

Research on the Application of Cancer-Assisted Diagnosis based on Machine Learning

Zhen Chen, Shuo Xu, Sijia Zhang

North China University of Science and Technology, Tangshan 063000, China;

Abstract

One of the most difficult medical problems is cancer, and the most effective means to cure cancer is early screening and early diagnosis. In the increasingly tense situation of cancer and the shortage of relevant staff, the intelligent cancer screening technology combined with big data technology and artificial intelligence provides an effective way to solve this problem. In recent years, artificial intelligence technology has become increasingly mature, in which neural networks have achieved more excellent results in various related fields, and the most excellent is the deep neural network algorithm. It needs to learn through huge amounts of data to train a good model, and the growing medical data and related data provide a lot of materials. This paper mainly studies the intelligent diagnosis of cancer. After in-depth understanding of the characteristics of relevant medical data and extensive analysis, this paper proposes a series of deep learning neural network models for assisting the intelligent detection of cancer, and performs performance evaluation and verification on several data sets. The main research work and achievements of this paper include : Classify-Multi-view Convolutional Neural Network (CMVCNN) is innovated for intelligent diagnosis of lung nodules in lung CT. Lung CT images have three characteristics. First, the lesion only accounts for a small part of the whole CT, and it will affect the surrounding lung tissue. Therefore, the surrounding environment of the lesion can provide information for the model for intelligent diagnosis. Second, according to the different stages of cancer, the diameter and shape of lung nodules will be greatly different. Therefore, we improve the multi-field convolution neural network, segment the diameter of nodules in multiple ranges, and then segment the corresponding field of vision according to the corresponding proportion. Third, the lung CT image is composed of multiple slices, and a more three-dimensional organ structure can be formed after the slices are connected. According to its characteristics, the multi-view convolutional neural network after classification is used. The model uses multiple horizons and uses different horizons to diagnose lesions. Using the model in LIDC-IDRI (Lung Image Database Consortium and Image Database Resource Initiative) for two classification diagnosis, judge whether benign. The data and model are three-dimensional and three-dimensional, so that the convolution kernel can be moved from three dimensions, so as to extract the characteristics of lung nodules in three dimensions. The model was used for binary diagnosis on Luna16

Keywords

Intelligent Cancer Diagnosis Neural Network, Artificial Intelligence, Lung Nodules.

1. Overview

1.1. Research background and significance

According to MedSci survey, the number of new cancer cases increased from 18.7 million in 2010 to 23.6 million in 2019, an increase of 26.3 % worldwide. During this period, the incidence of age standard values remained basically unchanged.

Cancer is the second leading cause of death in the world, causing about 10 million deaths each year. About a sixth of the world ' s death is caused by cancer, anti-cancer and cancer prevention is imperative, and early diagnosis of cancer has a great role in cancer treatment. But the discipline requires sufficient expertise and diagnostic experience, with fewer professionals and an increasing number of cancer patients, resulting in a shortage of medical resources and even an increase in cancer misdiagnosis rates. Therefore, improving the efficiency and accuracy of cancer diagnosis has become the purpose of cancer intelligent diagnosis.

Improving the performance of cancer intelligent screening and diagnosis requires data and algorithms. In many cases, getting more data is more important than getting better algorithms[1]. With the deepening of medical informatization, medical industry data accounts for a significant proportion in the digital universe, and the amount of data increases by 48 % annually, which is one of the fastest growing industries [2] . This provides a good opportunity for the development of intelligent diagnosis of cancer.

In recent years, artificial neural network is one of the most active algorithms in the field of artificial intelligence. Its development is attributed to the massive growth of data, the rapid development of hardware performance and the neural network architecture such as AlexNet model. Compared with the traditional manual extraction of image features, convolutional Neural Network (CNN) local links and multi-layer architectures have more performance advantages.

