

# Evaluation of Lung Involvement in COVID-19 Pneumonia Based on Ultrasound Images

**Zhaoyu Hu**

Fudan University - Handan Campus: Fudan University

**Zhenhua Liu**

Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

**Yijie Dong**

Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

**Jianjian Liu**

Shanghai Public Health Clinical Center

**Bin Huang**

xixi hospital of hangzhou

**Aihua Liu**

the sixth hospital of wuhan

**Jingjing Huang**

Shanghai Public Health Clinical Center

**Xujuan Pu**

Shanghai Public Health Clinical Center

**Xia Shi**

Shanghai Public Health Clinical Center

**Jinhua Yu**

Fudan University - Handan Campus: Fudan University

**Yang Xiao** (✉ [yang.xiao@siat.ac.cn](mailto:yang.xiao@siat.ac.cn))

Shenzhen Institutes of Advanced Technology Chinese Academy of Sciences

**Hui Zhang**

Shanghai Public Health Clinical Center

**Jianqiao Zhou**

Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

---

## Research

**Keywords:** ultrasound, lung involvement, classification, COVID-19, neural network

**Posted Date:** December 17th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-127171/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on March 20th, 2021. See the published version at <https://doi.org/10.1186/s12938-021-00863-x>.

1 **Evaluation of Lung Involvement in COVID-19 Pneumonia Based on Ultrasound Images**

2 Zhaoyu Hu <sup>1</sup>, Zhenhua Liu <sup>2</sup>, Yijie Dong <sup>2</sup>, Jianjian Liu <sup>3</sup>, Bin Huang <sup>4</sup>, Aihua Liu <sup>5</sup>, Jingjing Huang <sup>3</sup>,  
3 Xujuan Pu <sup>3</sup>, Xia Shi <sup>3</sup>, Jinhua Yu <sup>1</sup>, Yang Xiao <sup>6\*</sup>, Hui Zhang <sup>3\*</sup>and Jianqiao Zhou <sup>2\*</sup>

4  
5 \* Correspondence:

6 Yang Xiao

7 Email: yang.xiao@siat.ac.cn

8 Institute of Biomedical and Health Engineering Shenzhen Institutes of Advanced Technology,  
9 Chinese Academy of Sciences

10  
11 Hui Zhang

12 Email: zhang.hui@zs-hospital.sh.cn

13 Department of Ultrasound, Shanghai Public Health Clinical Center, Shanghai 201508, China

14  
15 Jianqiao Zhou

16 Email: zhousu30@126.com

17 Department of Ultrasound, Ruijin Hospital, Shanghai Jiaotong University School of Medicine,  
18 Shanghai 200025, China

19 Full list of author information is available at the end of the article

20

21

22

23

24

25

26

## 1 Abstract

2

3 **Background:** Lung ultrasound (LUS) can be an important imaging tool for the diagnosis and  
4 assessment of lung involvement. Ultrasound sonograms have been confirmed to illustrate damage to  
5 a person's lungs, which means that the correct classification and scoring of a patient's sonogram can  
6 be used to assess lung involvement.

7

8 **Methods:** The purpose of this study was to establish a lung involvement assessment model based on  
9 deep learning. A novel multimodal channel and receptive field attention network combined with  
10 ResNeXt (MCRFNet) was proposed to classify sonograms, and the network can automatically fuse  
11 shallow features and determine the importance of different channels and respective fields. Finally,  
12 sonogram classes were transformed into scores to evaluate lung involvement from the initial  
13 diagnosis to rehabilitation.

14

15 **Results and conclusion:** Using multicenter and multimodal ultrasound data from 104 patients, the  
16 diagnostic model achieved 94.39% accuracy, 82.28% precision, 76.27% sensitivity, and 96.44%  
17 specificity. The lung involvement severity and the trend of COVID-19 pneumonia were evaluated  
18 quantitatively.

19

20 **Keywords:** ultrasound; lung involvement; classification; COVID-19; neural network

21

22

23

24

25

26

## 1 **1. Background**

2 As of December 3, 2020, 64 469 705 confirmed cases of novel coronavirus 2019 (COVID-19)  
3 disease and 1 492 029 COVID-19-related deaths had been reported worldwide. COVID-19 has been  
4 reported in 212 countries [1]. The World Health Organization classified COVID-19 as a global  
5 pandemic [2]. Due to the increasing number of cases, it is necessary to rapidly diagnose patients and  
6 assess the severity of the disease.

7 The current gold standard for the diagnosis of COVID-19 is reverse transcription-polymerase  
8 chain reaction (RT-PCR) analysis of respiratory specimens. However, due to inaccurate methods  
9 when collecting samples via nasopharyngeal swabs, the false negative rate is high [3]. Delays in the  
10 diagnosis of COVID-19 will cause the spread of the disease and the aggravation of a patient's  
11 condition. Computed tomography (CT) is the main method for diagnosing and evaluating the severity  
12 of COVID-19 pneumonia [4, 5]. However, CT also has the following problems in lung-related  
13 diagnosis. First, CT-based diagnosis is costly and involves radiation [6]. In unstable and critically ill  
14 patients, CT is not easy to perform. In addition, patients who are sensitive to radiation, such as  
15 pregnant women, need to avoid the radiation caused by CT. Second, it is of great clinical significance  
16 to determine whether there is pulmonary airway obstruction in patients with COVID-19 pneumonia.  
17 CT can only obtain static images and cannot evaluate the movement of gas in the bronchi and  
18 bronchiole in real time.

19 As a nonradiation medical imaging method, ultrasound is highly sensitive to the diagnosis of  
20 various lung diseases [7]. Studies have shown that lung ultrasonography (LUS) can be an important  
21 imaging tool for diagnosing pneumonia and assessing the degree of lung involvement [8]. For  
22 example, Liu et al. proved the effectiveness of bedside LUS in the diagnosis of community-acquired  
23 pneumonia. With CT as the gold standard, the diagnosis of community-acquired pneumonia by LUS  
24 has reached 96.1% accuracy, and the diagnostic efficiency of LUS far exceeds that of chest X-ray [9].  
25 Bouhemad et al. [10] and Silvia et al. [11] reported that LUS has the potential to be a key tool for early  
26 ventilator-associated pneumonia (VAP) diagnosis. Within the technical period required for  
27 bronchoalveolar lavage analysis, LUS is an ideal decision-making tool for antimicrobial  
28 management. LUS has been used as an effective imaging method for diagnosing pneumonia in many

1 institutions. The advantages of LUS are that it is inexpensive, does not involve radiation, is easy to  
2 obtain, and can be checked at the bedside, which is especially useful for patients with severe  
3 pneumonia [12, 13].

4 An increasing number of researchers have used deep learning to complete COVID-related  
5 diagnosis, segmentation, and classification tasks. Mpesiana et al. [14] completed automatic diagnosis  
6 through X-ray images of COVID-19 patients, and the transfer learning based on ResNet50 used by  
7 the authors reached 98% accuracy compared with three other models. Chawki et al. [15] utilized  
8 CNN, DenseNet201, and several other popular architectures based on X-ray to classify patients as  
9 having a disease or being disease-free. The author also achieved a high accuracy rate of more than  
10 96%. To segment the regions of interest (ROIs), Zheng et al. [16-19] used U-Net, improved U-Net  
11 such as U-Net++, VB-Net, and the attention mechanism, which is an efficient and explainable method  
12 in COVID-19 applications.

13 Generally, current studies are more focused on the diagnosis and segmentation tasks of images.  
14 Few researchers examine the impact of COVID-19 on internal organ damage among patients,  
15 especially based on ultrasound images; internal organ damage is equally important for understanding  
16 COVID-19. The current challenge for COVID-19 is not the diagnosis but the severity and drug  
17 intake. In general, the limitations and challenges of the current research are as follows: 1) method  
18 validation of multicenter data; 2) sufficient utilization of ultrasound data; and 3) evaluation of the  
19 impact of COVID-19 on internal organs and bodily functions.

20 To address the issues presented above, we proposed a novel multimodal channel and receptive  
21 field attention (MCRF) block for assessing lung damage in COVID-19 patients. Ultrasound  
22 sonograms have been confirmed to illustrate damage to a person's lungs [12], and thus, accurate  
23 identification of sonograms is a prerequisite for evaluation. To extract the shallow features of the  
24 sonograms, K-means clustering and gradient field were used as two new modalities and combined  
25 with the original input by a fusion module. For the extraction of deep features, we used ResNeXt as  
26 the backbone of the overall network. Moreover, a novel channel and receptive field attention block  
27 was designed to focus on the correct modal channel and patch. Finally, each frame of the patient's

1 ultrasound video was classified and scored, and its correlation with a pathological indicator was  
 2 calculated. In summary, the contributions of this paper are as follows:

3 1) We assessed lung damage in COVID-19 patients through more common ultrasound images  
 4 and verified our method on three machines from four centers.

5 2) A brand new fusion network was designed to produce the multimodal input. The traditional  
 6 method and ResNeXt were used to extract the shallow and deep features of the ultrasound image,  
 7 respectively, so we fully combined the advantages of traditional methods and deep learning.

8 3) The attention mechanism we proposed makes the network focused on the correct modal  
 9 channels and regions, and Grad-Cam++ heatmaps were introduced to enhance the interpretability of  
 10 the network.

