International Journal of Computer Assisted Radiology and Surgery (2023) 18:303–312
https://doi.org/10.1007/s11548-022-02765-2

REVIEW ARTICLE

Automatic diagnosis of pneumothorax with M-mode ultrasound images based on D-MPL

Tao Zhang[1] · Shiju Yan[1] · Gaofeng Wei[2] · Linying Yang[1] · Tianxiang Yu[1] · Yibo Ma[3]

Received: 16 December 2021 / Accepted: 19 September 2022 / Published online: 1 November 2022

Abstract

Purpose To address the difficulties of M-mode ultrasound images classification in pneumothorax diagnosis and the shortcomings of existing neural network algorithms in this field, we proposed an M-mode ultrasound images classification model based on Disturbed Meta-Pseudo-Labels (D-MPL).

Methods An M-mode ultrasound image augmentation system was designed to make the model more robust and generalizable. In D-MPL, teacher-generated pseudo-labeling was first taught to students through a soft mask, and additional disturbance data were added to the teacher network. As the loss of the teacher network continues to decline, disturbance data were injected to improve the generalization of the model to cope with image differences across patients in clinical settings.

Results We compared the proposed model with four commonly used models, including MPL, EfficientnetB2, Inception V3, and Resnet101, in order to confirm its efficacy. Our model has an average specificity of 98.28%, sensitivity of 98.22%, F1-score of 98.23%, and AUC of 98.10%, according to the experiment findings, and its comprehensive performance is better than the above four models.

Conclusion The results demonstrated our model's superiority over the competition and its greater. The model proposed in this study is expected to assist doctors in the diagnosis of pneumothorax as an auxiliary mean.

Keywords Pneumothorax · M-mode ultrasound images · Deep learning · MPL · D-MPL

Introduction

Pneumothorax, a respiratory emergency that poses a life-threatening risk [1, 2], requires immediate diagnosis and treatment [3]. Bedside lung ultrasound is used to visualize the lungs at the bedside and is increasingly employed in critical care, which is widely used for the early diagnosis of traumatic pneumothorax and primary pneumothorax [4]. Ultrasonic signs, such as lung point, are typically used by clinicians to determine whether pneumothorax has occurred [5] so that bedside lung ultrasound can be adopted for rapid diagnosis of pneumothorax with the specificity and sensitivity of 90.9% and 98.2% in the presence of trained clinicians who recognize ultrasound images [6], respectively. Lung sliding abolition and lung point, especially the latter, are the most common signs in the process of diagnosing pneumothorax using ultrasound, and M-mode ultrasound offers distinct benefits for identifying the two types of indications [7]. The existence of lung point (or lung sliding abolition) is verified, and the patient is diagnosed with pneumothorax, if M-mode ultrasound exhibits the beach and streak sign (or stripe sign). If not, the patient is identified as not having a pneumothorax. Therefore, to some extent, the M-mode ultrasound is the gold standard for the diagnosis of pneumothorax. However, the effectiveness of bedside lung ultrasound entirely rests on the expertise doctors, and real-time diagnosis greatly challenges the doctor's ability to continuously and subtly observe. Since the ultrasonic pneumothorax image is not typical enough and sometimes the imaging is not clear enough, it is difficult to process effectively, and there are

✉ Shiju Yan
yanshiju@usst.edu.cn

✉ Gaofeng Wei
highpeak8848@163.com

1 School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

2 Naval Medical Department, Naval Medical University, Shanghai 200433, China

3 Department of Ultrasound, The Third Affiliated Hospital of Soochow University, Changzhou 213003, Jiangsu, China

few effective ultrasonic image classification algorithms for pneumothorax diagnosis [6]. Tony Lindsey et al. used the Bayesian approach, pre-trained architectures, and applied transfer learning to achieve 98.3% accuracy in M-mode classification [8]. Courosh Mehanian et al. used deep learning architecture to make the AUC of pneumothorax detection reach 0.83 [9]. Although simply classifying pneumothorax and non-pneumothorax and using data from porcine lung ultrasonography, their study served as a valuable resource for the researchers.

Deep learning has achieved great success in the field of medical image processing [10]. Nevertheless, the annotation of numerous medical photographs is still a time-consuming undertaking for medical professionals. To make up for the shortage of labeled data, semi-supervised learning has gained increasing traction [11, 12]. Meta-learning [13] may be utilized to enhance the generalization capabilities of the model and create a generic classification algorithm model [14] to make up for the data scarcity brought on by few samples. Meta-pseudo-labels (MPL) [15] have made an important breakthrough in the construction of general algorithm models. MPL has a teacher network for generating pseudo-labels on unlabeled data to teach the student network, and the learning performance of the student on the labeled data is fed back to the teacher network as a reward, and teacher network will generate a better pseudo-label to teach the student network. In the experiments using MPL for image classification, the performance of MPL still has much space for improvement. In this study, we proposed an efficient and generalized Disturbance-meta-pseudo-labels (D-MPL) and the soft mask and disturbance data [16] were put on top of the MPL training to provide the model with a stronger capacity for learning. The model training is time consuming, but the model testing is very efficient for detecting pneumothorax. The infer speed of the trained model is very high, and the time consumed by inferring a single image is less than 15 ms, which is much shorter than the time for a doctor to read a single M-mode ultrasound image.