The application of artificial neural network in intelligent diagnosis of cancer has the following advantages :

1. Deep convolution neural network has achieved great success in the field of image processing, and medical image data occupy a large part of medical data, which improves the performance of the model for cancer diagnosis through image data.
2. Strong feature extraction capability. The multi-layer architecture of neural network can extract the advanced features hidden in the original data, so as to design a better model.
3. High performance ceilings. Ordinary machine learning algorithms, performance has the corresponding upper limit, to a certain extent, even if the amount of data can not improve performance (Figure 1-1) [3] . Using AlexNet architecture, as long as the amount of data increases, the performance of neural networks can continue to rise.

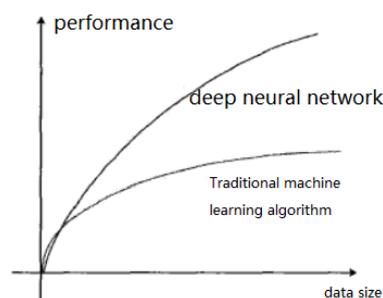
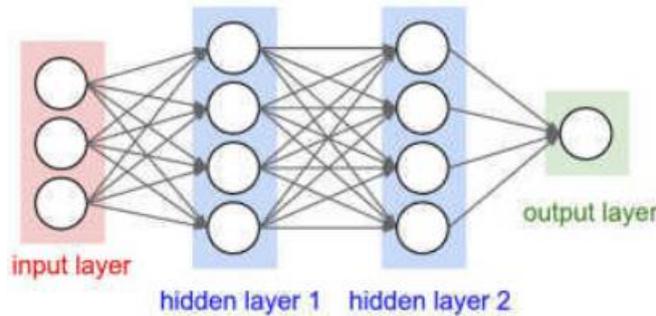


Fig. 1 - 1 : The performance change of deep neural network with the amount of data

2. Introduction of convolutional neural network

Convolutional Neural Networks, referred to as CNN, is a kind of Feedforward Neural Networks with convolution calculation and deep structure. It has the ability to represent learning and can classify the input information according to its special structure.

The overall structure of convolutional neural network is input layer, hidden layer and output layer. Figure :



2.1. Input layer

Most of the data can be processed, and standardized processing is needed to improve the learning efficiency and performance of convolutional neural networks.

2.2. Hidden layer

the hidden layer generally includes three common structures : convolution layer, pooling layer and full connection layer.

2.3. Convolution layer

The function of convolution layer is to extract the characteristics of the input data, and its interior often contains multiple convolution kernels, and each element constituting the convolution kernel corresponds to a weight coefficient and deviation. When the convolution kernel works, it will regularly sweep the input characteristics, and sum the input characteristics by matrix element multiplication and add bias.

$$Z^{l+1}(i, j) = [Z^l \otimes w^{l+1}](i, j) + b = \sum_{k=1}^{K_l} \sum_{x=1}^f \sum_{y=1}^f [Z_k^l(s_0 i + x, s_0 j + y) w_k^{l+1}(x, y)] + b$$

$$(i, j) \in \{0, 1, \dots, L_{l+1}\} \quad L_{l+1} = \frac{L_l + 2p - f}{s_0} + 1$$

The convolution kernel parameters include the size, step size and filling of the convolution kernel, which jointly determine the size of the feature map output by the convolution kernel and are the super parameters of the convolution neural network. The convolution kernel size can be specified as any value smaller than the input image size. The larger the convolution kernel, the more complex the input features can be extracted. [4]With the accumulation of the convolution layer, the size of the feature map will gradually shrink.

The convolution step length defines the distance of the position when the convolution kernel scans the feature map twice adjacently. When the step length is N, the next scan will skip N - 1 pixels.

2.4. Pooling layer

The feature map extracted from the convolution layer is passed to the pooling layer, so as to select features and filter useless information. The pooling layer contains a pre-set pooling function that replaces the entire feature area by delimiting a point in the feature area through a specific selection method. The steps of choosing pooling region in pooling layer are the same

as those of choosing convolution region in convolution layer, but the convolution operation is changed into pooling operation.

