



BSD: A multi-task framework for pulmonary disease classification using deep learning

Sanli Yi ^{a,b,*}, Shenglin Qin ^a, Furong She ^a, Dangguo Shao ^a

^a School of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China

^b Key Laboratory of Computer Technology Application of Yunnan Province, Kunming, Yunnan, China



ARTICLE INFO

Keywords:

Bone suppression
Segmentation network
Classification network
Pulmonary disease detection

ABSTRACT

The diagnosis of pulmonary diseases using deep learning on chest X-ray images can be affected by the bone structures, the tissue in regions outside the lungs, and the characteristics of the images as burrs, blurring and complex pulmonary structures. To address these issues, a new pulmonary diseases diagnostic framework named BSD is proposed, the innovations of which are: Firstly, three steps of bone suppression, pulmonary parenchyma extraction and pulmonary diseases diagnosis are included. The first two steps can eliminate the influence of bone structures and other tissues. Therefore, better diagnostic results can be obtained based on the extracted boneless pulmonary parenchyma. Secondly, aiming at the characteristics of the chest X-ray image, a new segmentation network named IU-Net is proposed for the pulmonary extraction, in which the original convolution layers of U-Net are replaced by our new designed function modules. Finally, based on the characteristic of boneless pulmonary parenchyma, an enhanced functional module RCA which can better extract the image features is designed, therefore improved the disease diagnosis. In the experiments on the chest X-ray pneumonia datasets, the framework achieved satisfied results in the binary classification with an accuracy and Kappa of 98.73% and 96.80%, respectively. And in the multi-classification task on the COVID-19 radiology database, it achieved an accuracy and Kappa of 95.76% and 95.75%, respectively, which is better than the state-of-the-art. Therefore, BSD framework can better facilitate physicians to diagnose pulmonary disease and increase the diagnostic capabilities.

1. Introduction

Pulmonary disease ranks third among all diseases that cause death, and according to the World Health Organization, more than 5 million people die each year worldwide (Bhandary et al., 2020; Han et al., 2020). Early diagnosis and treatment of pulmonary diseases is one of the important ways to reduce mortality.

Chest X-ray image is a widely used pathological image in the clinical diagnosis of pulmonary diseases (Parveen and Sathik, 2011). However, such images usually have burr, blurring and interference factors of bone structures, which leads to the diagnosis of pulmonary diseases such as pneumonia and COVID-19 to be very easily confused with other diseases, thus affecting the diagnostic accuracy of physicians (Stefan et al., 2022). The use of computer-aided diagnosis (CAD) can help doctors obtain more accurate results objectively and efficiently (Katsuragawa et al., 1989). And in the diagnosis of pulmonary diseases, many researchers have recently made different contributions using CAD, with

deep learning techniques playing a major role.

Deep learning has developed rapidly in recent years. In 2012, Krizhevsky et al. (2012) proposed AlexNet. In 2015, VGG (Simonyan and Zisserman, 2015) was proposed. And then Szegedy et al. (2015) proposed Inception structure can extract multi-scale features. Ronneberger et al. (2015) proposed U-shaped network, which has been widely employed in medical image segmentation. In 2016, ResNet was introduced, which connects the input and output of convolutional layers by skip connection to overcome the network's degradation (He et al., 2016). In 2017, Huang et al. (2017) proposed DenseNet by using dense connections to improve network performance. In 2018, MobileNet (Sandler et al., 2018) used deep separable convolution and point convolution to form a lightweight network. In 2019, Tan (2019) proposed EfficientNet to enhance the capability of the network by setting the scaling factors to balance the relationship between network depth, width and resolution. In 2021, Wang et al. (2021), believed that improving the input image's resolution could enhance the network's

* Corresponding author at: School of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China.

E-mail addresses: 152514845@qq.com (S. Yi), 1213763114@qq.com (S. Qin), furong.she@qq.com (F. She), litch2004@sina.com (D. Shao).

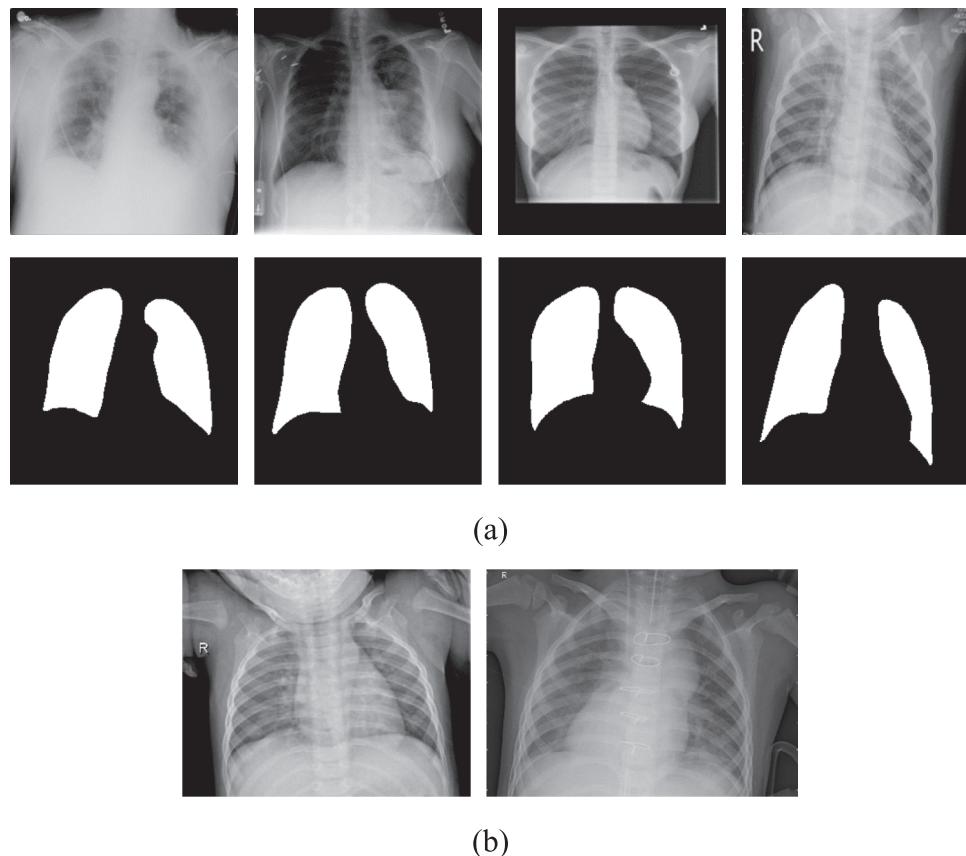


Fig. 1. Sample Graphs of dataset: (a) the first row shows example figures of D1, COVID-19, Lung Opacity, Normal, Viral Pneumonia from left to right, and the second row shows their corresponding pulmonary parenchyma masks. (b) example figures of D2, Normal, Pneumonia from left to right.

capability of extracting features. In 2023, Zhao et al., 2022 proposed a step-by-step gradient penalty with similarity calculation summary generation model, which can reduce the training time. For further enhancements to the existing network, many functional modules have been recently introduced and applied. For example, in 2021, the Coordinate Attention module (CA) proposed by Hou et al. (2021) can recalibrate the most critical information among many feature information, which enhance the feature representation and the transmission of feature information. In 2023, Zhao et al., 2023 proposed T-softmax to improve the accuracy of the models.

As the deep learning technology is developed, it is increasingly used in the diagnosis of pulmonary diseases. Including using radiology reports and medical images. For instance, based on radiology reports, Zhao, S et al., (2023) using a natural language processing method to extract key information of doctors' diagnosis. In the field of medical image, Li et al. (2019) identified pneumonia using a convolutional neural network, with an accuracy of 92.79 %. Subsequently, many researchers used improved CNN to diagnose pneumonia by employing chest X-ray images (Racic et al., 2021; GM et al., 2021). Aiming at the phenomenon that pulmonary diseases show blurring and opacity in X-ray images, Rajpurkar et al. (2017) performed multi-classification diagnosis of multiple types of pulmonary diseases by using DenseNet121 network with an average accuracy of 84.13 %. Kundu et al. (2021) used multiple networks to extract features of chest X-ray images and then by using the hyperbolic tangent function to select the best result with an accuracy 98.81 %. Later, in 2018 a study (Hussain et al., 2018) proves that the model's capability of image classification could be enhanced by using transfer learning. Since then, an extensive use has been made of it for the medical images' processing. For example, Ayan and Unver (2019) combined Xception, VGG19 network with transfer learning to detect pneumonia, and the accuracy reached 85 % and 87 %,

respectively. Rahman et al. (2020) combined multiple convolutional neural networks with transfer learning to realize pneumonia diagnosis and obtained better diagnostic results. Along as deep learning applications in the medical field develop, it is increasingly becoming a research hotspot to construct deep learning diagnostic models based on lung image features and pathological knowledge. Rahman et al. (2020) considered that other tissues in chest X-ray images would affect the extraction of pulmonary feature information. Based on this view, they constructed a new diagnostic model, which firstly used U-Net network to segment the pulmonary parenchyma, and then used DenseNet201 to diagnose pulmonary tuberculosis based on the segmented pulmonary parenchyma image. Based on the same view, Manickam et al. (2021) designed a diagnostic model for pneumonia and Sharma and Mishra (2022) built a diagnostic model for COVID-19. Above studies could obtain better results for their models were constructed on the images of pulmonary parenchyma which are firstly segmented from original pulmonary images. However, these pulmonary parenchyma images still have bone structures that affect the diagnosis. Rajaraman et al. (2021) constructed a model based on CNN to remove bone structures in chest X-ray images, and then diagnose pulmonary tuberculosis by using the network. While removing the bone structures from the chest X-ray images and then building a diagnostic model led to better diagnostic results, but they didn't consider the impact of other tissue in the image on the diagnosis of pulmonary disease.