Materials and methods

Data

The data were collected in collaboration with the doctors in the Third Affiliated Hospital of Soochow University. A total number of 204 patients were included from July 2020 to November 2021. We organized 2000 M-mode ultrasound images for experiments from these patients. The patients' age ranged from 5 to 60 years. It consisted of three categories, namely lung point (pneumothorax) [17], lung sliding abolition (pneumothorax) [18], and lung sliding (normal) [19]. The images of each type are classified into two subcategories, typical and atypical, by experienced ultrasound physicians according to whether the sign was typical or atypical. Fig. 1a and b shows atypical and typical images of lung point, c and d are atypical and typical images of lung sliding abolition, and e and f are atypical and typical images of lung sliding, respectively. The red box indicates the location of the corresponding pleural line. In the process of collecting data, the cooperating doctors acquired more data from patients with lung sliding abolition and lung point and fewer data from the patients without pneumothorax. Balancing these three types of data will be considered in future studies.

Methods

In this section, we firstly introduced a systematic augment of the M-mode ultrasound images, followed by D-MPL and a detailed description of the training process for D-MPL.

Our experimental sample size was small and the typicality of image signs varied. We designed a systematic scheme for M-mode ultrasound images augmentation for better generality. We used 450 of 2000 original images for testing and did not include them in the data preprocessing. The rest of the images (1550) were used for data expansion. As shown in Fig. 2, the original images were subjected to denoising, amplification and adding noise processing, respectively. Firstly, DNCNN [20], wavelet transform [21], and Gaussian smoothing denoising [22] were performed on the original data, followed by mirror transform [23], random clipping [24] and the resulting images were taken as the training set. Finally, the original images were subjected to additive noise [25], multiplicative noise [26] processing, and the processed images were taken as disturbance data. After the data expansion, 3550 groups of data for training and 450 groups of data for testing were finally generated, with the same labels as the original data.

Figure 3 is a schematic diagram showing the denoising principle of DNCNN for M-mode ultrasound images. The original images were input to DNCNN, and the generated residual images were taken as the approximate noise. After continuous learning and training, the noise distribution of the original images was obtained. Finally, the generated noise images were subtracted from the original images to obtain the denoised images by DNCNN.

MPL is a semi-supervised model for teacher network and student network learning, where pseudo-labels generated by teacher network are directly taught to the student network, which in turn provides feedback to the teacher network. In this study, we proposed an improved D-MPL, which was designed based on MPL. In D-MPL, the pseudo-labels generated by the teacher were not directly taught to the student, but via a soft mask. Besides, additional disturbance data were added to the teacher network. As the loss of the teacher model continued to decline, disturbance data were injected

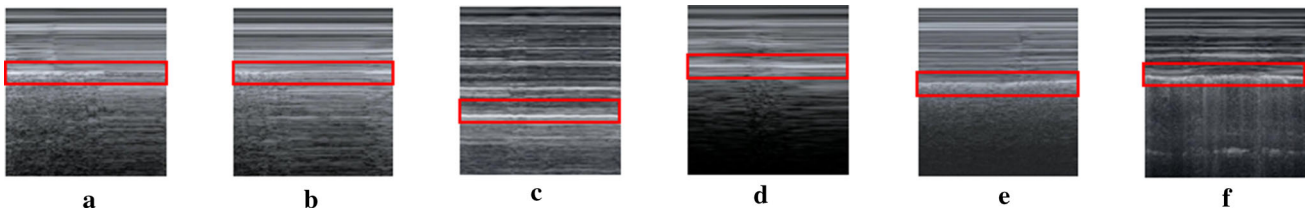

Fig. 1 M-mode ultrasound images. **a** and **b** are atypical and typical images of lung point, respectively, **c** and **d** are atypical and typical images of lung sliding abolition, respectively, and **e** and **f** are atypical and typical images of lung sliding, respectively