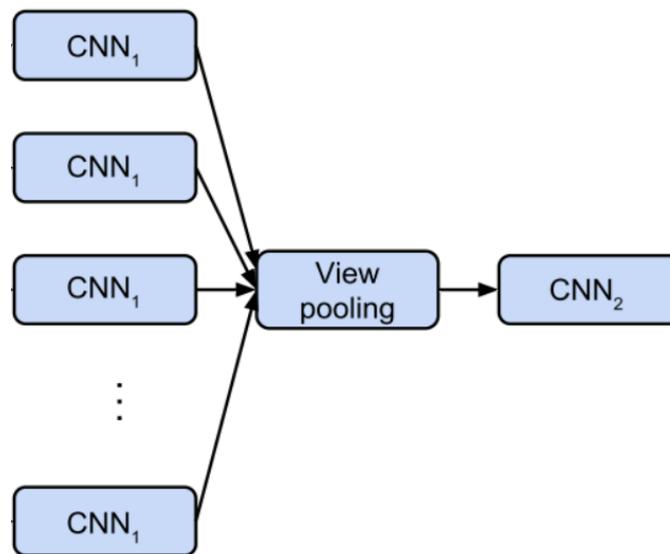
2.5. Full connection layer

In convolutional neural networks, the full connection layer is equivalent to the hidden layer in the traditional feedforward neural network. The full connection layer is located in the last part of the hidden layer of the convolutional neural network, and only transmits signals to other full connection layers. The feature map will lose the spatial topological structure in the full connection layer, and will be expanded into vectors and passed through the excitation function[4].

3. Method

3.1. Post-classification 3D multi-view convolutional neural networks

The essence of lung nodule diagnosis is to extract the features of lung nodule images and compare the judgment of feature similarity. It can be regarded as the feature extraction task with input CT image K and nodule position $[x, y, z, d]$. Where x, y, z represent the center coordinates of the lung nodule cube in the image, d represents the diameter of the lung nodule. To achieve this goal, a 3D multi-view convolutional neural network after classification is proposed



The network is composed of basic model CMVCNN.

3.2. CMVCNN

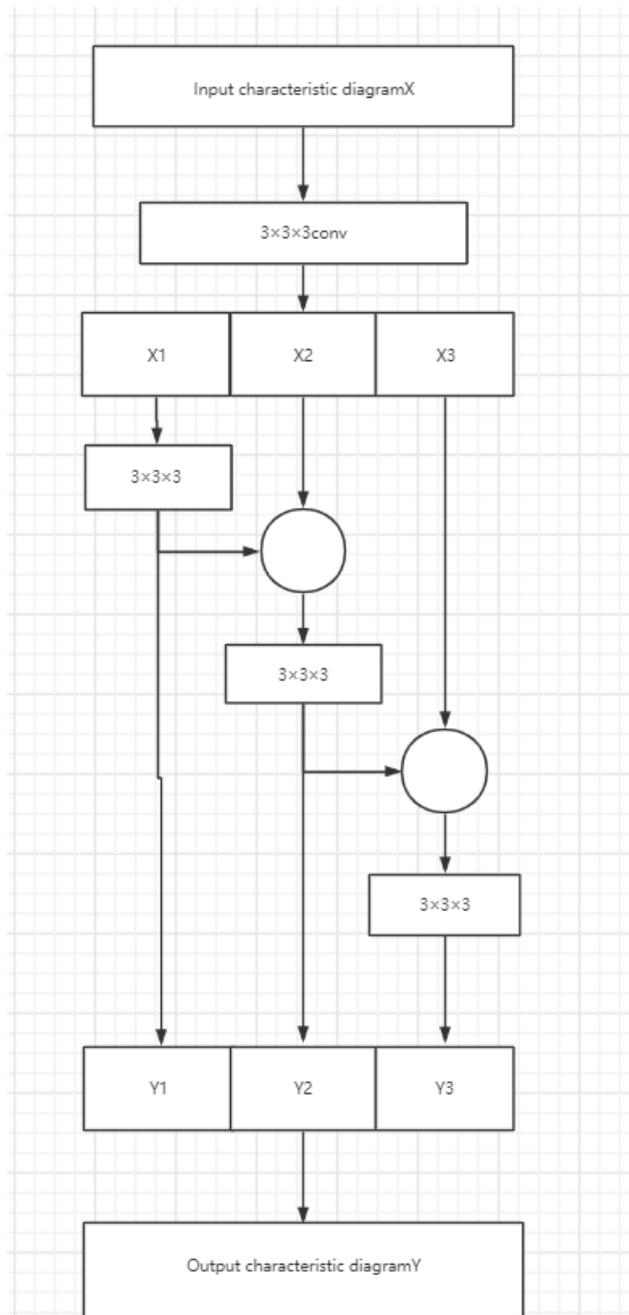
CMVCNN is composed of MVCNN and classification module units, and the module structure is shown in Figure 2. X_i, Y_i ($i = 1, 2, 3$) represent the feature maps of different views, $3 \times 3 \times 3$ represent the convolution kernel size of the convolution layer is $3 \times 3 \times 3$, and each convolution layer has a pooling layer and a Relu layer.

