Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion

Lingzhi Kong, Jinyong Cheng

School of Computer Science and Technology, Shandong Academy of Sciences, Jinan 250000, China

ARTICLE INFO

Keywords:
COVID-19
Feature Fusion
DenseNet
VGG16
Image classification

ABSTRACT

Since December 2019, the novel coronavirus disease (COVID-19) caused by the syndrome coronavirus 2 (SARS-CoV-2) strain has spread widely around the world and has become a serious global public health problem. For this high-speed infectious disease, the application of X-ray to chest diagnosis plays a key role. In this study, we propose a chest X-ray image classification method based on feature fusion of a dense convolutional network (DenseNet) and a visual geometry group network (VGG16). This paper adds an attention mechanism (global attention machine block and category attention block) to the model to extract deep features. A residual network (ResNet) is used to segment effective image information to quickly achieve accurate classification. The average accuracy of our model in detecting binary classification can reach 98.0%. The average accuracy for three category classification can reach 97.3%. The experimental results show that the proposed model has good results in this work. Therefore, the use of deep learning and feature fusion technology in the classification of chest X-ray images can become an auxiliary tool for clinicians and radiologists.

1. Introduction

Pneumonia-type illnesses are more contagious during the flu season [1,2]. Chest X-rays (CXRs) play an important role in patient care. Radiologists can use CXR features to determine the type of pneumonia and the underlying pathogenesis [3]. X-ray is one of the most common radiological examination methods for screening and diagnosing chest diseases, as well as the main means of classifying and screening pneumonia, tuberculosis and breast cancer, and is a painless and noninvasive examination method suitable for high populations with relatively low costs [4].

The pandemic of global concern caused by COVID-19 has also brought enormous challenges to governments and the healthcare industry [5-7]. The outbreak was declared a Public Health Emergency of International Concern on 30 January 2020. It was named COVID-19 by the World Health Organization (WHO) in February 2020; around March 2020, the World Health Organization announced that the disease has affected the whole world and is a global pandemic disease [8,9]. The characteristics of COVID-19 are diverse and unpredictable. The common clinical symptoms are mainly respiratory symptoms, and some patients may have gastrointestinal symptoms [10,11]. Real-time polymerase chain reaction (RT–PCR), loop-mediated isothermal amplification (LAMP), antigen testing and other methods can be used to detect COVID-19. Although the specificity of RT–PCR is sufficiently high for COVID-19, its sensitivity is relatively low in detecting COVID-19 [6,12]. LAMP technology has high sensitivity, fast reaction rate and strong specificity, but the design of primers is very complicated, and it is easy to produce nonspecific amplification [13]. Although the antigen test has a relatively fast detection speed, its sensitivity is poor.

Therefore, CXR, as a sensitive method to detect COVID-19 as well as other chest problems, plays an integral role in the early diagnosis and treatment of the disease [12,14]. Previous studies have shown that CXR images have specific differences in the imaging manifestations of common pneumonia and COVID-19. These differences or subtle features can be detected by artificial intelligence. For example, as shown in Fig. 1 [21], it can assist doctors in achieving more accurate classification and diagnosis.

In this study, we reviewed the relevant literature and work on CXR classification. At the same time, artificial intelligence methods were used to efficiently and quickly identify different cases of common pneumonia and COVID-19 as well as to distinguish healthy patients. Therefore, helping doctors to distinguish more accurately among ordinary pneumonia, COVID-19, and healthy patients could lead to more targeted treatment for patients and reduce the duration of illness.

* Corresponding author.
E-mail address: cjy@qlu.edu.cn (J. Cheng).

https://doi.org/10.1016/j.bspc.2022.103772
Received 2 January 2022; Received in revised form 22 April 2022; Accepted 27 April 2022
Available online 8 May 2022
1746-8094/© 2022 Elsevier Ltd. All rights reserved.
The following is the main contribution and summary of the work:

- Prior to this work, an image preprocessing method was applied to CXR data, and segmented images were selected to be sent into the model for more accurate and effective data analysis.
- In this work, we introduced the fine-tuned global attention block (GAB) and category attention block (CAB) for the imbalance of data and distribution to obtain more detailed information of small lesions.
- We fused the DenseNet and VGG models and fine-tuned the model to better detect different pneumonia diseases with convenience, speed and precision.
- We conducted two-category work for healthy and pneumonia patients and three-category work for healthy, common pneumonia, and COVID-19 patients. Compared with other advanced methods, the results show that our model can classify chest X-ray data with high accuracy.