11

## 12 **2. Results**

### 13 *2.1. MCRFNet Classification Accuracy*

14 We employed four widely used metrics to quantitatively evaluate the COVID-19 lung sonogram  
 15 classification performances, including accuracy (ACC), precision (PRE), sensitivity (SEN), and  
 16 specificity (SPE). In general, a better classification performance will have higher ACC, PRE, SEN,  
 17 and SPE. ACC describes the proportion of correctly classified images, which is expressed as follows:

$$18 \quad ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

19 where TP, TN, FP, and FN represent the number of true positive predictions, true negative predictions, false  
 20 positive predictions, and false negative predictions, respectively. PRE is useful for measuring the proportion of  
 21 true positive predictions of overall positive images and is defined as:

$$22 \quad PRE = \frac{TP}{TP + FP} \quad (2)$$

23 SEN, which is a measure of the number of true positive predictions and false negative predictions, is defined as:

$$24 \quad SEN = \frac{TP}{TP + FN} \quad (3)$$

25 Finally, SPE considers the number of true negative predictions and false positive predictions, which is defined  
 26 as:

$$SPE = \frac{TN}{TN + FP} \quad (4)$$

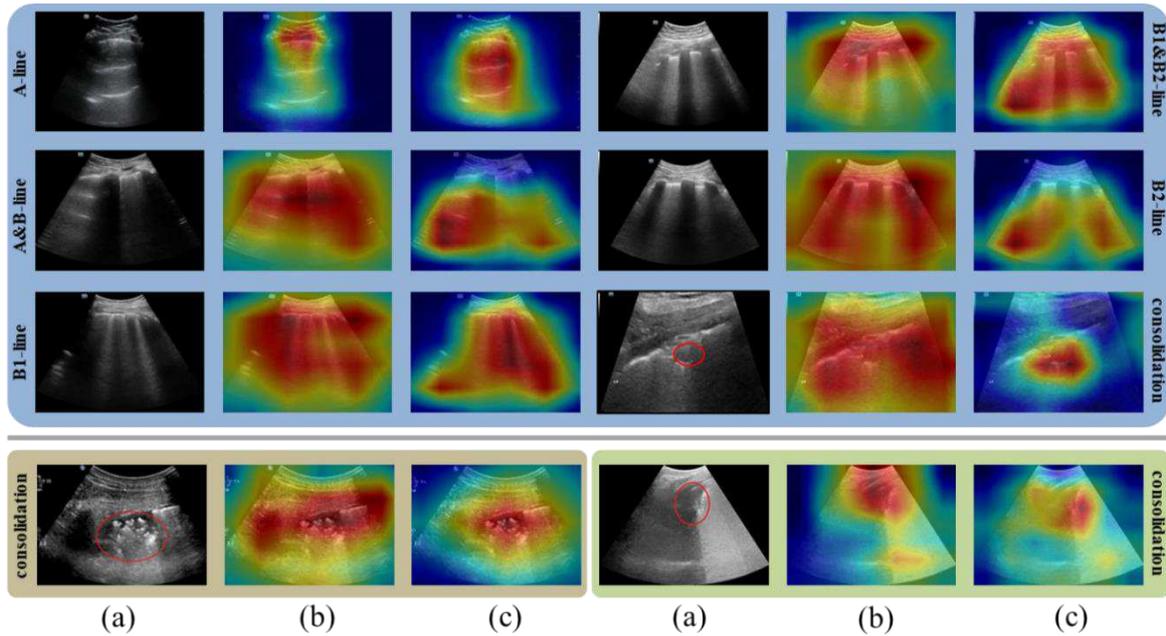
Table 1 summarizes the method comparison and ablation experiment. We evaluated the classic classification of deep neural networks and several attention-oriented architectures. It can be observed that our proposed MCRFNet has noticeably higher performance than other models, achieving ACC, PRE, SEN, and SPE values of 94.39%, 82.28%, 76.27% and 96.44% on the three datasets, respectively. Compared with classic models, our model is specially designed for COVID-19 lung sonogram images, and attention-guided architecture has advantages in ultrasound images. Attention-oriented models show better performance than classic models; however, no previous models have fused multimodel information or considered attention combining, which impacted their classification performance.

**Table 1.** Method comparison and ablation experiment on different datasets.

Method Comparison												
Method	S (%)				S&M (%)				S&M&P (%)			
	ACC	PRE	SEN	SPE	ACC	PRE	SEN	SPE	ACC	PRE	SEN	SPE
VGG[20]	93.15	73.07	74.62	96.09	89.87	76.36	73.05	94.06	88.81	73.5	67.37	93.19
ResNet[21]	93.19	73.56	75.55	96.06	90.63	77.77	74.94	94.51	89.41	74.56	68.78	93.56
ResNeXt (baseline)[22]	93.15	77.82	74.62	96.31	91.7	79.98	77.57	94.94	90.46	75.92	70.9	94.04
SENet-50[23]	92.5	72.16	73.33	95.65	93.64	83.16	<b>83.74</b>	96.28	92.37	79.71	75.93	95.33
SKNet-50[24]	92.8	<b>86.32</b>	76.97	95.79	92.12	80.45	78.63	95.4	91.03	77.21	72.55	94.55
Ablation Experiment												
w/o. CRFA & w/o. fusion module	93.15	77.82	74.62	96.31	91.7	79.98	77.57	94.94	90.46	75.92	70.9	94.04
w/o. CRFA module	93.19	78.09	75.55	96.34	92.12	81.21	78.63	95.17	90.93	77.14	71.94	94.3
MCRFNet	<b>97.73</b>	85.72	<b>88.06</b>	<b>98.69</b>	<b>96.25</b>	<b>87.2</b>	83.59	<b>97.74</b>	<b>94.39</b>	<b>82.28</b>	<b>76.27</b>	<b>96.44</b>

S = Stork dataset accuracy, M = Mindray dataset accuracy, P = Philips dataset accuracy, ACC=accuracy, PRE=precision, SEN=sensitivity, SPE=specificity.

Table 1 shows the results of a comparison between models trained with or without the CRFA module and fusion module. Specifically, the model trained with the CRFA and fusion module yields a 0.028 improvement in ACC, a 0.0636 improvement in PRE, a 0.0537 improvement in SEN, and a 0.024 improvement in SPE, which greatly surpasses the baseline model.

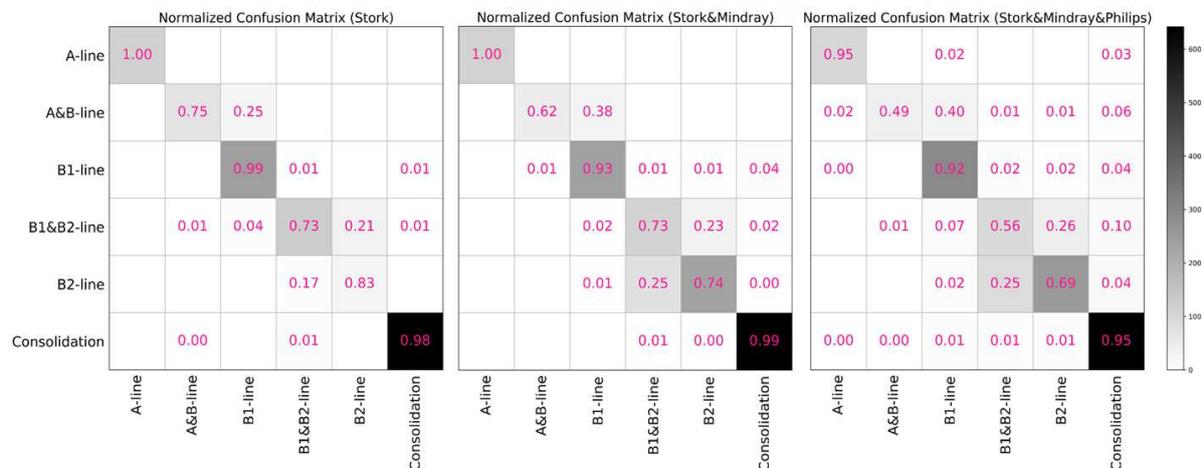


**Fig. 1.** The Grad-CAM++ visualization results. The blue, red, and green backgrounds represent the Mindray dataset, Stork dataset, and Philips dataset, respectively. (a) Original input (the actual data are not marked with a red circle), (b) ResNeXt, and (c) MCRFNet. The red circle is the lung consolidation area.

To intuitively understand the fusion module, channel, and receptive field attention capability of MCRFNet, we used the Grad-CAM++ method to visualize the class activation mapping of the backbone and our proposed network [25]. Grad-CAM++ is commonly used to locate discriminative feature regions for perception, which makes the model interpretable. As shown in Fig. 1, the areas with bright colors indicate that the current region contributes the most to the classification. The results show that the backbone network focuses on the recognition of the B-line in a wider area, and our proposed network can better allow the network to focus on the line-shaped area or lung consolidation area. The samples are well recognized by ResNeXt, but there is still room for improvement. However, our MCRFNet can adaptively select the appropriate modal channel and suitable convolution kernel size. Finally, our MCRFNet performs better than ResNeXt due to the influence of the MCRF module.

The normalized confusion matrixes of classification on three datasets by our proposed method are provided in Fig. 2. We observe that B2-line, B1&B2-line, and B1-line are partly misclassified as lower or higher level severity. Especially in the three datasets, misclassification is more obvious. The reasons for the performance degradation of multiple datasets will be discussed in the *Discussion* section. In short, the classification accuracy of the A-line and consolidation of the three datasets can reach nearly

1 100%. Misclassification generally occurs in the fusion classification of two categories: A&B-line and  
 2 B1&B2-line.

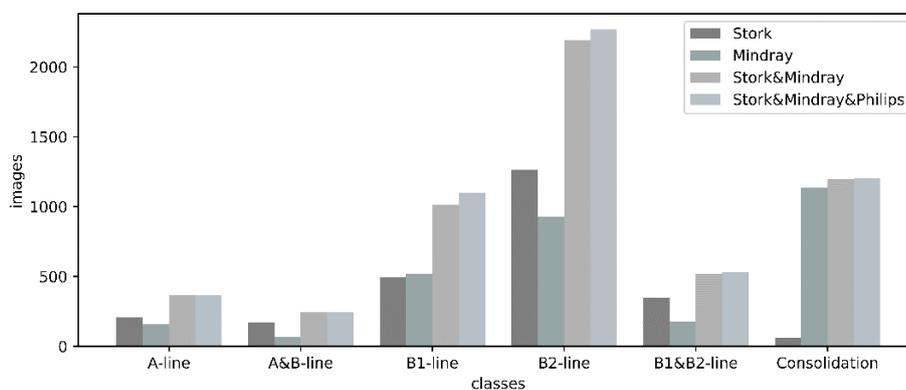


3  
 4 **Fig. 2.** The normalized confusion matrix of classification on three datasets by MCRFNet. The abscissa is the predicted label,  
 5 and the ordinate is the true label.

6

## 7 2.2. Distribution of datasets

8 Fig. 3 depicts the distribution of expert manual label classification of 5704 ultrasound images of  
 9 108 cases in three datasets. Specifically, the Stork, Mindray and Philips datasets have 2541 images of  
 10 43 cases, 2985 images of 51 cases, and 178 images of 14 cases, respectively, and we merged the first  
 11 two and the first three datasets as two new datasets (Stork & Mindray, Stork & Mindray & Philips)  
 12 to avoid imbalanced samples. These three datasets (Stork, Stork & Mindray, Stork & Mindray & Philips)  
 13 are divided into training and test sets at a ratio of 2:1 for each category. The number of cases in the  
 14 training set of the three datasets are: 29, 63, and 72, and the number of cases in the test set are 14, 31,  
 15 and 36.



