

A Deep Learning Model for the Diagnosis and Discrimination of Gram-Positive and Gram-Negative Bacterial Pneumonia for Children Using Chest Radiography Images and Clinical Information

Ru Wen^{1-3,*}, Peng Xu^{3,*}, Yimin Cai^{1,*}, Fang Wang³, Mengfei Li³, Xianchun Zeng², Chen Liu³

¹Medical College, Guizhou University, Guizhou, 550000, People's Republic of China; ²Department of Medical Imaging, Guizhou Provincial People Hospital, Guiyang City, Guizhou Province, 550000, People's Republic of China; ³Department of Radiology, Southwest Hospital, Army Medical University (Third Military Medical University), Chongqing, 400038, People's Republic of China

*These authors contributed equally to this work

Correspondence: Chen Liu, Department of Radiology, Southwest Hospital, Army Medical University (Third Military Medical University), 30 Gao Tan Yan St, Chongqing, 400038, People's Republic of China, Tel +8613108968808, Email liuchen@aifmri.com; Xianchun Zeng, Department of Medical Imaging, Guizhou Provincial People Hospital, No. 83, East Zhongshan Road, Nanming District, Guiyang City, Guizhou Province, 550000, People's Republic of China, Tel +8615286061024, Email zengxianchun04@foxmail.com

Purpose: This study aimed to develop a deep learning model based on chest radiography (CXR) images and clinical data to accurately classify gram-positive and gram-negative bacterial pneumonia in children to guide the use of antibiotics.

Methods: We retrospectively collected CXR images along with clinical information for gram-positive (n=447) and gram-negative (n=395) bacterial pneumonia in children from January 1, 2016, to June 30, 2021. Four types of machine learning models based on clinical data and six types of deep learning algorithm models based on image data were constructed, and multi-modal decision fusion was performed.

Results: In the machine learning models, CatBoost, which only used clinical data, had the best performance; its area under the receiver operating characteristic curve (AUC) was significantly higher than that of the other models ($P < 0.05$). The incorporation of clinical information improved the performance of deep learning models that relied solely on image-based classification. Consequently, AUC and F1 increased by 5.6% and 10.2% on average, respectively. The best quality was achieved with ResNet101 (model accuracy: 0.75, recall rate: 0.84, AUC: 0.803, F1: 0.782).

Conclusion: Our study established a pediatric bacterial pneumonia model that utilizes CXR and clinical data to accurately classify cases of gram-negative and gram-positive bacterial pneumonia. The results confirmed that the addition of image data to the convolutional neural network model significantly improved its performance. While the CatBoost-based classifier had greater advantages owing to a smaller dataset, the quality of the Resnet101 model trained using multi-modal data was comparable to that of the CatBoost model, even with a limited number of samples.

Keywords: pediatrics, antibiotics, clinical data, X-ray, multi-modal data

Introduction

Pneumonia is the most significant cause of death in children in developing countries.^{1,2} The most common causes of pneumonia are viral and bacterial pathogens. Based on the Global Burden of Disease data, in 2016, approximately 64% of children under the age of five with pneumonia died of bacterial pneumonia.³ Several types of bacteria can cause pneumonia, including *Streptococcus pneumoniae*, *Klebsiella pneumoniae*, *Haemophilus influenzae*, and *Pseudomonas aeruginosa*.^{4,5}

Rapid and accurate identification of pneumonia pathogens, such as gram-positive and gram-negative bacteria, is essential for antibiotic selection. Appropriate drug treatment and limiting the unreasonable use of antibiotics play an important role in improving the survival or cure rates of bacterial pneumonia.⁶ This may improve outcomes and slow the development of antibiotic resistance. Early and accurate identification of bacterial pneumonia pathogens presents a significant challenge for successful treatment.⁷ Symptoms of pneumonia caused by gram-positive and gram-negative bacteria usually overlap, and the etiological diagnosis is mainly based on clinical and diagnostic imaging findings. Accurate classification of pneumonia based on imaging tests alone can be challenging, especially when different types of pneumonia have similar imaging features. Chest radiography (CXR) is a standard diagnostic tool for pneumonia; however, the sensitivity and specificity are not high,⁸ and the interpretation of chest X-ray characteristics varies among clinicians with different levels of experience and ultimately depends on the expertise of the radiologist or clinical physician.⁹ The identification of pneumonia pathogens relies heavily on isolation and culture techniques, which include sputum smears, alveolar lavage, bronchial brushings, and antigen and virus testing. As a standard diagnostic tool for the detection of pathogens, the isolation and culturing process is complex and time-consuming, has many detection restrictions, and has low sensitivity and specificity.⁹ Blood or pleural cultures are highly specific for the diagnosis of bacterial infections, but are less sensitive and take a long time, usually several days.¹⁰ Emerging microbial detection techniques, such as matrix-assisted laser desorption ionization time-of-flight mass spectrometry, can be used to rapidly identify pathogenic microorganisms, such as bacteria, fungi, mycobacteria, and viruses, as well as their drug resistance. However, the technical cost is high and it can only be carried out in qualified units.¹¹ Therefore, early identification of the pathogenic bacteria responsible for pneumonia is particularly important.