Therefore, in this work, we propose a method based on feature fusion that can more accurately distinguish among healthy, common pneumonia, and COVID-19 patients using chest X-ray images. The main structure of the paper is as follows: Section II discusses the advanced work on using X-ray images for COVID-19. Section III shows data preprocessing of the datasets applied in the work. Section IV provides an in-depth exploration of the fusion models and experiments. Section V obtains and analyzes the experimental results. Artificial intelligence methods were used for the efficient and rapid identification of different cases of general pneumonia and COVID-19 as well as for healthy patients. Section VI analyzes the limitations of this work. Section VII discusses the research topic and prepares for future work.

2. Related work

In recent years, the development of artificial intelligence has been effective in a range of medical fields. The use in healthcare is also on the rise, particularly in medical imaging [15]. Radiological imaging technology, such as chest X-ray scans, can protect patients more effectively, isolate infected patients in time, and distinguish pneumonia types more accurately. A. Narin et al. [16] applied five convolutional neural network models (ResNet50, ResNet101, ResNet152, InceptionV3 and Inception-ResNetV2) to detect patients infected with coronavirus pneumonia using chest X-rays. M. Turkoglu [17] used transfer to learn the features of the convolutional and fully connected layers of the AlexNet model. The SVM classifier was used to detect and classify the important features identified by Relief. The VGG16 pretraining model combined with data enhancement and patching (RICAP) was used to improve robustness and assess the healthy population and COVID-19 [18]. Khan et al. [19] proposed the application of the CoroNet deep convolutional neural network model for automatic chest X-ray detection. Moreover, COVID-19, pneumonia and healthy patients were also classified into three categories, and the classification accuracy of the proposed model was 95%, which greatly promoted the detection of chest diseases. Ouchicha et al. [20] proposed the deep convolutional neural network (CNN) model of CVDNet, which uses chest X-ray images to classify COVID-19 infection from normal and other cases of pneumonia. The architecture is based on a residual neural network, employing two parallel layers of different convolution kernel sizes to capture the local and global features of the input. Through the study and research of the above methods, it was found that although the accuracy of the classification was more accurate in the process of binary classification (healthy population vs. common pneumonia patients, normal vs. COVID-19 patients, and common pneumonia patients vs. COVID-19 patients), in the process of three category classification (healthy, common pneumonia, and COVID-19 patients), the accuracy is generally low, and the three categories cannot be more accurately classified.

3. Dataset

3.1. Dataset

This study used two different publicly available datasets to collect chest X-ray images to create one dataset. The dataset contains a total of 6518 images, and the test set data account for 20% of the total data. The first is from a publicly available dataset. Chest X-ray images (before and after) were selected from pediatric patients aged 1 to 5 years at the Guangzhou Women’s and Children’s Medical Center [21]. All chest X-ray imaging was performed as part of the patient’s routine clinical care. To analyze chest X-ray images, all chest X-ray images were screened for quality control by removing all low-quality or unreadable scans. The second dataset is the COVID-19 X-ray image database developed by Joseph et al. [22], which utilizes images from various open access sources. The authors collected images relevant to radiology from various authentic sources (Radiology Society of North America (RSNA), Encyclopedia of Radiology, etc.). Most research on COVID-19 has used images from this source. The repository contains an open database of COVID-19 cases with chest X-ray images and is being updated regularly.

3.2. Data preprocessing

Low-dose X-ray images suffer from blurred edges, low contrast due to objective factors, and a low signal-to-noise ratio of projections. To better extract chest X-ray image features, image segmentation technology was used to remove the background noise and retain only the effective chest shadow area. In the course of our experimental study, data enhancement techniques were applied to enrich medical image datasets. In the images of the datasets used in this study, the size, shape, shadow area and location of the chest lesions precisely varied from patient to patient. (1) All images in the training set were traversed, and the “inference” function was called. ResNet34 [23] was used for semantic segmentation.
learning, and the segmentation threshold was set to 0.2.

