Benign and Malignant Classification Model of Pulmonary Nodules Based on Residual Neural Network

Zhenzhe Lin[1], Guitang Wang[1,*], Qinshen Fu[1] and Guozhen Wang[2]

[1]School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou Guangdong, China
[2]Beijing Normal University-Hong Kong Baptist University United International College, Zhuhai Guangdong, China
*Corresponding author

Abstract—Computer-assisted diagnosis is of significance in the timely treatment of lung cancer with classifying benign and malignant pulmonary nodules. Aiming at improving the low accuracy rate of benign and malignant pulmonary nodules and reducing the misdiagnosis rate and wrong-diagnosis rate in computer-aided diagnosis system, a classification model of pulmonary nodules based on residual network was proposed. Firstly, selected some lung CT images from LIDC-IDRI as a data set, amplified the data by horizontal flipping, and then converted them into single channel images. After cropping and normalization, the data was finally divided into training set and test set (7:3), and used to train and test a residual network (ResNet-26). After training, test results represent that the model accuracy rate, sensitivity and specificity are 97.53%, 97.91% and 97.18%. By comparing various methods, the raised method performs better than others according to accuracy, sensitivity and specificity, which demonstrates that it has the ability to help doctors in diagnosis.

Keywords—deep learning; residual network; pulmonary nodules; benign and malignant classification

I. Introduction

Lung cancer is one of the leading causes of cancer deaths in humans across the world. It is evidenced that conducting detections and treatments in the early stage could improve the survival rate of patients who suffer from lung cancers [1]. Accordingly, Computed Tomography (CT) has rapidly grown to an effective approach to detecting lung cancers. However, doctors may provide misdiagnoses due to tiredness caused by high workload and subjective biases. Therefore, given that machines and algorithms would not distract and are less likely to be affected by biases, they have great effects on classification of benign and malignant pulmonary nodules. It could reduce the workload of human experts, whose time then can be spent on practical treatments. Aiden Nibali et al. [2] put forward a residual network to classify lung cancers using CT images, obtaining an accuracy rate of 89.9%. Such a relatively low accuracy rate may be caused by the shallow architecture of the network or the insufficient features. Dai et al. [3] proposed a network called DenseNet (161 layers) to classify lung nodules, using the intermediate density projection method to embed 3D information of lung nodules into a convolutional neural network (CNN), achieving an accuracy rate of 89.93%. Another recent paper obtained a higher accuracy of 92.5%, taking as input the most informative areas truncated out of the original CT images [4].

All of the applications mentioned above preprocessed the image data to small images of the lung nodules with a size of 64×64 pixels before classification. This is problematic, as they cannot extract features of the entire lung image, losing other potentially relevant information. To this end, this article raises another approach to further promote the performance of lung cancer classification technique.

II. Residual Network-based Classification

This paper proposes a residual network-based method to make a pulmonary nodules classification between benign and malignant, including image preprocessing and residual convolutional network modeling.

A. Image Preprocessing

This paper makes use of a part of the LIDC-IDRI [5, 6] dataset with a total of 10,402 images. The proportion of benign and malignant is approximately 1:1. In order to learn better feature representations from meaningful areas of images, data needs to be pre-processed before training. The pre-processing include image amplification, single channel processing, center area cropping, and batch normalization.

1) Image Amplification

Image amplification is performed in order for the model to generalize better and overcome over-fitting [7]. It simply swaps each image about the vertical central axis. Let $P(x0, y0)$ denote a pixel location in an image with a width of W, the corresponding point after flipping is $(W-x0, y0)$. Matrix expression of the transformation is given as Eq. (1); image features remain the same after flipping, as shown in Fig. 1.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & W \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} \tag{1}$$

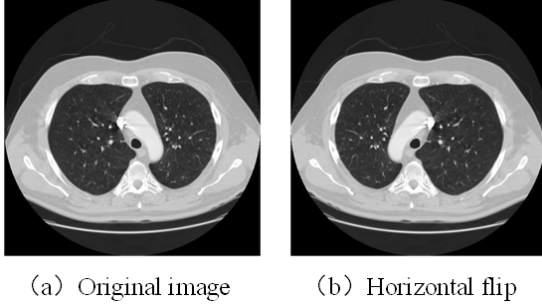(a) Original image   (b) Horizontal flip

FIGURE I. COMPARISON OF ORIGINAL IMAGE AND HORIZONTAL FLIP

2) Single Channel Processing

We also convert the three-channel (RGB) CT images into single-channel grayscale images to reduce parameters of the network by:

$$X^i \leftarrow \frac{1}{3}(X_R^i + X_G^i + X_B^i) \tag{2}$$

where i is the index of pixels; $R, G, B \in [0, 255]$ are pixel values of the RGB channels. Note that, the conversion has no impact on the original features.

