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

Medical Image Analysis 69 (2021) 101975

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Modality alignment contrastive learning for severity assessment of COVID-19 from lung ultrasound and clinical information

Wufeng Xue[a,b,1], Chunyan Cao[c,d,1], Jie Liu[c,d], Yilian Duan[c,d], Haiyan Cao[c,d], Jian Wang[a,b], Xumin Tao[a,b], Zejian Chen[a,b], Meng Wu[e], Jinxiang Zhang[f], Hui Sun[g], Yang Jin[h], Xin Yang[a,b], Ruobing Huang[a,b], Feixiang Xiang[c,d], Yue Song[c,d], Manjie You[c,d], Wen Zhang[c,d], Lili Jiang[c,d], Ziming Zhang[c,d], Shuangshuang Kong[c,d], Ying Tian[c,d], Li Zhang[c,d], Dong Ni[a,b,*], Mingxing Xie[c,d,**]

[a] School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China
[b] National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, China
[c] Department of Ultrasound, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[d] Hubei Province Key Laboratory of Molecular Imaging, China
[e] Department of Ultrasound, Zhongnan Hospital of Wuhan University, Wuhan, China
[f] Department of Emergency Surgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[g] Department of Endocrinology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[h] Department of Respiratory, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

ARTICLE INFO

Article history:
Received 3 June 2020
Revised 17 December 2020
Accepted 15 January 2021
Available online 20 January 2021

MSC:
00-01
99-00

Keywords:
Lung ultrasound
Multiple instance learning
Multi-modality
Contrastive learning

ABSTRACT

The outbreak of COVID-19 around the world has caused great pressure to the health care system, and many efforts have been devoted to artificial intelligence (AI)-based analysis of CT and chest X-ray images to help alleviate the shortage of radiologists and improve the diagnosis efficiency. However, only a few works focus on AI-based lung ultrasound (LUS) analysis in spite of its significant role in COVID-19.

In this work, we aim to propose a novel method for severity assessment of COVID-19 patients from LUS and clinical information. Great challenges exist regarding the heterogeneous data, multi-modality information, and highly nonlinear mapping. To overcome these challenges, we first propose a dual-level supervised multiple instance learning module (DSA-MIL) to effectively combine the zone-level representations into patient-level representations. Then a novel modality alignment contrastive learning module (MA-CLR) is presented to combine representations of the two modalities, LUS and clinical information, by matching the two spaces while keeping the discriminative features. To train the nonlinear mapping, a staged representation transfer (SRT) strategy is introduced to maximumly leverage the semantic and discriminative information from the training data.

We trained the model with LUS data of 233 patients, and validated it with 80 patients. Our method can effectively combine the two modalities and achieve accuracy of 75.0% for 4-level patient severity assessment, and 87.5% for the binary severe/non-severe identification. Besides, our method also provides interpretation of the severity assessment by grading each of the lung zone (with accuracy of 85.28%) and identifying the pathological patterns of each lung zone. Our method has a great potential in real clinical practice for COVID-19 patients, especially for pregnant women and children, in aspects of progress monitoring, prognosis stratification, and patient management.

1. Introduction

To date, the COVID-19 pandemic has led more then 200 million confirmed cases and 743,487 reported deaths as of August

* Corresponding author at: School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China.
** Corresponding author at: Department of Ultrasound, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
E-mail addresses: xuewf@szu.edu.cn (W. Xue), nidong@szu.edu.cn (D. Ni), xiemx@hust.edu.cn (M. Xie).

[1] They contribute equally important to this work.

https://doi.org/10.1016/j.media.2021.101975
1361-8415/

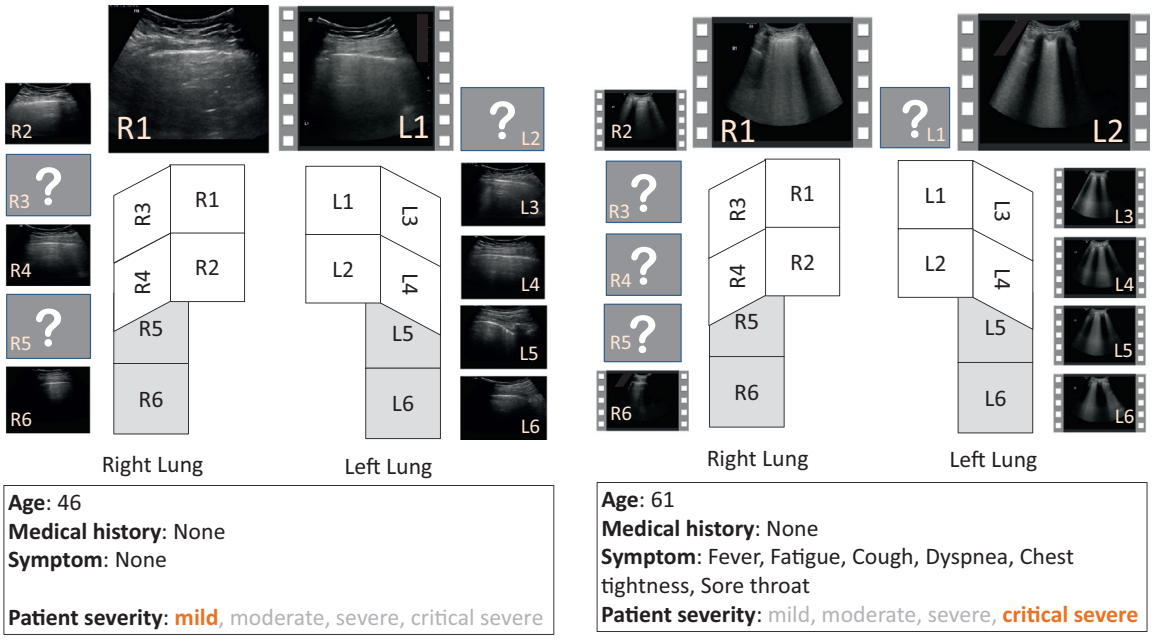


Fig. 1. Two examples for patient LUS data and clinical information. Unknown number of missing lung zones and the different format of LUS videos and images for the rest zones make the data highly heterogeneous; the clinical information provides helpful cues for analysis and also brings difficulty of combining two totally different sources of modalities; Learning the highly nonlinear mapping between the patient severity and the input LUS data is of great challenge.

13, 2020 (WHO). Given the high contagious rate, long incubation period and relatively high mortality rate, both early identification of infected patients and effective monitoring of critically severe patients are very important to control the pandemic and reduce mortality. While viral nucleic acid detection using real-time polymerase chain reaction (RT-PCR) remains the standard of reference, chest CT has been proved of high sensitivity (Ai et al., 2020; Fang et al., 2020) and become an important tool for early diagnosis of COVID-19. However, CT scan is not suitable for children and pregnant women, and is inconvenient for severely ill patients, due to its radiation and non-portability.

Lung ultrasound (LUS), due to its portability, real-time imaging capability and avoidance of radiation exposure, has become a safe bedside alternative imaging method for COVID-19 patients (Smith et al., 2020). LUS refelecting the state of lung aeration by the generated acoustic artifacts can be used as a diagnostic and respiratory monitoring tool in the ICU (Via et al., 2012). It has been proved to be highly sensitive and specific(Laursen et al., 2014) for patients with respiratory symptoms, and has been used for daily monitoring of patients (Mongodi et al., 2017; Laursen et al., 2014), and for diagnosis and follow up of community acquired pneumonia (CAP) (Reissig et al., 2012).

During the pandemic, LUS has been employed for combating the spreading of COVID-19 across the world. A review of these early experiences in LUS in the diagnosis and management of COVID-19 can be found in (Sultan and Sehgal, 2020). LUS is sensitive for detecting COVID-19 lung abnormalities and has similar accuracy compared to chest CT (de Gracia et al., 2020; Fiala, 2020). Correlated characteristic findings between CT and ultrasound were also summarized in Sultan and Sehgal (2020).

In summary, LUS can be used for 1) diagnosing symptomatic patients (Buonsenso et al., 2020b; Poggiali et al., 2020; Bar et al., 2020), 2) prognostic stratification and monitoring (Buonsenso et al., 2020b; Tung-Chen, 2020), 3) management of ICU patients with regard to treatment resources (Via et al., 2012; Volpicelli et al., 2020), and 4) monitoring the effect of therapeutic measures (Via et al., 2012; Tung-Chen, 2020; Soldati et al., 2020b). Besides, LUS is also a favorable techinique for pregnant women and children (Buonsenso et al., 2020a). Given the role of LUS for COVID-19, standardization (Soldati et al., 2020a), protocol (Manivel et al., 2020) and guidelines (Moore and Gardiner, 2020) of the use of LUS for COVID-19 patients are proposed.

When the imaging data are available, efficient analysis is a key factor for rapid decision making and treatment. To reduce the heavy workload of radiologists and improve the efficiency of the diagnosis pipeline, artificial-intelligence (AI) has been quickly introduced for analyzing COVID-19 images. Many efforts have been devoted into this area (Shi et al., 2020a; Dong et al., 2020) and achieved promising results. However, most of them focused on CT images, while only a few work contributed to intelligent analysis of LUS. Due to the dynamic scanning procedure of ultrasound, AI-based analysis for LUS can not only improve the work efficiency and alleviate the pressure on sonographers, but also reduce the observer variation and provide stable results to clinicians.