Based on the above statements, we can see: (1) The bone structures in chest X-ray images will affect the diagnosis of pulmonary diseases, therefore removing them will make the diagnosis more effective; (2) The extraction of pulmonary parenchyma from chest X-ray images can obtain richer and more effective features, which can better enable the diagnosis of pulmonary diseases; (3) Based on the characteristics of pulmonary images, i.e., the presence of burr, blurring and complex

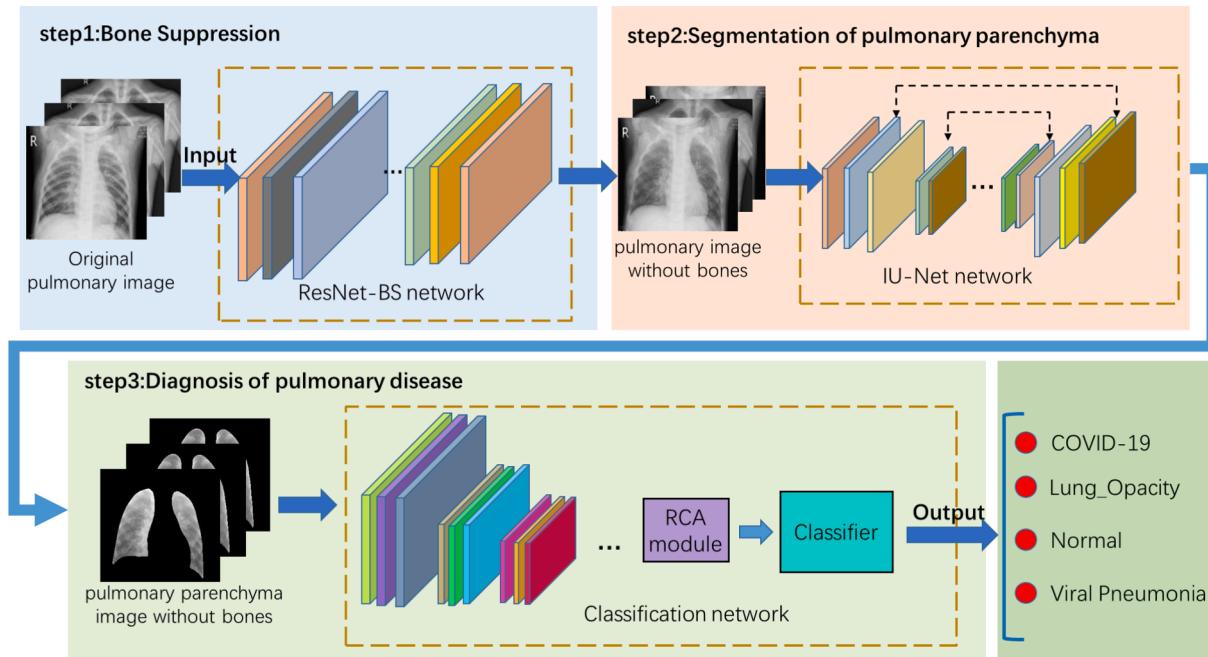


Fig. 2. Architecture of BSD framework: step1 Bone Suppression; step2 Segmentation of pulmonary parenchyma; step3 Diagnosis of pulmonary disease.

pulmonary structures, well-designed segmentation network and classification network are needed to better diagnose such disease, among which, segmentation network can better extract pulmonary parenchyma, which is more helpful for the diagnostic task, and classification network can better diagnose pulmonary disease.

In this paper, a new multi-step pulmonary disease diagnosis framework named BSD is proposed, based on which the main contributions of it are as following:

1. In order to eliminate the influence of bone structures and other tissues, the framework is composed of 3 processing steps: Step1, a bone suppression process that can eliminate the interference of chest bone structures is performed to obtain the images without bone structures; Step2, pulmonary parenchyma segmentation is performed to reduce the interference of other tissue; Step3, the obtained boneless pulmonary parenchyma image is input into the classification network for better diagnosis of the pulmonary disease.
2. In step2, aiming to resolve the above problems of Chest X-ray images, a new segmentation network named IU-Net is proposed. Based on U-shaped structure, IU-Net uses the new designed function modules to replace the original convolution layers of U-Net, which have the enhanced functions of containing multi-scale, large receptive fields, and channel attention to better extract image features.
3. In step3, to better perform the diagnosis process on the boneless pulmonary parenchyma image, a new enhanced functional module RCA is proposed, which is added behind the pre-trained CNNs and can better extract the image features both in space and channel.

The rest parts of this article are described below: [Section 2](#) gives the proposed method in detail; [Section 3](#) describes the experiments and results to verify the method; [Section 4](#) summarizes the significance and contribution of this work.

2. Materials and Methods

In this part, the datasets are introduced, followed by the data pre-processing steps and a detailed description of the framework's architecture.

2.1. Dataset

The two datasets used in this paper are downloaded from the kaggle platform, which are denoted as D1 and D2, respectively.

D1 presents COVID-19 radiography-database ([Rahman et al., 2021](#)), which was established by Qatar University and Dhaka University. It contains four types of images, including 3616 COVID-19 positive cases, 10,192 normal cases, 6012 lung opacity cases (non-COVID lung infection), 1345 viral pneumonia images, and corresponding pulmonary parenchyma masks. Based on the four types of chest X-ray images and corresponding pulmonary parenchyma masks, D1 can be used for the multi-classification task and pulmonary parenchyma segmentation network by using our proposed pulmonary disease diagnostic framework. [Fig. 1\(a\)](#) shows an example diagram for each category in this dataset.

D2 presents Chest X-Ray Images ([Kermany et al., 2018](#)), which is a pneumonia datasets including chest X-ray images from children aged 1–5 years. The dataset contained 4,273 images of pneumonia, 1,583 normal images, and a total of 5,856 chest X-ray images. Based on these two types of chest X-ray images, D2 can be used for the binary classification task by using our proposed pulmonary disease diagnostic framework. [Fig. 1\(b\)](#) shows an example diagram of each category in this dataset.

2.2. The proposed diagnose framework

The proposed pulmonary diagnostic framework BSD can better diagnose pulmonary disease on chest X-ray image. [Fig. 2](#) illustrates the framework's architecture, which is mainly composed of three steps: In step 1, bone suppression is performed on pulmonary images to eliminate the interference of chest bone structures; In step 2, to reduce interference from other tissue, the pulmonary parenchyma is extracted from the image without bone obtained in step 1, for which a new segmentation network IU-Net is proposed; Step 3 is the diagnoses of pulmonary disease, in which the pulmonary parenchyma image without bone obtained in step 2 is used and a new function module RCA is designed. Each step will be described in the subsequent sections.

2.2.1. Bone suppression

Due to the overlap with the bone structures, the pulmonary region in

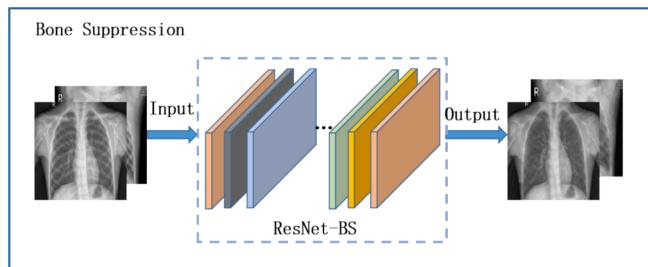


Fig. 3. The architecture of Bone Suppression.

the whole chest X-ray image usually contains subtle abnormalities. Therefore, the first step in our BSD framework is to remove the bone structures from the chest X-ray images. Rajaraman et al. (2021) proposed four convolutional neural networks to remove bone structures from chest X-ray images,¹ among which ResNet-BS network can effectively remove it from chest X-ray images by using residual structure to build a deeper network. Therefore, we applied this network to step 1 of BSD to remove it. Fig. 3 illustrate its processing flow, in which the original image is the input, and the pulmonary image without bone structures is the output.

2.2.2. Segmentation of pulmonary parenchyma

The step 2 of the BSD framework is to segment the pulmonary parenchyma from images without bone structures. U-shaped network (Ronneberger et al., 2015) is widely used in image segmentation for its unique structure of encoding, decoding and jump connection. However, when segmenting the pulmonary parenchyma, there exists several problems. First, the chest X-ray image contains lots of burrs and blurring which make the networks difficult to segment local areas and small contours accurately. Second, the complex structure of pulmonary images makes the segmentation task more difficult than ordinary images, such as the segmentation task is often influenced by the costophrenic angle and diaphragm.

Therefore, we designed a new segmentation network IU-Net, its structure is shown in Fig. 4: Based on the U-shaped structure, the network is composed of encoding part and decoding part, each of them has 4 layers, where the original convolution layers are replaced by new function modules as Module A, Module lite-A, Module B, and Module C. The detailed description is as below:

The Encoding part is mainly constructed by Module A and Module lite-A. Both of them are composed of inception block, scSE block and RD block, their functions are: (1) Feature maps of previous layer are firstly sent into inception block, which has 3 branches with different convolution kernel size of 1×1 , 3×3 , and 5×5 . Therefore, enhance the network's ability of extract image features in multi-scale, and obtained feature maps with sufficient information. (2) To extract useful features from lots of information in these maps more efficiently, scSE block is added behind it, which can strengthen key features and suppress useless features. (3) Then RD block is added sequentially, where three convolution layers that make the network deeper are applied for better extracting deep semantic information of the feature maps. Furthermore, we use a depthwise convolution with kernel of 7, which has a large receptive field and a smaller number of parameters, therefore enabling better extracting global features without increasing the network parameters. The difference of Module A and Module lite-A are shown in Fig. 4(b), the fusion operation of three branches in Inception block of Module A is concatenate, while Module lite-A is add operation with channels much fewer than Module A. The reason is that as the network deepens, the operation of concatenate multiplies the computation without increasing efficiency. Therefore, we use Module A in the first

three layers for better maintain useful features and Module lite-A in the fourth layer for efficiency.

In the Decoding part, each layer is constructed by Module B and Module C, which is shown in Fig. 4(c). Module B is firstly concatenate with Module A from the same layer in the encoding part for more feature details and then their combination are sent into Module C for better feature extraction. Module B is composed of UpSampling, Conv2d and BN, which can expand the resolution of the feature maps of the previous layer by using bilinear interpolation to reduce parameters. Module C is composed of Conv2d, RD block, BN and scSE block, based on which useful feature can be better extracted from the output of the combination.

2.2.3. Classification of pulmonary disease

After removing bone structures and segmenting the pulmonary parenchyma from chest X-ray images, i.e. step 1 and step 2 of the BSD framework, the pulmonary parenchyma image without bone structures and other tissues is segmented from the original image, then it is put into step 3 of the framework to perform the classification of pulmonary disease. During this procedure, due to the blurring and burr of chest X-ray images, effective and accurate diagnostic results cannot be achieved using traditional convolutional neural networks. Thus, for extracting the image's features better, we enhanced the traditional convolutional network by employing a new enhancement module, RCA module. Then the network added with RCA module is used in the step 3 of our framework to realize the diagnosis of pulmonary diseases.

Fig. 5(a) illustrated the classification network's construction, which is designed for step 3: First, the pulmonary parenchyma data is sent into the traditional classification network. Since our classification task is based on the small dataset, which would cause overfitting and inefficiency, we pre-trained the network on the ImageNet dataset. Second, the network's output is sent into the RCA module to strengthen the feature extraction capability of the network. Last, we put the extracted feature maps into the classifier to realize the classification of pulmonary diseases.

