Biomedical Data Science
2025, VOL. 5, NO. 2, 183-187
DOI: 10.47297/wspbdsWSP2752-630534.20250502

C2PSA-Enhanced YOLOv11: Boosting Pneumonia Detection in Chest X-rays Through Pseudo-Color Attention Mechanisms

Yuchen Li

Guangdong Polytechnic Normal University, Guangzhou, Guangdong, 517500, China

ABSTRACT
Pneumonia, a common respiratory infection, has a serious impact on human lives worldwide. In this paper, the editors propose a Yolov11 deep learning model that introduces an attention mechanism and present a data enhancement technique for medical imaging. After training and tuning on a public dataset and doing comparative experiments with other deep learning models, the model can efficiently detect lung X-images and determine whether it has pneumonia or not. The experimental data show that the accuracy, precision, recall, and F1-score of the model are 90.12%, 91.45%, 90.66%, and 91.05%, respectively, and have excellent results in the comparison experiments with four common architectures, namely, Yolov11, Yolov8, ResNet18, and ResNet50, which confirms the superiority of the model. By comparing the performance of the model before and after data enhancement, it was concluded that pseudo-coloring of gray medical images can improve the accuracy of the model by 4-7%.

KEYWORDS
Yolov11; C2PSA; OpenCV; pseudo-color

1 Introduction

Pneumonia is a common lower respiratory infection that poses a major threat to global public health. According to a study on the epidemiology of pneumonia by Xingyuan Zhou et al.[1] from the School of Public Health, Department of Medicine, Ningbo University, pneumonia causes a huge disease burden and economic burden worldwide. The epidemiological characteristics of pneumonia vary in different regions, with men, the elderly and children being at high risk of pneumonia morbidity and mortality. In addition, age, smoking, alcohol consumption, air pollution, comorbidities and vaccination status are the main factors influencing pneumonia morbidity and mortality.

Globally, cardiovascular disease related mortality increased by 18.7% from 2010 to 2020 and according to Yeshen Zhang et al.[2] it can be found that pneumonia is one of the important causes of death after cardiovascular disease. In the United Kingdom, an increased risk of pneumococcal pneumonia in patients with immune-mediated inflammatory diseases was learned from a statistical analysis by NAKAFERO G et al.[3] In China, the prevalence of community-acquired pneumonia varies in different regions and is more common in the elderly and children.

Pneumonia not only has an immediate impact on the patient's health, but can also lead to a range of long-term health problems. In order to get patients treated as soon as possible and detect the disease in a timely manner, doctors can determine whether a patient is suffering from pneumonia by looking at an X-ray of the patient's lungs. However, traditional testing methods rely on the doctor's experience, are highly subjective and susceptible to factors such as fatigue, resulting in inconsistent diagnostic results, and are inefficient, making it difficult to meet the demand for large-scale screening, especially during outbreaks, and failing to quickly and accurately identify whether or not a patient is suffering from pneumonia.

With the breakthrough of artificial intelligence in medical image analysis, image features can be automatically extracted, reducing manual intervention and improving diagnostic efficiency and accuracy. The method of detecting pneumonia through artificial intelligence can quickly analyze X-ray images and assist doctors in diagnosis, especially in resource-limited areas, which has important application value.

The goal of this paper is to apply artificial intelligence in the field of medicine to build models that can find lesions in the lungs by recognizing lung X-ray images, ultimately enabling the detection of pneumonia.

2 Literature Review

2.1 Deep learning based pneumonia detection

In recent years, deep learning techniques have been widely used in medical image analysis. Wenyi Yang[4] proposed a new CG2F-YOLO-based algorithm for the detection of neocoronary pneumonia, which improves the accuracy and efficiency of the detection by improving the YOLO network structure. Fengqi Cui et al. [5], on the other hand, used

Correspondence: 1161496417@qq.com

traditional image feature extraction methods, especially Local Binary Pattern (LBP), combined with Support Vector Machine (SVM) to achieve effective classification of viral pneumonia. Shenghui Huang[6] explored the effectiveness of the application in the recognition of CT images of neocoronary pneumonia based on a variety of deep learning models (e.g., Moco-resnet50, ResNet50, Moco-vit, and Vit), and the results showed that these models have better recognition ability.

