 2 pixels in the X, Y, and Z directions; then a random shear is performed in X, Y, and Z directions. The test set contains 192 unexpanded real data. Table 3 showed the distribution of four pathological types in the expanded dataset used in the experiment.

Table 4
Comparison of experimental results between single-resolution models and Mr-Mc model.

Model type	Acc	AUC
Archi-Low resolution	0.735	0.755
Archi-Med resolution	0.771	0.841
Archi-High resolution	0.798	0.850
Mr-Mc	0.810	0.876

Table 5
Results for various 3D deep learning models on the test set.

Models	Acc	AUC	Se	Pr	F1
Shen et al. [14]	0.747	0.829	0.247	0.200	0.221
Zhu et al. [17]	0.773	0.864	0.335	0.324	0.329
Xu et al. [19]	0.798	0.875	0.340	0.344	0.342
Dey et al. [21]	0.780	0.848	0.331	0.288	0.308
Mr-Mc (Proposed)	0.810	0.876	0.346	0.355	0.350

In Table 3, “Original nodules” represents the original pulmonary nodules directly extracted from lung CT without expansion. The numbers in parentheses in the expanded training set and expanded validation set, as well as the numbers in the test set represent the number of unexpanded original nodules, and the sum of the three items is equal to the number of original nodules.

5.2.2. Experiment settings

The total parameter of Mr-Mc is 11.4×10^6 , and the size is 89.7 MB. We used Keras framework to implement Mr-Mc on dual NVIDIA GTX1080ti GPUs (11GB). Mr-Mc was initialized with He initialization [45], and the SGD algorithm (momentum parameter is 0.5) was adopted as optimizer. When training Mr-Mc, we set the initial learning rate to 0.003, and reduced it to 70% every 100 epochs to accelerate convergence, and the batch size was 16. After grid search, the fusion weights were finally determined to be $\lambda_1 = 0.3$, $\lambda_2 = 0.3$, $\lambda_3 = 0.4$. It should be noted that three single resolution networks were trained simultaneously.

5.2.3. Results of Mr-Mc on LCID

After Mr-Mc was trained 600 epochs, the Acc of Inf, SCC, Ade, and OBD on the test set is 0.84, 0.82, 0.88 and 0.81. The average AUC is 0.8755. In addition, we compared Mr-Mc with low-resolution model, medium-resolution model, and high-resolution model. The parameters in these three single-resolution models are all the same as the Mr-Mc model. The Acc and AUC of these three single-resolution models and the Mr-Mc model on the test set are shown in Table 4, and the average ROC curve of four categories are shown in Fig. 7. In Table 4, the bold values represent the best value.

To further verify the proposed model, we compared Mr-Mc with some existing methods [14,17,19,21] and the average results of four categories were shown in Table 5. The bold values in Table 5 represent the best average value among the five methods.

In Table 5, we showed the average sensitivity (Se) and precision (Pr) of Inf, SCC, Ade and OBD. From Table 5 we can see that all models' performance on the sensitivity and precision metrics were

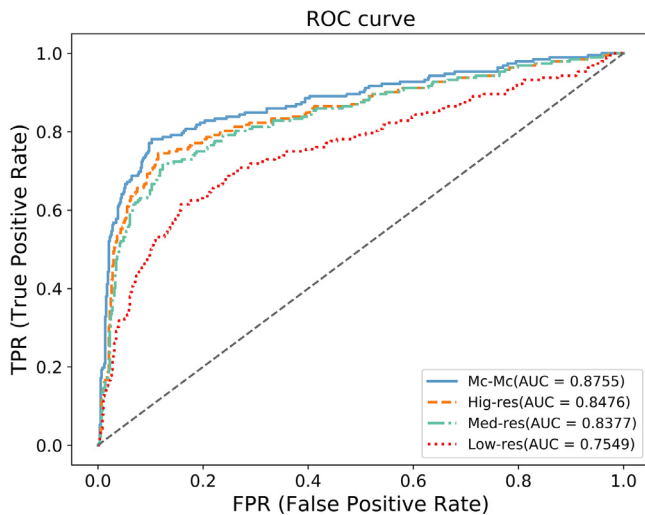


Fig. 7. ROC curves of single-resolution models and the Mr-Mc model.

Table 6
Distribution of data in LTBD.