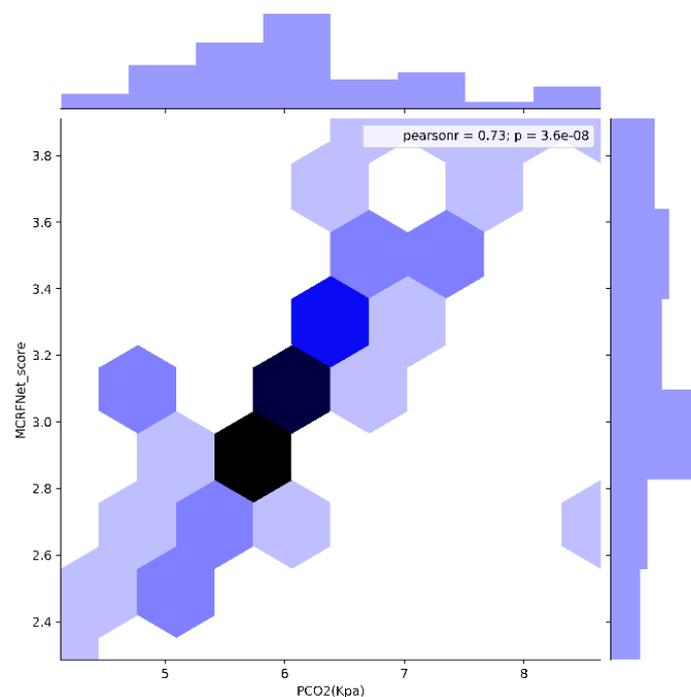
16

17

**Fig. 3** Distribution of expert manual labels of 5704 ultrasound images in three datasets.

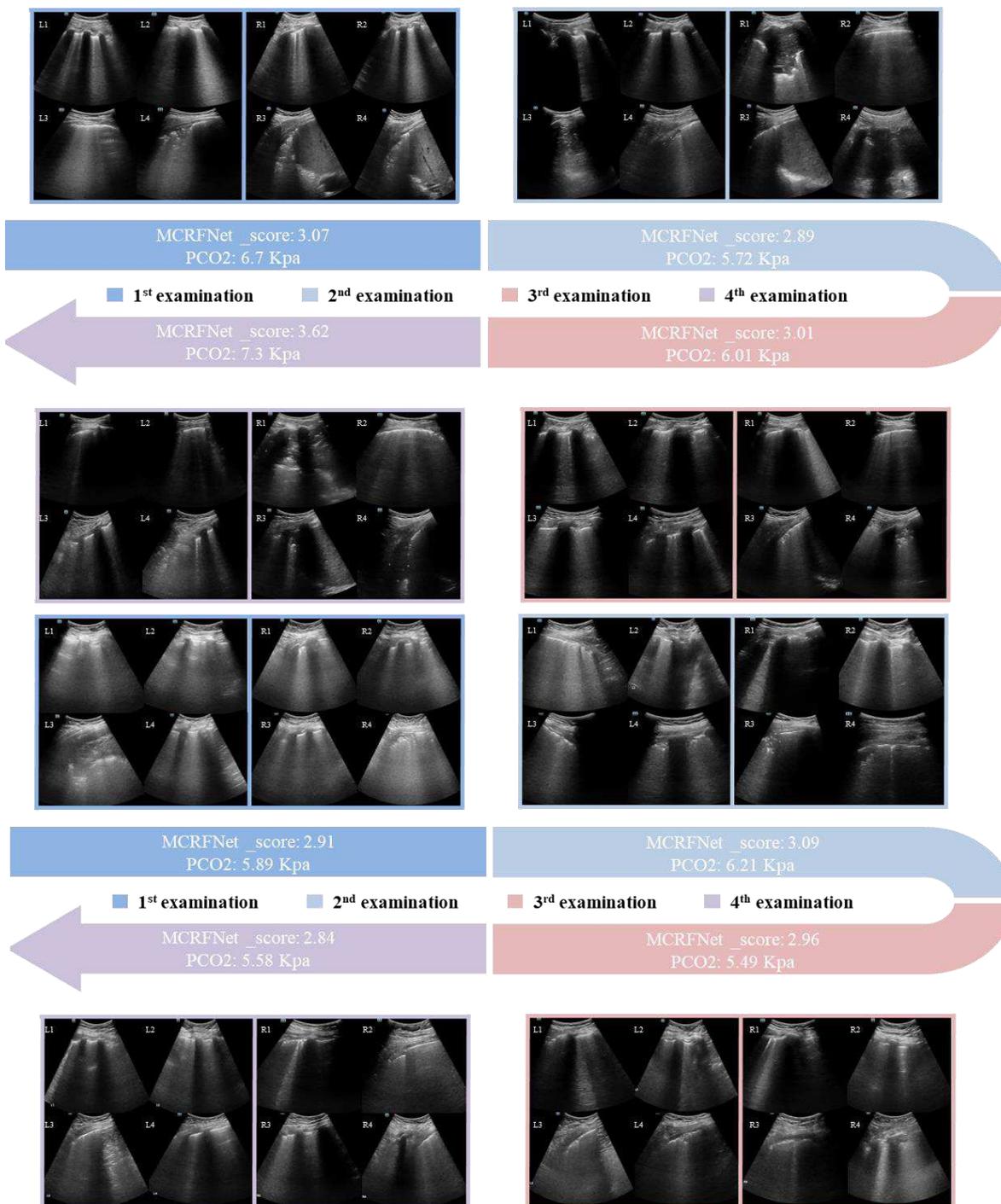
### 2.3. Evaluation of the trend in the degree of lung involvement

After obtaining the trained MCRFNet, we independently tested the videos of 8 patients (additional collected data), which were examined multiple times (4 times or more) from the initial diagnosis to rehabilitation, and we performed classification according to the method in the Establishment of Scoring Standards section. CO<sub>2</sub> partial pressure (PCO<sub>2</sub>) is a great indicator of respiratory function and is closely related to acid-base homeostasis, reflecting the amount of acid in the blood. The correlation between the score obtained and PCO<sub>2</sub> was analyzed by Pearson correlation analysis, and the correlation is shown in Fig. 4. The Pearson correlation coefficient was 0.73 ( $P < .001$ ). In the graph, darker colors indicate a higher frequency of occurrence. The graph shows that the score of MCRFNet is in the range of 2.7-3.4, which has a higher correlation with PCO<sub>2</sub>.



11  
12 **Fig. 4.** Scatter plot between the classification score of MCRFNet and CO<sub>2</sub> partial pressure (PCO<sub>2</sub>).

13 In addition, two patients with multiple examinations of the MCRFNet score and PCO<sub>2</sub> are shown  
14 in Fig. 5. We followed the three lines of the parasternal line (PSL), anterior axillary line (AAL), and  
15 posterior axillary line (PAL) in [26] to divide the left and right sides of the lungs into four areas (L1-L4  
16 and R1-R4). Only one picture is shown in the figure, but in the actual scoring, we averaged the scores of  
17 multiple pictures after framing to obtain the specific score in the figure.



1  
 2 **Fig. 5.** Two patients had multiple examinations of the MCRFNet score and PCO2 from the initial diagnosis to  
 3 rehabilitation.

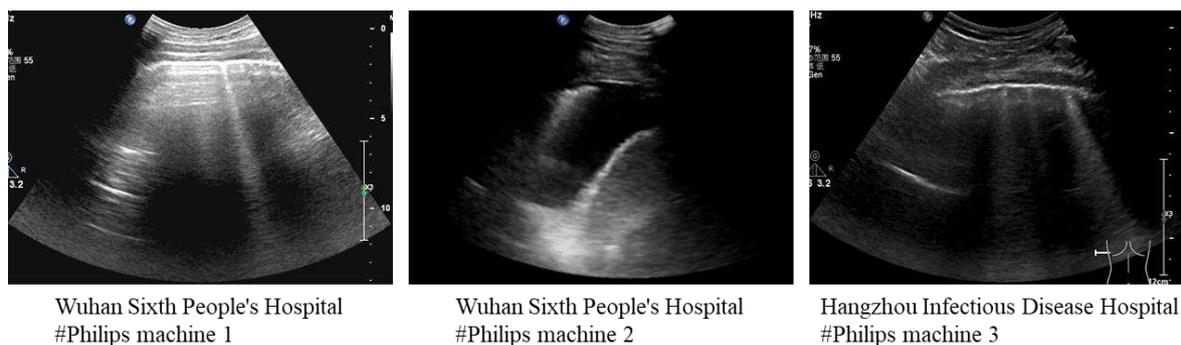
4 In short, our classification and scoring system not only reflects the degree of lung involvement of a  
 5 patient but also helps doctors combine this score with other indicators to evaluate the patient’s lung  
 6 disease and even the entire person’s condition. It is more beneficial to use this scoring system to  
 7 improve the nursing level of patients with pneumonia and enhance their support for the clinical  
 8 decision-making process for the pneumonia-management cascade.

### 1 3. Discussion

2 The current challenge of COVID-19 is not its diagnosis but its impact on internal organs. Recently,  
3 an increasing number of COVID-19-related diagnostic and segmentation studies have been published,  
4 but the damaging effects of COVID-19 on the internal organs of the human body are more important. In  
5 this paper, we proposed a novel multimodal channel and receptive field attention network (MCRFNet)  
6 for assessing lung damage in COVID-19 patients.

7 Rouby et al. assessed lung involvement by scoring eight areas' sonograms [27], while the  
8 sonograms need to be manually identified by doctors, which is time-consuming and labor-intensive.  
9 There are also some studies on the automatic classification of sonograms, but most of them are only for  
10 detecting the B-line of sonograms [28, 29]. Our MCRFNet can achieve fully automatic assessment of  
11 lung involvement in COVID-19 patients. The lung ultrasound images of these patients were classified  
12 into six types of sonograms, and the classification results were quantitatively scored to obtain the total  
13 scores of 8 regions. Then, a correlation analysis between the scores and PCO<sub>2</sub>, which is the most  
14 relevant to lung involvement, was obtained. Finally, a Pearson correlation coefficient of 0.73 was  
15 calculated, indicating that our classification scores can reflect the lung involvement of COVID-19  
16 patients. It is useful to choose the correct treatment method based on the severity of the situation.

17 The reason for the performance decline after adding the Philips dataset is shown in Fig. 6. The  
18 Philips dataset comes from three different machines in two centers. Due to its contrast and resolution  
19 disparity with the Stork and Mindray datasets, our model misclassified some images into incorrect  
20 categories. It also means that the robustness and cross-domain adaptability of our model are not  
21 perfect; this limitation serves as a direction for future improvement. In terms of this problem, to make  
22 the classification model more robust, we used traditional methods to extract shallow features that are  
23 not sensitive to imaging parameters and observed great performance.