(2) Images of different sizes were scaled to $512 \times 512$. The mask area marked with red and green was returned.

(3) The original image data value was converted into a 255-level gray value image by enhancing the information of the lung shadow area and surrounding tissues.

(4) Finally, the effective region was segmented by the dot product of the grayscale image and the mask data matrix. The redundant parts were removed, and chest X-ray images without background were obtained.

Therefore, we created a new chest X-ray image dataset: the chest X-ray images were horizontally flipped, and 3x3 Gaussian blur was added to reduce the overfitting of the model during training and effectively achieve the invariant stability of the model learning process. The segmented images not only have the noise redundant background regions removed but also help the model to accurately analyze images of healthy, common pneumonia, and COVID-19 patients. Fig. 2 shows the step-by-step segmentation of normal subject chest X-ray images (X-ray image, recognition of the chest region image, and chest image after removing the background).

We unified the segmented images in the form of .jpg and classified them by type. Eighty percent of the data were used for training, and the remaining 20% were used for testing. That is, the total amount of data in the training set was 5230, and the total amount of data in the test set was 1288. The specific situation of the chest X-ray dataset prepared by us is presented in Table 1.

### 4. Method

#### 4.1. Model connection and feature fusion

##### 4.1.1. DenseNet

For our first piece of model, we use a 201-layer dense connected

| Labels       | Train (80%) | Test (20%) | Total (100%) |
|--------------|-------------|------------|--------------|
| COVID-19     | 460         | 116        | 576          |
| HEALTHY      | 1266        | 317        | 1583         |
| PNEUMONIA    | 3418        | 855        | 4273         |
| Total        | 5230        | 1288       | 6518         |

![Fig. 3. Structure diagram of the DenseNet201 model.](image)

![Fig. 4. VGG16 model structure diagram.](image)
convolutional network (DenseNet). As shown in Fig. 3, the network layer of DenseNet201 is densely linked by four network blocks, and the dense blocks are uniformly connected by transition layer pooling. Finally, the feature map of the last layer is pooled by global average pooling to form a feature point. These feature points constitute the final feature vector, which is calculated in SoftMax [24]. The DenseNet model mainly realizes feature reuse through feature connection on the channel. There are two basic methods in feature reuse: bypass and concatenation. Since gradient disappearance usually occurs in the deep layer of the network, it is more appropriate to place bypass layers at the beginning of the deep layer of the network [25]. The dense connection mechanism is that each layer will be connected with all the previous layers on the channel dimension as input to the next layer. It is more efficient to achieve feature reuse. In the traditional network, the output of the l layer is represented as, $x_l = H_l(x_{l-1})$ while in DenseNet, all layer dense connection modes are represented as $x_l = H_l([x_0, x_1, ..., x_{l-1}])$. $H_l(\cdot)$ stands for nonlinear transformation. The structure of Batchnorm + ReLU + 3×3Conv can be used to obtain more input features and improve the efficiency of feature reuse. It not only greatly reduces the number of network parameters but also alleviates the vanishing gradient problem to a certain extent [26].

4.1.2. Vgg16

The second model used in this work is the VGG network structure. The VGG16 network can not only increase the network depth but also improve the performance more effectively. The simple module is composed of a small convolution kernel, small pooling kernel and ReLU. As shown in Fig. 4, there are 5 convolutional layers, 3 fully connected layers and a softmax output layer. Max pooling is used to separate the layers, and the ReLU function is used for the activation units of all hidden layers [27]. Therefore, one of the great advantages of VGG networks is that they simplify the structure of neural networks. The obtained $7 \times 7 \times 512$ feature map is fully connected, and then softmax activation is carried out to output the recognition results of the three objects.