Datasets	Inf	SCC	Ade	OBD	Total
Fold1	24(4)	20(2)	35	24(4)	103(45)
Fold2	24(4)	20(2)	35	18(3)	97(44)
Fold3	24(4)	20(2)	34	18(3)	96(43)
Fold4	18(3)	20(2)	34	18(3)	90(42)
Fold5	18(3)	20(2)	34	18(3)	90(42)
Original LTBs (total)	18	10	172	16	216
Expanded LTBs (total)	108	100	172	96	476

low and the proposed Mr-Mc model achieved the best result. Regarding the poor performance of all models on the sensitivity and precision metrics, we think this is caused by the data imbalance of the LCID dataset we constructed, and has nothing to do with models themselves. In the future, we plan to collect more data to solve this problem.

Moreover, in order to prove that Mr-Mc can effectively extract the features of pulmonary nodules with different pathological types, we used class activation mapping (CAM) [46] to visualize attention regions in feature map in Fig. 8. We showed 2D plane images of the middle layer in the 3D volume data of pulmonary nodules in Fig. 8. It should be noted that since the high-resolution network contributes most to the performance of the Mr-Mc model, we used the last convolutional layer in the high-resolution network to obtain CAMs.

5.3. Experimental analysis of MLP model

5.3.1. Data preprocessing of LTBD

Due to the small number of samples of Inf, SCC and OBD in the 216 available LTBs data, the SMOTE algorithm [47] was used to expand these three types of data. After sampling by the SMOTE algorithm, the total sample reached 476 samples, in which Inf expanded 5 times, totaling 108 samples; SCC expanded 9 times, totaling 100 samples; OBD expanded 5 times, totaling 96 samples.

5.3.2. Experimental settings

We trained and tested the model using 476 samples obtained after the original data was expanded by the SMOTE algorithm. In order to eliminate the influence of accidental factors on the experimental results, we repeated five times of five-fold cross-validation in the experiments. The data distribution of each five-fold cross-validation was detailed shown in Table 6. It should be noted that in the 5-fold data division process, we first divided the data set,

Table 7
Experimental results between combined LTBs and complete LTBs.

Different combinations of LTBs	Acc	AUC
CEA, NSE, Cyfra21-1, SCCA, ProGRP	0.662	0.86
CEA, NSE, Cyfra21-1, SCCA, ProGRP, CA125	0.649	0.91
CEA, NSE, Cyfra21-1, SCCA, ProGRP, CA724	0.683	0.90
CEA, NSE, Cyfra21-1, SCCA, ProGRP, CA50	0.798	0.94
CEA, NSE, Cyfra21-1, SCCA, ProGRP, CA242	0.695	0.90
All	0.887	0.98

and then expanded the data. The SGD algorithm was used to minimize the cross-entropy loss to update the parameters of the MLP model. The initial learning rate was 0.001 and the batch size was 16. We used a five-fold cross-validation for experiments, and the Acc and AUC of the MLP model are the average values of the five-fold cross-validation.

In Table 6, “original LTBs” represents the original LTBs we collected without expansion. The numbers in parentheses represent the number of unexpanded original LTBs.

In order to verify the performance of the MLP model, we first used the complete 9 kinds of LTBs data for experiments. Since the training set and the test set in the LTBD dataset were randomly divided, the results of each cross-validation are accidental. In addition, in order to verify that combined multiple LTBs can more effectively improve the accuracy of MLP, we also focused on the impact of combined LTBs data on the classification of multiple pathological types of pulmonary nodules.

5.3.3. Results of MLP on LTBD

Fig. 9 showed the five-fold cross-validation ROC curves, and Fig. 9(a) shows the four-type average ROC curves for each fold after five-fold cross-validation and the average ROC curve (AVE) of the five-fold cross-validation. After five-fold cross-validation, the Acc of the MLP model was 0.887 and the AUC was 0.98. Fig. 9(b) shows the average ROC curve obtained after repeating five times of five-fold cross-validation. After 5 repeated verifications, the average Acc was 0.87 and the average AUC was 0.97.

Repeated cross-validation for 5 times was performed under the condition that all the parameters of the MLP model remain unchanged. It can be seen from Fig. 9(b) that after repeated cross-validation, the AUC of the MLP model was still above 0.95, which proved that the influence of contingency factors on the model was excluded and the MLP model had high stability and reliability.