Artificial intelligence (AI) can effectively improve the accuracy and speed of pneumonia diagnosis.^{12–14} Existing research on the application of AI technology in the field of pneumonia is still mainly focused on the detection and segmentation of pneumonia lesions and differential diagnosis of pneumonia. A deep learning algorithm-based chest X-ray diagnosis method, CheXNeXt, has an accuracy of diagnosis for 11 common chest diseases, which is comparable to or even better than that of practicing radiologists.¹⁵ Another study presented a deep-learning model for lung-thoracic segmentation based on clinical pediatric CXR of 1135 patients, which achieved an area under the receiver operating characteristic curve (AUC) of 0.95 and a sensitivity of 88.7%. Compared with this algorithm, radiology residents showed lower sensitivity but higher specificity.¹⁶ However, most of the current studies are based on the differential diagnosis between viral pneumonia, especially COVID-19, and other types of pneumonia, such as interstitial pneumonia and community-acquired pneumonia.^{17,18} There are only a few studies on the differential diagnosis of bacterial pneumonia. One study used a computerized tomography image dataset in children (AUC: 0.75, accuracy: 0.58).¹⁹ However, this study was only based on imaging data and did not combine imaging examination and other indicators. We hypothesize that deep learning could be a valuable tool for the differential diagnosis of gram-positive and gram-negative pneumonia, based on previous research demonstrating the effectiveness of machine learning algorithms in medical diagnosis, which is of great importance to guide the choice of empirical anti-infective therapy and reduce antibiotic resistance, thus reducing pneumonia-associated mortality and sequelae in children. AI technology has been proven to improve diagnostic rates by fusing multi-modal information, which can overcome the problem of sensitivity and specificity that could not be balanced by previous technology.^{20,21} Compared with the single mode (X-ray only), the AUC of the existing multi-modal X-ray diagnosis in COVID-19 and non-COVID-19 pneumonia increased from 0.89 to 0.93.²² Therefore, whether multi-modal information including X-ray data can improve the differential diagnosis of gram-positive or gram-negative bacteria requires further investigation.

Therefore, this study aimed to develop a learning model to identify gram-positive and gram-negative bacterial pneumonia in children based on X-ray data with clinical data to effectively improve the accuracy of diagnosis of pneumonia in children as well as to shorten the time to diagnosis, which may be useful for guiding empirical antibiotic treatment.

Materials and Methods

Study Sample

This study included 842 patients with community-acquired pneumonia aged <14 years who were admitted to our institution between January 2016 and June 2021. Precise clinical diagnosis, radiological diagnosis, and clear etiological examination results were available for all patients. Sputum and pleural effusion cultures were used to test the etiological

results of bacterial pneumonia. A chemiluminescence immunoassay quantitative diagnostic kit (Chongqing Bypass Medical Instrument Co. Ltd.; Chongqing, China) was used. The inclusion and exclusion criteria for pneumonia cases were first screened on the data platform of our institution. We selected patients diagnosed with pneumonia by chest X-ray examination and who had undergone etiological examination (n=1055). Patients with mixed (n=116), viral (n=55), and fungal (n=42) infections were excluded, leaving 447 cases of gram-positive bacterial pneumonia and 395 cases of gram-negative bacterial pneumonia (Figure 1). Patient consent was waived by the ethical review committee of Southwest Hospital, Army Medical University, as the study was a retrospective analysis of de-identified data. This study was conducted in accordance with the Declaration of Helsinki, and all patient data were kept confidential. The waiver of consent was reviewed and approved by the institutional review board of Southwest Hospital, Army Medical University (approval number: KY2020277).

Processing of Data

We conducted a series of preprocessing of the clinical and imaging data used in the experiment to meet the requirements of model training and evaluation. For image data, the size of each image was adjusted to a unified 256×256-pixel resolution, and its intensity value was linearly normalized into the interval [0,1]. In the case system, we collated 8 indicators of clinical signs, including cough and sputum, 5 patient baseline characteristics, 4 previous history, and 27 indicators of laboratory tests. A total of 44 clinical indicators were identified. The clinical data were categorized into categorical and numerical data. The categorical data were converted into numerical data using dummy variable coding.

