Pneumothorax Recognition Neural Network Based on Feature Fusion of Frontal And Lateral Chest X-ray Images

Jia Xin Luo, Wu Feng Liu, and Liang Yu

Henan University of Technology, Zhengzhou 450001, China

Corresponding author: Wu Feng Liu (e-mail: lwf@haut.edu.cn).

This research was supported by the National Natural Science Foundation of China (11005136).

ABSTRACT Pneumothorax is a potentially life-threatening disease that requires urgent diagnosis and treatment. Clinically, a chest X-ray examination is the first choice for diagnosing pneumothorax. However, it is difficult to diagnose pneumothorax by only frontal chest X-ray imaging when the lesion area is only composed of a small amount of air. Therefore, we propose a pneumothorax diagnosis neural network based on feature fusion, where frontal and lateral X-ray information are fused. In this network, there are two inputs and three outputs. The two inputs are the frontal chest X-ray image and the lateral chest X-ray image. The three outputs are the classification results of the frontal chest X-ray image, the classification results of the lateral chest X-ray image, and the classification results integrating the characteristics of the fused frontal chest X-ray image and lateral chest X-ray image. Our algorithm considers the vanishing gradient problem in the pneumothorax recognition model and introduces the residual block to alleviate this problem. Because of the large number of channels in this model, we also utilize channel attention mechanisms to improve the model’s performance. Our comparative experiments show that neural network fusion of frontal and lateral chest image features can achieve higher accuracy than the single task model. Using only image-level annotation, our pneumothorax model can achieve high recognition accuracy.

INDEX TERMS Convolutional neural network, pneumothorax, chest X-ray images, computer-aided diagnosis, multiple input network.

I. INTRODUCTION

Pneumothorax (lung collapse) occurs when excessive air accumulates in the pleural cavity between the lung and the chest wall. This air pressure causes the lung to collapse. The main symptoms of pneumothorax are chest tightness, shortness of breath, and cough. Pneumothorax may be caused by physical factors, such as chest trauma or impact, smoking, and pulmonary diseases [1, 2]. The diagnosis of pneumothorax is very complex and is usually determined by radiologists based on a chest X-ray examination. However, using a chest X-ray to diagnose pneumothorax is challenging for radiologists. Even experienced radiologists need to carefully adjust image display settings, such as the window width, window level, and image contrast, to make the correct diagnosis of the disease. This work requires a large amount of clinical experience and patience. Sometimes, fatigued doctors will make incorrect judgments. The diagnostic accuracy of pneumothorax highly depends on the expertise of the attending radiologist [3, 4].

In the absence of trained radiologists, the correct diagnosis and treatment of pneumothorax are often delayed, which may cause serious injury or even death in patients. The above situation is common in some undeveloped countries and regions. Therefore, there is a need for a computer-aided diagnostic [5, 6] (CAD) tool to help doctors accurately diagnose and detect pneumothorax. Deep learning-based technology is a popular choice for image segmentation and classification. Great success has been achieved using deep learning in different fields, such as natural scene image understanding[7], geographic exploration[8], and medical image recognition [9, 10]. The popularization of object detection, semantic segmentation, and disease classification based on deep learning has greatly relieved doctors of tedious work and improved the diagnosis efficiency[11-14]. Multitask learning strategies[15, 16], such as Siamese networks[17] and auxiliary tasks [18], have attracted increasing attention in computer-aided diagnostics.
In engineering, when designing an outstanding pneumothorax auxiliary diagnosis algorithm, the following challenges are usually encountered. The resolution of a chest X-ray is limited, and it is difficult to distinguish when the lesion area contains only a small amount of air. Air accumulation areas can be scattered at various locations and appear in multiple shapes. Pixel-level medical image annotation (i.e., strong supervision) is expensive and difficult to obtain. Image level medical image annotation (i.e., weak supervision) is relatively easy to obtain, but it is not easy to use to achieve high recognition accuracy [19].

To alleviate these problems, we propose a multitask learning model [15] (i.e., a fusion model) for the automatic detection of pneumothorax. The model can be used to fuse information from frontal and lateral chest X-ray images to realize high-precision automatic recognition of pneumothorax. We invited experienced radiologists to build a pneumothorax image dataset (Haut-NY). This dataset contains not only frontal chest X-ray images but also corresponding lateral chest X-ray images, which is meaningful. Many well-known pneumothorax datasets, such as the National Institutes of Health chest X-ray dataset [19, 20] and the Society for Imaging Informatics in Medicine (SIIM) Pneumothorax Challenge Dataset [21], do not contain lateral chest X-ray images. The main contributions of this work are summarized as follows.

- Different from previous works that only use frontal chest X-ray images to identify pneumothorax, we propose a multi-input multi-output neural network that can fuse information from frontal and lateral chest X-ray images. Experiments show that the model's accuracy is higher than that of the model using only frontal or lateral chest X-ray images.
- The fusion model includes residual block and channel attention mechanism. The residual block alleviates the vanishing gradient problem, and the channel attention mechanism gives different weights to different feature maps. Experiments show that those strategies can improve the pneumothorax recognition accuracy of the fusion model.
- The fusion model proposed in this paper only needs image-level annotation to achieve high pneumothorax recognition accuracy.

The rest of this study is organized as follows. In Section II, the literature on the automatic diagnosis of pneumothorax is introduced. In Section III, the specific structure and classification results of the fusion model, including the details of the dataset, the hyperparameters, and the comparison with the single input model, are described. In Section IV, a comparison of our model to models from excellent articles in recent years is given. Discussions, conclusions, and future work are given in Sections V and VI.

