Pneumonia Recognition Based on Convolutional Neural Network Feature Map Fusion

Xiaozhong Liu*, Zaixing Wang, Lijun Zheng, Jinhui Gao

School of Electronics and Information Engineering, Lanzhou Jiaotong University, No88, Anlingxi Road, Lanzhou, Gansu 730070, China

*rossi_liu66@gmail.com

Abstract. Pneumonia recognition has important research significance in computer-aided diagnosis, and there is a problem of low accuracy for pneumonia recognition. In this paper, an improved network is based on the convolutional neural network AlexNet, the AlexNet_Branch network. The AlexNet_Branch network adds a parallel branch convolutional neural network to AlexNet, and it connects AlexNet and the branch convolutional neural network at the fully connected layer. During training, the same image is simultaneously obtained by AlexNet and the branch convolutional neural network to obtain different feature maps, and then the feature maps are merged together at the fully connected layer to improve the accuracy of recognition. Through design experiments, different AlexNet_Branch networks composed of different layers of branch convolutional neural networks were built, and the network was trained and tested on the chest X-ray image set respectively. The results show that the addition of a branch convolutional neural network greatly improves the accuracy of pneumonia recognition, and the AlexNet_Branch network test accuracy consisting of a 16-layer branch convolutional neural network is 98.01%.

Keywords: Convolutional Neural Network, Pneumonia Image Recognition, Feature Map Fusion

1. Introduction
Pneumonia is a serious respiratory infection. Pneumonia can cause lung tissue necrosis and lead to lung abscess. If the infection further expands, it will cause various specific complications and make people difficult to breathe [1]. At present, pneumonia can be diagnosed based on clinical manifestations and imaging examinations [2]. Among them, imaging examinations are mainly chest X-rays. Diagnosing pneumonia through X-rays requires doctors to have corresponding knowledge of pneumonia recognition and high X-ray interpretation and judgment ability. In 1966, Computer Aided Diagnosis [3]
was proposed, which uses imaging, medical image processing technology, and other means combined with computer analysis to improve the accuracy of diagnosis and shorten the time of diagnosis. The use of computer-aided diagnosis technology in the diagnosis of pneumonia can reduce misdiagnosis, missed diagnosis, shorten diagnosis time, and improve diagnostic accuracy.

With the development of artificial intelligence, artificial intelligence has also made major breakthroughs in medical diagnosis. Many domestic and foreign researchers have carried out research in the field of pneumonia image analysis. Li Yang et al. proposed a hybrid kernel SVM method, which combines Gaussian and polynomial kernel functions to identify lung diseases. Although the SVM algorithm saves memory and has advantages in processing high-dimensional samples when the number of samples is too large, it will consume a lot of machine memory and calculation time, and because the data set is non-linear, the SVM algorithm will have an overfitting problem [4]. With the development of machine learning, Deep Learning represented by Convolutional Neural Networks (CNN) is widely used in various fields, including the application of CNN in the medical field. Xiang Wenbo et al. used a CNN composed of three convolutional layers, three pooling layers, and a fully connected layer to identify pneumonia, and the recognition accuracy reached 86.75% compared to the 81.4% accuracy of the SVM algorithm, the accuracy rate is significantly improved. Daniel S. Kermany et al. [5] used the transfer learning method to train the network using the Optical Coherence Tomography data set and then micro-call the network to identify pneumonia, with an accuracy rate of 92.80%. Compared with SVM and simple CNN algorithm, this method has significantly improved accuracy. However, due to the large difference between the data set used in migration learning and the pneumonia data set, the improvement of network accuracy is limited.

Therefore, this paper proposes an improved network structure AlexNet_Branch network based on AlexNet [6] structure, which is used for X-ray image recognition of pneumonia. The network structure proposed in this paper mainly improves AlexNet through the following aspects: (1) Use convolution kernels of different sizes to extract features, and fuse different features to improve classification accuracy, (2) Use Dropout layer [7] to prevent overfitting, (3) Add Batch Normalization (BN) [8] to all convolutional layers except the fully connected layer, to improves the accuracy of image recognition.

2. Pneumonia recognition model

In the 2012 ImageNet competition, AlexNet, designed by Hinton and Alex Krizhevsky, won the championship. The basic structure of AlexNet consists of 5 layers of convolutional layers and 3 layers of fully connected layers. The activation function uses ReLU and the Dropout layer is used for training. The basic structure of AlexNet is shown in Fig.1.

![Fig.1](image.png)

*Fig.1* The basic structure of AlexNet, conv represents the convolution kernel, FC represents the fully connected layer
As shown in Figure 1, AlexNet is composed of five convolutional layers and three fully connected layers. There are convolution kernels of size 11×11 and 5×5 in the convolution layer. Such a large-size convolution kernel is a relatively 3×3 convolution kernel that has the characteristics of a large number of parameters, complex calculations, and nonlinear differences [9].

Aiming at the characteristics of AlexNet's shallower network, large convolution kernel size, and single feature image obtained, this paper designs the AlexNet_Branch network. The structure of AlexNet_Branch is shown in Fig.2.

![Fig.2 The basic structure of the AlexNet_Branch network](image)

The AlexNet_Branch network is mainly composed of a branch network, a convolutional layer of AlexNet, and a three-layer fully connected layer, as shown in Figure 2. The branch network is composed of a multi-layer convolutional layer with a size of 3×3. Compared with the AlexNet convolutional layer, the branch network has a smaller size and a deeper network. Therefore, the branch network has a small amount of calculation, fast speed, and can get more abstract features. Combining different feature maps before the fully connected layer can get richer feature maps, which is beneficial to improve the image recognition rate [10].

The pneumonia image with a size of 227×227 is input into AlexNet_Branch, and two different feature maps are obtained after the convolution of the convolutional layer. The feature maps are fused, and then input into the fully connected layer, and finally recognized by the softmax layer. The AlexNet_Branch network uses the dropout method during training to prevent overfitting, uses the BN layer to regularize the data, uses Stochastic gradient descent (SGD) to optimize the loss function, and branches the convolutional neural network with a 3×3 volume It consists of multi-layer, maximum pooling layer, and fully connected layer. The output result expression of the convolutional layer is shown in formula (1).

\[
X_n^l = f \left( \sum_{i \in M} W_i^l \ast X_{n}^{l-1} + b_i^l \right)
\]  (1)
In formula (1), $X_n^L$ represents the nth feature map of the Lth layer of the convolutional layer, $M$ represents the feature image set, $W^L_i$ represents the i-th weight of the Lth layer, the symbol * represents convolution, $b^L_i$ represents the i-th bias of the L-th layer of the convolutional layer, and the function $f(\cdot)$ is the activation function. This paper uses the non-linear function ReLU function as the activation function. Compared with the sigmoid function and the tanh function, it has less calculation and overcomes the disappearance of the gradient, and alleviates the advantages of over-fitting, the mathematical expression of the ReLU function is shown in formula (2)

$$f(z) = \max(0, z)$$

(2)

The feature image output from the convolutional layer enters the pooling layer. The role of the pooling layer is to remove redundant information, compress the feature image, simplify the complexity of the network, and achieve nonlinearity. The output expression of the pooling layer is shown in formula (3)

$$X^i_i = down\left( X^{i-1}_i \right)$$

(3)

In formula (3), $X^i_i$ represents the i-th feature image of the L layer of the convolutional layer, down (•) represents downsampling, and the pooling layer in the AlexNet_n network uses maximum pooling.

In order to prevent overfitting and improve accuracy during training, a dropout structure is added to the AlexNet_Branch network. When training the AlexNet_Branch network, we use the mini-batch SGD algorithm. Due to the difference in data distribution in the batch, the weights obtained after training will fluctuate in a large range, resulting in a decrease in the accuracy of the model [8]. So in order to improve the accuracy, the BN layer is used in the AlexNet_N network. The BN layer is implemented as

$$\mu_B = \frac{1}{m} \sum_{i=1}^{m} x_i$$

(4)

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^{m} (x_i - \mu_B)^2$$

(5)

$$x^* = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$$

(6)

$$y_i = y^*{x^*} + \beta = BN_{\gamma, \beta}(x_i)$$

(7)
Equation (4) is to average the eigenvalues in a batch, where m is the number of eigenvalues in a batch and $\chi_i$ is $i$-th eigenvalue in a batch. Equation (5) is to calculate the variance of the eigenvalues in a batch. Equation (6) is to normalize the input data. Since the input data is forcibly normalized, it will destroy the features learned in the previous layer. Therefore, the learnable reconstruction parameters $\gamma, \beta$ are introduced to let the eigenvalues after normalization are restored to be close to the original distribution.

3. Experimental settings and results

3.1. Data set

The pneumonia image dataset in this article comes from the 2018 Kaggle competition dataset Chest X-Ray Images (Pneumonia). The chest X-ray images (front-rear) in this dataset are patients from Guangzhou Medical Center in Guangzhou. The data set has 5863 chest X-ray images, including 4,273 confirmed pneumonia X-ray images, and 1590 normal chest X-ray images. Some images in the pneumonia dataset are shown in Fig.3.

![Partial images of pneumonia image data set](image)

**Fig.3** Partial images of pneumonia image data set, (a) Chest X-ray picture of pneumonia, (b) Normal chest X-ray image

3.2. AlexNet/Branch network

In AlexNet/Branch, the number of convolutional layers of the branch network is different, the extracted features are different, and the corresponding network effects are different. In the design of this pneumonia recognition network, we must first determine the optimal N value in AlexNet/Branch used for pneumonia recognition, which is the number of convolutional layers of the branch network.

In the experiment, the dropout layer and the BN layer were added to the AlexNet and AlexNet/Branch networks. The value of N in AlexNet/Branch was 10, 13, 16, and 19 respectively. The image input size during training is 227×227, the dropout value is 0.5, the learning rate is $1\times10^{-4}$, and the stochastic gradient descent algorithm is used to update the weights. Each batch has a value of 32, and each batch is an iteration after training. After the training starts, a model is output every 20 iterations and
tested on the test set. AlexNet and AlexNet_Branch were trained on the training set 8,000 times, and the test results are shown in Tab.1.

| Number of iterations | 2K    | 4K    | 6K    | 8K    |
|----------------------|-------|-------|-------|-------|
| AlexNet              | 75.61%| 75.42%| 82.18%| 91.14%|
| N=10                 | 94.01%| 95.54%| 95.31%| 96.44%|
| N=13                 | 94.21%| 94.83%| 95.64%| 96.72%|
| N=16                 | 95.12%| 95.34%| 97.48%| 97.84%|
| N=19                 | 95.21%| 95.98%| 97.23%| 97.88%|

Tab.1 The accuracy of different networks under different iteration times

In Tab.1, AlexNet represents only the network structure of AlexNet, and N=10, 13, 16, 19 represents the AlexNet_Branch network structure, where the value of N represents the number of branches convolutional layers. It can be seen from Table 1 that when AlexNet is tested after 2,000 iterations, the accuracy is 75.61%, and after 8,000 iterations, the accuracy can reach 91.14%. The AlexNet_Branch network composed of 10 branch convolutional layers has an accuracy rate of 94.01% after 2,000 iterations. Compared with AlexNet after 2,000 iterations, the accuracy rate has increased by 18.4%. After 8,000 iterations, the accuracy rate is 96.44%. After 8,000 iterations of AlexNet, the accuracy rate increased by 5.3%, indicating that the addition of the branch network made the test accuracy of AlexNet_Branch improved compared to AlexNet. Because the branch network is composed of a multi-layer 3×3 convolutional layer, the deeper the network depth, the more abstract the feature image, and the feature map information after the AlexNet convolution will get more feature map information, thus improving the recognition of pneumonia The accuracy rate.

In the AlexNet and AlexNet_Branch networks, the Dropout layer and BN layer are added to improve the accuracy of the network. The influence of the Dropout layer, BN layer, and branch network on the accuracy of the network will be determined through experiments. Design four types of networks: AlexNet_Branch, AlexNet_Branch(a), AlexNet_Branch(b), and AlexNet_Branch(c) to compare the influence of branch network, Dropout layer, and BN layer on the accuracy of pneumonia recognition. Among them, AlexNet_Branch adds the Dropout layer and the BN layer. And the branch network composed of 16-layer branch convolutional layer, AlexNet_Branch(a) only added the Dropout layer and BN layer, AlexNet_Branch(b) only added the BN layer and the branch composed of 16-layer branch convolutional layer Network, AlexNet_Branch(c) only added the Dropout layer and the branch network composed of 16 branch convolutional layers. The result is shown in Fig.4.
Fig 4. The relationship between branch network, Dropout layer and BN layer and accuracy in the AlexNet_Branch network

After 14,000 iterations, the highest accuracy of AlexNet_Branch is 98.01%, the highest accuracy of AlexNet_Branch(a) is 94.10%, the highest accuracy of AlexNet_Branch(b) is 98.05%, and the accuracy of AlexNet_Branch(c) is 97.56%. Compared with AlexNet_Branch, the accuracy of AlexNet_Branch(a) is 3.91% lower, that is, the addition of the branch network improves the recognition accuracy. Compared with AlexNet_Branch(c), the accuracy of AlexNet_Branch is 0.45% higher, that is, the addition of the BN layer improves the accuracy but not much. It can be seen from Fig.4 that AlexNet_Branch, AlexNet_Branch(b), and AlexNet_Branch(c) converge faster than AlexNet_Branch(a), and the accuracy of AlexNet_Branch(a) is lower than other networks, namely branch networks, under the same number of iterations. The addition of the convergence speed of the network is accelerated and the accuracy is also improved. Since the addition of the branch network makes the network wider and deeper, the convergence speed will be accelerated. Because the convolution size and depth in the branch network are different from those in AlexNet Therefore, the convolution size and depth will obtain different feature maps, which is beneficial to the improvement of accuracy.

4. Conclusion
Aiming at the problem of low recognition accuracy in computer recognition of pneumonia X-ray images, this paper proposes a feature fusion convolutional neural network AlexNet_Branch based on AlexNet, which can extract different feature maps from the same image and then classify them. The basic process of the AlexNet_Branch network is to design a network consisting of a branch network and an AlexNet convolutional layer in parallel. Perform feature extraction on the same pneumonia image to obtain two different feature maps, which are merged together before the fully connected layer, the fully connected layer then performs classification and identification. AlexNet_Branch, composed of 16 branch
convolutional layers, has an accuracy of 98.01% for pneumonia image recognition. In future research work, AlexNet_Branch will be further optimized to continue to improve accuracy and reduce training time.

References
[1] McLuckie A 2009 Respiratory disease and its management (London: Springer)
[2] Hongli X, Zongli D and Suxia M 2009 Pulmonary manifestations of Mycoplasma pneumoniae pneumonia. *J. Chinese Medical Journal* **11** pp 14-16
[3] Gilbert F J and Lemke H 2005 Computer-aided diagnosis Br. *J. Radiol* **78**(suppl_1) pp 1-2
[4] Yang L, Dunwei W and Ki W 2013 Mixed kernel function SVM for pulmonary nodule recognition *Tsinghua Sci. Technol* **18**(1) pp 449-458
[5] Kermany D S, Goldbaum M and Cai W 2018 Identifying medical diagnoses and treatable diseases by image-based deep learning Cell **172**(5) pp 1122-1131
[6] Krizhevsky A, Sutskever I and Hinton G 2012 ImageNet classification with deep convolutional neural networks. NIPS. *Curran Associates Inc.* (Lake Tahoe)
[7] Hinton G, Srivastava N, Krizhevsky A and Sutskever I 2012 Improving neural networks by preventing coadaptation of feature detectors. *Computer ence* **3** pp 212-223
[8] Gao H, Zhuang L and Laurens VDM 2017 Densely connected convolutional networks *CVPR* (Hawaii)
[9] Pham V, Bluche T and Kermorvant C 2014 Dropout improves recurrent neural networks for handwriting recognition *ICFHR 14th* (Greece)
[10] Ioffe S and Szegedy C 2015 Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*