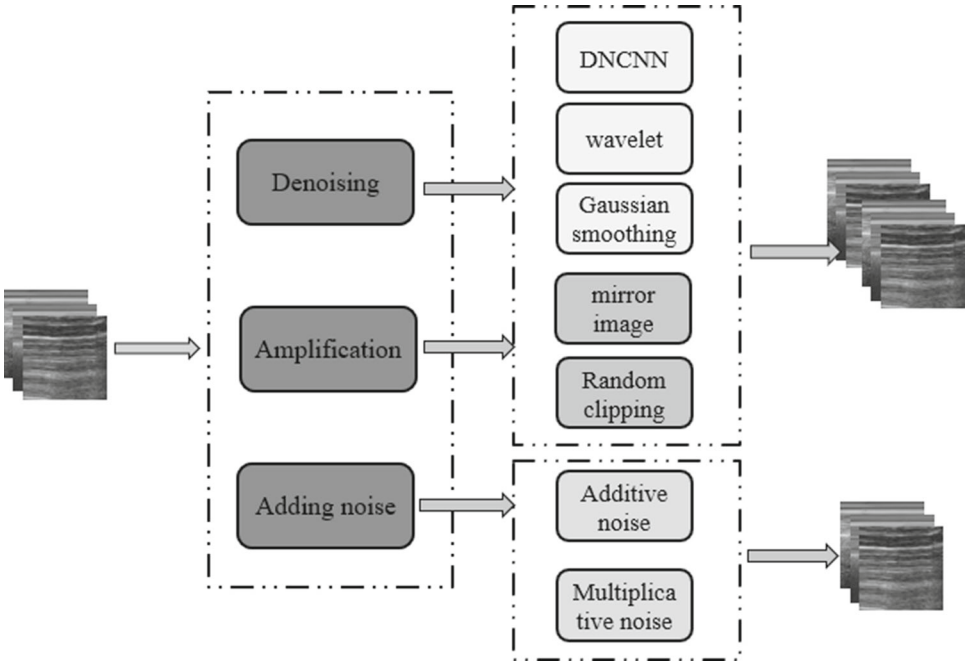

Fig. 2 M-mode ultrasound images preprocessing flow

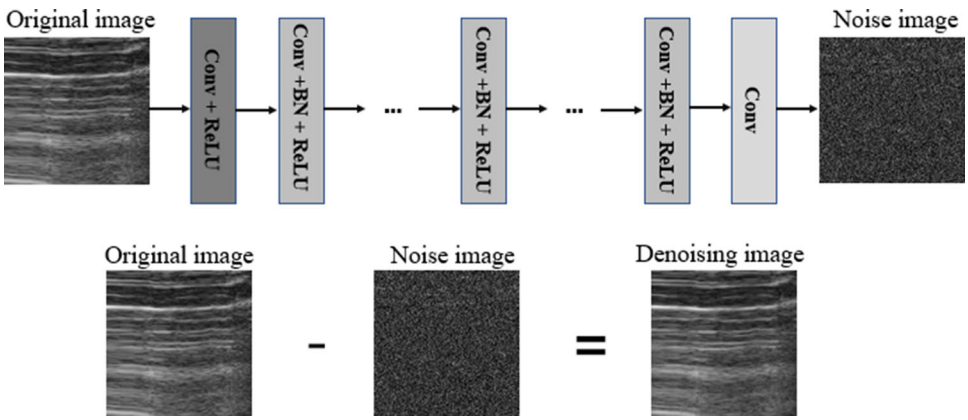

Fig. 3 Schematic diagram of DNCNN noise removal

to improve the generalization of the model to address more complex situations in clinical practice. Figure 4 shows the difference between MPL and D-MPL. In D-MPL, the soft mask is present to filter out the less certain information from the knowledge taught to the students by the teacher, as we want the student network to acquire more certain knowledge early in the learning process. In the later stage of training, the soft mask can be turned off or reduced to filter the uncertain information in the teacher network, because at this point the student network has a strong classification ability, and we hope that the student network can receive more difficult knowledge from the teacher network. This is very in line with the reality of the teacher's teaching process and student learning process. In the process of teaching the student network by the teacher network, if the loss continues to decline, a round