CMVCNN first classifies the data set annotation files by classification module units, and then uses a set of $3 \times 3 \times 3$ convolution kernels to classify the input feature graph K_i ($i = 1, 2, 3, \dots$) The output feature map is divided into three groups according to the visual field channel, and then the processed feature map X_i is added to the output Y_{i-1} of the previous filter Z_{i-1} and sent to the filter Z_i to obtain the output feature map Y_i , which is specific to

$$Y_i = \begin{cases} Z_i * X_i, & i = 1 \\ Z_i * (Y_{i-1} + X_i) & 2 \leq i \leq 3 \end{cases}$$

Finally, splice three Y together.

The compression process can output global features through adaptive average pooling. The following figure.

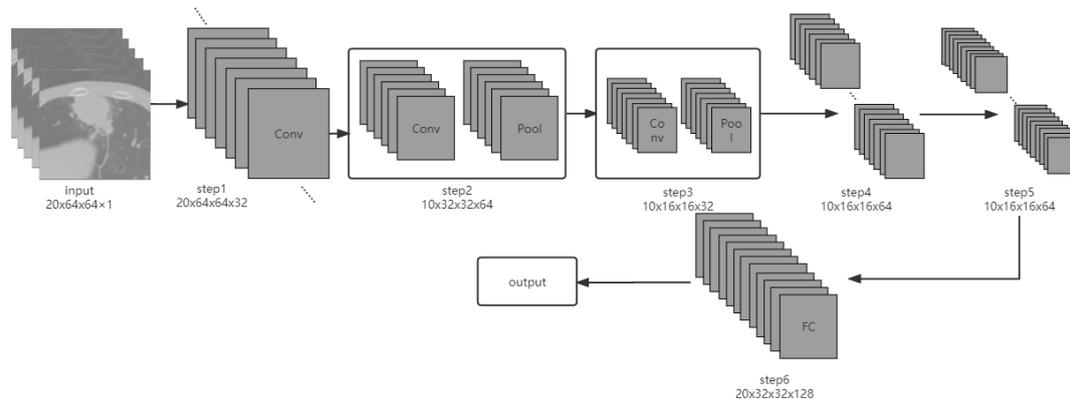


3.3. Network architecture

The 3D classified multi-view convolutional neural network proposed in this paper adopts the Alex-like network structure. According to the data characteristics after classification, the selected area is cut into 40 %, 80 % and 100 % of the maximum diameter of the classification as the field size. The network consists of two parts, labeled reading unit and MVCNN. The input is the image cube of lung nodules after cutting, and the dimension is (length × width × height × field of vision), that is (N1, N2, N3,1).

$$N_k(k = 1,2,3) = L_i * T_j(i = 1,2,3,4,5,6, j = 1,2,3)$$

Li is the maximum diameter after classification, Tj is the visual field selection ratio, Nk is the visual field.



The core MVCNN is divided into five stages. The first stage is a convolution layer and a Dropout layer, and convolution is performed with multiple $3 \times 3 \times 3$ convolution cores. The second and third stages include the convolution layer, the average pooling layer and the Dropout layer, respectively. Thus, the specific template is shown in the above figure. The fourth stage is used to switch the field of vision, and finally in the fifth stage, the full connection is carried out to predict the image, so as to predict the diagnosis.

Among them, the role of Dropout layer is to prevent over-fitting, and the output results of Dropout layer are input into the next layer of the network. The global output result is the possibility that the lung nodules at the current label position are false positive.