In addition, in order to analyze the differences in the diagnosis of multiple pathological types of pulmonary nodules by using the complete 9 kinds of LTBs and combined LTBs, we also conducted comparative experiments. NACB clearly stated that CEA, NSE, Cyfra21-1, SCCA, and ProGRP have a significant effect on the early diagnosis of pulmonary nodules. Therefore, in this comparative experiment, the following methods were adopted: fixing the information of five LTBs of CEA, NSE, Cyfra21-1, SCCA, and ProGRP, and then individually increasing the information of CA125, CA724, CA50 and CA242, and finally using the combined data of 6 kinds of LTBs to train 5 models (MLP-6) separately. The detailed results of Acc and AUC were shown in Table 7.

It can be seen from Table 5 that the Acc of the model built with the complete 9 kinds of LTBs information (shown as “All” in Table 5) was at least 10% higher than the Acc of any model built with combined LTBs information, indicating that combining multiple LTBs information can improve the accuracy of the model. As far as we know, this is the first study to use the complete 9 kinds of LTBs to diagnose multiple pathological types of benign and malignant pulmonary nodules.

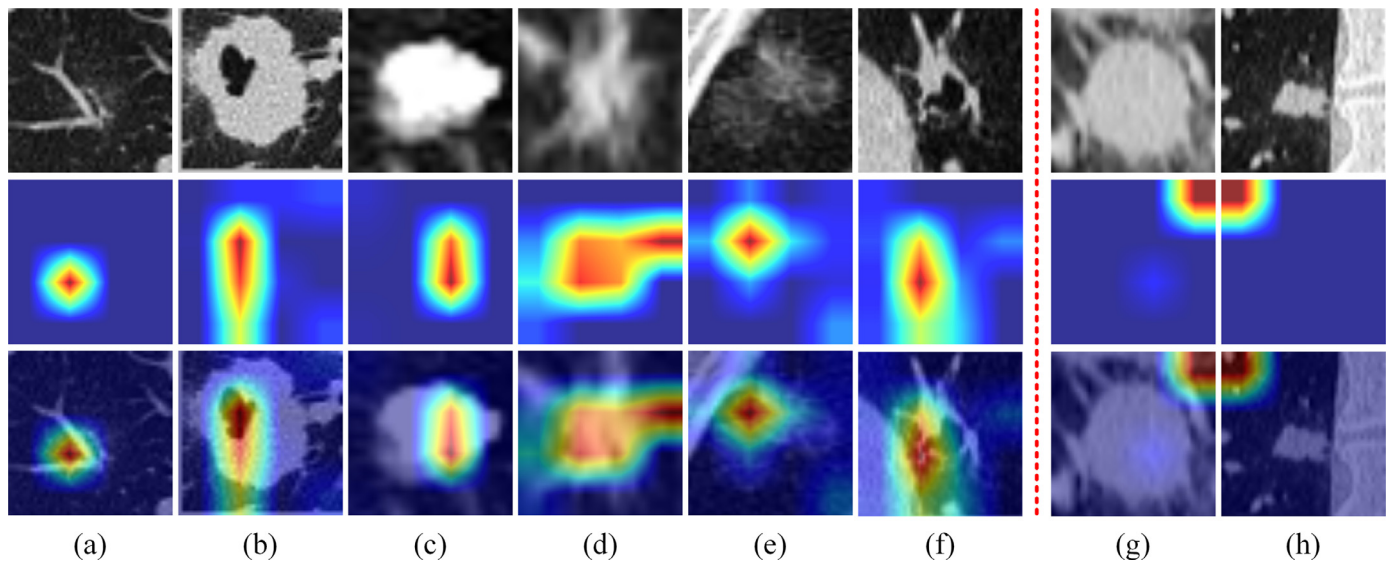


Fig. 8. Visualization of pulmonary nodules and the corresponding CAMs. Top row are original CT images of pulmonary nodules, middle row are CAMs, and bottom row are images obtained by superimposing CAMs on corresponding CT images of pulmonary nodules. In Fig. 8, we visualized the pulmonary nodules that Mr-Mc predicts correctly ((a)-(f)) and the pulmonary nodules that predict incorrectly ((g), (h)). Note that you can find CAM visualization images of all pulmonary nodules in the test set at the following website: <https://github.com/fuyu-sdu/CAM-visualization-data/blob/main/CAM.zip>. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