3) Center Area Cropping

In this experiment, the CT image is truncated into a size of 384×384 pixels from the original 512×512 images. Such a truncation retains the most informative areas, while reduces the number of network parameters and saves a significant amount of computational costs. Images before and after cropping are shown in Figure 2.
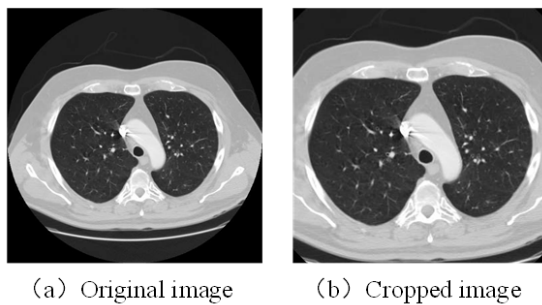

(a) Original image   (b) Cropped image

FIGURE II. COMPARISON OF ORIGINAL IMAGE AND AFTER CROPPING

4) Batch Normalization

In a neural network with many hidden layers, the data distribution would be shifted after a certain steps of computations. Ioffe et al. proposed batch normalization [8] (BN) to solve this covariance migration problem that affects the training process of deep neural networks. The algorithm is shown in Figure 3.
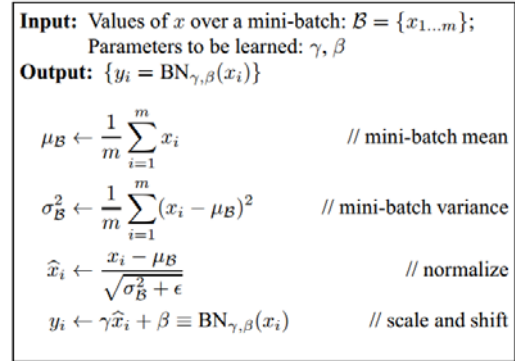

FIGURE III. ALGORITHM OF BATCH NORMALIZATION

B. Structural Design of Residual Network

1) Residual neural network

In recent years, convolutional neural networks have been progressed rapidly, with unprecedented performance achieved on many benchmark datasets. Meanwhile, many excellent CNN models have emerged in the famous ImageNet competition, such as AlexNet [9], VGGNet [10] and GoogLeNet [11], etc. However, as the number of layers increases, gradients tend to disappear or explode more easily.

Therefore, He et al. put forward a residual unit that could prevent gradients from diminishing. As well as solving the issue of network degradation, it promotes the accuracy of the networks [12].

Specifically, a residual unit is composed of a convolutional layer, a batch normalization (BN) layer and a nonlinear activation function (Relu). Assuming that the input of a residual unit is xl, a residual unit can be represented by the following Eq. (3). where, $F(\cdot)$ is the residual function, W_l is the weight vector.
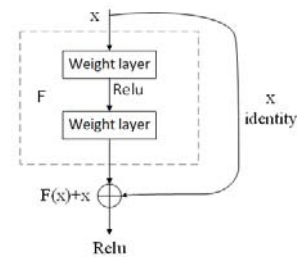
$$x_{l+1} = x_l + F(x_l, W_l) \tag{3}$$

FIGURE IV. RESIDUAL BLOCK

2) *Designed residual network structure*

The ResNet-26 model designed in this paper adds a BN layer after each convolutional layer, so that the back propagation algorithm trains the network more effectively. After training, validation is performed with the model. The parameters in the network are shown in Table 1. The overall network structure is shown in Figure 5.