Aiming at the problem that COVID-19 image features are difficult to detect, WANG Xin[7] proposes a detection model based on improved YOLOv8 and verifies the key features learned by the model by combining with Grad-CAM heat map visualization technology. Wen Xingkai et al.[8] investigated a machine learning-based method for recognizing lung CT images of COVID-19 patients and found that the LassoNet model had better interpretability, while the Residual Network (RN) was the best in terms of prediction performance. Murat Canayaz[9], on the other hand, combines multiple deep learning models (e.g., AlexNet, VGG19, GoogleNet, ResNet) and meta-heuristic algorithms (e.g., Binary Particle Swarm Optimization and Binary Graywolf Optimization) for feature selection and SVM for classification, resulting in a final model accuracy of 99.38%.

2.2 Pneumonia detection based on feature fusion and optimization

Narinder Singh Punn and Sonali Agarwal[10] used various deep learning methods such as ResNet, Inception-v3, Inception ResNet-v2, DenseNet169, and NASNetLarge to automatically diagnose COVID-19 in limited chest X-ray images. He Xiang[11] proposed a pneumonia diagnostic model based on multi-scale feature fusion to address the problem of small number of features in lung images and difficulty in training.Ikechukwu A V et al.[12] conducted a study using pre-trained VGG-19 and ResNet-50 models based on data augmentation and introduced regularization to reduce network overfitting, and the experimental results showed that the recall rate reached 92.03%.

2.3 Pneumonia detection based on a novel network structure

Xie Feng[13] proposed a densely connected self-attentive convolutional neural network (DSANet) to realize the automatic diagnosis of neocoronary pneumonia and the automatic segmentation function of lesion area. The experimental results show that DSANet has an accuracy of 95.01%, an F1-score of 94.99%, a precision of 95.13%, a recall of 94.99%, and an AUC of 99%.Kumari et al.[14] combined the VGG-19 convolutional neural network with a random forest classifier for extracting features from X-ray images and classifying them, and the model accuracy is 90.03%, precision was 93%, F1-score was 91% and area under the curve was 0.95.

However, when training the model, problems such as insufficient data volume and inaccurate data labeling can occur, which can affect the training results and generalization ability of the model, and deep learning is susceptible to overfitting due to changes in data distribution. In addition, the robustness of the model needs to be paid attention to for pneumonia, a disease that will continue to deteriorate, and there are some small lesions that are difficult to recognize. For these problems, it is important to design a model for high accuracy recognition.

3 Method

3.1 Yolov11-based X-ray recognition model for pneumonia

In order to solve the problem of recognizing the lesions of lung X-ray, this paper proposes a model of convolutional neural network, yolov11, as shown in Figure 1.

