Article

Research on the Application of Artificial Intelligence in Public Health Management: Leveraging Artificial Intelligence to Improve COVID-19 CT Image Diagnosis

Tiancheng He [1], Hong Liu [1,2], Zhihao Zhang [3], Chao Li [4,*] and Youmei Zhou [5,*]

1 Department of Political Party and State Governance, East China University of Political Science and Law, Shanghai 201620, China
2 Teacher Work Department of the Party Committee, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China
3 College of Computer Science and Technology, Shanghai University of Electric Power, Shanghai 200090, China
4 Department of Computer Science, Zhijiang College of Zhejiang University of Technology, Hangzhou 310024, China
5 Department of Landscape Architecture, College of Architecture and Urban Planning, Tongji University, Shanghai 200092, China
* Correspondence: cynthia0217@163.com (C.L.); 20310231@tongji.edu.cn (Y.Z.)

Abstract: Since the start of 2020, the outbreak of the Coronavirus disease (COVID-19) has been a global public health emergency, and it has caused unprecedented economic and social disaster. In order to improve the diagnosis efficiency of COVID-19 patients, a number of researchers have conducted extensive studies on applying artificial intelligence techniques to the analysis of COVID-19-related medical images. The automatic segmentation of lesions from computed tomography (CT) images using deep learning provides an important basis for the quantification and diagnosis of COVID-19 cases. For a deep learning-based CT diagnostic method, a few of accurate pixel-level labels are essential for the training process of a model. However, the translucent ground-glass area of the lesion usually leads to mislabeling while performing the manual labeling operation, which weakens the accuracy of the model. In this work, we propose a method for correcting rough labels; that is, to hierarchize these rough labels into precise ones by performing an analysis on the pixel distribution of the infected and normal areas in the lung. The proposed method corrects the incorrectly labeled pixels and enables the deep learning model to learn the infected degree of each infected pixel, with which an aiding system (named DLShelper) for COVID-19 CT image diagnosis using the hierarchical labels is also proposed. The DLShelper targets lesion segmentation from CT images, as well as the severity grading. The DLShelper assists medical staff in efficient diagnosis by providing rich auxiliary diagnostic information (including the severity grade, the proportions of the lesion and the visualization of the lesion area). A comprehensive experiment based on a public COVID-19 CT image dataset is also conducted, and the experimental results show that the DLShelper significantly improves the accuracy of segmentation for the lesion areas and also achieves a promising accuracy for the severity grading task.

Keywords: public health; COVID-19; artificial intelligence; automatic segmentation; CT image diagnosis

1. Introduction

Since the start of 2020, the outbreak of the Coronavirus disease (COVID-19) has been a globally public health emergency, and has caused unprecedented economic and social disaster [1]. It features a number of symptoms including endothelial barrier disruption, dysfunctional alveolar-capillary oxygen transmission, reduced oxygen diffusion capacity, alveolar wall thickening, increased vascular permeability and pulmonary oedema [2]. As a major global public health emergency, COVID-19 has once again proved that human beings

live in a "global risk society" with a common destiny, reminding us to be more alert to new and recurrent infectious diseases and to build a strong public health system to provide strong protection for people's health. There is no doubt that the development of COVID-19 has exceeded most people's expectations. Because of the rapid spread of COVID-19, the timely detection of the COVID-19 infection is essential to carry out the prompt isolation and treatment of COVID-19 patients. At present, reverse transcription-polymerase chain reaction (RT-PCR) is the most widely adopted method for the detection of COVID-19. However, RT-PCR suffers from several limitations: (1) it is time-consuming (requiring over 3 h to complete the detection process); (2) there is limited supply of test kits; (3) poor sampling quality causes false negatives [3]. Nowadays, CT plays an important role for detecting COVID-19 [3], and bilateral patchy shadows or ground-glass opacity in the lung can be clearly identified from chest-computed tomography (CT) images captured from COVID-19 patients [4]. In addition, compared with RT-PCR, the operation of chest CT is easy, and with chest CT we can judge the severity of the disease. Therefore, CT could serve as a practical method for the diagnosis of COVID-19. Moreover, to assess the severity of COVID-19, contouring the infected area is an essential procedure for an image diagnosis. However, the traditional manual contouring operation is tedious and time-consuming, and it heavily depends on the clinical experience of physicians. With the increase in the number of infected patients, the workload of radiologists has significantly increased; hence, an automatic CT image segmentation method for COVID-19 diagnosis is urgently expected.

Deep learning technology has been widely adopted in medical image segmentation due to its capability of feature extraction [5]. Deep learning methods show excellent performance in the task of the lesion segmentation of COVID-19, but large-scale labelled samples must be available prior to applying these deep learning methods [6–9]. The task of collecting sufficient COVID-19 CT images and accurately labelling them at a pixel level is time-consuming and costly. To tackle this issue, some methods employ data augmentation [10,11] and image synthesis [12,13] to extract the information from limited labeled images, but they usually suffer from poor generalization on different datasets. Other methods applying semi-supervised [14,15] and unsupervised learning [16,17] fail to achieve good performance due to the large variations of infection on CT images, such as irregular shapes and ambiguous boundaries [18].

Not only the quantity but also the quality of the pixel-level label restricts the training process of deep learning methods. By reviewing these dominant public COVID-19 CT image datasets, we found that: (1) the quality of the dataset is uneven because it is susceptible to the experience of the physician; and (2) the translucent ground-glass characteristics of the infected area are hard to accurately identify, and further lead to the labeling of some non-infected areas as infected areas (such as lung parenchyma and pulmonary vessels). In this study, we aim to correct these mislabeled labels to provide well-labelled datasets for model training. Rough labels (with mislabeled ones) are hierarchized according to the pixel distribution of the infected and normal areas in the lung image. The proposed method reassigns the mislabeled pixels and enables the deep learning model to learn the infected degree of each infected pixel. With the hierarchical labels, we propose a deep learning-based aiding system (named DLShelper) for COVID-19 diagnosis. The DLShelper performs lesion (in lung) segmentation from CT images, as well as the task of severity grading. A multi-layer preceptor (MLP) is used as a classifier. The proportion of the lesion to the lung and the proportion of each grade in the lesion are used as input features. Rich auxiliary diagnostic information (e.g., the severity grade, the proportion of the infected area and the visualization of the infected area) are provided for the physicians in clinic. The main contributions in this paper are as follows:

- In order to improve the performance of segmentation on COVID-19 infection, a label refinement method is proposed to refine the existing labels from rough to precise. The refinement reassigns the incorrectly labeled pixels and enables the network to learn the infection degree of each infected pixel.

