Quantitative Analysis and Automated Lung Ultrasound Scoring for Evaluating COVID-19 Pneumonia With Neural Networks

Jiangang Chen, Chao He, Jintao Yin, Jiawei Li, Xiaoqian Duan, Yucheng Cao, Li Sun, *Member, IEEE*, Menghan Hu, Wenfang Li, and Qingli Li, *Senior Member, IEEE*

Abstract— As being radiation-free, portable, and capable of repetitive use, ultrasonography is playing an important role in diagnosing and evaluating the COVID-19 Pneumonia (PN) in this epidemic. By virtue of lung ultrasound scores (LUSS), lung ultrasound (LUS) was used to estimate the excessive lung fluid that is an important clinical manifestation of COVID-19 PN, with high sensitivity and specificity. However, as a qualitative method, LUSS suffered from large interobserver variations and requirement for experienced clinicians. Considering this limitation, we developed a quantitative and automatic lung ultrasound scoring system for evaluating the COVID-19 PN. A total of 1527 ultrasound images prospectively collected from 31 COVID-19 PN patients with different clinical conditions were evaluated and scored with LUSS by experienced clinicians. All images were processed via a series of computer-aided analysis, including curve-to-linear conversion, pleural line detection, region-of-interest (ROI) selection, and feature extraction. A collection of 28 features extracted from the ROI was specifically defined for mimicking the LUSS. Multilayer fully connected neural networks, support vector machines, and decision trees were developed for scoring LUS images using the fivefold cross validation. The model with 128 × 256 two fully connected layers gave the best accuracy of 87%. It is concluded that the proposed method could assess the ultrasound images by assigning LUSS automatically with high accuracy, potentially applicable to the clinics.

Index Terms— Automated scoring, COVID-19 pneumonia, lung ultrasound, quantitative analysis.

I. Introduction

ULTRASONOGRAPHY is a radiation-free, easy to use, and portable imaging modality that competes the magnetic resonance imaging (MRI) and computed tomography (CT) in emergency and intensive care [1]. In this epidemic of COVID-19 Pneumonia (PN), ultrasonography was the only imaging modality with the access to the intensive care unit (ICU) and used as bedside supports in infected areas. Lung ultrasound (LUS) is appealing for PN diagnosis with the sensitivity and specificity superior to bedside chest X-ray, even close to CT [1], [2]. LUS has been intensively explored in evaluation of lung properties and increasingly applied in diagnosis of various lung diseases, including COVID-19 PN [3]–[10]. There existed several protocols for LUS evaluation of the COVID-19 PN, the impacts of which were analyzed by Mento *et al.* [11]. Lung ultrasound score (LUSS) is utilized in clinics for semiquantitative assessment of pulmonary edema with good correlation with excessive lung fluid [12], [13], which is an important clinical manifestation of COVID-19 PN [14].

In clinical practice, there existed different methods to calculate the LUSS. As the most exhaustive method, the 28-sector approach accumulated all the numbers of B-lines in the 28 intercostal spaces to achieve a score [11]. Alternatively, a method calculated the LUSS by dividing the chest wall into eight zones, including two anterior and two lateral zones of each hemithorax, with the B-lines evaluated in one space of each zone [15]. In other practices, the clinicians assigned one of the four numbers (0, 1, 2, and 3) to an ultrasound image as the LUSS, depending on the observation of B-lines (appearance, number, and confluence) [16], [17]. Other methods for calculating the LUSS could be referred to [18] and [19]. Although the methods for assigning LUSS provided semiobjective quantitative assessment for pulmonary edema, they suffered from subjectivity and large interobserver variations [16], [20], [21]. In addition, accurate assignment of the

Jiangang Chen, Jintao Yin, Xiaoqian Duan, Yucheng Cao, Li Sun, Menghan Hu, and Qingli Li are with the Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China (e-mail: jgchen@cee.ecnu.edu.cn).

Chao He and Wenfang Li are with the Department of Emergency and Critical Care, Changzheng Hospital, Naval Medical University, Shanghai 200433, China.

Jiawei Li is with the Fudan University Shanghai Cancer Center, Department of Medical Ultrasound, Fudan University, Shanghai 200433, China, and also with the Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200433, China.

LUSS depended on the experience of clinicians, which limited the efficacy of LUSS in clinics, especially for caring patients with COVID-19 PN where the medical staffs are lacking [21].

Recently, some computer-aided methods have been proposed for quantitative analysis on LUS images for objective assessment of the lung conditions [9], [20]–[23]. Brattain et al. [23] developed an image processing algorithm to detect the number of B-lines along angular slices to determine LUSS. Corradi et al. [21] proposed to quantitatively analyze the LUS images containing consolidation based on frequency distribution of gray scales. Brusasco et al. [20] developed an image segmentation method for automated detection of B-lines. Zong et al. [16] developed an animal model for quantitative evaluation of the pulmonary water content using LUSS. Recent endeavors were focused on the application of different classifiers for assigning the LUSS, including machine learning and deep learning [24], [25]. Although existing techniques provided objective lung evaluations with success to different extents, there is a lack of methods for automated B-line scoring that may be most potentially applicable to the clinics.

In view of this, we developed a quantitative and automatic LUS scoring system using multilayer fully connected neural networks for evaluating the LUS of COVID-19 PN.

II. MATERIALS AND METHODS

A. Experiment Design

From February 23, 2020 to April 2, 2020, 31 patients (age: 55 ± 21 years old, male: 19, and female: 12) admitted to the ICU ward of Huoshenshan Hospital, a newly built hospital specialized for caring patients with COVID-19 PN in Wuhan, were included. All the patients had the symptoms of fever and dyspnea. Chest CT (United Imaging, uCT760) showed bilateral ground-glass opacities with peripheral, posterior, and basal predominance, which was in line with the international agreement [26]. The patients were confirmed as affected by the COVID-19 virus with a positive RT-PCR test after admission to the hospital. In addition, the patients were confirmed with COVID-19 PN of different conditions: critical (ten cases, 32.3%), severe (nine cases, 29.0%), common (seven cases, 22.6%), and mild (five cases, 16.1%) according to the diagnosis and treatment standard for COVID-19 PN in China [27]. The height and weight of the patients were recorded for the calculation of body mass index (BMI) (height: 168 ± 13 cm and weight: 70 ± 18 Kg). All patients were monitored in terms of oxygen index, positive end-expiratory pressure, static lung compliance, respiratory index, blood pressure, and body temperature. Conditions of all patients were evaluated daily by the clinicians. The study was approved by the Ethics Committee of Huoshenshan Hospital, Wuhan, China (Approval number: HSSLL030). All patients provided written informed consents by themselves or family members.