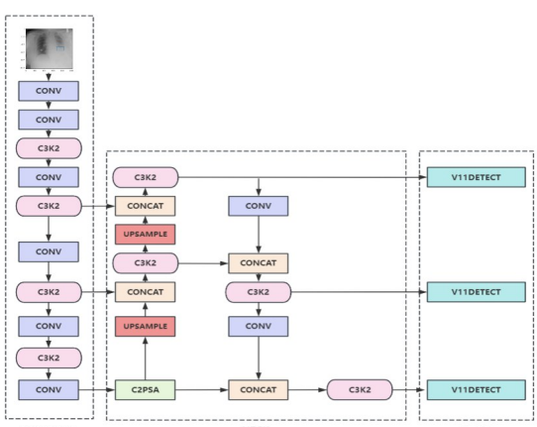

Figure 1 Network structure of yolov11

The structure of this model is divided into Backbone, Neck, and Head parts. Firstly, the lung X-ray image is input to the Backbone part, which consists of multiple convolutional layers (CONV) and C3K2 module for extracting features from the image. The C3K2 module may contain multiple convolutional layers and residual connections to enhance the feature extraction capability. After completing the feature extraction, the extracted feature maps enter the Neck section where the feature maps are processed for multi-scale feature fusion through feature map stitching (CONCAT), upsampling (UPSAMPLE) and further convolutional processing. This process enhances the model's ability to detect targets of different sizes and provides richer information for the final detection results. Finally, the fused feature maps are passed to the Head section for target detection via the V11DETECT module.The V11DETECT module is a YOLOv11-specific detection header used to output the final detection results, including bounding box prediction and category prediction. It contains multiple branches, each responsible for different scales of detection.

Yolov11 has the following significant advantages in identifying lung lesions: its powerful feature extraction capability and optimized architectural design enable the model to reduce the number of parameters while maintaining high accuracy, which significantly improves computational efficiency; it supports multi-tasking and cross-environmental deployment, which is able to adapt to different hardware conditions; it possesses real-time and highly efficient processing capability, which is suitable for rapid detection; through generalization capability and preprocessing steps (e.g., image enhancement, noise filtering, data expansion), it is able to learn multiple features of lung foci, reducing misdiagnosis and omission; ultimately, its high efficiency and reliability provide strong technical support for clinical diagnosis and reduce the workload of doctors, while improving the accuracy and efficiency of diagnosis.

3.2 Introducing the attention mechanism

In order to improve the precision and recognition accuracy of the model, so that the model can better recognize the key lesions, the feature weights of the key regions (e.g., the edge of the lung) can be enhanced through the position-sensitive mechanism by introducing Cross-Stage Partial Spatial Attention (CSPA) in the Backbone of YOLOv11.

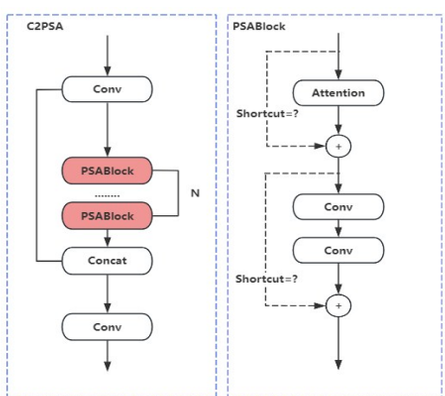

Figure 2 Network structure diagram of the C2PSA module

The above figure2 shows the network structure of the C2PSA module, which starts with a convolutional layer (Conv) for channel dimensionality adjustment, followed by multilevel feature refinement through N series-connected PSABlocks, and generates spatial attention weights through the Attention mechanism within each PSABlock, which dynamically enhances the pneumonia focal areas (e.g., lobular infiltration, ground-glass shadow) of the feature responses, and using the Shortcut connection of conditional judgment (the symbol "Shortcut = ?") (the symbol "Shortcut = ?" indicates that the residual connection is selected adaptively according to the feature matching), local detail enhancement is achieved by two convolutions while retaining the initial lung X-ray image information, and then the multi-scale features output from multiple PSABlocks are reorganized into high-discriminative features by terminal convolution after fusion of the Concat layer.

The core advantage of introducing this method for pneumonia X-ray recognition is that the Attention mechanism focuses on the key lesion areas in the lung field through position-sensitive weights, effectively suppressing the interference of irrelevant structures such as ribs and trachea, and the conditionalized Shortcut connection preserves the fine texture features of the lesions in the deep network, alleviating the problem of disappearing gradients and accelerating the convergence of the model.