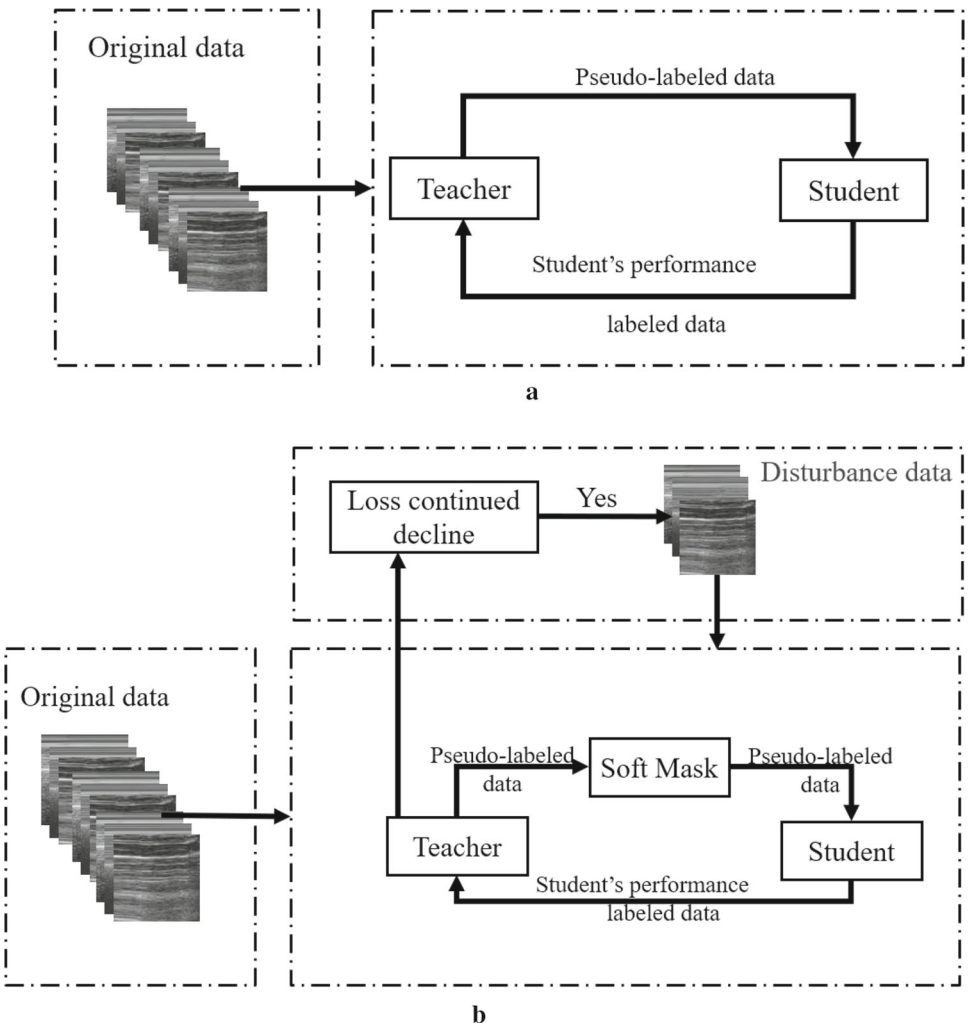

Fig. 4 Differences between MPL and D-MPL **a**. MPL **b**. D-MPL

of disturbance data is injected to improve the learning difficulty of the student network. Specifically, the disturbance data serve to prevent the student network from having difficulty satisfying the learning process, so it may be difficult for the student network to differentiate once challenging questions appear on the test. This is also in line with the reality of the teacher teaching process, and the teacher will not always let the students go smoothly in their studies. Occasionally, the teacher will give some small questions to test the students.

Train processing

In D-MPL, T denotes the teacher network, S denotes the student network, (x_l, y_l) denotes labeled data,x_l denotes a training sample, y_l denotes a label corresponding to x_l, x_u denotes unlabeled data. θ_{T} denotes the parameters of the teacher network,θ_S denotes the parameters of the student network, $T(x_l, \theta_T)$ denotes the prediction labels of the teacher network on the labeled data set, and $S(x_l, \theta_S)$ denotes the prediction labels of the student network on the labeled data set. $T(x_u, \theta_T)$ represents the prediction pseudo-labels of the teacher network on the unlabeled dataset, and $S(x_u, \theta_S)$ represents the prediction labels of the student network on the unlabeled dataset. $CE(y_l, T(x_l, \theta_T))$ represents the loss of the teacher network, $CE(y_l, S(x_l, \theta_S))$ represents the loss of the student network on the labeled data, and $CE(T(x_u, \theta_T), S(x_u, \theta_S))$ represents the loss of the student network on the pseudo-labeled data generated by the teacher network. We further process the pseudo-labels generated by the teacher network. When the confidence level of the pseudo-labels generated by the teacher network is greater than 0.7, the generated pseudo-labels are passed to the student network. The student network needs as much precise information as possible from the teacher at the beginning of learning since it lacks classification capacity and using the student network below the 0.7 thresholds also performed best in the test. At the later stage of the training, the threshold can be appropriately reduced or directly closed, because at that time the student network already has a strong classification ability. This is congruent with how students learn, allowing

them to receive more straightforward data during early network training and more challenging data during later training. So $T(x_u, \theta_T)$ is rewritten as $T(x_u, \theta_T) * mask$. Mask is 1 when the confidence level of $T(x_u, \theta_T)$ is greater than 0.7 and is 0 when $T(x_u, \theta_T)$ is less than 0.7.

The parameter update of the student network depends only on the pseudo-label data generated by the teacher network. The feedback given by the student network to the teacher network is from the loss of the student network on the labeled data in the previous round minus the loss of the current student network on the labeled data. The purpose of this operation is to determine whether the student is in the online learning process. If there is progress, the difference must be positive and then, positive feedback is given to the teacher network. If there is no progress in this round of learning, then the network will provide negative feedback to the teacher. The loss is calculated based on the labeled data, not the pseudo-label, because we want the student network to test on real data. The parameters of the teacher network are updated depending on the loss of the teacher network's labeled data and the feedback from the student network. The above operation is very in accordance with the actual relationship between teachers and students.

In the training process, the pseudo-labels $T(x_u, \theta_T)$ generated by the teacher network are first calculated. Then, we calculate$S(x_l, \theta_S)$, through the expression (1) calculation to get the optimal$\theta^{'}{}_S$:

$$\theta'_S = argmin_{\theta_S} CE(T(x_u, \theta_T) * mask, S(x_u, \theta_S)) \tag{1}$$

Thus, the loss value $\mathcal{L}(x_l, \theta_S)$ of the student network on the labeled data can be obtained through the expression (2):

$$\mathcal{L}_S(x_l, \theta_S) = CE(y_l, S(x_l, \theta_S)) - CE\big(y_l, S\big(x_l, \theta'_S\big)\big) \tag{2}$$

The student network feeds back the loss value $\mathcal{L}(x_l, \theta_S)$ to the teacher network so that the teacher network can better teach the beneficial knowledge to the student network. In order to better adjust the network, we also add an adjustable super parameter α to $\mathcal{L}(x_l, \theta_S)$. The parameter θ_T of the teacher network can be updated by Expression (3):

$$\theta_T = argmin_{\theta_T}(CE(y_l, T(x_l, \theta_T)) + \mathcal{L}_S(x_l, \theta_S) * \alpha) \tag{3}$$