The RCA module is shown in Fig. 5(b). It is a new enhanced functional module based on the RA module structure (Yi et al., 2021), in which the GC module is replaced by the CA module, which can fuse channel attention and spatial attention to highlight important features and suppress useless features for better image feature acquisition. The pipeline of RCA is detailed as follows: Firstly, the input feature maps F_{in} is fed into two branches, and then two feature maps F_1 and F_2 are obtained respectively. Among which, F_1 contains the local information by using three convolution layers, and F_2 contains the global information by using a 1×1 convolution layer. Secondly, F_1 is input to the CA module, which has the abilities mentioned above. Finally, the output F_{out} is obtained by adding F_2 and F_3 .

3. Experiments

We firstly describe the experimental setup and performance metrics. After that, we performed experiments to verify the function of each step of BSD and to verify the whole framework's efficiency. The details of these experiments are as follows: Step 1 is the processing of bone suppression; Step 2 is segmentation experiment, in which the ablation experiments and comparison experiments are performed to testify the efficiency of the IU-Net network; Step 3 is classification experiment, in which the ablation experiments are performed to testify efficiency of RCA module; In BSD experiment, ablation experiments and comparison experiments are performed to verify the whole framework.

3.1. Environmental settings and performance metrics

The experiments are performed on Intel Core™ i9-10900 k CPU@3.70 GHz, 64.0 GB RAM, RTX3080. The integrated development environment is PyCharm by Python, choosing Keras based on

¹ <https://github.com/sivaramakrishnan-rajaraman/CXR-bone-suppression>.

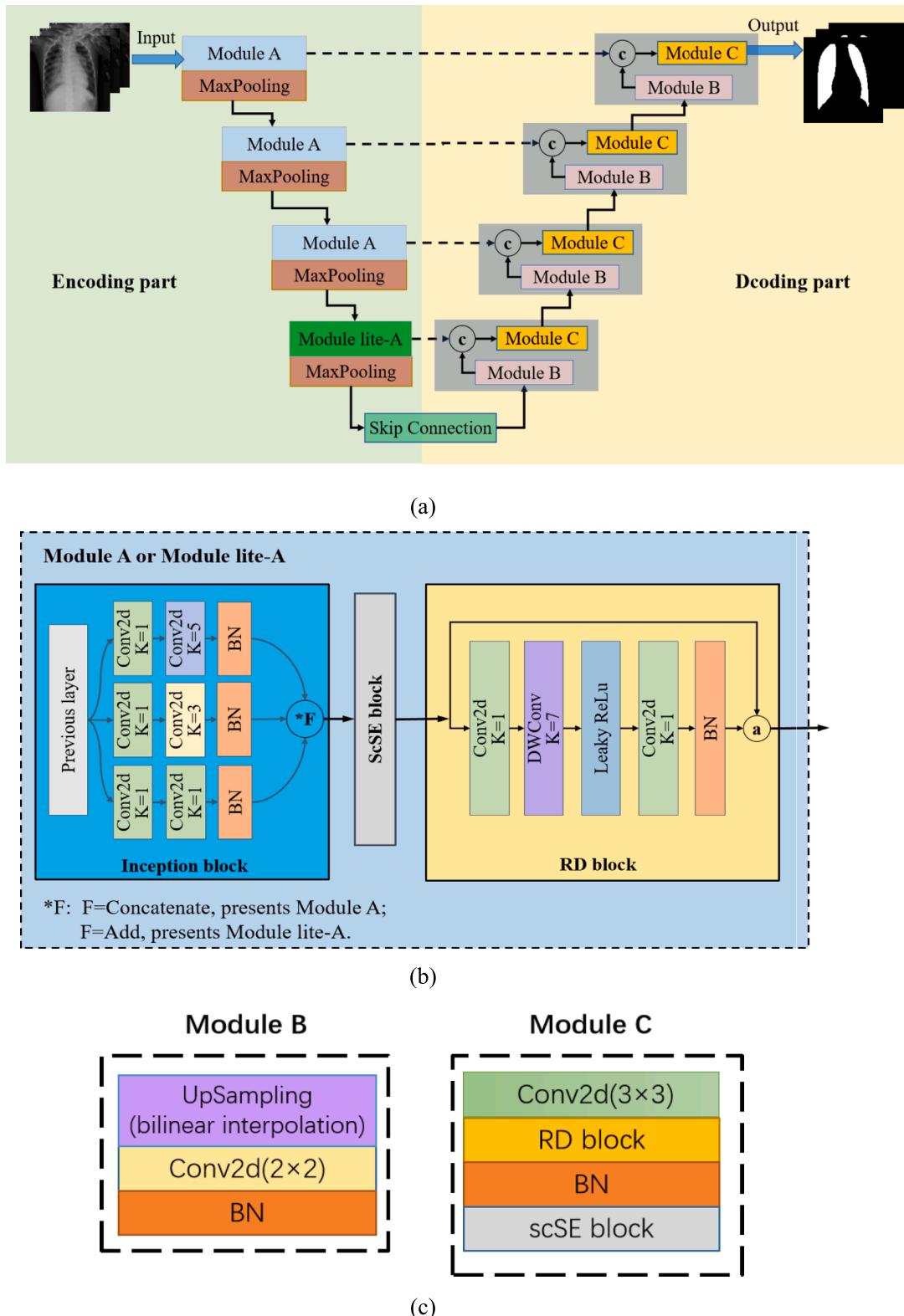


Fig. 4. Structure of IU-Net: (a) the whole Structure of IU-Net; (b) Module A and Module lite-A; (c) Module B and Module C.

Tensorflow.

In order to testify our BSD framework, D1 and D2 are used in our experiments, among which D1 is used for the pulmonary parenchyma segmentation and multi-classification task, and D2 is used for the binary classification task. They are split into a dataset of training and testing in the ratio of 8:2, respectively. Further explanation of the data partition is

given in Table 1. By using the dataset of training to train the networks adopted in our framework, the trained networks are obtained and then used in testing set to testify them. To accelerate the network convergence, the images of these two datasets are resized and normalized.

The performance of the pulmonary disease diagnostic framework is evaluated based on the metrics as accuracy, precision, recall, F1 score,

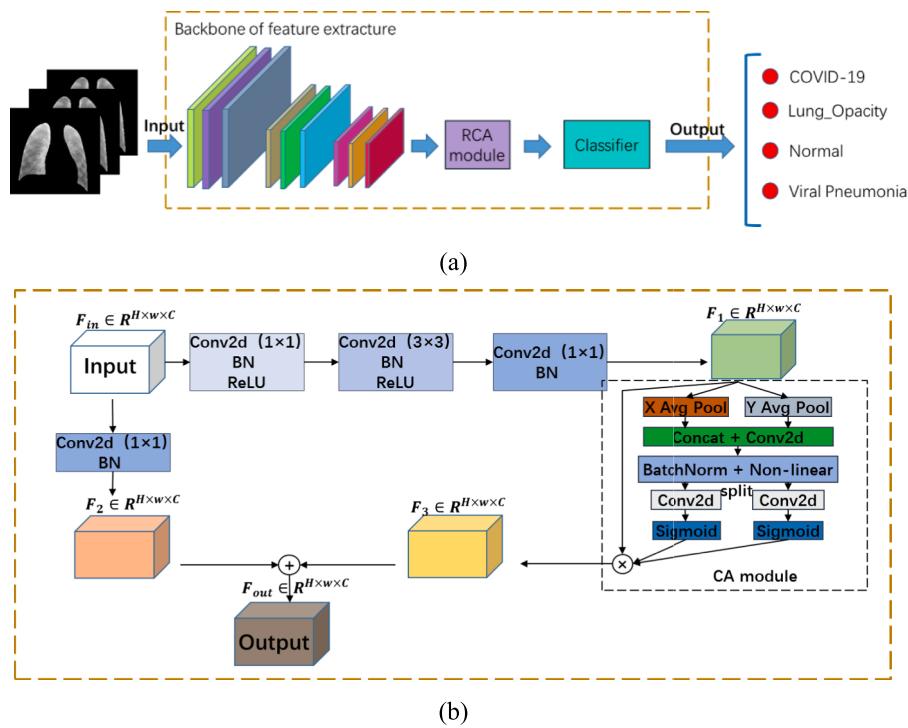


Fig. 5. Classification network of pulmonary disease: (a) The flow chart of classification model; (b) RCA module.

Table 1
The number of each category in D1 and D2.

Dataset	Category	Training	Testing	Total
D1 (COVID-19 radiography-database)	COVID-19	2892	724	3616
	Lung Opacity	4810	1202	6012
	Normal	8154	2038	10,192
	Viral Pneumonia	1076	269	1345
D2 (Chest X-Ray Images)	Normal	1267	316	1583
	Pneumonia	3419	854	4273

and so on, which are used to compare the performances and are defined as follows:

TP represents Sum of all True Positives.

FP represents All False Positives.

FN represents All False Negatives.

TN represents Sum of all True Negatives.

K represents Category.

P_o denotes the accuracy, a_n denotes the real samples' number of a certain class, b_n denotes the predicted samples' number of a certain class, and M denotes the samples' total number.

For step2: Segmentation experiment, we use accuracy, sensitivity, specificity, miou and dsc evaluation metrics, and the formula for each metrics is as follows:

$$\text{Accuracy} = \frac{\text{No. of images correctly classified}}{\text{Total no. of images}} \quad (1)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (3)$$

$$\text{Miou} = \frac{1}{k+1} \sum_{i=0}^k \frac{\text{TP}}{\text{FN} + \text{FP} + \text{TP}} \quad (4)$$

$$\text{Dice} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \quad (5)$$

For step3: classification experiment, we use accuracy, precision, recall, F1 and kappa evaluation metrics, and the formula for each metrics is as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

$$\text{F1} = \frac{2\text{TP}}{2\text{TP} + \text{FN} + \text{FP}} \quad (8)$$

$$\text{Cohen's kappa} = \frac{\text{P}_o - \text{P}_e}{1 - \text{P}_e} \quad (9)$$

$$\text{P}_e = \frac{a_1 \times b_1 + a_2 \times b_2 + \dots + a_n \times b_n}{M} \quad (10)$$

3.2. Step1: Bone suppression

Step 1 in our BSD framework is the process of removing bone structures. Thus, the input of BSD framework is firstly fed into ResNet-BS (Rajaraman et al., 2021) network to obtain the image without it. Fig. 6 shows the original image of the two datasets (D1 and D2) and the image without bone structures.