All patients underwent LUS examinations for 12 standard fields on both hemithorax, including the upper and lower halves of the anterior, lateral, and posterior fields [28]. Repeated examinations were performed for part of the patients whose conditions changed during the treatment.

A total of 45 patient times were finally performed with eight patients scanned twice and three patients scanned three times. The ultrasound equipment LOGIQ e (GE Healthcare, Wauwatosa, WI, USA) was utilized with a curved array low-frequency transducer (1–5 MHz) with acquisition details as image depth of 15 cm, focal depth of 7.5 cm, mechanical index (MI) of 1.2, thermal index (TI) of 0.7, and operation mode: penetration. Three ultrasound images that contained A-lines or B-lines were collected in each field and stored in the DICOM format. All scans were operated in the transverse plane to avoid the acoustic shadows of ribs. The time gain compensation was maintained the same for all patients in order to minimize variations in intensity among the collected images. A total of 1620 images were collected and scored blindly by two experienced clinicians (>6 years in using LUS) following the criteria as Score 0: normal defined as absence of B-lines and appearance of A-lines, Score 1: septal syndrome defined as B-lines at regular distances of about 7 mm, Score 2: interstitial-alveolar syndrome defined as the B-line distance less than 7 mm with some confluent, and Score 3: white lung designated for coalesced B-lines with the confluence more than 80% [17]. As an assistant, the distance between two adjacent B-lines whose distance was visually the maximum among others on an image was measured when determining different scores, in particular for Scores 1 and 2. The measurement was performed by placing two points on the middle of the B-linewidth along the middle line of the image height on the ultrasound machine. Images that were assigned with different scores by the two clinicians were discarded (see Fig. 1). Finally, there were 1527 images assigned with scores and included in the study with the composition of Score 0: 413 (27.0%), Score 1: 370 (24.2%), Score 2: 417 (27.3%), and Score 3: 327 (21.4%). The data will be provided on our lab website in the future (https://bio-hsi.ecnu.edu.cn/). Fig. 2(a)–(d) representatively shows the ultrasound images with Scores 0, 1, 2, and 3, respectively.

B. Ultrasound Image Processing and Automated Scoring

The collected LUS images were automatically processed via four steps: 1) curve-to-linear conversion; 2) pleural line detection; 3) region-of-interest (ROI) selection; and 4) feature extraction, to extract features for the subsequent classification task.

1) Curve-to-Linear Conversion: The ultrasound images (see Fig. 2) were converted to the linear format through the curve-to-linear conversion in an automated manner using a customized MATLAB code. Considering a matrix containing a curved ROI denoted as Cmatrix [see Fig. 3(a)] and its corresponding linear matrix denoted as Lmatrix [see Fig. 3(b)]. The conversion from Cmatrix to Lmatrix is realized by the following steps:

1) To extract the ROI of Cmatrix from the image, namely Region A-B-C-D.
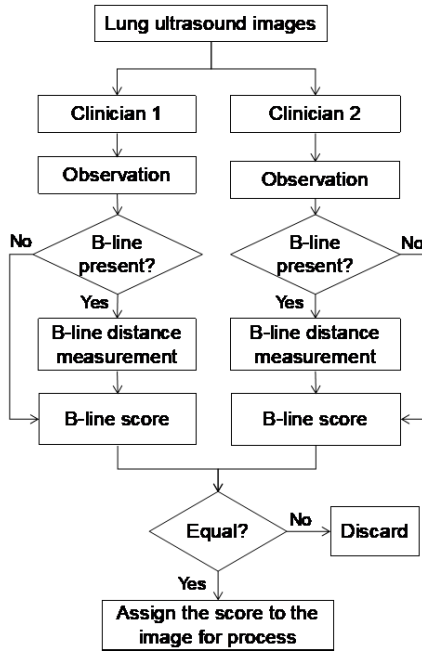2) To extract the coordinates of four endpoints A, B, C, and D of the ROI.
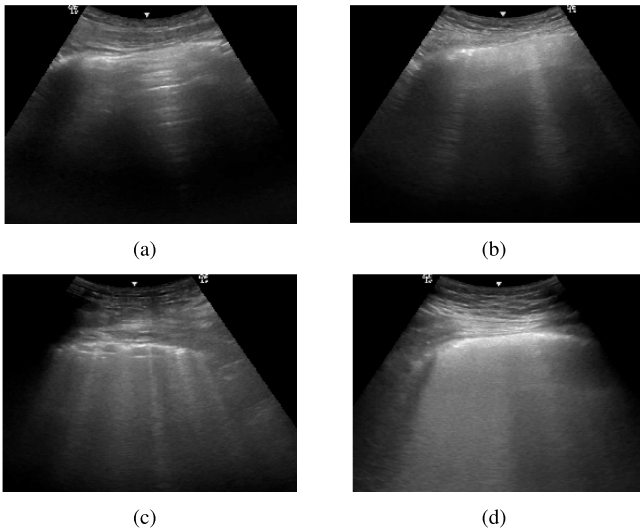
Fig. 1. Procedure of LUSS assignment.

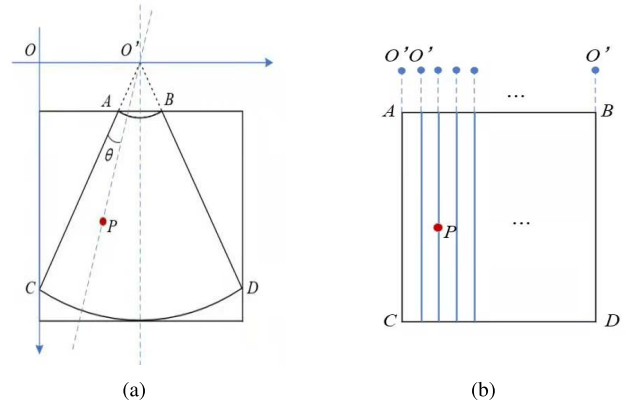Fig. 2. Ultrasound images of COVID-19 PN with LUSS of (a) 0, (b) 1, (c) 2, and (d) 3.

