*Computer Methods in Medicine and Health Care*
*G. Casalino and A. Nayyar (Eds.)*


*doi:10.3233/ATDE220552*

# Application of Convolutional Neural Network in Lung Cancer Pathological Image Recognition

Kezhi ZHANG[a,1], Guoqiang WEI[a,1], Shengbang ZHOU[a], Qinghua MENG[a], Ge ZHAO[a], Jianbo LI[a], Enshuang GAO[b,2] and Ze FENG[c,2]
[a] *School of Physics and Electronics, Nanning Normal University, Nanning 530001, Guangxi, China*
[b] *School of Environment and Life Science, Nanning Normal University, Nanning 530001, Guangxi, China*
[c] *Guangxi Vocational Institute of Technical School, Chongzuo 532200, Guangxi, China*

**Abstract.** Lung cancer is a high incidence disease, which seriously affects people's health. The pathological section of lung cancer can determine the type and differentiation of lung cancer cells, so as to provide an important basis for the selecting treatment options. In recent years, researchers focus on using convolutional neural network (CNN) algorithm to assist doctors and improve the recognition of cancerous regions in pathological images. In this paper, the CNN models were used to identify the cancerous region of lung cancer pathological images. The public data set was selected to train the AlexNet, GoogLeNet and ResNet34 models, and adjust the relevant parameters to improve the accuracy and specificity of recognition as much as possible. The experimental results showed that the accuracy and specificity of ResNet34 model were 98.9% and 99.0%, respectively, indicating that the model could effectively assist doctors to identify cancerous regions in lung cancer pathological images.

**Keywords.** Convolutional neural network, lung cancer, pathological images, ResNet34

## 1. Introduction

According to the statistics of International Agency for Research on Cancer, there are about 1930 million new cancer cases and nearly 10 million cancer deaths worldwide in 2020. Lung cancer accounts for 11.4% of new cancer cases, ranking second, while it accounts for 18% of cancer deaths, ranking first [1]. This shows that lung cancer seriously affects people's health. The overall 5-year relative survival rate of diagnosed lung cancer patients is 19% [1]. Therefore, early identification and diagnosis are of great significance. The diagnosis of lung cancer in modern medicine can be divided into

---

[1] Kezhi Zhang and Guoqiang Wei contributed equally to this work.
[2] Corresponding Author, Enshuang GAO, School of Environment and Life Science, Nanning Normal University, Nanning 530001, Guangxi, China; Email: gaoenshuang@163.com. Corresponding Author, Ze FENG, Guangxi Vocational Institute of Technical School, Chongzuo 532200, Guangxi, China; Email: 1789459645@qq.com.

analytical diagnosis and pathological diagnosis, and the latter is called the 'gold standard' for lung cancer diagnosis because of its objectivity and accuracy.

In recent years, artificial intelligence technology has developed rapidly, and more and more research has started to penetrate the healthcare field. Computer vision aided detection using deep learning is a research hotspot. In 2012, the famous scholar Krizhevsky [2] proposed the AlexNet network, which further improved the performance of CNN in image recognition tasks. Although its overall framework is similar with LeNet-5, it deepened the hierarchy, along with the inclusion of ReLu [3] activation function and the Dropout method [4] to avoid overfitting. So, it had achieved good results. In 2017, Atsushi [5] designed an automatic classification model for lung cancer cell pathology by deep convolution neural network. The CNN model composed of three convolution layers, three pooling layers and two fully connected layers. It had an accuracy of 89.0%, 60.0%, and 70.3% in the classification and diagnosis of adenocarcinoma, squamous carcinoma, and small cell carcinoma, respectively, with an overall accuracy of 71.1%, and the results were comparable to those of pathologists.

In 2018, Li [6] applied CNN models to lung cancer histopathology images and compared the performance of SqueezeNet, ResNet50, AlexNet, and VGG16 models for predicting malignant tumors on image blocks of $256 \times 256$ size. The accuracy rate can reach 88.1% to 91.19%, and AlexNet has achieved the best results. This study shows that CNN model is helpful for rapid tumor detection and is expected to provide auxiliary diagnosis and decision support for lung cancer. In the same year, Coudray [7] used the Inception V3 model and migration learning to automatically classify lung cancer pathology images into normal tissue, squamous cell carcinoma, and adenocarcinoma based on image morphological features. It shows considerable high performance compared with the diagnosis of pathologists. In 2019, the Google Brain team proposed the Efficient model [8], which proposes dimensions of depth, width, and resolution scaled proportionally. Compared with other networks, the network has not only improved greatly in speed, but also achieved high-precision experimental results. In 2020, Wang [9] proposed CNN pathological image classification based on hard instance guidance, using a re-weighting training algorithm to improve learning accuracy by increasing the weights of difficult examples. In 2021, Liu [10] proposed breast cancer pathological image classification based on VGG16 feature connection, and obtained excellent performance.

In conclusion, CNN model can be used in lung cancer pathological image recognition. However, the accuracy and specificity of CNN model for lung cancer pathology image recognition need to be further improved. Therefore, in the paper we expect to find a CNN model, so that its recognition of lung cancer pathological images can ensure high accuracy and high specificity.

## 2. Dataset

The dataset used in this paper was a part of the dataset of “Jiangsu Big Data Development and Application Competition - Intelligent Diagnosis of Cancer Risk” in 2020, with a total of 1770 images, each of which was annotated by experienced doctors. Each image is accurately labeled by experienced doctors. Examples of the images were shown in Figure 1.

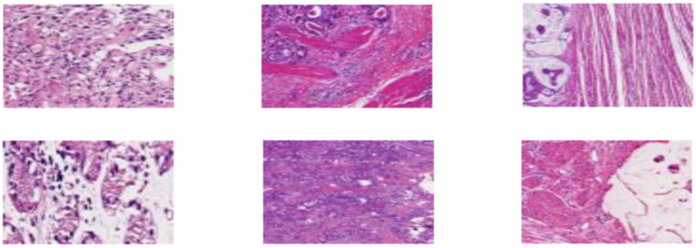

**Figure 1.** Examples of lung cancer pathology images.

There are the image masks along with the image dataset. The white part of the mask corresponds to the cancerous area, and the black part corresponds to the normal area. The cancerous region and normal region in the dataset images are separated, and the cancer images and normal images were obtained. After separation, these pathological images of lung cancer were divided into three categories, one of which is normal image marked as '0' and the other two are different cancerous modalities. The squamous cell carcinoma image was marked as ‘1’, and the adenocarcinoma image was marked as ‘2’.

## 3. Methods

In the paper, AlexNet, GoogLeNet and ResNet were selected according to the data characteristics and detection targets.

### 3.1. AlexNet Model

The AlexNet model has good image recognition ability. It has eight weighted layers, including five convolution layers and three fully connected layers, and multiple convolution cores of different sizes. Therefore, it can achieve deeper learning. The network structure is shown in Figure 2. It has better performance than LeNet model by using data enhancement and Drop-Out method [4] to suppress overfitting, using ReLU activation function [3] to reduce gradient disappearance, and using GPU parallel computing.

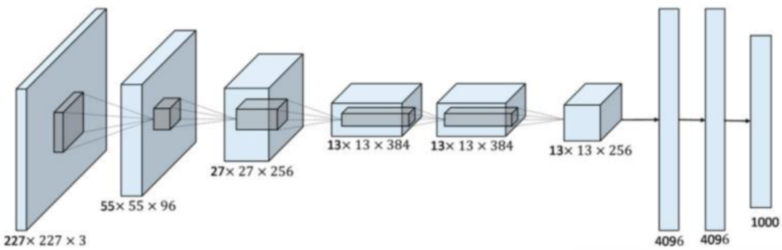


**Figure 2.** AlexNet structure diagram.

### 3.2. GoogLeNet Model

The Inception structure uses different sized convolutional kernels for the synthesis of the previous layer of the network to extract features and combines maximum pooling for feature fusion The Inception structure is shown in Figure 3 [7]. To avoid causing feature map redundancy, a 1x1 convolution is added before 3x3, before 5x5, and after max pooling, respectively. Nine Inception modules, five pooling layers, and two convolutional layers are used in the GoogLeNet model.

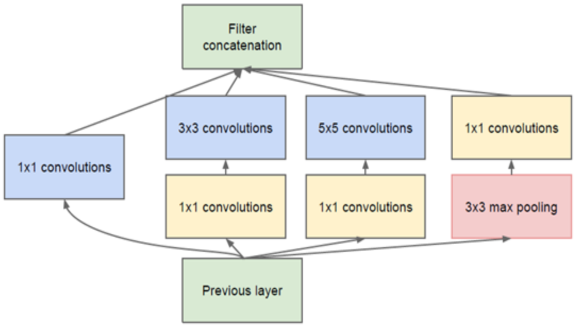


**Figure 3.** Inception structure diagram.

### 3.3. ResNet34 Model

The main idea of ResNet is to add a direct connection channel in the network, allowing the original input information to be directly transmitted to the later layer, while retaining a certain proportion of the output of the previous network layer. ResNet protects the integrity of the information by directly bypassing the input information to the output, and the whole network only needs to learn the part of the input and output residuals, simplifying the learning objectives, difficulty and complexity. There are many kinds of ResNet, mainly due to the difference in the number of layers. In this paper, the 34-layer ResNet model is adopted, and the corresponding BasicBlock structure is shown in Figure 4.

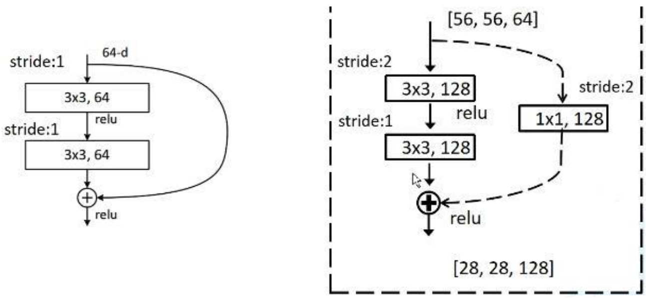


**Figure 4.** BasicBlock structure diagram.

## 4. Experiments

### *4.1. Experimental Parameters and Evaluation Indexes*

batch_size: the number of images imported per batch. epochs: the number of iterations; lr: the learning rate. loss: the average loss rate per iteration.

Also, some evaluation metrics [11] need to be introduced to evaluate the performance of our network quantitatively and objectively after obtaining the experimental results. According to the research content, we reasonably set the statistical data and evaluation indexes required for the experiment, as shown in Table 1 and Table 2, respectively.

**Table 1.** Confusion matrix of experimental statistical data.

| | | Predicted Value | |
|---|---|---|---|
| | | Positive Sample | Negative sample |
| Actual Value | Positive Sample | True Positive (TP) | False Negative (FN) |
| | Negative sample | False Positive (FP) | True Negative (TN) |

**Table 2.** Evaluation indicators.

| Evaluation Metrics | Calculation Formula | Evaluation Meaning |
|---|---|---|
| Precision (P) | $Precision = \frac{TP}{TP+FP}$ | The percentage of all predictive positive samples correctly identified as positive |
| Specificity (SPE) | $Specificity = \frac{TN}{TN+FP}$ | The percentage of all actual negative samples correctly identified as negative |
| Accuracy (Acc) | $Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$ | The percentage of samples with correct identification results among all samples |

### *4.2. AlexNet Model Training and Analysis*

The AlexNet class is defined and the AlexNet model is built. Then, Conv2d, ReLu and Maxpool2d were used to build the feature extraction part, and Dropout and Linear are used to build the full connection layer.

The values of epochs, lr and batch_size were set to 100, 0.000001 and 16, respectively. The model was trained, and the trend of Acc and loss with epochs on the validation set was obtained and shown in Figure 5. As the number of epochs increases, the value of Acc increases rapidly, and after 40 iterations, the increase is slow, while the value of loss decreases rapidly. When epochs reach 100, the value of Acc still has more room for improvement, while the value of loss has more room for decrease. Therefore, the values of lr and epochs were adjusted to 0.000006 and 150, respectively. The new trend was obtained, as shown in Figure 6. After 120 iterations, the value of Acc is stable at about 95.6%, and the value of loss is stable at about 0.12, indicating the AlexNet model has good generalization ability.

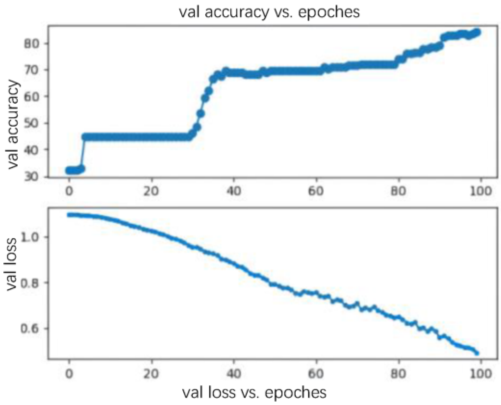


**Figure 5.** Trends of acc and loss on validation set before parameter optimization in AlexNet.

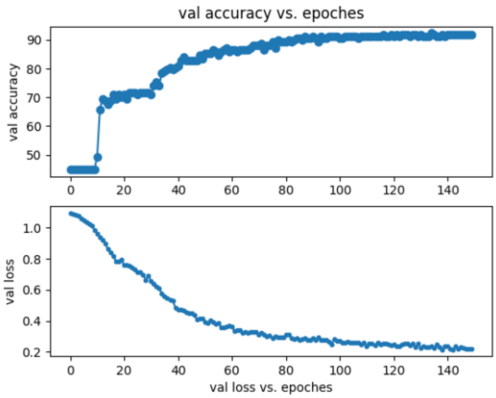


**Figure 6.** Trends of acc and loss on validation set after parameter optimization in AlexNet.

### *4.3. GoogLeNet Model Training and Analysis*

In the average pooling layer, an adaptive average pooling sampling operation was adopted, without limiting the size of the input image. Setting the parameters, the values of epochs, lr and batch_size were 100, 0.000001 and 8. After training the model, the trends of Acc and loss with epochs on the verification set were obtained, as shown in Figure 7. With the increase of the number of iterations, the value of Acc increased rapidly, increased slowly after 60 times, and stabilized at about 92.8% after 100 times; while the value of loss decreases rapidly, and it decreases very slowly after 10 times. After 80 iterations, it basically remains stable, and its value is around 0.32.

The values of epochs and lr were increased to 150 and 0.00003, respectively. The new trend graphs were shown in Figure 8. After 100 times, the value of Acc stabilized at about 97.5%, while the value of loss stabilized at about 0.10, indicating the GoogLeNet model also has good generalization ability.

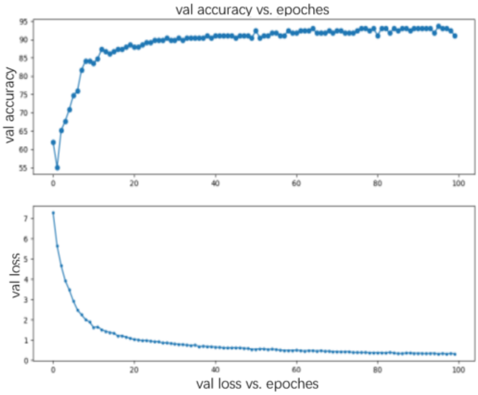


**Figure 7.** Trends of acc and loss on validation set before parameter optimization in GoogLeNet.

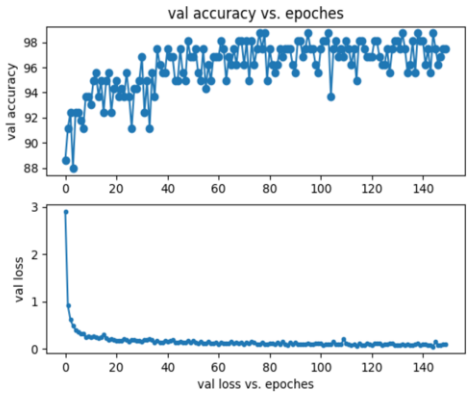


**Figure 8.** Trends of acc and loss on validation set after parameter optimization in GoogLeNet.

### *4.4. ResNet34 Model Training and Analysis*

The ResNet34 model was built up according to the idea of 3.3. The model was trained by setting lr to 0.0001, batch_size to 8, epochs to 50. The trend of Acc and loss with epochs on the verification set were shown in Figure 9. It is obvious that Acc quickly reaches a high value and then keeps fluctuating around a certain value, while the loss decreases in a sawtooth pattern. Therefore, the values of lr, epochs and batch_size were adjusted to 0.000005, 150 and 16, respectively. The new trends were obtained, as shown in Figure 10. After 120 times, the value of Acc stabilizes at 99.4%, while the value of loss stabilized at 0.029, indicating the ResNet34 model has very good generalization ability.

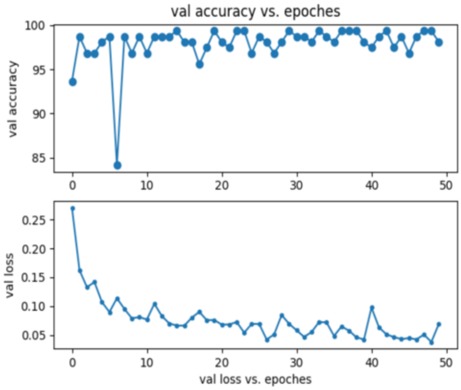


**Figure 9.** Trends of acc and loss on validation set before parameter optimization in ResNet34.

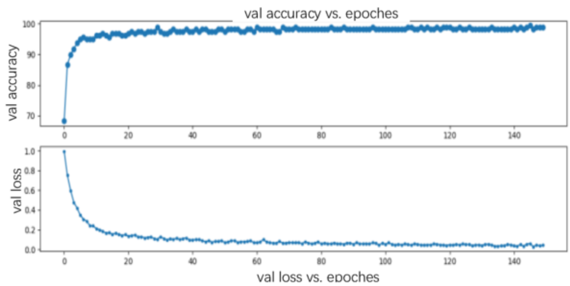


**Figure 10.** Trends of acc and loss on validation set after parameter optimization in ResNet34.

### *4.5. Evaluation and Comparative Analysis of the Models*

In order to preferably evaluate and compare the three models, them were tested under the optimal parameters. The relevant evaluation metrics of the tests are included in Table 3. It can be seen that the three models have very good recognition for non-cancerous images. For the recognition of two types of lung cancer images, ResNet34 model is much higher than AlexNet model and GoogLeNet model. The accuracy and specificity of ResNet34 model were 98.9% and 99.0%, respectively, indicating that the model has the potential to assist doctors in identifying cancerous regions in lung cancer pathological images.

**Table 3.** Comparison of evaluation index results of three models.

| | AlexNet | | | GoogLeNet | | | ResNet34 | | |
|---|---|---|---|---|---|---|---|---|---|
| | P | SPE | Acc | P | SPE | Acc | P | SPE | Acc |
| 0 | 1.0 | 1.0 | | 1.0 | 1.0 | | 1.0 | 1.0 | |
| 1 | 0.875 | 0.934 | 0.936 | 0.840 | 0.971 | 0.942 | 0.975 | 0.993 | 0.989 |
| 2 | 0.802 | 0.774 | | 0.823 | 0.768 | | 0.987 | 0.990 | |

In this paper, three models of AlexNet, GoogLeNet and ResNet were used to carry out experiments on lung cancer pathological image recognition. The three models were trained on the public dataset, and the parameters were constantly adjusted to make them have higher accuracy and better generalization ability. Finally, the models were evaluated. The diagnosis of pathological image recognition based on CNN in medical image processing is an important topic. In terms of the current development status, the research in this area still needs to continue to be explored and discovered. Of course, the experiments in this paper have shortcomings, for example, the number of images in the dataset is not large enough, the extensiveness of dataset is not high, and the categories are too single. Therefore, this experiment is only a simple test of the feasibility of CNN in lung cancer pathological image recognition, and the real realization needs the concerted efforts of all elites.

## Acknowledgements

This research was funded by Guangxi Science and Technology Base and Talent Special Project (No. AD22080022), Ministry of Science and Technology "Science and Technology for Economy 2020" Key Project (No. 17) and Project of Improving the Basic Ability of Scientific Research of Young and Middle-aged Teachers in Guangxi Colleges and Universities (No. 2020KY09028).

## References

[1] Sung, H, Ferlay J, Siegel R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2021;71(3):209-249.

[2] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems. 2012;25(2).

[3] Nair V, Hinton G E. Rectified linear units improve restricted Boltzmann machines. Proc of 27th International Conference on Machine Learning. Haifa: Omnipress, 2010;807-814.

[4] Hinton G E, Srivastava N, Krizhevsky A, et al. Improving neural networks by preventing co-adaptation of feature detectors. Computer Science. 2012;3(4):212-223.

[5] Atsushi T, Tetsuya T, Yuka K, et al. Automated classification of lung cancer types from cytological images using deep convolutional neural networks. BioMed Research International. 2017;2017:1-6.

[6] Zhang L, Hu Z Y, Xu J L, et al. Computer-aided diagnosis of lung carcinoma using deep learning-a pilot study. arXiv preprint arXiv:1803.05471. 2018.

[7] Coudray N, Ocampo P S, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nature Medicine. 2018;24(10):1559-1567.

[8] Tan M X, Le Q V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. 2019.

[9] Wang Y, Peng T, Duan J, et al. Pathological image classification based on hard example guided CNN. IEEE Access, 2020;(99):1-1.

[10] Liu M, Yi M, Wu, M H, et al. Breast pathological image classification based on VGG16 feature concatenation. Journal of Shanghai Jiaotong University (Science). 2021;1-12.

[11] Hossin M, Sulaiman M N. A review on evaluation metrics for data classification evaluations. International Journal of Data Mining & Knowledge Management Process.2015;5(2):01-11.