**Fig. 6.** Comparison of B1-line of Philips dataset from three different machines.

Some attempts were made to verify the performance of the model. We tried to train with data from a single center, while data from another center were used for independent testing. However, the Stork dataset has only 58 consolidation images; if we use the trained Stork model to predict Mindray's consolidation data (1136 images), the accuracy of independent testing will be greatly reduced in this category of classification and vice versa.

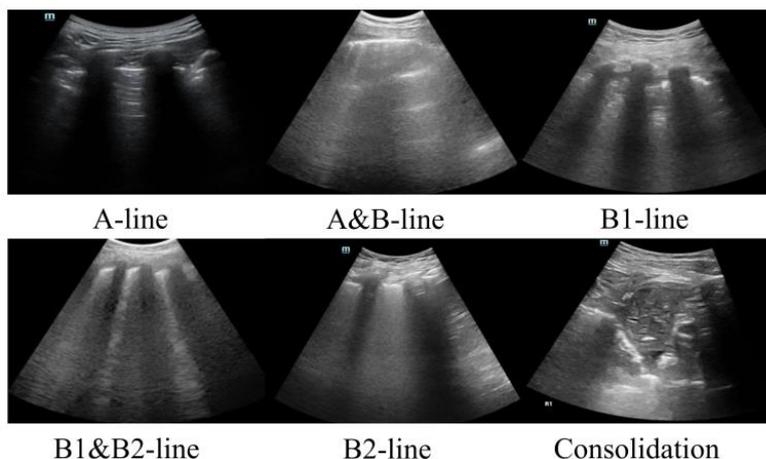
#### 4. Conclusions

In this paper, we proposed a novel classification network named MCRFNet that utilizes multimodal fusion and channel and receptive field attention to classify lung sonograms. In addition, we scored the predicted categories that reflect the degree of lung involvement in the patient and helped doctors to combine other indicators to assess disease trends in COVID-19 patients.

#### 5. Methods

##### 5.1. Ultrasound Data Acquisitions

In ultrasound imaging, the degree of lung involvement is related to several typical sonograms. The A-line is a horizontal reverberation artifact of the pleura caused by multiple reflections, representing the normal lung surface [30]. The B-line represents the interlobular septum, which is denoted by a discrete laser-like vertical hyperechoic artifact that spreads to the end of the screen, and it can be represented as the B1-line [31]. The fusion B-line is a sign of pulmonary interstitial syndrome, which shows a large area filled with the B-line in the intercostal space, and it can be represented as the B2-line [26]. Pulmonary consolidation is characterized by a liver-like echo structure of the lung parenchyma, with a thickness of at least 15 mm [27], as shown in Fig. 7.

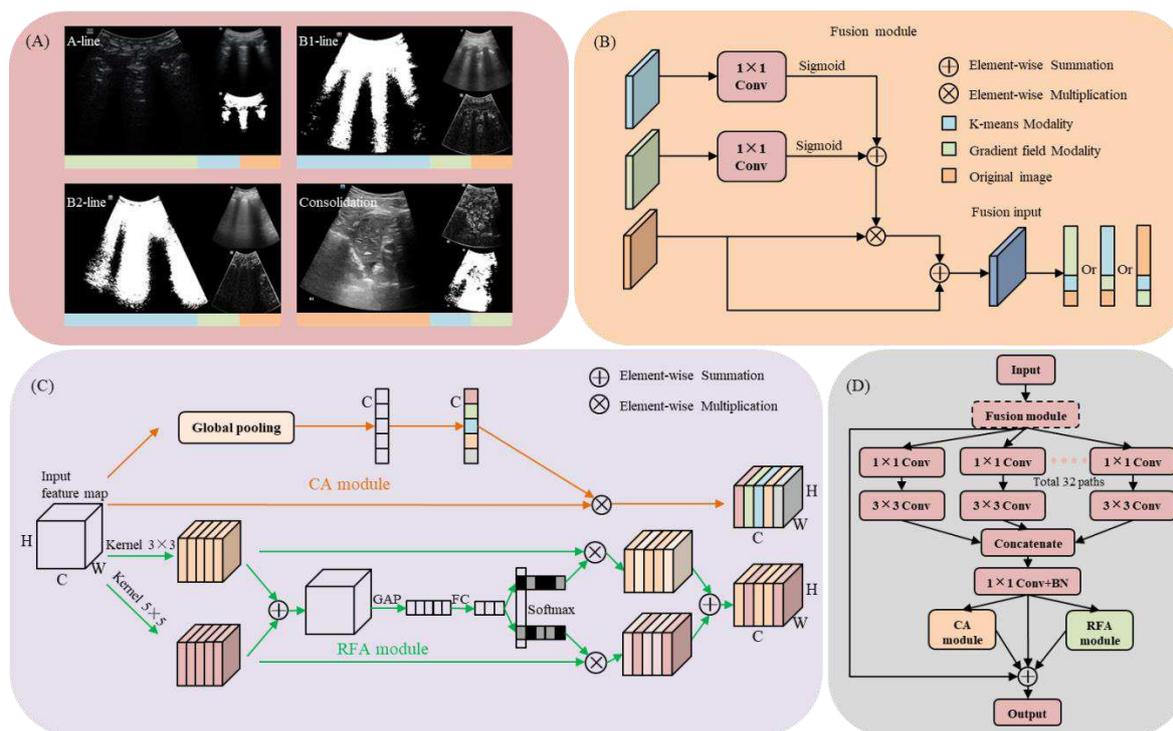


**Fig. 7.** Different ultrasound sonograms in lung examination.

We used three datasets from four medical centers to build and evaluate the model: ultrasound images collected by the Stork ultrasound system (Stork Healthcare Co., Ltd. Chengdu, China) at Ruijin Hospital, Mindray ultrasound system (Mindray Medical International Limited, Shenzhen, China) at Shanghai Public Health Center, Philips ultrasound system (Philips Medical Systems, Best, the Netherlands) at Wuhan Sixth People's Hospital and Hangzhou Infectious Disease Hospital. The Stork dataset was collected with an H35C (2-5 MHz) convex array transducer, the Mindray dataset with an SC5-1 (1-5 MHz) convex array transducer, and the Philips dataset with an Epiq 5, Epiq 7 C5-1 (1-5 MHz) convex array transducer.

### 5.2. Multimodal generation and fusion

According to doctors' experience in recognizing sonograms, parallel echo rays of the A-line, beam-like echo rays of the B-line, and the accumulation of exudate of lung consolidation are used as markers for classification. The gradient field is highly sensitive to the parallel echo rays of the A-line, and K-means clustering can better highlight the beam-like echo rays of the B-line [28]. As shown in Fig. 8(A), we produced the gradient field and K-means clustering images as two new modalities for extracting shallow features.



1  
2 **Fig. 8.** (A) Generated gradient field and K-means clustering modalities. The largest picture represents the most  
3 sensitive modal. (B) The proposed fusion module. (C) Details of CA and RFA module. (D) The MCRF block is  
4 integrated with a ResBlock in ResNeXt. The fusion module is only used in the first ResNeXt layer.

5        There are many methods to fuse multimodal inputs, and concatenate-based fusion is an intuitive  
6 fusion method [32], but this method is more suitable for situations where each modality is equally  
7 important for classification. Extracting features first and then concatenating is also a very popular  
8 fusion method [33], while the number of parameters and GPU memory limit its application. In this  
9 paper, we proposed a brand-new fusion network, as shown in Fig. 8(B); this network used the  
10 minimum network parameters to achieve multimodal automatic weight distribution, thus underlining  
11 the embedding of the other two modalities on the original image. Two  $1 \times 1$  convolutions were used to  
12 update the weights of the K-means modality and gradient field modality in the easiest way. After  
13 elementwise summation, we highlighted it on the original image by elementwise multiplication with  
14 the original image and finally added it to the original image to obtain the final fusion input.

15

### 1 5.3. ResNeXt with CRF attention block for classification

2 Shallow features were extracted by the traditional methods in section 5.2. To extract deep features  
 3 more effectively, we chose deep and wide ResNeXt as the backbone network for classification.  
 4 ResNeXt [22] is a combination of ResNet [21] and Inception [34], which improves accuracy through  
 5 wider or deeper networks. Each of its blocks is a measurable dimension in addition to the width and  
 6 depth dimensions. It inherits the strategy of repeating layers of ResNet but increases the number of  
 7 paths and uses split conversion and merge strategies in a simple and scalable manner. ResNeXt with  
 8 the CRF attention building block is shown in Fig. 8 (D). Our whole network replaces the building  
 9 block in ResNeXt with our CRF attention building block. In detail, there is one first layer and three  
 10 residual layers in our network, and every first layer and residual layer has one and three grassroots  
 11 CRFA building blocks, respectively.

12 The CRF attention module is comprised of channelwise and receptive field attention modules,  
 13 denoted as CA and RFA, respectively (Fig. 8[C]). The CA module attempts to assist the learning of  
 14 layer-specific features and explores channelwise dependencies for the selection of useful features.  
 15 Specifically, given an intermediate input feature channel set  $U \in R^{H \times W \times C}$ , a squeeze operation is  
 16 performed on the input image  $F_{sq}(U)$ , that is, global average pooling (GAP), to encode the entire  
 17 spatial feature on a channel as a global feature.

$$F_{sq}(U) = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H U(i, j), U \in R^{W \times H \times C} \quad (5)$$

18 The squeeze operation obtains the global description feature, and another operation is required to  
 19 capture the relationship between the channels, namely, the excitation operation  $F_{ex}(F_{sq}(U))$ .

$$F_{ex}(F_{sq}(U)) = \sigma(W_2 \delta(W_1 F_{sq}(U))) \quad (6)$$

20 where  $\delta(\cdot)$  and  $\sigma(\cdot)$  are the ReLU activation and sigmoid function, respectively.  $W_1 \in R^{\frac{C}{16} \times C}$  and  
 21  $W_2 \in R^{\frac{C}{16} \times C}$  are the learning weights of the two fully connected layers. The excitation operation can  
 22 learn the nonlinear relationship between channels. Finally, the learned activation value of each  
 23 channel (sigmoid activation) is multiplied by the original feature on  $U$  :

$$\hat{U}_C = F(U, F_{ex}(F_{sq}(U))) = U \cdot F_{ex}(F_{sq}(U)), U \in R^{W \times H \times C} \quad (7)$$

1 given the same input feature channel set  $U \in R^{H \times W \times C}$ , we first conducted two transformations  
 2  $\hat{F}: U \rightarrow \hat{U} \in R^{H \times W \times C}$  and  $\hat{F}: U \rightarrow \hat{U} \in R^{H \times W \times C}$  with kernel sizes of 3 and 5, respectively. Then,  
 3 the results of multiple branches are combined by summing the elements as follows:

$$4 \quad U = \hat{U} \oplus \hat{U} \quad (8)$$

5 For the output features  $\hat{U}$  and  $\hat{U}$ , squeeze and excitation are performed, respectively, as in  
 6 Equation 2. Additionally, we used soft attention across channels to select different spatial scales of  
 7 information, which is guided by the compact feature descriptor  $\mathbf{z}$ :

$$8 \quad \hat{U}_{RF} = \frac{e^{A \cdot \mathbf{z}}}{e^{A \cdot \mathbf{z}} + e^{B \cdot \mathbf{z}}} F_{ex}(F_{sq}(\hat{U})) + (1 - \frac{e^{A \cdot \mathbf{z}}}{e^{A \cdot \mathbf{z}} + e^{B \cdot \mathbf{z}}}) F_{ex}(F_{sq}(\hat{U})), U \in R^{W \times H \times C} \quad (9)$$

9 where  $A, B \in R^{\frac{C}{16} \times C}$ . The final feature map of RFA is obtained through the attention weights on  
 10 various kernels as in the above equation.