The loss function used in this network is categorical_crossentropy for single label classification.

$$\text{Loss} = - \sum_{i=1}^{\text{output size}} y_i \cdot \log \hat{y}_i$$

3.4. Experiments

3.4.1. Experimental set

The data of LUNA16 come from a larger dataset LIDC-IDRI, which has a total of 1018 CT scans, i. e., 1018 cases. Each CT image has an xml label file. The data of this dataset come from seven different academic institutions, and the scanners and related parameters are not the same. Therefore, the distribution of 1018 images is uneven.

The LUNA16 dataset removes CT with slice thickness greater than 3 mm, and removes CT with slice space inconsistency and missing slice. Finally, 888 CTs are generated, which constitutes LUNA16. However, there are two problems.

1. Excluding more than 3mm CT is because the thinner the slice, the better the effect, which is not difficult to understand, the lung scanning, must be the more data obtained by scanning, if only one slice of the whole lung, then even 3D data are not available.
- 2.Slice space is the sliced data is 3D, that is, the two-dimensional slices are integrated together, there are z, y, x three dimensions, spacez refers to the slice thickness, spacey and spacex refers to each slice of a single pixel represents the actual width and height of lung tissue. All CT images are stored in. mhd format.

3.4.2. Data processing stage

Firstly, the existing CT images were classified according to the diameter of lung nodules and positive and false positive. The classification range was below 3cm, 3cm to 6cm, 6cm to 12cm, 12cm to 18cm, 18cm to 24cm and 24cm to 32cm. After classification, the lung nodule center

was segmented according to the maximum range diameters of 40 %, 80 %, and 100 %. Then, each lung nodule image after cutting was rotated by 90°, 180°, and 270°, respectively, and the images with four angles were obtained, which again increased the amount of data.

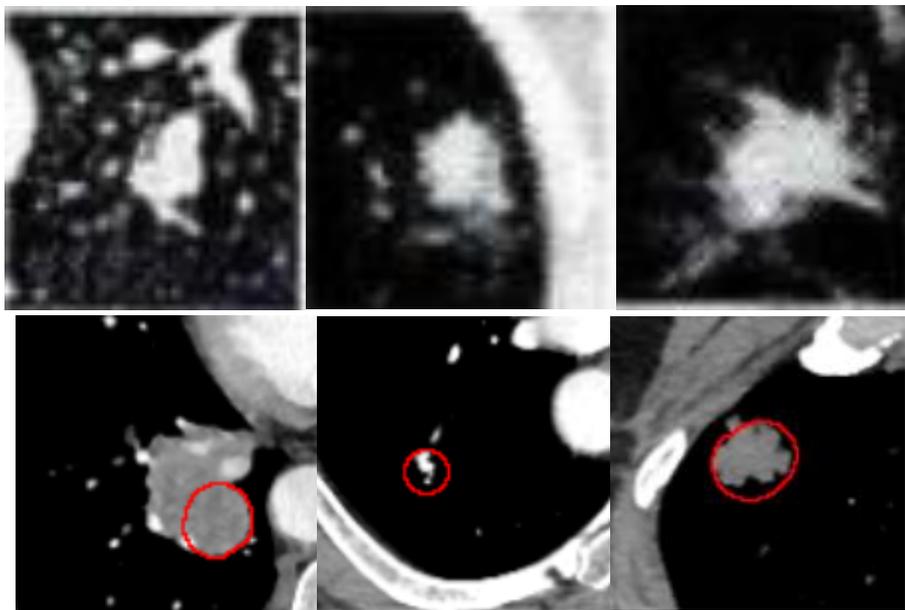
At the same time, in order to improve the performance of image features, the voxel range of lung nodules was normalized from [- 1200, 600] to [0, 1].

Specifically

$$\text{val} = \begin{cases} 0 & \text{val} < -1200 \\ \frac{\text{val} - (-1200)}{600 - (-1200)} & -1200 \leq \text{val} \leq 600 \\ 1 & \text{val} > 600 \end{cases}$$

Then, the pixel spacing in X, Y and Z directions of CT image was uniformly adjusted to 1 mm.