Fig. 3. Illustration of the curve-to-linear conversion. (a) Curved image. (b) Linear format after conversion.

gray values after the conversion involved bilinear interpolation [29]. Thereby, the image quality remained. Applied the curve-to-linear conversion, the images in the curve format in Fig. 2 were converted to their corresponding linear format, as shown in Fig. 4.

2) Pleural Line Detection: Our previous work described an automated method for the pleural line detection in the linear images as simplified in the following steps [30].

1) To apply Radon transform to the ultrasound image to extract all line objects serving as candidates for the pleural line. The resultant of the Radon transform is an intensity image with peaks condescending to the line objectives in the original image. The position of the peak reflected the depth and angle of the line.
2) The peak corresponding to the pleural line was picked up from all the candidates by considering the following preknowledge:
   a) The pleural line was normally more continuous than other candidates. As a result, the peak in the Radon-transformed intensity image corresponding to the pleural line normally has high value, ranking top positions among all the candidates.
   b) The depth of the pleural line is equal to the chest wall thickness (CWT) [31]. A personalized CWT could be estimated via a linear relationship of the BMI (defined as weight [in kilograms]/(height × height) [in square meters]) and CWT (details please refer to Graphs 1 and 2 in [31]). With the range of the possible depths of the pleurae, the searching range of the pleural line could be much shrunk.
3) The depth of the pleural line was then determined based on the position of the peak selected for the pleural line in the previous steps.

3) ROI Selection: According to the depth of the pleural line, a rectangular ROI (120 × 200 pixels, corresponding to a physical area of 55.4 × 92.3 mm², to cover most B-lines in the images) was automatically selected (see Fig. 4) via two steps: 1) the vertical position of the rectangular was 5 pixels below the pleural line and 2) the lateral position of the rectangular was determined where the average gray

3) To calculate the angle of the ROI by the slope of Lines A–C and B–D, denoted as ∠AO'B.
4) To calculate the distance between pixels by

$$D_p = L_{A-C}/(N_{A-C} - 1) \qquad (1)$$

where D_p, L_{A-C}, and N_{A-C} is the distance between adjacent pixels, length of Line A-C, and number of pixels along Line A-C, respectively.
5) To project the pixels in Cmatrix in polar coordinates [see Fig. 3(a)] to those in Lmatrix in Cartesian coordinates [see Fig. 3(b)], e.g., projecting the arcs A–B and C–D in Fig. 3(a) to Lines A–B and C–D in Fig. 3(b), respectively. The point P in Fig. 3(a) and (b) demonstrated the position of one pixel before and after the conversion. The calculation of the pixels'
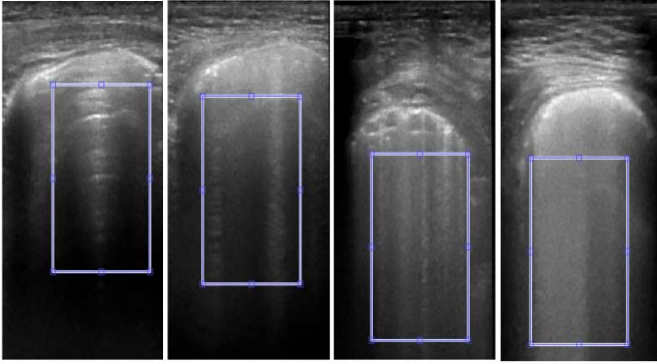
Fig. 4. Ultrasound images of COVID-19 PN with ROI selected in the linear format with LUSS of (a) 0, (b) 1, (c) 2, and (d) 3.
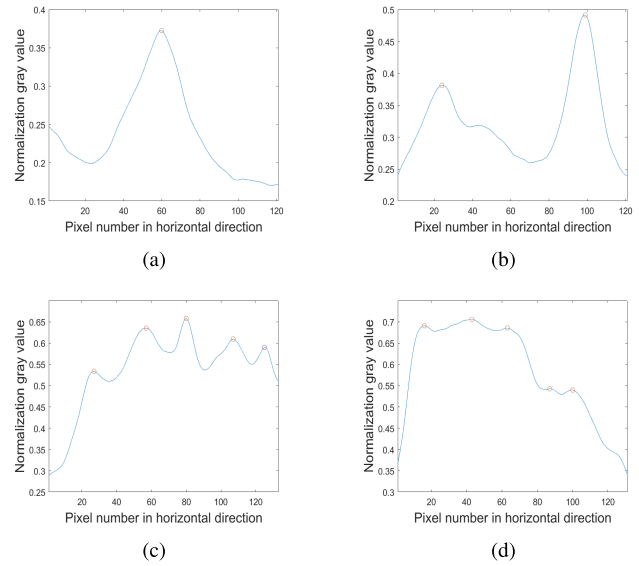

Fig. 5. Projection curves of the ROI in the vertical direction with LUSS of (a) 0, (b) 1, (c) 2, and (d) 3, corresponding to the graphs in Fig. 4. The curve is normalized with respect to the maximum intensity of the ROI.

value inside the rectangle was the maximum when shifting the rectangle from the left to right with the depth of the rectangle kept constant as determined in Step 1. The vertical position was chosen 5 pixels below the pleural line with two considerations: 1) it could avoid covering part of the pleural line which had certain thickness and 2) 5 pixels could be long enough to make the ROI away from the pleural line.

4) Ultrasound Image Feature Extraction: Two curves were created by averaging all the pixels' intensity of the ROI along each row and column in horizontal and vertical directions and normalized with respect to the maximum intensity of the ROI, denoted as C_H and C_V, respectively. The normalization was intended to avoid the influence of variations in intensity level among the images. Figs. 5 and 6 show the curves of C_H and C_V correspondingly obtained from the ROIs in Fig. 4, respectively. The peaks of the curves were automatically detected and indicated by red circles in the figures. Recognizing that the bases for LUSS include the presence of A-lines (Score: 0) and B-lines (Scores: 1-3), as well as the confluent degree of B-lines are a kind of 1-D problem; 28 features from the C_H, C_V, and the ROI were creatively extracted in an automated manner and grouped into four categories: peak information in the horizontal direction, peak information in the vertical direction, area under the curve, and image feature, as listed in Table I. Noted that the peaks were detected via several steps: 1) filtering the projection curves using "Gaussian" filter, making the curve smooth; 2) removing the mean value or linear trend of the signal using MATLAB command: "detrend;" and 3) detecting the peaks by moving a window (three points in size) along the curve to detect the peaks. A peak was detected when the value of the middle point was higher than that of the left and right points with a predefined constraint, i.e., the peak value should be higher than 0.1 to avoid the influence of noises.