4.1.3. Model fusion

Since we have adopted datasets of different scales and styles in our work, we use different network models to extract and identify their application performance. Therefore, to achieve the best quality model classification, we used the methods of model connection [28] and feature fusion [29] to build the model in this study, which are mainly applied to the two network structures. DenseNet and VGG16 have a clear division of labor. Using the characteristics of the two networks, corresponding adjustments are made in the input of the dataset. The original image is first fed into the DenseNet network model architecture, which slows down the vanishing gradient problem and enhances feature propagation [30]. Then, the dataset generated by ResNet34 is sent to the VGG16 network model. Increasing the network depth affects the final performance of the network to a certain extent. The mobility is enhanced while reducing the error rate, and the generalization to other image data is also good [31]. Finally, ensemble learning is used to fuse the features, and two attention mechanisms (global attention block and category attention block) are used to solve the problem of the weak generalization ability of a single neural network, as shown in Fig. 5.
4.2. Sample balance and attention mechanism addition

From Fig. 6, we can clearly find that there is a serious imbalance in the proportion of images obtained in different datasets, which can easily lead to problems such as a decline in the predictive ability of the model and a large error [32]. In response to the imbalance of data samples, two main parts are added to the work to equalize the data samples: (1) the “class_weight” function in the Keras library and (2) the attention mechanism.

4.2.1. Sample balance

Using the class_weight function in the Keras library changes the range of loss, which may affect the stability of training [33]. We can choose “balanced” to let the library automatically increase the weight of illegal user samples. “Balanced” improves the weight of certain categories so that more sample categories will be classified into high-weight categories than without considering the weight, thus balancing the number of samples in the dataset. The weight calculation formula under “balanced” can be expressed as:

\[ W_i = \frac{\text{Sum}(\text{Class} \times \text{Sum})}{\text{Class} \times \text{Sum}} \]  

(1)

Where, \( W_i \) indicates the final calculated weight value of each category, and \( i \) represents the number of \( k \) categories. \( \text{Sum} \) represents the size of all samples in this type of dataset. \( \text{Class} \) represents the number of categories in the total sample. \( \text{Sum} \) indicates the number of samples corresponding to class.

4.2.2. Attention mechanism

For the problems in Fig. 6, the uneven distribution of COVID-19, healthy, and pneumonia data leads to a high level of concern in the model. The sample imbalance greatly affects the performance of the final classification. As mentioned in the literature [34], CBAM is a lightweight general module that can be better applied to any CNN architecture and has a small sales volume, thus playing a significant role in the application of GAB and CAB [35]. GAB can be used to preserve detailed pathological information in pneumonia images and suppress color features and brightness features of similar parts. CAB can be used to learn distinguishing features to better solve the problem of low accuracy caused by an uneven distribution of data.

In Eq. (2), \( Z_{\text{avg}} \) is used to calculate the attention feature of the channel, where \( H \) represents the height, \( W \) represents the width, and \( C \) represents the number of channels. \( \delta \) means using the sigmoid activation function, \( \text{GAP} \) means average pooling, and \( Z_{\text{IN}} \) uses a \( 1 \times 1 \) convolution to reduce the number of channels.

\[ Z_{\delta, \text{OUT}} = \delta(\text{Conv2}(\text{GAP}(Z_{\text{IN}})))) \otimes Z_{\text{IN}} \in R^{H \times W \times C}, Z_{\text{IN}} \in R^{H \times W \times C'} \]

\[ = C/2 \]  

(2)

\( C_G \) stands for cross-channel average pooling and saves more detailed information about small lesions in Eq. (3). Different lesions in chest X-ray images can be better divided into details as input of CAB. The number of channels required by \( Z \) to detect each category discrimination area is obtained. \( Z \) retains half of the features. The dropout function is removed, and all are the features that are predicted.

\[ Z_{\text{G-OUT}} = Z_{\delta, \text{OUT}} \otimes (\delta(C_G(Z_{\delta, \text{OUT}})))) \]  

(3)

The representative in Eq. (4) responds to the importance of each category feature map. GMP stands for global max pooling. By averaging the sum of \( Z \) pooling, the score \( S \) represents the true label of the original image, and \( y \) represents the predicted label. Since loss is a multioutput function, the calculation of loss is also a multiple process. For example, Eq. (9) is affected by the error, so when the error is large, the weight updates quickly, while when the error is small, the weight updates slowly.