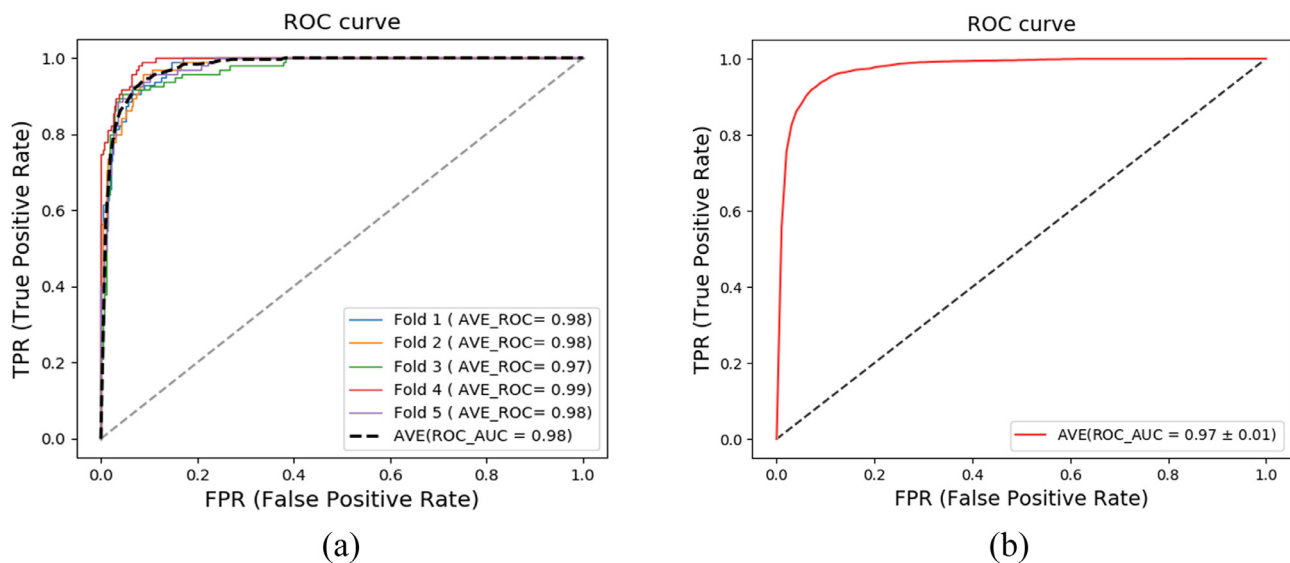


Fig. 9. ROC curves of five-fold cross validation.

5.4. Experimental analysis of fusion model

5.4.1. Data and experimental settings

The dataset used in the fusion model was constructed from the data of 64 cases with both CT images and LTBs in the test set of LCID, which included 12 cases of Inf, 6 cases of SCC, 41 cases of Ade, and 5 cases of OBD. In this experiment, we first excluded the 64 samples in the LTBD, and then used the remaining data in LTBD to construct a data set to retrain the MLP model, the detailed data distribution was shown in Table 8. The fusion model was constructed by the Mr-Mc model and the MLP model through weight fusion, and the weight coefficients are between (0,1), which are determined by grid search.

In Table 8, the numbers in parentheses represent the number of unexpanded original LTBs.

Table 8

Distribution of data used in fusion model.

Dataset types	Inf	SCC	Ade	OBD	Total
Total original LTBs	18	10	172	16	216
Expanded training set	30(5)	30(3)	99	54(9)	213(116)
Expanded validation set	6(1)	10(1)	32	12(2)	60(36)
Test set	12	6	41	5	64

5.4.2. Results of fusion model

After grid search, under different weight coefficients, the Acc and AUC corresponding to the fusion model are shown in Table 9, where “C” represents the Mr-Mc model based on CT images and “B” represents the MLP model based on LTBs.

Table 9
Results of fusion model with different weight coefficients.

Fusion models	C	0.9C+0.1B	0.8C+0.2B	0.7C+0.3B	0.6C+0.4B	0.5C+0.5B	0.4C+0.6B	0.3C+0.7B	0.2C+0.8B	0.1C+0.9B	B
Acc	0.811	0.828	0.836	0.844	0.891	0.898	0.891	0.898	0.906	0.891	0.888
AUC	0.855	0.862	0.898	0.919	0.933	0.942	0.946	0.950	0.950	0.940	0.930