If $CE(y_l, T(x_l, \theta_T)) + \mathcal{L}_S(x_l, \theta_S) * \alpha$ is found to continuously decline after ten rounds of continuous training, one round of labeled disturbance data (x_d, y_d) will be added. In addition to generalize the teacher and student networks, this also prevents the network performance deterioration brought on by adding too much disturbance data directly. The following expression is for the teacher network updating with the disturbed data (4):

$$\begin{aligned} &\mathcal{L}_S(x_d, \theta_S) = CE(y_d, S(x_d, \theta_S)) - CE\big(y_d, S\big(x_d, \theta'_S\big)\big), \\ &\theta_T = argmin_{\theta_T}(CE(y_d, T(x_d, \theta_T)) + \mathcal{L}_S(x_d, \theta_S) * \alpha) \end{aligned} \tag{4}$$

In this study, Adam serves as the loss function optimizer. 500 epochs are iterated with an initial value of 0.001 using the cosine annealing learning rate. We may extract the teacher network and the student network—the latter is utilized for classification—through that train of the two networks.

Experiments and results

Experimental setup

Konica Minolta SONIMAGE HS1 color ultrasonic diagnostic instrument was used for examination and acquiring ultrasonic images, and the probe frequency is 4–18 MHz. Our workstation for model training was DELL PC, equipped with Nvidia RTX3090(24G) GPU. The programming language was python 3.7, and pytorch 1.8 was used as the algorithm implementation framework.

Experimental method

To verify the classification performance of the model proposed in this study, the data were divided into a typical group (A) and an atypical group (B) by the doctors according to the typical degree of the image signs. The typical group (A) and the atypical group (B) were then merged into a mixed group (C). Each group included three types of images to be classified, namely, the existence of lung sliding (0), lung point (1), and lung sliding abolition (2). As shown in Table 1, the test set was obtained from the reserved 450 images in the Data Pre-section. The training set and the validation set were obtained from the 3550 images obtained after pretreatment. We divided the 3,550 images into a training set and a validation set at a 3:1 ratio for comparative trials.

Performance indexes

In this study, we used specificity, sensitivity [27], F1-score, and AUC [28] as evaluation indicators. Sensitivity is how many of all affected people obtained a positive result. Specificity refers to how many of all people without the disease get a negative result. F1-score is the harmonic average of precision rate and recall rate, which is used to measure the robustness of the model with a maximum value of 1 and a minimum value of 0. AUC is the area under the ROC curve and can be used to evaluate the pros and cons of the classifiers with a larger AUC value indicating better classifier

Table 1 Experimental scheme design and data division

Group	Class	Train	Validation	Test	Total
A	0	225	75	31	331
	1	600	200	86	886
	2	375	125	83	583
B	0	150	50	23	223
	1	800	250	121	1171
	2	525	175	106	806
C	0	375	125	54	554
	1	1400	450	207	2057
	2	900	300	189	1389

Table 2 Classified evaluation metrics

Assessments	Formula
Sensitivity (Sen)	$\frac{TP}{TP+FN}$
Specificity (Spe)	$\frac{TN}{TN+FP}$
F1-score	$2 \times \frac{P \times R}{P+R}$
AUC	$\frac{\sum_{ins_i \in positiveclass} rank_{ins_i} - \frac{M \times (M+1)}{2}}{M \times N}$

performance. The calculation formulas of these indicators are shown in Table 2.

Results

We firstly performed a detailed verification of D-MPL' s performance (both D-MPL' s teacher and student networks are made up of EfficientnetB2). As shown in Table 3, D-MPL performed well in the typical group (A), the atypical group (B), and the mixed group (C). The sensitivity to lung sliding was 100% in the typical and mixed groups, which was higher than that in the atypical group (more than 90.00%). The typical group, the atypical group, and the mixed group also showed high specificity for the three types of M-mode ultrasound images. The F1-score for lung sliding was 87.50% in the atypical group, and lower than 93.94% in the other categories. The F1-score for lung point reached 98.80% for the mixed group. Nearly all indicators were higher for the typical group than in the atypical group, except that the sensitivity for lung sliding abolition for the typical group was lower than that in the atypical group. We also found that the three indicators (Specificity, Sensitivity, F1-Score) of the three types of images in the mixed group were higher than those in the typical group and the atypical group.

The proposed model was compared with the existing typical algorithms, including MPL, EfficientnetB2 [29], Inception_V3 [30], Resnet101 [31]. Table 4 shows classification performance indicators of the five models, including mean sensitivity, mean specificity, mean F1-score, and AUC. In terms of average specificity, our model was superior to the other four models in all three groups (97.20%, 92.33%, and 98.28%, respectively). Our model outperformed the other models in terms of mean sensitivity in both the typical and atypical groups (97.00% and 94.55%, respectively), except for the mixed group where the values were lower than MPL and EfficientnetB2. In terms of mean F1-score, our model was also superior to other models in all three groups (97.03%, 93.19%, and 98.23%, respectively). The results also demonstrated that our model's AUC values were greater than those of the other four models (96.42%, 93.51%, and 98.10%, respectively). As shown in Fig. 5, the model with the best performance in each evaluation index is visualized by histogram. For the group A, B and C, except that the mean sensitivity in the group C experiment was lower than the EfficientnetB2, our D-MPL had obvious advantages at the four evaluation indexes.