3.3 Datasets

The dataset used for the experiment is from the RSNA Pneumonia Detection Challenge competition dataset on the official website of Kaggle, which is a medical image of the lung. In order to conduct the experiment, the author divided

the dataset into training set, test set, and validation set according to the ratio of 6:2:2, where the training set has 4809 images, of which half of them are normal images and half of them are pneumonia images, and the validation set has 1603 images and the test set has 1603 images. Then, the data enhancement operation is performed on the data, through OpenCV's multilevel pseudo-color enhancement based on tissue density, which converts the gray scale difference into saturation difference, so that the color contrast between the pneumonia lesions and the normal tissues is more significant, which improves the ability of the model to capture the tiny lesions, and enhances the model's resistance to luminance interference, and weakens the sensitivity to the luminance difference of the original X-ray film through color mapping, so that the model to pay more attention to the pathological characteristics of the tissue density itself. Figure 3 and Figure 4 show the comparison between before and after data enhancement.

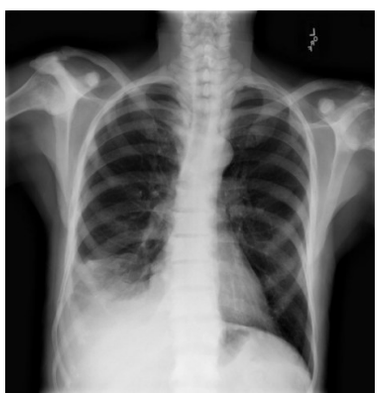

Figure 3 X-ray image before data enhancement

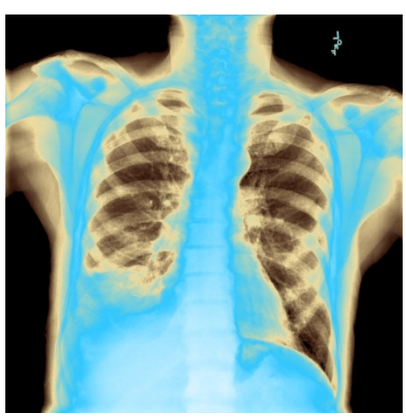

Figure 4 X-ray image after data enhancement

3.4 Evaluation of indicators

In order to be able to evaluate the performance of the model, the effectiveness of the model is evaluated by accuracy, recall, precision, and F1-score, and the related formulas are shown in (1)-(4).

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

$$Pre = \frac{TP}{TP + FP} \tag{2}$$

$$Rec = \frac{TP}{TP + FP} \tag{3}$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Re}\, call}{\text{Pr}\, ecision + \text{Re}\, call} \tag{4}$$

TP in the formula is True Positive used to indicate that the sample is positive and predicted to be positive; FP is False Positive used to indicate that the sample is negative and predicted to be positive; FN is False Negative used to indicate that the sample is positive but predicted to be negative; TN is True Negative used to indicate that the sample is negative and predicted to be negative.

4 Conclusion

The comparative results of the models before and after data enhancement are shown in Table 1 and Table 2.

Table 1 Experimental data before data enhancement

Models	Acc/%	Pre/%	Rec/%	F1/%
Yolov11	85.72	86.89	85.62	86.25
Yolov11-C2PSA	84.68	87.31	85.61	86.45
ResNet18	82.37	83.74	86.47	85.08
Retnet50	85.62	86.43	87.62	87.02
Yolov8	81.77	81.95	82.49	82.22

Table 2 Experimental data after data enhancement

Models	Acc/%	Pre/%	Rec/%	F1/%
Yolov11	88.96	90.62	89.74	90.18
Yolov11-C2PSA	90.12	91.45	90.66	91.05
ResNet18	89.11	89.98	91.14	90.57
Retnet50	89.56	90.26	90.58	90.42
Yolov8	86.93	89.25	88.74	88.99

Compared with other deep learning models, the Yolov11 recognition model with the introduction of the attention mechanism has an excellent performance in recognizing pneumonia X-ray images with an accuracy of 90.12% and an F1-value of 91.05, and comparing the experimental data before and after the data enhancement, it can be found that the accuracy of the model can be improved by 4-7% on average after pseudo-coloring the gray medical images. The model accuracy can be generally improved by 4~7% on average.

Overall, the research experiments on data enhancement of medical images and the introduction of an attention mechanism to the yolov11 model proved its effectiveness and accuracy in image recognition of pneumonia X-rays, which can provide a new automated technological tool for doctors to reduce their workloads again while improving the quality of healthcare services.

References

[1] ZHOU Xingyuan, HONG Hang, FANG Ting, XU Guozhang. Research progress on the epidemiology of pneumonia. Preventive Medicine, 2023, 35 (8): 682-686.

[2] Yeshen Zhang, Haobin Liu, Yining Dai, Fei Ye, Wenzhi Luo, Shan Tu, Weikun Chen, Siyu Kong, Yu He, Ning Tan, Zhihui Zhang, Pengcheng He, Yuanhui Liu, Evaluation of a bidirectional causal association between cardiovascular diseases and pneumonia: a Mendelian randomization study, European Journal of Preventive Cardiology, 2024;, zwae060, https://doi.org/10.1093/eurjpc/zwae060

[3] NAKAFERO G, GRAINGE MJ, CARD T, et al. Effectiveness of pneumococcal vaccination in adults with common immune-mediated inflammatory diseases in the UK: a case-control study[J]. Lancet Rheumatol. 2024 Sep;6(9):e615-e624. https://pubmed.ncbi.nlm.nih.gov/39067457/ DOI: 10.1016/S2665-9913(24)00128-0.

[4] Wengyi Yang. Research on X-ray Image Detection of COVID-19 Based on Deep Learning[D]. Chongqing University of Technology, 2024. DOI: 10.27753/d.cnki.gcqgx.2024.001568

[5] Fengqi Cui, Haoran Xiao,Yi Mou. A study on the classification of viral pneumonia X-ray images[J]. Scientific and Technological Innovation, 2021 (04): 56-58.

[6] Shenghui Huang. COVID-19 CT image recognition based on deep learning[D]. Shanghai University of Finance and Economics, 2023. DOI: 10.27296/d.cnki.gshcu.2023.002164.

[7] WANG Xin. Research on Deep Learning-Based Methods for COVID-19 ImageFeature Detection. 2023. Lanzhou University of Technology, MA thesis.doi:10.27206/d.cnki.ggsgu.2023.001378.

[8] Wen Xingkai, Qian Yishan, Liu Wenmin, et al. COVID-19 patient lung CT image recognition based on machine learning[J]. Technology Innovation and Application, 2023,13(14):1-6.DOI:10.19981/j.CN23-1581/G3.2023.14.001.

[9] Murat Canayaz. MH-COVIDNet: Diagnosis of COVID-19 using Deep Neural Networks and Meta-heuristic-based Feature Selection on X-ray Images[J].Biomedical Signal Processing and Control,2020(prepublish):102257-102257.

[10] Narinder Singh Punn;Sonali Agarwal. Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray images using fine-tuned deep neural networks[J].Applied Intelligence,2020(5):1-14.

[11] He Xiang.Deep learning-based CT-aided diagnosis ofpneumonia Technology Research[D].China University of Mining and Technology, 2022. DOI:10.27623/d.cnki.gzkyu.2022.000680.

[12] Ikechukwu A V, Murali S, Deepu R, et al. ResNet-50 vs VGG-19 vs training from scratch: Acomparative analysis of the segmentation and classification of Pneumonia from chest X-ray

[13] Xie Feng.Research on COVID-19 Aided Diagnosis andLesion Segmentation Method Based on DeepLearning[D]. Shenyang Ligong University, 2022.DOI:10.27323/d.cnki.gsgyc.2022.000320.

[14] Kumari, Kajal, Sahana, Sudip Kumar, Mustafi, Debjani (2025) Deep Learning Techniques for Pneumonia Detection on Chest X-Ray Images Communications in Computer and Information Science, Volume 2335 CCIS, Pages 152-167