11 With the CA and RFA modules, the CA and RFA results are further integrated with the add  
 12 operation, as shown in Fig. 6(D):

$$13 \quad \bar{U} = \hat{U}_C + \hat{U}_{RF} \quad (10)$$

14 Detailed procedures are as follows. (1) Extract the most common 6 types of datasets in Fig. 5  
 15 from the training set in equal proportions randomly to avoid an imbalanced sample and ensure that  
 16 each category can be learned. (2) Enhance the data by rotation and normalize the intensity of the  
 17 image. (3) Select the classifier with the best performance and test it on the test set to obtain the  
 18 corresponding prediction results.

19

#### 20 5.4. Establishment of scoring standards

21 We predicted the patient's per part ultrasound video of multiple examinations through the trained  
 22 MCRFNet and classified and scored sonograms according to the paper [35]. A-line indicates that the  
 23 patient is normally ventilated, with a score of 0; A & B-line indicates that the patient has mild lung  
 24 ventilation loss, with a score of 1; B1-line indicates that the patient has moderate lung ventilation loss,  
 25 with a score of 2; B1 & B2-line indicates that the patient has severe lung loss of ventilation, with a

1 score of 2.5; B2-line indicates that the patient has very severe loss of lung ventilation, with a score of  
2 3; Consolidation indicates that the patient has a solid lung change characterized by dynamic air  
3 bronchial signs, with a score of 4. After the classification result is quantified, the sum is divided by all  
4 the frames to obtain the final lung function severity score, which is 0 to 4.

### 6 *5.5. Training Strategy*

7 For the Stork, Mindray, Stork & Mindray, and Stork & Mindray & Philips datasets, we used an  
8 independence test to verify the performance of the classifier. All the images were resized to  $128 \times 128$ ,  
9 and a training batch consisted of 8 randomly selected images. We regularized the model by using  
10 dropout during training, and the neural network parameters were then trained by maximizing  
11 log-likelihood using the momentum optimizer with an initial learning rate of 0.1. Then, every 30  
12 epochs, the learning rate dropped by 10 times, stochastically minimizing the cross-entropy between  
13 annotated labels and predictions. Our experiments were implemented using TensorFlow on a PC with  
14 an Intel Xeon E5, 64G RAM, Nvidia TITAN Xp 12G.

### 16 **Abbreviations**

17 LUS: Lung ultrasound; COVID-19: novel coronavirus 2019; RT-PCR: reverse transcription  
18 polymerase chain reaction; CT: Computed tomography; VAP: ventilator-associated pneumonia;  
19 PCO<sub>2</sub>: CO<sub>2</sub> partial pressure; PSL: parasternal line; POI: points of interest; AAL: anterior axillary  
20 line; PAL: posterior axillary line.

### 21 **Ethics approval and consent to participate**

22 The Research Ethics Committee of the Ruijin Hospital granted approval for this study, and the  
23 collection and use of the data analyzed in this study.

### 24 **Consent for publication**

25 Agreed by the author.

### 26 **Availability of data and materials**

27 The datasets used and analysed during the current study are available from the corresponding author  
28 on reasonable request.

### 29 **Competing interests**

30 The authors declare that they have no competing interests.

### 31 **Funding**

32 This research was funded by the National Natural Science Foundation of China, grant number  
33 91959127 and by the Shanghai Science and Technology Action Innovation Plan, grant number  
34 19441903100.

### 35 **Author's contributions**

1 ZH and JhY suggested the algorithm for image analysis, ZH implemented it and analysed the  
 2 experimental results; ZL, HZ, JL, BH, AL, XS and YX collected experimental data and provided  
 3 clinical guidance; AL, JH, XP and YD consulted the obtained result. All authors read and approved  
 4 the final manuscript.

### 5 **Acknowledgements**

6 The authors are grateful to all study participants.

### 7 **Authors' information**

8 <sup>1</sup> Department of Electronic Engineering, Fudan University, Shanghai 200433, China.

9 <sup>2</sup> Department of Ultrasound, Ruijin Hospital, Shanghai Jiaotong University School of Medicine,  
 10 Shanghai 200025, China.

11 <sup>3</sup> Department of Ultrasound, Shanghai Public Health Clinical Center, Shanghai 201508, China.

12 <sup>4</sup> Department of Ultrasound, Xixi Hospital of Hangzhou, Hangzhou 310023, China.

13 <sup>5</sup> Department of Ultrasound, The Six Hospital of Wuhan, Affiliated Hospital of Jianghang University,  
 14 Wuhan, China.

15 <sup>6</sup> Institute of Biomedical and Health Engineering Shenzhen Institutes of Advanced Technology,  
 16 Chinese Academy of Sciences.

### 17 **References**

18  
 19 1. Coronavirus resource center JHUoM. COVID-19 Case Tracker: Follow global cases and trends. Updated  
 20 daily. <https://coronavirus.jhu.edu>. Published 2020.

21 2. Coronavirus disease (COVID-19) Pandemic WHO.  
 22 <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. Published 2020.

23 3. Hao W, Li M. Clinical diagnostic value of CT imaging in COVID-19 with multiple negative RT-PCR testing.  
 24 *Travel Med Infect Dis* 2020;101627. doi: 10.1016/j.tmaid.2020.101627

25 4. Bernheim A, Mei X, Huang M, Yang Y, Fayad ZA, Zhang N, Diao K, Lin B, Zhu X, Li K, Li S, Shan H, Jacobi A,  
 26 Chung M. Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of Infection.  
 27 *Radiology* 2020;200463. doi: 10.1148/radiol.2020200463

28 5. Pan F, Ye T, Sun P, Gui S, Liang B, Li L, Zheng D, Wang J, Hesketh RL, Yang L, Zheng C. Time Course of  
 29 Lung Changes at Chest CT during Recovery from Coronavirus Disease 2019 (COVID-19). *Radiology*  
 30 2020;295(3):715-721. doi: 10.1148/radiol.2020200370

31 6. Self WH, Courtney DM, McNaughton CD, Wunderink RG, Kline JA. High discordance of chest x-ray and  
 32 computed tomography for detection of pulmonary opacities in ED patients: implications for diagnosing  
 33 pneumonia. *Am J Emerg Med* 2013;31(2):401-405. doi: 10.1016/j.ajem.2012.08.041

34 7. Parra A, Perez P, Serra J, Roca O, Masclans JR, Rello J. Pneumonia and Lung Ultrasound in the Intensive Care  
 35 Unit. *Chest* 2014;145(3). doi: 10.1378/chest.1806646

36 8. Aghdashi M, Aghdashi M, Broofeh B, Mohammadi A. Diagnostic performances of high resolution  
 37 trans-thoracic lung ultrasonography in pulmonary alveoli-interstitial involvement of rheumatoid lung disease.  
 38 *Int J Clin Exp Med* 2013;6(7):562-566.

39 9. Liu XL, Lian R, Tao YK, Gu CD, Zhang GQ. Lung ultrasonography: an effective way to diagnose  
 40 community-acquired pneumonia. *Emerg Med J* 2015;32(6):433-438. doi: 10.1136/emermed-2013-203039

41 10. Bouhemad B, Dransart-Raye O, Mojoli F, Mongodi S. Lung ultrasound for diagnosis and monitoring of  
 42 ventilator-associated pneumonia. *Ann Transl Med* 2018;6(21):418. doi: 10.21037/atm.2018.10.46

43 11. Mongodi S, Via G, Girard M, Rouquette I, Missot B, Braschi A, Mojoli F, Bouhemad B. Lung Ultrasound for  
 44 Early Diagnosis of Ventilator-Associated Pneumonia. *Chest* 2016;149(4):969-980. doi: 10.1016/j.chest.2015.12.012