In this work, we build a systematic framework for severity assessment of COVID-19 patients from LUS data and clinical information. Specifically, when given the LUS images/videos from 12 (or less) lung zones and the clinical information of a COVID-19 patient (Fig. 1), the system is capable of 1) identifying the characteristic LUS patterns associated with COVID-19 infection; 2) evaluating the LUS zone score from these patterns, and 3) assessing the patient's severity by integration of LUS findings of all lung zones and clinical information. The first two tasks are auxiliary tasks and help extract discriminative features from the LUS data for the third task. Great challenges exist for this task, which can be summarized as:

- Data heterogeneous of LUS data for patients with various distributions and degrees of lung infection (Fig. 1). The LUS data for each lung zone can either be a dynamic video sequence, a static image or even no image. The system should be capable of dealing with the incomplete and inconsistent LUS data, and evaluating patient severity from them.
- Multi-modality of system input, which includes both the LUS data and the clinical information for each patient. Leveraging both, other than using one of them alone, can help improve the system performance.

- Highly nonlinear mapping between the high dimensional input of LUS data and the scalar output of patient severity. Effective learning of discriminative features from LUS data is the key point to achieve accurate predictions.

In this work, we overcome these challenges and make the following contributions:

- We conducted the first large scale research for AI-based LUS analysis of COVID-19 and built an effective framework, which can assess the clinical severity of patients into four types with accuracy of 75.0%, and into severe/non-severe types with accuracy of 87.5%. From the intermediate results, our method can identify the critical infected lung zones with LUS zone score (2-way accuracy of 85.28%) and indicate the characteristic pathological patterns (Hit Ratio of 80.4 for B-line).
- We designed a dual-level supervised attention-based multiple instance learning (DSA-MIL) module, which can flexibly combine the heterogeneous LUS images and videos, while leveraging supervision from both the LUS zone scores and the patient severity, therefore leading to discriminative LUS features.
- We proposed a modality alignment contrastive learning of representation (MA-CLR) module, which can force the representations of LUS data and clinical information to align with each other while keeping their respective discriminativeness, therefore combining the two modalities effectively for patient severity prediction;
- We learned the framework following a staged representation transfer (SRT) pipeline, which explores pixel-level semantic space, zone-level discriminative space, and then patient-level representations sequentially, therefore greatly alleviating and accelerating the learning procedure of the highly nonlinear mapping.

The rest of the paper is organized as follows. We first introduce existing work for AI-based analysis of COVID-19 imaging data in Section 2. Then in Section 3 we describe the materials we use for patient severity assessment, including the fundamentals of LUS for diagnosis of COVID-19, the clinical information, and the annotation procedure. In Section 4, the whole system for patient severity assessment is first briefed. Then details of the system, including the multiple instance learning, the contrastive learning and the staged training procedure, are given. Section 5 demonstrate the experiments and analyze the results. Section 6 concludes the paper and discusses the limitations.

2. Related work

Since the high sensitivity of CT image for diagnosis of COVID-19, the heavy load of the reading work, and the shortage of radiologists, urgent needs of reliable automatic systems for diagnosis of COVID-19 emerge across the whole world to combat the virus. Existing work of AI-based COVID-19 diagnosis from CT scans can be categorized into two groups: 1) identification of COVID-19 disease from other pneumonia or normal cases; 2) lesion quantification and severity assessment. We recommend the reader Shi et al. (2020a) and Dong et al. (2020) for more comprehensive review of existing AI applications for COVID-19.

2.1. Identification of COVID-19

By making use of the clinical diagnosis result, as well as existing CT images, much efforts have been devoted into identification of the novel pneumonia from normal cases and other types of pneumonia, either by deep learning models or by applying classifiers on hand-crated features. Wang et al. (2020a) conduct a proof-of-principle work for classification of CT images of COVID-19 versus viral pneumonia with Inception transfer learning on CT images. Xu et al. (2020) established a screening model to distinguish CT samples of COVID-19 from Influenza-A viral pneumonia and healthy cases, through a 3D convolution neural network (CNN) for candidate region segment ion and another location-attention model for classification. Shi et al. (2020b) proposed to screen COVID-19 from CAP based on handcrafted features from thin-section CT images. They designed an infection size aware random forest for patients with different range of infection size and obtained accurate results on a large scale dataset. Bai et al. (2020) utilized an EfficientNet B4 network (Tan and Le, 2019) for identification of COIVD-19 from other pneumonia with abnormal CT slices from a large dataset of 1186 patients, and achieved a test accuracy of 96%, sensitivity of 95%, and specificity of 96%. The system also helped improve radiologists' performance from accuracy of 85% to 90%.

There are models that take advantages of existing dataset for lung region segmentation, with which more discriminative features can be obtained, thereby leading to improved classification performance. Zheng et al. (2020) proposed COVID-19 detection method that predict the infectious probability from 3D CT volumes, with a pre-trained lung segmentation UNet and another 3D probability prediction network. Wang et al. (2020b) developed a DL system with large scale of CT scans to identify COVID-19 from other pneumonia and stratify patients into high-risk and low-risk groups, with a DenseNet121-FPN for lung segmentation and another DenseNet-like structure for diagnosis and prognosis. Li et al. (2020) employed a deep network to classify chest CT images from 3322 patients into COVID-19, CAP, and non-pneumonia (NP). The methods employed a U-net for lung ROI segmentation and a ResNet50 as the backbone for the prediction.

Without fine grained ground truth annotations, lung lesions can also be detected in the classification method as class activation map (CAM) through a gradient-based method. In Song et al. (2020) a pre-trained ResNet50 that equipped with the Feature Pyramid Network (FPN) and an attention module was used to the discriminating COVID-19 patients from others, and localize the GGO in CT images. The obtained image-level predictions were aggregated by mean pooling to achieve the final patient-level prediction. Hu et al. (2020) proposed to detect the infection region by identifying COVID-19 CT images from CAP and NP in a weak-supervised manner. A multi-view U-net trained with an open dataset was first used for inferencing the delineation of the lung area, which was then used for multi-scale classification. Categorical-specific saliency map was obtained through gradient method for lesion localization. Wu et al. (2020) proposed a joint classification and segmentation method to achieve explainable COVID-19 detection system by learning from image-level classification labels and pixel-level lesion annotations. Ouyang et al. (2020) proposed an online attention module in 3D CNN to identify COVID-19 from CAP with a dual sample strategy to tackle the imbalanced distributions of the size of infection regions. Di et al. (2020) proposed a uncertainty vertex-weighted hypergraph learning method for identifying COVID-19 from CAP with image features and radiometric features. Other classification work includes detection of COVID-19 from negative cases (Jin et al., 2020b; 2020a; Liu et al., 2020a), image-level and patient-level detection of COVID-19 (Chen et al., 2020a), radiomics models for prediction of hospital stay (long term or short term) (Qi et al., 2020).

Our work shares a similarity with the recent work Mei et al. (2020), which used artificial intelligence (AI) algorithms to integrate chest CT findings with clinical symptoms, exposure history, and laboratory testing for diagnosis of COVID-19. In their work, only directly concatenation of the CNN features for CT image and the clinical data was used. In our work, we

adaptively combine clinical information (age, symptoms, medical history) with a bag of LUS image/video for the severity assessment of COVID-19 patients. It's worth to note that the exposure history, and laboratory testing are not included in our work. This makes our task more difficult.

2.2. Quantification and severity assessment

Once the suspected cases are confirmed, it's important to assess the patient's severity, monitor the progress, and evaluate the effect of the treatment. Severity assessment therefore play significant role in these condition. Automatic AI-based severity assessment from imaging data can make the task more efficient. Shan et al. (2020) developed a deep learning (DL) based method VB-Net for automatic segmentation and quantification of infection lung regions from chest CT scans. A human-in-the-loop strategy was used to accelerate the manual delineation. Chaganti et al. (2020) proposed a deep learning method that can detect and quantify the abnormal tomographic patterns present in COVID-19, which include ground glass opacities (GGO) and consolidations. Two severity measure were also introduced for the extent of COVID-19 abnormalities and the presence of high opacities. Liu et al. (2020b) proposed to synthesize COVID-19 CT images by inpainting the main patterns of the disease on to CT images from control patients with 3D generative adversarial network. The obtained data were than used for performance improvement of lung segmentation, lesion segmentation and percentage of opacity prediction. Tang et al. (2020) proposed a random forest model to assess the severity (non-severe or severe) of 176 COVID-19 patients from of chest CT images using 63 quantitative features that were obtained by a COVID-19 chest CT analysis tool. In Shen et al. (2020), a scheme including segmentation of the lung and the pulmonary vessels, and detection of pneumonia was used to quantify parameters of lung volume, lesion volume, lesion percentage, and mean lesion density of the whole lung, right lung, left lung, and each lobe, which were then used to evaluate the severity of COVID-19. Gozes et al. (2020) proposed a pipeline for COVID-19 CT image analysis, which include a UNet architecture for lung segmentation, a ResNet-50 for classification of normal and abnormal cases, and a gradient-based method for lesion localization. Based on the 3D localization map, they proposed a quantitative *corona score* for severity estimation. He et al. (2020) proposed a combination framework of lung lobe segmentation and severity assessment, where volumetric CT image was represented by multiple 2D image patches, U-net structure was used for patch-wise segmentation, and hierarchy multiple instance learning was used for severity. Zhang et al. (2020) developed an AI system for classification of COVID-19 from CAP and normal controls, and analyzed important clinical markers that correlation with the lung lesions of COVID-19 patients. Zhu et al. (2020) predicted whether and when a COVID-19 patient would develop a severity symptoms from disease-related hand-crafted features using a sparse logic regression with class balancing strategy.