To verify the impact of our proposed data augmentation method on different classification models, we conducted three groups of controlled experiments, with F1-score as the evaluation index. Group A was a typical group with data augmentation, and group A' was a typical group without data augmentation, group B and group B' were atypical groups with and without using data augmentation, and group C and group C' were mixed groups with and without using data augmentation. As shown in Fig. 6, F1-scores of the models using the proposed data augmentation method were obviously higher than to those of the models without using data augmentation on both typical and atypical group. Models that used data augmentation also outperformed models that did not use data augmentation in the mixed groups. The experimental results show that the proposed data augmentation method can be used to improve the performance of the classification models.

Table 3 The specificity, sensitivity and f1-score of D-MPL in the test set

Group	Class	Spe (%)	Sen (%)	F1-score (%)
A	0	88.57	100.00	93.94
	1	98.82	97.67	98.25
	2	98.75	95.18	96.93
B	0	84.00	91.30	87.50
	1	95.87	95.87	95.87
	2	97.12	95.28	96.19
C	0	93.10	100.00	96.43
	1	98.56	99.03	98.80
	2	99.46	96.83	98.12

Table 4 Comparison of classification results of different models on test sets

Group	Model	Avg. Sep (%)	Avg. Sen (%)	Avg. F1-score (%)	AUC (%)
A	D-MPL	**97.20**	**97.00**	**97.03**	**96.42**
	MPL	96.03	95.50	95.63	95.14
	EfficientnetB2	94.57	95.77	95.14	95.03
	Inception_V3	87.94	92.18	89.34	92.22
	Resnet101	88.28	91.91	89.46	92.31
B	D-MPL	**92.33**	**94.55**	**93.19**	**93.51**
	MPL	91.25	93.29	91.00	92.89
	EfficientnetB2	90.16	94.42	92.00	92.66
	Inception_V3	85.07	91.79	87.44	91.89
	Resnet101	85.38	89.79	87.22	89.67
C	D-MPL	**98.28**	98.22	**98.23**	**98.10**
	MPL	97.06	98.30	97.65	97.73
	EfficientnetB2	96.83	**98.80**	97.75	97.64
	Inception_V3	95.63	96.71	96.13	96.25
	Resnet101	94.17	96.04	95.04	96.13