- 1 12. Berlet T, Etter R, Fehr T, Berger D, Sendi P, Merz TM. Sonographic patterns of lung consolidation in  
2 mechanically ventilated patients with and without ventilator-associated pneumonia: a prospective cohort  
3 study. *J Crit Care* 2015;30(2):327-333. doi: 10.1016/j.jcrc.2014.11.021
- 4 13. Xia Y, Ying Y, Wang S, Li W, Shen H. Effectiveness of lung ultrasonography for diagnosis of pneumonia in  
5 adults: a systematic review and meta-analysis. *J Thorac Dis* 2016;8(10):2822-2831. doi: 10.21037/jtd.2016.09.38
- 6 14. Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from X-ray images utilizing transfer  
7 learning with convolutional neural networks. *Phys Eng Sci Med* 2020;43(2):635-640. doi:  
8 10.1007/s13246-020-00865-4
- 9 15. El Asnaoui K, Chawki Y, Idri AJae-p. Automated Methods for Detection and Classification Pneumonia  
10 based on X-Ray Images Using Deep Learning. 2020;arXiv:2003.14363. Accessed March 01, 2020.
- 11 16. Zheng C, Deng X, Fu Q, Zhou Q, Feng J, Ma H, Liu W, Wang X. Deep Learning-based Detection for  
12 COVID-19 from Chest CT using Weak Label, 2020.
- 13 17. Chen J, Wu L, Zhang J, Zhang L, Gong D, Zhao Y, Hu S, Wang Y, Hu X, Zheng B, Zhang K, Wu H, Dong Z,  
14 Xu Y, Zhu Y, Chen X, Yu L, Yu H. Deep learning-based model for detecting 2019 novel coronavirus pneumonia  
15 on high-resolution computed tomography: a prospective study in 27 patients, 2020.
- 16 18. Shan F, Gao Y, Wang J, Shi W, Shi N, Han M, Xue Z, Shen D, Shi Y. Lung Infection Quantification of  
17 COVID-19 in CT Images with Deep Learning, 2020.
- 18 19. Gaál G, Maga B, Lukács A. Attention U-Net Based Adversarial Architectures for Chest X-ray Lung  
19 Segmentation. arXiv e-prints 2020;arXiv:2003.10304.
- 20 20. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv  
21 e-prints 2014;arXiv:1409.1556.
- 22 21. He KM, Zhang XY, Ren SQ, Sun J. Deep Residual Learning for Image Recognition. 2016 Ieee Conference on  
23 Computer Vision and Pattern Recognition (Cvpr) 2016:770-778. doi: 10.1109/Cvpr.2016.90
- 24 22. Xie SN, Girshick R, Dollar P, Tu ZW, He KM. Aggregated Residual Transformations for Deep Neural  
25 Networks. *Proc Cvpr Ieee* 2017:5987-5995. doi: 10.1109/Cvpr.2017.634
- 26 23. Hu J, Shen L, Albanie S, Sun G, Wu E. Squeeze-and-Excitation Networks. *IEEE Trans Pattern Anal Mach*  
27 *Intell* 2020;42(8):2011-2023. doi: 10.1109/TPAMI.2019.2913372
- 28 24. Li X, Wang W, Hu X, Yang J. Selective Kernel Networks. arXiv e-prints 2019;arXiv:1903.06586.
- 29 25. Chattopadhyay A, Sarkar A, Howlader P, Balasubramanian VN. Grad-CAM++: Generalized Gradient-Based  
30 Visual Explanations for Deep Convolutional Networks. 2018 IEEE Winter Conference on Applications of  
31 Computer Vision (WACV)2018; p. 839-847.
- 32 26. Volpicelli G, Elbarbary M, Blaivas M, Lichtenstein DA, Mathis G, Kirkpatrick AW, Melniker L, Gargani L,  
33 Noble VE, Via G, Dean A, Tsung JW, Soldati G, Copetti R, Bouhemad B, Reissig A, Agricola E, Rouby JJ, Arbelot  
34 C, Liteplo A, Sargsyan A, Silva F, Hoppmann R, Breikreutz R, Seibel A, Neri L, Storti E, Petrovic T, Icc-Lus I-L.  
35 International evidence-based recommendations for point-of-care lung ultrasound. *Intens Care Med*  
36 2012;38(4):577-591. doi: 10.1007/s00134-012-2513-4
- 37 27. Rouby JJ, Arbelot C, Gao Y, Zhang M, Lv J, An Y, Wang C, Bin D, Barbas CSV, Dexheimer Neto FL, Prior  
38 Caltabeloti F, Lima E, Cebey A, Perbet S, Constantin JM, group As. Training for Lung Ultrasound Score  
39 Measurement in Critically Ill Patients. *Am J Respir Crit Care Med* 2018. doi: 10.1164/rccm.201802-0227LE
- 40 28. Brusasco C, Santori G, Bruzzo E, Tro R, Robba C, Tavazzi G, Guarracino F, Forfori F, Boccacci P, Corradi F.  
41 Quantitative lung ultrasonography: a putative new algorithm for automatic detection and quantification of  
42 B-lines. *Crit Care* 2019;23(1):288. doi: 10.1186/s13054-019-2569-4

- 1 29. van Sloun RJG, Demi L. Localizing B-Lines in Lung Ultrasonography by Weakly Supervised Deep Learning,  
2 In-Vivo Results. *IEEE J Biomed Health Inform* 2020;24(4):957-964. doi: 10.1109/JBHI.2019.2936151
- 3 30. Liu J, Copetti R, Sorantin E, Lovrenski J, Rodriguez-Fanjul J, Kurepa D, Feng X, Cattaross L, Zhang H,  
4 Hwang M, Yeh TF, Lipener Y, Lodha A, Wang JQ, Cao HY, Hu CB, Lyu GR, Qiu XR, Jia LQ, Wang XM, Ren XL,  
5 Guo JY, Gao YQ, Li JJ, Liu Y, Fu W, Wang Y, Lu ZL, Wang HW, Shang LL. Protocol and Guidelines for  
6 Point-of-Care Lung Ultrasound in Diagnosing Neonatal Pulmonary Diseases Based on International Expert  
7 Consensus. *J Vis Exp* 2019(145). doi: 10.3791/58990
- 8 31. Francisco MJN, Rahal AJ, Vieira FA, Silva PS, Funari MB. Advances in lung ultrasound. *Einstein (Sao Paulo)*  
9 2016;14(3):443-448. doi: 10.1590/S1679-45082016MD3557
- 10 32. Hong D, Gao L, Yokoya N, Yao J, Chanussot J, Du Q, Zhang B. More Diverse Means Better: Multimodal  
11 Deep Learning Meets Remote Sensing Imagery Classification. *arXiv e-prints 2020:arXiv:2008.05457*.
- 12 33. Zhou T, Fu H, Zhang Y, Zhang C, Lu X, Shen J, Shao L. M2Net: Multi-modal Multi-channel Network for  
13 Overall Survival Time Prediction of Brain Tumor Patients. *arXiv e-prints 2020:arXiv:2006.10135*.
- 14 34. Szegedy C, Liu W, Jia YQ, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going  
15 Deeper with Convolutions. 2015 Ieee Conference on Computer Vision and Pattern Recognition (Cvpr) 2015:1-9.  
16 doi: DOI 10.1109/cvpr.2015.7298594
- 17 35. Ranzani OT, Taniguchi LU, Torres A. Severity scoring systems for pneumonia: current understanding and  
18 next steps. *Curr Opin Pulm Med* 2018;24(3):227-236. doi: 10.1097/Mcp.0000000000000468  
19

# Figures

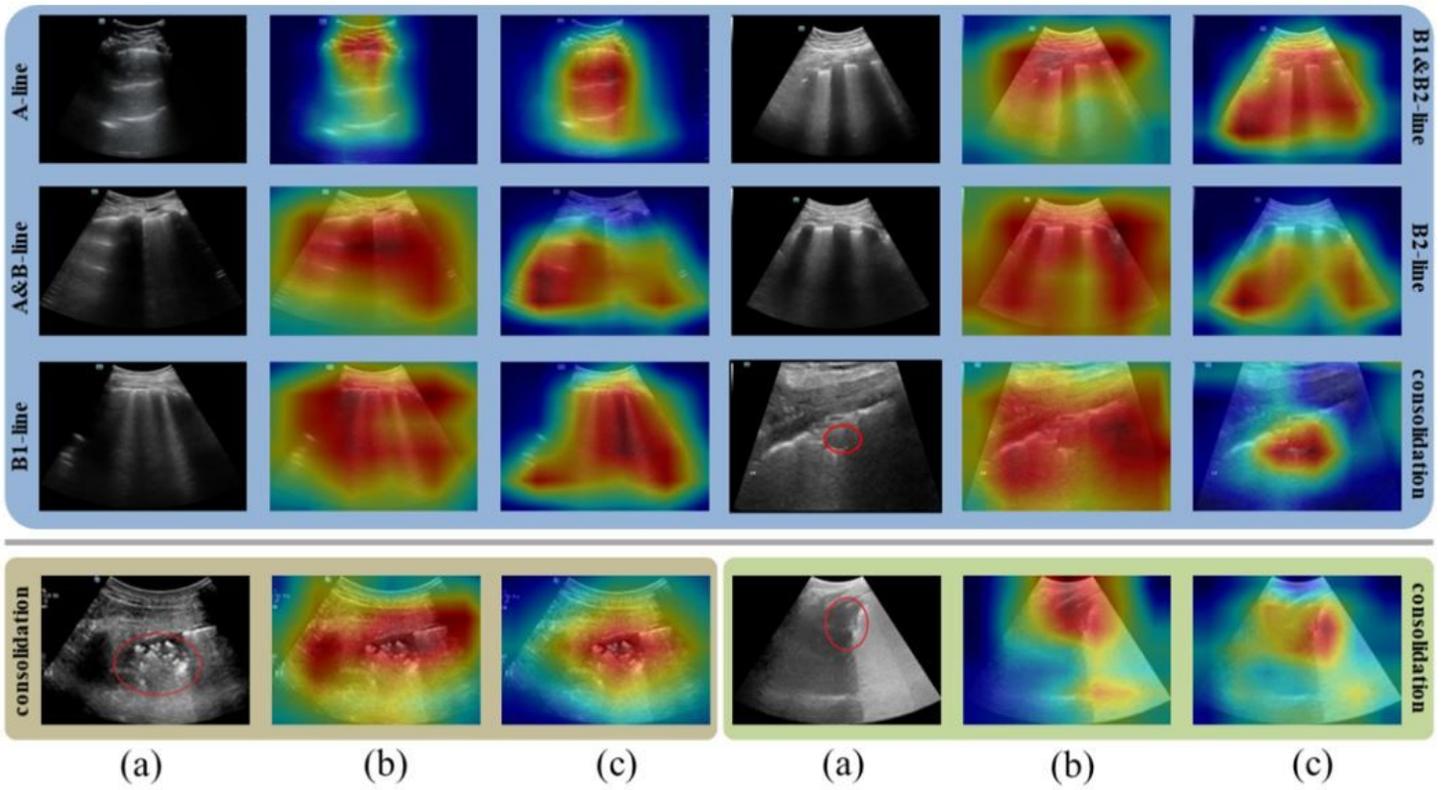


Figure 1

The Grad-CAM++ visualization results. The blue, red, and green backgrounds represent the Mindray dataset, Stork dataset, and Philips dataset, respectively. (a) Original input (the actual data are not marked with a red circle), (b) ResNeXt, and (c) MCRFNet. The red circle is the lung consolidation area.