The preprocessed image is as follows :



Among them, the images segmented with the maximum diameter of 20 % can extract the internal complex features of the lesion and improve the performance of the extracted features, so as to maximize the accuracy of the resolution of lung nodules. The images segmented with the maximum diameter of 50 % can extract the internal features of a few lesions and the influence of a large part of the lesions on the surrounding tissues, so as to improve the robustness and identification of the model. The images segmented with the maximum diameter of 100 % will extract the internal characteristics of all lesions and the influence of lesions on surrounding tissues, so as to obtain more information to improve the performance of the model and increase the amount of data.

3.4.3. Training phase

The training platform used in the experiment used an Intel® Core™ i7 - 9750H processor with a CPU of 2.6 GHz basic main frequency and 4.5 GHz maximum core frequency, and a memory of 8G. All experiments used the data preprocessing described in Section 2.2. All models and network structures are built with Python 3.7, and the framework is built with Tensorflow to accelerate on a GTX1650 card.

A set of chest CT after data processing will be composed of a series of slices in multiple directions. After stacking these slices in order, a three-dimensional chest projection image will be formed.

The processed three-dimensional images are grouped into six groups according to the nodule size above and output to the corresponding neural network in order for training. Batch size set

to 64, a total of 150 training. In the training to 50 times, 100 times and 150 times, the learning rate is adjusted to find a more appropriate learning rate.

3.4.4. Evaluation criteria

After processing, it is necessary to calculate the accurate value of the obtained results, and the calculated results can be used as the evaluation results of the algorithm performance. The accuracy calculation formula is as follows :

$$H = \frac{R}{R + W}$$

In this formula, R represents the number of true positive nodules, and W represents the number of false positive nodules. The criterion of true nodules is that the labelled nodules are positive nodules, and the false positive nodules are labeled with positive nodules.

3.4.5. Experimental results

In order to verify the rationality of CMVCNN in the proposed network structure, two models before improvement, namely 3DCNN and MVCNN, were introduced to conduct accuracy tests on Luna16, and comparative experiments were carried out. The experimental results are shown in Table 1. Experimental results show that.

Among the three network architectures, 3DCNN has the lowest diagnostic accuracy, but it still reaches 83.21 %, indicating that 3DCNN with simple structure has considerable performance. In order to verify the effectiveness of MVCNN, 3DCNN is modified to MVCNN. The experimental results show that the diagnostic accuracy of MVCNN increases by 1.06 %, indicating that MVCNN can effectively improve the diagnostic performance. In order to verify the performance of CMVCNN, CMVCNN is used to replace MVCNN. The experimental results show that the diagnostic accuracy of CMVCNN is 1.58 % higher than that of MVCNN, which proves the rationality of CMVCNN.

Table1 Performance comparison of different nodule convolutional neural networks

Network model	accuracy
3DCNN	83.21%
MVCNN	84.27%
CMVCN (This Paper)	85.85%

4. Conclusion

In this paper, a multi-view lung nodule diagnosis method based on convolutional neural network is proposed, which is composed of classification module unit and multi-view 3D convolutional neural network. In order to effectively extract the features of lung nodules in different stages, the classification module and multi-view 3D convolutional neural network are combined to build a multi-view convolutional neural network after classification. Compared with the existing lung nodule diagnosis methods, the proposed algorithm improves the diagnostic efficiency and optimizes the diagnostic performance, which has high value in the field of lung diagnosis.

References

- [1] Domingos P. A Few Useful Things to Know About Machine Learning[J]. Communications of the ACM, 2012, 55(10): 78-87.
- [2] Corporation ID. The digital universe driving data growth in healthcare[R]. Hopkinton, Massachusetts: IDC Analyze the Future, 2014:1-16.

- [3] Alom MZ, Taha TM, Yakopcic C, et al. The History Began from Alexnet: A Comprehensive Survey on Deep Learning Approaches[J]. aeXiv preprint arXiv:1803.01164,2018.
- [4] Goodfellow I, Bengio Y, Courville A. Deep learning (Vol. 1). Cambridge:MIT press, 2016