- Aiming to assist physicians in the efficient diagnosis of COVID-19, a deep learning-aided system (named DLSHELPER) using refined hierarchical labels is proposed. DLSHELPER provides rich auxiliary diagnostic information, including the proposed severity grade, proportion of infected area and infected area visualization.
- We validate the accuracy of our method for COVID-19 lesion segmentation and grading on public COVID-19 CT datasets.

At present, it is an important opportunity to change the public health governance system. This study takes the diagnosis of COVID-19 as an example to explore the enabling effect of AI in the management of public health emergencies.

The rest of the paper is organized as follows: Section 2 introduces the related work. Section 3 details the proposed method for COVID-19 CT image diagnosis. Section 4 presents the experiment and discussion. Finally, Section 5 concludes the study.

2. Related Work

In recent years, the intelligent analysis of medical images based on artificial intelligence has been extensively researched [19]. Santosh et al. [20] proposed a lung feature detection model based on multi-feature parameters, and it achieved an accuracy of up to 91%. Pratondo et al. [21] combined multiple machine learning models and a region-based contouring algorithm for the task of medical image segmentation. Ahmad et al. [22] used the Content-Based medical image retrieval algorithm for lung segmentation. However, its Jaccard similarity coefficient was only 0.870. Shepherd et al. [23] proposed a statistical model based on shape prior for segmentation combined with online/offline learning models. Xu et al. proposed a method for lung function assessment based on cough sound [24]. Shaukat et al. [25] developed a fully automated method to detect lung nodules using a hybrid feature set of SVM and achieved a promising accuracy. Souza et al. [26] proposed a Deep Convolutional Neural Network method (DCNN) for fully automated lung segmentation. Park et al. [27] used DCNN for lung CT image segmentation. Although DCNN is capable of learning complex data, it is overly dependent on the amount of data used in the training process. Besides, the size of the data also impacts the performance of the model.

The quality of these public CT image datasets is uneven because they are susceptible to the experience of the physician. In addition, in contrast to the semantically segmented objects, the COVID-19 lesion is translucent and of a low contrast with the surroundings. The labeling operation is conducted manually; hence, the labeling process unavoidably involves human errors. Some normal pixels are mislabeled as infected ones in the situations where: (1) the lung parenchyma pixels are entrapped between lesion pixels; (2) other areas are tissues such as pulmonary vessels. These mislabeled pixels will weaken the performance of model training.

3. Method

3.1. Overview

As shown in Figure 1, the functions of the proposed method include: (1) lung segmentation; (2) lesion label refinement; (3) lesion segmentation; and (4) severity grading. The original CT images and lung parenchyma labels are used to train a two-category semantic segmentation network, which is used for segmenting the lung parenchyma image. With these segmented lung parenchyma images and lesion labels, the infected and normal areas can be identified. By further analyzing the pixel distribution in these two areas, the mislabeled pixels can be corrected and these pixels can be hierarchized to different levels according to the value of each pixel so that these rough lesion labels are finally refined to accurate hierarchical labels. The "level" not only represents the value of a pixel, but also indicates the infected degree of the area in which the pixel is contained. The lung parenchyma images and refined hierarchical labels are used to train a multi-category semantic segmentation network, then we use it to segment the lesion areas. Different output lesions are converted into different colors to generate a hierarchical visual map that

provides intuitive information for auxiliary diagnosis. We calculate the proportion of three categories in the lesion area, respectively. Then, the total proportion of the lesion to the whole lung parenchyma is provided as other information for auxiliary diagnosis. Moreover, these four radiological features are used as input parameters for the severity grading, which is based on a three-layer multi-layer preceptor (MLP). In summary, there are three types of information provided to physicians by the proposed system: (1) the hierarchical visual map; (2) the proportion of the lesion in the lung area; and (3) the severity grade.
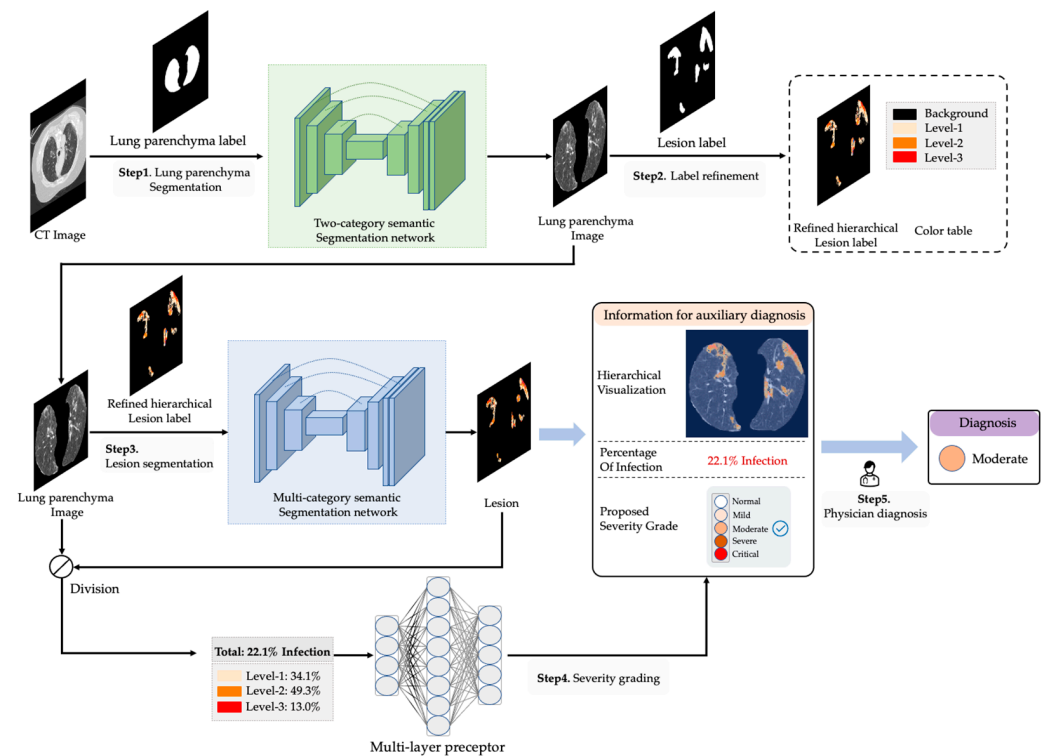