3.3. Step2: Segmentation experiment

The task of step 2 is to extract the pulmonary parenchyma for better diagnose of pulmonary disease, in which IU-Net is used. In this subsection, to verify that the network can better extract the pulmonary parenchyma, we compared the IU-Net with the currently commonly

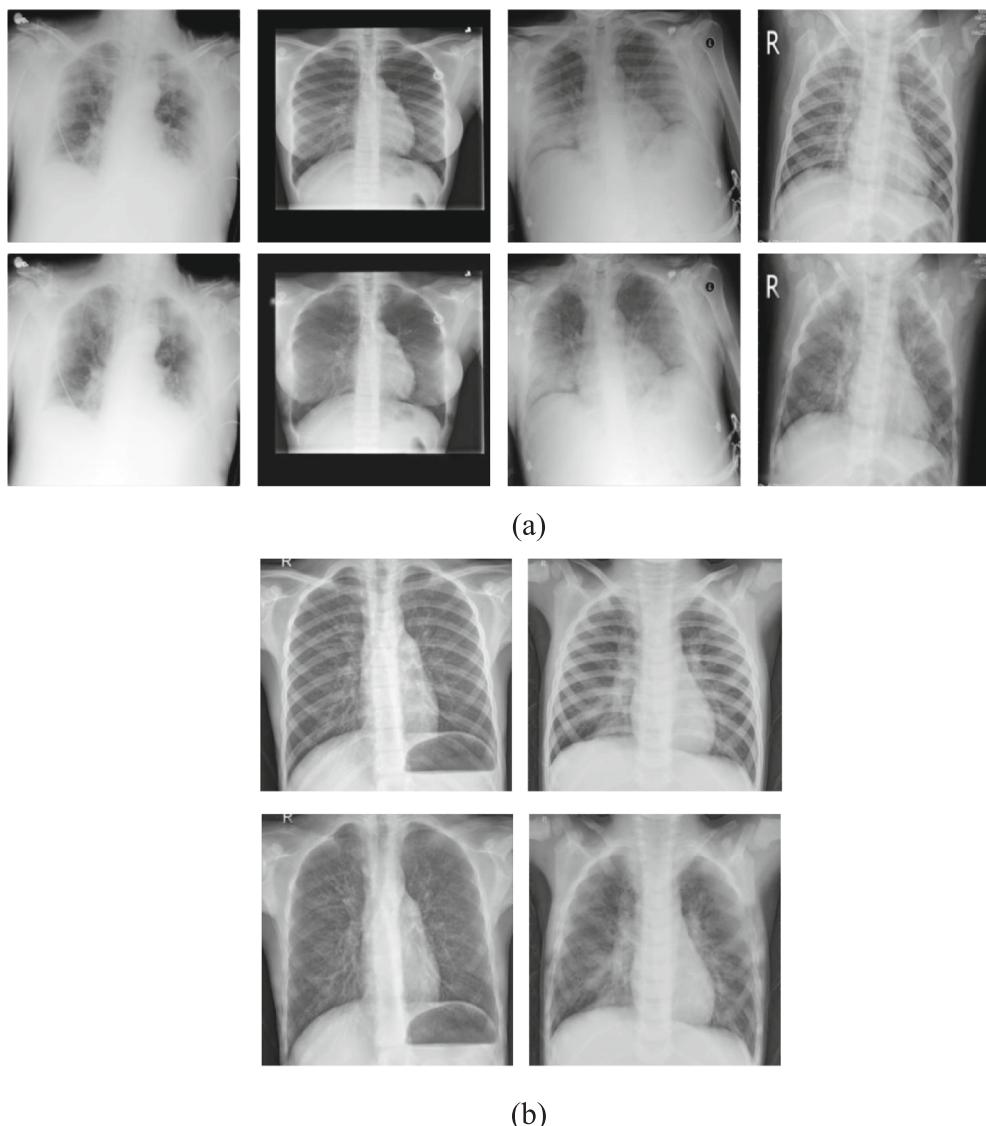


Fig. 6. Input and output diagram of step1: (a) The first line is the original image in D1, the second line is the Res-BS network's output; (b) The first line is the original image in D2, the second line is the Res-BS network's output.

Table 2
Results of different segmentation networks.

Network	Acc (%)	Sen (%)	Spe (%)	Miou (%)	Dsc (%)	Para
UNet (Ronneberger et al., 2015)	98.65	94.41	99.20	96.33	97.12	1.9 M
FCN-8S (Long et al., 2015)	98.71	94.65	99.04	96.49	97.25	5 M
SegNet (Badrinarayanan et al., 2017)	98.79	94.96	99.37	96.70	97.41	11 M
IU-Net	99.06	96.08	99.41	97.43	98.00	2.3 M

used segmentation networks. Furthermore, to prove each function module's efficiency, we carried out ablation experiments. In the experiment, the initial learning rate is set to $6e^{-4}$, the epoch of 45, the batch size of 10, and the optimizer is Adam.

3.3.1. Comparison experiments

There are lots of networks which can achieve satisfactory results in

the segmentation of image and being widely used recently, such as U-Net (Ronneberger et al., 2015), FCN_8S (Long et al., 2015) and SegNet (Badrinarayanan et al., 2017). We compare our IU-Net with these networks, and Table 2 is the results of the experiment.

As can be seen from Table 2, the IU-Net obtained the best results with accuracy, miou and dsc of 99.06 %, 97.43 % and 98.00 %, respectively. Although, it has parameters of 2.3 M, which is the second best and 0.4 M more than U-Net, its other metrics are much better than U-Net with Miou and Dsc are improved by 1.1 % and 0.88 % respectively.

Fig. 7 illustrates the segmentation results of each network, and we can see the differences more clearly from the red boxes between the various networks and the ground truth (Fig. 7(b)). Through comparing the performance of each network, the performance of the proposed IU-Net is the best: In the first row of samples, U-Net (Fig. 7(c)) and FCN-8S (Fig. 7(d)) exhibits cracked and discontinuous in the right pulmonary segmentation, which do not occur in SegNet (Fig. 7(e)) and IU-Net (Fig. 7(f)); In the second row of samples, U-Net and SegNet appear over-segmented in the left and right lungs, respectively, FCN-8S exhibits discontinuous, all of which do not occur in IU-Net; In the third row of samples, U-Net, FCN-8S and SegNet exhibits over-segmentation, which do not occur in IU-Net. Based on above segmentation results, we can see

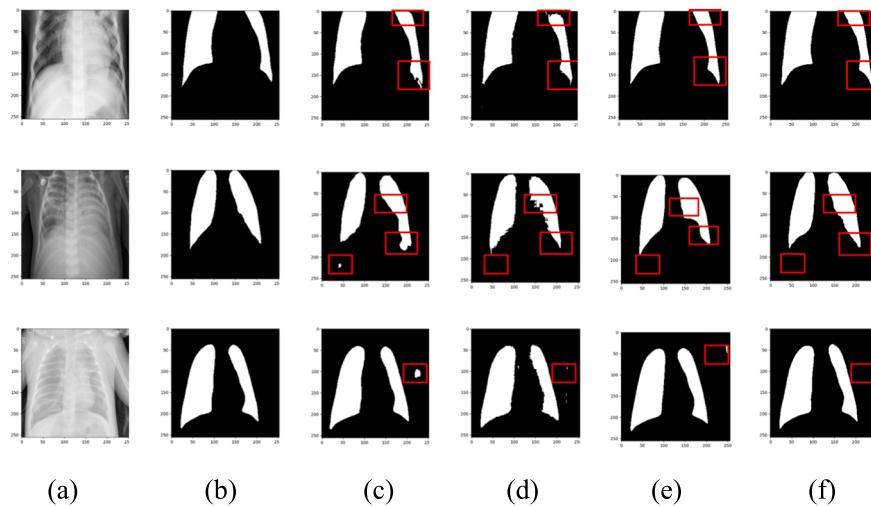


Fig. 7. Comparison of the segmentation: (a) Original image; (b) ground truth; (c) U-Net; (d) FCN-8S; (e) SegNet; (f) IU-Net.

Table 3
Results of ablation experiments.

Baseline	Model	A	lite-A	B	C	Acc	Sen	Spe	Miou	Dsc	Para
U-Net	Model1	✗	✗	✗	✗	98.65	94.41	99.20	96.33	97.12	1.9 M
	Model2	✓	✓	✗	✗	98.75	94.77	99.34	96.58	97.31	3.2 M
	Model3	✓	✓	✓	✗	98.99	95.73	99.55	97.21	97.82	2.8 M
	Model4	✓	✓	✓	✓	99.06	96.08	99.41	97.43	98.00	2.3 M

*A, lite-A, B, C represent Module A, Module lite-A, Module B and Module C, respectively;

Table 4
Results of Binary classification by using D2.

RCA module	Network	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	Kappa (%)
Without	MobileNetV2	95.21	95.20	95.21	95.19	87.99
	ResNet50	94.51	94.49	94.51	94.47	86.32
	InceptionV3	92.96	92.95	92.96	92.95	82.22
	EfficientNet	96.82	96.82	96.82	96.80	92.02
	DenseNet201	95.71	95.70	95.71	95.69	89.20
With	MobileNetV2 + RCA	98.71	98.72	98.71	98.72	96.74
	ResNet50 + RCA	96.05	96.14	96.05	96.08	90.13
	InceptionV3 + RCA	98.22	98.22	98.22	98.22	95.50
	EfficientNet + RCA	98.73	98.73	98.73	98.73	96.80
	DenseNet201 + RCA	98.36	98.39	98.36	98.36	95.89

that the IU-Net can perform segmentation of pulmonary parenchyma more effectively.

3.3.2. Ablation experiments

This section presents ablation experiments for validating the capabilities and performance of every module of IU-Net. The results are shown in [Table 3](#). Model1 represents U-Net; Model2 represents Model1 + Module A+Module lite-A, i.e. in the first three layers of the encoding part ordinary convolution layers are replaced by Module A and Module lite-A is used in the fourth layer to replace the ordinary convolution layer; Model3 represents Model2 + Module B, i.e. the transposed convolution of the decoding part of Model2 is replaced by Module B; Model4 represents Model3 + Module C, i.e. the ordinary convolution of the decoding part of Model3 is replaced by Module C.

From [Table 3](#), we can see that Model4, i.e. IU-Net performs the best

Table 5
Results of Multi-classification by using D1.