The features were intentionally to mimic the clinician's interpretation when assigning the B-line scores, e.g., the number of peaks in C_H represented the number of B-line; the average distance between peaks and average area between peaks in C_H reflected the convergence of B-lines.

5) Classification Model: With the features extracted from the ROI as input, we applied one- and two-layer fully connected neural networks to learn the data with a set of different number of hidden nodes, including one layer with 4,

25, 256, and 1024 hidden nodes, two layers with 4×8, 128×256, and 512×1024 hidden nodes. Adam optimizer was applied to fit the data with ReLU as the activation function, learning rate as 0.0001, batch size as 32, and epoch as 1000. In addition, machine learning models, including support vector machine (SVM) with linear and Gaussian cores and decision trees of 4 and 8 layers, were applied to the data. All the proposed models were evaluated via the fivefold cross validation, i.e., the data were randomly divided into five equal parts with four as the training data sets and the other one as the testing data set. Such a division between the training and testing data sets was repeated five times under each of the two configurations: 1) based on patient information, i.e., there were no data from the same patient falling in both the training and testing data sets (Config-1) and 2) based on all the data neglecting patient information (Config-2). The neural networks have been developed under both Config-1 and Config-2. In addition, machine learning models were trained under Config-1. The model was developed on the platform of Pytorch 1.4 with Python 3.7 and trained on a personal computer (CPU: amd4800u, RAM: 16 g 4266 MHz). As an example, the development of the 512×1024 neural network with 1000 epochs costs 198 s.

As a summary, the proposed method included five steps: 1) curve-to-linear conversion; 2) pleural line detection; 3) ROI selection; 4) ultrasound image feature extraction; and 5) classification. Each step of the process was essential to perform the whole analysis.

1) Step (1) transferred all the images from curve to linear, which made all B-lines parallel to each other, and A-lines straight. This would benefit for the subsequent feature extraction from the projection curves in the vertical and horizontal directions.
2) Step (2) performed the pleural line detection that supported for the succedent ROI selection.
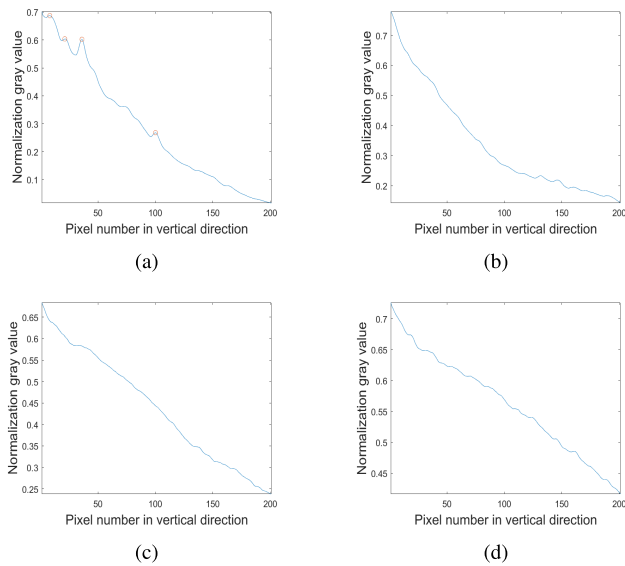
Fig. 6. Projection curves of the ROI in the horizontal direction with LUSS of (a) 0, (b) 1, (c) 2, and (d) 3, corresponding to the graphs in Fig. 4. The curve is normalized with respect to the maximum intensity of the ROI.

3) Step (3) selected the ROI from which the features were extracted for the classification task.
4) In Step (4), a total of 28 defined features that mimicked the human interpretation were extracted.
5) In Step (5), the deep learning model for automated lung ultrasound scoring was developed.

C. Statistics

Statistical analysis was performed using SPSS 22.0 for Windows system (SPSS Inc., Chicago, IL, USA). For a number of peaks, the Kruskal–Wallis test was used for the comparison among the four scoring subgroups, while one-way analysis of variance (ANOVA) was used for data comparison for all other quantitative parameters listed in Table I. The testing results of the fivefold cross validation of all the proposed models were compared using one-way repeated measures ANOVA. It was considered statistically significant if the P value was less than 0.05.

III. RESULTS

The statistical analysis indicated that all features showed a statistically significant difference among the four scoring subgroups ($P < 0.05$). Representatively, Fig. 7(a)–(c) shows the average number of peaks, average distance between peaks, and average area between peaks in C_H and C_V of all the images in terms of LUSS:$0 \sim 3$, respectively. The accuracies of different neural networks developed under Config-1 and Config-2 were plotted in terms of the computing epoch in Fig. 8(a) and (b), respectively. It is noted that the reported performances of the proposed models were the averages of all the testing results of the cross validation. Statistical analysis showed that the testing results of the cross validation of different models presented a significant difference ($p < 0.001$). In addition, the accuracies of different neural networks trained

TABLE I
FEATURES EXTRACTED FROM C_H, C_V, AND ROI

Feature category	Extracted features
Peak information in horizontal direction	Number of peaks,average distance between peaks, max distance between peaks,min distance between peaks,SD of the distances between peaks
Peak information in vertical direction	Average value of peaks,max value of peaks,min value of peaks,SD of values of peaks
Area under the curve	Area under the curve,average area under the curve between peaks,max area under the curve between peaks,min area under the curvebetween peaks,SD of areas under the curve between peaks
Image feature	Average intensity of pixels inside the ROI,SD of intensity of pixels inside the ROI

under Config-1 and Config-2 were compared in Fig. 8(c). For each neural network, the accuracy of the model trained under Config-2 was higher than that under Config-1. All the models trained under Config-1, including the SVM and decision tree, were compared in Fig. 9, in which the neural network with 128×256 neurons gave the highest accuracy of 87%.

IV. DISCUSSION

In this study, we proposed several classification models for automated scoring of LUSS using the fivefold cross validation. The ultrasound images were collected from 31 COVID-19 PN patients with different conditions and labeled as Scores $0 \sim 3$ by experienced clinicians. A collection of 28 features were predefined and extracted from the ROI and ROI-generated curves for the classification task. Several classification models were trained and evaluated. The performance of the model (two layers with 128×256 hidden neurons) in the testing data demonstrated high accuracy for scoring LUS, indicating the promise for clinical applications.