2.3. AI-based LUS analysis for COVID-19

No large scale research for AI-based LUS analysis of COVID-19 have been reported. Karakuş et al. (2020) proposed a algorithm for detection of B-line, which provides vital information for the stage and progression of COVID-19, through a non-convex regularization problem in the Radon transform domain, with a Cauchy-based penalty function. The method was validated with LUS images from 9 patients only, and cannot be used for detection of other artifacts associated with COVID-19. Roy et al. (2020) proposed a deep neural network based on Spatial Transformer network to predict the frame-level severity of LUS videos, and then aggregate them into video-level severity score. A separate UNet was trained for segmentation of COVID-19 lesion. In this work, an LUS dataset containing 277 videos from 35 patients was used, among which 80 videos from 11 patients were kept for test. It's worth to note that no patient-level severity score was obtained in this work, which plays critical role for treatment decision, progression monitoring, and management of medical resources (e.g, mechanical ventilator).

Our work differs from this work a lot and takes more ambitious steps toward AI-based LUS analysis in the following aspects: 1) Combination of multi-modality information, i.e., LUS and clinical information is used to achieve more reliable severity assessment; 2) Highly nonlinear mapping is build from the imperfect high dimensional heterogeneous LUS data to the scalar of patient severity; 3) Incorporating 313 COVID-19 patients (images or videos from 1791 lung zones), which is the largest scale of research on AI-based LUS analysis of COVID-19 patients.

3. Materials

3.1. Image acquisition

The dataset was retrospectively collected from multiple centers in Wuhan, including Cancer Center of Union Hospital, West of Union Hospital, Jianghan Cabin Hospital, Jingkai Cabin Hospital, Leishenshan Hospital, and with various ultrasound devices (Mindray M7, M8, M9 and GE Logiq E9, Logiq e Portable Ultrasound Machine). All the patients were previously diagnosed as COVID-19, with different degrees of severity. For each patient, up to twelve lung zones were examined, following the 12-zone protocol Kruisselbrink et al. (2017), with six chest areas per side based on a division of each hemithorax into anterior/lateral/posterior and upper/lower zones (Fig. 1). To improve the speed of examination for COVID-19 patients, it's assumed that none or only one static image was saved if a lung zone shows no abnormal findings. For those lung zones with abnormal findings, an ultrasound video with a duration of up to 3 respiratory cycles (about 10 s at 30 frames per second) was saved. Therefore, the collected LUS data is highly heterogeneous across different patients.

3.2. Fundamentals of LUS for COVID-19

The fundamental of LUS is based on the relative amounts of air and fluid, i.e, air/fluid ratio, which determines the characteristics of the image, and may vary according to the state of aeration of the lung. For normally aerated lung, the visceral pleura-lung air boundary reflects amost all of the incident ultrasound and generates a white hyperechoic horizontal line (pleura line). The reverberation between the pleural line and the probe yields transverse artifacts (A lines) parallel to the pleural line. Partial loss of aeration generates discrete 3D aerated structures that yield longitudinal laser-like artifacts (B lines). Complete absence of air beneath the visceral pleural makes favorable conditions for ultrasound transmission and generates a representation of lung tissue as consolidation. For COVID-19 patients, characteristics findings include thickened or unsmooth pleura line, discontinuous/fused B lines, and multiple consolidations (Sultan and Sehgal, 2020; Peng et al., 2020; Volpicelli et al., 2020).

3.3. Annotation

A total of 473 laboratory confirmed COVID-19 pneumonia patients were initially collected, among which 313 patients (male: 169; female: 144) were included in our work. The patients' age is from 17 to 97, with median of 59 years old. According to Chinese Clinical Guidance for COVID-19 Pneumonia Diagnosis and Treat-

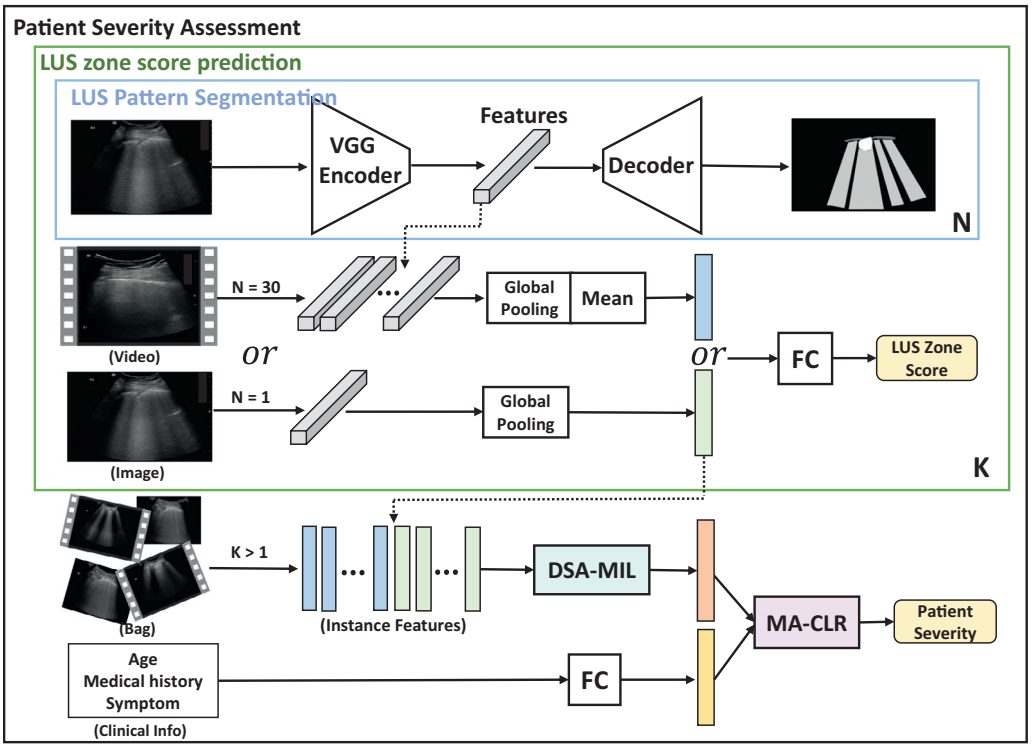


Fig. 2. Overall architecture. The representation learning follows a pipeline of LUS pattern segmentation, LUS zone score prediction, and patient-level severity assessment, during which the representation of LUS data transfers progressively into a discriminative feature. The variable at the bottom-right of the rectangle means the repetitions of the operation inside that rectangle. For video with N frames, the segmentation has to be repeated N times; for patients with data of K lung zones available, the LUS zone score prediction has to be repeated K times. The dashed line represents the two representation transfer in our training procedure.

ment[2] issued by the China National Health Commission, all patients were classified into four types (please refer to the guideline for more details):

- Mild: the clinical symptoms are mild, with no sign of pneumonia on chest imaging;
- Moderate: fever and respiratory symptoms, with positive radiological assessment of pneumonia;
- Severe: shortness of breath, less oxygen saturation, less alveolar oxygen partial pressure/fraction of inspiration or significant progression of lesion from pulmonary imaging with 24–48 h;
- Critical severe: respiratory failure requiring mechanical ventilation, shock, or failure of other ograns requiring ICU monitoring and treatment.

We denote this as patient severity (ps).

For each lung zone, an LUS zone score (zs) was given by clinicians to indicate the severity of that lung zone following the standard of: 0, normal; 1, presence of 3∼5 B-lines; 2, ≥6 B-lines or irregular pleura line; 3, fused B-lines or thickening pleura line; 4, consolidation. Similar scoring standard can be found in (Smith et al., 2020; Soldati et al., 2020a). For 173 of the patients, characteristic ultrasound patterns in each lung zone were manually delineated for up-to 10 frames of the video or for the static image. These patterns include pleura line, A-line, B-line, and lung consolidation. Detail introduction of them can be found in Kruisselbrink et al. (2017). Examples can be found in Fig. 7.

[2] http://kjfy.meetingchina.org/msite/news/show/cn/3337.html, accessed May 1, 2020.

3.4. Clinical information

Clinical information was also collected along with the LUS data for 443 patients, with records of patient's age, medical history of disease of cardiovascular, digestive, respiratory, and nervous systems, and symptom presence of fever, cough, dyspnea, pharyngalgia, diarrhea, headache, etc. This clinical information can provide complementary information to LUS for the task of patient severity assessment, and is included as part of the criterion for patient classification according to the guideline.

4. Method

To assess patient severity accurately using the challenging heterogeneous LUS data along with clinical information, we propose a staged representation transfer (SRT) framework to transfer discriminative features in a bottom-up manner. The whole pipeline consists of three sequentially training procedures: 1) frame-level LUS pattern segmentation, 2) zone-level LUS score prediction, and 3) patient severity assessment (see Fig. 2). Details of these procedures are presented below. The subscripts i, k, n indicate the nth frame of the kth lung zone of the ith patient. For the patient i, the number of lung zone is denoted as K_i, and the frame number of zone k is denoted as N_{ik}.

4.1. Pipeline of staged representation transfer for LUS

Utilizing high dimensional LUS data and clinical information to classify patients into four types of different severity is a non-trivial task. To tackle this, we propose to break down the problem into three tasks with increasing difficulty. The features and information learned from earlier tasks provide a scaffolding for later and more challenging task. Fig. 2 shows the overall architecture of our framework.

LUS pattern segmentation The characteristic findings, such as pleural line, B-line and lung consolidation, are indispensable to accurate diagnosis of COVID-19. To embed this knowledge in the final decision-making, we design a model to segment these LUS patterns. The model deploys a VGG encoder as the backbone for representation learning, on top of which three different heads are attached. The first one is a decoder head for semantic segmentation of these patterns. The encoder and decoder form a UNet-like architecture, and is trained with combination of Dice loss and cross-entropy loss. By this way, the obtained latent frame-level representation $f_{i,n}$ for the nth frame of patent i is endowed with rich semantic information and could be beneficial to the following LUS zone score prediction and patient severity assessment. To make the network capable of processing the huge amount of LUS data for one patient, we make the network a light-weight model.