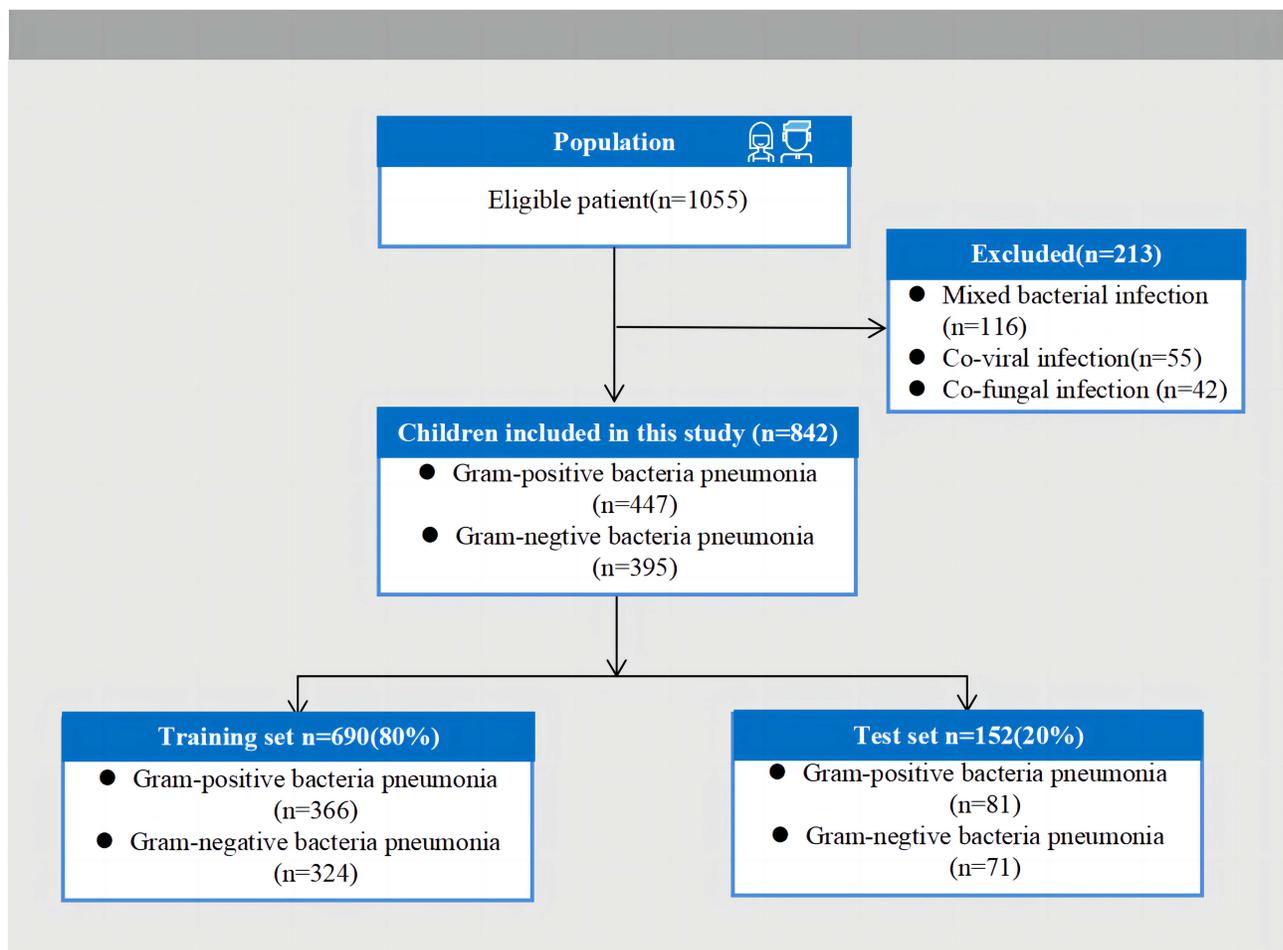


Figure 1 Patient selection flowchart.

These steps produced 842 matched cases for the final experimental dataset. To test the algorithm, we randomly selected 690 (80%) cases as the training set and 152 (20%) as the verification set.

Model Selection and Parameter Setting

To evaluate whether the fusion of imaging and non-imaging features can improve the classification results of childhood pneumonia, we established models based on imaging and clinical features, respectively. The CNN model consists of three convolution layers and three fully connected layers, each of which uses ReLU as the activation function, and the final output uses sigmoid as the activation function. The model for clinical features had four fully connected layers. Each of the first three layers is followed by BatchNorm1D and ReLU activation function construction, and the output of the last layer uses sigmoid as the activation function. In addition, we compared the most commonly used models for image and clinical feature classifications. We selected logistic regression (LR), support vector machines (SVM), and CatBoost as machine learning models for clinical data and ResNet, ResNeXt and MobileNetV3 as deep learning models for image data for the following reasons: LR and SVM are popular traditional machine learning algorithms that have been extensively used in medical research because of their interpretability and ability to handle high-dimensional data. CatBoost, on the other hand, is a gradient-boosting algorithm that has shown promising results in various applications, including medical clinical data analysis. ResNet, ResNeXt, and MobileNetV3 are deep learning models that have demonstrated exceptional performance in various image classification tasks. These models are known for their ability to handle high-dimensional data and learn complex representations of images. The feature fusion method proposed in this study concatenates image features extracted by a convolutional network using fully connected layers and clinical features. After the fusion of the image and clinical features, the feature space was learned and projected into the classification space to complete the classification task. The evaluation indicators of our classification algorithm were Precision, Recall, AUC, and F1.

Experimental Environment

The experiment was based on python3.7 language, in which the hardware environment used was RAM32G CPU: Intel(R) Xeon(R) Gold 6271C CPU @ 2.60GHz, NVIDIA V100 32G high-performance graphics card. Deep learning was prepared based on the libraries paddle 2.3.2, and machine learning was prepared based on the libraries CatBoost 1.0.6 and scikit-learn 0.22.

Statistics Analysis

Before building the model, we evaluated the differences in the clinical factors between gram-positive and gram-negative bacterial pneumonia. Student's *t*-test or Kruskal–Wallis *H*-test was used for continuous variables, and the χ^2 test or Fisher's exact test was used for categorical variables. Statistical analyses were performed using Statistical Package for the Social Sciences version 26.0 (IBM SPSS Statistics; Armonk, New York, USA). The level of statistical significance was set at $P < 0.05$.

Results

A total of 842 children were included in this study, including 447 cases of gram-positive and 395 cases of gram-negative bacterial pneumonia. Table 1 lists the characteristics of the patients in the training and test sets. There were significant differences in age, height, and body mass index between the gram-positive and gram-negative bacteria groups ($P < 0.05$). Among the clinical symptoms, cough, phlegm in the throat, foaming at the mouth, fever, groaning, nasal congestion, cyanosis, and poor mental state showed significant differences ($P < 0.05$). In the laboratory data, the levels of C-reactive protein, direct bilirubin, γ -glutamyltransferase, and alanine aminotransferase in the gram-positive group were significantly higher than those in the gram-negative group ($P < 0.05$).