TABLE I. NETWORK PARAMETERS

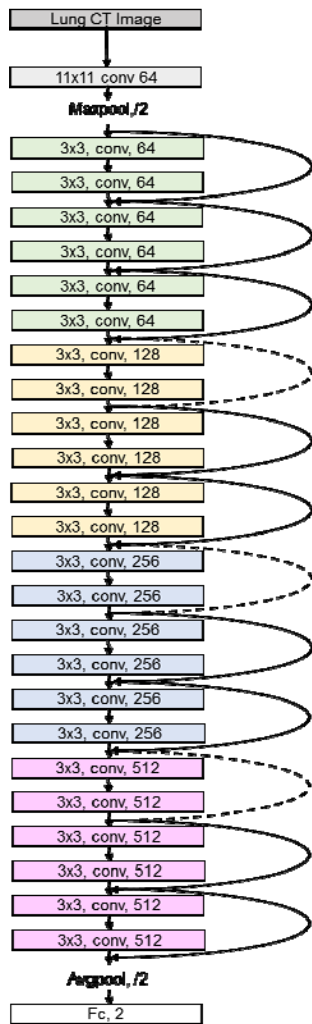| Category	Select type/value
Loss function	Cross entropy
Optimizer	Stochastic gradient descent
Activation function	Relu
Learning rate	0.00002（2e-5）
Training batch	32
Training iterations	40

FIGURE V. NETWORK STRUCTURE OF RESNET-26

III. ANALYSIS OF EXPERIMENTAL RESULTS

A. *Experimental Platform and Method*

Experimental hardware information: Intel® Core(TM) i7-7700K 4.20GHz (8CPUs) Processor, NVIDIA Geforce GTX 1080Ti 11GB Graphic Card. Software information: Ubuntu 16.04, Python 3.6, and Pytorch 0.4.1.

This paper mainly focuses on classification between benign and malignant pulmonary nodules, which are evaluated using the following three indicators. For a two-category problem, there are four cases of predictive results, as shown in Table 3. The correct situations are true positives and negatives.

TABLE II. CONFUSION MATRIX OF CLASSIFICATION RESULT

Forecast result / Actual situation	Positive	Negative
True	TP	FN
False	FP	TN

The four evaluation indexes are defined as follows:

(1) Accuracy Rate, AR: the proportion of correct classification of all instances:

$$AR = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

(2) True Positive Rate, TPR: the proportion of the true class that that are correctly classified out of all the instances:

$$TPR = \frac{TP}{TP + FN} \quad (5)$$

(3) True Negative Rate, TNR: the proportion of the false class that that are correctly classified out of all the instances:

$$TNR = \frac{TN}{TN + FP} \quad (6)$$

(4) Area Under Curve, AUC: The area under the Receiver Operating Characteristic Curve (ROC curve). The curve is based on different classification thresholds, with TPR values as y axis and (1-TNR) values as x axis. By definition, the larger the area under the curve, the higher the performance of the classification model.

B. *Experimental result*

In this paper, a total of 10,402 lung CT images from the LIDC-IDRI dataset are selected; the proportion of benign and malignant labels is close to 1:1; a total of 20804 images are obtained by mirror transformation; training set and test set are randomly assigned at a ratio of 7:3. The model used in this

paper is ResNet-26, which achieves a classification accuracy of 97.53%. The curve of accuracy is shown in Figure 11. The figure shows that the model has converged at around 20 epochs, and finally stayed in the range of 97.50% ± 0.05%. The result also shows that the model reached a TPR of 0.9791 and TNR of 0.9718, while the AUC was 0.958.
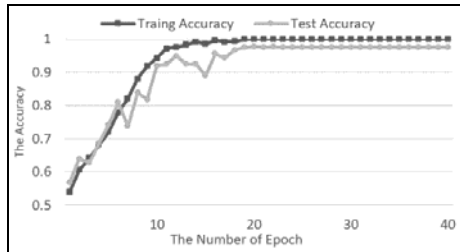