RCA module	Network	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	Kappa (%)
Without	MobileNetV2	90.54	90.60	90.54	90.50	85.46
	ResNet50	86.13	86.40	86.13	86.11	78.74
	InceptionV3	88.14	88.13	88.14	88.09	81.75
	EfficientNet	90.42	90.40	90.42	90.37	85.26
	DenseNet201	90.43	90.68	90.43	90.44	85.37
With	MobileNetV2 + RCA	94.26	94.27	94.26	94.24	91.20
	ResNet50 + RCA	94.71	94.70	94.71	94.69	91.88
	InceptionV3 + RCA	95.76	95.76	95.76	95.75	93.51
	EfficientNet + RCA	95.05	95.05	95.05	95.04	92.42
	DenseNet201 + RCA	95.62	95.62	95.62	95.61	93.29

with accuracy, sensitivity, specificity, miou, and dsc are 99.06 %, 96.08 %, 99.41 %, 97.43 %, and 98.00 %, respectively. All models of Model2, Model3 and Model4 outperformed Model1, which indicates that each module of the IU-Net can improve the segmentation.

3.4. Step3: Classification experiment on RCA

Here, the comparative experiments are performed to testify the function of the RCA module being added after the classification network. In these experiments, we choose five pre-trained neural networks: MobileNetV2, ResNet50, InceptionV3, EfficientNet, and DenseNet201 for comparison, use a 5 fold cross-validation technique to verify each network's capability, and in order to avoid the problem of unbalanced data volumes in each category we use a weighted average of the experimental results ([Chowdhury et al., 2020](#)). In the experiment, initial

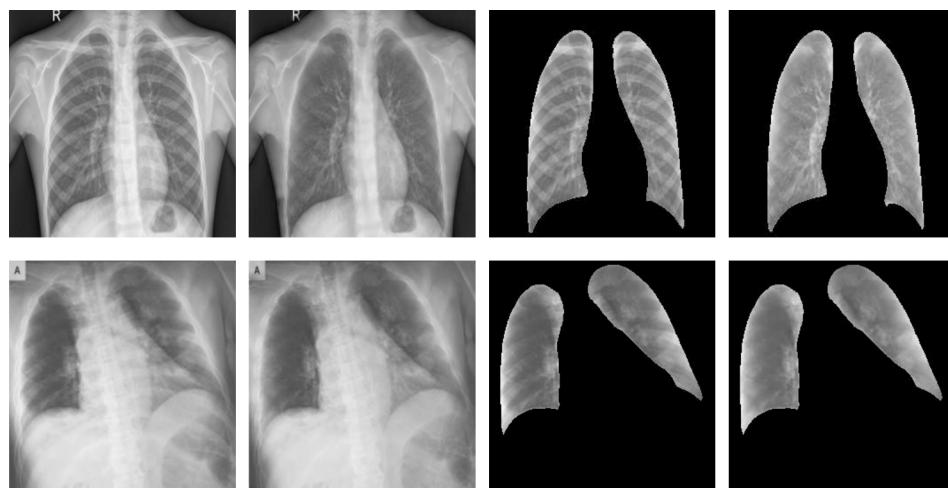


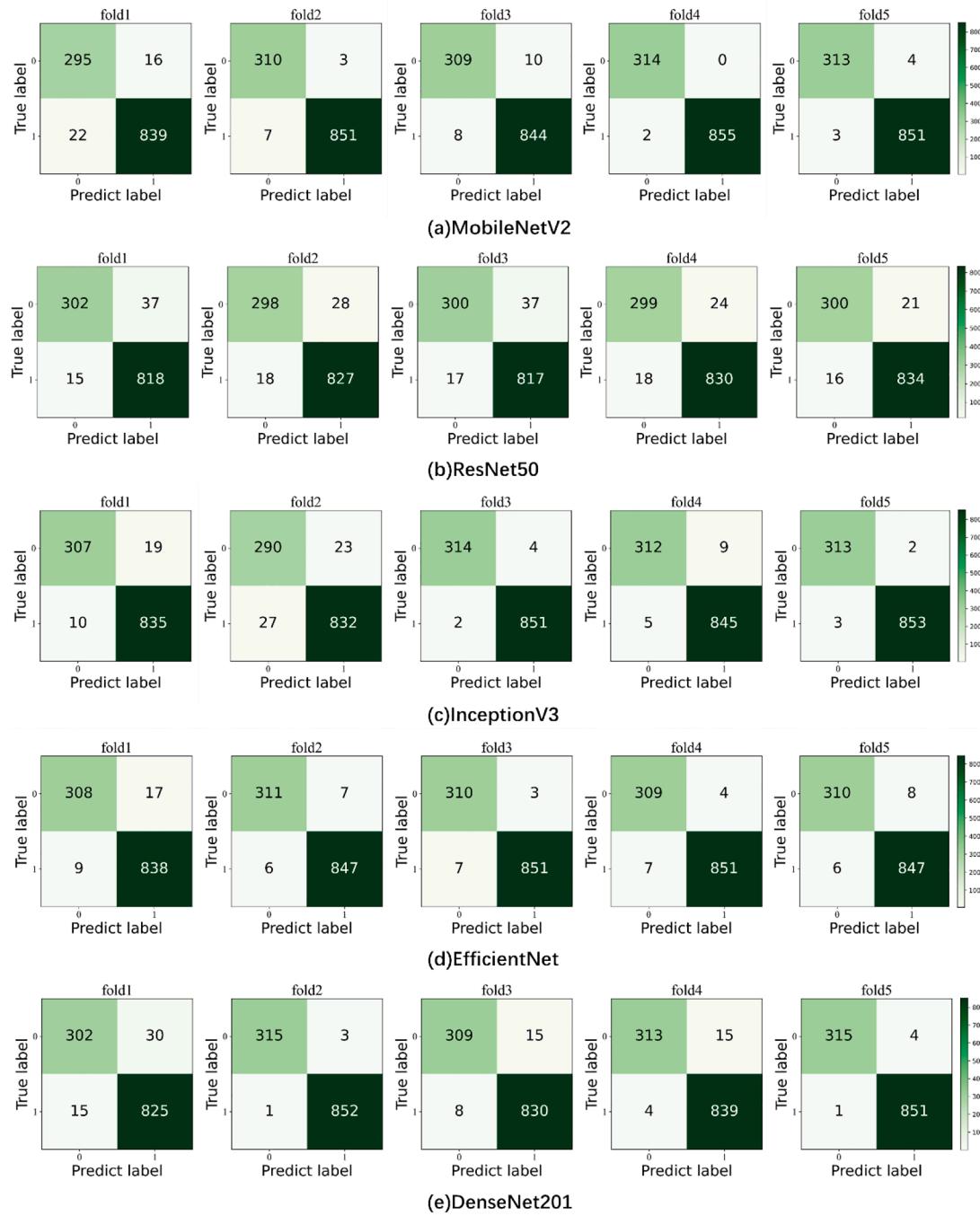
Fig. 8. Example images processed by different steps of the BSD framework: (a) original image; (b) the image processed by step1 solely; (c) the image processed by step2 solely; (d) the image processed by both step1 and step2.

Table 6
Binary classification results base on D2.

Classification Network	Steps of BSD	Step1	Step2	Step3	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	Kappa (%)
MobileNetV2	×	×	✓		92.95	92.92	92.95	92.89	81.87
	✓	×	✓		96.72	96.75	96.72	96.73	91.73
	×	✓	✓		97.93	97.93	97.93	97.93	94.76
	✓	✓	✓		98.71	98.72	98.71	98.72	96.74
ResNet50	×	×	✓		93.06	93.42	93.06	93.17	81.93
	✓	×	✓		95.01	94.99	95.01	94.96	87.10
	×	✓	✓		95.88	95.87	95.88	95.87	89.60
	✓	✓	✓		96.05	96.14	96.05	96.08	90.13
InceptionV3	×	×	✓		92.19	92.25	92.19	92.22	80.34
	✓	×	✓		95.47	95.48	95.47	95.46	88.55
	×	✓	✓		93.50	93.48	93.50	93.48	83.65
	✓	✓	✓		98.22	98.22	98.22	98.22	95.50
EfficientNet	×	×	✓		96.37	96.58	96.37	96.42	90.58
	✓	×	✓		97.83	97.92	97.83	97.83	94.54
	×	✓	✓		98.24	98.24	98.24	98.23	95.56
	✓	✓	✓		98.73	98.73	98.73	98.73	96.80
DenseNet201	×	×	✓		94.89	94.92	94.89	94.90	86.98
	✓	×	✓		97.38	97.39	97.38	97.38	93.37
	×	✓	✓		95.42	95.41	95.42	95.41	88.45
	✓	✓	✓		98.36	98.39	98.36	98.36	95.89

Table 7
Multi-classification result base on D1.

Classification Network	Steps of BSD	Step1	Step2	Step3	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	Kappa (%)
MobileNetV2	×	×	✓		91.40	91.40	91.40	91.32	86.74
	✓	×	✓		93.72	93.72	93.72	93.72	90.40
	×	✓	✓		92.85	92.96	92.85	92.84	89.05
	✓	✓	✓		94.26	94.27	94.26	94.24	91.2
ResNet50	×	×	✓		79.95	80.10	79.95	79.98	69.48
	✓	×	✓		81.61	81.68	81.61	81.62	71.96
	×	✓	✓		80.18	79.81	80.18	79.91	69.38
	✓	✓	✓		94.71	94.70	94.71	94.69	91.88
InceptionV3	×	×	✓		84.45	84.46	84.45	84.45	76.22
	✓	×	✓		85.18	85.19	85.18	85.17	77.29
	×	✓	✓		94.82	94.82	94.82	94.81	92.07
	✓	✓	✓		95.76	95.76	95.76	95.75	93.51
EfficientNet	×	×	✓		92.26	92.45	92.26	92.17	88.05
	✓	×	✓		94.85	94.87	94.85	94.84	92.09
	×	✓	✓		94.70	94.74	94.70	94.70	91.89
	✓	✓	✓		95.05	95.05	95.05	95.04	92.42
DenseNet201	×	×	✓		90.64	90.63	90.64	90.63	85.70
	✓	×	✓		93.82	93.81	93.82	93.82	90.57
	×	✓	✓		92.43	92.42	92.43	92.42	88.40
	✓	✓	✓		95.62	95.62	95.62	95.61	93.29



*In these maps: 0 for Normal, 1 for Pneumonia.

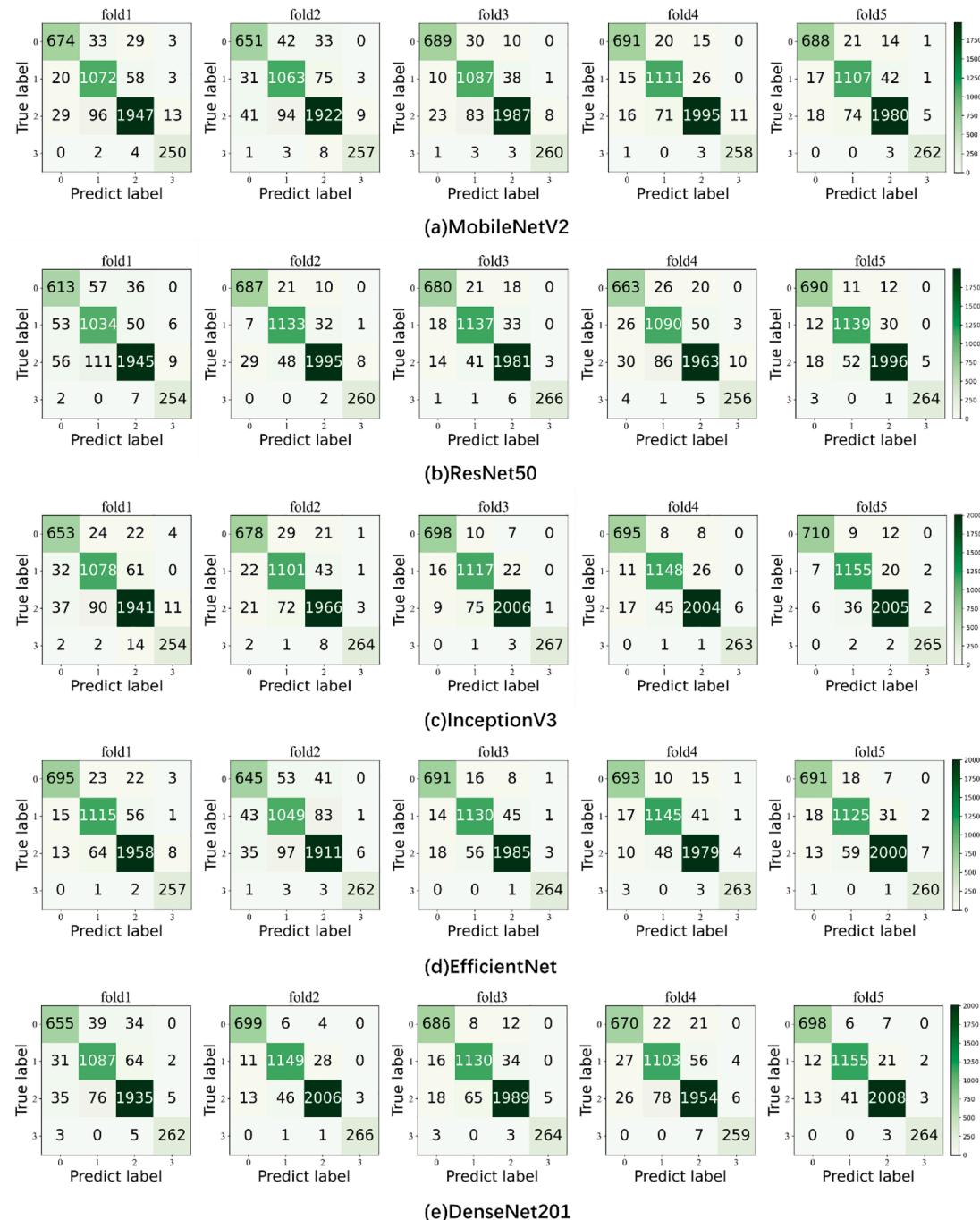
Fig. 9. Confusion matrixes of each network in the binary classification task. *In these maps: 0 for Normal, 1 for Pneumonia.

learning rate is set to 0.005, epoch is 50, batch size is 8, and the optimizer is SGD.

For verifying the capability of our RCA module in the classification network, comparative experiments are carried out. First, the datasets of D1 and D2 are fed into BSD framework, after step1 and step2 of the framework, the pulmonary parenchyma images without bone structures are obtained. Second, based on the obtained images, the classification network with or without RCA module are performed. Table 4 illustrate the binary classification experiments' results base on D2, and Table 5 lists the results of the multi-classification experiments base on D1.

Table 4 lists the performance of each network in the binary classification task. Among all the networks without RCA module, the best one

is EfficientNet, with accuracy, precision, recall, F1 value, and Kappa of 96.82 %, 96.82 %, 96.82 %, 96.80 %, and 92.02 %, respectively. By comparing them with the networks with RCA module, we can see that the performance of each network has improved, with MobileNetV2, ResNet50, InceptionV3, EfficientNet, and DenseNet201 increasing their accuracy by 3.5 %, 1.54 %, 5.26 %, 1.91 %, and 2.65 %, respectively. And we can see that among all the networks, EfficientNet added with RCA module performs best with accuracy, precision, recall, F1 value, and Kappa of 98.73 %, 98.73 %, 98.73 %, 98.73 %, 98.73 %, and 96.80 %, respectively, which means that EfficientNet added with RCA module is more suitable for the binary classification task of pulmonary disease based on chest X-ray image.



*In these maps: 0 for COVID-19, 1 for Lung_Opacity, 2 for Normal, 3 for Viral Pneumonia.

Fig. 10. Confusion matrixes of each network in the multi-classification task. *In these maps: 0 for COVID-19, 1 for Lung_Opacity, 2 for Normal, 3 for Viral Pneumonia.

Table 5 shows the performance of each network in the multi-classification tasks. Among all the networks without RCA module, MobileNetV2 performs better, with results for accuracy of 90.54 % and 85.46 % for Kappa, etc. By comparing them with the networks with RCA module, we can see that the performance of each network has improved, with MobileNetV2, ResNet50, InceptionV3, EfficientNet, and DenseNet201 increasing their accuracy by 3.72 %, 8.58 %, 7.62 %, 4.63 %, and 5.19 %. And we can see that among all the networks, InceptionV3 added with RCA module performs best with accuracy, precision, recall, F1 value, and Kappa of 95.76 %, 95.76 %, 95.76 %, 95.75 %, and 93.51

%, respectively, which means that InceptionV3 added with RCA module is more suitable for the multi-classification task of pulmonary disease based on chest X-ray image.

3.5. Experiment of BSD

To testify the efficiency of the whole framework, the ablation experiments and comparison experiments were carried out. In the framework: step 1 use ResNet-BS network; Step 2 use IU-Net; In step 3, the RCA module is added after all the classification networks used in the

Table 8

Comparison results with other literature base on binary classification.

Study	Method	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
Sharma et al. (2020)	CNN	90.68	—	—	—
Liang and Zheng (2020)	CNN	90.50	89.10	96.70	92.70
Ortiz-Toro et al. (2022)	Superpixels based histon	91.40	96.20	90.60	93.30
Manickam et al. (2021)	Modified resNet50	93.06	88.97	96.78	92.71
Victor Ikechukwu et al. (2021)	Iyke-Net	93.60	91.30	92.03	—
Dabre, K et al. (2024)	RAPID-Net	94.60	93.18	94.62	93.89
Bodapati and Rohith (2022)	ChxCapsNet	94.84	94.84	97.73	95.93
Kılıçarslan et al. (2023)	SupEx	95.37	—	—	—
Mittal et al. (2020)	E4CC	96.36	—	—	—
Balmuri et al. (2024)	ensemble transfer learning	97.10	97.10	97.10	97.10
Ours	BSD	98.73	98.73	98.73	98.73

Table 9

Comparison results with other literature base on multi-classification.

Study	Method	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
Azade and Anand (2022)	CNN	81.97	78.36	82.60	79.88
Ahmed Ali et al. (2022)	KNN	86.70	87.00	87.00	87.00
J. He. and Hu (2022)	CNN	89.07	90.62	89.68	90.11
Zhang, Z et al. (2023)	MCENet	89.70	92.00	90.00	91.00
Alablani, I. A et al. (2023)	COVID-ConvNet	90.93	—	—	—
Abd Elaziz et al. (2021)	MobileNetV3 + AO	92.40	86.60	92.40	89.40
Islam et al. (2023)	Xception	93.00	—	—	—
Alali and Ghatareh (2022)	GoogleNet	94.10	—	95.00	—
Kumara et al. (2023)	Xception	94.13	—	—	—
Boulila et al. (2022)	PPDL	94.20	—	—	—
Ours	BSD	95.76	95.76	95.76	95.75

experiment.

3.5.1. Ablation experiments

In this subsection, the ablation experiments are performed, and the example images of each processing step are shown in Fig. 8: Fig. 8(a) presents the original image; Fig. 8(b) presents the image processed by step 1 solely; Fig. 8(c) is the image processed by step 2 solely; and Fig. 8(d) is the image processed by both step 1 and step 2. The first line presents images from D2 and the second presents that of D1.

The ablation experiments are conducted on both binary classification based on D2 datasets and Multi-classification based on D1 datasets, and their results are shown in Table 6 and Table 7, respectively.

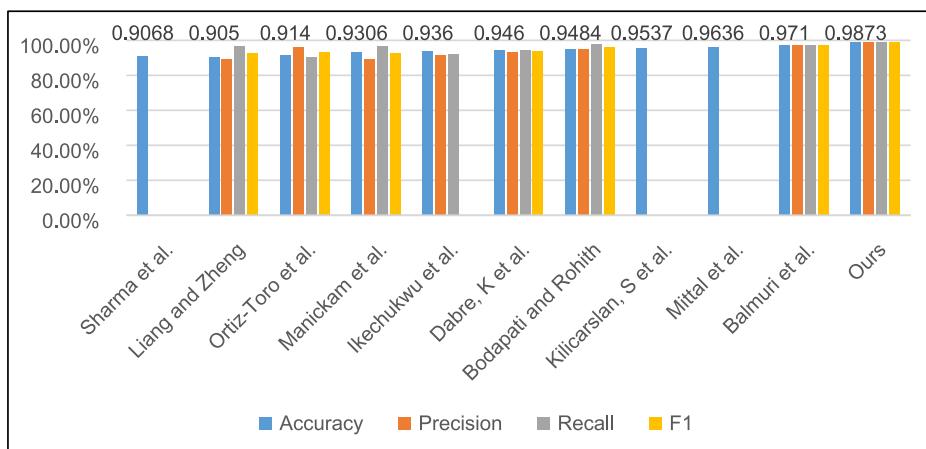
Through Table 6 we can draw conclusions as follows:(1) The classification using the image pre-processed by step 1 is more effective than that of only using the original image (i.e step3); (2) The classification using the image pre-processed by step 2 is more effective than that of only using the original image; (3) The classification using the image pre-processed both by step 1 and step 2, i.e BSD framework has the best classification effect. Furthermore, by comparing the results of all the networks, EfficientNet achieve the best result in binary classification task, with results for accuracy of 98.73 % and 96.80 % for Kappa.

Similarly, Table 7 shows the performance of the BSD framework in the multi-classification tasks. From the table, the same summary can be obtained that the results of both step 1 and step 2 are better than the original image. And our BSD framework performs the best with the most satisfying classification results. Furthermore, by comparing the results of all the networks, InceptionV3 based on the BSD framework achieves the best performance, such as its accuracy reaching 95.76 %.