The training and verification sets were divided based on a ratio of 8:2, wherein the final training set contained 690 cases. Among them, 366 were gram-positive and 324 were gram-negative. In total, 152 cases were verified. There were 81 gram-positive and 71 gram-negative patients. We achieved the most basic image binary classification in convolution and full connection. CNN-based refers to a simple CNN network structure. First, the binary task of gram-positive and

Table I Clinical Characteristics of Children (n=842)

| Characteristics | Gram-Negative Bacteria (n=395) | Gram-Positive Bacteria (n=447) | χ^2/t | P |
|---|--------------------------------|--------------------------------|------------|-------------------|
| Age (years, mean \pm SD) | 7.18 \pm 2.53 | 6.93 \pm 3.43 | 7.945 | P<0.001 |
| Sex, n (%) | | | | |
| Female | 138 (34.9) | 192 (43) | 5.654 | P<0.05 |
| Male | 257 (65.1) | 255 (57) | | |
| Height, weight, BMI (mean \pm SD) | | | | |
| Height | 69.26 \pm 14.43 | 70.31 \pm 15.59 | -8.395 | P<0.001 |
| Weight | 6.8 \pm 5.01 | 6.97 \pm 5.64 | 0.166 | P>0.05 |
| BMI | 5.25 \pm 8.11 | 7.25 \pm 11.57 | 6.431 | P<0.001 |
| Symptoms, n(%) | | | | |
| Cough | 81 (20.5) | 252 (56.4) | 112.851 | P<0.01 |
| Throat sputum | 50 (12.7) | 150 (33.6) | 50.57 | P<0.01 |
| Foaming at the mouth | 272 (68.9) | 374 (83.7) | 25.747 | P<0.01 |
| Fever | 43 (10.9) | 127 (28.4) | 39.971 | P<0.01 |
| Moaning | 74 (18.7) | 38 (8.5) | 19.041 | P<0.01 |
| Nasal congestion | 37 (9.4) | 68 (15.2) | 6.564 | P<0.01 |
| Cyanosis | 60 (15.2) | 38 (8.5) | 9.122 | P<0.01 |
| Poor mental health | 25 (6.3) | 74 (16.6) | 21.134 | P<0.01 |
| Past history of patients (%) | | | | |
| Disease history | 4 (1) | 6 (1.3) | 0.194 | P>0.05 |
| Surgical history | 1 (0.3) | 4 (0.9) | 1.463 | P>0.05 |
| History of blood transfusion | 6 (1.5) | 8 (1.8) | 0.094 | P>0.05 |
| History of infectious diseases | 9 (2.3) | 12 (2.7) | 0.142 | P>0.05 |
| Laboratory examination (mean \pm SD) | | | | |
| C-reactive protein (CRP):(mg/L) | 6.86 \pm 22.89 | 7.11 \pm 19.24 | -2.227 | P<0.05 |
| Alpha hydroxybutyrate dehydrogenase (α -HBDH):(U/L) | 291.51 \pm 203.57 | 274.58 \pm 139.73 | 3.669 | P<0.001 |
| α -hydroxybutyrate dehydrogenase (α 1-MG):(mg/L) | 15.31 \pm 3.41 | 15.78 \pm 3.98 | 0.36 | P>0.05 |
| β 2-microglobulin (β 2-MG):(μg/L) | 3 \pm 1.31 | 3.16 \pm 1.42 | 0.981 | P>0.05 |
| γ -glutamyltransferase (GGT):(U/L) | 95.5 \pm 103.83 | 111.79 \pm 124.14 | 7.619 | P<0.001 |
| Alanine aminotransferase (ALT):(U/L) | 24.04 \pm 31.45 | 31.59 \pm 59.1 | -5.228 | P<0.001 |
| Lactate dehydrogenase (LDH):(U/L) | 516.47 \pm 352.05 | 601.5 \pm 409.92 | 1.04 | P>0.05 |
| Total protein (TP):(g/L) | 58.68 \pm 8.55 | 57.89 \pm 9.41 | -8.942 | P<0.001 |
| Total calcium (Ca):(mmol/L) | 2.38 \pm 0.26 | 2.35 \pm 0.28 | -3.613 | P<0.001 |
| Albumin/globulin ratio (ALB/GLO) | 2.09 \pm 0.54 | 1.91 \pm 0.52 | 2.997 | P<0.01 |
| Albumen (ALB):(g/L) | 39 \pm 5.69 | 37.32 \pm 6.12 | -6.502 | P<0.001 |
| Direct bilirubin (DBIL):(μmol/L) | 6.15 \pm 14.02 | 8.76 \pm 31.47 | 3.87 | P<0.001 |
| Alkaline phosphatase (ALP):(U/L) | 225.62 \pm 103.83 | 220.36 \pm 121.77 | -1.382 | P>0.05 |
| Creatine kinase (CK):(U/L) | 314.5 \pm 511.4 | 284.37 \pm 441.36 | 3.253 | P<0.01 |
| Creatine kinase isoenzyme MB (CK-MB):(U/L) | 61.44 \pm 94.64 | 51.4 \pm 51.19 | 5.48 | P<0.001 |
| Hypersensitive C reactive protein (hs-CRP):(mg/L) | 1.5 \pm 0.8 | 1.55 \pm 0.65 | -2.847 | P<0.01 |
| Adenosine deaminase (ADA):(U/L) | 9.39 \pm 7.37 | 9.06 \pm 6.59 | -8.659 | P<0.001 |
| Procalcitonin (PCT):(ng/mL) | 1.23 \pm 4.74 | 0.71 \pm 2.32 | 1.917 | P>0.05 |
| Neutrophil count (Neut):(10 ⁹ /L) | 5.63 \pm 4.27 | 4.84 \pm 3.01 | 3.455 | P<0.01 |
| Monocyte count (Mono):(10 ⁹ /L) | 1.07 \pm 0.58 | 1.02 \pm 0.46 | 0.723 | P>0.05 |
| Basophil count (Baso):(10 ⁹ /L) | 0.03 \pm 0.07 | 0.03 \pm 0.06 | 1.862 | P>0.05 |
| Eosinophil count (Eos):(10 ⁹ /L) | 0.25 \pm 0.23 | 0.23 \pm 0.29 | -0.008 | P>0.05 |
| Mean red blood cell volume (MCV):(fL) | 94.66 \pm 12.04 | 95.86 \pm 12.35 | 10.874 | P<0.001 |
| Mean erythrocyte hemoglobin content (MCH):(pg) | 32.21 \pm 4.6 | 32.68 \pm 4.54 | 10.617 | P<0.001 |
| Mean erythrocyte hemoglobin concentration (MCHC):(g/L) | 339.33 \pm 12.9 | 340.11 \pm 11.01 | 3.208 | P<0.01 |
| Platelet count (PLT):(10 ⁹ /L) | 303.69 \pm 109.33 | 289.09 \pm 103.69 | -3.9 | P<0.001 |
| Hemoglobin (Hb):(g/L) | 142.24 \pm 26.44 | 140.4 \pm 23.53 | 7.148 | P<0.001 |