Figure 1. The full workflow of the proposed auxiliary diagnosis system.

3.2. Label Refinement

As discussed in Section 2, with the traditional method pixels marked as infected may contain normal pixels. Moreover, the traditional strategy for lesion labeling only includes the categories: infected (marked as 1) and normal (marked as 0), which ignores the information contained in the infected pixels; e.g., for each pixel in the infected area, the higher the value, the more serious the infection is.

For two lesions with the same area (assuming that the area of the lung parenchyma in which they are located is also equal and the value of pixel falls in the range (0–255)), the more the grayscale distribution approaches 255, the more serious the infection is in clinical diagnosis. As shown in Figure 2, we selected four CT images from different severity grades and calculated grayscale histograms of their lesions. The results reveal a positive correlation between the grayscale distribution of the lesion with its severity. However, there is no accurate metric to measure the grayscale distribution, Therefore, we hierarchize the infected area to a different level according to its pixel value, and the grayscale distribution can be described by the percentage of pixels at different levels in the lesion.
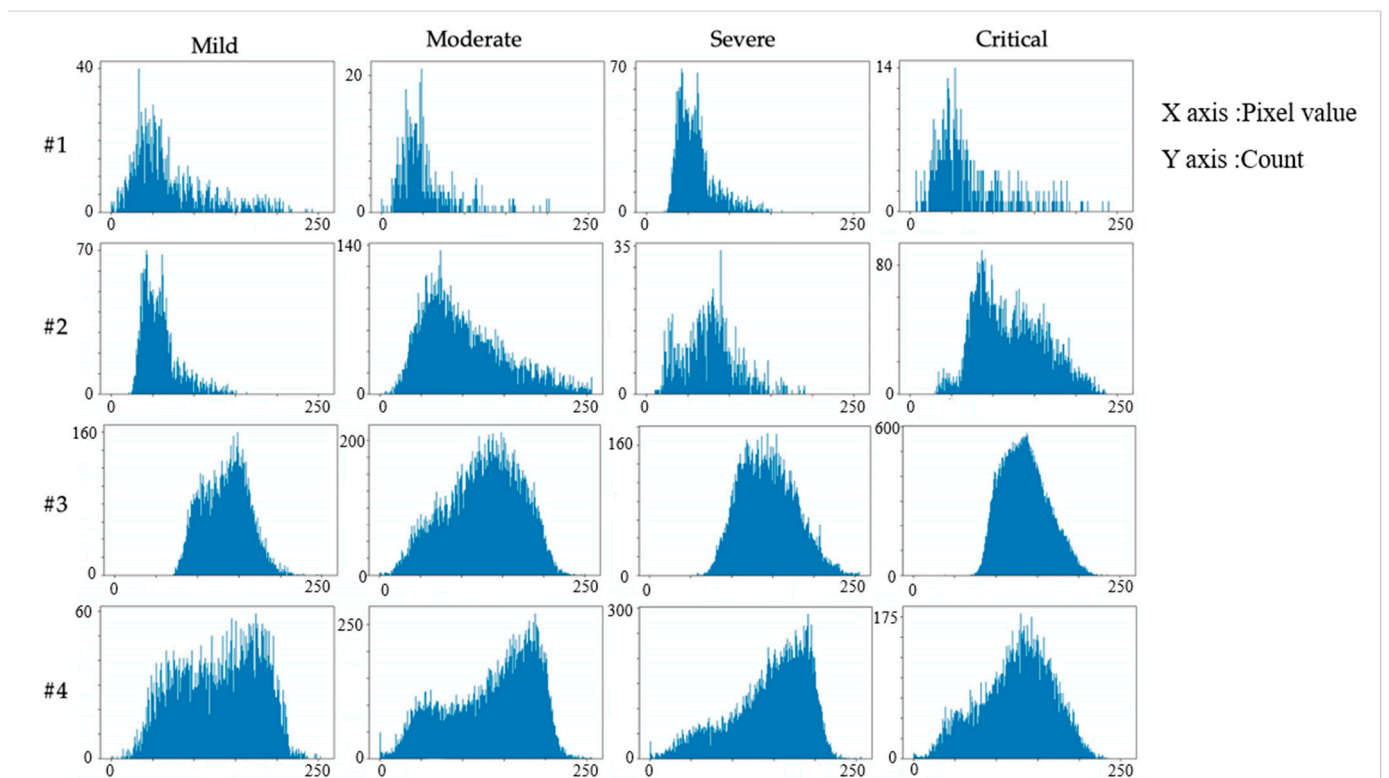
Int. J. Environ. Res. Public Health **2023**, *20*, 1158

5 of 12

Figure 2. The comparison of the grayscale histograms of four CT images (lesions) from different severity grades.