II. RELATED WORKS

Early automatic pneumothorax detection methods relied on traditional feature extraction techniques. Hough transform [22] was used to model the appearance of pneumothorax in X-ray images and local intensity histograms, image edge detection was used to catch the visceral pleural edge [23], and texture information was used to quantify pulmonary vascular markers [24]. Because the predefined appearance features cannot capture a variety of human lungs and pneumothorax, the diagnostic accuracy of such algorithms is still relatively low. The development of deep learning has introduced a new approach for the automatic diagnosis of pneumothorax. The deep learning algorithm can be used to train the model to classify X-ray images of lungs with pneumothorax and without pneumothorax [25, 26]. Cicero et al. [27] pioneered pneumothorax diagnosis. They used GoogLeNet to detect five common lung diseases using more than 35,000 adult chest X-ray images. The accuracy of pneumothorax detection was 0.86 AUC (area under the curve). Taylor et al. [28] compared the performance of Inception, VGG, and ResNet neural network architectures, and the AUC obtained for pneumothorax detection was 0.94. Rajpurkar et al. [29] later proved that the performance of ResNet is statistically equivalent to the ability of radiologists to diagnose pneumothorax through chest X-ray images. Park et al. [30] applied the YOLO series network to identify traumatic pneumothorax after chest puncture. Xiyue Wang et al. [14] used a multitask training strategy to improve the accuracy of pneumothorax recognition. The AUC obtained in pneumothorax detection was 0.9786. The highest AUC value was achieved by Xiasong Wang et al. [31], who combined medical reports with X-ray images from the same patient to achieve automatic pneumothorax classification. The AUC obtained in pneumothorax detection was 0.995. To protect the privacy of patients, they did not publish these reports.

| Related work          | Task                | Annotation            | AUC    |
|-----------------------|---------------------|-----------------------|--------|
| Cicero et al. [27]    | Multi-classification| Image level           | 0.86   |
| Taylor et al. [28]    | Binary classification| Image level           | 0.94   |
| Rajpurkar et al. [29] | Multi-classification| Image level           | 0.94   |
| Park et al. [30]      | Binary classification| Bounding box          | 0.905  |
| Xiyue Wang et al. [14]| Binary classification| Pixel level           | 0.9786 |
| Xiasong Wang et al. [31]| Multi-classification| Medical reports       | 0.995  |
| Our work              | Binary classification| Image level           | 0.979  |

III. EXPERIMENTS

A. OBJECTIVES

The work presented in [14] and [31] is better than that of most other work on pneumothorax recognition. The former used pixel-level annotation, while the latter used medical reports to detect pneumothorax. However, these two types of annotations are expensive. How to use low-cost image-level annotation to achieve high-precision pneumothorax recognition is our main objective. Different sizes of
pneumothorax may be presented in the front chest X-ray image and the lateral chest X-ray image of the same patient. Frontal chest X-ray images and lateral chest X-ray images have their own advantages. Therefore, we aim to use a neural network to fuse frontal chest X-ray image information and lateral chest X-ray image information to improve the accuracy of pneumothorax recognition.

B. DATASETS AND DATA ENHANCEMENT METHODS

1) DATASET
In cooperation with Nanyang Central Hospital in Henan Province, China, we collected 2,530 pairs of chest X-ray images from patients of different ages and gender (each pair consists of a frontal and lateral chest X-ray image from the same patient). The dataset contains a total of 5,060 digital imaging and communications in medicine (DICOM) files. Each DICOM file includes the patient’s protected health information (PHI), including name, sex, age, and image-related information. Specifically, there were 1,670 negative pneumothorax cases and 860 positive pneumothorax cases, which were labeled by experienced radiologists. Figure 1 shows X-ray images of one patient taken from different angles. The specific data distribution is shown in Figure 2 and Figure 3.

FIGURE 1. An example of the chest X-ray images of one patient from the Haut-NY dataset. The first line includes four frontal chest X-ray images. The second line contains four lateral chest X-ray images.

Gender and Pneumothorax distributions

![Gender and Pneumothorax distributions](image)

FIGURE 2. The distribution of sex and pneumothorax for the Haut-NY dataset.

![Gender and Pneumothorax distributions](image)

FIGURE 3. Age distribution for the Haut-NY dataset.
2) IMAGE AND METHOD ENHANCEMENTS

Data preprocessing: First, Pydicom was used to convert all DICOM files into PNG files. The original image resolution was 3200 × 3200 pixels. We zoomed in on the image to achieve a 224 × 224 pixel resolution. The advantage of this resolution is that relatively few model parameters are required and a fast training speed can be achieved. In addition, we also attempted to use resolutions of 768 × 768 and 1024 × 1024 pixels. We used bilinear interpolation[32, 33] to ensure the quality of the reduced image as much as possible.

Considering the limited number of chest X-ray images, we used the Albumentations image enhancement tool, which is a fast training data enhancement library for OpenCV, to enhance the image. It has a very simple and powerful interface that can be used for various tasks, such as classification, segmentation, and detection. In addition, it is easy to customize and add other frameworks and can be used to convert the dataset pixel by pixel. The specific image enhancement methods used are shown in Table 2.

Because our model needs to fuse chest X-ray images of two different views of the same patient (i.e., frontal and lateral views), the data enhancement of the two images must be consistent. Otherwise, convergence cannot occur in our deep learning model.

![Figure 4](image)

**FIGURE 4.** An example of chest X-ray images of a patient from the Haut-NY dataset. The first row shows the original frontal chest X-ray images, and the second row shows the frontal chest X-ray images after image enhancement. The third row shows the original lateral chest X-ray images, and the fourth row shows the lateral chest X-ray images after image enhancement. A consistent data enhancement method must be maintained for a pair of images (i.e., the frontal and lateral chest X-ray images of the same patient).

