Application of Artificial Intelligence Technology in Medical Imaging

Min Xu¹ and Chuanhai Jia²,*

¹School of Medical Imaging, Bengbu Medical College, Bengbu, Anhui Province, 233030, China
²Department of Medical Imaging, Changshu Hospital Affiliated to Xuzhou Medical University, Changshu, Jiangsu Province, 215500, China

*Corresponding author e-mail: 2547769174@bbmc.edu.cn

Abstract. Based on massive medical image data, through reasonable system design and parameter optimization, the AI-assisted diagnosis system can help doctors make more clinical decisions, reduce their daily work pressure, reduce unnecessary invasive tests, and thus improve the quality of life of patients. This paper mainly studies the application of artificial intelligence technology in the field of medical imaging. In this paper, deeplesion public data set is used as the training set to study and optimize the convolutional neural network structure and parameters of deep learning algorithm based on the basic convolutional neural network model, aiming at the pathological features of lung cancer in the data set, such as lobulation sign and burr sign. At the same time, combining with the characteristics of neural network recognition and data set, the data set is preprocessed to highlight the features of the image and improve the accuracy of deep learning algorithm. Moreover, an auxiliary recognition system is developed around the optimized algorithm, so that the deep learning algorithm can intuitively show the automatic recognition of lung cancer through the recognition system after a large amount of data training.

Keywords: Artificial Intelligence, Medical imaging, Deep Learning, Convolutional Neural Networks

1. Introduction
Artificial intelligence has now penetrated into our lives, from personal assistants, security and autonomous driving to healthcare, e-commerce, finance and education, AI is closely related to our lives. Artificial intelligence can realize automatic recognition and diagnosis function in the field of intelligent medical treatment. For example, as long as the image of the mole on the arm is obtained, the computer can clearly determine whether it is a mole or skin cancer, and a group of eyeball images can be taken to determine whether there is a risk of cataract or deepened myopia [1]. The development of artificial intelligence has made people's imagination unrestrained. On the one hand, medicine is slowly being miniaturized from large equipment in hospitals through intelligence, such as the skin cancer test mentioned above, where people collect images from their mobile phones and upload them...
to a cloud server so that a computer can make a diagnosis. On the other hand, the smarter a medical device is, the more complementary it is to a doctor. Intelligentization of medical treatment brings convenient and quick diagnostic assistance to doctors, reduces the working pressure of doctors, and brings efficient medical treatment and low price to patients [2]. In the face of complex diseases, experienced doctors will inevitably misdiagnose, while the new technology of intelligent medical treatment can extract accurate and effective information from complex diseases to assist doctors in efficient diagnosis.

Intelligent medicine is one of the most popular research directions in the medical field in recent years. Most of the algorithms adopt deep learning method. Deep learning algorithms have achieved good results in natural images, and deep learning algorithms have been gradually applied to various fields, including data mining, traffic services and medical diagnosis. In the field of medical diagnosis, Stanford University realized the recognition of skin cancer through deep learning algorithm, trained the CNN network through 129,450 clinical data, and the accuracy rate was very close to that of doctors [3]. In one study, the deep learning algorithm InceptionV3 was used to train 9649 IHC staining and 2490 H&E staining lung cancer tissue sections, and the accuracy of AI recognition of lung cancer reached 99%[4].

In most of the open deep learning algorithms, lung cancer deep learning algorithms are relatively few or not fully open compared with other fields. However, lung cancer images are quite different from other cancer images, which requires some optimization based on the algorithm. At the same time, the market lung cancer diagnosis system is expensive and requires a large number of comprehensive medical data, still many hospitals can not afford. Therefore, the development of medical imaging detection of lung cancer based on artificial intelligence technology has great theoretical significance and practicability.