We denote the CT image as *I* and its corresponding lesion label as *M*. *M* has the same size as *I*. We obtain the lung pixel from *I* by applying lung segmentation and denote it as $\mathcal{O}_{Lung}$. Then, the lesion in *I* is obtained by the mask operation in *I* and *M*, and we denote it as $\mathcal{O}_{Infected}$. The complement of $\mathcal{O}_{Infected}$ in $\mathcal{O}_{Lung}$ is $\mathcal{O}_{No\text{-}infected}$ (the normal pixel in the lung). These processes are formulized as:

$$I_{Lung} = N_{Two\text{-}catagory}(I) \otimes I, \ \mathcal{O}_{Lung} = \left\{ p \in I_{Lung} \mid I_{Lung}(p) > 0 \right\} \tag{1}$$

$$I_{Infected} = I \otimes M, \ \mathcal{O}_{Infected} = \left\{ p \in I_{Infected} \mid I_{Infected}(p) > 0 \right\} \tag{2}$$

$$\mathcal{O}_{No\text{-}infected} = \complement_{\mathcal{O}_{Lung}} \mathcal{O}_{Infected} \tag{3}$$

where $N_{Two\text{-}catagory}$ denotes the network for lung parenchyma segmentation, $\otimes$ denotes the element-wise multiplication and *p* denotes the pixel in the image; $\complement_a b$ denotes the complement of *a* in set B.

We denote the average value of $\mathcal{O}_{No\text{-}infected}$ as *a* and the maximum of $\mathcal{O}_{Infected}$ as *b*, respectively. Given the pixel of $\mathcal{O}_{Infected}$ can be divided into Grade-*g*, the pixel with a value less than *a* and greater than or equal to Grade-*g* will be reassigned as the background. These processes are formulized as:

$$a = \mathrm{Mean}\left(\mathcal{O}_{No\text{-}infected} \right) \tag{4}$$

$$b = \mathrm{Max}\left(\mathcal{O}_{Infected} \right) \tag{5}$$

$$s = (b - a)/(g + 1) \tag{6}$$

$$R_0 \in [0, a] \cup (b - s, 255] \tag{7}$$

$$R_i \in (a + (i - 1) * s, \ a + i * s], \ 0 < i \leq g \tag{8}$$

Int. J. Environ. Res. Public Health **2023**, 20, 1158

6 of 12

where s denotes the interval between grades. R_0 represents the range of pixel values of the background; meanwhile, R_i represents the range of pixel values of Grade-i ($0 < i \leq g$). Finally, we assign the pixel belonging to each grade in I_{Lung}; i.e., the value of the Grade-i pixel is set to i; the value of the background pixel is set to 0. Thus, a refined hierarchical label is generated. Of note, the reason for the pixel with a value greater than or equal to Grade-g being reassigned as the background is that lung trachea and blood vessels may be contained in these pixels. A value of g that is too small or too large will impact the accuracy of the label refinement; hence, we use it as a hyper-parameter and compulsorily set it to 3 (according to experimental results). As shown in Figure 3c, mislabeled infected pixels are corrected to normal ones, and these infected pixels are hierarchized to different grades.
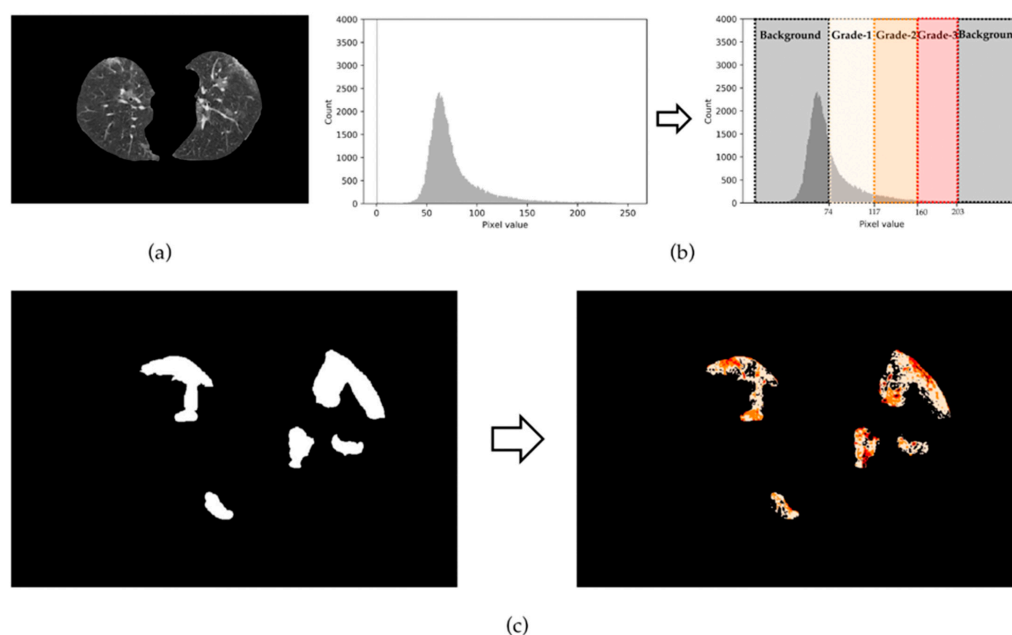

Figure 3. Visualization of label refinement (g = 3): (**a**) is the lung image segmented from a CT image. The left part of (**b**) is the histogram of (**a**) and we grade the pixels based on the pixel distribution of the infected and uninfected areas in the lung; (**c**) shows the lesion labels before and after the refinement.

3.3. Lung and Lesion Segmentation

Traditional segmentation models (especially UNet [28]) have achieved good performance on segmentation tasks for lung and COVID-19 lesions. UNet adopts symmetric encoding and decoding paths to aggregate semantic information and recover spatial information with the help of shortcut connections, and it is suitable for medical image segmentation. Thus, in this study, we adopt UNet for lung and lesion segmentation. In addition, we use a multiple-category training strategy (instead of the traditional two-category strategy) to learn the grades of pixels in the lesion.

With the completion of network training, a CT image will be input to the UNet to segment a lung image. Then, the obtained lung image is input to the multiple-category segmentation network to obtain the COVID-19 lesion. Based on these different categories of lesions, a colorful visualized map is generated hierarchically.