\[ Z_{\text{avg}} = \frac{1}{m} \sum_{i=1}^{m} \text{GMP}(z_i^e) \in \{ 1, 2, 3, ..., L \} S = \{ S_1, S_2, ..., S_L \} \]  

(4)

\[ Z_{\text{avg}} \] in Eq. (5) represents the mapping response of the semantic features of the i-th class. \( z_i^e \) represents the response of the j-th feature of the i-th class in \( Z \). The calculated scores for each class are multiplied, summed, and averaged by the semantic features of that class to obtain \( \text{GAB} \), which provides the area of diagnosis, as shown in Eq. (6). Finally, \( Z_{\text{G-OUT}} \) can be transformed into feature map \( Z_{\text{G-OUT}} \) through the category attention mechanism. We can obtain the classification situation after sample equilibrium more accurately. Compared with the balanced sample, the efficiency is up to 97.3%.

\[ Z_{\text{avg}} = \frac{1}{m} \sum_{i=1}^{m} z_i^e i \in \{ 1, 2, 3, ..., L \} \]  

(5)

\[ A_{\text{GAB}} = \frac{1}{L} \sum_{i=1}^{L} S_i Z_{\text{avg}} A_{\text{GAB}} \in R^{H \times W \times 1} \]  

(6)

\[ Z_{\text{G-OUT}} = Z_{\text{G-IN}} \otimes A_{\text{GAB}} \]  

(7)

4.3. Setting hyperparameters

In this paper, the hyperparameters of the model are tuned, and much work is mainly done on the batch size (Batch_size), optimizer (Optimizer), loss function (Loss), and normalization operation (BN). The best results of the work comparison are provided in Table 2. Although the sample of the initial dataset was unbalanced, the optimizer combined with Adam was faster and more efficient than other optimizers with the application of class weight and the adjustment of the attention mechanism [36]. The Adam optimizer has the highest accuracy and plays an indispensable role in other deep learning algorithms in the medical field [37]. The difference between the probability distribution trained by the cross entropy loss function (categorical_crossentropy) [38] and the true distribution is obtained. It describes the distance between the actual output (probability) and the expected output (probability); that is, the smaller the cross entropy value is, the closer the two probability distributions will be [39], as shown in Eq. (8). At the same time, the Label_smoothing function is used to set the parameter “Label_smoothing” to smooth the label. It increases the generalization ability of the model and prevents overfitting to some extent.

\[ \text{loss} = - \sum_{i=1}^{n} \hat{y}_i \log y_i + \hat{y}_2 \log y_2 + \cdots + \hat{y}_m \log y_m \]  

(8)

\[ \frac{\partial \text{loss}}{\partial y_i} = - \frac{\hat{y}_i}{y_i} \]  

\[ \frac{\partial \text{loss}}{\partial y_2} = - \frac{\hat{y}_2}{y_2} \]  

\[ \frac{\partial \text{loss}}{\partial y_m} = - \frac{\hat{y}_m}{y_m} \]  

(9)

\( n \) represents the number of samples, \( m \) represents the number of classifications, \( \hat{y} \) represents the true label of the original image, and \( y \) represents the predicted label. Since loss is a multioutput function, the calculation of loss is also a multiple process. For example, Eq. (9) is affected by the error, so when the error is large, the weight updates quickly, while when the error is small, the weight updates slowly.
5. Work results and analysis

5.1. Fivefold Cross-validation

To effectively adjust the volatility and stability of the detection results, the model in this paper adopts the k-fold cross-validation method. We randomly divided the dataset into m equal parts, which gives $D_1, ..., D_m$. For example,

$$\bigcup_{i=1}^{m} D_i = D_{\text{dataset}} \forall 1 \leq i \leq m$$  \hspace{0.5cm} (10)

In this work, we set $m = 5$ and perform the following steps:

Step 1: Divide the data sample into 5 equal parts.