FIGURE VI.   TRAIN / TEST ACCURACY RATE OVER EPOCH

C.   Experimental analysis

To estimate the capability of the model, other models with were selected and compared. The ROC curve is shown in Fig. 13; Methods, the sizes of the images, the AR, TPR, TNR and AUC values, of these papers are shown in Table 4.
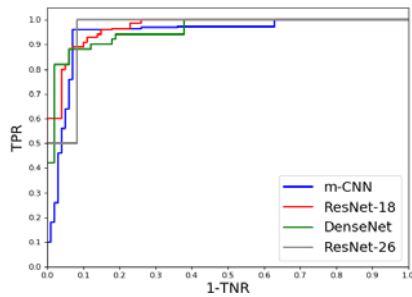

FIGURE VII.   COMPARISON OF DIFFERENT MODEL ROC CURVES

According to Table 4, the accuracy obtained in this paper is higher than that of other methods. It has certain advantages in judging the benign and malignant of pulmonary nodules in CT images. Our method also achieved higher TPR and TNR, reducing the rate of misdiagnoses.

TABLE III.   COMPARISON OF CLASSIFICATION PERFORMANCE OF DIFFERENT METHODS OF PULMONARY NODULES

Source of the model	Input size / pixel	AR/%	TPR/%	TNR /%	AUC
Paper [2]	64×64	89.90	91.07	88.64	0.945
Paper [3]	112×112	89.93	91.14	89.07	0.947
Paper [4]	64×64	92.50	94.00	91.00	0.930
ResNet26	**384×384**	**97.53**	**97.91**	**97.18**	**0.958**

IV.   CONCLUSION

In this paper, we put the ResNet-26 CNN model into the use of the lung cancer classification problem. To improve performance, data drawn from the LIDC-IDRI datasets were amplified, transformed into single channel, cropped, and normalized. The model results in an accuracy of 97.53%, a true

positive rate of 97.91% and a true negative rate of 97.18%, an AUC value of 0.958. Our method outperforms other methods applied to the same problem using the same dataset. The main difference that underlies the better results is that we used the CT images of a whole lung instead of segmenting each nodule for training. Besides, our method is also less time-consuming compared to others. Future works include extending the two-category problem to multi-category, increasing the size of dataset, implementing more complex network architecture to enable broader functionalities.

REFERENCES

[1] Jemal A, Bray F, Center Mm, Ferlay J, Ward E, Forman D. Global cancer statistics [J]. Cancer Journal for Clinicians, 2011, 61(2):59-90.

[2] Nibali A, He Z, Wollersheim D. Pulmonary nodule classification with deep residual networks [J]. International Journal of Computer Assisted Radiology and Surgery, 2017, 12(10):1799-1808.

[3] Dai Y J, Yan S J, Song C L. Benign or malignant lung nodules classification model based on modified DenseNet [J]. Chinese Journal of Medical Imaging Technology, 2018, 34(07):1104-1109. Guan S, Zhang Q Y, Xie H W, Qiang Y, Cheng Z. Convolutional neural network model of CT images recognition [J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30(08):1530-1535.

[4] Charles D. F, Samuel A, Denise A, Matthew B, Claudia H, Michael M, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans [J]. Medical Physics, 2011, 38(2): 915-931.

[5] Mcnittgray M F, Armato S G, Meyer C R, Anthony P.R, Geoffrey M, Richie C. P, et al. The Lung Image Database Consortium (LIDC) data collection process for nodule detection and annotation [J]. Acad Radiol, 2007, 14(12):1464-1474.

[6] Akbar S, Peikari M, Salama S, S. Nofech-Mozes, Anne M. The transition module: a method for preventing overfitting in convolutional neural networks [J]. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 2019, 7(3):260-265.

[7] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]// International Conference on International Conference on Machine Learning. JMLR.org: Lille, France, 2015:448-456.

[8] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[C]// NIPS: Conference and Workshop on Neural Information Processing Systems. Lake Tahoe, Nevada Curran Associates Inc., 2012:1097-1105

[9] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint arXiv:1409.1556, 2014.

[10] Christian S, Liu W, Jia Y, Pierre S, Scott R, Dragomir A, et al. Going deeper with convolutions[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA: IEEE, 2015:1-9.

[11] He K , Zhang X , Ren S , et al. Deep Residual Learning for Image Recognition[C]// 2016 CVPR: 2016 Conference on Computer Vision and Pattern Recognition. IEEE Computer Society: Las Vegas, 2016:770-778.