3.4. Severity Grading

As reported in [26], the number, quadrant and area of lesions in CT images are important factors to determine the severity of the COVID-19 case. However, as the area of lung parenchyma in a volume of continuous CT image slices is different, it is inappropriate to use a fixed value as the threshold to determine the grade of severity. In this work, we calculated the proportion of all lesions in the lung parenchyma to address this issue. We found that the higher the value, the whiter the pixel appears in the lesion area, so the

level of the "white" pixel in the lesion area and the density of the white pixel can be taken as indicators for determining the grade of severity. With regard to these indicators, a multilayer perceptron (MLP) is used as a classifier for severity grading.

The multilayer perceptron is a feedforward artificial neural network that uses supervised back-propagation, which is widely used for nonlinear classifications. As shown in Figure 4, the MLP in the proposed method consists of an input layer, a hidden layer and an output layer. The Relu function is used as the activation function of the hidden layer and the SoftMax function is used as the activation function of the output layer for the classification. The number of neurons in the hidden layer is determined by an empirical formula:

$$k = \sqrt{m + n} + a \qquad (9)$$

where k denotes the number of neurons in the hidden layer, n denotes the number of neurons in the input layer, m denotes the number of neurons in the output layer, and a denotes a constant between 1 and 10.
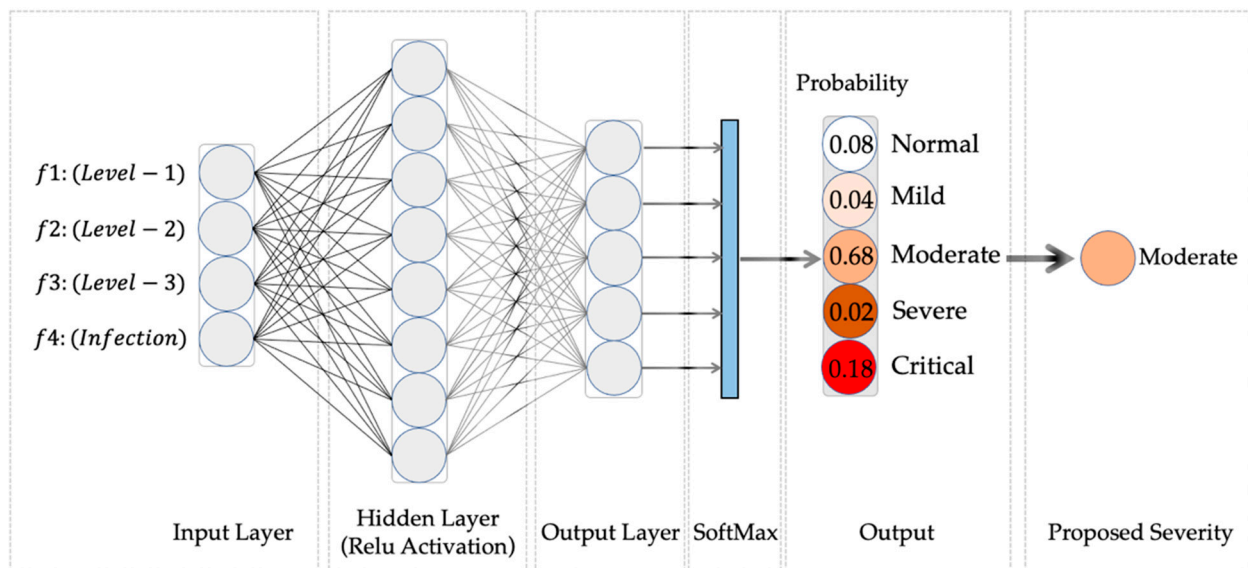

Figure 4. Structure of the multi-layer preceptor network (g = 3). f_{1-4} denotes the input features of the MLP network. f_{1-3} denotes the proportion of infected pixels of three levels in all infected pixels. f_4 denotes the proportion of infected pixels in lung parenchyma pixels.

4. Experiments and Analysis

The dataset [29] used in this study contains about 3500 CT image slices and corresponding lung and lesion segmentation labels. In addition, we recruited a radiology graduate student to label each CT image with a grade of severity (e.g., normal, mild, moderate, severe and critical). The labels were then verified by an experienced radiology specialist for reliability.

4.1. Implementation and Evaluation

A two-stage training strategy is adopted in this experiment: (1) training the segmentation of lung and COVID-19 lesions; (2) oversampling the training set, and finally training the MLP. We reproduced all the related networks and modules in the Pytorch framework. When training the segmentation network, we set the number of the batch size to 1, then initialize the network weights with Kaiming initialization, set network biases to zero and train the positive/negative samples alternately. In addition, the training set is shuffled in each iteration. We use different metrics in different stages. Intersection over union (IoU), sensitivity (SEN), specificity (SPE) and Dice similarity coefficient (DSC) are used to evaluate the accuracy of the lung segmentation. Besides, mean intersection over union (mIoU), mean pixel accuracy (mPA) and class pixel accuracy (CPA) are used to evaluate the accuracy

Int. J. Environ. Res. Public Health **2023**, *20*, 1158

8 of 12

of COVID-19 lesion segmentation using original labels and refined hierarchical labels. Precision is used to evaluate the accuracy of the severity grading. The above-mentioned metrics can be calculated as:

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{10}$$

$$\text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{11}$$

$$\text{SPE} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{12}$$

$$\text{DSC} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \tag{13}$$

$$\text{mIoU} = \frac{1}{k+1} \sum_{i=0}^{k} \frac{\text{TP}}{\text{FN} + \text{FP} + \text{TP}} \tag{14}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{15}$$

$$\text{mPA} = \frac{1}{k+1} \sum_{i=0}^{k} \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{16}$$

where TP denotes true positives, TN denotes true negatives, FP denotes false positives and FN denotes false negatives.