Notes: The bold value indicated a statistical significance with $p < 0.05$.

gram-negative bacterial pneumonia based on image data was established. Then, the binary task of clinical data based on LR, SVM, CatBoost, and full-connection network was established. Finally, the clinical and image features were fused and a multi-modal binary classification task based on deep learning was established. The model fusion strategy is shown in Figure 2. We used four evaluation indicators (accuracy (ACC), AUC, Recall, and F1) to evaluate the diagnosis of bacterial pneumonia in the training and verification sets of the model. Table 2 lists the comprehensive prediction performance of each classifier in the test set. Among the four evaluation indexes, CatBoost had the best performance, and its AUC was significantly higher than that of the LR, SVM, and fully connected network models. In the classification task based on clinical data, we used LR, SVM, CatBoost, and four fully connected networks. SVM had the best accuracy

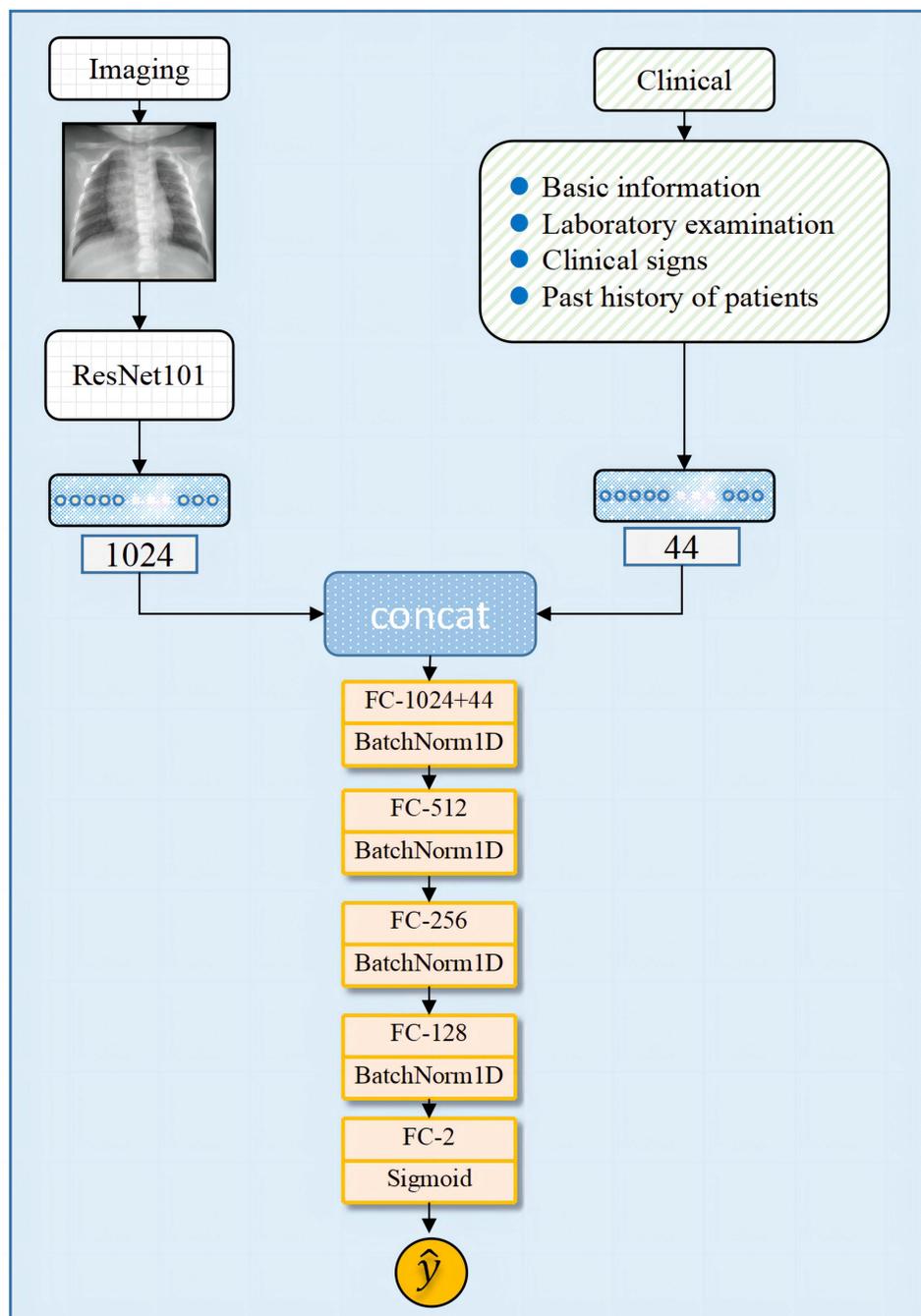


Figure 2 Fusion strategies using deep learning.