Step 2: Take one copy of each work for testing, and use the rest for training.

Step 3: Average five times.

As shown in Fig. 7, for the first time, the work takes the first copy as the test set and the rest as the training set. For the second time, the work takes the second copy as the test set and the rest as the training set. The result is averaged by Eq. (11).

$$\text{Accuracy}_{\text{test}} = \frac{1}{m} \sum_{i=1}^{m} \text{Accuracy}_i$$  \hspace{0.5cm} (11)

From Table 3 and Table 4, the work of fivefold cross-validation is carried out for two classes and three classes, respectively. Finally, the performance indicators (precision, recall and F1-score) are considered, and the classified reports are given. The overall performance is obtained by averaging each fold. Through the corresponding confusion matrices in Figs. 11 and 12, the classification performance can be analyzed better.

5.2. Ablation experiment

To evaluate the effectiveness of model fusion with the addition of an attention mechanism, we conducted additional ablation experiments on chest X-ray, as shown in Table 5. The work application model is mainly composed of DenseNet and VGG16. One component is deleted at a time, including model actions, attention actions, and fusion actions. First, the attention mechanism is removed, and the accuracy, precision, recall and F1-score under a single model are calculated separately. The accuracy index decreased by approximately 0.04 at most, and other evaluation criteria also declined. Second, the overall performance of the two models is slightly improved by combining the relevant features. The accuracy increased by approximately 0.03. Finally, the attention mechanism is added to the basic model, and the overall index rises by approximately 0.01, indicating that the attention module plays an important role in enhancing relevant features. Therefore, the proposed feature fusion model can improve the efficiency of chest X-ray image classification to a certain extent.

5.3. Comparison of work effects

In this paper, the work environment carried out research and discussion on multiple classifications of two different datasets: healthy vs. pneumonia-type patients and healthy vs. pneumonia vs. COVID-19 patients.

5.3.1. Work analysis of HEALTHY patients and patients with pneumonia (Two Categories)

Chest X-ray images were used to distinguish the work of healthy patients from those of pneumonia patients (bacterial, viral, and COVID-
From Fig. 8(a)(b), we can clearly see that 80 iterations can make the accuracy stable. In the first 20 iterations, the growth rate of accuracy increased significantly, and then the growth rate gradually showed a slow upward trend. In comparison with other advanced models, work in this field has achieved good results, as shown in Table 6. The results in Table 7 show that the accuracy effect calculated by this work model is relatively successful. The confusion matrix in Fig. 11 clearly shows the classification effect of the binary classification. In addition, it also has a relatively high result in other evaluation indicators, and the model has obvious advantages in classification.

5.3.2. Work analysis of Healthy, common pneumonia and COVID-19 patients (Three Categories)

To verify the generalization ability of the model, we also conducted research on the three categories of healthy, common pneumonia and COVID-19 patients. The comparison of loss and val_loss in three categories is shown in Fig. 9(a)(b). The results in Table 8 show that the accuracy effect calculated by this work model is relatively successful. The confusion matrix in Fig. 11 clearly shows the classification effect of the three categories. In addition, it also has a relatively high result in other evaluation indicators, and the model has obvious advantages in classification.
COVID-19 patients. Different forms of datasets are input for different models, and detailed segmentation is used to better optimize the X-ray images of the chests and extract key features. From the work results in Fig. 9(a)(b), the fit of the training data and the test data curve is good. During the first 15 iterations, the overall upward trend was faster. We also compare this field with other advanced models, as shown in Table 8. The accuracy can reach 97.3%.

Similar datasets are compared in Table 9. The accuracy of the evaluation index is generally stable and high. Compared with the recall data in Ref. [19], the value is relatively low, but the average value among other data is relatively stable and above 0.95. According to the model commonly used in medical images, a comparison is made. Fig. 10 and Table 10 show the comparison of accuracy, loss and time consumption of different models on the same dataset. Under different models, our model training and test sets have the highest accuracy and the lowest loss. Although slightly longer in time consumption, the overall performance is the best. The confusion matrix in Fig. 12 clearly shows the classification effect of the three category classification.

