ISSN:2790-1688 DOI: 10.56028/aetr.4.1.217.2023

Performance study of automatic detection of pulmonary nodules based on convolutional neural network

Youli Ren, Mian Huang*, Shunyou Liu

Yunnan Land And ResourcesVocational College

201810005@yngtxy.edu.cn*

Abstract. In the medical field, the continuous monitoring of the probability of lung cancer incidence can be known that the actual growth rate is becoming faster and faster, and the complications become more and more serious. According to the research and analysis of experts in various fields, it is found that haze is one of the main factors leading to the occurrence of lung cancer in human body, especially in the urban environment. How to use the theory of artificial intelligence technology to design the automatic detection system of lung nodules is the core issue discussed in the field of medical science and technology at present. On the basis of understanding the status quo of convolutional neural network and lung cancer diagnosis and treatment, this study uses CAD system to complete the detection work from four steps, and combines practical cases to judge the detection performance of this technology, in order to lay a foundation for the development of technical innovation in the medical field.

Keywords：Convolutional neural network; Pulmonary nodules; Lung cancer; CAD system; Artificial intelligence.

1. Introduction

Lung cancer has one of the highest mortality rates of all cancers in the world, with a survival rate of only 10 to 100 percent within five years of diagnosis. According to the time survey, the reason why the survival rate of lung cancer is too low is that the disease cannot be detected in the early stage, which is specifically manifested as abnormal areas and irregular growth in lung tissue, among which lung nodules are the most common contents. Therefore, some scholars have proposed in the research that how to accurately detect lung nodules in human body can help better judge the condition of lung lesions. Develop effective treatment plans early. Computed tomography (CT) is an imaging technique with high sensitivity in the detection of pulmonary nodules, which provides an effective reference for the diagnosis and treatment of medical staff. Compared with other imaging techniques, CT imaging has higher quality, but it cannot show pulmonary nodules, and each imaging includes 150 to 500 sections, which will increase the work burden of radiologists, and is easily affected by the subjective consciousness, work experience, knowledge level and other factors, and cannot guarantee the final diagnosis. According to the relevant data survey, the annual growth rate of medical image data in China can reach 30%, while the number of radiology staff increases by only 4%. In other words, the medical field faces great work burden when processing medical image data. Therefore, in order to further reduce the working pressure of staff and improve the efficiency and quality of diagnosis and treatment, computer aided detection technology should be built based on automatic detection of tuberculosis, so as to help them more efficient and accurate diagnosis and treatment.[1-3]

At present, the integration of cutting-edge technologies in the field of artificial intelligence and medical image big data is the core content of research in the medical field. As an important representative of this kind of technology, deep learning can automatically learn the best features from a large amount of data, without manually designing complex features, and ensure the flexibility and versatility of learning features. Compared with other machine learning algorithms, the application performance of deep learning algorithms will continue to improve with the continuous increase of data volume, which is also the unique advantage of the development of medical image big data fusion. At present, the combination of deep learning and automatic detection of lung nodules based on CT image faces many opportunities and challenges, but it can

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.217.2023
give full play to its corresponding value in the medical field as long as technical research and training of professionals are strengthened. After designing the CAD system, this paper mainly analyzed from the four aspects of lung parenchyma segmentation, extraction of areas of interest, computational features and detection of lung nodules, and verified it with practical cases, so as to clarify the application performance of the automatic detection technology of lung nodules with convolutional neural network as the core.[4-6]

2. Methods

This paper mainly uses computer image processing technology to create an automatic detection system of lung nodules for lung cancer diagnosis in the medical field, so as to improve the probability of disease diagnosis and reduce the pressure of staff. The overall system is mainly operated from four aspects. The specific process is shown in Figure 1 below:

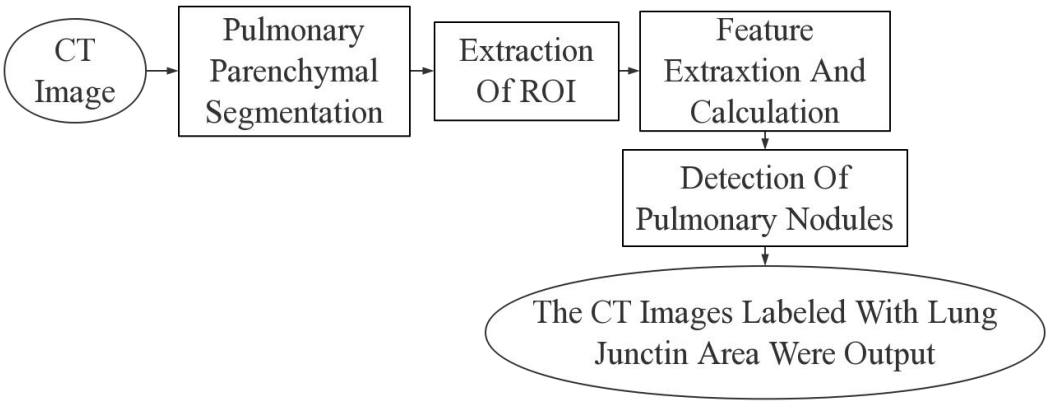

Figure 1 Flow chart of system detection

First, the lung parenchyma is segmented. Taking a CT image as an example, after extracting the torso contour, the interference of the trachea should be removed. These two tasks refer to the preprocessing of CT images. In the extraction of lung parenchyma, the neighborhood shape contour should be extracted from the original CT image, the trunk module should be mastered, and the inner region of the template should be defined to obtain lung parenchyma. At present, the common segmentation and extraction algorithms of lung parenchyma include mathematical morphology operation, threshold segmentation, regional growth method, etc.[7-9]

Secondly, the region of interest is extracted. In the process of image processing, the area of interest is to select an image area in the image, which refers to the focus of image analysis and is also the main problem discussed by medical staff. In this link, the region of interest will be defined first, so as to reduce the data processing time and improve the accuracy of information detection. The input image of this part of the experiment is the image of lung parenchyma obtained after processing in the previous step. In order to avoid omission problem as much as possible, dot filter with mathematical Hessian matrix as the core should be used in this link for detection and analysis, and the region of interest can be automatically obtained.

Thirdly, the features are extracted and calculated. In the previous step, although the area of interest was marked, it not only contained pulmonary nodules, but also overlaps soft tissue and blood vessels. After image processing, it showed the same brightness effect, so there were still many false positives of pulmonary nodules after the completion of the previous step. In the face of these problems, classification is needed to extract important features that must correctly distinguish pulmonary nodules from non-pulmonary nodules. Actual detection extraction features are shown in Table 1:

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.217.2023

Table 1 Feature extraction

Characteristic content	characteristic	Feature description
Visual characteristics	Roundness	Features reflect the degree to which the target is close to a circle.
	compactness	Describe regional characteristics
	consistency	When all gray values are equal, the metric is the largest and decreases from here.
	evenness	Relative Smoothness Measure of Brightness in a Region
	curvature	Boundary curvature
	Boundary irregularity	The relationship between the distance from the boundary point to the center of gravity
Statistical characteristics	variance	Average brightness measurement
	maximum	
	minimum value	
	Third moment	Average contrast measure
	Invariant moment	The third moment of the mean, measuring the skewness of the histogram.
Transform coefficient characteristics	Fourier descriptor	Seven invariant moments that are insensitive to translation, scaling, mirroring and rotation.

Finally, detection analysis. After defining the corresponding characteristics, the classifier of BP neural network should be used for detection and analysis.

Based on the analysis of the figure above, it can be seen that the feature vector obtained in the previous step should be regarded as input data, and then the classification test and analysis should be carried out. It should be noted that, before input, principal component analysis algorithm must be used to optimize the original features. This kind of algorithm will reduce the dimension of data sources on the basis of maintaining the original data information, which can not only solve the problems caused by traditional feature selection, but also improve the classification level of neural network. The optimization results of actual feature vectors are shown in Table 2 below:[10-13]

Table 2 Optimization result

serial number	1	2	3	4	5	6	7	8
characteristic	average /mean value	Third moment	Roundness	Irregularity	First-order invariant moment	Second-order invariant moment	Circular order invariant moment	Fourier descriptor

In this paper, CAD technology is mainly used to build a computer-aided diagnosis system with artificial intelligence as the core, which is convenient for the staff to use image processing algorithms to calculate and analyze the CT images of the lungs, accurately judge the lung nodules contained in the images, and facilitate the detection and analysis of patients' pathological conditions, so as to optimize the clinical diagnosis and treatment effect. This research has laid a foundation for technological innovation in the medical field in the new era, and its application in clinical diagnosis can not only optimize the current level of detection technology, but also reduce the incidence of lung cancer.