Table 2 Performance Comparison of Machine Learning Algorithm Models Based on Clinical Data

| Model | Clinical Data | | | |
|-------------------------|---------------|--------|-------|-------|
| | Accuracy | Recall | AUC | F1 |
| LR | 0.661 | 0.804 | 0.760 | 0.722 |
| SVM | 0.756 | 0.804 | 0.813 | 0.723 |
| CatBoost | 0.738 | 0.837 | 0.852 | 0.778 |
| Fully connected network | 0.692 | 0.733 | 0.728 | 0.717 |

Table 3 Deep Learning Model Based on Imaging Data Alone and Its Performance Based on the Fusion of Imaging and Clinical Data

| Model | Image Data | | | | Imaging Data + Clinical Data | | | |
|-------------|------------|--------|-------|-------|------------------------------|--------|-------|-------|
| | Accuracy | Recall | AUC | F1 | Accuracy | Recall | AUC | F1 |
| Resnet18 | 0.680 | 0.772 | 0.748 | 0.683 | 0.632 | 0.914 | 0.796 | 0.726 |
| Resnet34 | 0.714 | 0.587 | 0.744 | 0.647 | 0.638 | 0.864 | 0.790 | 0.720 |
| Resnet50 | 0.699 | 0.630 | 0.759 | 0.652 | 0.684 | 0.877 | 0.771 | 0.747 |
| Resnet101 | 0.694 | 0.533 | 0.742 | 0.609 | 0.750 | 0.839 | 0.803 | 0.782 |
| ResNeXt50 | 0.714 | 0.543 | 0.742 | 0.629 | 0.671 | 0.827 | 0.795 | 0.728 |
| ResNeXt101 | 0.689 | 0.478 | 0.719 | 0.579 | 0.665 | 0.840 | 0.803 | 0.727 |
| MobileNetV3 | 0.714 | 0.533 | 0.709 | 0.624 | 0.632 | 0.827 | 0.795 | 0.705 |

Table 4 Optimization of the Resnet101 Model

| Resnet101 | Imaging Data | | | |
|---------------|--------------|--------|--------|--------|
| | Accuracy | Recall | AUC | F1 |
| Base | 0.6941 | 0.5326 | 0.7421 | 0.6086 |
| +aug | 0.7087 | 0.6413 | 0.7672 | 0.6629 |
| +pretrain | 0.7184 | 0.424 | 0.7139 | 0.5735 |
| +resize (512) | 0.7184 | 0.4457 | 0.7685 | 0.5857 |

Abbreviations: Base, base models; +aug, add data augment; + pretrain, add pre-training model; + resize (512), resize input image as 512×512.

(0.756), CatBoost had the best recall rate and AUC (0.738 and 0.852, respectively), and SVM had the best AUC. After a comprehensive evaluation based on the F1 value, the best was CatBoost (0.778).

In Table 3, we comprehensively compare the enhancement effects of the clinical feature fusion image models. The input image size resize of its unified model was 256×256, the learning rate was 1e-3, the epoch was 30, the batch size was 8, the stochastic gradient descent (SGD) optimizer was used, and the mean squared error (MSE) loss function was used. Because the CNN-based network does not converge after adding clinical features, it is not considered in statistics. Comparing the results of the seven models, ResNet18, ResNet34, ResNet50, ResNet101, ResNeXt50, ResNeXt101, and MobileNetV3, the comprehensive index AUC and F1 increased by an average of 5.6% and 10.2%, respectively. This shows that introducing clinical data can significantly improve the model performance. Finally, the best performance of the Resnet101 model accuracy was 0.75, the recall rate was 0.84, the AUC was 0.803, and F1 was 0.782. In Table 4, we employed several common methods to optimize the model performance, including data augmentation techniques such as

horizontal flipping, vertical flipping, 90° rotation, and image resizing. Furthermore, we utilized a pre-trained model based on ImageNet to initialize our model, which helped to accelerate the model's convergence. The results show that not all optimizations are effective; data enhancement is a relatively general optimization scheme.

Discussion

Using gram-positive and gram-negative bacterial pneumonia data, we proved that the multi-modal fusion model based on X-ray images and non-image features is significantly better than image-only and clinical data-only models in predicting the types of bacterial pneumonia. When classifying based on image features, adding clinical information can further improve the performance of the model analysis. On average, the F1 and AUC improved by 5.6% and 10.2%, respectively. Therefore, our results show that the fusion of image data and clinical data can significantly improve the prediction performance of deep learning methods for automatic diagnosis. Conversely, owing to the limited amount of data in this study, the deep learning model offered a manageable advantage. However, the best Resnet101 model still performed as well as the best machine-learning model.^{23–27} The optimization scheme of deep learning during image training cannot be directly used in the multi-modal decision fusion of clinical and image data. Data enhancement is a more suitable model optimization scheme. In this study, we focused on a binary classification problem, in which each image was classified as gram-positive or gram-negative. It should be noted that our current results are not particularly ideal, compared with traditional imaging, where it is difficult for radiologists to distinguish the types of bacterial pneumonia using only CXR, which may be related to the pathological manifestations of bacterial pneumonia.²⁸ This deep learning model can provide a reference for continuous optimization in the future.²⁹

