International Journal of Biology and Life Sciences
ISSN: 2957-9511 | Vol. 4, No. 3, 2023

Research on Classification of ConvNeXt Chest X-ray Images Based on Attention Mechanism

Baoyu Peng [1], Yimin Miao [1], Changyou Fu [1], Yi Wang [2, *]

[1] Sichuan University of Science & Engineering, School of Computer Science & Engineering, Yibin 644000, China
[2] The Fourth People's Hospital of Zigong City, Zigong 643000, China

*** Corresponding author:** Yi Wang

Abstract: The global novel coronavirus pandemic has become the norm, and lung disease is one of the most important diseases at presentation In order to improve the reading efficiency and the accuracy of disease classification, this paper proposes a transfer learning model based on attention mechanism: CS ConvNeXt In the data preprocessing stage, the generalization ability of the model is improved by data augmentation, and in the model learning stage, the ConvNeXt network with attention mechanism is used to study and migrate to the proposed dataset COVIDx, and the model is tested by the COVIDx dataset in the testing stage Experiments show that the CS ConvNeXt transfer learning method can improve the accuracy of CXR image classification, and the average accuracy of Chest X-ray images for three categories reaches 98.18%, of which the accuracy for COVID-19 reaches 99.52%.

Keywords: Attention Mechanism; Transfer Learning; Chest X-ray; Disease Classification; Data Enhancement.

1. Introduction

With the development of information technology, medical imaging has become an important carrier for obtaining clinical information. Due to the much higher amount of radiation detected by CT compared to X-rays, Chest X-ray (CXR) images are still one of the most important ways to assist doctors in diagnosing lung diseases. Discriminating X-ray scanning image data requires doctors to have a high level of professional expertise and rich medical experience. When the number of visits suddenly increases and leads to a sharp increase in image data, imaging doctors may also miss or misdiagnose due to prolonged mechanical, lengthy, and repetitive film reading. Based on these issues, it is necessary to develop an efficient and automatic classification method to assist doctors in clinical diagnosis.

In recent decades, artificial intelligence has shone brightly in medical image analysis, and deep learning technology has been applied in the field of medical images [1] Many scholars have improved classification algorithms in convolutional neural networks and achieved good experimental results. Xu Baohong [2] summarized the progress of various neural networks in the diagnosis of Alzheimer's disease (AD), and summarized the problems and research directions of using neural networks in this field; Zhaohan Xiong [3] developed AtriaNet on 154 3-DLGE-MRI, which can automatically segment the left atrium (LA) skeleton and endocardium. AtriaNet performed better than state-of-the-art CNN, with DICE scores of 0.940 and 0.942 for LA adrenal and endocardium, respectively, and a patient to patient variance of<0.001; Wu Ying [4] used transfer schools to classify ultrasound images of benign and malignant breast tumors, and the sensitivity and accuracy of this method in diagnosing breast malignant tumors reached 96.04% and 97.67%, respectively; K Padmavathi et al. [5]used serial fusion and PCA (Principal Component Analysis) based feature selection methods to enhance feature vectors. The performance of the DL frame was tested using LIDC-IDRI benchmark lung cancer CT images, with a classification accuracy of 97.27% and a high score in Auc Roc and accuracy. The use of this method can significantly improve the diagnostic performance of metastatic cancer. In the detection and classification tasks of lung diseases, Wang et al. [6] released the Chest14 dataset, which includes 14 common lung diseases, providing data support for subsequent research. Yao et al. [7]combined ResNet and DenseNet at multiple resolutions and proposed a network structure for classification. This method outperforms high-resolution chest X-ray images, but lacks attention to feature channel information, resulting in low accuracy. Zhang et al. [8] added dense compressed excitation blocks to the basic network of DenseNet to enhance the information expression ability of feature channels. Introducing compression excitation modules in a densely connected manner in the network to form a feature channel attention module, enhancing effective feature information in the channel while suppressing invalid feature information. Ozturk et al. [9] proposed a DarkCovidNet model for COVID-19 based on Darknet-19, which improves network performance while reducing model complexity. Cheng Wenjuan et al. [10] proposed a CXR image classification algorithm based on the XDense RC net network. This method improves the DenseNet model by introducing a newly proposed spatial attention mechanism in the original dense connection layer to achieve feature extraction and fusion, optimize the Transition module of DenseNet, and use two different pooling strategies to enhance the model's anti-interference ability. Shao Lingyun et al. [11] proposed a chest X-ray disease classification algorithm DECA Net (Dense Efficient Channel Attention Network) based on efficient channel attention. The efficient channel attention module is added to the basic feature extraction network in a densely connected manner to enhance the transmission of effective information in the feature channels and suppress the transmission of invalid information; Using asymmetric convolutional blocks to improve network feature extraction capability; Using multi label loss function to solve the problem of multi label and data imbalance. Pang Yu et al. [12] proposed a chest X-ray image classification algorithm based on DuaLNet network: using two symmetric networks AC ResNet as feature extractors for DuaLNet, the AC ResNet network integrates asymmetric convolutional blocks to improve the accuracy of residual network disease classification.