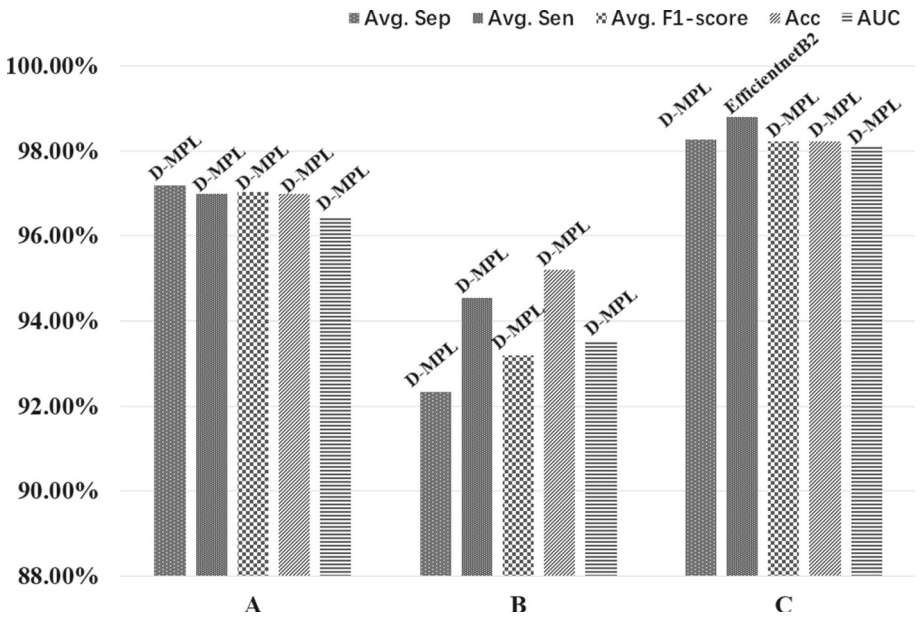

Fig. 5 Performance comparison of different models

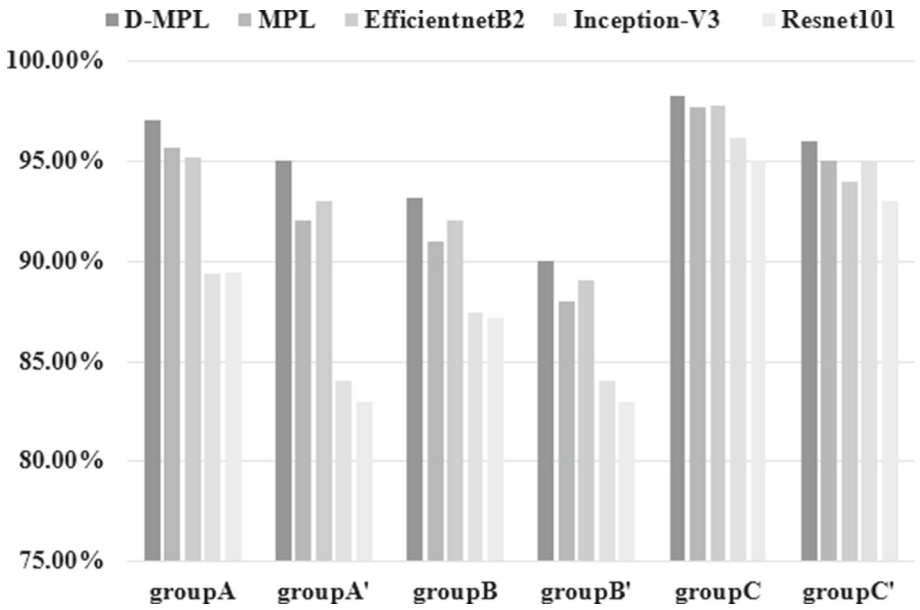

Fig. 6 Influence of data augmentation on model performance

Discussion

In this study, we conducted an in-depth study on the accurate classification of the M-mode ultrasound images. First, we proposed a set of methods to enhance M-mode ultrasound image data. Second, an improved D-MPL was proposed to classify M-mode ultrasound images. The performance of the proposed D-MPL model and the applicability of the data enhancement method were verified by comparison experiments.

In the data enhancement section, we performed a systematic enhancement of M-type ultrasound images, showing that the five models using our proposed enhancement method were of higher performance. By analyzing the reasons, we concluded that the proposed data enhancement method was well adapted to the characteristics of the original data, the total amount of samples was amplified and the characteristics of the samples were more diverse, which enables the models to learn more characteristics of the images and thus, enhance the robustness of the models.

To classify M-mode ultrasound images accurately, the D-MPL model was proposed in this study, which inherits the excellent performance of MPL and further enhances the classification accuracy of the model. In MPL, the teacher network and the student network collaborate to learn, improving the categorization performance. In contrast to the MPL, the proposed D-MPL incorporates a soft mask and a data disturbance mechanism. The soft mask can be used to filter the poor pseudo-labels generated by the teacher network so that the student network can better learn the knowledge taught by the teacher network. When loss consistently decreased during D-MPL training, the disturbance data were introduced into the network, enabling the generalization of the model to deal with clinically complicated and unpredictable circumstances.

In order to verify the performance of the D-MPL and the applicability of the data augmentation method proposed in this study, a number of experiments were conducted. In the D-MPL performance verification, we performed a systematic evaluation of the D-MPL' s performance. The specificity of our approach might eliminate individuals without pneumothorax by up to 99.46%, which indicates that medical resources might be avoided. Because of the high sensitivity and low rate of missed diagnoses, our approach may be used to more accurately identify patients with pneumothoraxes and avoid delaying the onset of therapy. F1-score can reach up to 98.80%, which indicates good comprehensive performance as well as good robustness and applicability. D-MPL also offers several benefits over other models when comparing the performance of various models indicating an Avg. Sep up to 98.28%, Avg. Sen up to 98.22%, Avg. F1-score up to 98.23%, Acc up to 98.22%, and AUC up to 98.10%. We also found that the precision of the typical group was higher than that of the atypical group, and the model precision on the mixed group was higher than that of the typical group and the atypical group, whether data augmentation was performed or not. Therefore, we further concluded that training models using typical data are superior to using atypical data, and training models using a large number of M-type ultrasound images mixed are superior to using only typical or atypical data. We supposed that there are two main reasons for this phenomenon: 1. the limited number of images used for the experiment and 2. the imbalance between the number of typical images and the number of atypical images. These reasons may cause that the model learns more features from the mixed group data.

The data we used in this study are not sufficiently balanced. During the process of data collection from different patients, the doctors collected more data from patients with lung sliding abolition and lung point. In future studies, we will consider balancing the three categories of data used in the study to make it more suitable for clinical diagnosis. In this study, the diagnosis of pneumothorax was primarily performed using M-mode ultrasound. And we will continue to investigate a new algorithm that combines M-mode ultrasound and B-mode ultrasound to diagnose pneumothorax. To reach higher diagnostic standards for clinical pneumothorax, we will continue to experiment with more data augmentation approaches and more models for control trials in future research while also optimizing the D-MPL.

Conclusion

In this study, D-MPL and enhancement of the M-mode ultrasound images were proposed. The D-MPL was compared with MPL, EfficientnetB2, Inception_V3, and Resnet101 in the experiments showing good robustness. The comprehensive performance of our model was superior to four models at four indicators such as average sensitivity, average specificity, average F1-score, and AUC. The proposed method and model are expected to assist doctors in the diagnosis of pneumothorax as an auxiliary mean.

Acknowledgements M-mode ultrasound image data used in this study are all from the Third Affiliated Hospital of Soochow University.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all participants included in the study.