We optimize the lung and lesion segmentation networks using a binary cross-entropy L_{lung} and a multi-category cross-entropy loss L_{lesion}, respectively, using a mean-squared error loss to train the MLP L_{mlp}.

$$L_{lung}(a, b) = -[b log(a) + (1 - b) \log(1 - a)] \tag{17}$$

$$L_{lesion}(a, b) = -\sum_{m=0}^{g} b_m log f(a)_m \tag{18}$$

$$f(a)_m = \frac{e^{a_m}}{\sum_{n=0,\ n \neq m}^{g} e^{a_n}} \tag{19}$$

$$L_{mlp}(a, b) = \|a - b\|^2 \tag{20}$$

$$L = L_{lung} + L_{lesion} + L_{mlp} \tag{21}$$

where a is the ground truth and b is the predicted result. g is the number of grades in refined hierarchical labels.

4.2. Evaluation of Lung Segmentation

As shown in Table 1 and Figure 5, lung segmentation with UNet works efficiently, with DSC up to 96%. Besides, IoU, SEN and SPE all surpass 90%. The accurate segmentation of the lung parenchyma ensures the quality of subsequent COVID-19 lesion segmentation and severity grading.

Table 1. Obtained result of lung segmentation metrics.

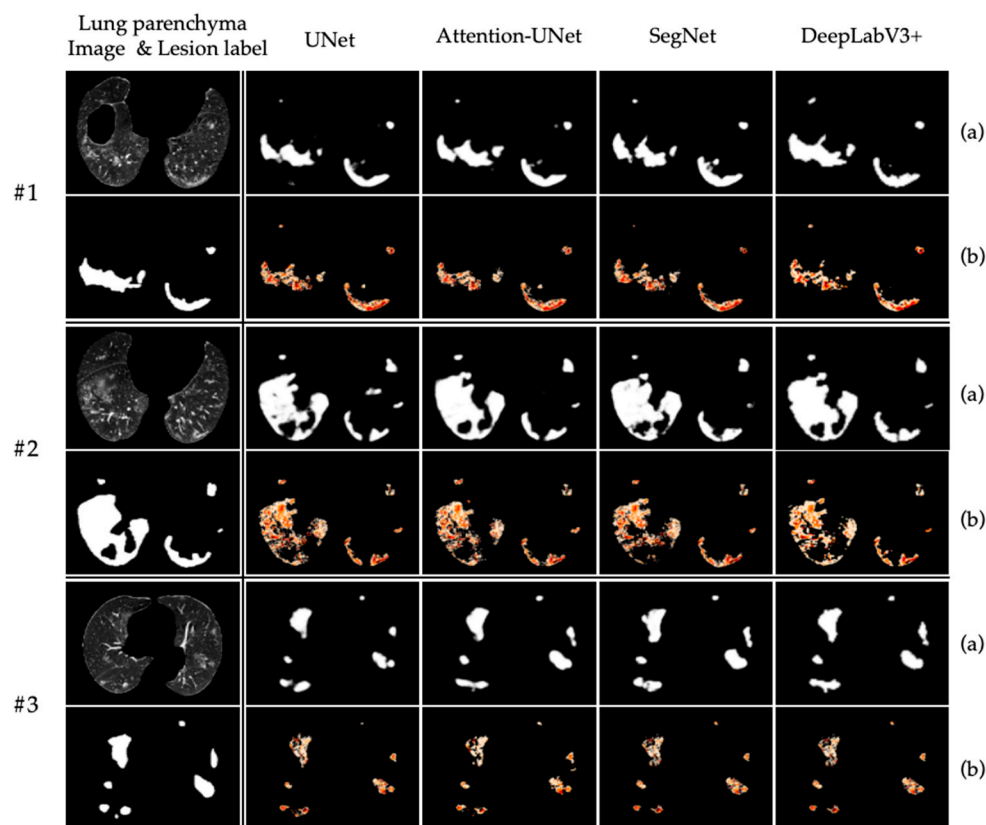| Metric	IoU	DSC	SEN	SPE
Value	0.94	0.96	0.97	1.00

Figure 5. Visualization of the lesion segmentation using original label and refined hierarchical label. (**a**) and (**b**) illustrate the segmentation results using original lesion labels and refined hierarchical labels, respectively.

4.3. Evaluation of COVID-19 Lesion Segmentation Using Refined Hierarchical Labels

To evaluate the performance of refined hierarchical labels for COVID-19 lesion segmentation, four state-of-the-art networks are selected and trained with original labels and refined hierarchical labels, respectively (as shown in Figure 5). With regard to these widely used metrics (e.g., IoU, DSC, SEN and SPE) for medical image segmentation, an evaluation is carried out. Table 2 shows the values of these four metrics of the model trained with original labels and refined hierarchical labels. Table 3 shows the values of these four metrics, the CPA of each level and the MIoU and MPA of the model trained with refined hierarchical labels.

With the original labels, DeepLabV3+ achieves the best DICE of 82.94% among all the networks. Meanwhile, UNet achieves the worst performance. However, we find that the area marked as a lesion by the original labels in the input image contains many normal pixels such as lung parenchyma and pulmonary vessels. As illustrated in Figure 5, the #2 image is the most mislabeled. By introducing the refined hierarchical labels, the segmentation network can not only accurately identify the infected pixels, but also filter out these mislabeled pixels. Besides, as shown in Table 2, with the introduction of refined hierarchical labels, the model achieves better performance, such as the DSC of UNet and Attention-UNet reaching 83.47% and 82.35%, respectively. As shown in Table 3, the performance of pixel segmentation with UNet (2) is the best. Because the ground truth used is different, we cannot directly compare the performance of models training on original labels and refined hierarchical labels. Experienced radiologists from a hospital in Zhejiang Province confirm that refined hierarchical labels bring more precise results.

Int. J. Environ. Res. Public Health **2023**, 20, 1158

10 of 12

Table 2. Comparison of lesion segmentation performance using original labels and refined hierarchical labels.

| Network | Labels | | IoU (%) | DSC (%) | SEN (%) | SPE (%) |
	Original	Refined				
UNet [28]	√		68.02	78.36	82.31	99.74
		√	73.54	83.47	87.75	99.83
Attention-UNet [30]	√		71.89	82.13	85.51	99.78
		√	71.82	82.35	82.37	99.87
SegNet [31]	√		68.98	79.57	83.38	99.74
		√	68.49	79.54	78.85	99.87
DeepLabV3+ [32]	√		72.43	82.94	85.08	99.79
		√	68.71	80.48	79.89	99.84

Table 3. Detailed performance using refined hierarchical labels.

Network	g	CPA (%)	IoU (%)	DSC (%)	SEN (%)	SPE (%)	MIoU (%)	MPA (%)
UNet	1	77.30	56.52	70.71	77.30	99.87		
	2	81.31	65.47	77.20	81.31	99.92	61.04	75.73
	3	75.33	61.13	72.93	75.33	99.95		
Attention-UNet	1	70.30	53.24	67.80	70.30	99.88		
	2	73.95	62.82	75.17	73.95	99.94	58.27	71.56
	3	70.41	58.75	70.81	70.41	99.96		
SegNet	1	55.58	41.31	56.20	55.58	99.84		
	2	71.21	52.55	66.77	72.22	99.86	51.96	67.74
	3	66.59	49.40	63.18	66.59	99.92		
DeepLabV3+	1	55.21	40.62	56.20	55.21	99.85		
	2	61.62	44.43	60.37	61.62	99.86	41.69	57.05
	3	54.34	40.03	55.11	54.34	99.91		

4.4. Evaluation of COVID-19 Severity Grading

There are few samples of mild, severe and critical cases in the dataset. As shown in Table 4, the classification accuracy of these categories is very low, even as low as 0 (mild). To solve this category imbalance problem, we applied the Synthetic Minority Oversampling Technique (SMOTE) [33] to minority classes. With the operation of oversampling the dataset, all the categories of samples reached a balance, the classification accuracy of mild reached 100% and the classification accuracy of severe and critical increased by 19.81% and 10.25%, respectively. Moreover, the overall accuracy was 98.82%.

Table 4. Precision of Severity Grading.

Severity	Precision (%)	Precision (%) (With Operation of Oversampling)
Normal	1	1
Mild	0	1
Moderate	99.37%	99.38%
Severe	62.69%	82.50%
Critical	79.49%	89.74%
Total	96.49	98.82

Int. J. Environ. Res. Public Health **2023**, *20*, 1158

11 of 12

5. Conclusions

In this study, we propose a method for refining lesion labels from rough to precise. Then, a deep learning-based aiding system for CT image diagnosis using refined labels is developed. It performs lung and lesion segmentation from CT images, as well as severity grading. A multi-layer preceptor is used as a classifier, and the proportion of the lesion to the lung and the proportion of each grade in the lesion are used as input features. Auxiliary diagnostic information including the severity grade, proportion of infected area and visualization of the infected area are provided by the DLShelper for physicians in clinic. A comparative experiment based on public datasets is carried out, and the experimental results show that the proposed method achieves better accuracy in comparison with several state-of-the-art networks. Besides, the proposed method achieves a high accuracy for severity grading. In future, we will develop a new metric to describe the grayscale distribution features so as to further improve the performance.

In COVID-19 prevention and control, while developing AI and playing its positive role, we should be alert to the social risks and ethical challenges brought by AI itself, carry out responsive and principled scientific and technological governance, and strengthen ethical review and data legislation under the principle of "harmony, friendship, fairness, inclusiveness and sharing, respect for privacy, security and controllability, shared responsibility, open cooperation, and agile governance".

Author Contributions: Conceptualization, Z.Z. and T.H.; methodology, H.L.; software, Z.Z.; validation, C.L.; writing—original draft preparation, Z.Z.; writing—review and editing, C.L.; visualization, Y.Z.; project administration, Y.Z. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data are available by request to the corresponding authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, P.; Zhong, Y.; Deng, Y.; Tang, X.; Li, X. CoSinGAN: Learning COVID-19 Infection Segmentation from a Single Radiological Image. *Diagnostics* **2020**, *10*, 901. [CrossRef] [PubMed]
2. Bertolini, M.; Brambilla, A.; Dallasta, S.; Colombo, G. High-quality chest CT segmentation to assess the impact of COVID-19 disease. *Int. J. Comput. Assist. Radiol. Surg.* **2021**, *16*, 1737–1747. [CrossRef]
3. Ai, T.; Yang, Z.; Hou, H.; Zhan, C.; Chen, C.; Lv, W.; Tao, Q.; Sun, Z.; Xia, L. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases. *Radiology* **2020**, *296*, E32–E40. [CrossRef] [PubMed]
4. Adams, H.J.; Kwee, T.C.; Yakar, D.; Hope, M.D.; Kwee, R.M. Chest CT imaging signature of coronavirus disease 2019 infection: In pursuit of the scientific evidence. *Chest* **2020**, *158*, 1885–1895. [CrossRef] [PubMed]
5. Shi, F.; Wang, J.; Shi, J.; Wu, Z.; Wang, Q.; Tang, Z.; He, K.; Shi, Y.; Shen, D. Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19. *IEEE Rev. Biomed. Eng.* **2020**, *14*, 4–15. [CrossRef]
6. Zhang, K.; Liu, X.; Shen, J.; Li, Z.; Sang, Y.; Wu, X.; Zha, Y.; Liang, W.; Wang, C.; Wang, K.; et al. Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. *Cell* **2020**, *181*, 1423–1433. [CrossRef] [PubMed]
7. Wu, Y.H.; Gao, S.H.; Mei, J.; Xu, J.; Fan, D.P.; Zhang, R.G.; Cheng, M.M. JCS: An Explainable COVID-19 Diagnosis System by Joint Classification and Segmentation. *IEEE Trans. Image Process.* **2020**, *30*, 3113–3126. [CrossRef]
8. 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* **2020**, arXiv:2003.04655.
9. Qiu, Y.; Liu, Y.; Xu, J. MiniSeg: An Extremely Minimum Network for Efficient COVID-19 Segmentation. *Proc. Conf. AAAI Artif. Intell.* **2020**, *35*, 4846–4854. [CrossRef]
10. Çiçek, Ö.; Abdulkadir, A.; Lienkamp, S.S.; Brox, T.; Ronneberger, O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Cham, Switzerland, 2016; pp. 424–432.

Int. J. Environ. Res. Public Health **2023**, 20, 1158

12 of 12

11. Zhang, L.; Wang, X.; Yang, D.; Sanford, T.; Harmon, S.; Turkbey, B.; Roth, H.; Myronenko, A.; Xu, D.; Xu, Z. When unseen domain generalization is unnecessary? rethinking data augmentation. *arXiv* **2019**, arXiv:1906.03347.

12. Jin, D.; Xu, Z.; Tang, Y.; Harrison, A.P.; Mollura, D.J. CT-realistic lung nodule simulation from 3D conditional generative adversarial networks for robust lung segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Cham, Switzerland, 2018; pp. 732–740.

13. Shin, H.-C.; Tenenholtz, N.A.; Rogers, J.K.; Schwarz, C.G.; Senjem, M.L.; Gunter, J.L.; Andriole, K.P.; Michalski, M. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In *International Workshop on Simulation and Synthesis in Medical Imaging*; Springer: Cham, Switzerland, 2018; pp. 1–11.

14. Xu, Z.; Niethammer, M. DeepAtlas: Joint semi-supervised learning of image registration and segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Cham, Switzerland, 2019; pp. 420–429.

15. Fan, D.-P.; Zhou, T.; Ji, G.-P.; Zhou, Y.; Chen, G.; Fu, H.; Shen, J.; Shao, L. Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images. *IEEE Trans. Med. Imaging* **2020**, *39*, 2626–2637. [CrossRef] [PubMed]

16. Shan, S.; Yan, W.; Guo, X.; Chang, E.I.; Fan, Y.; Xu, Y. Unsupervised end-to-end learning for deformable medical image registration. *arXiv* **2017**, arXiv:1711.08608.

17. de Vos, B.D.; Berendsen, F.F.; Viergever, M.A.; Sokooti, H.; Staring, M.; Išgum, I. A deep learning framework for unsupervised affine and deformable image registration. *Med. Image Anal.* **2019**, *52*, 128–143. [CrossRef] [PubMed]

18. Wang, G.; Liu, X.; Li, C.; Xu, Z.; Ruan, J.; Zhu, H.; Meng, T.; Li, K.; Huang, N.; Zhang, S. A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT images. *IEEE Trans. Med. Imaging* **2020**, *39*, 2653–2663. [CrossRef] [PubMed]

19. Chen, J.; Li, Y.; Guo, L.; Zhou, X.; Zhu, Y.; He, Q.; Han, H.; Feng, Q. Machine learning techniques for CT imaging diagnosis of novel coronavirus pneumonia: A review. *Neural Comput. Applications.* **2022**. [CrossRef] [PubMed]

20. Santosh, K.; Antani, S. Automated chest X-ray screening: Can lung region symmetry help detect pulmonary abnormalities? *IEEE Trans. Med. Imaging* **2017**, *37*, 1168–1177. [CrossRef]

21. Pratondo, A.; Chui, C.K.; Ong, S.H. Integrating machine learning with region-based active contour models in medical image segmentation. *J. Vis. Commun. Image Represent.* **2017**, *43*, 1–9. [CrossRef]

22. Ahmad, W.; Zaki, W.; Fauzi, M. Lung segmentation on standard and mobile chest radiographs using oriented Gaussian derivatives filter. *Biomed. Eng. Online* **2015**, *14*, 20. [CrossRef]

23. Shepherd, T.; Prince, S.J.; Alexander, D.C. Interactive lesion segmentation with shape priors from offline and online learning. *IEEE Trans. Med. Imaging* **2012**, *31*, 1698–1712. [CrossRef]

24. Xu, W.; He, G.; Pan, C.; Shen, D.; Zhang, N.; Jiang, P.; Liu, F.; Chen, J. A Forced Cough Sound based Pulmonary Function Assessment by Using Machine Learning. *Front. Public Health.* **2022**, *10*, 1015876. [CrossRef]

25. Shaukat, F.; Raja, G.; Gooya, A.; Frangi, A. Fully automatic detection of lung nodules in CT images using a hybrid feature set. *Med. Phys.* **2017**, *44*, 3615–3629. [CrossRef] [PubMed]

26. Souza, J.C.; Diniz, J.O.B.; Ferreira, J.L.; da Silva, G.L.F.; Silva, A.C.; de Paiva, A.C. An automatic method for lung segmentation and reconstruction in chest X-ray using deep neural networks. *Comput. Methods Programs Biomed.* **2019**, *177*, 285–296. [CrossRef] [PubMed]

27. Park, B.; Park, H.; Lee, S.M.; Seo, J.B.; Kim, N. Lung Segmentation on HRCT and Volumetric CT for Diffuse Interstitial Lung Disease Using Deep Convolutional Neural Networks. *J. Digit. Imaging* **2019**, *32*, 1019–1026. [CrossRef] [PubMed]

28. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Navab, N., Hornegger, J., Wells, W., Frangi, A., Eds.; MICCAI 2015, Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2015; Volume 9351. [CrossRef]

29. Jun, M.; Cheng, G.; Yixin, W.; Xingle, A.; Jiantao, G.; Ziqi, Y.; Jian, H. *COVID-19 CT Lung and Infection Segmentation Dataset (Verson 1.0) [Data set]*; Zenodo: Geneva, Switzerland, 2020.

30. Oktay, O.; Schlemper, J.; Le Folgoc, L.; Lee, M.; Heinrich, M.; Misawa, K.; Rueckert, D. Attention U-Net: Learning Where to Look for the Pancreas. *arXiv* **2018**, arXiv:1804.03999.

31. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]

32. Firdaus-Nawi, M.; Noraini, O.; Sabri, M.Y.; Siti-Zahrah, A.; Zamri-Saad, M.; Latifah, H. DeepLabv3+ _encoder-decoder with Atrous separable convolution for semantic image segmentation. *Pertanika J. Trop. Agric. Sci.* **2011**, *34*, 137–143.

33. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]