2. CT Images Processing Based on Convolutional Neural Network

2.1 Overview of Convolutional Neural Networks

Convolutional Neural Network (CNN) is a major research direction of deep learning. It is the first successfully trained neural network containing multi-layer structure, widely used in the processing of one-dimensional sound signals, two-dimensional images, three-dimensional image sequences and video signals, especially excellent in the processing of two-dimensional images. In recent years, in international conference papers such as ICCV, CVPR, NIPS and ECCV, as well as in ILSVRC competition, solutions to computer vision problems have been involved, and almost all models have been completed based on CNN [5-6].

The structure of CNN can be understood as a multi-layer feedforward network, each layer is composed of multiple two-dimensional planes, each plane is composed of multiple neurons, and its basic structure consists of four parts: Input layer, convolutional layer, pooling layer and output layer [7].

The input layer can directly receive two-dimensional visual patterns, multi-channel images and image sequences without manual selection or design of appropriate features as input.

In CNN, the convolution of the convolution layer is general image convolution, namely two-dimensional discrete convolution. The formula of image convolution is as follows:

\[
W(x, y) \ast F(x, y) = \sum_{u}^{U} \sum_{v}^{V} W(u, v) F(x - u, y - v)
\]

(1)

Where I = F(x, y) represents an image, F(x, y) is the gray value of the pixel in row x and column y of the image, and W(u, v) is the gray value of the pixel in row u and column v of the convolution kernel.

What is special about convolution in CNN is that its convolution can be trained.

The down-sampling layer, also known as the pooling layer, is generally alternately connected with the convolutional layer. The core of the down-sampling layer is to conduct sub-sampling of images, so as to reduce the amount of calculation and retain key information [8].
In CNN, the convolution layer, pooling layer and full connection layer usually introduce nonlinear factors through activation function, so that the model can better solve complex nonlinear problems [9]. Common activation functions include Sigmoid, Tanh and ReLU.

The Sigmoid function is a monotone and continuous function, and its expression is shown in Equation (2). The input can be mapped to (0,1) and the derivative is easy to be obtained. However, the soft saturation of the function itself makes it easy for the gradient to disappear in the training process.

\[ f(x) = \frac{1}{1 + e^{-x}} \]  

Equation (2)

The tanh function expression is shown in Equation (3), which is a zero-centered output function. Although the convergence rate is improved compared with Sigmoid function, it does not solve the problem of gradient disappearance in the training process.

\[ \tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \]  

Equation (3)

Relu is a linear unsaturated function. Compared with the previous two functions, its biggest advantage is that the computational complexity is reduced and the optimization speed is very fast. On the one hand, the unsaturation of ReLU improves the sparsity of the network, but once suppressed (zero) neurons may not be activated again [10]. Different activation functions have their advantages and disadvantages, and scholars have been working on the research of activation functions with better performance.

2.2 Image Processing

The spectrum map is intercepted according to the coordinates of the region where the spectrum map is located provided in the Deeplesion file, and the region information in the Deeplesion file is obtained by referring to the relevant information of the US region calibration module in the information module specification published by Deeplesion.

Because the coordinate Settings of the enrolled spectrum are inconsistent, the peak information in the spectrum waveform cannot be reflected in the spectrum image. In order to avoid the loss of the peak information, we need to normalize the spectrum first. In the normalized spectral image, the peak information of the waveform can be reflected by the height of its highest point. Specific operation: look for the minimum scale of spectrum in DICOM data set (the minimum interval in X direction is 0.004s, and the minimum interval in Y direction is 0.078cm/s), and normalize all spectrum graphs under this scale.

Normalized spectral images contain different number of signal cycles and different image sizes, so we standardized the images. Specific operation: First, the signal of two cycles is extracted from the spectrum diagram for analysis. Secondly, we construct the background image with the maximum image size (1700x1700, all pixel values are 0), and place all normalized spectral images in the center position of the background image with this size.

In order to further reduce the amount of image data, simplify the image, and highlight the spectral waveform contour, we normalized the unified size of spectral images for image binarization.

Binary image plays an important role in image classification, image binarization of the original image can be turned into black and white images, simplify the image data at the same time highlight the outline of main goal, on the one hand able to more clearly reflect the characteristics of the local and overall, on the other hand, the amount of data to reduce the network can also speed up the model of training, shorten the learning time.