3. Result analysis

The data set used in this study came from the pulmonary nodules Detection and Classification competition in 2016, which mainly included 1018 CT images, each of which corresponded to a case.

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.217.2023
In addition to the CT images, the data set also included the files marked by four professional doctors for pulmonary nodules, which were divided into two parts: On the one hand, each doctor would mark the CT images separately and label all candidate nodules as large nodules, small nodules and non-nodules. On the other hand, doctors have to check the information provided by other doctors and determine the final result through communication and discussion. After a joint review by four dimensional doctors, the number of missed medical CT images will be significantly reduced, and the actual results have high confidence.

After mastering the experimental data, it is necessary to do a good job in the first place, and then carry out data enhancement to effectively alleviate the imbalance between positive and negative samples. The ten-fold cross-validation method is used for experimental analysis, and the data set is divided into 10 subsets. Only nine subsets need to be selected each time as the training set, while the other subsets are used as the test set. It can not only make full use of the sample data, but also guarantee the final data results. The test results are shown in Table 3 below:[14-15]

Table 3 Type analysis of test results

Prediction result / Real result	positive	negative	total
positive	True positive（TP）	False Negative（FN）	TP+PN
negative	False positive（FP）	True Negative（TN）	FP+TN
total	TP+FP	FN+TN	TP+FP+TN+FN

Combined with the above table analysis, it can be found that positive images represent pulmonary nodules, while negative images represent no pulmonary nodules. If the doctor marks the result as a nodule and the actual test result is also a nodule, then the result is true positive. If the doctor marks no nodules, but the actual test results are nodules, then the result is a false positive. If the doctor marks the presence of nodules, but the actual test results show no nodules, then the result is false negative. If the doctor notes no nodules and the actual test results are no nodules, then the result is a true negative.

In this study, not only the ablation experiment was designed, but also the ten-fold cross-validation was carried out. The former is to use multi-scale features, extract modules and attention modules to comprehensively improve the detection accuracy of model applications and judge the effectiveness of each model application. The latter will summarize the verification results on the basis of intuitive display of the average detection effect of application performance. The final experimental results prove that the multi-scale feature extraction module and attention module studied in this paper are effective in the automatic detection system of pulmonary nodules. Therefore, future scientific research should apply the convolutional neural network to the detection model, so as to dig more valuable data information deeply. Although the algorithm studied in this paper has achieved excellent results, there are still many improvements in the experimental design, which are reflected in the following points: First, the detection accuracy of the algorithm studied in this paper needs to be improved. Although a representative data set is selected, the CPM score of the algorithm can reach 0.851, which is highly competitive in the medical field. However, compared with other excellent algorithms, there is still a big gap in practical application. Therefore, in the future research discussion, scholars should actively seek effective improvement schemes in the aspects of loss function and network structure, so as to ensure that the research algorithm in this paper maintains the advanced level. Secondly, the amount of data in the experiment is not enough, and the types and formats of images are not perfect. Therefore, in the following experimental analysis, verification and analysis should be carried out in combination with the data sets provided by other competitions. A good cooperative relationship can also be established with hospitals, and relevant medical image data can be regarded as the main basis, so as to improve the generalization

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.217.2023
ability and detection accuracy of the application algorithm. Finally, the research focus of this paper is mainly on the detection of lung nodules, and the classification of benign and malignant is not carried out. In the follow-up research and development of technology, a more complete automatic detection and classification system of lung nodules can be built, so as to facilitate the application and analysis of staff.