At present, deep learning has achieved good results in CXR image classification, but objective factors still limit the effectiveness of deep neural learning networks. For example, CXR only has a frontal image, with a single manifestation and some features not obvious. Moreover, the pathological features of different diseases in the chest cavity are extremely similar, making it difficult to distinguish between different types of diseases. In order to further improve the accuracy and robustness of classification, this paper adds an attention mechanism to the ConvNeXt model and uses transfer learning methods to construct our classification method. This method has the following three improvements: (1) using random erasing (RE) data augmentation to improve network performance. (2) Add an attention mechanism module to the backbone network to obtain more detailed features of classified images. (3) The random erasure enhancement method and attention mechanism module were applied to ConvNeXt, and transfer learning was applied to it. The main process of this article's method is as follows: flipping and standardizing CXR images to enhance data; By using the transferred ConvNeXt model to train the classification network, the initial performance of the model can be improved, the speed of the model can be improved, and the model can converge better without relying on specific modules such as window attention and relative position shift; Finally, obtain the output of the classification, update the network parameters through backpropagation, and obtain the final classification model.

2. Based on ConvNeXt Classification Algorithm

The ConvNext model applies the most advanced structures and training strategies in computer vision to Resnet (Residual Neural Network)[13]. The deeper the convolutional network, the deeper the features can be extracted, and the more problems of gradient vanishing and exploding may occur. In Resnet, this degradation phenomenon is solved by adding residual units to the deep neural network. ConvNeXt is a transformer like network structure that has undergone extensive exploration in training parameters and strategies, ultimately resulting in a convolutional neural network structure. The backbone network of this article uses ConvNeXt T, as shown in Fig 1

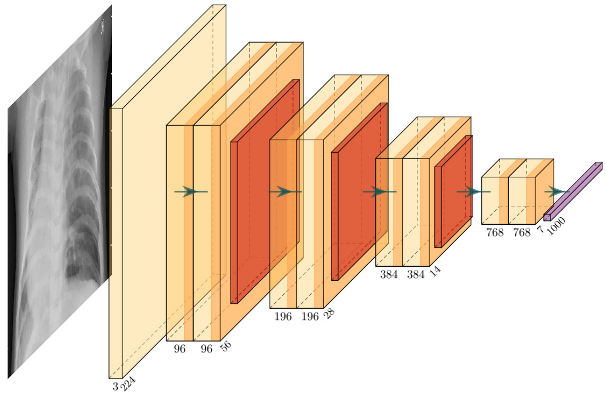


Fig 1. ConvNeXt-T model structure diagram

2.1. Attention Mechanism

2.1.1. Sub Heading

Attention mechanism is the study of human vision. In cognitive science, due to the bottleneck of information processing, humans selectively focus on a portion of all information while ignoring other visible information. The attention mechanism mainly has two aspects: determining which part of the input needs to be focused on; Allocate limited information processing resources to important parts. In deep learning, attention can be achieved through importance weight vectors: when predicting or inferring an element, such as a pixel in an image or a word in a sentence, we use attention vectors to determine how strongly it is related to other elements, and then sum the weighted vectors to approximate the final target value. Attention is the process of filtering out a small amount of important information from a large amount of information and focusing on these important pieces of information, ignoring most unimportant ones. The larger the weight, the more focused it is on its corresponding Value, that is, the weight represents the importance of information, and the Value is its corresponding information.

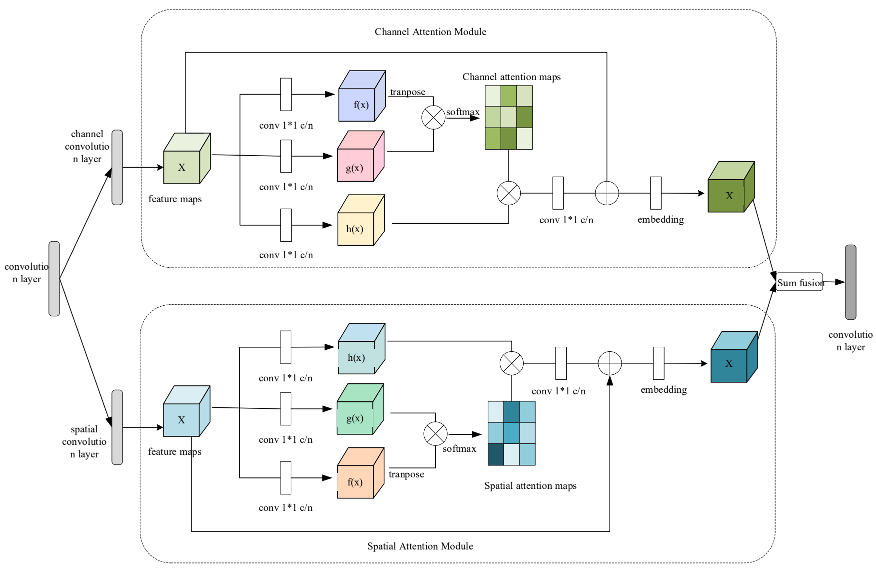


Fig 2. Attention Mechanism Module

The attention mechanism used in this article is shown in Fig. 2, which uses channel attention module and spatial attention module to filter features respectively. The Spatial Attention Module and Channel Attention Module both use self attention method, but their positions of action are different. The idea of self attention is to compress the input feature map through three sets of 1 * 1 convolutions with the same number, while preserving the channel dimension to flatten the width and height into one dimension. This is mainly to reduce the information redundancy of the input feature map and reduce the complexity of subsequent similarity calculations. Secondly, transpose the feature maps of branch f (x), then multiply the matrix with the feature maps of branch g (x), and normalize the results through softmax. Finally, the normalized output attention matrix is multiplied by the feature map obtained from branch h (x), and the channel is further expanded to the number of channels in the input feature map through softmax and 1 * 1 convolution. At this point, the key detail features in the output feature map are more fully expressed compared to the original feature map, thus achieving attention reassignment. Finally, the channel domain attention and spatial domain attention are combined to obtain the final feature output.

By adding an attention mechanism module to ConvNeXt, medical images may not only have obvious features of existing diseases, but also have non prominent features that are difficult to detect in the early stages. The spatial attention mechanism can allocate weights based on the importance level of each feature point on the feature layer, obtaining deeper features. Each feature point on the resulting feature map is obtained by weighting each feature point on the original feature map. The channel attention mechanism uses autocorrelation to better focus on the relationships in the channel dimension. Firstly, space is compressed, convolution is performed on the channel dimension for feature learning, and weights are assigned to each channel to obtain the importance of each channel. The use of attention mechanism can locate useful features, suppress useless features, and improve the convergence speed of the model.

2.2. Model Architecture

CS ConvNeXt adds an attention mechanism module before downsampling in the ConvNeXt backbone network. This attention module uses channel attention mechanism and spatial attention mechanism in parallel, which can obtain fused features in the image from both channel and spatial levels, further enhancing feature representation, extracting richer contextual and deeper features, and improving classification accuracy. The specific configuration table is shown in Table 1.

CS-ConvNeXt consists of multiple convolution and attention layers. The network first reduces the floating point number of operations by sampling four times through a convolution layer with a convolution kernel size of 4x4 and a step length of 4. Conv2_x consists of three ConvNeXtBlock residuals that are used to extract features, use Layer Normalization in the residuals module to speed up convergence of the network to reduce overfitting, use GELU activation functions, and then input into the next layer structure. The Attention module is used to extract the features of space and channels, assign attention weights and keep the number of channels constant. Then it is sent into a separate subsampling layer for subsampling, and after output, it enters the next residual block for feature extraction. Finally, average pooling is used to map input features, full connection layer is used to classify features, and softmax is used to normalize the three classifications.

Table 1. CS ConvNeXt Network Configuration Table

Layer name	Output size	ConvNeXt
Conv1	56 × 56	$4\times4,96,\mathit{stride}\ 4$
Conv2_ X	56 × 56	$\begin{bmatrix} d7\times7,96 \\ 1\times1,384 \\ 1\times1,96 \end{bmatrix}\times3$
Attention	56 × 56	Attention
Conv3_ X	28 × 28	$\begin{bmatrix} d7\times7,192 \\ 1\times1,768 \\ 1\times1,192 \end{bmatrix}\times3$
Attention	28 × 28	Attention
Conv4_ X	14 × 14	$\begin{bmatrix} d7\times7,384 \\ 1\times1,1536 \\ 1\times1,384 \end{bmatrix}\times27$
Attention	14 × 14	Attention
Conv5_ X	7×7	$\begin{bmatrix} d7\times7,768 \\ 1\times1,3072 \\ 1\times1,768 \end{bmatrix}\times3$
FC	1×1	Average pool, 9-d fc, softmax

2.3. Transfer Learning

The concept of transfer learning is the process of using the similarity between models, tasks, or data to apply the original domain model in a new field for learning. Given the source domain Ds and source task Ts, target domain Dt and target task Tt, transfer learning is the use of the content of Ds or Ts to improve the prediction performance of target task learning when Ds ≠ Dt or Ts is not equal to Tt.

Considering that deep learning models have many hidden layers and parameters, training from scratch requires significant computational and time costs, and actual data is limited, making it difficult to train effective deep models. Model based transfer learning can not only effectively solve the above problems, but also accelerate the convergence speed of the model and improve its classification performance.

The specific implementation of transfer learning is as follows:

The domain is represented by D={X, P (X)}, where X is the feature space; P (X) probability distribution, and X={x1,,, xn} ∈ X; The task is represented by T={Y, F (X)}, where Y is the label space; F (X) objective prediction function.

Using the pre trained CS ConvNeXt model on nine types of images, remove the existing fully connected and softmax layers, retain the input layer, pooling layer, and 13 convolutional layers, and maintain the input image format of 224x224 while retaining its weights. Transfer the convolutional and pooling layers to the chest X image classification model, and add 3 fully connected layers at the back of the model, with each layer's activation function using the ReLU function; The final layer is fully connected, and the activation function uses the Sogmoid function to obtain three classifications. Finally, the trained network is used to fine tune the training of the chest X image database.

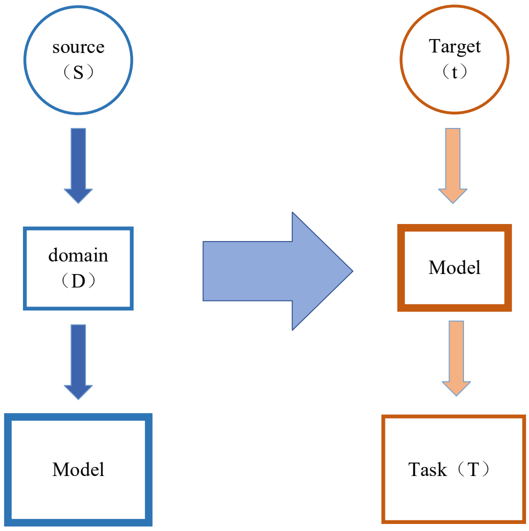


Fig 3. Schematic diagram of transfer learning

2.4. Random Erasure Data Augmentation

Data augmentation is a data augmentation technique that can utilize existing data to create more equally effective data. Traditional data augmentation methods include rotation, cropping, adding noise, etc. Random erasure is a lightweight method that does not require additional parameter learning or consumes more memory. It can be integrated into the network structure of this article, which can improve the performance of the model. The specific implementation method is as follows:

Input the image I , with a erase probability of p , the scale range of the erase area is from s_l to s_h , and the scale range of the aspect ratio is from r_1 to r_2 . Firstly, p determine whether an image needs to be erased based on probability:

$$p_1 = Rand\left(0,1\right) \tag{1}$$

Among them p_1 is the probability of random generation. If it is $p_1 > p$, the image will not be processed, otherwise it needs to be erased.

Based on the input image H , the I length, width, W and area of the image can be obtained S . Based on the $Rand\left(s_1,s_h\right)*S$ obtained erasure area S_e , where s_1 is the random erasure length ratio, s_h is the random erasure width ratio, and the erasure area length and width are obtained from the following equation:

$$\begin{aligned} r_e &= Rand\left(r_1,r_2\right) \\ H_e &= \sqrt{S_e * r_e} \\ W_e &= \sqrt{\frac{S_e}{r_e}} \end{aligned} \tag{2}$$

Where, r_e is the random aspect ratio, H_e is the length of the erase area, and W_e is the width of the erase area.

Enter the erased top-left position x_e in picture I . y_e can be obtained from $Rand\left(0,H\right)$ and $Rand\left(0,W\right)$. By adding the length and width, the position of the erasing area can be obtained. There may be $y_e + H_e > H$ situations where $x_e + W_e > W$ and, if so, re implement the above algorithm until either $x_e + W_e \le W$ or is met $y_e + H_e \le H$.

The data augmentation methods used by ConvNeXt in the preprocessing stage are Cutout and Mixup. Cutout randomly cuts out some areas of the sample and fills them with 0 pixel values, keeping the classification results unchanged; Mixup involves mixing two random samples in proportion, and distributing the classification results proportionally. Cutout directly fills with 0 pixel values, which does not match the actual chest X-ray image. Mixup mixes the two images proportionally, which will affect the classification results and does not meet the requirements of single label classification. The ConvNeXt backbone network increases the depth of the network, which to some extent increases the computational complexity and memory usage of the model. In addition, medical images often have blurred boundaries, noise, and the inability to determine the location of lesions, which can affect the classification performance of images. To effectively solve the above problems, this paper uses random elimination for data augmentation in the experiment. Random erasure allows for the random selection of a matrix region of an image and the random erasure of pixel values, simulating the situation where the current part is obstructed by other parts in medical images. This method is also a lightweight enhancement method that does not add additional parameter techniques and memory. Place the data augmentation method before the model performs feature extraction to improve its learning ability.

2.5. Evaluation Indicators

In order to objectively evaluate the method proposed in this paper, for single label multi classification, accuracy (Acc), area under ROC curve (AUC) under receiver operating characteristics (ROC), and confusion matrix were used to demonstrate the recognition and classification ability of the method for the three categories. Using AUC requires defining false positive rate (FPR) and true positive rate (TPR), and using confusion matrix requires defining precision and recall.

$$\begin{cases} \mathrm{Acc} = \dfrac{TP+TN}{TP+FP+TN+FN} \\ FNR = \dfrac{FN}{TP+FN} \\ FPR = \dfrac{FP}{FP+TN} \\ \mathrm{Pr}\,ecision = \dfrac{TP}{TP+FP} \\ \mathrm{Re}\,call = \dfrac{TP}{TP+FN} \end{cases} \tag{3}$$

Where, TP(True positives) is the true rate, representing the number of positive samples recognized as positive samples. TN(Ture negatives) is true negative ratio, which represents the number of negative samples identified as negative samples. FP(False positives) is the false positive rate, representing the number of negative samples identified as positive samples. FN(False negatives) is the false negative rate, which represents the number of positive samples that are identified as negative samples.

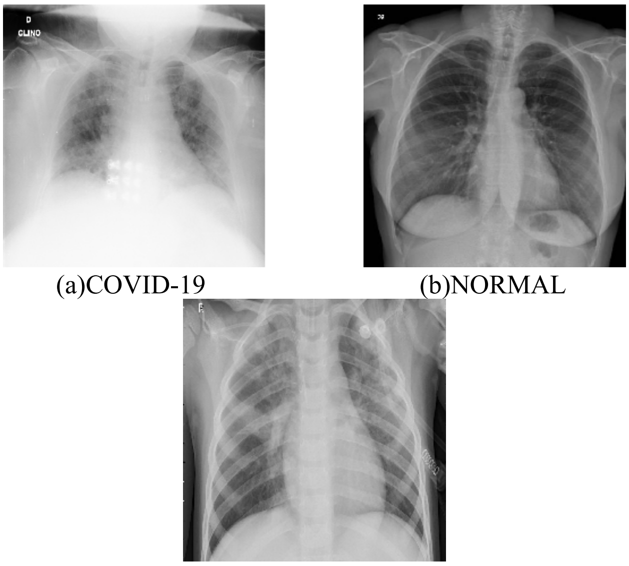
The higher the accuracy, the closer the ROC curve is to the upper left corner, the larger the AUC value, and the closer the value is to 1. The larger the main diagonal of the confusion matrix, the better the classification performance of the algorithm. Based on the validation set, save the model weights with the best accuracy for test level prediction. In the experiment, draw confusion matrices and ROC curves for each disease, and calculate the accuracy of prediction and AUC scores for each category to evaluate classification performance.

3. Experiments and Analysis

3.1. Random Erasure Data Augmentation

This experiment uses the COVIDx [8] dataset, mainly targeting databases of pneumonia cases, viral pneumonia cases, and normal images. It is a medical image directory structure divided into three sub folders (COVID-19, NORMAL, PNEUMOIA), which contain chest X-ray (CXR) images. COVID-19: 1626 images, normal: 1802 images, pneumonia: 1800 images. This article randomly divides the dataset into three parts: training set, validation set, and test set, with each part having a ratio of 7:2:1.

The CXR image in the COVIDx dataset is shown in Figure 4. From left to right, it is COVID-19, normal, and viral pneumonia. Due to the input standard of the convolutional neural network used in this article being 224 * 224 * 3, scaling down the 256 * 256 image to 224 * 224 can reduce computational complexity; Using RE enhancement methods for data augmentation; Then, convert the CXR image to RGB 3-channel tensor format, which can create higher dimensional matrices; Finally, normalize the image values of CXR images and convert the data to a standard normal distribution.



(a)COVID-19 (b)NORMAL

(c)PNEUMONIA

Fig 4. Data Category Diagram

3.2. Random Erasure Data Augmentation

The experiment was conducted on Windows 10 using the Anaconda environment and the deep learning framework using pytorch, running on the GPU of NVIDIA RTX3060. In terms of parameter settings, the optimization algorithm uses Adaptive Moment Estimation (Adam) with 100 iterations, a batch size of 16, and an initial learning rate of 0.001. The learning rate is set to 0.1 times the original value for every 20 epochs of iteration.

The confusion matrix of the classification experiment in three categories using this method is shown in Fig From the graph, it can be seen that the values of the confusion matrix are mostly distributed on the main diagonal, indicating that the classification model performs well.

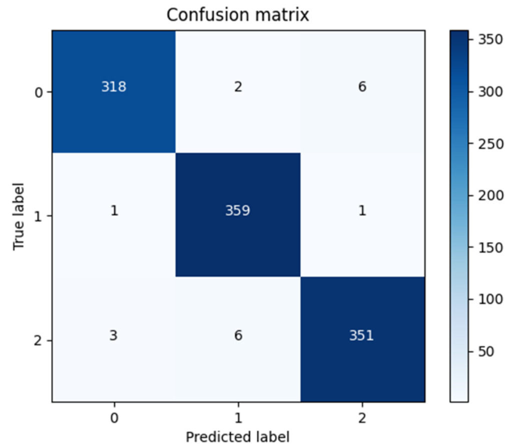


Fig 5. ConvNeXt Confusion Matrix

The ROC curve shown in Fig depicts the classification performance of our method on three types of CXR images on the dataset used. The closer the curve is to the upper left corner and the area is 1, the better the overall classification performance of the method.

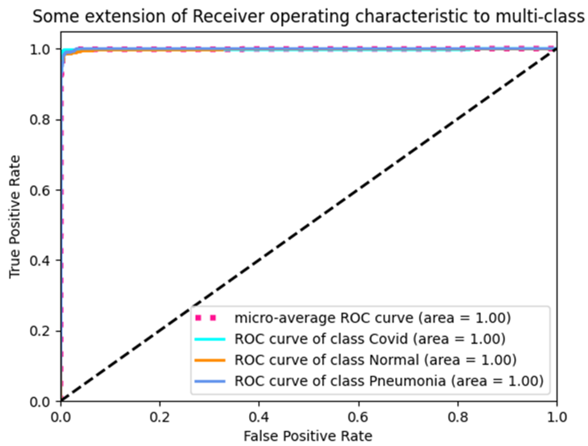


Fig 6. ROC curve graph

Table 2. Dataset Classification Accuracy Table

method	Acc	Pmacro	Rmacro	Fmacro
Wang et al.[6]	89.00	87.68	90.67	89.15
Pham et al. [14]	93.00	91.94	93.17	92.55
Ozturk et al.[9]	93.75	92.60	94.33	93.46
SE Net[15]	93.50	92.03	93.67	92.84
CBAM[16]	92.25	90.97	93.00	9.97
XDense RC net [10]	96.75	95.77	96.50	96.13
CS ConvNeXt	98.18	99.51	98.22	98.18

Table 2 shows the classification accuracy on this dataset. From Table 2, it can be seen that compared with XDense RC net, our method has improved accuracy (Acc), accuracy (Pmacro), recall (Rmacro), and F1 (Fmacro) by 1.43, 3.74, 1.72, and 2.05, respectively.

Table 3 presents the classification results for three categories, from which it can be seen that the proposed method achieves an accuracy of over 98% in each category.

Table 3. Precision Table for Each Classification of the Dataset

category	Accuracy	Precsion	Recall	F1
COVID-19	0.9952	0.9908	0.9938	0.9923
Nomal	0.9837	0.9673	0.9861	0.9766
Pneumonia	0.9847	0.9886	0.9666	0.9775

To verify the effectiveness of the fine-tuning model, ablation experiments were conducted on both ConvNeXt's basic model and the model using transfer learning. As shown in Table 4, RE_ ConvNeXt indicates that using data augmentation and random erasure in data augmentation resulted in an increase of 0.67%, 0.32%, 0.65%, and 0.70% in Accuracy, Precsion, Recall, and F1, respectively. This demonstrates that random erasure is suitable for CRX images and can improve the model's generalization ability. AT_ ConvNeXt showed that adding attention modules to the backbone network improved various indicators by 0.86%, 0.66%, 0.7%, and 0.93%, respectively. The experiment proved that using attention mechanism methods can optimize the performance of the model and improve the accuracy of the classification network. CS_ ConvNeXt is the method used in this article. In the feature extraction stage, random erasure is added for data augmentation, and attention mechanism is added to the backbone network. Transfer learning is used to improve various indicators by 4.68%, 1.68%, 4.56%, and 4.73%, respectively, greatly improving classification accuracy.

Table 4. Ablation Test

method	Accuracy	Precsion	Recall	F1
ConvNeXt	93.50	97.83	93.66	93.45
RE_ ConvNeXt	94.17	98.15	94.31	94.15
AT_ ConvNeXt	94.36	98.49	94.36	94.38
CS_ ConvNeXt	98.18	99.51	98.22	98.18

4. Summary

To effectively improve the classification accuracy and discrimination ability of convolutional neural networks for pulmonary inflammation, this paper proposes an attention mechanism based convolutional neural network model CS ConvNeXt. This model uses random erasure for data augmentation to improve the generalization ability of the model. The application of transfer learning technology effectively improves the classification accuracy of the model and enhances the network's feature extraction ability for diseases, thereby classifying COVID-19 and PNEUMONIA in chest X-ray images. Experimental results have shown that the model can accurately classify COVIDx, with an accuracy rate of 98.18% compared to other existing methods, demonstrating the effectiveness of the proposed method.

Acknowledgments

Sichuan University of Science & Engineering Graduate Innovation Fund Project (Y2022175); China University Industry University Research Innovation Fund (2021Z YA1102); Key Laboratory of Higher Education of Sichuan Province for Enterprise Informationalization and Internet of Things (2023WYY01).

References

[1] Y. Yizhou, S. Dejun, M. Jiechao, and Z. Zhen, "Advances in application of artificial intelligence in medical image analysis," Chinese Journal of Medical Imaging Technology, Vol. 35(2019) No. 12, p. 1808-1812.

[2] X. Baohong, D. Chong, and X. Guizhi, " research on the application of convolution neural network in the diagnosis of alzheimer's," disease Journal of Biomedical Engineering, Vol. 38(2021) No. 01, p. 169-177.

[3] Z. Xiong, V. V. Fedorov, X. Fu, E. Cheng, and R. Macleod et al., "Fully automatic left atrium segmentation from late gadolinium enhanced magnetic resonance imaging using a dual fully convolutional neural network," Ieee Transactions On Medical Imaging, Vol. 38(2019) No. 2, p. 515-524.

[4] W. Ying, L. Liangping, X. Bo, J. Huang, and Z. Luyu, "Classification and diaanosis of ultrasound images with breast tumors based on transfer learning," Chinese Journal of Medical Imaging Technology, Vol. 35(2019) No. 03, p. 357-360.

[5] K. Padmavathi, M. V. Karki, and M. Bhat, "Medical image fusion of different modalities using dual tree complex wavelet transform with pca," in 2016 International Conference on Circuits, Controls, Communications and Computing (I4C), IEEE, Bangalore, India2016, p. 1-5.

[6] X. Wang, Y. Peng, Le Lu, Z. Lu, and M. Bagheri et al., "Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Honolulu, HI, USA, 2017, p. 3462-3471.

[7] L. Yao, J. Prosky, E. Poblenz, B. Covington, and K. Lyman, "Weakly supervised medical diagnosis and localization from multiple resolutions," arXiv e-prints,2018, p. 1803-7703.

[8] Z. Zhirui, L. Qiang, and G. Xin, "Multilabel chest x-ray disease classification based on a dense squeeze-and-excitation network," Journal of lmage and Graphics, Vol. 25(2020) No. 10, p. 2238-2248.

[9] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, and O. Yildirim et al., "Automated detection of covid-19 cases using deep neural networks with x-ray images," Computers in Biology and Medicine, Vol. 121(2020), p. 103792.

[10] C. Wenjuan and Y. Guoqing, "Lmproved cxr image classification algorithm based on xdense-rc-net," Application Research of Computers, Vol. 39(2022) No. 12, p. 3803-3807.

[11] S. Lingyun, L. Qiang, G. Xin, and D. Xuewen, "Disease classification algorithm of chest -ray based on efficient channel attention," Laser & Optoelectronics Progress, Vol. 60(2023) No. 12, p. 360-369.

[12] P. Yu, B. Rui, W. Huiqian, and H. Zhiwei, "Research progress of antiphage defense systems in bacteria," Life Science Instruments, Vol. 20(2022) No. 02, p. 44-50.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, NV, USA, 2016, p. 770-778.

[14] T. D. Pham, "Classification of covid-19 chest x-rays with deep learning: new models or fine tuning?" Health Information Science and Systems, Vol. 9(2020) No. 1, p. 2.

[15] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, Salt Lake City, UT, USA, 2018, p. 7132-7141.

[16] S. Woo, J. Park, J. Lee, and I. S. Kweon, "Cbam: convolutional block attention module," in Computer Vision – ECCV 2018, Springer International Publishing, Cham,2018, p. 3-19.