A. Ultrasound Image Feature Selection

The statistical analysis indicated that each feature used in the study was sensitive to different subgroups and may be used to statistically discriminate images with Scores $0 \sim 3$. As representatives, three features (number of peaks, average distance between peaks, and average area between peaks) were plotted in Fig. 7. The number of peaks in C_H and C_V in Fig. 7(a) potentially mirrored the number of B-lines and A-lines in the ROI, respectively. In Fig. 7(a), significant distinction can be found between Score 0 and Scores 1–3 for the number of A-lines (i.e., peaks in C_V) and between Scores 1 and 2 for B-lines (i.e., peaks in C_H). A-lines demonstrates as parallel lines horizontally, resulting in multiple distinct peaks in C_V, as shown in Figs. 4(a) and 6(a). While displayed as vertical lines/beams, B-lines rarely show peaks in C_V, as shown in Figs. 4(b) and (c) and 6(b) and (c). The significant difference between A-lines and B-lines in the horizontal direction guaranteed the obvious distinction in peak number in C_V for different LUSSs as learned from Fig. 7(a). The same pattern also applied to the vertical direction for both A-lines and B-lines.
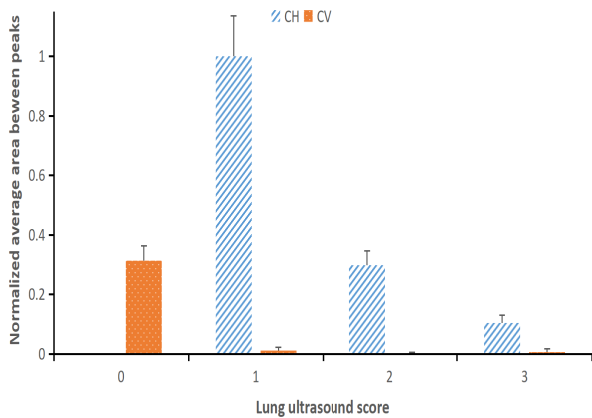
The observations form Fig. 7(a) may also be applied to Fig. 7(b) and (c), which shows the average distances and areas between peaks in C_H and C_V, respectively. It is deserved to
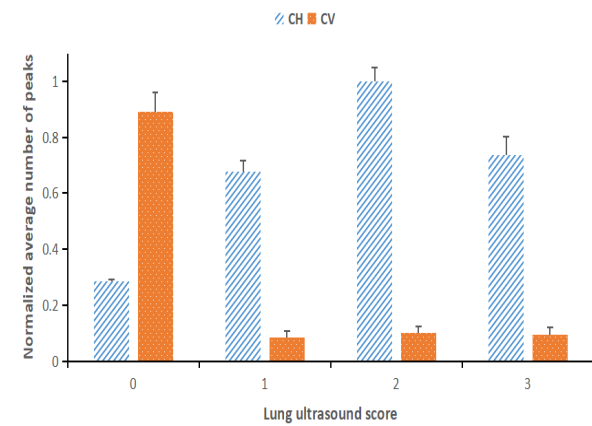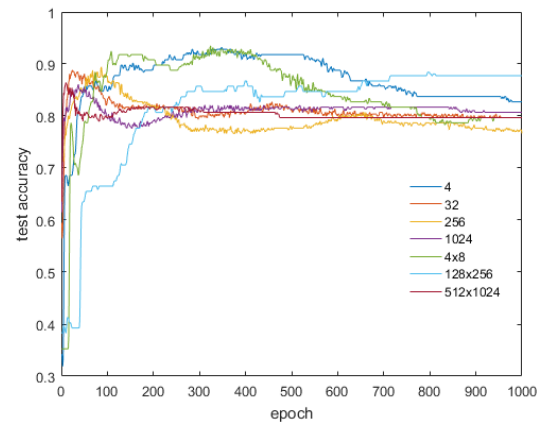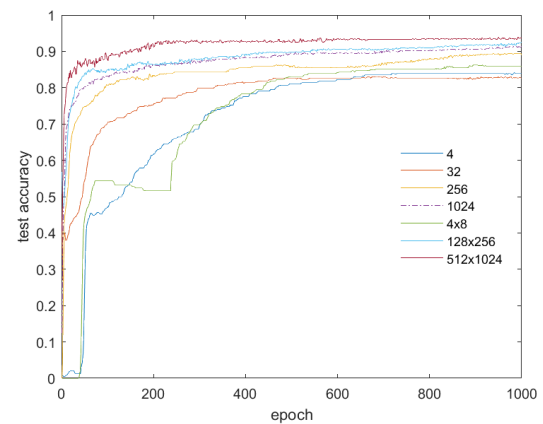
(a)
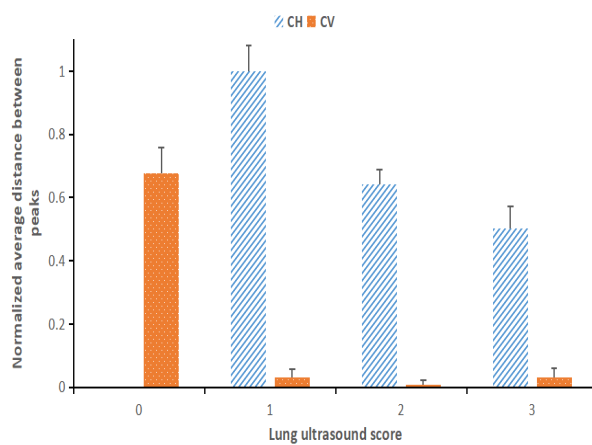
(b)

(c)

Fig. 7. (a) normalized number of peaks, (b) normalized average distance between peaks, and (c) normalized area between peaks in C_H and C_V for different LUSSs.

(a)

(b)

(c)

Fig. 8. Accuracies of different models: (a) models trained under Config-1, (b) models trained under Config-2 (legend shows the number of hidden neurons in each layer of the models), and (c) models under both Config-1 and Config-2.
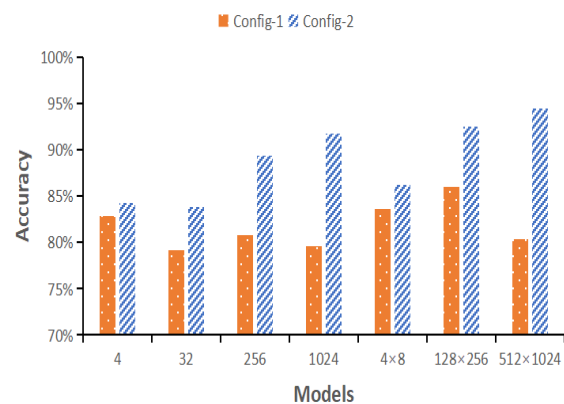
discuss that the average distances and areas between peaks in C_H decreased as the LUSS increased from 1 to 3. Such a phenomenon was because B-lines were getting closer and more confluent when the LUSS increased (i.e., more edema), as shown in Fig. 5(b)–(d). In addition, B-lines rarely had peaks in C_V, resulting in almost zero value for average distance and area between peaks in C_V, as shown in Fig. 7(b) and (c).