Conclusion

In summary, by integrating the existing research experience at home and abroad, it can be seen that there are many types of pulmonary nodules with different sizes and shapes and only variable positions. Therefore, how to efficiently and accurately detect the content in diagnosis and treatment is a major issue for medical innovation in the new era. According to the application status of automatic detection technology of lung nodules, the main content of deep learning technology theory is clarified, and the automatic detection system of lung nodules with convolutional neural network as the core is constructed. It can not only quickly discover hidden safety problems in human body, but also excavate the correlation information between image features from various perspectives, so as to facilitate medical personnel to accurately detect patients' lungs. Reduce the risk of lung cancer. Therefore, according to the existing technical theories and application experience, researchers should continue to deeply explore the automatic detection technology and main modules of lung nodules with convolutional neural network as the core, so as to optimize the level of lung cancer diagnosis and treatment.

References

[1] Ping 'an Wang. Medical image analysis and surgical simulation: Application of artificial intelligence and Virtual reality in medicine [J]. Optics and Opto-Electronics Technology, 2021, 19(6):10.

[2] Qihao Xu,Bo Li. Pulmonary nodule detection based on NRU network [J]. Computer Engineering and Applications, 2021, 57(4):83-90. (in Chinese)

[3] Hong Huang,Rongfei Lv,Junli Tao, et al. Segmentation algorithm of pulmonary nodules in CT image based on improved U-net ++ [J]. Acta Photonica Sinica, 2021, 50(2):65-75.

[4] Lin Yao,Binbin Gu,Xinghua Shen, et al. Treatment experience of extracorporeal membrane oxygenation combined with mechanical ventilation in the treatment of severe pulmonary tuberculosis during pregnancy [J]. Chinese Clinical Case Outcome Database, 2023, 05(01):E00275-E00275.

[5] Mengnan Wang, Xuanyuan Zhao,Ning Xiao, et al. False positive recognition of pulmonary nodules based on Deformable convolutional Neural Network [J]. Computer Engineering and Design, 2022(006):043. (in Chinese)

[6] Yudu Zhao,Zhenwei Peng,Jun Ma, et al. 3D Convolutional Neural network for pulmonary nodule Detection Based on Multi-scale attention Mechanism [J]. Journal of Biomedical Engineering, 2022(002):039.

[7] Zhichao Hou,Yang Yang,XiaoqinLi, et al. Reduction of false positives in pulmonary nodule detection based on 3D convolutional neural network. Beijing Biomedical Engineering, 2022, 41(4):9.

[8] Huaijin Yang,Kewen Xia,Fangyuan Liu, et al. Pulmonary nodule detection based on convolutional neural network and manifold learning [J]. Science Technology and Engineering, 2021, 21(1):9.

[9] Zhi Wang, Yanfei Wu, Xiaofeng Li. Prediction of nodule attributes in lung CT of patients with lung adenocarcinoma by 3D convolutional neural network [J]. Journal of Practical Oncology, 2021, 36(3):7.

[10] Dongyun Huang,Xin Wang, Bin Qin. Pulmonary nodule detection based on 3D convolutional neural network [J]. Journal of Hunan University of Technology, 2023, 37(1):8.

[11] Zhichao Hou, Yang Yang,Xiaoqin Li. Detection and localization of pulmonary nodules based on 3D convolutional neural network [J]. Bioinformatics, 2022, 20(1):7.

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.217.2023
[12] Qi Ai,Jun Wang,Fuquan Ren, et al. Detection of pulmonary nodules in chest CT images based on dimensional adaptive deep neural network [J]. Chinese Journal of Biomedical Engineering, 2021, 40(6):10.

[13] Yujie Zhang,Xining Ye. Pulmonary nodule image segmentation algorithm based on WU-Net Network [J]. Application Research of Computers, 2022(003):039.

[14] Study on lung CT image classification algorithm based on convolutional neural network [J]. Electronic Design Engineering, 2022, 30(21):6.

[15] Xuheng Jiang,Zongqiang Chang,Yingjie Ding. Segmentation of lung parenchyma in CT images based on U-net++ convolutional neural network [J]. Electronic Technology and Software Engineering, 2021(023):000.