Although computerized tomography has a better advantage in diagnosing pneumonia, CXR is the best choice for diagnosing pneumonia in children, given the child's level of cooperation and radiation exposure.³⁰ The diagnosis of pneumonia is multi-modal. Doctors need to combine imaging findings with data on various clinical risk factors to make a diagnosis and use antibiotics empirically.³¹ Deep learning methods for the automatic classification of pneumonia usually use only image or non-image clinical data, and do not use both. Previous studies have classified childhood bacteria, viral pneumonia, and malaria based solely on clinical data, achieving an overall sensitivity and specificity for bacterial diagnosis of 96% and 86%, respectively.³² Some studies have also realized the classification of gram-positive, gram-negative, and atypical bacterial pneumonia based only on image data, with a sensitivity of 0.57 and specificity of 0.78.¹⁹ In this study, by integrating a deep learning algorithm based on imaging and clinical data, the classification accuracy of gram-negative pneumonia was 0.75, and the recall rate was 0.84. We also investigated the effect of reducing the partitioning of training sets on the classification performance of multi-modal models including clinical and imaging data. Specifically, we reduced the training set from an 8:2 partition to a 6:4 partition, resulting in a decrease in the number of training set samples. In the experiments, the data augmentation techniques were used to complement the training data. The imaging data consisted of images with a size of 256×256, and the base size was set to 1e-4. We performed ablation experiments to assess the effect of reducing the training set size. Additionally, we assessed the accuracy of a CNN model containing images. The clinical data was used to build the model and compared the performance of the multi-modal model with that of a model constructed solely from clinical data using the catBoost algorithm. We found that reducing the partition of the training set from 8:2 to 6:4 led to a decrease in the AUC of the model's classification performance. The catboost algorithm was more stable with a smaller amount of data ([Supplementary Table 1](#)). The accuracy of the CNN model containing images also decreased significantly ([Supplementary Table 2](#)). We observed that the model constructed from clinical data was more stable when the data changed compared with the deep learning model constructed from multi-modal data. This suggests that the classification performance of the multi-modal model mainly relies on image features. Consequently, subsequent research should focus on improving the importance of clinical information features in the model. Reducing the partition of the training set can have a significant impact on the classification performance of multi-modal models. The study also suggests that clinical information features should be given more importance in the development of multi-modal models. Future research should aim to improve the stability and accuracy of such models by incorporating clinical data in a more effective manner.

The current study has several limitations. First, only a single-center database was used. Second, the study population was from China, and patients from other countries were not considered. Third, we excluded patients with low immunity

or immunodeficiency to avoid including patients with mixed infections. All patients were isolated and samples cultured to determine the pathogen species; however, the inclusion of potentially co-infected patients could not be entirely excluded. Randomized clinical trials are still needed to test the multi-modal classification algorithm model on a larger scale. Although our research has demonstrated the advantages of multi-modal information in convolutional neural networks, the limited availability of data remains a challenge for achieving optimal results with multi-modal models. In order to further enhance the performance of the model, we hope the following areas for future research: (1) training predictive models on a larger and more diverse dataset that includes both images and clinical data and (2) focusing on optimizing the model structure for clinical data and developing more effective methods for integrating information from multiple modalities.

Conclusion

The multi-modal decision of integrating deep learning-based image data into clinical data has good performance in the classification of pediatric bacterial pneumonia, which may preliminarily guide clinical empirical antibiotic therapy. This algorithm is expected to be useful as a clinical diagnostic assistant in the real world.

Abbreviations

CXR, Chest radiography; AUC, Area under the receiver operating characteristic curve; CNN, Convolutional neural network; MSE, Mean squared error; SGD, Stochastic gradient descent.

Funding

This study has received funding from major science and technology projects of Chongqing City (Grant No. cstc2018jszxcyztzxX0017), Young and Middle-aged Medical Talents Foundation Project of Chongqing (Grant No. 414Z395), National Natural Science Foundation of China (NSFC) (grant No. 82071910), Emergency Project for Technological Breakthrough in Clinical Treatment of Hospital-acquired COVID-19 Infection in 2023(2023XGII07), the Youth Training Project of Medical Science and Technology (Grant No. 20QNPY012).

Disclosure

The authors of this article declare no relationships with any companies, whose products or services may be related to the subject matter of the article. The authors declare that they have no competing interests.

References

1. Liu L, Johnson HL, Cousens S, et al. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *Lancet*. 2012;379:2151–2161.
2. Jain S, Williams DJ, Arnold SR, et al. Community-acquired pneumonia requiring hospitalization among U.S. children. *N Engl J Med*. 2015;372:835–845. doi:10.1056/NEJMoa1405870
3. Wang H, Abajobir AA, Abate KH, et al. Global, regional, and national under-5 mortality, adult mortality, age-specific mortality, and life expectancy, 1970–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet*. 2017;2017:1.
4. Jones RN. Microbial etiologies of hospital-acquired bacterial pneumonia and ventilator-associated bacterial pneumonia. *Clin Infect Dis*. 2010;51 Suppl 1:S81–S87. doi:10.1086/653053
5. Levy SB, Marshall B. Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med*. 2004;10:S122–S129. doi:10.1038/nm1145
6. Donnelly JP, Baddley JW, Wang HE. Antibiotic utilization for acute respiratory tract infections in U.S. emergency departments. *Antimicrob Agents Chemother*. 2014;58:1451–1457. doi:10.1128/AAC.02039-13
7. Pervaiz F, Chavez MA, Ellington LE, et al. Building a prediction model for radiographically confirmed pneumonia in Peruvian children: from symptoms to imaging. *Chest*. 2018;154:1385–1394.
8. Yan C, Hui R, Lijuan Z, Zhou Y. Lung ultrasound vs. chest X-ray in children with suspected pneumonia confirmed by chest computed tomography: a retrospective cohort study. *Exp Ther Med*. 2020;19:1363–1369. doi:10.3892/etm.2019.8333
9. Edwards M, Lawson Z, Morris S, et al. The presence of radiological features on chest radiographs: how well do clinicians agree? *Clin Radiol*. 2012;67:664–668. doi:10.1016/j.crad.2011.12.003
10. Murdoch DR, O'Brien KL, Scott JA, et al. Breathing new life into pneumonia diagnostics. *J Clin Microbiol*. 2009;47:3405–3408. doi:10.1128/JCM.01685-09
11. Kratz AMP, Sullivan K, Gallagher J. Clinical impact of matrix-assisted laser desorption ionization time-of-flight mass spectrometry for the management of inpatient pneumonia without additional antimicrobial stewardship support. *Infect Control Hosp Epidemiol*. 2019;40(9):1053–1055. doi:10.1017/ice.2019.191