Note that the appearance of B-lines may be affected by the imaging frequency and bandwidth of the ultrasound probes/scanners [32]. Although this problem was mitigated in the study since the same probe was used, it is worth noting that counting the artifacts does not provide an absolute measure but only for a relative one. The change in B-line number may be more important than the number itself in the clinical practice.
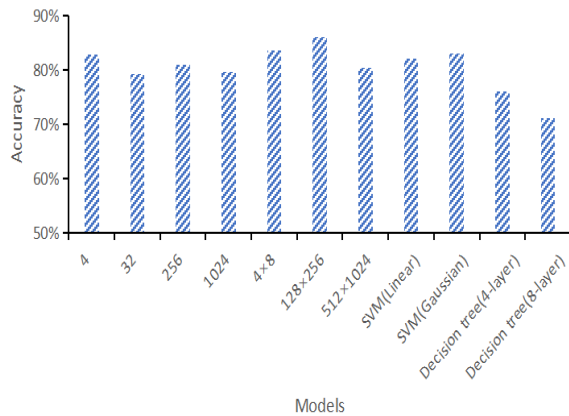
Fig. 9. Accuracies of different models including neural networks, SVM, and decision tree trained under Config-1.

It is noticed from Fig. 8(c) that the accuracy of the model trained under Config-2 was higher than that under Config-1 for each neural network. Such a phenomenon may be attributed to the fact that in Config-2, the training and testing data sets were divided without considering patient information, making those trained under Config-2 overfitting, since the data collected from the same patient may be of high similarity and grouped into the training and testing data sets. With such a consideration, the models trained under Config-1 were preferred. Among all the models trained under Config-1, the neural network with 128×256 neurons gave the highest prediction accuracy of 87% (see Fig. 9).

One innovation of the study is the introduction of features extracted from 1-D curves obtained by projecting the pixels' intensities in the ROI in horizontal and vertical directions. Such an operation is based on the instinctive observation from the LUS images, e.g., B-lines are shown as parallel vertical lines/beams that can be imagined as a curve with several crests if looking from the bottom edge of the image toward the pleural line; the same applies to the A-line if looking from the left to the right side of the image, as shown in Figs. 4–6. In this study, 28 features that were well defined by mimicking the visual observations of the clinicians were extracted for training the classification models. In virtue of such an operation, the proposed model showed good performance based on the simple model structure and limited training data. It is noted that the curve-to-linear conversion is an important process to transfer all the curved images to linear, making the B-lines vertical and parallel to each other. The curve-to-linear conversion made the 2-D problem of the B-line evaluation in curved images to the 1-D problem of the projection curve analysis in linear images.

B. Advantage of Neural Networks for Classification

Classification approaches have been applied to analyze ultrasound images and recently utilized to detect and evaluate the B-lines in an automated manner [20], [23], [33]–[36]. Brusasco *et al.* [20] applied the K-means classification to divide pixels in a filtered image into two subsets, i.e., B-lines

and no B-lines. However, the method was validated on a small number of subjects, weakening its potential for real clinical applications. Brattain *et al.* [23] detected the B-lines via classification based on five features extracted from the angular slices. The limited number of features was not enough to depict the characteristics of the B-lines. In addition, the study did not mention any intelligent models for the classification. Soummer *et al.* [28] developed a sophisticated deep learning model for classification and localization of ultrasound COVID-19 markers by training on both ultrasound frames and videos. The model contained many hidden neurons which required data of a large number and powerful computing capability for training.

In this study, we tried different classification models with artificial feature extraction as an anterior step to undertake the task of the classification with the advantages of: 1) no dependence on arbitrary and human-decided thresholds; 2) capability of getting robust classification models based on limited training data sets; and 3) avoiding overfitting by considering several parameters including the activation function, learning rate, batch size, and epoch. The accuracy of the neural network with 128×256 two fully connected layers trained under Config-1 was as high as 87% after the 700th epoch, indicating that the neural network could provide good performance for scoring LUS with high accuracy and low computation load. The model could then be potentially integrated into mobile ultrasound for clinical applications.

C. Limitations and Future Study

There are several limitations of the study. First, the LUSS used in the study was only focused on the B-lines without considering the subpleural consolidations, which is not specially designed for the COVID-19 PN and may not fully reflect the status of the disease. Second, the study relied on visual determination of the B-score that is not a reliable gold standard. Third, the features to portray the LUS images were limited to 28 predefined parameters, which may be not enough for the current classification task although the proposed model showed good accuracy. In future studies, the scoring system developed specifically for COVID-19 PN will be applied as suggested by Perrone *et al.* [37]. In addition, the proposed technique needs to be improved with validation on reliable gold standard, e.g., using the postmortem gravimetry in animal studies or with the thermodilution in human studies as the reference in the future [20], [38].

V. CONCLUSION

In this study, we proposed an automated scoring method to evaluate the ultrasound images of COVID-19 PN using different classification models. The results unveiled that the proposed method could assign LUSS for LUS images of COVID-19 PN with high accuracy, making it promise for automated LUS scoring. In addition, the proposed automated LUSS model has the potential to be integrated into portable and mobile ultrasound equipment for clinical use in hospitals of different levels as well as prehospital settings such as the ambulance.

ACKNOWLEDGMENT

The authors would like to thank the reviewers in advance for their comments and suggestions.

REFERENCES

[1] P. Mayo *et al.*, "Thoracic ultrasonography: A narrative review," *Intensive care Med.*, vol. 45, pp. 1200–1211, Sep. 2019.

[2] Q.-Y. Peng, X.-T. Wang, and L.-N. Zhang, "Findings of lung ultrasonography of novel corona virus pneumonia during the 2019–2020 epidemic," *Intensive Care Med.*, vol. 46, no. 5, pp. 849–850, May 2020.