LUS zone score prediction A lung can be partitioned into 12 zones (Kruisselbrink et al. (2017)), and it is important to recognize the impairment of each of them. To achieve this, we devise a classification model, within which a optional aggregating module is used to transform the obtained frame-level features $[f_{i,k,1}, f_{i,k,2}, \dots f_{i,k,N_{ik}}]$ into the same space as the image-level feature, therefore make our method capable of dealing with heterogeneous data in a mixture of images and videos. Temporal average pooling is used to aggregating frame-level features into zone-level features, which is denoted as $x_{i,k}$. Finally, $x_{i,k}$ is passed to a fully-connected layer and categorizes the input into five class. Through this procedure, the infection degree of each zone is further incorporated into $x_{i,k}$ and could be transfer to the next level. A bag of zone-level representations can be obtained for each patient to predict the patient severity. Details of this model can be found in the middle part of Fig. 2.

Patient severity assessment The last model predict the severity of a patient using both LUS data and clinical information. The zone-level features of all available zones, i.e, $x_{i,k}$ for all zones, are first fused together with a multiple instance learning module (DSA-MIL). The resulted patient-level LUS feature is then aligned and combined with the clinical information by a contrastive learning (CMA-CLR) module to predict the patient severity. These two novel modules will be explained in details below.

4.2. Dual-level supervised attention-based multiple instance learning

Predicting the patient severity ps_i of patient i from a bag of lung zone representation $X_i = \{x_{i,k}, k = 1 \dots K_i\}$ can be viewed as a classical multiple instance learning (MIL) problem (Chen et al., 2006):

$$ps_i = g(\sigma(\{f(x_{i,1}), f(x_{i,2}), \dots, f(x_{i,K_i})\})). \tag{1}$$

where σ is a permutation-invariant MIL operator that give the bag prediction or representation from the instance score or representations, and can be implemented as max-pooling, average-pooling or recurrent neural network. Two different cases exist for f. When f is a instance-level classifier, g is an identity function. When f is a nonlinear function that transforms input feature into a low dimensional space, g acts as a bag-level classifier. MIL has been employed in CT-based COVID-19 patient screening (Han et al., 2020) and severity assessment (He et al., 2020), with 2D CT patches and 3D CT patches respectively. In He et al. (2020), two-level MIL was used to obtain the volume feature for severity assessment.

In the work of Ilse et al. (2018), the attention mechanism was introduced into MIL model and used for average-pooling of the instance representation with adaptive weights. The weight of each instance was obtained through an attention module and provided interpretation of the instance-bag relationship. The attention-based

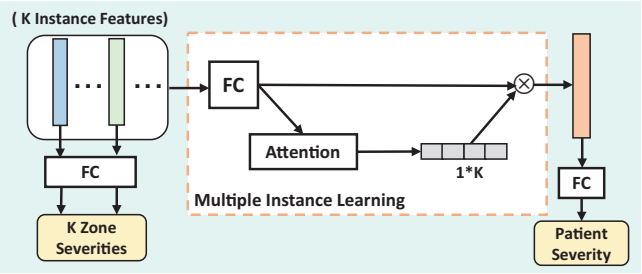


Fig. 3. Dual-level supervised multiple instance learning. Both the instance-level and the bag-level supervision are used to help patient representation learning.

MIL can be formulated as:

$$b_i = \sum_{k=1}^{K_i} a_k f(x_{i,k}) \tag{2}$$

where

$$a_k = \frac{\exp(W_a^T max(V_a f(x_{i,k}), 0))}{\sum_{l=1}^{K_i} \exp(W_a^T max(V_a f(x_{i,l}), 0))} \tag{3}$$

is the attention vector and can be used for interpreting the importance of each instance. W_a and V_a are the parameters of two liner transformations, between which the ReLU nonlinear function is used.

In our work, we aim to learn a bag-level representation with maximal discriminativeness from patient's LUS data of all lung zones to achieve a reliable severity assessment. To this end, we propose a dual-level supervised attention-based MIL module, or DSA-MIL, as shown in Fig. 3. Different from existing MIL method that only predict the bag label from b_i, we propose to supervise the learning of MIL with both the bag label and instance labels, i.e, ps and zs in our case. Therefore, a prediction of LUS zone score probability $\hat{zs}_{i,k}$ from each $f(x_{i,k})$ can be obtained by a linear mapping W_z and a softmax function. The probability of patient severity $\hat{ps}_i$ can be obtained from b_i by a linear mapping W_{p1} and a softmax function.The loss function of the dual-level supervision is:

$$\mathcal{L}^{DS} = \frac{1}{S}\sum_{i=1}^{S}\left[CE(ps_i, \hat{ps}_i) + \frac{1}{K_i}\sum_{k=1}^{K_i} CE(zs_{i,k}, \hat{zs}_{i,k})\right] \tag{4}$$

where CE is the cross-entropy loss for classification, S is the total number of patients. Note that when the clinical information is used, $\hat{ps}_i$ is obtained from combination of the two modalities, which is described below. The key difference between the MIL module of our method and He et al. (2020) is that both the instance-level and bag-level supervision are used in our MIL module.

4.3. Modality alignment contrastive learning of representation

While LUS itself can provide a good severity assessment, we try our best to combine both the patient's LUS data and the clinical information, with the expectation of further improving the performance. However, the clinical information is very different from the LUS data and may be recorded at a different time from the acquisition of LUS. Misalignment may exist between the two modalities and undermine the combination of them, as can be seen in row CONCAT of Table 4 that direct concatenation leads to obvious performance decrease. Thus it's important to design a module that can effectively combine them.

We propose a novel modality alignment contrastive learning of representation (MA-CLR), which can make the two modalities align to each other, and in the meantime maintain the dis-

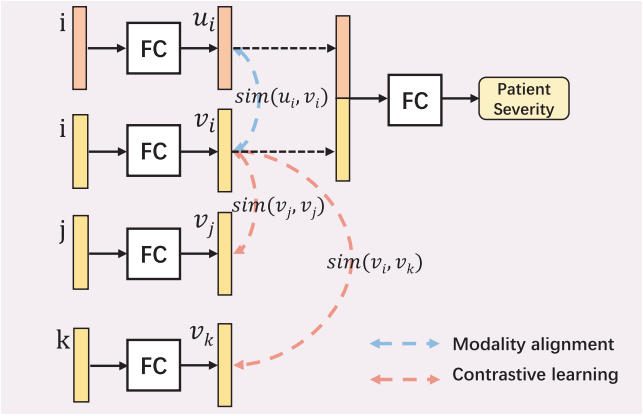


Fig. 4. Modality alignment contrastive learning of representation. Patients i, j, k are of different severities. The blue arrow maximizes the agreement between LUS and clinical information of the same patient, therefore make the two modalities aligned; the red arrow minimize the inter-class agreement of two patients with different severities, therefore keep the discriminativeness; an additional constraint $sim(\mathbf{v}_i, \mathbf{v}_k) \leq sim(\mathbf{v}_i, \mathbf{v}_j)$ makes this contrastive pattern consistent. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

criminativeness of themselves. The module is inspired by the recent research on self-supervised contrastive learning of representation (Chen et al., 2020b), where discriminative representations are learned by maximizing the agreement between positive pair samples. Differing from this work that learn effective representation in a self-supervised way from one modality, we employ the idea of contrastive learning for combination of two modalities, with the contrastive learning guided by the label information. Details of MA-CLR can be found in Fig. 4. The LUS feature and the clinical feature are first transformed with a linear mapping into the same space, which are denoted as $\mathbf{u}_i, \mathbf{v}_i$ for patient i, respectively. Then the probability of patient severity $\hat{ps}_i$ is predicted from concatenation of them $[\mathbf{u}_i, \mathbf{v}_i]$ via a linear mapping W_{p2} and a softmax function. The contrastive loss in our proposed MA-CLR contains to three parts, which respectively cope with the modality alignment $\mathcal{L}^{MA}$, inter-class contrastive learning $\mathcal{L}^{CL}$, and contrastive continuity $\mathcal{L}^{CC}$. The proposed MA-CLR can be implemented by optimizing the following objective:

$$\mathcal{L}^{MACLR} = \frac{1}{S}\sum_{i=1}^{S}\left[\mathcal{L}_i^{MA} + \mathcal{L}_i^{CL} + \mathcal{L}_i^{CC}\right] \tag{5}$$

where

$$\mathcal{L}_i^{MA} = 1 - sim(\mathbf{u}_i, \mathbf{v}_i) \tag{6}$$

$$\begin{aligned}\mathcal{L}_i^{CL} = &[sim(\mathbf{u}_i, \mathbf{v}_j) + sim(\mathbf{u}_i, \mathbf{v}_k)] \\ &+ [sim(\mathbf{v}_i, \mathbf{v}_j) + sim(\mathbf{v}_i, \mathbf{v}_k)]\end{aligned} \tag{7}$$

$$\begin{aligned}&\forall i, \exists j, |ps_i - ps_j| = 1 \\ \mathcal{L}_i^{CC} = &\max(sim(\mathbf{v}_i, \mathbf{v}_k) - sim(\mathbf{v}_i, \mathbf{v}_j), 0), \\ &\forall i, \exists j, k, |ps_i - ps_j| < |ps_i - ps_k|\end{aligned} \tag{8}$$

where $sim(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u}^T\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}$ denotes the cosine similarity between two vectors $\mathbf{u}$ and $\mathbf{v}$. The first term aims to maximize the agreement between LUS and clinical representations. The second term minimizes the inter-class agreement for, either LUS or clinical information, therefore maintaining the discriminativeness. The third term keeps the continuity of this discriminativeness across different classes. For the second term and the third term, clinical information of two additional patients with different patient severities will be randomly sampled to help the contrastive learning during each training step.