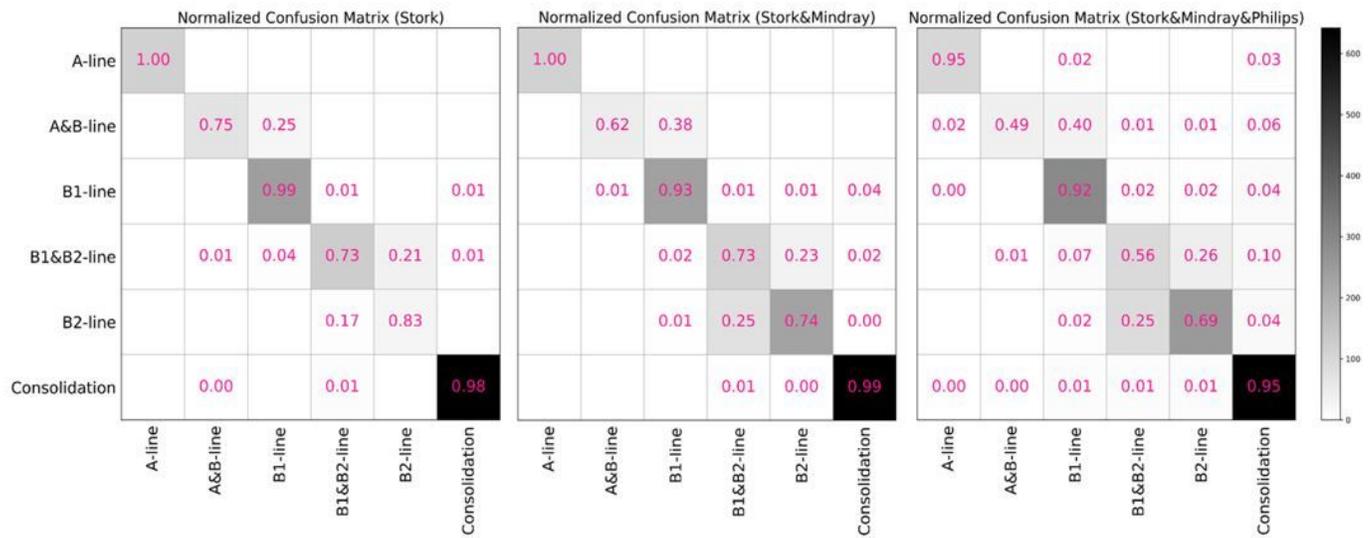
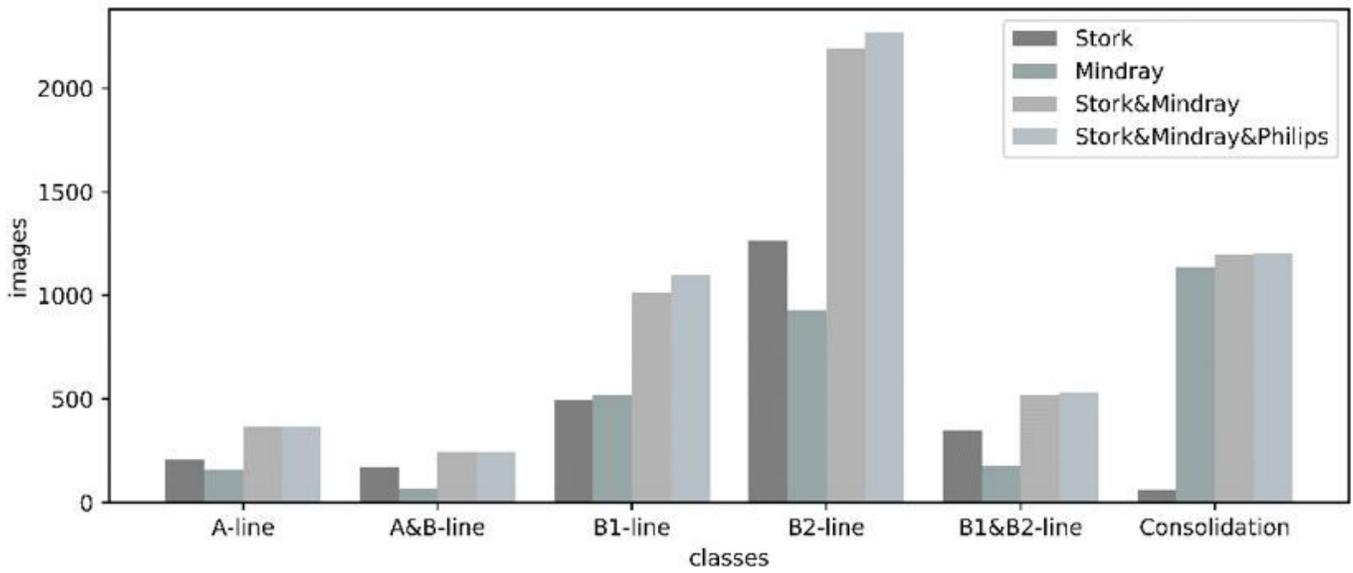


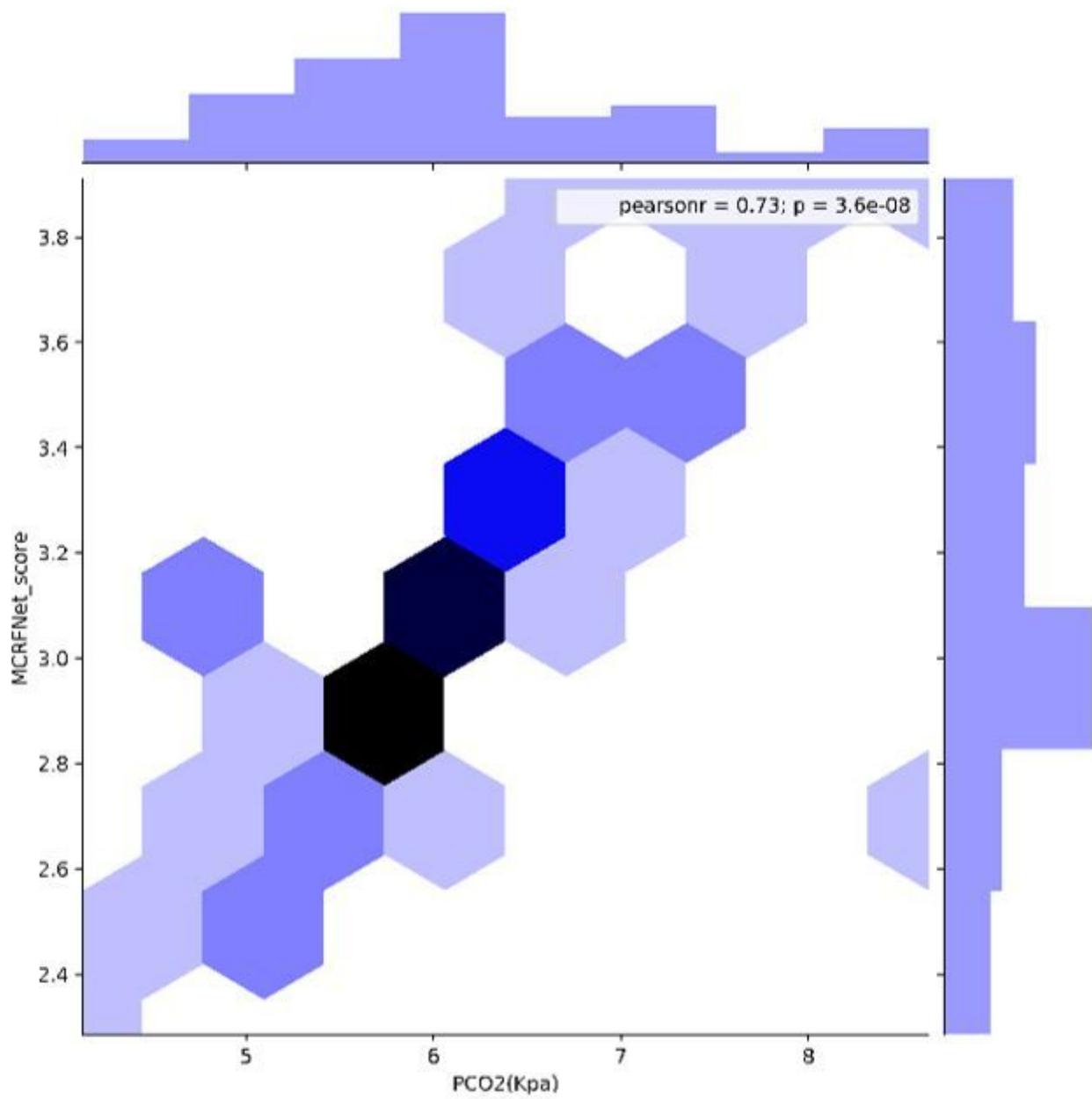
Figure 2

The normalized confusion matrix of classification on three datasets by MCRFNet. The abscissa is the predicted label, and the ordinate is the true label.



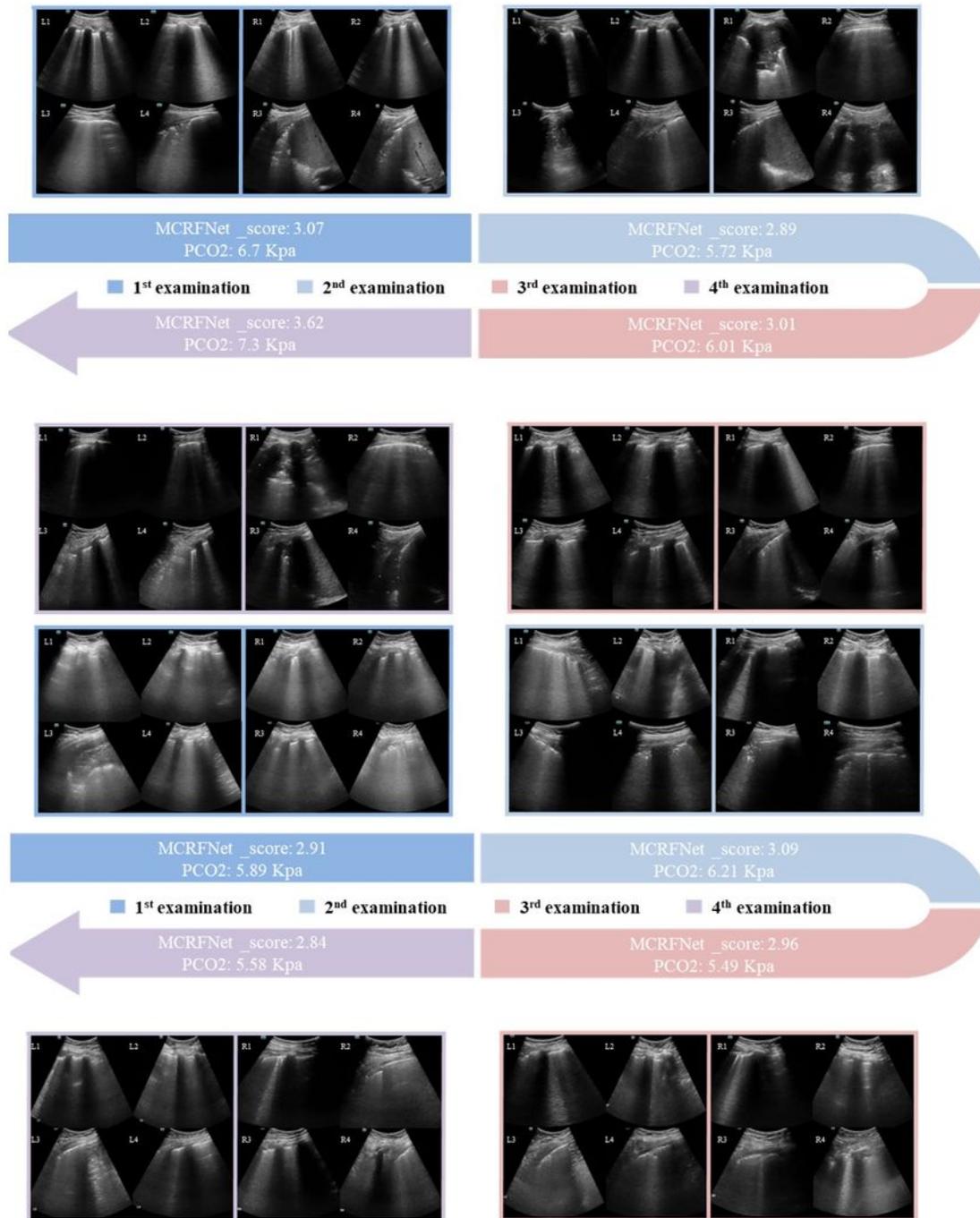
**Figure 4**

Distribution of expert manual labels of 5704 ultrasound images in three datasets.