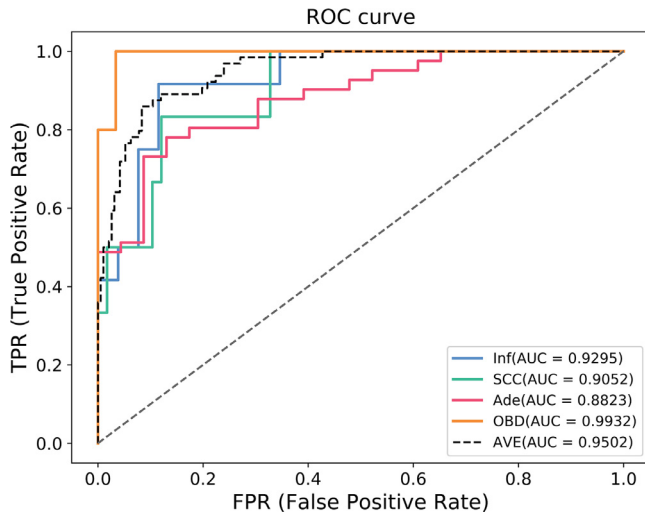


Fig. 10. ROC curves of the fusion model.

It can be seen from Table 6 that when the weight factor of the Mr-Mc model was 0.2 and the weighting factor of the MLP model was 0.8, the average Acc of the fusion model was up to 0.906, and the AUC also reached the highest value of 0.950. The classification effect of the fusion model is better than the two single-modal models of Mr-Mc and MLP, the main reason is that the fusion model not only used the Mr-Mc model to extract the complete 3D CT features of pulmonary nodules, but also used the MLP model to utilize relatively complete information of LTBs, which enriches the number of features that contribute to the classification of pathological types of pulmonary nodules. Fig. 10 showed ROC curves and average ROC curve of four categories of the fusion model (0.8L+0.2C), where “AVE” represents the average AUC off our pathological types.

Compared with the literature [30] which also carried out the multi-classification of pulmonary nodules, although the SVM model constructed in [30] achieved a higher Acc, the SVM model was not suitable for dealing with multi-classification problems. So, the author used “one-against-all” method to build 5 SVM binary-classification models, which led to a more complicated training process. The fusion model we constructed only needs to train the Mr-Mc model and the MLP model once to achieve multi-classification, and the Acc is high, so it is more suitable for clinical diagnosis of pulmonary nodules.

6. Discussion and conclusion

The accurate diagnosis of early pulmonary nodules is the key to fight lung cancer. Nevertheless, early diagnosis of pulmonary nodules in clinic requires medical workers to have rich clinical experience, which means they should make their best effort to identify specific pathological types of benign and malignant pulmonary nodules based on the morphological characteristics of pulmonary nodules and the LTBs. The continuous development of CT imaging examination has led to a sharp increase in the amount of CT imaging data, which has greatly increased the workload of medical workers. When faced with a large amount of image data, medi-

cal workers are prone to fatigue and reduce diagnostic efficiency. This urgently requires computer-aided diagnosis (CAD) to reduce the workload of medical workers so that they can focus on more difficult cases to diagnose.

In this paper, the LIDC-CISB dataset was first created by our team in order to construct a model for intelligent diagnosis of multiple pathological types of pulmonary nodules. Using LIDC-CISB, we constructed a Multi-resolution 3D Multi-class deep learning model (Mr-Mc) based on lung CT image information and a Multi-Layer Perceptron multi-classification machine learning model (MLP) based on LTBs information. In order to simultaneously use the double modal information of CT images and LTBs for early detection of pulmonary nodules, an intelligent fusion model for diagnosing multiple pathological types of benign and malignant pulmonary nodules was constructed through transfer learning.

Regarding the problem of the weight coefficients of the Mr-Mc model and the MLP model in constructing the fusion model, it is obvious that the MLP model provides most of the classification performance of the fusion model. The reason is that MLP model is a machine learning method. The advantage of Mr-Mc model is to extract complex features of pulmonary nodules. Compared with Mr-Mc model, MLP is more suitable for small sample data sets with fewer feature values. However, due to the small amount of data in Inf\SCC\OBD in the LCID data set, the Mr-Mc model cannot play an advantage in predicting these three categories, and the MLP model is just suitable for the classification of small sample sets. By comparing the results of the MLP and Mr-Mc models, we can see that the imbalance problem has less impact on the MLP model based on LTBs, and has a greater impact on the Mr-Mc model based on CT. Generally speaking, the main problems that affect the accuracy of model prediction are the small size of the data set and the imbalance of the data in each category. In order to further verify the contribution of the Mr-Mc and MLP to the fusion model, we conducted a detailed analysis of 64 samples in the test set. We found that Mr-Mc tended to predict Inf\SCC\OBD as Ade, which will lead to a high recall rate of Ade. We also found that MLP had a lower predictive effect on Ade than Mr-Mc, but it had a better predictive effect on Inf\SCC\OBD. Due to the higher prediction accuracy of Mr-Mc for Ade and the higher accuracy of MLP for Inf\SCC\OBD, the performance of the fusion model is better than that of the Mr-Mc and MLP model, which shows that the fusion of CT and LTBs can be used to better diagnose multiple pathological types of pulmonary nodules.

In this paper, we only used pulmonary nodule samples with clear pathological information, however, we also collected some samples that were not marked with pathological information in the process of collecting data. We plan to use these unlabeled pulmonary nodules through semi-supervised adversarial method [48] to improve the performance of Mr-Mc model. Besides, we will adopt the multi-view patches extraction method proposed in literature [49] to enhance the richness of input data, which we believe will further improve the four-classification performance of the Mr-Mc model. In this paper, the double modal information of lung CT images and LTBs were fused by the weighted fusion method in transfer learning. In the future work, we will explore some new fusion methods to realize the deep fusion of multimodal information. In addition, because the LIDC-CISB dataset we constructed is a small sample and unbalanced data set, the next step is to use attention modules to design small networks and design loss func-

tions that solve the problem of data imbalance to improve the diagnosis effect of multiple pathological types of pulmonary nodules. At the same time, we will collect more data to train deep learning models with better performance and be able to diagnose more pathological types of pulmonary nodules.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported in part by the [Fundamental Research Funds for the Central Universities](#) (China), in part by the [National Natural Science Foundation of China](#) under Grant 62171261, [81671848](#) and [81371635](#), in part by the Key Research and Development Project of Shandong Province under Grant [2019GGX101022](#).

Appendix A

Frequent abbreviations and corresponding explanations.

Abbreviations	Explanations	Abbreviations	Explanations
CT	Computer Tomography	CA724	Carbohydrate Antigen 724
LCID	Lung CT Image Dataset	GAP	Global Average Pooling
LTBD	Lung Tumor Biomarkers Dataset	GMP	Global Max Pooling
LTBs	Lung Tumor Biomarkers	3DPN	3D Dual Path Network
MLP	Multi-Layer Perceptron	Ade	Adenocarcinoma
CEA	Carcinoembryonic Antigen	SCC	Squamous Cell Carcinoma
Cyfra21-1	Cytokeratin 19 Fragment	Inf	Inflammation
SCCA	Squamous Cell Carcinoma Antigen	OBD	Other Benign Diseases
NSE	Neurospecificity Enolase	Acc	Accuracy
ProGRP	Pro-Gastrin-Releasing Peptide	AUC	Area under The Receiver Operating Characteristic Curve
CA50	Carbohydrate Antigen 50	Se	Sensitivity
CA125	Carbohydrate Antigen 125	Pr	Precision
CA242	Carbohydrate Antigen 242	CAM	Class Activation Mapping
LIDC-CISB	Lung Information Dataset Containing 3D CT Images and Serum Biomarkers	LIDC-IDRI	Lung Image Database Consortium and Image Database Resource Initiative
Mr-Mc	Multi-resolution 3D Multi-classification Deep Learning Model		

References

- [1] J. Ferlay, et al., Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods, *Int. J. Cancer* 144 (8) (Apr. 2019) 11953–12941.
- [2] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2020, *CA Cancer J. Clin.* 70 (1) (Jan. 2020) 7–30.
- [3] D. Sarikaya, J.J. Corso, K.A. Guru, Detection and localization of robotic tools in robot-assisted surgery videos using deep neural networks for region proposal and detection, *IEEE Trans. Med. Imaging* 36 (7) (Jul. 2017) 1542–1549.
- [4] D. Kaul, et al., Prognostic indices in stereotactic radiotherapy of brain metastases of non-small cell lung cancer, *Radiat. Oncol.* 10 (1) (Nov. 2015) 10–244.
- [5] Y. Peng, et al., Utility of multiple increased lung cancer tumor markers in treatment of patients with advanced lung adenocarcinoma, *Chin. J. Lung Cancer* 20 (10) (2017) 1999–6187.
- [6] S. Ando, et al., Optimal combination of seven tumour markers in prediction of advanced stage at first examination of patients with non-small cell lung cancer, *Anticancer Res.* 21 (4B) (Jul. 2001) 3085–3092.
- [7] K. Yonemori, et al., Development and validation of diagnostic prediction model for solitary pulmonary nodules, *Respirology* 12 (6) (Nov. 2007) 856–862.
- [8] C.V. Pecot, et al., Added value of a serum proteomic signature in the diagnostic evaluation of pulmonary nodules, *Cancer Epidemiol. Biomarkers Prev.* 21 (5) (May. 2012) 786–792.
- [9] Q.F. Tang, et al., Value of serum marker HE4 in pulmonary carcinoma diagnosis, *Int. J. Clin. Exp. Med.* 8 (10) (Oct. 2015) 19014–19021.
- [10] E.T. Korkmaz, et al., Triple test with tumor markers CYFRA 21.1, HE4, and Pro-GRP might contribute to diagnosis and subtyping of lung cancer, *Clin. Biochem.* 58 (Aug. 2018) 15–19.
- [11] F. Han, et al., Texture feature analysis for computer-aided diagnosis on pulmonary nodules, *J. Digit. Imaging* 28 (1) (Feb. 2015) 99–115.
- [12] C. Jacobs, et al., Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images, *Med. Image Anal.* 18 (2) (Feb. 2014) 374–384.
- [13] K. Suzuki, et al., Computer-aided diagnostic scheme for distinction between benign and malignant nodules in thoracic low-dose CT by use of massive training artificial neural network, *IEEE Trans. Med. Imaging* 24 (9) (Sep. 2005) 1138–1150.
- [14] W. Shen, et al., Multi-scale Convolutional Neural Networks for Pulmonary nodule Classification, in: *Information Processing in Medical Imaging*, Cham, Jun. 2015, pp. 588–599.
- [15] W. Shen, et al., Multi-crop Convolutional Neural Networks for pulmonary nodule malignancy suspiciousness classification, *Pattern Recognit.* 61 (61) (Jan. 2017) 663–673.
- [16] A.A.A. Setio, et al., Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks, *IEEE Trans. Med. Imaging* 35 (5) (May. 2016) 1160–1169.
- [17] W. Zhu, et al., DeepLung: deep 3D dual path nets for automated pulmonary nodule detection and classification, in: *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Lake Tahoe, NV, USA, Mar. 2018, pp. 673–681.
- [18] L. Liu, et al., Multi-task deep model with margin ranking loss for pulmonary nodule analysis, *IEEE Trans. Med. Imaging* 39 (3) (Aug. 2020) 718–728.
- [19] X. Xu, et al., MSCS-DeepLN: evaluating pulmonary nodule malignancy using multi-scale cost-sensitive neural networks, *Med. Image Anal.* 65 (Jul. 2020) 101772.
- [20] M. Sergeeva, et al., Classification of pulmonary nodules on computed tomography scans. evaluation of the effectiveness of application of textural features extracted using wavelet transform of image, in: *18th Conference of Open-Innovations-Association-FRUCT and Seminar on Information Security and Protection of Information Technology (ISPIT)*, St Petersburg, RUSSIA, Apr. 2016, pp. 291–299.
- [21] R. Dey, Z. Lu, Y. Hong, Diagnostic classification of pulmonary nodules using 3D neural networks, in: *Proceedings of the IEEE 15th International Symposium on Biomedical Imaging (ISBI)*, Washington, DC, USA, April 2018, pp. 774–778.
- [22] M. Zhang, et al., Establishment of a mathematic model for predicting malignancy in solitary pulmonary nodules, *J. Thorac. Dis.* 7 (10) (Oct. 2015) 1833–1841.
- [23] J. Chen, et al., Clinical value of serum tumor markers, beta 2 macroglobulin and interieukin-6 in the diagnosis of patients with a solitary pulmonary nodule, *Int. J. Clin. Exp. Med.* 11 (8) (Aug. 2018) 8509–8515.
- [24] Z.F. Jiang, M. Wang, J.L. Xu, Thymidine kinase 1 combined with CEA, CYFRA21-1 and NSE improved its diagnostic value for lung cancer, *Life Sci.* 194 (Feb. 2018) 1–6.
- [25] W. Wang, et al., Clinical utility of haptoglobin in combination with CEA, NSE and CYFRA21-1 for diagnosis of lung cancer, *Asian Pac. J. Cancer Prev.* 15 (22) (Jan. 2014) 9611–9614.
- [26] W. Wang, et al., Analysis of the discriminative methods for diagnosis of benign and malignant solitary pulmonary nodules based on serum markers, *Oncol. Res. Treat.* 37 (12) (Nov. 2014) 740–746.
- [27] W.A. Song, et al., Utility of squamous cell carcinoma antigen, carcinoembryonic antigen, Cyfra 21-1 and neuron specific enolase in lung cancer diagnosis: a prospective study from China, *Chin. Med. J.* 124 (20) (Oct. 2011) 3244–3248.
- [28] F. Guida, et al., Assessment of lung cancer risk on the basis of a biomarker panel of circulating proteins, *JAMA Oncol.* 5 (12) (Apr. 2018) 1811.
- [29] J. Gong, et al., Fusion of quantitative imaging features and serum biomarkers to improve performance of computer-aided diagnosis scheme for lung cancer: a preliminary study, *Med. Phys.* 45 (12) (Dec. 2018) 5472–5481.

