X-ray Image Recognition of Pneumonia Based on Three Different Neural Networks

Haoran Xiong*

Guangdong University of Technology School of Automation, Guangdong, Guangzhou, 510006, China

*Corresponding author: 3121006856@mail2.gdut.edu.cn

Abstract. With the outbreak of new coronary pneumonia, a variety of pneumonia diseases are emerging, doctors often have misdiagnoses and omissions. To assist doctors in better clinical diagnosis, a variety of high-performance neural network structures have been applied to the X-ray recognition of pneumonia, and the X-ray image recognition accuracy will be improved, will effectively improve the efficiency of the hospital detection, at the same time, can also avoid the patient to miss the best rescue time. To this end, this paper compares and analyses the recognition performance of the three currently used neural network structures in the recognition of pneumonia X-ray images, uses some publicly available datasets and a mixed set composed of different datasets to conduct experiments, and summarises the advantages of the models and the direction of possible improvement. The three neural network models are presented, compared, and analysed to suggest useful references for the recognition of pneumonia X-ray images. Finally, an analysis and outlook for future neural network models for pneumonia x-ray image recognition is presented.

Keywords: Pneumonia x-ray; neural networks; image recognition.

1. Introduction

Pneumonia is a highly contagious respiratory disease. Since the end of 2019, with the global spread of the neo-collateral pneumonia pandemic and the emergence of various types of pneumonia, such as mycoplasma pneumonia, viral pneumonia, bacterial pneumonia, etc., the contagiousness and complexity of pneumonia, as a serious respiratory disease, has increasingly attracted global health attention. These pneumonias are difficult to differentiate based on symptom presentation. To improve the accuracy of clinical diagnosis and prevent missed treatment due to misdiagnosis, x-ray image recognition has been introduced into the diagnosis of pneumonia, it also effectively alleviates the problem of lack of experience in hospital manpower, while compared to the naked eye observation of the treatment has also played a very important role in assisting the doctor's treatment, but the accuracy of the x-ray image recognition needs to be further improved, and the image analysis to provide a more accurate and efficient recognition will be of great significance. In which the algorithmic model of image recognition plays a very vital role in improving the accuracy. Currently, common algorithms in the field of recognition and classification include Convolutional neural network (CNN) frameworks such as AlexNet, VGGNet, ResNet, Inception, and MobileNet [1]. Ibrahim et al. studied four different architectures for detecting and diagnosing diseases affecting the human lungs. Among all the models used in their study, CNN and VGG19 models performed the best with 98.05% accuracy. Using GRU + ResNet152V2, an accuracy of 96.09% was achieved [2]. In addition, an article on "Pneumonia x-ray image recognition algorithm based on improved DenseNet network" highlighted the great potential of DenseNet in this field [3]. However, there is a research gap on the difference in performance of different neural network structures in the recognition task of pneumonia x-ray images.

Based on this, to better optimise the neural network structure for pneumonia X-ray image recognition, this paper aims to analyse the performance of three neural network structures based on DenseNet, ResNet, and Visual Geometry Group (VGG) for recognising pneumonia-related X-ray images. The comparison of experimental results provides a useful reference in providing an automated recognition method for pneumonia X-ray images.