2.3 Model Building

This paper adopts the most basic CNN model for deep learning. The established network model is trained and tested with data samples.
Convolutional neural network is composed of a series of operation layers. In this paper, a total of 6 operation layers are used in the network, namely Conv2D, Activation, MaxPooling2D, Ddropout, Dense, and Flatten.

Conv 2D is the convolution layer: the convolution operation is performed on the input tensor (usually the matrix of HxWxC). A convolution layer usually contains several convolution kernels. Each convolution check is used to carry out local weighting operation on the input image, and the parameters of weighting operation are learnable. The main function of the convolutional layer is to extract image feature maps. After the number of kernels passing through this convolutional layer, several feature maps will be generated. For example, feature map A may extract image edge features, feature map B may extract image color features, and so on.

A nonlinear operation is performed on the input tensor (usually the matrix of HXWXC). The nonlinear operation used in this network is mainly Max (x, 0), where x is each element in the matrix. The main function of the activation layer is that the samples are usually not linearly separable in the feature space, so it is necessary to carry out nonlinear transformation to make the sample points linearly separable in the transformed space.

MaxPooling2D is a pooling layer: the input tensor (usually the matrix of HXWXC) is locally maximized, equivalent to a local point to represent the entire region.

Dropout is the discard layer: Due to the large number of parameters and strong fitting ability of convolutional neural network, overfitting is easy to occur in the case of limited samples. Dropout is a means to reduce overfitting. The Dropout layer will randomly discard some elements (set to 0) according to a certain proportion of the input tensor. Since the discarded position is random each time, it requires the subsequent computing layer to use more global abstract features to classify, similar to a piece of incomplete jigsaw puzzle to obtain the overall information.

Flatten is the unwrapping layer: the input tensor (usually an HXWXC matrix) is unrolled into a one-dimensional vector (the feature that is ultimately used for classification).

Dense is the full connection layer: the input one-dimensional vector is linearly weighted to get the final classification probability.

3. Identification Process and Parameter Setting of Convolutional Neural Network

In this identification experiment, NewVGG neural network structure (JSON file) and corresponding weights (H5 file) were selected to reconstruct the neural network.

The recognition data set is divided into two parts. One part is 1000 lung cancer images and 1000 normal lung images selected from the public data set Deeplesion, a total of 2000 images. The other part is the desensitization CT data of the hospital. According to the PET-AC images corresponding to the CT, 200 CT images of suspected lung cancer are found for recognition. The format of the input image is 512*512 gray image of single channel.

The pre-processing Deeplesion recognition set uses the same processing method as the training data set, and finally makes a single channel image of 8bit grayscale in PNG format. In the preprocessing of hospital images, the Radiant Dicom Viewer (Radiant Dicom Viewer) is used to save the image of suspected lung cancer as a PNG format, and then the image is normalized.

Batch identification can be easily achieved by combining the recognition function with a simple cycle function. PREDICT outputs 0 as lung cancer and 1 as normal, and records the statistical accuracy.

4. Identification Results and Analysis

4.1 Results of Public Data Sets
As shown in Table 1 and Figure 1, since 2000 images are known to include 1000 lung cancers and 1000 normal lungs, ACC (TP)=89.4% and ACC (TN)=83.2% can be calculated through the accuracy formula of statistical recognition, and the recognition accuracy is 86.3%, namely, 274 images are judged wrong.

Through observation, most of the misidentified images are not obvious after preprocessing. If lung window or mediastinal window can be used to make lung cancer more obvious, the recognition rate can be further improved. Therefore, in the recognition of hospital data, each image has been processed by adding window width and window level. The other part of data features have certain specificity, and the corresponding training sample size is relatively small, which leads to failure to learn the corresponding eigenvalues when training the network model. Most of the clinical manifestations of lung cancer are: vesicle sign, lobulation sign, burr sign, pleural depression sign, vascular cluster sign, small spinous process sign and calcification. Most of the image shapes such as vesicle sign, lobulation sign and burr sign can be well recognized.