- [30] Z. Zhao, et al., Establishing assistant diagnosis models of solitary pulmonary nodules based on intelligent algorithms, *Cell. Physiol. Biochem.* 35 (6) (Apr. 2015) 2463–2471.
- [31] B. Yang, L. Ying, J. Tang, Artificial neural network enhanced bayesian PET image reconstruction, *IEEE Trans. Med. Imaging* 37 (6) (Jun. 2018) 1297–1309.
- [32] J. Lyu, et al., KerNL: kernel-based nonlinear approach to parallel MRI reconstruction, *IEEE Trans. Med. Imaging* 38 (1) (Aug. 2018) 312–321.
- [33] P. Xue, E. Dong, H. Ji, Lung 4D CT image registration based on high-order markov random field, *IEEE Trans. Med. Imaging* 39 (4) (Apr. 2020) 910–921.
- [34] S.G. Armato III, et al., The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of pulmonary nodules on CT scans, *Med. Phys.* 38 (2) (Jan. 2011) 915–931.
- [35] S. Liu, et al., Hierarchical nonlocal residual networks for image quality assessment of pediatric diffusion mri with limited and noisy annotations, *IEEE Trans. Med. Imaging*, early access, doi:10.1109/TMI.2020.3002708.
- [36] Z. Tang, et al., Deep learning of imaging phenotype and genotype for predicting overall survival time of glioblastoma patients, *IEEE Trans. Med. Imaging* 39 (6) (Jan. 2020) 2100–2109.
- [37] Y. Fu, et al., Deep model with Siamese network for viable and necrotic tumor regions assessment in osteosarcoma, *Med. Phys.* 47 (10) (Oct. 2020) 4895–4905.
- [38] Q. Dou, et al., Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection, *IEEE Trans. Biomed. Eng.* 64 (7) (Jul. 2017) 1558–1567.
- [39] Z. Luo, M.A. Brubaker, M. Brudno, Size and texture-based classification of lung tumors with 3D CNNs, in: *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Santa Rosa, CA, USA, Mar. 2017, pp. 806–814.
- [40] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, *Science* 313 (5786) (Jul. 2006) 504–507.
- [41] Y. Chen, et al., “Dual path networks,” 2017, arXiv:1707.01629. [Online]. Available: <http://arxiv.org/abs/1602.07261>.
- [42] K. He, et al., Deep residual learning for image recognition, in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.
- [43] G. Huang, et al., Densely connected convolutional networks, in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2261–2269.
- [44] S.J. Pan, Q. Yang, A survey on transfer learning, *IEEE Trans. Knowl. Data Eng.* 22 (10) (Oct. 2010) 1345–1359.
- [45] K. He, et al., Delving deep into rectifiers: surpassing human-level performance on imagenet classification, in: *IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, Dec. 2015, pp. 1026–1034.
- [46] B. Zhou, et al., Learning deep features for discriminative localization, in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2921–2929.
- [47] N.V. Chawla, et al., SMOTE: synthetic minority over-sampling technique, *J. Artif. Int. Res.* 16 (1) (Jun. 2002) 321–357.
- [48] Y. Xie, J. Zhang, Y. Xia, Semi-supervised adversarial model for benign-malignant lung nodule classification on chest CT, *Med. Image Anal.* 57 (Oct. 2019) 237–248.
- [49] Y. Xie, et al., Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT, *IEEE Trans. Med. Imaging* 38 (4) (Apr. 2019) 991–1004.