The total objective function of patient severity assessment is therefore summation of $\mathcal{L}^{DS}$ and $\mathcal{L}^{MACLR}$.

Table 1
Dataset splitting for patient severity assessment.

Patient severity	1	2	3	4
Training	12	145	63	13
Test	20	20	20	20

Table 2
Distribution of LUS zone score for the training and test dataset. According to the split in Table 1, the LUS zone score also follow an even distribution.

Zone severity	0	1	2	3	4
Training	697	476	206	113	34
Test	55	53	53	52	52

Table 3
Numbers of images/frames with presence of the corresponding ultrasound patterns. A: A-line; P: pleural line; B: B-line; C: consolidation.

LUS patterns	A	P	B	C	Total
Training	666	4338	3611	340	4398
Test	205	2438	2295	481	2528

5. Experiments and analysis

5.1. Configurations

Following the stages representation transfer procedure, we split the dataset into training set and test set and then apply them for all of the three tasks. In our experiments, LUS data from 313 patients are used and augmented to keep the data distribution balance. The details of the splitting and the data distributions of patient severity, LUS zone score, and LUS patterns are demonstrated in Tables 1, 2, 3. The augmentation we used include affine transformations (translation, rotation, scaling, shearing), reflection, contrast change, Gaussian noise, and Gaussian filtering.

During the training procedure of LUS zone score and patient severity, we randomly sampled 30 frames from each video. For the segmentation task, we sampled the 10 key frames where these patterns were manually delineated by clinicians, and their two neighboring frames. For inference, we randomly sampled 30 frames from each video. Each patient is tested five times and the max-vote is used for final assessment.

In our experiments, we focus on the main task of patient severity assessment. We first conduct extensive ablation studies for our method with different configurations to validate the role of the staged representation transfer learning procedure, the benefits of the dual-level supervision of DSA-MIL and the contrastive learning of MA-CLR. Besides the four-way severity assessment task, we also demonstrate the results of the 2-way task, i.e., whether the patient is non-severe or severe, as commonly did in existing work (Tang et al., 2020; Shen et al., 2020; He et al., 2020). Evaluation criteria include accuracy, precision, recall and macro F1-score. We also demonstrate the results of LUS zone score and pattern segmentation, which can provide more detailed indications of which zone and where in that zone COVID-19 infections can be found.

5.2. Evaluation of patient severity assessment

We design the following ablation studies to validate the effectiveness of staged learning procedure, the dual-level supervised MIL module and the contrastive learning module. Specifically, the

Table 4
Patient severity results (%). For each configuration, average performance of 10 learned model is reported. The standard deviations of the accuracies are also displayed. These results reveal the effectivenesses of the staged learning procedure, the DSA-MIL module, and the MA-CLR module. CI: clinical information.

config.	modality	Patient severity		Severe or non-severe			
		accuracy	F1-score	accuracy	recall	precision	F1-score
MZS	LUS	49.38±1.35	45.89	79.50±0.87	78.50	80.11	79.29
A-MIL1	LUS	64.25±1.34	66.05	84.0±2.99	76.75	**90.13**	82.57
A-MIL2	LUS	64.25±1.47	64.57	84.75±2.02	**86.25**	83.77	84.91
DSA-MIL	LUS	67.63±1.38	67.91	84.88±1.50	78.75	89.83	83.84
MLP	CI	56.75±0.65	56.15	79.88±1.24	80.75	79.60	80.03
CONCAT	LUS and CI	55.25±0.99	57.37	83.13±1.47	80.25	85.36	82.52
MA-CLR	LUS and CI	72.75±0.53	72.31	86.5±0.53	82.25	89.91	85.90
MA-CLR (max-vote)	LUS and CI	**75.00**	**74.44**	**87.5**	85.00	89.47	**87.18**

following different configurations are employed during our experiments:

- MZS: inferring the patient severity by averaging the predicted zone severity $\hat{ps}$.
- A-MIL1: prediction of patient severity from a bag of zone-level representations with the attention-based multiple instance learning module. No pretrained model is used here.
- A-MIL2: pre-trained network weights from the task of LUS zone score (which also employs pre-trained weights from the segmentation task) is used.
- DSA-MIL: dual-level supervised information is used in the A-MIL2 module.
- MLP: prediction of patient severity from the clinical information only using a multilayer perceptron (MLP).
- CONCAT: directly concatenate the representations of LUS and clinical information for prediction of patient severity.
- MA-CLR: combine the two modality with the proposed contrastive learning module MA-CLR for prediction of patient severity.
- MA-CLR (max-vote): MA-CLR is repeated five times for test, and the max-vote result is adopted as the final prediction.

In Table 4 we demonstrate the performance for patient severity assessment with these configurations. The results were obtained by averaging of 10 repetitions. Overall, our method can predict the patient severity from combination of LUS data and clinical information very well, with accuracy of 75.0% for severity assessment, and 87.5% for binary severe/non-severe classification. The small standard deviations indicate that the performance of our method is of low variation.

From the table, we can have a clear observation of how the SRT procedure and the DSA-MIL module alternatively improve the performance of the 4-way and the 2-way classification tasks. Firstly, only 49.38% accuracy can be obtained by MZS, which implies the necessity of learning the instance combination adaptively. A-MIL learns the weight parameter for each instance and therefore leads to improvement with large margin. Secondly, the LUS zone score pre-trained weights can bring clearly improvement for the 2-way task with the transferred discriminative information, while keeping the performance of the more difficult 4-way task. It also makes the learning curve converge more quickly (reduce 50% of training epochs, from 60 epochs to 30 epochs) and stable. Then, the proposed DSA-MIL, which learns the instance combination procedure with two-level supervision, can enhance the discriminativeness of the instance presentation first, and then combine them with the attention vector. Therefore, the performance of the 4-way task is improved by a large margin, while the performance of the 2-way task is maintained.

The effectiveness of the proposed MA-CLR module can be revealed by the last four rows of the table. Due to the inconsistency of acquisition time, the representation of LUS and clinical information may be not well aligned and make it extremely hard to effectively combine them to improve the performance of patient severity prediction. As can be seen in the table, CONCAT even leads to a significant performance decrease for the 4-way classification performance, and a mild decrease for the binary task. This reveals that the two modalities may align well at that border of the severe/non-severe group, but compromise each other for the fine-grained discrimination of mild/moderate groups, and severe/critical severe groups. This can be clearly revealed by the off-diagonal elements of the confusion matrices. With MA-CLR, the representations of the two modalities can be well-aligned, and the inter-class contrastive pattern and intra-class consistency can be well enhanced, therefore performances of both the 4-way and the 2-way tasks are significantly improved. It's worth noting that the LUS images in our dataset were acquired with multiple devices, which usually raises the issue of style difference for learning-based methods. MA-CLR can be used not only as a general multi-modality combination method, but also as a domain adaption way to alleviate the style differences for medical images.

In Table 5, the p-values of paired t-test were demonstrated for the accuracy in Table 4. p-value less then 0.05 means the performance difference of the two methods is significant. We can clearly observe that the differences between our method MA-CLR and the rest methods are significant, which further validates the effectiveness of the modules in our method.

From the confusion matrix in Fig. 5, more interesting observations can be drawn: 1) when only clinical information is used, MLP can identify well the severe patients from the non-severe ones, but cannot probe the minor difference between mild and moderate groups, and between severe and critical severe groups. 2) direct concatenation of the representations of LUS data and clinical information cannot help differentiate these minor differences. Therefore, the proposed MA-CLR greatly reduces these misclassifications with the contrastive learning mechanism. It'a also interesting to note that when combined with LUS and age, the symptom information delivers better accuracy than medical history (71.25% vs. 70.38%).

The performance of existing work on binary COVID-19 severity prediction from CT images is listed in Table 6. Our method achieved better accuracy than Yang et al. (2020), and comparable performance with Tang et al. (2020), revealing the great potential for AI-based LUS analysis.

The inference time of our severity assessment model was tested on a platform with Intel(R) Xeon(R) Silver 4114 CPU, RTX2080Ti GPU, and Ubuntu 180.04. Given the fact that the LUS data is a mixture of images and videos of unknown number, two cases are tested. For patients with 12 static images, the average inference time is 6.74 ms. For patients with 12 LUS videos (each with 30 frames), the average inference time is 76.98 ms.