In order to show the classification results of each category more clearly and the capability of the network more intuitively, we use confusion matrixes to display the results of binary classification (Fig. 9) and multi-classification (Fig. 10) with 5 fold cross-validation, and each fold of the network is presented on the confusion matrix to represent their effects.

Fig. 9 shows the performance of each network in the binary classification task. From the figure we can see that EfficientNet (Fig. 9(d)) performs best: The class 0 (Normal) in fold1, 308 images are categorized accurately, a higher percentage than the others. The class 1 (Pneumonia) in fold3, 851 images are categorized accurately, a higher percentage than the others.

Fig. 10 shows the performance of each network in the multi-classification task. From the figure we can see that InceptionV3 (Fig. 10(c)) performs best: The class 0 (COVID-19) in fold3, 698 images are categorized accurately, a higher percentage than the others. In the class 2 (Normal), 2006 images are categorized accurately, a higher

**Fig. 11.** Visualization of the comparison results based on the binary classification task.

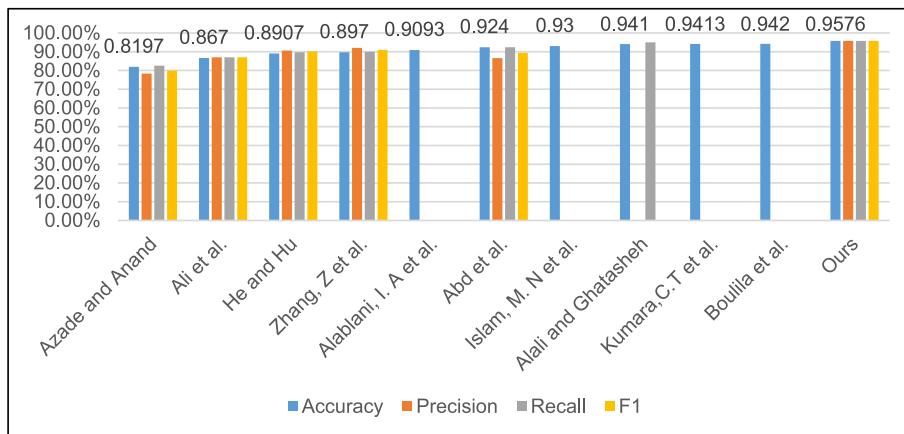


Fig. 12. Visualization of the comparison results based on the multi-classification task.

percentage than the others. In fold5, the number of images correctly classified in all classes is higher than others.

According to the above statements, both the experimental Table and confusion matrix verified the efficiency of the BSD framework.

3.5.2. Comparison experiments

In this section, for further verification of the diagnostic capability of BSD framework, it is compared to some existing literature. According to the ablation experiments, Efficient Net performs best in the binary classification task, therefore its results are used here to compare with other existing literature and the results are shown in Table 8. Meanwhile, InceptionV3 performs best in the multi-classification task, therefore its results are used to compare with other existing literature and Table 9 displayed the results.

Table 8 shows the performance of our BSD framework and other literature on binary classification tasks. The accuracy, precision, recall, and F1 value of our BSD framework reach 98.73 %, 98.73 %, 98.73 %, and 98.73 %, respectively. Our BSD framework obtains the best classification performance among all compared literature. Fig. 11 is the visualization of the comparison results on this task, from which the same conclusion as Table 8 can be drawn.

Table 9 shows the performance of other literature with our BSD framework on the multi-classification task, and our BSD framework achieves accuracy of 95.76 %, precision of 95.76 %, recall of 95.76 %, and F1 value of 95.75 %. Our BSD framework achieves the best results. Fig. 12 is the visualization of the comparison results on this task, from which the same conclusion as Table 9 can be drawn.

4. Discussion

In order to validate the effectiveness of our proposed BSD framework, experiments of each steps and the whole framework are carried out. Based on the above experimental results, the discussions are as follows:

- (1) The experiment of step1 verifies the efficiency of removing bone structures, based on which experiments of following steps are performed.
- (2) The experiment of step2 verifies the efficiency of the proposed segmentation network, IU-Net, for which comparison experiments and ablation experiments are performed. Table 2 and Fig. 7 is the results of comparison experiments. Table 2 shows that the proposed network outperforms others on all the listed metrics. And detailed discussion can be given based on Fig. 7: Firstly, in its first row, IU-Net and SegNet outperforms others with lower cracked and discontinuous in the segmented pulmonary. The reason is both of them used function modules deeper and wider in

the encoding part to obtain more image features, with IU-Net using Module A and Module lite-A, and SegNet using Vgg blocks. Secondly, in the second and third rows, IU-Net performed the best than all the compared networks with less oversegmentation. The reason is that in the decoding part, IU-Net using the combination of Module B and Module C to replace ordinary convolution layer. Among them, Module B can better fuse with Module A from the same layer with less parameters, and obtained feature maps with sufficient image features. Module C can better extract the useful information from these feature maps. Table 3 is the results of ablation experiments, which verifies that each designed module of IU-Net can improve its performance. Therefore, we can draw the same conclusions as above.

- (3) The experiment of step3 verifies the efficiency of the proposed RCA module, which is added after the classification network. The results for both binary classification (Table 4) and multi-classification (Table 5) illustrate that the classification networks being added with RCA module performed better than those without it. The reason is that it enhances the backbone networks through two branches, which can reduces the feature loss by fusing the global features and local features and extract useful features better.
- (4) The experiment of BSD verifies the efficiency of the whole framework, by performing ablation experiments and comparison experiments. In the ablation experiments, Table 6 and Table 7 show the performance of the BSD framework in binary classification and multi-classification tasks, respectively. Both tables demonstrate that each step of the framework i.e the process of bone suppression and pulmonary parenchyma extraction, improved the diagnosis process by delete the influence of bone structures and other tissues. Meanwhile, the confusion matrix of them are shown in Fig. 9 and Fig. 10, which further verified above conclusions. In the comparison experiments: Table 8 and Fig. 11 present the results of experiment on binary classification task, where the proposed BSD performed the best, owing to its multiple processing steps. For example, Mittal et al. (2020) and Dabre et al. (2024), they used a well-designed classification network to obtain better results, but they did not consider the effect of other tissues, thus is inferior to ours. Manickam et al. (2021) and Ortiz-Toro et al. (2022), they extracted the pulmonary parenchyma and then performed classification process without considering the effect of bone artifacts, and so is inferior to ours. Table 9 and Fig. 12 present the results of experiment on multi-classification task, based on which the same conclusions can be drawn.

5. Conclusion

In this article, a BSD framework for better diagnose pulmonary diseases that can address the influence of bone and other tissue is proposed, in which: First, three steps of bone suppression, pulmonary parenchyma extraction and pulmonary diseases diagnosis are included, which can eliminate the influence of bone structures and other tissues for better diagnose performance. Second, to settle the problems of the images burrs, blurring and complex pulmonary structures, we designed a new segmentation network IU-Net, which is adopted in step2 of BSD. Finally, to better extract the feature of the boneless pulmonary parenchyma images, an enhanced functional module RCA is proposed, which is used in step3 of BSD and improves the diagnostic performance.

Through a series of experiments, the efficiency of BSD framework is verified. In binary classification task, it obtained satisfied results with an accuracy of 98.73 % and a Kappa of 96.80 %, etc. In multi-classification task, it obtained the results of an accuracy of 95.76 % and a Kappa of 93.51 %, etc. These results demonstrate that it is more efficient in the diagnose of pulmonary disease than the state-of-the-art.

The BSD framework can obtain satisfactory results, but still has some limitations: Firstly, each step of the framework needs to be run independently, thus increasing the complexity of performance; Secondly, bone suppression can delete the impact of the bone artifacts, but also removes the information behind the bone structures. In the future, we will propose a model that can automatically perform all steps sequentially and design a recovery algorithm to recover the lost pulmonary parenchyma information.

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.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported in part by National Natural Science Foundation of China, General Project (No. 62266025) and National Natural Science Foundation of China (No. 22174057).