**TABLE 2.** The image enhancement methods used in this experiment are as follows. The parameter "p" represents probability. For example, the probability of "blur" is 0.8, and the probability of no "blur" is 0.2. For the training dataset, we used nine methods to increase the model’s generalization ability. For the test dataset, only the normalization method was used because for a real diagnosis, medical images will not be rotated or blurred.

| Method            | Related parameters | p     |
|-------------------|--------------------|-------|
| RandomGamma       | gamma_limit=(60, 120) | 0.5   |


RandomBrightness limit=0.2 p=0.8
RandomContrast limit=0.2 p=0.8
Blur blur_limit=4 p=0.8
MotionBlur blur_limit=4 p=0.8
HorizontalFlip clip_limit=4.0 tile_grid_size=(4, 4) p=0.5
CLAHE clip_limit=4.0 tile_grid_size=(4, 4) p=0.5
ShiftScaleRotate shift_limit=0.2 scale_limit=0.2 rotate_limit=20 p=0.5
Normalize mean=(0.485, 0.456, 0.406) std=(0.229, 0.224, 0.225) max_pixel_value=255.0 p=1

C. EVALUATED INDICATORS

True positive (TP): Positive samples are correctly classified by the model;
False negative (FN): Positive samples are incorrectly classified by the model;
False positive (FP): Negative samples are incorrectly classified by the model;
True negative (TN): Negative samples are correctly classified by the model;
The area under the curve (AUC) is defined as the area under the ROC curve. Researchers often use the AUC value as the evaluation standard of the model because the ROC curve cannot clearly explain which classifier is better. In contrast, the classifier with a larger AUC is better.

Formulas 1-6 are the six indicators used in this experiment to evaluate the performance of the classification model.

\[
\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}
\]

\[
\text{Recall} = \frac{TP}{TP+FN} \tag{2}
\]

\[
\text{Precision} = \frac{TP}{TP+FP} \tag{3}
\]

\[
\text{Macro avg} = \frac{P_{no}+P_{yes}}{2} \tag{4}
\]

\[
F1\text{- Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}
\]

\[
\text{Weighted avg}=P_{no} \times (support_{no} + support_{all}) + P_{yes} \times (support_{yes} + support_{all}) \tag{6}
\]

D. SINGLE INPUT MODEL

In this experiment, we only used 2,530 frontal chest X-ray images. The label of the positive pneumothorax image is set to 1, and the negative pneumothorax image is set to 0. The model only used the frontal branch (i.e., single input, single output), as shown in Figure 6, to input the frontal chest X-ray images and output the binary classification results. We attempted to
use different pretraining networks. We found that ResNet-50 always achieved the highest accuracy (Table 3). Therefore, we used ResNet-50 pretrained on ImageNet as the backbone. The input size was 224 \times 224 \times 3. The data were flattened using the flattening layer. Then, the image information was parsed using four fully connected layers with the ReLU activation function. The parameter of the final fully connected layers was 2 (representing the binary classification), and the softmax activation function was used to convert the output value of the binary classification into probability distributions in the range of [0,1]. The loss function was binary cross-entropy.

After the first set of experiments, we used only 2,530 lateral chest X-ray images for the next experiment. The model only used the lateral branches (i.e., single input and single output), as shown in Figure 6, to input the lateral chest X-ray images and output the binary classification results. The training details of the model were consistent with the previous single input single output forward chest X-ray image model.

E. MULTI-INPUT MODEL

Different from traditional learning only through frontal chest X-ray images or lateral X-ray images, we wanted to design a multi-input network that could integrate frontal chest X-ray image and chest lateral X-ray image information to diagnose pneumothorax. The model pretrained on ImageNet was used as a tool to extract image features. We attempted to use several different pretraining models as the backbone. We found that when ResNet-50 was used as the backbone, the model's accuracy was the highest (Table 6). We also found that the characteristics of the ResNet-50 network determine the accuracy. In training the model, we found that one of the challenges is the vanishing gradient problem[34], which occurs when the network is deep. The deeper the network is, the more obvious the vanishing gradient, and the poorer the training effect of the network. However, the shallow network cannot significantly improve network performance. This is a contradictory problem, but the residual block[35] (an important module in ResNet-50.) effectively alleviates the vanishing gradient in a deeper network.

![Figure 5. Skip connection in the residual network.](image)

The relevant equations of the channel attention mechanism are as follows:

\[
\tilde{u}_c = v_c \ast X = \sum_{s=1}^{C} V_s \ast X^s 
\]

(10)

\[
z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i, j) 
\]

(11)

\[
s = F_{scale}(z, W) 
\]

(12)

\[
\hat{X}_c = F_{scale}(u_c, s_c) 
\]

(13)

In our work, frontal and lateral chest X-ray images were combined and two-channel attention mechanisms were used. In the first step, the global spatial characteristics of each channel were taken by the squeeze operation as the representation of the channel to form a channel descriptor. In the second step, the dependence on each channel was learned and the feature map was adjusted according to the dependence. The modified feature map was the output of the SE block. The benefits of SE block reprofiling could be accumulated throughout the network. The concat function was used to fuse information of frontal and lateral chest X-ray images, and
finally, a flatten layer was used to flatten the data. Fully connected layers (FCs) played the role of classifiers in the whole convolutional neural network. Finally, the binary classification results were output using the softmax function (Formula 14).

\[
P_C = \frac{\exp(f_x)}{\sum_c \exp(f_c)}
\] (14)

Our model has two inputs and three outputs. The two inputs are the frontal chest X-ray image and the lateral chest X-ray image. The three outputs are the classification results of the frontal chest X-ray image, the classification results of the lateral chest X-ray image, and the classification results integrating the characteristics of the fused frontal chest X-ray image and lateral chest X-ray image. Our fusion model can fuse frontal feature maps and lateral feature maps of chest X-ray images and then learn them.

\[
F_{\text{Total}} = F_{\text{Flatten}}\left[\text{Concat}(F_{\text{F-feature map}}, F_{\text{L-feature map}})\right]
\] (15)

The loss function was binary cross-entropy (Formula 16). Using cross-entropy as the loss function can alleviate the imbalance between positive and negative samples to a certain extent, and the calculated gradient is more stable[37]. The loss function adopted by the three branches was binary cross-entropy, but the weighting indices of the three branches are different (Formula 17).

\[
L_{\text{BCE}} = -\sum_i [\hat{y}_i \log(y_i) + (1 - \hat{y}_i) \log(1 - y_i)]
\] (16)

\[
L_{\text{Total}} = \lambda_{\text{Fused}} L_{\text{Fused-BCE}} + \lambda_{\text{Frontal}} L_{\text{Frontal-BCE}} + \lambda_{\text{Lateral}} L_{\text{Lateral-BCE}}
\] (17)

FIGURE 6. Pneumothorax recognition neural network based on feature fusion.