Table 5
p-value of significant test of the severity assessment performance for different configuration. *p*-value less then 0.05 means the performance difference of the two methods is significant.

p-value	MZS	A-MIL1	A-MIL2	DSA-MIL	MLP	CONCAT
accuracy of patient severity						
A-MIL1	3.12E-11					
A-MIL2	1.64E-10	1				
DSA-MIL	4.54E-10	1.97E-04	7.72E-05			
MLP	2.61E-07	2.96E-07	2.96E-07	3.36E-10		
CONCAT	8.42E-07	1.35E-07	1.35E-07	1.60E-09	2.56E-03	
MA-CLR	2.15E-12	5.37E-08	1.47E-08	1.84E-06	2.80E-13	1.24E-12
accuracy of severe/non-severe						
A-MIL1	1.73E-03					
A-MIL2	5.74E-05	5.09E-01				
DSA-MIL	7.81E-06	4.42E-01	9.06E-01			
MLP	3.94E-01	6.29E-03	1.18E-04	3.18E-06		
CONCAT	3.13E-05	3.90E-01	7.68E-02	4.45E-02	7.46E-04	
MA-CLR	1.96E-08	1.68E-02	2.05E-02	1.33E-02	3.01E-07	2.09E-05

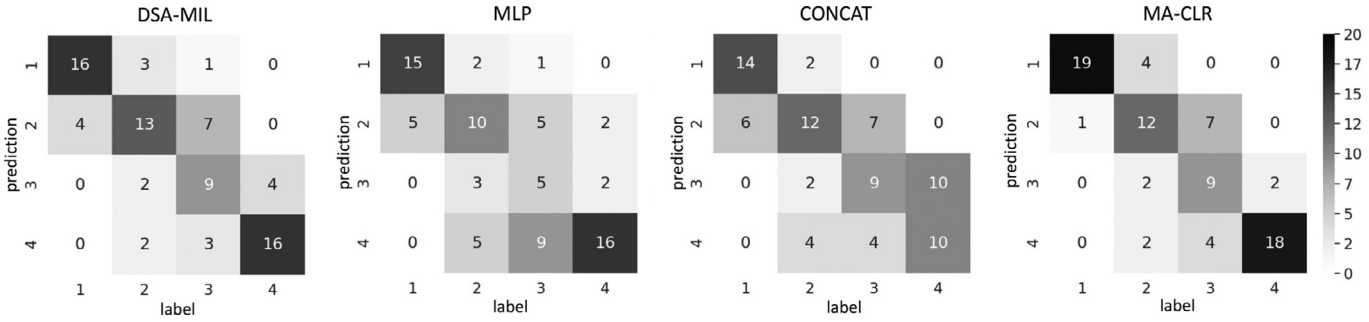


Fig. 5. Confusion matrix for patient severity assessment under different configurations. 1, mild; 2, moderate; 3, severe; 4, critical severe.

Table 6
Performance comparison with existing work for binary severe/non-severe prediction (%).

References	Modality	Patient	accuracy	recall	precision
Yang et al. (2020)	CT	102	-	83.3	75.0
Tang et al. (2020)	CT	176	87.5	93.3	-
He et al. (2020)	CT	242	98.5	95.2	97.5
ours	LUS+CI	313	87.5	85.0	89.47

Table 7
Evaluation of LUS segmentation task. A: A-line; P: pleural line; B: B-line; C: consolidation.

	Hit Ratio$_{50}$	Dice	mIoU	Precision	Recall
A	0.284	0.258	0.302	0.444	0.280
P	0.704	0.620	0.649	0.730	0.645
B	0.804	0.715	0.695	0.743	0.737
C	0.160	0.170	0.201	0.379	0.173

5.3. Evaluation of LUS zone score prediction

The LUS zone score indicates which regions the lung are infected, and to which extent the infection is. This provides more interpretation of the patients overall severity. We evaluated the performance of our method for LUS zone score prediction with and without the pre-trained segmentation model. Similar to the task of patient severity, we merge the zone scores $\{0, 1\}$ as 0 representing non-severity and $\{2, 3, 4\}$ as 1 representing severity. Then the performance for this binary classification is also reported here.

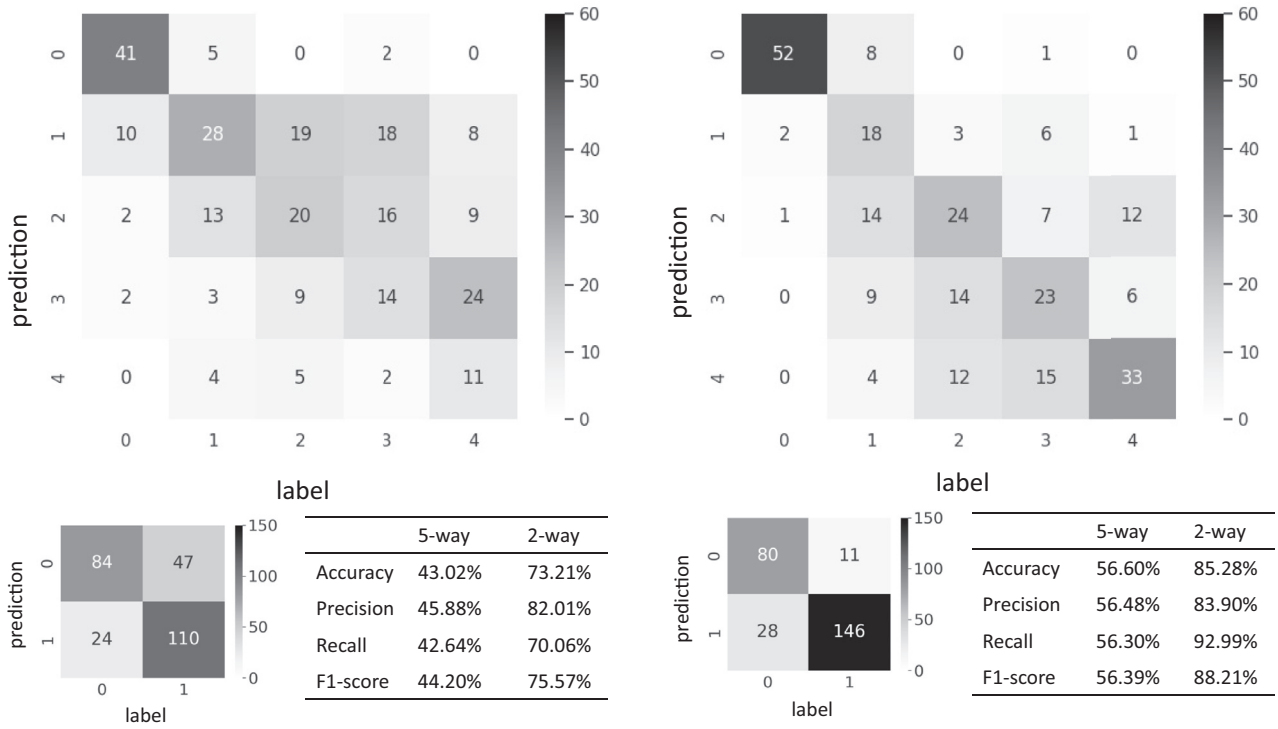
From Fig. 6, it can be seen that when the pre-trained segmentation model is used, the accuracy of both 5-way and 2-way classification tasks can be obviously improved: from 43.02% to 56.60% for the 5-way task, and from 73.21% to 85.28% for the binary severity/non-severity task. The F1-score is also improved from 44.20% to 56.39% for the 5-way task. The reduce of the off-diagonal elements prove that representation transfer can greatly reduce the mis-classifications between groups of 1, 2,3 and groups of 3, 4. This again validates the effectiveness of our SRT training procedure. For comparison, the only one existing work on AI-based LUS severity assessment (Roy et al., 2020) achieved F1-score of 57.9% for 4-way video based classification.

For identifying the severely infected zones (binary classification), our method can obtain a high accuracy of 85.28%, recall of 92.99%, precision of 83.90%, and F1-score of 88.21%.

5.4. Evaluation of LUS pattern segmentation

Here we show the result for segmentation of ultrasound patterns, which can provide a clearly identification of the pathological regions that are associated with the COVID-19. Examples of five LUS images with LUS zone score from 0 to 4 are illustrated in Fig. 7. The segmentation head of our method can provide accurate segmentations of B-lines, and pleural lines. For A-lines and consolidation, they only appear in an extremely small portion of our training set, therefore the network failed to output good contours.

In Table 7 the quantitative evaluation for different type of patterns is demonstrated. The metrics include the commonly used ones for segmentation, such as Dice, mIoU, precision, recall, and the Hit Ratio, which indicate how frequently the method can detect the corresponding pattern, with IoU more than 50%. For the



	5-way	2-way
Accuracy	43.02%	73.21%
Precision	45.88%	82.01%
Recall	42.64%	70.06%
F1-score	44.20%	75.57%

	5-way	2-way
Accuracy	56.60%	85.28%
Precision	56.48%	83.90%
Recall	56.30%	92.99%
F1-score	56.39%	88.21%

Fig. 6. Performance and confusion matrix for LUS zone score prediction with (right) and without (left) the pre-trained segmentation model. Both 5-way and 2-way classifications are evaluated. It's obvious that including the rich semantic information in the pre-trained model brings great performance improvement for LUS zone score prediction.

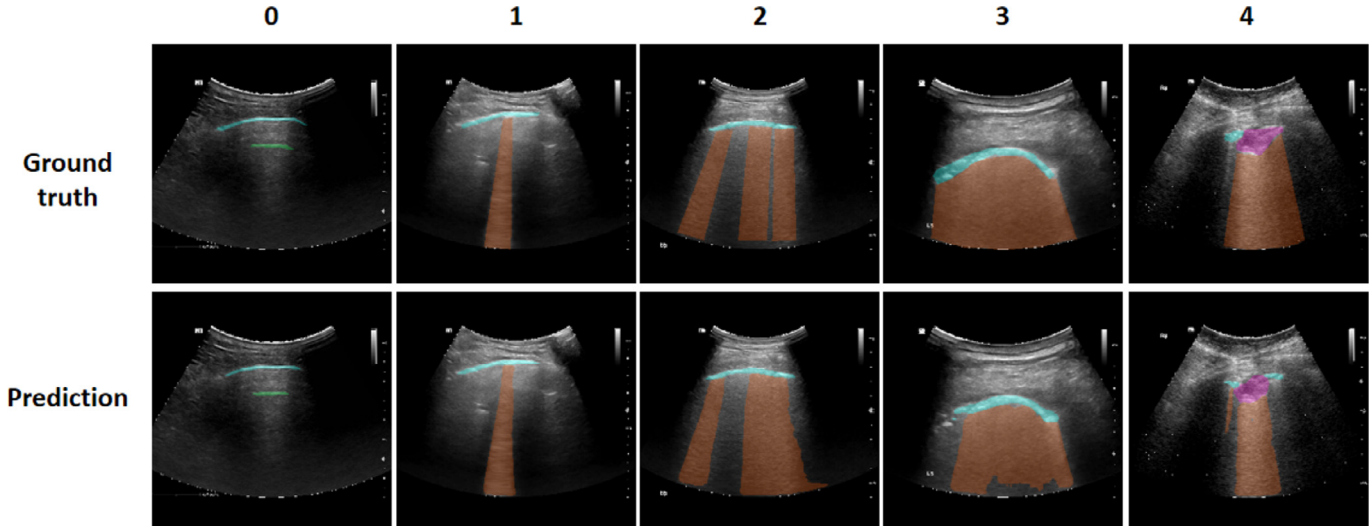


Fig. 7. Segmentation examples for a representative frame from five lung zones with different severity. Cyan: pleural line, Green: A-line, Brown: B-line, Purple: consolidation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

two prevalence patterns in our dataset, the method can achieve good segmentation, with Dice of 0.620, and Hit Ratio of 0.704 for pleural line, and Dice of 0.715, Hit Ratio of 0.804 for B-line. For the two rare ones, the method detects them with very low Hit Ratio. We leave this to our future work.

6. Conclusion and discussion

To improve the efficiency and reliability of LUS in assessing and monitoring patients with COVID-19 while the novel coronavirus is still spreading worldwide, we proposed a novel and effective method that can accurately predict the patient severity, according to the non-standard heterogeneous LUS images and videos, and as while with combination of the clinical information, such as age, medical history, symptoms. A DSA-MIL module was proposed to predict the patent severity from heterogeneous LUS data of multiple lung zones, and a contrastive learning module MA-CLR was proposed for combination of the LUS data and the clinical information. When validated with 80 patients, the method can achieve high accuracy for both 4-way severity assessment (75.0%) and two-way classification (87.5%). As the by-product, our method can provide interpretable clues for the patient severity, by predicting the LUS zone scores and identifying the infected regions (B-lines) in each lung zone. This endows the method a great potential

in progress monitoring, treatment planning, therapy evaluation and patient management in clinical practice.

There are still limitations that the proposed method failed to overcome. Firstly, the distribution of lung zone information is ignored in our method, which reflects the evolvement of infection across different lung zones and is beneficial for patient assessment. In our dataset, unknown number of images and videos constitute the LUS data of a patient, therefore the distribution of the lung zone information can not be well represented and modeled. When LUS data from all lung zones are available, such information can be incorporated by prior-based constraints, correlation modeling or graph networks. Secondly, quantitative analysis of the characteristic patterns in LUS data may correlate well with the zone score. For example, a quantitative measure of the B lines with proper normalization may be an effective indicator of the infection extent. We leave this to future work.

Declaration of Competing Interest

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

CRediT authorship contribution statement

Wufeng Xue: Conceptualization, Methodology, Writing - original draft. **Chunyan Cao:** Data curation, Formal analysis. **Jie Liu:** Data curation, Formal analysis. **Yilian Duan:** Data curation. **Haiyan Cao:** Data curation. **Jian Wang:** Investigation, Validation. **Xumin Tao:** Investigation, Validation. **Zejian Chen:** Investigation, Validation. **Meng Wu:** Data curation. **Jinxiang Zhang:** Data curation. **Hui Sun:** Data curation. **Yang Jin:** Data curation. **Xin Yang:** Methodology, Project administration. **Ruobing Huang:** Writing - review & editing. **Feixiang Xiang:** Data curation. **Yue Song:** Data curation. **Manjie You:** Data curation. **Wen Zhang:** Data curation. **Lili Jiang:** Data curation. **Ziming Zhang:** Data curation. **Shuangshuang Kong:** Data curation. **Ying Tian:** Data curation. **Li Zhang:** Conceptualization, Resources, Writing - review & editing, Project administration. **Dong Ni:** Supervision, Conceptualization, Writing - review & editing. **Mingxing Xie:** Supervision, Conceptualization, Resources, Writing - review & editing.

Acknowledgment

The work is partially supported by the National Key R&D Program of China (No. 2019YFC0118300), Natural Science Foundation of China (No. 81727805, 81530056, 61801296), Natural Science Foundation of Shenzhen (JCYJ20190808115419619), Shenzhen Peacock Plan (No. KQTD2016053112051497, KQJSCX20180328095606003), Medical Scientific Research Foundation of Guangdong Province (No. B2018031).

References

Ai, T., Yang, Z., yan Hou, H., Zhan, C., Chen, C., Lv, W., Tao, Q., Sun, Z., Xia, L., 2020. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology.

Bai, H.X., Wang, R., Xiong, Z., Hsieh, B., Chang, K., Halsey, K., Tran, T.M.L., Choi, J.W., Wang, D.-C., Shi, L.-B., et al., 2020. AI augmentation of radiologist performance in distinguishing COVID-19 from pneumonia of other etiology on chest CT. Radiology 201491.

Bar, S., Lecourtois, A., Diouf, M., Goldberg, E., Bourbon, C., Arnaud, E., Domisse, L., Dupont, H., Gosset, P., 2020. The association of lung ultrasound images with COVID-19 infection in an emergency room cohort. Anaesthesia.

Buonsenso, D., Pata, D., Chiaretti, A., 2020a. COVID-19 outbreak: less stethoscope, more ultrasound. Lancet Respir. Med. 8 (5), e27.

Buonsenso, D., Raffaelli, F., Tamburrini, E., Biasucci, D.G., Salvi, S., Smargiassi, A., Inchingolo, R., Scambia, G., Lanzone, A., Testa, A.C., Moro, F., 2020b. Clinical role of lung ultrasound for the diagnosis and monitoring of COVID-19 pneumonia in pregnant women. Ultrasound Obstet. Gynecol..

Chaganti, S., Balachandran, A., Chabin, G., Cohen, S., Flohr, T., Georgescu, B., Grenier, P., Grbic, S., Liu, S., Mellot, F., et al., Quantification of tomographic patterns associated with COVID-19 from chest CT. arXiv:2004.01279

Chen, J., Wu, L., Zhang, J., Zhang, L., Gong, D., Zhao, Y., Hu, S., Wang, Y., Hu, X., Zheng, B., et al., 2020a. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study. medRxiv.

Chen, T., Kornblith, S., Norouzi, M., Hinton, G., A simple framework for contrastive learning of visual representations. arXiv:2002.05709

Chen, Y., Bi, J., Wang, J.Z., 2006. Miles: Multiple-instance learning via embedded instance selection. IEEE Trans. Pattern Anal. Mach.Intell. 28 (12), 1931–1947.

Di, D., Shi, F., Yan, F., Xia, L., Mo, Z., Ding, Z., Shan, F., Li, S., Wei, Y., Shao, Y., et al., Hypergraph learning for identification of COVID-19 with CT imaging. arXiv:2005.04043

Dong, D., Tang, Z., Wang, S., Hui, H., Gong, L., Lu, Y., Xue, Z., Liao, H., Chen, F., Yang, F., et al., 2020. The role of imaging in the detection and management of COVID-19: a review. IEEE Rev. Biomed. Eng..

Fang, Y., Zhang, H., Xie, J., Lin, M., Ying, L., Pang, P., Ji, W., 2020. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. Radiology 200432.

Fiala, M., 2020. Ultrasound in COVID-19: a timeline of ultrasound findings in relation to CT. Clin. Radiol..

Gozes, O., Frid-Adar, M., Sagie, N., Zhang, H., Ji, W., Greenspan, H., Coronavirus detection and analysis on chest CT with deep learning. arXiv:2004.02640

de Gracia, M.M., DiezTascon, A., Agudo-Fernandez, S., Alonso-Gonzalez, R., Fuertes, P.R., Parra-Gordo, L., Ossaba-Velez, S., Fuentes, R.L., et al., 2020. Correlation between chest computed tomography and lung ultrasonography in patients with coronavirus disease 2019 (COVID-19). medRxiv.

Han, Z., Wei, B., Hong, Y., Li, T., Cong, J., Zhu, X., Wei, H., Zhang, W., 2020. Accurate screening of COVID-19 using attention based deep 3D multiple instance learning. IEEE Trans. Med. Imaging.

He, K., Zhao, W., Xie, X., Ji, W., Liu, M., Tang, Z., Shi, F., Gao, Y., Liu, J., Zhang, J., et al., Synergistic learning of lung lobe segmentation and hierarchical multi-instance classification for automated severity assessment of COVID-19 in CT images. arXiv:2005.03832

Hu, S., Gao, Y., Niu, Z., Jiang, Y., Li, L., Xiao, X., Wang, M., Fang, E.F., Menpes–Smith, W., Xia, J., et al.. Weakly supervised deep learning for COVID-19 infection detection and classification from CT images.

Ilse, M., Tomczak, J.M., Welling, M., 2018. Attention-based deep multiple instance learning. In: 35th International Conference on Machine Learning, ICML 2018. International Machine Learning Society (IMLS), pp. 3376–3391.

Jin, C., Chen, W., Cao, Y., Xu, Z., Zhang, X., Deng, L., Zheng, C., Zhou, J., Shi, H., Feng, J., 2020a. Development and evaluation of an ai system for COVID-19 diagnosis. medRxiv.

Jin, S., Wang, B., Xu, H., Luo, C., Wei, L., Zhao, W., Hou, X., Ma, W., Xu, Z., Zheng, Z., et al., 2020b. AI-assisted CT imaging analysis for COVID-19 screening: building and deploying a medical ai system in four weeks. medRxiv.

Karakuş, O., Anantrasirichai, N., Aguersif, A., Silva, S., Basarab, A., Achim, A., Line artefact quantification in lung ultrasound images of COVID-19 patients via non-convex regularisation. arXiv:2005.03080

Kruisselbrink, R., Chan, V., Cibinel, G.A., Abrahamson, S., Goffi, A., 2017. I-AIM (indication, acquisition, interpretation, medical decision-making) framework for point of care lung ultrasound. Anesthesiology 127 (3), 568–582.

Laursen, C.B., Sloth, E., Lassen, A., dePont Christensen, R., Lambrechtsen, J., Madsen, P.H., Henriksen, D.P., Davidsen, J.R., Rasmussen, F.N., 2014. Point-of-care ultrasonography in patients admitted with respiratory symptoms: a single-blind, randomised controlled trial.. Lancet Respir. Med. 2 (8), 638–646.

Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., Bai, J., Lu, Y., Fang, Z., Song, Q., et al., 2020. Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. Radiology 200905.

Liu, B., Liu, P., Dai, L., Yang, Y., Xie, P., Tan, Y., Du, J., Shan, W., Zhao, C., Zhong, Q., Lin, X., Guan, X., Xing, N., Sun, Y., Wang, W., Zhang, Z., Fu, X., Fan, Y., Li, M., Zhang, N., Li, L., Liu, Y., Xu, L., Du, J., Zhao, Z., Hu, X., Fan, W., Wang, R., Wu, C., Nie, Y., Cheng, L., Ma, L., Li, Z., Jia, Q., Liu, M., Guo, H., Huang, G., Shen, H., An, W., Li, H., Zhou, J., He, K., 2020a. Assisting scalable diagnosis automatically via CT images in the combat against COVID-19. medRxiv doi:10.1101/2020.05.11.20093732.

Liu, S., Georgescu, B., Xu, Z., Yoo, Y., Chabin, G., Chaganti, S., Grbic, S., Piat, S., Teixeira, B., Balachandran, A., et al., 3D Tomographic pattern synthesis for enhancing the quantification of COVID-19. arXiv:2005.01903

Manivel, V., Lesnewski, A., Shamim, S., Carbonatto, G., Govindan, T., 2020. CLUE: COVID-19 lung ultrasound in emergency department. Emerg. Med. Aust.

Mei, X., Lee, H.-C., Diao, K., Huang, M., Lin, B., Liu, C., Xie, Z., Ma, Y., Robson, P.M., Chung, M., et al., 2020. Artificial intelligence-enabled rapid diagnosis of COVID-19 patients. medRxiv.

Mongodi, S., Pozzi, M., Orlando, A., Bouhemad, B., Stella, A., Tavazzi, G., Via, G., Iotti, G.A., Mojoli, F., 2017. Lung ultrasound for daily monitoring of ARDS patients on extracorporeal membrane oxygenation: preliminary experience. Intensive Care Med. 44, 123–124.

Moore, S., Gardiner, E., 2020. Point of care and intensive care lung ultrasound: a reference guide for practitioners during COVID-19. Radiography. (London, England: 1995)

Ouyang, X., Huo, J., Xia, L., Shan, F., Liu, J., Mo, Z., Yan, F., Ding, Z., Yang, Q., Song, B., et al., Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia. arXiv:2005.02690

Peng, Q., Wang, X., Zhang, L., 2020. Findings of lung ultrasonography of novel corona virus pneumonia during the 2019-2020 epidemic. Intensive Care Med. 1–2.

Poggiali, E., Dacrema, A., Bastoni, D., Tinelli, V., Demichele, E., Ramos, P.M., Marcianò, T., Silva, M., Vercelli, A., Magnacavallo, A., 2020. Can lung us help critical care clinicians in the early diagnosis of novel coronavirus (COVID-19) pneumonia? Radiology.

Qi, X., Jiang, Z., Yu, Q., Shao, C., Zhang, H., Yue, H., Ma, B., Wang, Y., Liu, C., Meng, X., et al., 2020. Machine learning-based CT radiomics model for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: a multicenter study. medRxiv.

Reissig, A., Copetti, R., Mathis, G., Mempel, C., Schuler, A., Zechner, P.M., Aliberti, S., Neumann, R., Kroegel, C., Hoyer, H., 2012. Lung ultrasound in the diagnosis and follow-up of community-acquired pneumonia: a prospective, multicenter, diagnostic accuracy study. Chest 142 (4), 965–972.

Roy, S., Menapace, W., Oei, S., Luijten, B., Fini, E., Saltori, C., Huijben, I., Chennakeshava, N., Mento, F., Sentelli, A., et al., 2020. Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound. IEEE Trans. Med. Imaging.

Shan, F., Gao, Y., Wang, J., Shi, W., Shi, N., Han, M., Xue, Z., Shen, D., Shi, Y., Lung infection quantification of COVID-19 in CT images with deep learning. arXiv:2003.04655

Shen, C., Yu, N., Cai, S., Zhou, J., Sheng, J., Liu, K., Zhou, H., Guo, Y., Niu, G., 2020. Quantitative computed tomography analysis for stratifying the severity of coronavirus disease 2019. J. Pharm. Anal..

Shi, F., Wang, J., Shi, J., Wu, Z., Wang, Q., Tang, Z., He, K., Shi, Y., Shen, D., 2020a. Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for COVID-19. IEEE Rev. Biomed. Eng..

Shi, F., Xia, L., Shan, F., Wu, D., Wei, Y., Yuan, H., Jiang, H., Gao, Y., Sui, H., Shen, D., Large-scale screening of COVID-19 from community acquired pneumonia using infection size-aware classification. arXiv:2003.09860

Smith, M., Hayward, S., Innes, S., Miller, A., 2020. Point-of-care lung ultrasound in patients with COVID-19–a narrative review. Anaesthesia.

Soldati, G., Smargiassi, A., Inchingolo, R., Buonsenso, D., Perrone, T., Briganti, D.F., Perlini, S., Torri, E., Mariani, A., Mossolani, E.E., Tursi, F., Mento, F., Demi, L., 2020a. Proposal for international standardization of the use of lung ultrasound for COVID-19 patients; a simple, quantitative, reproducible method.. J. Ultrasound Med..

Soldati, G., Smargiassi, A., Inchingolo, R., Buonsenso, D., Perrone, T., Briganti, D.F., Perlini, S., Torri, E., Mariani, A., Mossolani, E.E., et al., 2020b. Is there a role for lung ultrasound during the COVID-19 pandemic? J. Ultrasound Med..

Song, Y., Zheng, S., Li, L., Zhang, X., Zhang, X., Huang, Z., Chen, J., Zhao, H., Jie, Y., Wang, R., et al., 2020. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. medRxiv.

Sultan, L.R., Sehgal, C.M., 2020. A review of early experience in lung ultrasound in the diagnosis and management of COVID-19. Ultrasound Med. Biol.

Tan, M., Le, Q., 2019. EfficientNet: rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning, pp. 6105–6114.

Tang, Z., Zhao, W., Xie, X., Zhong, Z., Shi, F., Liu, J., Shen, D., Severity assessment of coronavirus disease 2019 (COVID-19) using quantitative features from chest CT images. arXiv:2003.11988

Tung-Chen, Y., 2020. Lung ultrasound in the monitoring of COVID-19 infection.. Clin. Med..

Via, G., Storti, E., Gulati, G., Neri, L., Mojoli, F., Braschi, A., 2012. Lung ultrasound in the ICU: from diagnostic instrument to respiratory monitoring tool. Minerva Anestesiol. 78 (11), 1282–1296.

Volpicelli, G., Lamorte, A., Villén, T., 2020. What's new in lung ultrasound during the COVID-19 pandemic. Intensive Care Med. 1–4.

Wang, S., Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., Cai, M., Yang, J., Li, Y., Meng, X., et al., 2020a. A deep learning algorithm using CT images to screen for corona virus disease (COVID-19). MedRxiv.

Wang, S., Zha, Y., Li, W., Wu, Q., Li, X., Niu, M., Wang, M., Qiu, X., Li, H., Yu, H., et al., 2020b. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. medRxiv.

WHO,. Coronavirus disease (COVID-19) pandemic. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019.

Wu, Y.-H., Gao, S.-H., Mei, J., Xu, J., Fan, D.-P., Zhao, C.-W., Cheng, M.-M., JCS: an explainable COVID-19 diagnosis system by joint classification and segmentation. arXiv:2004.07054

Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., Yu, L., Chen, Y., Su, J., Lang, G., Deep learning system to screen coronavirus disease 2019 pneumonia. arXiv:2002.09334

Yang, R., Li, X., Liu, H., Zhen, Y., Zhang, X., Xiong, Q., Luo, Y., Gao, C., Zeng, W., 2020. Chest CT severity score: an imaging tool for assessing severe COVID-19. Radiology 2 (2), e200047.

Zhang, K., Liu, X., Shen, J., Li, Z., Sang, Y., Wu, X., Cha, Y., Liang, W., Wang, C., Wang, K., et al., 2020. Clinically applicable ai system for accurate diagnosis, quantitative measurements and prognosis of COVID-19 pneumonia using computed tomography. Cell.

Zheng, C., Deng, X., Fu, Q., Zhou, Q., Feng, J., Ma, H., Liu, W., Wang, X., 2020. Deep learning-based detection for COVID-19 from chest CT using weak label. medRxiv.

Zhu, X., Song, B., Shi, F., Chen, Y., Hu, R., Gan, J., Zhang, W., Li, M., Wang, L., Gao, Y., et al., Joint prediction and time estimation of COVID-19 developing severe symptoms using chest CT scan. arXiv:2005.03405