4.2 Hospital Dataset Results
A total of 500 images were collected from the hospital dataset. The 500 patients were double-blind tested, and the images were confirmed with the assistance of doctors at the hospital.
As shown in Figure 2, in the 300 hospital images, the true positive rate was 85.3%, and the true negative rate was 53.5%. The combination of MSCT and magnetic resonance imaging (MRI) can reduce the missed diagnosis rate of lung cancer to 5 to 10 percent, compared with 15 to 20 percent for experienced physicians in clinical diagnosis. Through the analysis of the above test results, the missed diagnosis rate of the system in the public data set was 10.6%, and the missed diagnosis rate in the hospital data set was 14.7%, both lower than the clinical missed diagnosis rate. In this paper, the neural network is trained with public data set as training group and has been preprocessed to a certain extent. Compared with some hospital data, the specificity of some hospital data is strong. Among the 300 cases collected, there are some cases that do not have public data set. From the perspective of learning and research, this system can achieve a higher accuracy and a lower rate of missed diagnosis. From the perspective of clinical application, a large amount of hospital data is still needed to train the neural network to optimize the accuracy and missed diagnosis rate of the system.

5. Conclusions
With the continuous improvement of modern medical imaging technology and medical means, computer-aided diagnosis and computer-guided surgical treatment have been more and more applied in clinical practice. At the same time, it also put forward higher requirements for the corresponding computer technology, faster detection, more accurate location of the lesion area is the pursuit of the goal of the skill field. In this paper, CNN (Convolutional Neural Network) is used to design intelligent recognition of lung cancer. Its purpose is to provide an auxiliary diagnosis reference for examining doctors in the Department of Ultrasound and for imaging students in clinical practice. In the process of model design and training, image processing is adopted and the method of model building is used to improve the results of network classification. The final results show that deep learning can effectively identify lung cancer. Therefore, the application of deep learning in the clinical application of lung cancer has great prospects and is worth further exploration and research.

References
[1] Marie W A , Yong L , Regner K R , et al. Artificial Intelligence, Physiological Genomics, and Precision Medicine[J]. Physiological Genomics, 2018, 50(4):237-243.
[2] Sammour T , Bedrikovetski S . Radiomics for Diagnosing Lateral Pelvic Lymph Nodes in Rectal Cancer: Artificial Intelligence Enabling Precision Medicine?[J]. Annals of Surgical Oncology, 2020, 27(11):4082-4083.
[3] Arimura, Hidetaka, Soufi, et al. Radiomics with artificial intelligence for precision medicine in radiation therapy[J]. Journal of radiation research, 2019, 60(1):150–157.

[4] Buch V H, Ahmed I, Maruthappu M. Artificial intelligence in medicine: current trends and future possibilities[J]. Br J Gen Pract, 2018, 68(668):143-144.

[5] Jaderberg M, Simonyan K, Vedaldi A, et al. Reading Text in the Wild with Convolutional Neural Networks[J]. International Journal of Computer Vision, 2016, 116(1):1-20.

[6] Tajbakhsh N, Shin J Y, Gurudu S R, et al. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?[J]. IEEE Transactions on Medical Imaging, 2016, 35(5):1299-1312.

[7] Pereira S, Pinto A, Alves V, et al. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images[J]. IEEE Transactions on Medical Imaging, 2016, 35(5):1240-1251.

[8] Selim A, Elgharib M, Doyle L. Painting style transfer for head portraits using convolutional neural networks[J]. Acm Transactions on Graphics, 2016, 35(4):1-18.

[9] Korfiatis P, Kline T L, Lachance D H, et al. Residual Deep Convolutional Neural Network Predicts MGMT Methylation Status[J]. Journal of Digital Imaging, 2017, 30(5):622-628.

[10] Cai N, Su Z, Lin Z, et al. Blind inpainting using the fully convolutional neural network[J]. Visual Computer International Journal of Computer Graphics, 2017, 33(2):249-261.