2. Introduction to Neural Network Structure

2.1. Origin of CNNs and Introduction

CNN, as a model of deep learning, is mainly composed of a convolutional layer, pooling layer, activation function, and complete connected layer. The convolution layer is usually used to extract the features of the input data, while the pooling layer filters the features on this basis, sends them into the activation function for calculation, and finally gets the output result. The complete connectivity layer is mainly used to connect the output results of convolutional networks with the input structural data, to generate images with rich features. As an important learning method, Convolutional neural networks are widely used in various tasks in the field of computer vision. At the same time, it lays a good foundation for the subsequent problem of image object detection based on classification and regression. The prototype of CNN can be traced back to the 1980's. Yann LeCun et al. proposed the LeNet network in 1989, which contains the basic components of the convolutional layer, pooling layer, and fully connected layer, which is widely used in computer vision tasks, and it also lays the foundation for the deeper and more complex CNN model later on and thus begins the modern road of CNN [4]. Later, Yann LeCun et al. proposed the deeper LeNet5 CNN network with multiple convolutional and pooling layers, which can be used for more complex image recognition tasks. In 2012, Alex Krizhevsky et al. proposed AlexNet, which achieved remarkable success and marked the rise of deep learning, with the rise of deep learning, CNNs have made a series of breakthroughs in image processing, computer vision, and other fields [4]. In 2014, a variety of advanced frameworks appeared, including AlexNet's improved version of ZFNet which adjusts the convolution kernel of the first convolutional layer with the paraplegic step size to increase the feature extraction capability; Christian's GoogLeNet makes use of 1x1 convolution for channel dimensionality reduction, which solves the problem of parameter proliferation due to network deepening; Simonyan k's VGGNet has good migration learning ability, and its 3x3 convolution has become the standard for later convolutional neural network structures. A variety of network structures continue to deepen, and different frameworks are also optimised and improved, and some models have appeared to improve network performance and training effect by introducing different structures and connections. For example, DenseNet, ResNet, VGG, Inception, mobilenet, GoogleNet, and so on.

2.2. Inception-ResNet-v2 Neural Network

He Ming's team found in the study that the degradation problem that has existed before in deep convolutional neural networks can be solved by ResNet, a novel deep convolutional neural network they proposed, which can effectively solve the degradation problem of deep convolutional neural networks that has always existed and has not been solved, and its residual connections allow skipping information directly from the previous layer to the subsequent layers, which leads to a deeper network that is easier to train [5]. The first Inception network appeared after GoogleNet was proposed by Google in 2014. There are four iterations of the Inception network up to now. By now there are four versions in iteration. Inception-ResNet network is to introduce the residual structure of ResNet in the Inception module in the iteration process. The ResNet structure in the neural network can both accelerate training speed and improve performance in the whole process, while the Inception module can obtain sparse or non-sparse features of the same layer [6]. Of the two versions in the Inception-ResNet network, Inception-ResNet-v2 in the structure of Fig. 1 better integrates the advantages of Inception and ResNet to achieve better performance.

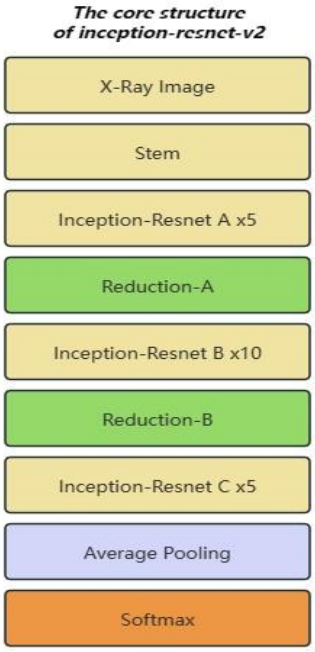


Fig. 1 Inception-ResNet-v2 [7]

2.3. DenseNet-121 Neural Network

Later, Huang Gao and others at Cornell University proposed the densely connected convolutional neural network DenseNet, which is a bolder densely connected network than ResNet. To improve the model's anti-fitting ability, DenseNet uses an architecture to connect the data between the layers so that the data between the layers can be transferred to the maximum extent. Unlike other networks, DenseNet contains only a small number of training samples, so the network can output narrower and fewer parameters. DenseNet improves the efficiency of feature maps and gradient transformation, which greatly improves the learning efficiency of the network. Traditional deep networks transmit inputs and gradients hierarchically, that is, only the upper layer is optimized, which leads to gradient loss, while DenseNet overcomes the problem of gradient loss by associating inputs at each level with loss functions. DenseNet's dense connectivity mechanism allows each layer to have direct access to the features of all the previous layers, which facilitates the flow of information and gradient Propagation. This structure achieves significant improvements in parameter efficiency and model performance. Among them, are multiple dense blocks in DenseNet-121 (121 layers deep), as shown in Figure 2, each dense block contains multiple small convolution layers, batch normalization and ReLU activation, and a transition layer. These dense blocks are connected to other layers through a connector channel [8].