As shown in Fig. 13(a)(b), the ROC distribution diagram also shows that the accuracy of the ROC in the second classification is as high as 99%, and the accuracy of different diseases in the third classification fluctuates at 99%. It can also be better explained that the higher the index is, the higher the accuracy of the model diagnosis.

6. Limitations

Due to the limited number of COVID-19 datasets now available for public research applications, the sources for obtaining COVID-19 images are constantly updated. The research object of the work is relatively limited. Two publicly available datasets were used in this study. How to

| Reference | Accuracy | Precision | Recall | F1-score |
|-----------|----------|-----------|--------|----------|
| [50]      | 0.910    | 0.920     | 0.870  | 0.880    |
| [51]      | 0.940    | 0.913     | –      | –        |
| [19]      | 0.95     | 0.950     | 0.969  | 0.956    |
| Our Model | 0.973    | 0.955     | 0.948  | 0.951    |

Fig. 10. (a): Comparison of accuracy and val_accuracy of different models under the same dataset; (b): Comparison of loss and val_loss of different models under the same dataset.

Fig. 11. (a): Confusion matrix for binary classification (HEALTHY vs. PNEUMONIA vs. COVID-19); (b): Confusion matrix (percentage) for binary classification (HEALTHY vs. PNEUMONIA vs. COVID-19).
better assist doctors in becoming an important part of the clinical work testing process requires more data and image learning and research.

7. Conclusion

At a time when an increasing number of people are infected with COVID-19, using AI methods to create effective and rapid tests will certainly reduce the workload of healthcare workers. In this paper, a feature fusion-based chest X-ray image classification method is proposed. Successful detection of healthy, common pneumonia, and COVID-19 cases from chest images (X-ray) is demonstrated. By using ResNet34 to effectively process the segmentation of the dataset, the extraction of features is more efficient. Based on the fusion of DenseNet and VGG16, GAB and CAB attention blocks can be added to carry out detailed feature extraction of regions. Due to the highly similar X-ray results of COVID-19 and other common pneumonia diseases, doctors may have some misdiagnoses in diagnosis. It is therefore crucial to make a precise distinction between ordinary pneumonia and COVID-19. The work showed that when distinguishing ordinary pneumonia, COVID-19 and healthy patients, the accuracy, precision, recall and F1-score reached 97.3%, 95.5%, 94.8% and 95.1%, respectively. Compared with those with pneumonia, the accuracy, precision, recall and F1-score reached 97.9%, 95.0%, 96.0% and 96.0%, respectively. The model was tested using X-ray images of tuberculosis to see if it had a strong generalization ability [52]. The accuracy, precision, recall and F1-score reached 99%, 99%, 99% and 99%. Despite the positive results, more clinical trials and studies are needed to test this model. Better auxiliary imaging for doctors, improvement of the accuracy of diagnosis, and a more in-depth study of the differences between different pneumonias in a timely manner so that patients could avoid the pain of the disease are needed. Considering the present work, there are still gaps in some research fields, which can be discussed and realized in future work:

1. To better realize clinical applications, it is necessary to expand the number of multiclass datasets so that the model can accurately judge chest X-ray and provide a more accurate classification diagnosis.
2. The proposed model should be further optimized and improved and combined with an SVM classifier to improve the matrix disease diagnosis ability.