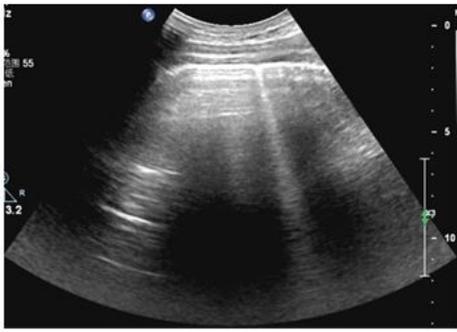
**Figure 5**

Scatter plot between the classification score of MCRFNet and CO2 partial pressure (PCO2).



**Figure 6**

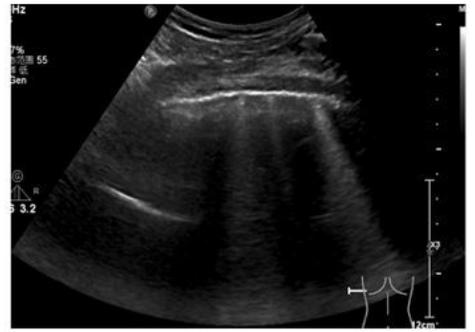
Two patients had multiple examinations of the MCRFNet score and PCO2 from the initial diagnosis to rehabilitation.



Wuhan Sixth People's Hospital  
#Philips machine 1



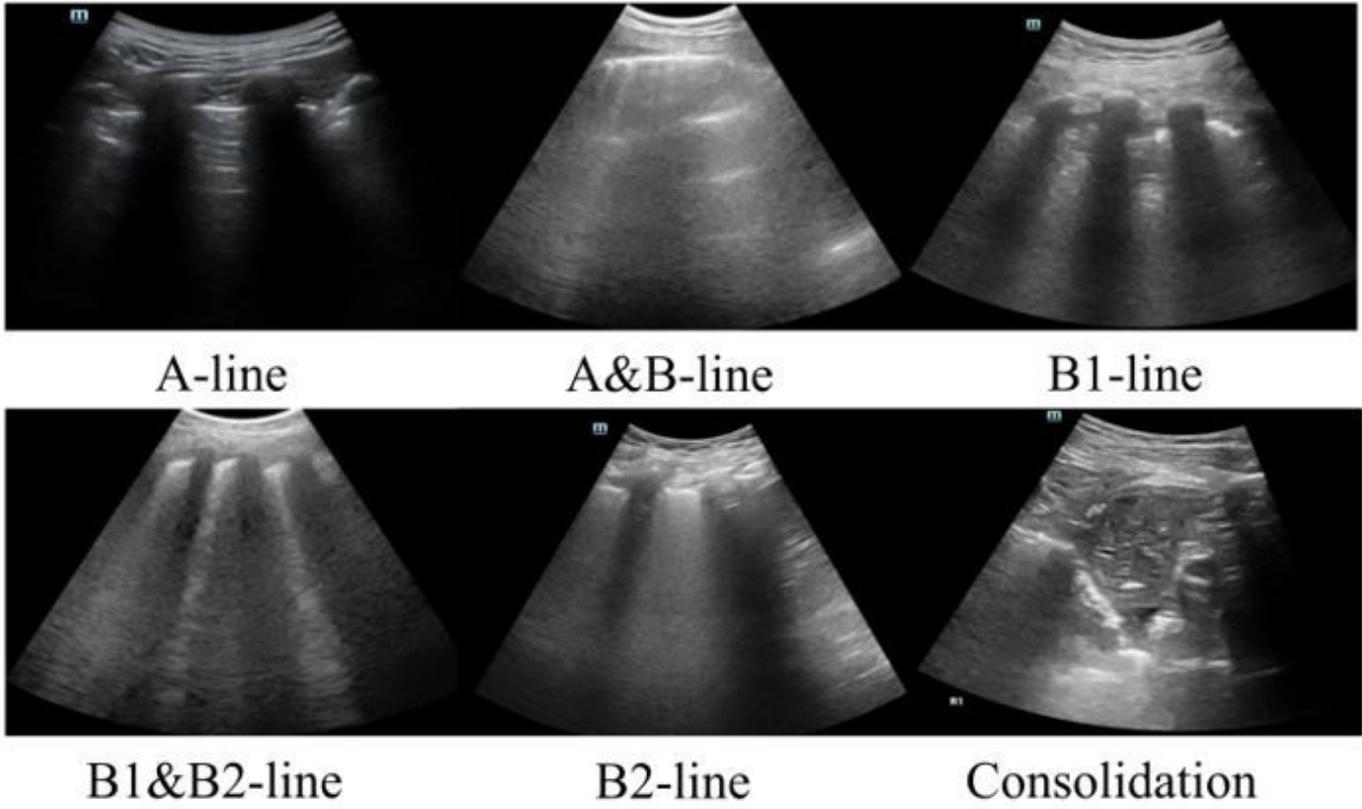
Wuhan Sixth People's Hospital  
#Philips machine 2



Hangzhou Infectious Disease Hospital  
#Philips machine 3

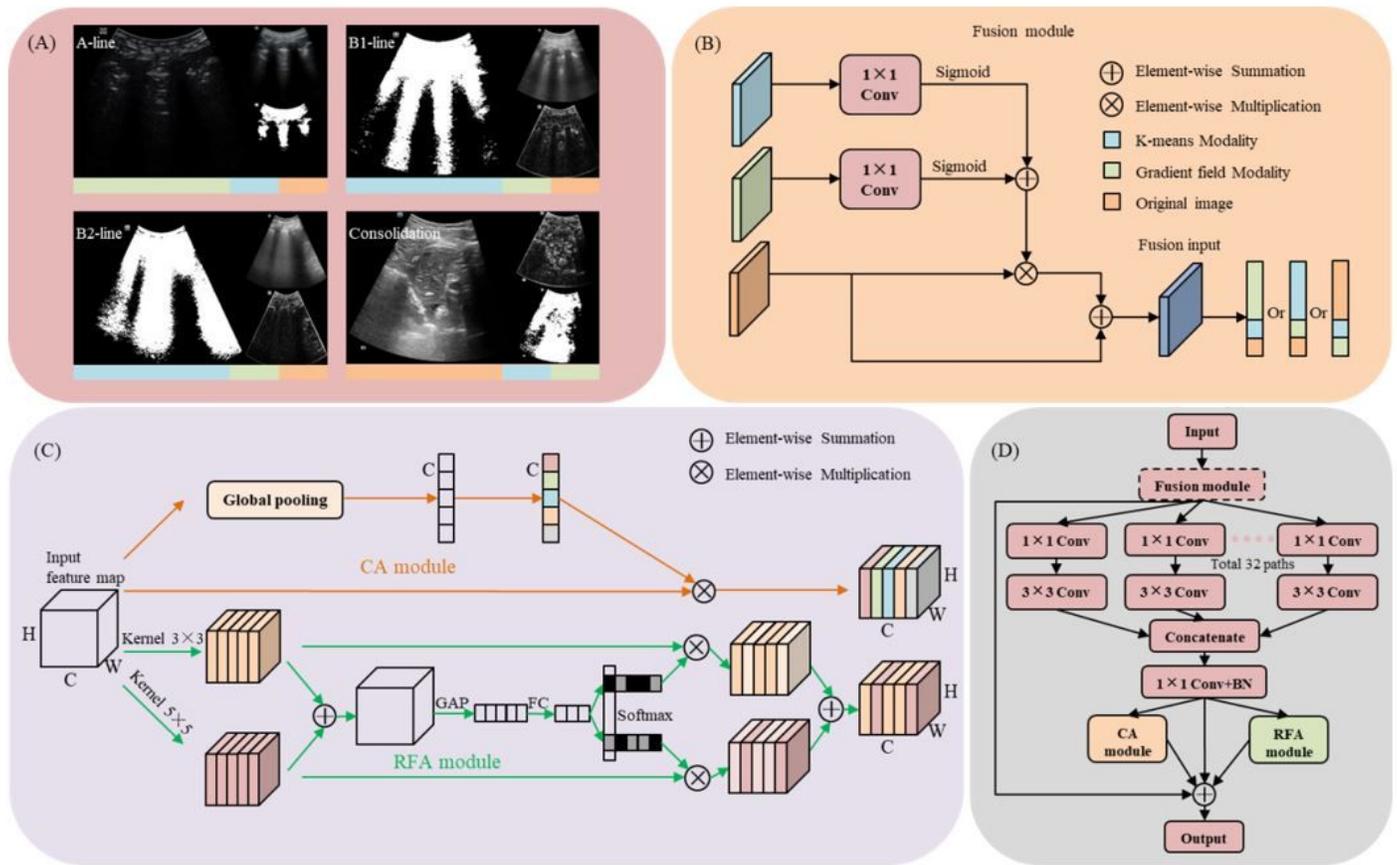
**Figure 6**

Comparison of B1-line of Philips dataset from three different machines.



**Figure 8**

Different ultrasound sonograms in lung examination.



**Figure 9**

(A) Generated gradient field and K-means clustering modalities. The largest picture represents the most sensitive modal. (B) The proposed fusion module. (C) Details of CA and RFA module. (D) The MCRF block is integrated with a ResBlock in ResNeXt. The fusion module is only used in the first ResNeXt layer.