FIGURE 7. Squeeze-and-Excitation Module and Concat.
All experiments in this study adopted the 5-fold-cross-validation method, where the data were divided into five equal parts on average. Each data point was completely independent and did not cross other data points. One part of the data was taken for testing in each experiment, and the rest was used for training. The average value was obtained for five experiments.

**Algorithm: 5-fold cross-validation**

**Input:** Preprocess pneumothorax dataset, then the dataset was randomly divided into 5 parts on average:

\[ D = D_1 \cup D_2 \cup D_3 \cup D_4 \cup D_5 \]

For \( \alpha \leq 5 \)

1. Train model using \( D - D_\alpha \);
2. Test model using \( D_\alpha \);
3. Save score as \( R \);
4. \( \alpha++ \);
5. Output: \( \sum R_i \)

---

**F. EXPERIMENTAL DETAILS AND OPTIMIZATION METHODS**

The following experimental hyperparameter settings were utilized. The Adam optimizer was used, the batch size was 32, and the initial learning rate was 0.0001. The epoch was 60. If the model’s performance was not improved after five epochs were trained continuously, the learning rate was decreased to one-tenth of its original value. If the performance after 15 epochs was not enhanced, “early stop” was initiated to prevent overfitting. All networks were implemented based on the TensorFlow framework and trained using NVIDIA GeForce RTX 3080 Ti GPU cards.

**G. CLASSIFICATION RESULTS**

1) **SINGLE INPUT MODEL CLASSIFICATION RESULTS**

The classification results of frontal chest X-ray images are shown in Table 4. First, for the binary classification of the frontal images, based on the 334 nonpneumothorax images, a recognition precision of 0.89 was achieved, and based on the 172 pneumothorax images a recognition precision of 0.82 was achieved. The recognition accuracy was 0.87.

The classification results of lateral chest X-ray images are shown in Table 5. First, for the binary classification of the frontal images, based on the 334 nonpneumothorax images, a recognition precision of 0.90 was achieved, and based on the 172 pneumothorax images, a recognition precision of 0.77 was achieved. The recognition accuracy was 0.85.

**TABLE 3.** Prediction results of single input models under different pre-training networks (AUC was the indicator).

| Pretraining model                        | AUC  |
|-----------------------------------------|------|
| Resnet50 (frontal image)                | 0.932|
| Resnet50 (lateral image)                | 0.924|
| VGG16 (frontal image)                   | 0.918|
| VGG16 (lateral image)                   | 0.920|
| EfficientNet-b4 (frontal image)         | 0.924|
| EfficientNet-b4 (lateral image)         | 0.924|
| Inceptionv3 (frontal image)             | 0.922|
| Inceptionv3 (lateral image)             | 0.919|
| Inceptionv4 (frontal image)             | 0.930|
| Inceptionv4 (lateral image)             | 0.923|
| DenseNet-121 (frontal image)            | 0.931|
| DenseNet-121 (lateral image)            | 0.923|

**TABLE 4.** Comprehensive performance index of the pneumothorax binary classification for frontal chest X-ray images (single input, single output).

|               | Precision | Recall | F1-score | Number |
|---------------|-----------|--------|----------|--------|
| Nonpneumothorax | 0.89      | 0.91   | 0.90     | 334    |
| Pneumothorax   | 0.82      | 0.79   | 0.81     | 172    |
| Macro avg      | 0.86      | 0.85   | 0.86     | 506    |
| Weighted avg   | 0.87      | 0.87   | 0.87     | 506    |

**TABLE 5.** Comprehensive performance index of the pneumothorax binary classification for lateral chest X-ray images (single input, single output).

|               | Precision | Recall | F1-score | Number |
|---------------|-----------|--------|----------|--------|
| Nonpneumothorax | 0.90      | 0.87   | 0.89     | 334    |
| Pneumothorax   | 0.77      | 0.81   | 0.79     | 172    |
| Macro avg      | 0.83      | 0.84   | 0.84     | 506    |
| Weighted avg   | 0.85      | 0.85   | 0.85     | 506    |

2) **MULTI-INPUT MODEL CLASSIFICATION RESULTS**

For the case of multiple inputs, the classification results of the frontal branch are shown in Table 7. For the 334 images from patients without pneumothorax, a recognition precision of 0.93 was achieved, and for the 172 images from patients with pneumothorax, a recognition precision of 0.85 was achieved. The recognition accuracy was 0.91. In the case of multiple inputs, the classification results of the lateral branch are shown in Table 8. For the 334 images from patients without pneumothorax, a recognition precision of 0.91 was achieved, and for the 172 images from patients with pneumothorax, a recognition precision of 0.87 was achieved. The recognition accuracy was 0.90. Compared with a single input, all values were improved to varying degrees.

The most obvious improvement in the accuracy was a result of the fusion branch. In the case of multiple inputs (without
channel attention), the classification results of the frontal branch are shown in Table 9. For the 334 images from patients without pneumothorax, a recognition precision of 0.95 was achieved, and for the 172 images from patients with pneumothorax, a recognition precision of 0.88 was achieved. The recognition accuracy was 0.92. Both the macro avg and weighted avg indices improved to varying degrees. In the case of combining the multiple inputs model and channel attention, the classification results of the frontal branch are shown in Table 10. For the 334 images from patients without pneumothorax, a recognition precision of 0.93 was achieved, and for the 172 images from patients with pneumothorax, a recognition precision of 0.95 was achieved. The recognition accuracy was 0.94. Both macro avg and weighted avg indices improved to varying degrees.

**TABLE 7.** The comprehensive performance index of the pneumothorax binary classification for frontal chest X-ray images (two inputs, three outputs).

|                  | Precision | Recall | F1-score | Number |
|------------------|-----------|--------|----------|--------|
| Nonpneumothorax  | 0.93      | 0.92   | 0.93     | 334    |
| Pneumothorax     | 0.85      | 0.87   | 0.86     | 172    |
| Accuracy         |           |        | 0.91     | 506    |
| Macro avg        | 0.89      | 0.90   | 0.89     | 506    |
| Weighted avg     | 0.91      | 0.91   | 0.91     | 506    |

**TABLE 8.** The comprehensive performance index of the pneumothorax binary classification for lateral chest X-ray images (two inputs, three outputs).

|                  | Precision | Recall | F1-score | Number |
|------------------|-----------|--------|----------|--------|
| Nonpneumothorax  | 0.91      | 0.94   | 0.92     | 334    |
| Pneumothorax     | 0.87      | 0.83   | 0.85     | 172    |
| Accuracy         |           |        | 0.90     | 506    |
| Macro avg        | 0.89      | 0.88   | 0.89     | 506    |
| Weighted avg     | 0.90      | 0.90   | 0.90     | 506    |

**TABLE 9.** The comprehensive performance index of the pneumothorax binary classification for fused images (two inputs, three outputs, and without channel attention).

|                  | Precision | Recall | F1-score | Number |
|------------------|-----------|--------|----------|--------|
| Nonpneumothorax  | 0.95      | 0.94   | 0.94     | 334    |
| Pneumothorax     | 0.88      | 0.90   | 0.89     | 172    |
| Accuracy         |           |        | 0.92     | 506    |
| Macro avg        | 0.91      | 0.92   | 0.92     | 506    |
| Weighted avg     | 0.93      | 0.92   | 0.93     | 506    |

**TABLE 10.** The comprehensive performance index of the pneumothorax binary classification for fused images (two inputs, three outputs, and with channel attention).

|                  | Precision | Recall | F1-score | Number |
|------------------|-----------|--------|----------|--------|
| Nonpneumothorax  | 0.93      | 0.98   | 0.95     | 334    |
| Pneumothorax     | 0.95      | 0.87   | 0.91     | 172    |
| Accuracy         |           |        | 0.94     | 506    |
| Macro avg        | 0.94      | 0.92   | 0.93     | 506    |
| Weighted avg     | 0.94      | 0.94   | 0.94     | 506    |
TABLE 11. Area under the curve (AUC), 95% confidence interval (CI), and convergence time with different strategies.

| Method | Mean AUC (95%CI) | Mean Convergence Time |
|--------|-----------------|-----------------------|
| (1) Frontal chest X-ray image (Single input model) | 0.932 (0.928-0.935) | 5.5 hours |
| (2) Lateral chest X-ray image (Single input model) | 0.924 (0.921-0.927) | 5.7 hours |
| (3) Frontal chest X-ray image (Frontal branch) | 0.956 (0.953-0.959) | 10.3 hours |
| (4) Lateral chest X-ray image (Lateral branch) | 0.950 (0.948-0.952) | 10.2 hours |
| (5) Feature fusion (Fused branch) | 0.962 (0.958-0.964) | 10.7 hours |
| (6) Feature fusion + Channel attention (Fused branch) | 0.975 (0.974-0.977) | 11.1 hours |

FIGURE 8. Pneumothorax classification confusion matrix of the frontal images (Single input), lateral images (Single input) and fused branch (With channel attention).

H. SOME FACTORS AFFECTING THE ACCURACY OF THE MODEL

1) INFLUENCE OF MODEL TRAINING SEQUENCE ON ACCURACY

We also found that training three branches in different sequences will affect the performance of the fusion model. We trained our fusion model with four different sequences.

1) Three branches were trained at the same time (train together).
2) First, the frontal branch was trained, and then the lateral branch and fused branch were trained (1-2-3).
3) First, the lateral branches were trained, and then the frontal branch and fused branch were trained (2-1-3).
4) The frontal and lateral branches were trained together, and then fused branch was trained (1 2-3)

TABLE 12. AUC under different training sequences (ResNet-50 was used as the backbone).

| Training sequence | Frontal branch (1) | Lateral branch (2) | Fused branch (3) |
|-------------------|--------------------|--------------------|-----------------|
| Train together    | 0.956              | 0.950              | 0.975           |
| 1-2-3             | 0.960              | 0.949              | 0.971           |
| 2-1-3             | 0.953              | 0.954              | 0.969           |
| 1 2-3             | 0.952              | 0.955              | 0.973           |

2) INFLUENCE OF WEIGHT RATIO OF LOSS FUNCTION ON MODEL ACCURACY

Because our model contains three outputs, the weight ratio between the three loss functions must be considered. In fact, in our experiment, different weight ratios affect the accuracy of the model to a certain extent.

TABLE 13. AUC under different training weight ratios. (ResNet-50 was used as the backbone, and the training sequence was trained together).

| Weight ratio (Fused: Frontal: Lateral) | Frontal branch | Lateral branch | Fused branch |
|---------------------------------------|---------------|---------------|-------------|
| 1:0:2:0.2                             | 0.906         | 0.892         | 0.951       |
| 1:0:4:0.4                             | 0.906         | 0.896         | 0.960       |
| 1:0:6:0.6                             | 0.956         | 0.950         | 0.975       |
| 1:0:8:0.8                             | 0.959         | 0.955         | 0.963       |
| 1:1:1                                 | 0.955         | 0.956         | 0.959       |

We found that when the weight of the loss function is set to 1:0:6:0.6 (fused:frontal:lateral), the fused branch achieved the highest AUC.
We attempted to analyze the connection between the frontal branch, lateral branch, and fusion branch from the perspective of statistics. Softmax was used in the process of binary classification to map the output of multiple neurons into the (0,1) interval, which can be understood as probability, to carry out binary classification when we input a chest X-ray image into the model. The softmax function output of the frontal branch was [0.49, 0.51], which means that the probability of the corresponding chest X-ray image showing nonpneumothorax is 0.49, and the probability of pneumothorax is 0.51. The softmax function output of the lateral branch is [0.01, 0.49], and the probability of pneumothorax is 0.51. The softmax function output of the lateral branch is [0.01, 0.49], which means that the probability that the lateral branch assesses the corresponding chest X-ray image is nonpneumothorax is 0.01, and the probability of pneumothorax is 0.99.

Let the softmax function output of the frontal branch be [X\text{Frontal}, Y\text{Frontal}], the softmax function output of the lateral branch be [X\text{Lateral}, Y\text{Lateral}], and the softmax function output of the fusion branch be [X\text{Fused}, Y\text{Fused}]. The connection between the three is shown in Table 14, where \( P (Y_{\text{Fused} > 0.5}) \) represents the probability that the corresponding image of fusion branch recognition is positive pneumothorax.

### TABLE 14. The relationship between the activation functions of the three branches.

| Y\text{Frontal} + Y\text{Lateral} | P (Y_{\text{Fused} > 0.5}) |
|----------------------------------|---------------------------|
| (0,02)                           | 0                         |
| (0.02,0.2)                      | 0.08                      |
| (0.2,0.4)                        | 0.21                      |
| (0.4,0.6)                        | 0.36                      |
| (0.6,0.8)                        | 0.44                      |
| (0.8,1.0)                        | 0.63                      |
| (1.0,1.2)                        | 0.89                      |
| (1.2,1.4)                        | 0.92                      |
| (1.4,1.6)                        | 0.95                      |
| (1.6,1.8)                        | 0.97                      |
| (1.8,2.0)                        | 0.99                      |

### IV. COMPARISON WITH OTHER WORK

TABLE 15. Comparison with the work of other researchers in recent years (binary classification of pneumothorax).

| Articles | Datasets and details | Methods | Image size | AUC | Accuracy | Precision | Recall | F1-score |
|----------|----------------------|---------|------------|-----|----------|-----------|--------|----------|
| Taylor et al.[38] (2018) | Private dataset, Image level annotation | VGG19 | 512*512 | 0.94 | --- | --- | --- | --- |
| Wang Y et al.[39] (2019) | Private dataset, ChestX-ray14, Image level annotation, Only frontal images | VGG16 + Image augmentation | 480*480 | 0.9006 | 0.9451 | --- | --- | --- |
| X Wang et al.[14] (2020) | SIIM-ACR, Image level annotation, Only frontal images | Multitask + EfficientNet-B3 | 1024*1024 | 0.9786 | 0.9467 | 0.8206 | 0.9189 | 0.8670 |
| This article | Image level annotation, Frontal image + Lateral image | Fusion model | 768*768 | 0.9763 | 0.9446 | 0.8068 | 0.9212 | 0.8601 |

Relevant articles from recent years on pneumothorax classification were divided into two categories. The first category was multiclassification, which usually predicts 5-14 different chest diseases, and the AUC for pneumothorax recognition ranges from 0.80 to 0.92[19, 27, 40-42]. The second category of articles was the binary classification of pneumothorax[14, 38, 39]. The work related to the binary classification of pneumothorax was the main comparison for this experiment.

Taylor et al.[38] established a private pneumothorax dataset. They used several different pretraining networks for
comparative experiments. The average AUC obtained in pneumothorax detection was 0.94. Their main contribution was to analyze the differences in recognizing various severities of pneumothorax, such as nonpneumothorax, trace pneumothorax, moderate or large pneumothorax. In the case of unbalanced pneumothorax datasets, Wang Y et al.[39] used the positive image augmentation method to increase the positive pneumothorax cases and improve recognition accuracy. When the ratio of the original positive pneumothorax sample to the negative pneumothorax sample was 1507:26891, the model's accuracy was only 50.08%. Then, they increased the data of positive pneumothorax cases to 27126:26891 (no new positive images were added, only image augmentation was used), and the model's accuracy improved to 88.52%. However, only using image augmentation to expand the positive samples may risk overfitting without adding new images. It may only improve the recognition accuracy of the current dataset, while in the face of new datasets, the recognition accuracy of the model may be reduced. The strategy by X. Wang et al.[14] was to use pixel-level pneumothorax annotation combined with a multitask learning model to achieve high-precision pneumothorax classification. Their work was very comprehensive. Among the five evaluated indicators in Table 15, AUC is the most valuable because AUC can resist the imbalance of the number of samples to a certain extent. The accuracy, precision, recall, and F1-score are vulnerable to data imbalance (i.e., the number of positive and negative pneumothorax cases).

V. DISCUSSION

When accuracy was the evaluated indicator, using the fusion network improved the accuracy by approximately 7% compared with using only the frontal chest X-ray image information and approximately 9% more accurate than using only lateral chest X-ray image information. When AUC was the evaluated indicator, our fusion model was 4% higher than that using only frontal or lateral models. In fact, by comparing the data in Table 11, we found that no matter which performance index was used as the evaluation, the performance of the fusion network was the best.

Our further experiments showed that the accuracy of pneumothorax recognition was related to the training sequence of the model and the weight ratio of the loss function. When the training sequence was three branches trained simultaneously and the weight of the loss function was 1:0:6:0.6 (fused:frontal: lateral), the model's accuracy was the highest.

Referring to Table 15, using only image-level annotation, our model still achieved high pneumothorax recognition accuracy.

VI. CONCLUSION AND FUTURE DIRECTIONS

We proposed a pneumothorax binary classification neural network based on feature fusion. This was meaningful because most of the literature on pneumothorax recognition only considers frontal chest X-ray images. Our model could be used to fuse frontal and lateral X-ray information to achieve higher precision pneumothorax recognition. The model design considered the phenomenon of the vanishing gradient in deeper neural networks, so we introduced the residual block to alleviate it. There were too many channels in the feature map after feature fusion, so the channel attention mechanism SeNet was used to adjust the feature map. Comparative experiments showed that the accuracy of this method was higher than that of the traditional single task pneumothorax recognition network. The main value of our work is that only using image-level datasets can achieve high pneumothorax recognition accuracy. However, we need pairs of image-level annotations (frontal annotation + lateral annotation) rather than only front images. Therefore, our dataset is more expensive than the dataset containing only frontal images. This is a limitation of our method. However, the cost is still lower than that of pixel-level annotations, even for paired image-level annotations. Therefore, the proposed method may assist radiologists with the prompt and accurate diagnosis of pneumothorax and precise treatment planning.

Future work is as follows. First, we will improve the dataset, expand the number of images, invite experienced radiologists to add pixel-level annotations to our dataset, and conduct in-depth research around visualization techniques. Second, we will combine our model with the expert system[43] and the fuzzy consensus[44], which play an important role in artificial intelligence-aided diagnosis and the internet of medical things.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (11005136). We thank Nanyang Central Hospital for providing us with the dataset.

REFERENCES

[1] K. M. Yeow et al., "Risk factors of pneumothorax and bleeding: multivariate analysis of 660 CT-guided coaxial cutting needle lung biopsies," *Chest*, vol. 126, no. 3, pp. 748-754, 2004.

[2] M. Blaivas, M. Lyon, and S. J. A. E. M. Duggal, "A prospective comparison of supine chest radiography and bedside ultrasound for the diagnosis of traumatic pneumothorax," *Academic Emergency Medicine*, vol. 12, no. 9, pp. 844-849, 2005.

[3] T. Roth, M. Alifano, O. Schussler, P. Magdaleinat, and J. F. J. T. A. o. T. S. Regnard, "Catamenial pneumothorax: Chest X-ray sign and thoracoscopic
treatment," *The Annals of Thoracic Surgery*, vol. 74, no. 2, pp. 563-565, 2002.

[4] A. J. Salazar, D. A. Aguirre, J. Ocampo, J. C. Camacho, and X. D. J. o. D. Imaging, "Evaluation of Three Pneumothorax Size Quantification Methods on Digitized Chest X-ray Films Using Medical-Grade Grayscale and Consumer-Grade Color Displays," *Journal of Digital Imaging*, vol. 27, no. 2, pp. 280-286, 2014.

[5] K. Suzuki, L. Feng, S. Sone, and K. J. I. T. o. M. I. Doi, "Computer-aided diagnostic scheme for distinction between benign and malignant nodules in thoracic low-dose CT by use of massive training artificial neural network," *IEEE Transactions on Medical Imaging*, vol. 24, no. 9, pp. 1138-1150, 2005.

[6] S. Sanada, K. Doi, and H. J. M. P. Macmahon, "Image feature analysis and computer - aided diagnosis in digital radiography: Automated detection of pneumothorax in chest images," *Medical Physics*, vol. 19, 1992.

[7] J. Gao, J. Yang, J. Zhang, and M. Li, "Natural scene recognition based on Convolutional Neural Networks and Deep Boltzmann Machines," in *IEEE International Conference on Mechatronics & Automation*, 2015.

[8] D. Połap, N. Wawrzyniak, M. J. I. T. o. G. Włodarczyk-Sielicka, and R. Sensing, "Side-scan sonar analysis using roi analysis and deep neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, 2022.

[9] L. Wang, C. Wang, Z. Sun, and S. J. I. A. Chen, "An Improved Dice Loss for Pneumothorax Segmentation by Mining the Information of Negative Areas," *IEEE Access*, vol. 8, no. 99, pp. 167939 - 167949, 2020.

[10] S. Bharati, P. Podder, and M. J. I. i. M. U. Mondal, "Hybrid deep learning for detecting lung diseases from X-ray images," *Informatics in Medicine Unlocked*, vol. 20, 2020.

[11] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481-2495, Dec 2017.

[12] D. A. J. o. M. I. Moses and R. Oncology, "Deep learning applied to automatic disease detection using chest X - rays," *Journal of Medical Imaging and Radiation Oncology*, vol. 65, no. 5, pp. 498-517, 2021.

[13] M. Yahyatabar, P. Jouvet, and F. Cheriet, "Dense-Unet: a light model for lung fields segmentation in Chest X-Ray images," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2020.

[14] X. Wang, S. Yang, J. Lan, Y. Fang, X. J. I. T. o. C. Han, and D. Systems, "Automatic Segmentation of Pneumothorax in Chest Radiographs Based on a Two-stage Deep Learning Method," *IEEE Transactions on Cognitive and Developmental Systems*, no. 99, pp. 1-1, 2020.

[15] Z. Yu and Y. Qiang, "A Survey on Multi-Task Learning," *IEEE Transactions on Knowledge and Data Engineering*, 2017.

[16] A. Malhotra et al., "Multi-task driven explainable diagnosis of COVID-19 using chest X-ray images," *Pattern Recognition*, vol. 122, p. 108243, 2022.

[17] A. Ms and C. J. P. R. Mshb, "MetaCOVID: A Siamese neural network framework with contrastive loss for n - shot diagnosis of COVID-19 patients," *Pattern Recognition*, vol. 113, 2020.

[18] K. He, C. Lian, B. Zhang, X. Zhang, and D. J. I. T. o. M. I. Shen, "HF-UNet: Learning Hierarchically Inter-Task Relevance in Multi-Task U-Net for Accurate Prostate Segmentation in CT images," *IEEE Transactions on Medical Imaging*, vol. PP, no. 99, pp. 1-1, 2021.
[19] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. J. I. Summers, "ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
[20] https://www.kaggle.com/nih-chest-xrays/data.
[21] https://www.kaggle.com/c/siim-acr-pneumothorax-segmentation.
[22] S. Sanada, K. Doi, and H. J. M. P. Macmahon, "Image feature analysis and computer-aided diagnosis in digital radiography: Automated detection of pneumothorax in chest images," Medical Physics, vol. 19, no. 5, pp. 1153-1160, 1992.
[23] O. Geva, G. Zimmerman-Moreno, S. Lieberman, E. Konen, and H. Greenspan, "Pneumothorax detection in chest radiographs using local and global texture signatures," in Medical Imaging 2015: Computer-Aided Diagnosis, 2015.
[24] Y. H. Chan, Y. Z. Zeng, H. C. Wu, M. C. Wu, and H. M. J. O. H. E. Sun, "Effective Pneumothorax Detection for Chest X-Ray Images Using Local Binary Pattern and Support Vector Machine," Journal of Healthcare Engineering, vol. 2018, pp. 1-11, 2018.
[25] G. Kitamura and C. J. C. I. Deible, "Retraining an open-source pneumothorax detecting machine learning algorithm for improved performance to medical images," Clinical Imaging, vol. 61, pp. 15-19, 2020.
[26] R. W. Filice, A. Stein, C. C. Wu, V. A. Arteaga, and G. J. J. O. D. I. Shih, "Crowdsourcing pneumothorax annotations using machine learning annotations on the NIH chest X-ray dataset," Journal of Digital Imaging, vol. 33, no. 4, pp. 490-496 2019.
[27] Cicero et al., "Training and Validating a Deep Convolutional Neural Network for Computer-Aided Detection and Classification of Abnormalities on Frontal Chest Radiographs," Invest Radiol, vol. 52, no. 5, pp. 281-287, 2017.
[28] A. G. Taylor, C. Mielke, and J. J. P. M. Mongan, "Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study," PLOS Medicine, vol. 15, no. 11, 2018.
[29] P. Rajpurkar et al., "Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists," PLOS Medicine, vol. 15, no. 11, 2018.
[30] S. M. L. Sohee Park "Application of deep learning-based computer-aided detection system: detecting pneumothorax on chest radiograph after biopsy," European radiology, vol. 29, no. 10, pp. 5341-5348, 2019.
[31] X. Wang, Y. Peng, L. Lu, Z. Lu, and R. M. Summers, "TieNet: Text-Image Embedding Network for Common Thorax Disease Classification and Reporting in Chest X-rays," in IEEE CVPR 2018, 2018.
[32] E. J. Kirkland, "Bilinear interpolation," in Advanced Computing in Electron Microscopy: Springer, 2010, pp. 261-263.
[33] Guo and S.-M. J. O. E. Imaging, "Fast pixel-size-based large-scale enlargement and reduction of image: adaptive combination of bilinear interpolation and discrete cosine transform," Journal of Electronic Imaging, vol. 20, no. 3, pp. 51-51, 2011.
[34] M. Roodschild, J. Gotaysardias, and A. J. P. I. A. I. Will, "A new approach for the vanishing gradient problem on sigmoid activation," Progress in Artificial Intelligence, vol. 9, no. 4, pp. 351-360, 2020.
[35] K. He, X. Zhang, S. Ren, and J. J. I. Sun, "Deep Residual Learning for Image
Recognition," in Conference on Computer Vision and Pattern Recognition 2016.

[36] H. Jie, S. Li, S. Gang, S. J. I. T. o. P. A. Albanie, and M. Intelligence, "Squeeze-and-Excitation Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PP, no. 99, 2017.

[37] A. Semenov, V. Boginski, and E. L. Pasiliao, "Neural Networks with Multidimensional Cross-Entropy Loss Functions," in International Conference on Computational Data and Social Networks, 2019, pp. 57-62: Springer.

[38] A. G. Taylor, C. Mielke, and J. J. P. M. Mongan, "Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study," PLOS Medicine, vol. 15, 2018.

[39] Y. Wang, L. L. Sun, Q. J. I. A. T. o. C. B. Jin, and Bioinformatics, "Enhanced Diagnosis of Pneumothorax with an Improved Real-time Augmentation for Imbalanced Chest X-rays Data Based on DCNN," IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 18, no. 3, pp. 951 - 962, 2019.

[40] B. Chen, J. Li, X. Guo, G. J. B. s. p. Lu, and control, "DualCheXNet: dual asymmetric feature learning for thoracic disease classification in chest X-rays," Biomedical signal processing and control, vol. 53, no. Aug., pp. 101554.1-101554.10, 2019.

[41] P. Kumar, M. Grewal, and M. M. Srivastava, "Boosted cascaded convnets for multilabel classification of thoracic diseases in chest radiographs," in International conference image analysis and recognition, 2018, pp. 546-552: Springer.

[42] C. Yan, J. Yao, R. Li, Z. Xu, and J. Huang, "Weakly supervised deep learning for thoracic disease classification and localization on chest x-rays," in

Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics, 2018, pp. 103-110.

[43] A. Saibene, M. Assale, and M. J. E. S. w. A. Giltri, "Expert systems: definitions, advantages and issues in medical field applications," Expert Systems with Applications, vol. 177, p. 114900, 2021.

[44] D. J. I. A. Połap, "Fuzzy Consensus With Federated Learning Method in Medical Systems," IEEE Access, vol. 9, pp. 150383-150392, 2021.