[3] G. Soldati *et al.*, "Proposal for international standardization of the use of lung ultrasound for patients with COVID-19: A simple, quantitative, reproducible method," *J. Ultrasound Med.*, vol. 39, no. 7, pp. 1413–1419, Jul. 2020.

[4] F. Mento and L. Demi, "On the influence of imaging parameters on lung ultrasound B-line artifacts, *in vitro* study," *J. Acoust. Soc. Amer.*, vol. 148, no. 2, pp. 975–983, Aug. 2020.

[5] G. Soldati *et al.*, "Is there a role for lung ultrasound during the COVID-19 pandemic?" *J. Ultrasound Med.*, vol. 39, pp. 1459–1462, Jul. 2020.

[6] A. Smargiassi *et al.*, "Lung ultrasound for COVID-19 patchy pneumonia: Extended or limited evaluations?" *J. Ultrasound Med.*, vol. 40, no. 3, pp. 521–528, Mar. 2021.

[7] M. Demi, R. Prediletto, G. Soldati, and L. Demi, "Physical mechanisms providing clinical information from ultrasound lung images: Hypotheses and early confirmations," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 3, pp. 612–623, Mar. 2019.

[8] L. Demi, M. Demi, R. Prediletto, and G. Soldati, "Real-time multifrequency ultrasound imaging for quantitative lung ultrasound–first clinical results," *J. Acoust. Soc. Amer.*, vol. 148, no. 2, pp. 998–1006, Aug. 2020.

[9] R. Clay *et al.*, "Assessment of interstitial lung disease using lung ultrasound surface wave elastography," *J. Thoracic Imag.*, vol. 34, no. 5, pp. 313–319, 2019.

[10] K. Mohanty, J. Blackwell, T. Egan, and M. Müller, "Characterization of the lung parenchyma using ultrasound multiple scattering," *Ultrasound Med. Biol.*, vol. 43, no. 5, pp. 993–1003, May 2017.

[11] F. Mento *et al.*, "On the impact of different lung ultrasound imaging protocols in the evaluation of patients affected by coronavirus disease 2019: How many acquisitions are needed?" *J. Ultrasound Med.*, vol. 9999, pp. 1–4, Nov. 2020.

[12] V. E. Noble, A. F. Murray, R. Capp, M. H. Sylvia-Reardon, D. J. Steele, and A. Liteplo, "Ultrasound assessment for extravascular lung water in patients undergoing hemodialysis: Time course for resolution," *Chest*, vol. 135, no. 6, pp. 1433–1439, 2009.

[13] F. Corradi *et al.*, "Assessment of extravascular lung water by quantitative ultrasound and CT in isolated bovine lung," *Respiratory Physiol. Neurobiol.*, vol. 187, no. 3, pp. 244–249, Jul. 2013.

[14] Y.-H. Jin *et al.*, "A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version)," *Mil. Med. Res.*, vol. 7, no. 1, p. 4, 2020.

[15] K. L. Anderson, J. M. Fields, N. L. Panebianco, K. Y. Jenq, J. Marin, and A. J. Dean, "Inter-rater reliability of quantifying pleural B-lines using multiple counting methods," *J. Ultrasound Med.*, vol. 32, no. 1, pp. 115–120, Jan. 2013.

[16] H. Zong, G. Guo, J. Liu, L. Bao, and C. Yang, "Using lung ultrasound to quantitatively evaluate pulmonary water content," *Pediatric Pulmonol.*, vol. 55, no. 3, pp. 729–739, Mar. 2020.

[17] H. Li *et al.*, "A simplified ultrasound comet tail grading scoring to assess pulmonary congestion in patients with heart failure," *BioMed Res. Int.*, vol. 2018, Jan. 2018, Art. no. 8474839.

[18] L. Gargani *et al.*, "Ultrasound lung comets in systemic sclerosis: A chest sonography hallmark of pulmonary interstitial fibrosis," *Rheumatology*, vol. 48, no. 11, pp. 1382–1387, Nov. 2009.

[19] D. A. Lichtenstein and G. A. Mezière, "Relevance of lung ultrasound in the diagnosis of acute respiratory failure: The BLUE protocol," *Chest*, vol. 134, no. 1, pp. 117–125, Jul. 2008.

[20] C. Brusasco *et al.*, "Quantitative lung ultrasonography: A putative new algorithm for automatic detection and quantification of B-lines," *Crit. Care*, vol. 23, no. 1, pp. 1–7, Dec. 2019.

[21] F. Corradi *et al.*, "Computer-aided quantitative ultrasonography for detection of pulmonary edema in mechanically ventilated cardiac surgery patients," *Chest*, vol. 150, no. 3, pp. 640–651, Sep. 2016.

[22] F. Corradi, G. Via, F. Forfori, C. Brusasco, and G. Tavazzi, "Lung ultrasound and B-lines quantification inaccuracy: B sure to have the right solution," *Intensive Care Med.*, vol. 2, pp. 1–3, Mar. 2020.

[23] L. J. Brattain, B. A. Telfer, A. S. Liteplo, and V. E. Noble, "Automated B-Line scoring on thoracic sonography," *J. Ultrasound Med.*, vol. 32, no. 12, pp. 2185–2190, Dec. 2013.

[24] S. Roy *et al.*, "Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2676–2687, Aug. 2020.

[25] L. Carrer *et al.*, "Automatic pleural line extraction and COVID-19 scoring from lung ultrasound data," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 11, pp. 2207–2217, Nov. 2020.

[26] C. Jalaber, T. Lapotre, T. Morcet-Delattre, F. Ribet, S. Jouneau, and M. Lederlin, "Chest CT in COVID-19 pneumonia: A review of current knowledge," *Diagnostic Interventional Imag.*, vol. 101, nos. 7–8, pp. 431–437, Jul. 2020.

[27] Released by National Health Commission & National Administration of Traditional Chinese Medicine on March 3, "Diagnosis and treatment protocol for novel coronavirus pneumonia (trial version 7)," *Chin. Med. J.*, vol. 133, no. 9, pp. 1087–1095, 2020.

[28] A. Soummer *et al.*, "Ultrasound assessment of lung aeration loss during a successful weaning trial predicts postextubation distress," *Crit. Care Med.*, vol. 40, no. 7, pp. 2064–2072, Jul. 2012.

[29] S. Gao and V. Gruev, "Bilinear and bicubic interpolation methods for division of focal plane polarimeters," *Opt. Exp.*, vol. 19, no. 27, pp. 26161–26173, 2011.

[30] J. Chen, J. Li, C. He, W. Li, and Q. Li, "Automated pleural line detection based on radon transform using ultrasound," *Ultrason. Imag.*, vol. 43, no. 1, pp. 19–28, Jan. 2021.

[31] A. R. McLean, M. E. Richards, C. S. Crandall, and J. L. Marinaro, "Ultrasound determination of chest wall thickness: Implications for needle thoracostomy," *Amer. J. Emergency Med.*, vol. 29, no. 9, pp. 1173–1177, Nov. 2011.

[32] F. Mento, G. Soldati, R. Prediletto, M. Demi, and L. Demi, "Quantitative lung ultrasound spectroscopy applied to the diagnosis of pulmonary fibrosis: The first clinical study," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 11, pp. 2265–2273, Nov. 2020.

[33] S. Liu *et al.*, "Deep learning in medical ultrasound analysis: A review," *Engineering*, vol. 5, no. 2, pp. 261–275, Apr. 2019.

[34] R. J. G. van Sloun and L. Demi, "Localizing B-Lines in lung ultrasonography by weakly supervised deep learning, *in-vivo* results," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 4, pp. 957–964, Apr. 2020.

[35] C. Baloescu *et al.*, "Automated lung ultrasound B-Line assessment using a deep learning algorithm," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 11, pp. 2312–2320, Nov. 2020.

[36] R. Moshavegh, K. L. Hansen, H. Møller-Sørensen, M. B. Nielsen, and J. A. Jensen, "Automatic detection of B-lines in *in vivo* lung ultrasound," *IEEE Trans. Ultrason., Ferroelectr., Freq. control*, vol. 66, no. 2, pp. 309–317, 2018.

[37] T. Perrone *et al.*, "A new lung ultrasound protocol able to predict worsening in patients affected by severe acute respiratory syndrome coronavirus 2 pneumonia," *J. Ultrasound Med.*, vol. 9999, pp. 1–9, Nov. 2020.

[38] S. L. Vrancken *et al.*, "Estimation of extravascular lung water using the transpulmonary ultrasound dilution (TPUD) method: A validation study in neonatal lambs," *J. Clin. Monitor. Comput.*, vol. 30, no. 6, pp. 985–994, Dec. 2016.

Jiangang Chen received the Ph.D. degree from the Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hong Kong, China, in 2011.

He worked as a Postdoctoral Research Scientist at Columbia University, New York, NY, USA, from 2012 to 2013, a Senior Scientist at Philips Research China, Shanghai, China, from 2013 to 2017, and a Senior Research Scientist at the Oxford Suzhou Center for Advanced Research, Suzhou, China, in 2019. He is currently an Associate Professor with the Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai, China. His current research interests include emergency and critical care ultrasound, obstetric ultrasound, handheld ultrasound systems, and artificial intelligence for medical image analysis.

Chao He received the bachelor's degree from the Department of Clinical Medicine, Bengbu Medical College, Bengbu, Anhui, China, in 2005, and the master's degree in emergency medicine from Naval Medical University, China, in 2013.

He has been working at the Emergency Department, Changzheng Hospital, Shanghai, China, since 2005. His research interests include severe ultrasound, comprehensive treatment of central nervous system dysfunction, and sepsis.

Jintao Yin received the B.E. degree in communication engineering from East China Normal University, Shanghai, China, in 2018, where he is currently pursuing the master's degree in communication and information system with a focus on medical imaging and image processing.

Jiawei Li received the Ph.D. degree from the Department of Anesthesia and Intensive Care, The Chinese University of Hong Kong, Hong Kong, in 2012.

In 2014, she joined the Department of Medical Ultrasound, Fudan University Shanghai Cancer Center. She is currently an Associate Professor specialized in oncology ultrasonography. Her research interests include sonographic assessment for breast and thyroid, ultrasound radiomics, and artificial intelligence used for breast cancers.

Xiaoqian Duan received the B.E. degree in communication engineering from the Heilongjiang University of China, Harbin, China, in 2019. She is currently pursuing the master's degree in electronics and communication engineering with East China Normal University, Shanghai, China, with a focus on ultrasound image processing.

Yucheng Cao received the B.E. degree in electronic information science and technology from the Ocean University of China, Qingdao, Shandong, China, in 2019. He is currently pursuing the master's degree in electronics and communication engineering with East China Normal University, Shanghai, with a focus on ultrasound image processing.

Li Sun (Member, IEEE) received the Ph.D. degree from the School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an, China, in 2011.

He worked as a Visiting Researcher of intelligent autonomous systems, Technical University of Munich, Munich, Germany, from 2007 to 2009. From 2011 to 2013, he worked as a Postdoctoral Research Associate with the Institute of Information and Communication Technologies, Electronics and Applied Mathematics, Université Catholique de Louvain, Leuven, Belgium. He is currently an Associate Professor with the Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai, China. His current research interests include computer vision, image processing, and deep learning.
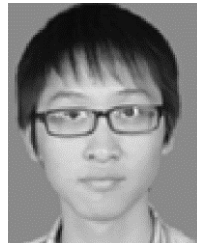
Menghan Hu received the Ph.D. degree (Hons.) in biomedical engineering from the University of Shanghai for Science and Technology, Shanghai, China, in 2016.

From 2016 to 2018, he was a Postdoctoral Researcher with Shanghai Jiao Tong University, Shanghai. He is currently an Associate Professor with the Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai.

Wenfang Li received the Master of Medicine and M.D. degrees from The Second Military Medical University, Shanghai, China, in 1997 and 2010, respectively.

He is currently the Head of the Emergency Department and the Deputy Director of the Chinese People's Emergency Medical Center, Changzheng Hospital, Shanghai, China. His research interests include disaster emergency, shock first-aid, and emergency immediate detection.

Qingli Li (Senior Member, IEEE) received the B.S. and M.S. degrees in computer science and engineering from Shandong University, Jinan, China, in 2000 and 2003, respectively, and the Ph.D. degree in pattern recognition and intelligent system from Shanghai Jiao Tong University, Shanghai, China, in 2006.

Since March 2012, he has been a Visiting Scholar with the Medical Centre, Columbia University, New York, NY, USA. He is currently a Professor with the Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai, where he is engaged in research on medical imaging, pattern recognition, and image processing.