References

- Abd Elaziz, M., Dahou, A., Alsaleh, N. A., Elsheikh, A. H., Saba, A. I., & Ahmadein, M. (2021). Boosting COVID-19 image classification using MobileNetV3 and Aquila optimizer algorithm. *Entropy*, 23(11), 1383. <https://doi.org/10.3390/e23111383>
- Ahmed Ali, H., Hariri, W., Smaoui Zghal, N., & Ben Aissa, D. (2022). A comparison of machine learning methods for best accuracy COVID-19 diagnosis using chest X-ray images. 2022 IEEE 9th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications. doi:10.1109/setit54465.2022.9875477.
- Alali, A. M., & Ghatareh, M. S. (2022). Diagnosing chest x-rays for early detection Of COVID-19 and distinguishing it from other pneumonia using deep learning networks. *Webology*, 19(2). <https://www.webology.org/abstract.php?id=1513>.
- Alablani, I. A., & Alenazi, M. J. (2023). COVID-convnet: A Convolutional neural network classifier for diagnosing COVID-19 infection. *Diagnostics*, 13(10), 1675. <https://doi.org/10.3390/diagnostics13101675>
- Ayan, E., & Unver, H. M. (2019). Diagnosis of pneumonia from chest X-ray images using deep learning. 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT). doi:10.1109/ebbt.2019.8741582.
- Azade, A., & Anand, K. M. (2022). Impact of image augmentation in COVID-19 detection using chest X-ray images. 2022 IEEE Delhi Section Conference (DELCON). doi:10.1109/delcon54057.2022.9752785.
- Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A deep Convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12), 2481–2495. <https://doi.org/10.1109/tipami.2016.2644615>
- Bhandary, A., Prabhu, G. A., Rajinikanth, V., Thanaraj, K. P., Satapathy, S. C., Robbins, D. E., & Raja, N. S. (2020). Deep-learning framework to detect lung abnormality – A study with chest X-ray and lung CT scan images. *Pattern Recognition Letters*, 129, 271–278. <https://doi.org/10.1016/j.patrec.2019.11.013>
- Bodapati, J. D., & Rohith, V. (2022). ChxCapsNet: Deep capsule network with transfer learning for evaluating pneumonia in paediatric chest radiographs. *Measurement*, 188, Article 110491. <https://doi.org/10.1016/j.measurement.2021.110491>
- Boulila, W., Ammar, A., Benjdira, B., & Koubaa, A. (2022). Securing the classification of COVID-19 in chest X-ray images: A privacy-preserving deep learning approach. 2022 2nd International Conference of Smart Systems and Emerging Technologies. doi:10.1109/smartech54121.2022.000055.
- Balmuri, K. R., Konda, S., Mamidala, K. K., Gunda, M., Rani, B., & S. (2024). Automated and reliable detection of multi-diseases on chest X-ray images using optimized ensemble transfer learning. *Expert Systems with Applications*, 246, Article 122810. <https://doi.org/10.1016/j.eswa.2023.122810>
- Chowdhury, N. K., Rahman, M. M., & Kabir, M. A. (2020). PDCOVIDNet: A parallel-dilated convolutional neural network architecture for detecting COVID-19 from chest X-ray images. *Health Information Science and Systems*, 8, 27. <https://doi.org/10.1007/s13755-020-00119-3>
- Dabre, K., Varma, S. L., & Patil, P. B. (2024). RAPID-net: Reduced architecture for pneumonia in infants detection using deep convolutional framework using chest radiograph. *Biomedical Signal Processing and Control*, 87, Article 105375. <https://doi.org/10.1016/j.bspc.2023.105375>
- GM, H., Gourisaria, M. K., Rautaray, S. S., & Pandey, M. (2021). Pneumonia detection using CNN through chest X-ray. *Journal of Engineering Science and Technology (JESTEC)*, 16(1), 861–876.
- Han, T., Nunes, V. X., De Freitas Souza, L. F., Marques, A. G., Silva, I. C., Junior, M. A., & Filho, P. P. (2020). Internet of Medical Things—Based on deep learning techniques for segmentation of lung and stroke regions in CT scans. *IEEE Access*, 8, 71117–71135. <https://doi.org/10.1109/access.2020.2987932>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition. doi:10.1109/cvpr.2016.90.
- Hou, Q., Zhou, D., & Feng, J. (2021). Coordinate attention for efficient mobile network design. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. doi:10.1109/cvpr46437.2021.01350.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected Convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition. doi:10.1109/cvpr.2017.243.
- Hussain, M., Bird, J. J., & Faria, D. R. (2018). A Study on CNN Transfer Learning for Image Classification. In *UKCI 2018. Advances in Intelligent Systems and Computing*, 840. Springer, Cham. doi:10.1007/978-3-319-97982-3_16.
- Islam, M. N., Alam, M. G., Apon, T. S., Uddin, M. Z., Allheeb, N., Menshawi, A., & Hassan, M. M. (2023). Interpretable differential diagnosis of Non-COVID viral pneumonia, lung opacity and COVID-19 using tuned transfer learning and explainable AI. *Healthcare*, 11(3), 410. <https://doi.org/10.3390/healthcare11030410>
- J. He, A., & Hu, H. (2022). Identifying COVID-19 chest X-rays by image-based deep learning. 2022 7th International Conference on Machine Learning Technologies (ICMLT). doi:10.1145/3529399.3529418.
- Kermany, D., Zhang, K., & Goldbaum, M. (2018). Labeled optical coherence tomography (oct) and chest x-ray images for classification. *Mendeley data*, 2(2), 651. <https://doi.org/10.17632/rscbjbr9sj.2>
- Katsuragawa, S., Doi, K., & MacMahon, H. (1989). Image feature analysis and computer-aided diagnosis in digital radiography: Classification of normal and abnormal lungs with interstitial disease in chest images. *Medical Physics*, 16(1), 38–44. <https://doi.org/10.1118/1.596412>
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). Imagenet classification with deep convolutional neural networks. *Adv. Neural Informat. Process. Syst.*, 25(2). <https://doi.org/10.1145/3065386>
- Kundu, R., Das, R., Geem, Z. W., Han, G., & Sarkar, R. (2021). Pneumonia detection in chest X-ray images using an ensemble of deep learning models. *PLoS One*, 16(9), e0256630.
- Kumara, C. T., Pushpakumari, S. C., Udyhani, A. J., Aashiq, M., Rajendran, H., & Kumara, C. W. (2023). Image enhancement CNN approach to COVID-19 detection using chest X-ray images. 2023 IEEE 5th Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability. doi:10.3390/engproc2023055045.
- Kılıçarslan, S., Közkurt, C., Baş, S., & Elen, A. (2023). Detection and classification of pneumonia using novel Superior Exponential (SupEx) activation function in convolutional neural networks. *Expert Systems with Applications*, 217, Article 119503. <https://doi.org/10.1016/j.eswa.2023.119503>
- Li, Z., Yu, J., Li, X., Li, Y., Dai, W., Shen, L., ... Pu, Z. (2019). PNet: An efficient network for pneumonia detection. 2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). doi:10.1109/cisp-bmei48845.2019.8965660.
- Liang, G., & Zheng, L. (2020). A transfer learning method with deep residual network for pediatric pneumonia diagnosis. *Computer Methods and Programs in Biomedicine*, 187, Article 104964. <https://doi.org/10.1016/j.cmpb.2019.06.023>
- Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition. doi:10.1109/cvpr.2015.7298965.
- Manickam, A., Jiang, J., Zhou, Y., Sagar, A., Soundrapandian, R., & Dinesh Jackson Samuel, R. (2021). Automated pneumonia detection on chest X-ray images: A deep learning approach with different optimizers and transfer learning architectures. *Measurement*, 184, Article 109953. <https://doi.org/10.1016/j.measurement.2021.109953>

- Mittal, A., Kumar, D., Mittal, M., Saba, T., Abunadi, I., Rehman, A., & Roy, S. (2020). Detecting pneumonia using convolutions and dynamic capsule routing for chest X-ray images. *Sensors*, 20(4), 1068. <https://doi.org/10.3390/s20041068>
- Ortiz-Toro, C., García-Pedrero, A., Lillo-Saavedra, M., & Gonzalo-Martín, C. (2022). Automatic detection of pneumonia in chest X-ray images using textural features. *Computers in Biology and Medicine*, 145, Article 105466. <https://doi.org/10.1016/j.combiomed.2022.105466>
- Parveen, N. R., & Sathik, M. M. (2011). Detection of pneumonia in chest X-ray images. *Journal of X-Ray Science and Technology*, 19(4), 423–428. <https://doi.org/10.3233/xst-2011-0304>
- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... & Ng, A. Y. (2017). CheXnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225. doi:10.48550/arXiv.1711.05225.
- Racic, L., Popovic, T., Cakic, S., & Sandi, S. (2021). Pneumonia detection using deep learning based on Convolutional neural network. 2021 25th International Conference on Information Technology (IT). doi:10.1109/it51528.2021.9390137.
- Rahman, T., Chowdhury, M. E., Khandakar, A., Islam, K. R., Islam, K. F., Mahbub, Z. B., & Kashem, S. (2020). Transfer learning with deep Convolutional neural network (CNN) for pneumonia detection using chest X-ray. *Applied Sciences*, 10(9), 3233. <https://doi.org/10.3390/app10093233>
- Rahman, T., Khandakar, A., Kadir, M. A., Islam, K. R., Islam, K. F., Mazhar, R., & Chowdhury, M. E. (2020). Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. *IEEE Access*, 8, 191586–191601. <https://doi.org/10.1109/access.2020.3031384>
- Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Abul Kashem, S. B., & Chowdhury, M. E. (2021). Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. *Computers in Biology and Medicine*, 132, Article 104319. <https://doi.org/10.1016/j.combiomed.2021.104319>
- Rajaraman, S., Zamzmi, G., Folio, L., Alderson, P., & Antani, S. (2021). Chest X-ray bone suppression for improving classification of tuberculosis-consistent findings. *Diagnostics*, 11(5), 840. <https://doi.org/10.3390/diagnostics11050840>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science*, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *ICLR Conference*. <https://doi.org/10.48550/arXiv.1409.1556>
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. doi:10.1109/cvpr.2018.00474.
- Sharma, A., & Mishra, P. K. (2022). Covid-Manet: Multi-task attention network for explainable diagnosis and severity assessment of COVID-19 from CXR images. *Pattern Recognition*, 131, Article 108826. <https://doi.org/10.1016/j.patcog.2022.108826>
- Sharma, H., Jain, J. S., Bansal, P., & Gupta, S. (2020). Feature extraction and classification of chest X-ray images using CNN to detect pneumonia. 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence). doi:10.1109/confluence47617.2020.9057809.
- Stefan, M. S., Eltanbedawi, A., Devoe, N. C., Khan, S., Zhou, Y., Lateef, T., & Lagu, T. C. (2022). Death among patients hospitalized with symptomatic COVID-19: Implications for high-risk patients. *Journal of Hospital Medicine*, 17(4), 252–258. <https://doi.org/10.1002/jhm.12805>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition*. doi:10.1109/cvpr.2015.7298594.
- Victor Ikechukwu, A., Murali, S., Deepu, R., & Shivamurthy, R. (2021). Resnet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of pneumonia from chest X-ray images. *Global Transitions Proceedings*, 2 (2), 375–381. <https://doi.org/10.1016/j.gltcp.2021.08.027>
- Tan, M., & Le, Q. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. Proceedings of the 36th International Conference on Machine Learning, in Proceedings of *Machine Learning Research* 97:6105–6114. Available from <https://proceedings.mlr.press/v97/tan19a.html>.
- Wang, X., Xu, M., Zhang, J., Jiang, L., & Li, L. (2021). Deep multi-task learning for diabetic retinopathy grading in fundus images. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(4), 2826–2834. <https://doi.org/10.1609/aaai.v35i4.16388>
- Yi, S., Yang, X., Wang, T., She, F., Xiong, X., & He, J. (2021). Diabetic retinopathy diagnosis based on RA-efficientnet. *Applied Sciences*, 11(22), 11035. <https://doi.org/10.3390/app112211035>
- Zhao, S., Li, Q., He, T., & Wen, J. (2022). A step-by-step gradient penalty with similarity calculation for text summary generation. *Neural Processing Letters*, 55(4), 4111–4126. <https://doi.org/10.1007/s11063-022-11031-0>
- Zhang, Z., Ma, Y., & Li, K. (2023). An improved X-ray image diagnosis method for COVID-19 pneumonia on a lightweight neural network embedded device. Proceedings of the 2023 3rd International Conference on Bioinformatics and Intelligent Computing. doi:10.1145/3592686.3592749.
- Zhao, S., Li, Q., Yang, Y., Wen, J., & Luo, W. (2023). From Softmax to Nucleusmax: A novel sparse language model for Chinese radiology report summarization. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 22(6), 1–21. <https://doi.org/10.1145/3596219>
- Zhao, S., Liang, Z., Wen, J., & Chen, J. (2023). Sparsening and smoothing for the seq2seq models. *IEEE Transactions on Artificial Intelligence*, 4(3), 464–472. <https://doi.org/10.1109/taii.2022.3207982>