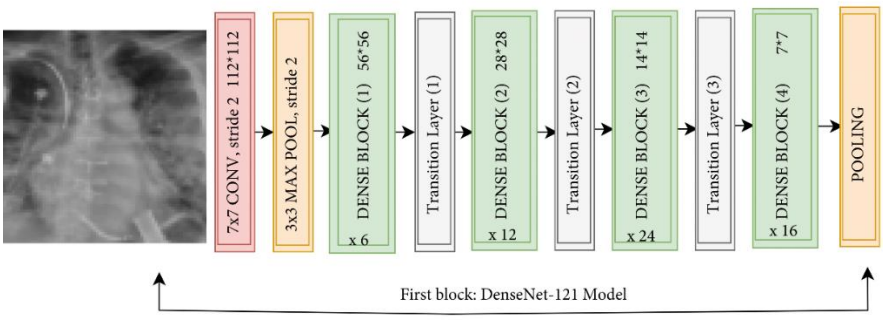


Fig. 2 DenseNet-121 [9]

2.4. VGG-16 Neural Network

In 2014, the runner-up structure of lmageNet, VGGNet, implemented a further improvement of the convolutional neural network, using several consecutive small convolutional kernels instead of a

large convolutional kernel, and its core idea is to improve the performance by increasing the depth of the network, and its structure, as seen in Fig. 3, which contains 13 convolutional layers and three fully-connected layers The VGG-16 network model divides the hierarchical structure into five convolutional blocks, the first two blocks consisting of two convolutional layers and one pooling layer, and the last three blocks consist of three convolutional layers and one pooling layer, which makes the overall structure more simple and standardised and makes the VGGNet one of the network structures with better performance in the history of convolutional neural network development [10].

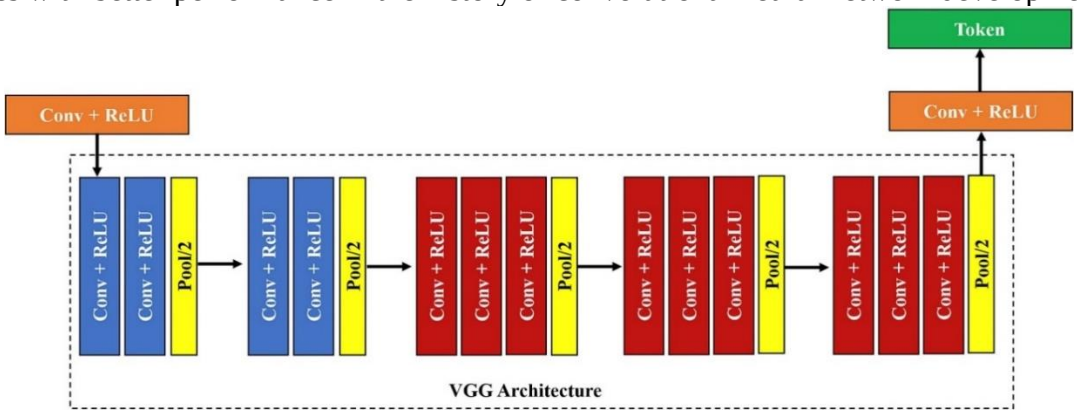


Fig. 3. VGG-16 Architecture [11]

To better complete the image classification task, the VGG-16 network is suitable for the image classification task, while the Inception- ResNet-v2 model has higher performance, and the DenseNet-121 network has a superior structure, which also performs well in the image recognition task. So in this paper, the VGG-16 network, Inception-ResNet-v2 network, and DenseNet-121 network will be used for the comparison and analysis of pneumonia X-ray images.

3. Experiment and Analysis

3.1. Medical Image-Based Profiling

Through the use of x-ray images through the lungs different densities and thicknesses of tissue can be observed. Pneumonia will show up as a patchy shadow of increased density on the chest radiograph, turning an otherwise black area white. In the early stages of novel coronavirus pneumonia, radiographic examination of the lungs may show multiple small patches, mild interstitial changes, and distinct extrapulmonary bands. As the disease progresses and the damage to the lungs worsens, multiple gross glass shadows, infiltrating shadows, and solid changes in the lungs may appear on chest radiographs, giving the appearance of white lungs. Unlike bacterial pneumonia, the lesions of novel coronavirus pneumonia are located closer to the pleura, whereas bacterial pneumonia is more frequently distributed around the bronchi or fine bronchi. In addition, there is a difference in the presentation of the two types of pneumonia on medical images, with novel coronavirus pneumonia showing ground glass-like changes, while bacterial pneumonia mainly shows large solid shadows. The imaging features of different cases of pneumonia are demonstrated by visualisation [10].

3.2. Model Comparison Between Inception-ResNet-v2 and VGG-16

3.2.1. About Datasets

In this paper, we selected 5,856 lung X-ray images published by Kaggle and divided them into two groups: the normal group and the pneumonia group. On this basis, the patients with pneumonia were divided into three test samples, which were 80%, 10% and 10% respectively [2].

3.2.2. Comparison of models

Taking the accuracy of the trained model as the evaluation index, in which the Bacteria and Virus datasets are refined from the pneumonia dataset, the models of Inception-ResNet-v2 and VGG-16 are

obtained through the experiments for the recognition of the pneumonia X-ray images obtained from the data in Table 1. It can be seen that the precision of VGG-16 is better than Inception-Resnet-v2.

Table 1. Comparative analysis of the model performance of Inception-ResNet-v2 and VGG-16 [2]

Network model	Categorical	Accuracy	Precision
Inception-ResNet -v2	Bacteria	79.1%	77. 4%
	Normal		91. 9%
	Virus		67. 2%
VGG-16	Bacteria	79.5%	82. 1%
	Normal		94. 2%
	Virus		61.0%

3.3. Model Comparison Between Inception-ResNet-v2 and VGG-16

3.3.1. About datasets

Because there are too few publicly available datasets on bacterial pneumonia and novel coronavirus pneumonia and the amount of data is not large enough, image sets blended from multiple datasets were used as new datasets [12]. For the comparison of model performance between DenseNet-121 and VGG-16, the COVID-19 radiography database dataset was created using the COVID-19 radiography database dataset created by the researchers from the Qatar University in Doha, Qatar, and the University of Dhaka in Dhaka, Bangladesh, together with the Pakistani and Malaysian COVID-19 radiography database dataset created by the collaborators in collaboration with doctors and a mixed set of various datasets containing COVID-19 images [13].

3.3.2. Comparison of models

Taking the accuracy of the trained model as an evaluation metric, the performance of the models of DenseNet-121 and VGG-16 for pneumonia X-ray image recognition is obtained through experiments with mixed datasets augmenting the amount of data as shown in Table II, and the accuracy of VGG-16 is slightly better than that of DenseNet-121.

Table 2. Comparative analysis of the model performance of DenseNet-121 and VGG-16 [13]

Network model	Dataset	Accuracy	Advantage
DenseNet-121	Mixed datasets	92.00%	Automatic feature extraction
VGG-16	Mixed datasets	97.36%	Data augmentation

3.4. Results

Based on the comparative analysis of the above experimental results, it can be found that VGG-16 performs slightly better than Inception-ResNet -v2 and DenseNet-121 in recognising X-ray images of the lungs. However, VGG-16 does not perform well in the binary classification problem of distinguishing whether a new crown pneumonia disease is present or not, but its high model expression ability, small convolutional kernel, and modular structure still give it an excellent performance in the recognition of pneumonia X-ray images, which makes it possible to use VGG-16 or a neural network structured on VGG-16 for the recognition of pneumonia X-ray images. Optimised structure of the network can be trained for clinical use.

4. Conclusion

The results of the study show that for the three network structures, VGG-16 network, Inception-ResNet -v2 network, and DenseNet-121 network, the VGG-16 network performs the best in the recognition of pneumonia X-ray medical images, so it is more suitable to be applied in the problem of classification of pneumonia X-ray medical images, and the structure can be optimised in the VGG-

16 network in the future to better assist physicians to diagnose the disease. In the future, we can optimise the structure of the VGG-16 network to better assist physicians in diagnosing the disease.

The first is the network structure, the simple neural network network structure has no way to classify the medical images of pneumonia, with the continuous Computer vision development, deep learning in the field of graph classification continues to progress, most of the COVID-19 diagnostic models based on deep learning have achieved significant results, but there is still room for improvement. Meanwhile, unsupervised learning is also the development direction of medical image recognition of pneumonia.

The second dataset aspect is that because the number of publicly available datasets of medical images of new coronary pneumonia is not very large, the machine learning is not very effective, resulting in overfitting and poor robustness of the training results, and therefore the training results are not universally applicable in clinical conditions. Moreover, there is an algorithmic bias, also known as algorithmic discrimination, because some of the datasets taken by the model are focused on children and women, and these datasets will make the model less generalisable for medical diagnostic use and may harm those who have not yet included their data in the training set, so it is very important to select and process the datasets, and in the future, we hope to obtain datasets with a large amount of data to the model.

References

[1] Z. Jing, N. Changrui, Y. Zhiyong. A review of target detection algorithms based on convolutional neural network. Journal of Weapons and Equipment Engineering, 2022, 43(06): 37-47.

[2] Y. Junhao, M. Zhiqing, W. Guohui, et al. Recognition and classification of childhood pneumonia based on improved Inception-ResNet-v2. Advances in Lasers and Optoelectronics, 2023, 60(14): 85-92.

[3] M. Jinlin, Q. Shuo, M. Ziping, et al. A review of deep learning diagnostic methods for novel coronavirus pneumonia. Computer Engineering and Applications, 2022, 58(12): 51-65.

[4] T. Agrawal, P. Choudhary. Segmentation and classification on chest radiography: a systematic survey. the Visual Computer. 2023.

[5] D. Sarwindaa, R. H. Paradisaa, A. Bustamama. Pinkie Anggiab.Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of Colorectal Cancer. Procedia Computer Science, 2021, 3.

[6] J. H. Yang, Z. Q. Ma, G. H. WeiV, et al. Recognition and classification of childhood pneumonia based on improved Inception-ResNet-v2. Advances in Lasers and Optoelectronics, 2023, 60(14): 2.

[7] S. Dash, P. K. Sethy, S. K. Behera. Cervical Transformation Zone Segmentation and Classification based on Improved Inception-ResNet-V2 Using Colposcopy Images. Cancer Informatics, 2023, 22.

[8] W. Rongjian, S. Jianfei. X-ray image recognition algorithm for pneumonia based on improved DenseNet network. Television Technology, 2021, 45(6): 140-143.

[9] Z. G. A. Mekhlafi, E. M. Senan, J. S. Alshudukhi, et al. Hybrid Techniques for Diagnosing Endoscopy Images for Early Detection of Gastrointestinal Disease Based on Fusion Features. International Journal of Intelligent Systems, 2023, 20.

[10] Y. Su. Research on deep learning-based recognition of chest radiographs of new coronary pneumonia. Master's thesis, Central University of Finance and Economics. 2022.

[11] C. Weimin, A. Muhammad, L. Mengyun, et al. A fusion of VGG-16 and ViT models for improving bone tumour classification in computed tomography. 2023.

[12] W. Qiyao, W. Jianqing. Research on deep learning-based CT image recognition of new coronary pneumonia. Information and Computer. 2020.

[13] M. Jinlin, Q. Shuo, M. Ziping, et al. A review of deep learning diagnostic methods for novel coronavirus pneumonia. Computer Engineering and Applications, 2022, 58(12): 51-65.