12. Soomro TA, Zheng L, Afifi AJ, Ali A, Yin M, Gao J. Artificial intelligence (AI) for medical imaging to combat coronavirus disease (COVID-19): a detailed review with direction for future research. *Artif Intell Rev.* 2022;55:1409–1439. doi:10.1007/s10462-021-09985-z
13. Shamman AH, Hadi AA, Ramul AR, Zahra MMA, Ghani HM. The artificial intelligence (AI) role for tackling against COVID-19 pandemic. *Mater Today Proc.* 2021;80:3663–3667. doi:10.1016/j.matpr.2021.07.357
14. L E, Zhao B, Guo Y, et al. Using deep-learning techniques for pulmonary-thoracic segmentations and improvement of pneumonia diagnosis in pediatric chest radiographs. *Pediatr Pulmonol.* 2019;54:1617–1626. doi:10.1002/ppul.24431
15. Rajpurkar P, Irvin J, Ball RL, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med.* 2018;15:e1002686. doi:10.1371/journal.pmed.1002686
16. Hwang EJ, Nam JG, Lim WH, et al. Deep learning for chest radiograph diagnosis in the emergency department. *Radiology.* 2019;293:573–580. doi:10.1148/radiol.2019191225
17. Wang J, Bao Y, Wen Y, et al. Prior-attention residual learning for more discriminative COVID-19 screening in CT images. *IEEE Trans Med Imaging.* 2020;39:2572–2583. doi:10.1109/TMI.2020.2994908
18. Xu GX, Liu C, Liu J, et al. Cross-site severity assessment of COVID-19 from CT images via domain adaptation. *IEEE Trans Med Imaging.* 2022;41:88–102. doi:10.1109/TMI.2021.3104474
19. Zhang M, Yu S, Yin X, et al. An AI-based auxiliary empirical antibiotic therapy model for children with bacterial pneumonia using low-dose chest CT images. *Jpn J Radiol.* 2021;39:973–983. doi:10.1007/s11604-021-01136-2
20. Xu M, Ouyang L, Han L, et al. Accurately differentiating between patients with COVID-19, patients with other viral infections, and healthy individuals: multimodal late fusion learning approach. *J Med Internet Res.* 2021;23:e25535. doi:10.2196/25535
21. Sheu RK, Chen LC, Wu CL, et al. Multi-modal data analysis for pneumonia status prediction using deep learning (MDA-PSP). *Diagnostics.* 2022;13:12. doi:10.3390/diagnostics13010012
22. Tan T, Das B, Soni R, et al. Multi-modal trained artificial intelligence solution to triage chest X-ray for COVID-19 using pristine ground-truth, versus radiologists. *Neurocomputing.* 2022;485:36–46. doi:10.1016/j.neucom.2022.02.040
23. Woan Ching SL, Lai KW, Chuah JH, et al. Multiclass convolution neural network for classification of COVID-19 CT images. *Comput Intell Neurosci.* 2022;2022:9167707. doi:10.1155/2022/9167707
24. Showkat S, Qureshi S. Efficacy of transfer learning-based ResNet models in chest X-ray image classification for detecting COVID-19 pneumonia. *Chemometr Intell Lab Syst.* 2022;224:104534. doi:10.1016/j.chemolab.2022.104534
25. Rasheed J, Shubair RM, Lunghi C, Dupuis D. Screening lung diseases using cascaded feature generation and selection strategies. *Healthcare.* 2022;11:10. doi:10.3390/healthcare11010010
26. Ksibi A, Zakariah M, Ayadi M, et al. Improved analysis of COVID-19 influenced pneumonia from the chest X-rays using fine-tuned residual networks. *Comput Intell Neurosci.* 2022;2022:9414567. doi:10.1155/2022/9414567
27. Dhiman G, Chang V, Kant Singh K, Shankar A. ADOPT: automatic deep learning and optimization-based approach for detection of novel coronavirus COVID-19 disease using X-ray images. *J Biomol Struct Dyn.* 2022;40:5836–5847. doi:10.1080/07391102.2021.1875049
28. Rangarajan AK, Ramachandran HK. A preliminary analysis of AI based smartphone application for diagnosis of COVID-19 using chest X-ray images. *Expert Syst Appl.* 2021;183:115401. doi:10.1016/j.eswa.2021.115401
29. Jia R, Yang J, Cui Y, Guo D, Li T. Gene expression analysis for pneumonia caused by Gram-positive bacterial infection. *Exp Ther Med.* 2018;15:3989–3996. doi:10.3892/etm.2018.5904
30. Agweyu A, Lilford RJ, English M; Clinical Information Network Author G. Appropriateness of clinical severity classification of new WHO childhood pneumonia guidance: a multi-hospital, retrospective, cohort study. *Lancet Glob Health.* 2018;6:e74–e83. doi:10.1016/S2214-109X(17)30448-5
31. van Vugt SF, Verheij TJ, de Jong PA, et al. Diagnosing pneumonia in patients with acute cough: clinical judgment compared to chest radiography. *Eur Respir J.* 2013;42:1076–1082. doi:10.1183/09031936.00111012
32. Valim C, Ahmad R, Lanaspas M, et al. Responses to bacteria, virus, and malaria distinguish the etiology of pediatric clinical pneumonia. *Am J Respir Crit Care Med.* 2016;193:448–459. doi:10.1164/rccm.201506-1100OC

Infection and Drug Resistance

Dovepress

Publish your work in this journal

Infection and Drug Resistance is an international, peer-reviewed open-access journal that focuses on the optimal treatment of infection (bacterial, fungal and viral) and the development and institution of preventive strategies to minimize the development and spread of resistance. The journal is specifically concerned with the epidemiology of antibiotic resistance and the mechanisms of resistance development and diffusion in both hospitals and the community. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/infection-and-drug-resistance-journal>