References

1. Duclos G, Bobbia X, Markarian T, Muller L, Cheyssac C, Castillon S, Ressequier N, Boussuques A, Volpicelli G, Leone M, Zieleskiewicz L (2019) Speckle tracking quantification of lung sliding for the diagnosis of pneumothorax: a multicentric observational study. Intensive Care Med 45(9):1212–8. https://doi.org/10.1007/s00134-019-05710-1
2. Weissman J, Agrawal R (2021) Dramatic complication of pneumothorax treatment requiring lifesaving open-heart surgery. Radiol Case Rep 16:500–3
3. Lichtenstein DA (2015) BLUE-protocol and FALLS-protocol: two applications of lung ultrasound in the critically ill. Chest 147(6):1659–1670
4. Rovida S, Orso D, Naeem S, Vetrugno L, Volpicelli G (2022) Lung ultrasound in blunt chest trauma: a clinical review. Ultrasound 30(1):72–79
5. Bouhemad B, Zhang M, Lu Q, Rouby J-J (2007) Clinical review: Bedside lung ultrasound in critical care practice. Crit Care 11(1):205
6. Alrajhi K, Woo MY, Vaillancourt C (2012) Test characteristics of ultrasonography for the detection of pneumothorax: a systematic review and meta-analysis. Chest 141(3):703–708
7. Santos-Silva J, Lichtenstein D, Tuinman PR, Elbers PW (2019) The lung point, still a sign specific to pneumothorax. Intensive Care Med 45(9):1327–1328
8. Lindsey T, Lee R, Grisell R, Vega S, and Veazey S (2018) Automated pneumothorax diagnosis using deep neural networks. In: Iberoamerican congress on pattern recognition (pp 723–731). Springer, Cham. Available from: https://doi.org/10.1007/978-3-030-13469-3_84.
9. Mehanian C, Kulhare S, Millin R, Zheng X, Gregory C, Zhu MSS (2019) Deep learning-based pneumothorax detection in ultrasound videos. In: Smart ultrasound imaging and perinatal, preterm and paediatric image analysis (pp 74–82). Springer, Cham. doi.org/https://doi.org/10.1007/978-3-030-32875-7_9.
10. Lundervold AS, Lundervold A (2019) An overview of deep learning in medical imaging focusing on MRI. Z Med Phys 29(2):102–27
11. Singh A (2021) Clda: contrastive learning for semi-supervised domain adaptation. Adv Neural Inform Process Syst 34:5089–5101
12. Ding R, Zhou Y, Xu J, Xie Y, Liang Q, Ren H, Wang Y, Chen Y, Wang L, Huang M (2021) Semi-supervised optimal transport with self-paced ensemble for cross-hospital sepsis early detection. arXiv preprint arXiv:2106.10352.
13. Wang JX (2021) Meta-learning in natural and artificial intelligence. Curr Opin Behav Sci 38:90–5
14. Peng H (2021) A Brief Summary of Interactions Between Meta-Learning and Self-Supervised Learning. arXiv preprint arXiv:2103.00845.
15. Pham H, Dai Z, Xie Q, and Le QV (2021) Meta pseudo labels. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp 11557–11568).
16. Yang W, Zhou Y, Hu M, Wu D, Zheng JX, Wang H, Guo S. (2021) Gain without Pain: offsetting DP-injected Nosies Stealthily in Cross-device Federated Learning. IEEE Internet of Things Journal.
17. Lichtenstein D, Mezière G, Biderman P, Gepner A (2000) The "lung point": an ultrasound sign specific to pneumothorax. Intensive Care Med 26(10):1434–1440
18. Lenoir V, Kohler R, Montet X (2013) The empty azygos fissure. J Radiol Case Rep 7(4):10–15
19. Oizumi H, Kato H, Endoh M, Inoue T, Watarai H, Sadahiro M (2014) Slip knot bronchial ligation method for thoracoscopic lung segmentectomy. Ann Thorac Surg. 97(4):1456–8
20. Zhang K, Zuo W, Chen Y, Meng D, Zhang L (2017) Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. IEEE Trans Image Process 26(7):3142–3155
21. Lang M, Guo H, Odegard JE, Burrus CS, Wells RO (1996) Noise reduction using an undecimated discrete wavelet transform. IEEE Signal Process Lett 3(1):10–12
22. Tay MKC, Laugier C (2008) Modelling Smooth Paths Using Gaussian Processes. In: Laugier C, Siegwart R (eds) Field and Service Robotics. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 381–390. https://doi.org/10.1007/978-3-540-75404-6_36

23. Stach S, Giurfa M (2001) How honeybees generalize visual patterns to their mirror image and left–right transformation. Anim Behav 62(5):981–91
24. Wang M, Luo C, Hong R, Tang J, Feng J (2016) Beyond object proposals: random crop pooling for multi-label image recognition. IEEE Trans Image Process 25(12):5678–5688
25. Smilkov D, Thorat N, Kim B, Viégas F, and Wattenberg M (2017). Smoothgrad: removing noise by adding noise. arXiv preprint arXiv: 1706.03825.
26. Buades A, Coll B, Morel JM (2005) A review of image denoising algorithms, with a new one. Multiscale Model Simul 4(2):490–530
27. Gelman A, Carpenter B (2020) Bayesian analysis of tests with unknown specificity and sensitivity. J Roy Stat Soc: Ser C (Appl Stat) 69(5):1269–1283
28. Fawcett T (2006) An introduction to ROC analysis. Pattern Recogn Lett 27(8):861–874
29. Tan M, and Le Q (2019) Efficientnet: rethinking model scaling for convolutional neural networks. In: International conference on machine learning (pp 6105–6114). PMLR.
30. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, and Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition (pp 2818–2826).
31. He K, Zhang X, Ren S, and Sun J (2016) Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp 770–778).