CRediT authorship contribution statement

Lingzhi Kong: Conceptualization, Methodology, Software, Writing – original draft. Jinyong Cheng: Supervision, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Scohy, A. Anantharajah, M. Bod, A new coronavirus of probable bat origin, Nature. 579 (7798) (2020) 270–274.
[2] M. Guan, L. Zhou, Y. Wu, X. Li, X. Ou, C. Dai, J. Shen, et al., Clinical features of 152 predicted novel coronavirus pneumonia in China, Lancet. 395 (10223) (2020) 507–514.
[3] X. Wang, Y. Peng, L. Lu, Z. M. Bagheri, R.M. Summers, ChestX-ray8: Hospital-localized chest X-ray image dataset with annotations, in: Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR, 2017, https://doi.org/10.1109/CVPR.2017.369.
[4] Jinyong Cheng, Eus, B. Kabamba-Mukadi, A. Verroken, COVID-19 and what are their advantages and disadvantages? Cent, Evidence-Based Med, 2020.
[5] A.I. Khan, J.L. Shah, M.M. Bhat, CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images, Comput. Methods Programs Biomed. 196 (2020) 105809.
[6] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proc. - 36th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR, 2017. https://doi.org/10.1109/CVPR.2017.243.
[7] W. Hao, S. Zhili, Research on image classification based on improved DenseNet, ACM Int. Conf. Proceeding Ser, 2021, https://doi.org/10.1145/3477124.
[8] X. Chang, J. Wu, T. Yang, G. Feng, DeepFake Face Image Detection based on Improved VGG Convolutional Neural Network, in: Chinese Cont. Conf. 2020, https://doi.org/10.23919/C00060.2020.9189596.
[9] J. Luojia-García, C. Várez-Márquez, Y. Villarreal-Soza, O. Camacho-Nieto, A transfer learning model based on concatenation approach for the diagnosis of brain tumor, IEEE Access. 8 (2020) 5514–55144.
[10] Y. Cheng, J. Feng, K. Jia, A Lung Disease Classification Based on Feature Fusion Convolutional Neural Network with X-ray Image Enhancement, in: 2018 Asia-Pacific Signal Inf. Process. Assoc. Ann. Summit Conf. APSIPA ASC 2018 - Proc., 2019, https://doi.org/10.23919/APSIPA.2018.8652970.
[11] C. Huang, Y. Wang, X. Li, Y. Ren, X. Zhang, Y. Chen, et al., A multi-scale cascaded method for detection of coronavirus (Covid-19) from chest x-ray images, Chaos, Solitons and Fractals. 140 (2020) 110245.
[12] L. Kong and J. Cheng: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia.
[45] M. Hu, H. Lin, Z. Fan, W. Gao, L. Yang, C. Liu, Q. Song, Learning to recognize chest-xray images faster and more efficiently based on multi-kernel depthwise convolution, IEEE Access. (2020), https://doi.org/10.1109/ACCESS.2020.2974242.

[46] E. Ayan, B. Karabulut, I.M. Ünver, Diagnosis of pediatric pneumonia with ensemble of deep convolutional neural networks in chest x-ray images, Arab. J. Sci. Eng. (2022), https://doi.org/10.1007/s13369-021-06127-z.

[47] S. Heidarian, P. Afsar, N. Enshaei, F. Naderkhani, M.J. Raffee, F. Bahaki Fard, K. Samimi, S.F. Atashzar, A. Oikonomou, K.N. Plataniotis, A. Mohammadi, Covid-19 and pneumonia classification using deep learning for identification of COVID-19 cases from chest CT Scans, Front. Artif. Intell. (2021), https://doi.org/10.3389/frai.2021.598932.

[48] M. Alotaibi Detection of COVID-19 using deep learning on X-ray images Intell. Autom. Soft Comput. 2021 https://doi.org/10.32604/iasc.2021.018350.

[49] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Q. Ni, Y. Chen, J. Su, G. Lang, Y. Li, H. Zhao, J. Liu, K. Xu, L. Ruan, J. Sheng, Y. Qiu, W. Wu, T. Liang, L. Li, A deep learning system to screen novel coronavirus disease 2019 pneumonia, Engineering, (2020), https://doi.org/10.1016/j.eng.2020.04.010.

[50] J.E. Luján-García, M.A. Moreno-Ibarra, Y. Villuendas-Rey, C. Yáñez-Márquez, Fast COVID-19 and pneumonia classification using chest X-ray images, Mathematics. (2020), https://doi.org/10.3390/MATH8091423.

[51] A.U. Ibrahim, M. Ozsoz, S. Serte, F. Al-Turjman, P.S. Yakoi, Pneumonia classification using deep learning from chest X-ray images during COVID-19, Cognit. Comput. (2021), https://doi.org/10.1007/s12559-020-09787-5.

[52] Tuberculosis (TB) Chest X-ray Database. https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset.