Richard Xue

Shanghai American School Puxi Campus

Shanghai, China

Professor Feng Rui, Professor Guo Yi

Fudan University

Multi-DeepNet: A Novel Weakly-Supervised Multi-Task and Multi-View-Oriented Convolution Neural Network for COVID-19 Diagnosis from CT Images

Multi-DeepNet: A Novel Weakly-Supervised Multi-Task and Multi-View-Oriented Convolution Neural Network for COVID-19 Diagnosis from CT Images

Abstract—Currently, manual analysis performed by professional radiologists is required for COVID-19 diagnosis given the patient's chest Computed Tomography (CT) images, but this process is inefficient and costly. Deep learning methods can provide computer vision-based solutions to help guide radiologists perform faster and more accurate diagnosis. However, current well performed methods require training on large and balanced datasets with pixel level lung lesion annotations, both of which are not easily accessible. Moreover, visual similarities between COVID-19 and other pneumonia in CT scans make it difficult to learn their distinguishing features. To address these issues, we propose a novel weakly-supervised deep learning model, named Multi-DeepNet, that can be well trained to perform fine-grained classification on small and imbalanced datasets. Specifically, a multi-task pre-training module is introduced to better extract distinguishing features between COVID-19 and other similar pneumonia. Furthermore, a multi-view-oriented classifier is proposed to extract complimentary information from the axial, coronal and sagittal planes. Experimental results demonstrate that our Multi-DeepNet achieves superior sensitivities, specificity, and accuracies compared to state-of-the-art methods. zzzz

Author: Richard Xue

Keywords— COVID-19 Diagnosis, Deep Learning, Multi-Task, Multi-View

Table of Contents

I. Introduction ***3***
II. Methodology ***4***
A. Image Preprocessing **4**
B. Multi-Task Pre-Training Module **5**
C. Multi-View-Oriented Classifier **6**
III. Experimental Setups ***8***
A. Dataset **8**
B. Implementation Details **8**
C. Evaluation Metrics **8**
IV. Results ***8***
V. Discussion and Conclusion ***10***
VI. References ***12***
VII. Acknowledgements ***13***

I. Introduction

Since December 2019, a novel coronavirus known as COVID-19 has plagued the world. According to the World Health Organization (WHO), human-to-human transmissions resulted in a rapid spread globally. By June 15th, 2021, 175,847,347 people have been infected, causing 3,807,276 deaths. Currently, the reverse transcription polymerase chain reaction test (RT-PCR) is used as the golden standard for COVID-19 diagnosis. However, in its beginning phases, the sensitivity of the RT-PCR test had been relatively low at 59% [1]. In addition, lack of testing kits in some regions made diagnosis extremely difficult. By contrast, Chest Computer Tomography (CT) scanning can yield higher sensitivities of 97% while providing more specific information regarding the lung lesions of the patients [1]. But CT imaging is not a flawless method either. To illustrate this, Figure 1 compares a typical COVID-19 and CAP CT scan. As shown, the similarities between the visual features of COVID-19 and other pneumonia patients makes accurate diagnosis a difficult and time-consuming task for radiologists. Usually, it will take 10-15 minutes for a radiologist to analyze the CT scans of a single patient. Considering the huge number of suspected cases, radiologists often suffer from visual fatigue in reading these scans, which may lead to erroneous diagnosis. Therefore, it is extremely meaningful to invent a reliable COVID-19 diagnosis algorithm based on CT images to aid the radiologists.

Existing studies [2]-[6] have demonstrated the promising performance of applying deep learning methods for COVID-19 diagnosis. These recent works typically fall into two categories.

The first category of methods relies on a large and balanced dataset to construct their learning-based models. For example, Ouyang et al. [2] uses a large dataset consisting of 4982 CT scans from 3645 patients to train a generalizable model. Li et al. [3] introduces the COVNet model, which performs training on a large dataset of 1165 COVID-19, 1560 CAP, and 1193 Non-Pneumonia CT volumes. However, during an unexpected outbreak, hospitals are commonly overwhelmed with patients, and radiologists do not have time to create labeled CT images and process the data for deep learning scientists. In addition, limited resources in certain areas and hospitals can result in the difficulty with providing sufficient CT samples for various illnesses, creating a largely imbalanced dataset that will negatively affect training and model generalizability. Therefore, a screening model that can achieve a high accuracy using a small and unbalanced dataset should also be proposed.

The second category of methods uses the segmented infection regions in CT images for COVID-19 diagnosis prediction. For instance, Wang et al. [4] proposed a 3D CNN (DeCoVNet) to detect COVID-19, which took a CT volume with its lung segmentation as input and output the probabilities of COVID-positive and COVID-negative. Jin et al. [5] proposed a segmentation-classification method, in which a large dataset of CT scans and their corresponding lung lesion masks were used to perform segmentation that helped guide the classification. However, it is time-consuming for radiologists to precisely annotate the lung or infection regions manually. This inspires us to develop a model that works under weakly-supervised setting without any lung or infection segmentations.

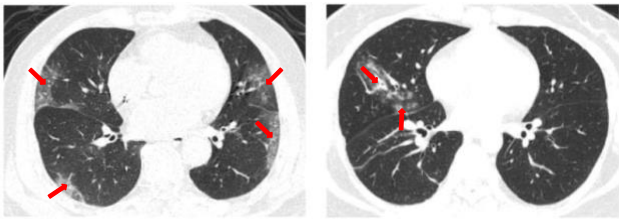


Fig. 1. Comparison of COVID-19 and CAP CT scans.

Among previous works, few studies focused on the classification between COVID-19 and Community-Acquired Pneumonia (CAP), a common lung illness that exhibits similar CT manifestations to COVID-19. Consequently, misdiagnosis remains a severe problem that presents significant risks to patients while undermining pandemic control efforts. Therefore, it is necessary to develop a model that is not only able to judge between healthy and COVID-19 patients but is also capable of distinguishing COVID-19 from CAP.

In this study, we aim to address these limitations and propose a deep learning model that is able to classify between healthy, COVID-19 and CAP patients after being trained on a imbalanced dataset without lung lesion masks. Specifically, our contributions come in three-fold:

(1) We develop a model that can achieve a high accuracy after being trained on a small and unbalanced dataset. We also provide visualizations of the meaningful features extracted by the network to demonstrate the effectiveness of our method.

(2) We introduce a novel multi-task pre-training method, which guides the model to learn the distinguishing features between similar or fine-grained classes, specifically COVID-19 and CAP.

(3) We propose a multi-view-oriented classifier, which allows the model to consider information from all axial, coronal, and sagittal planes. This help information not exhibited in one perspective can be extracted from another plane, allowing the model to capture more features to guide its classification.

II. Methodology

Figure 2 is the framework of our proposed Multi-DeepNet, which consists of four steps: (1) image preprocessing, (2) multi-task pre-training, (3) feature extraction and (4) multi-view-oriented classifier. Specifically, the proposed Multi-DeepNet takes a 3D CT volume as input and outputs a classification prediction and a Grad-CAM map for visualization. In the first step, the chest CT scan is preprocessed. Afterwards, a novel multi-task pre-training module is introduced to pre-train the feature extraction layers of the shared deep learning model. The final model integrates the pre-trained feature extraction layers with our proposed multi-view-oriented classifier. Further training of this final network yields the Multi-DeepNet.

A. Image Preprocessing

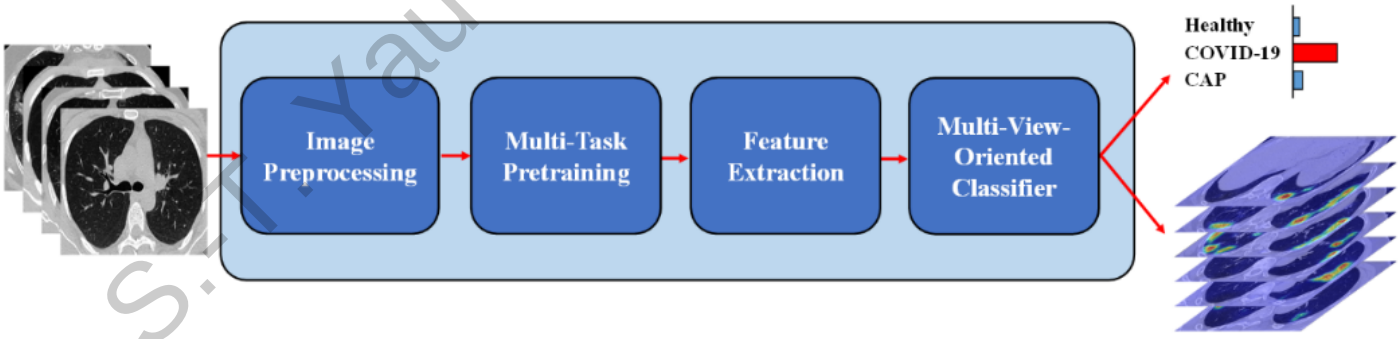


Fig. 2. The framework of our Multi-DeepNet.

The pixel value of CT images reflects the absorption rate of different human tissues to X-Rays, which is measured by Hounsfield Unit (HU). However, different equipment for CT scanning varies in slice thickness or pixel value. Therefore, the use of raw images to train a classification network will result in inconsistencies and affect the robustness of the model. To address this issue, we introduce an effective preprocessing approach. The details of the algorithm are as follows:

(a) Using linear interpolation, all CT scans are resampled to (1,1,1) mm spacing to maintain consistency in CT resolution and slice thickness throughout the dataset.

(b) Circular masks around the lungs are adopted to remove some background information that may impede upon the lung segmentation process afterwards.

(c) The lung segmentation algorithm in [13] is applied to extract a basic lung mask. However, blood vessels and lesions near the outer walls of the lungs may be omitted from this rudimentary mask.

(d) Convex hull and dilation method are used to help refine the lung mask.

(e) Using the lung mask, a bounding box is generated to crop out the lung region and remove unnecessary background information.

(f) The CT volume dimensions are rescaled to 128×256×256.

(g) The image pixels are linearly transformed to the interval [0,1] to normalize intensity. The Window Width and Window Level of all CT scans are uniformly set to 600 and -600, respectively.

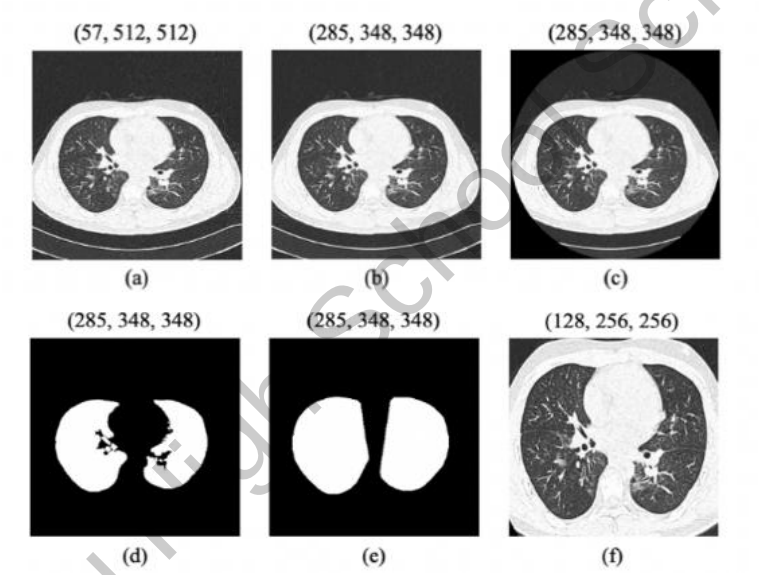


Fig. 3. An example of a healthy patient in the preprocessing procedure. The tuple above each figure indicates the shape (z, x, y) of the 3D volume at each step.

Figure 3 is an example of a healthy patient in the preprocessing procedure.

B. Multi-Task Pre-Training Module

Classification between COVID-19 and CAP patients is especially difficult for two reasons. First, inherent similarities between COVID-19 and CAP CT manifestations makes the classification a challenging task. Commonly, both exhibit signs of ground-glass opacity (GGO), consolidation, and crazy-paving pattern. Thus, the extraction of basic lesion information may not be enough for accurate diagnosis. Second, limited availability and imbalance of data exacerbates the problem. The model would struggle to learn distinctive features on the CAP class, since it is not provided with enough samples.

To address these issues, we propose the multi-task pre-training module, as detailed by Figure 4. Instead of performing three-class classification directly, we partition the original problem into two-class tasks first. Specifically, a shared 3D ResNet18

backbone for feature extraction is trained on three tasks simultaneously: classification of healthy, COVID; classification of healthy, CAP; and classification of COVID, CAP. We formulate the loss function in cross-entropy form for each task:

$$Loss_{task\,i} = -[y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \qquad (1)$$

where y_i is a label valued as 0 or 1; p_i is the SoftMax probability prediction. Each loss guides the model to learn the most discriminating features between the two classes in the respective problem. Finally, to account for all 3 tasks, the final loss is computed as:

$$Loss_{final} = \lambda_1 Loss_{task\,1} + \lambda_2 Loss_{task\,2} + \lambda_3 Loss_{task\,3} \qquad (2)$$

Empirically, we set λ_1=λ_2=λ_3=1. During the training phase, this loss is backpropagated throughout the network to update model parameters. Sub-tasks of the three-class problem guides the model to learn distinguishing features for every category. This prevents the underrepresented CAP class from being neglected during training. Therefore, the multi-task pre-training module allows us to pre-train a network that achieves considerable performance when distinguishing between any two classes. Further training of the model on three-class classification yields exceptional results.

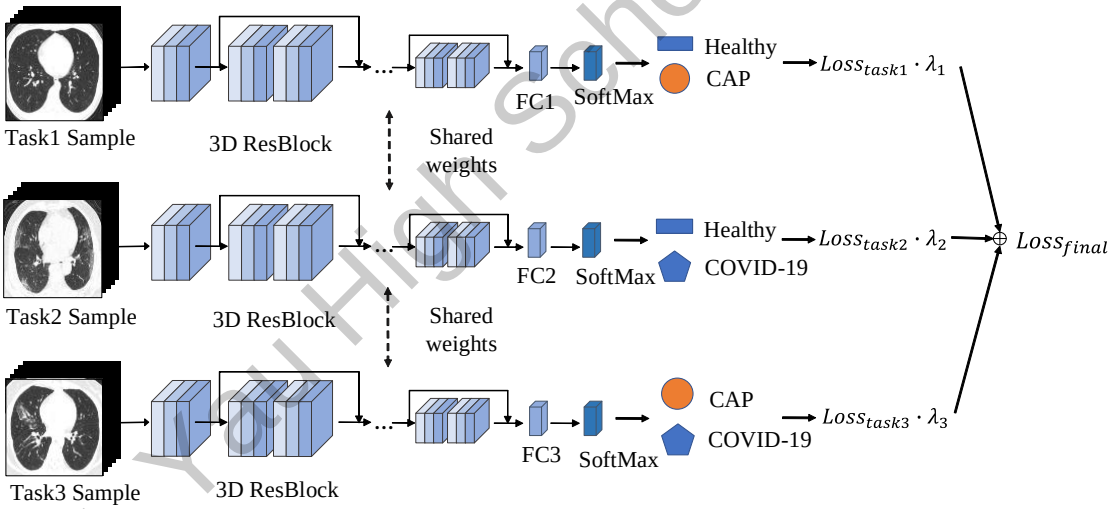


Fig. 4. The details of our multi-task pre-training module.

C. Multi-View-Oriented Classifier

When diagnosing, it is important to gain comprehensive information regarding the CT manifestations of each patient. A single plane may not provide complete information regarding a specific lesion, and in earlier phases of COVID-19, a manifestation may not be observable from some perspectives.

To address this issue, professional radiologists observe the axial, coronal, and sagittal planes of a CT volume simultaneously. This provides them with more information regarding the characteristics of the lesion. In addition, analyzing from three planes at once can help doctors gain a more comprehensive understanding of a manifestation's spatial position within the lung. This is crucial

to accurate diagnosis, because even though COVID-19 and CAP exhibit similar types of manifestations, the spatial distribution of these lesions is a discerning factor. Specifically, it is reported that COVID-19 demonstrates a higher prevalence of lesion distribution in the peripheral zone and involvement of upper and middle lung lobes compared to other pneumonia [1]. Therefore, being able to extract and learn this spatial information can be essential to successfully diagnosing COVID-19.

To take advantage of these findings, we propose the multi-view-oriented classifier, which imitates the radiologists' process in analyzing CT images, as illustrated in Figure 5. We achieve this by first transposing the feature map output of the last convolutional layer:

$$\begin{aligned} A_{\text{axial}} &= \mathrm{M}(\mathrm{d} \times \mathrm{l} \times \mathrm{w}) \\ A_{\text{coronal}} &= \mathrm{M}(\mathrm{w} \times \mathrm{l} \times \mathrm{d}) \\ A_{\text{sagittal}} &= \mathrm{M}(\mathrm{l} \times \mathrm{d} \times \mathrm{w}), \end{aligned} \tag{3}$$

where A represents the transposed planes, M(x, y, z) represents the dimensions of the original input array, w represents the original width, l represents the original length, and d represents the original depth.

For each transposed plane, a fully connected layer and SoftMax function are applied to generate the class probability. The final prediction is given by:

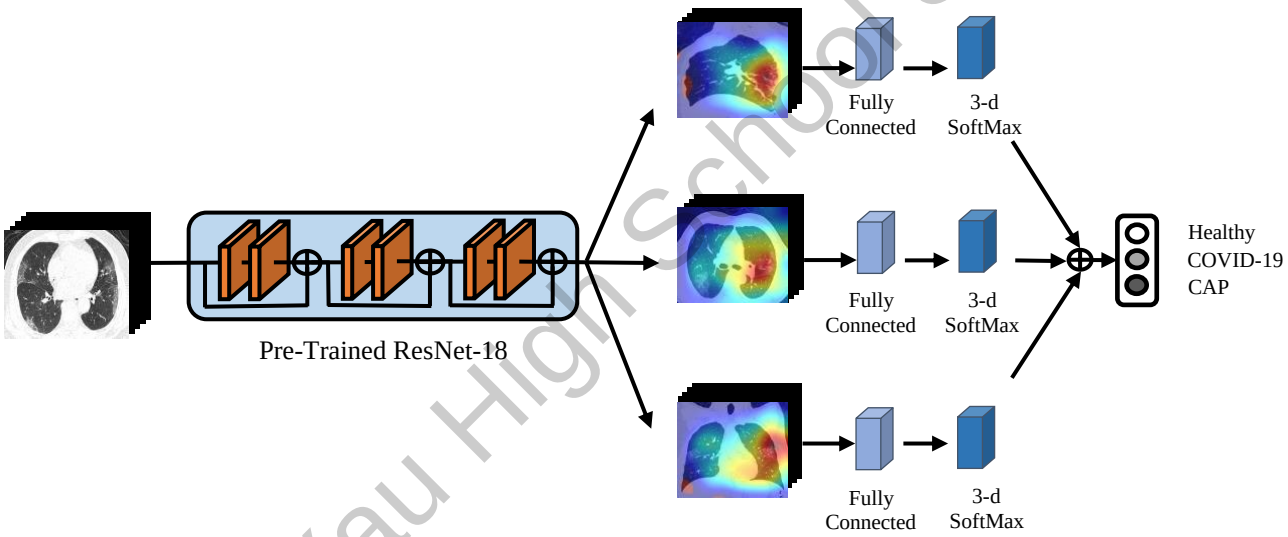


Fig. 5. The details of our proposed Multi-View-Oriented Classifier.

$$P_{\text{final}} = P_{\text{axial}} + P_{\text{coronal}} + P_{\text{sagittal}} \tag{4}$$

By evaluating and predicting based on all 3 planes, our multi-view-oriented classifier allows the model to extract important spatial information regarding the CT manifestations of each patient. We would show in our experiment that this improves the model performance significantly.

III. Experimental Setups

A. Dataset

All data in this research is provided by the Shanghai Public Health Clinical Center (SPHCC) from February 9th to November 24th, 2020. The images are collected with institutional board approval, including a waiver of informed consent from all patients. Since the SPHCC is mainly focused upon the COVID-19 virus, data of other diseases such as CAP was provided in smaller numbers. This resulted in an imbalanced dataset. Specifically, this dataset consists of CT scans from a total number of 247 healthy people, 118 CAP patients, and 223 COVID-19 patients. Table 1 is the distribution of the dataset.

Table 1: Training/Testing Set Distribution Over Various Illnesses By Number of CT Volumes

	Healthy	**CAP**	**COVID-19**	**Total**
Training Set	180	90	160	430
Testing Set	67	28	63	158

B. Implementation Details

Using three-fold cross validation, we trained the model for 300 epochs, and we set the batch size to 6. To address the imbalance of data across classes, we adopt an oversampling strategy on our training dataset. We optimize the network using Stochastic Gradient Descent (SGD) with a momentum of 0.9 and a weight decay of 10-5. We also used a step-decay learning rate, setting the initial value at 0.001 and dividing it by 10 at 30% and 80% of the total number of epochs.

Our proposed model was implemented in PyTorch and trained with three NVIDIA GeForce GTX1080Ti GPUs.

C. Evaluation Metrics

To assess model performance, we use accuracy, sensitivity and specificity to measure the model's ability to correctly identify positive and negative cases. Area under curve (AUC) scores demonstrate the model's ability to distinguish each class from the remaining categories.

IV. Results

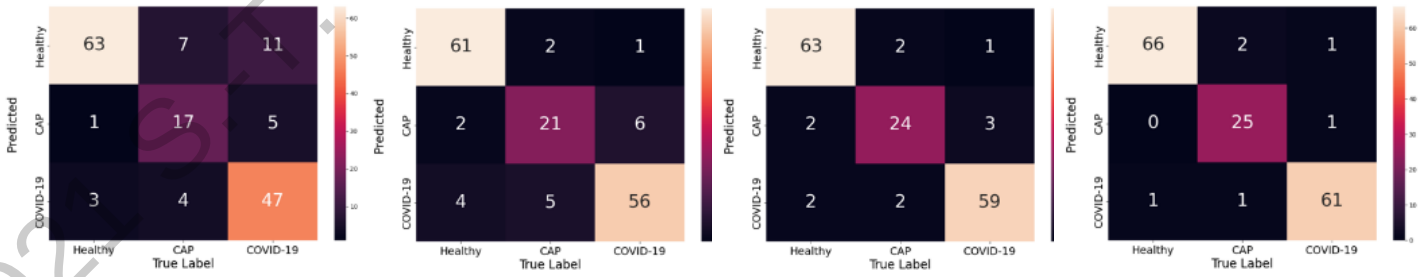
The experimental results are shown in Table 2 and Figure 6. Firstly, we evaluate the performance of the baseline ResNet18 model. For the healthy class, the network achieved a high sensitivity of 94.03% but a low specificity of 77.78%. This reveals that the baseline model struggled to differentiate between ill and healthy patients, frequently misclassifying sick cases as healthy. An analysis of the confusion matrix shows a similar conclusion. For the CAP class, 7 of 28 patients are misclassified as healthy. Furthermore, 11 of 63 COVID-19 cases are also incorrectly categorized as normal. Such inaccuracies reflect that the model is unable to extract relevant information regarding the CT manifestations of each patient as signs of illness.

TABLE 2: ABLATION STUDY ANALYZING THE EFFECTS OF THE MULTI-TASK PRE-TRAINING MODULE AND MULTI-VIEW-ORIENTED CLASSIFIER ON MODEL PERFORMANCE

Modifications to Baseline 3D ResNet18	Healthy		CAP		COVID-19		Average Sensitivity	Average Specificity	Accuracy	AUC
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity				
Baseline	94.03	77.78	60.71	73.91	74.60	87.04	76.45	79.58	80.39	85.28
Baseline + Multi-View	91.05	95.31	75.00	72.41	88.89	86.15	84.98	84.62	87.42	90.51
Baseline + Multi-Task	94.03	95.46	85.71	82.76	93.65	93.65	91.13	90.62	92.68	94.30
Our Method	98.51	95.65	89.29	96.15	96.83	96.83	94.88	96.21	96.20	97.15

Application of the multi-view-oriented classifier results in significant improvements. Although sensitivity for the healthy class decreased from 94.03% to 91.05%, the specificity increased from 77.78% to 95.31%. Sensitivity for the CAP class also increased from 60.71% to 75.00%, and sensitivity for the COVID-19 class increased from 74.60% to 88.89%. Finally, a higher AUC score of 90.51 shows improved performance in identifying and differentiating between classes. Similar trends are identified in the confusion matrix. Only 2 cases from CAP and 1 case from COVID-19 were incorrectly labeled as healthy, compared to 7 misclassified CAP and 11 misclassified COVID-19 cases for the baseline model. These conclusions support the hypothesis that the multi-view-oriented classifier helps the model extract more spatial and feature information regarding the CT manifestations of each sample, improving its ability to recognize infected cases and distinguish it from healthy patients.

The multi-task pre-training module also improves the model performance drastically. Most notably, sensitivity and specificity for the CAP class increases to 85.71% and 82.76%, and for the COVID-19 class, sensitivity and specificity both increase to 93.65%. In addition, the AUC score increases to 94.30%. These changes reflect a clear improvement in the model's ability to distinguish between the three classes, especially for the two illnesses. Such conclusions are also demonstrated in the confusion matrix. While the multi-view-oriented classifier model incorrectly categorizes 5 CAP cases as COVID-19, the multi-task model only makes 2 errors. For the COVID-19 cases, the multi-view-oriented classifier model misclassifies 6 samples as CAP, but the multi-task model



a) Baseline ResNet18 b) Multi-View-Oriented Classifier c) Multi-Task Pre-Training d) Multi-DeepNet

Fig. 6. Confusion matrices are generated to analyze the specific performance of different models on each of the three classes (from left to right: baseline ResNet18, ResNet18 with Multi-View-Oriented Classifier, ResNet18 with Multi-Task Pre-Training, and our proposed Multi-DeepNet.

only misclassifies 3. These results show that our method aligns with the original purpose of the multi-task pre-training module, which is to guide the model to learn distinguishing features between every two categories, especially the similar or fine-grained classes of COVID-19 and CAP.

Finally, our Multi-DeepNet, which combines the multi-task pre-training module and the multi-view-oriented classifier, yields the highest accuracy of 96.20%. For all three classes, sensitivity and specificity for our model reached the highest performance. Out of 63 COVID-19 cases, only 2 were misclassified, and for the CAP class, only 1 patient was incorrectly categorized as COVID-19. Therefore, the methods proposed in this study greatly improves the baseline model's ability to distinguish between healthy, CAP, and COVID-19 patients.

To offer interpretability to our model and confirm that the network is extracting reasonable CT manifestation information to perform classification, we utilize Grad-CAM maps to visualize our network's predictions [14]. We expect to use the gradients of the predicted class flowing into the last convolutional layer to highlight areas of interest that guided the model to make its final classification. Figure 7 shows typical examples of Grad-CAM maps among Healthy, CAP and COVID-19 patients. As shown, our model was able to extract the most relevant information to perform accurate diagnosis. In the healthy case, the model does not focus on any incorrect lesion information, allowing it to correctly classify the sample. In the CAP case, the model successfully concentrates on the GGO information in the right lung (boxed in red), which provides necessary information for the network to diagnose accurately. In the COVID-19 case, the model focuses on the GGO information in the left lung (boxed in red), which also allows it to correctly classify the sample.

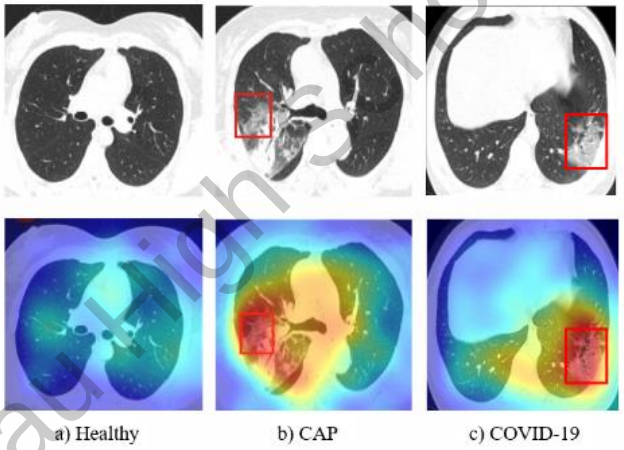


Fig. 7. Examples of a) Healthy, b) CAP, and c) COVID-19 Grad-CAM maps.

V. Discussion and Conclusion

In this paper, we develop an efficient and reliable method for COVID-19 diagnosis under imbalanced dataset. During pandemic scenarios, it is necessary to develop immediate solutions to minimize the effects of viral spread. Despite initial limitations in terms of data availability, which is a practical problem in this situation, we develop an automated CT analysis algorithm based on deep learning networks. Specifically, to guide the training on a small and imbalanced dataset, we propose a multi-task pre-training module. To extract relevant special information that aids model classification, we created the multi-view-oriented classifier. With a 96.20% accuracy and clear Grad-CAM attention results, our Multi-DeepNet can generate crucial imaging information to the radiologists in just 6 seconds.

There are still multiple limitations needing to be addressed. Firstly, our model is not trained in an end-to-end manner, which reduces its convenience and efficiency. Currently, data pre-processing occurs separately from model training, and the selection of a best epoch after multi-task pre-training requires manual intervention. Therefore, improving the model to be an end-to-end network is an important future step. Furthermore, we hope to apply our model in more clinical settings. We have begun conducting clinical experiments at the Shanghai Public Health Clinical Center to assess the real-time performance of the Multi-DeepNet. Moreover, we hope to further cooperate with more hospitals to further analyze the generalization ability of our method in diagnosing COVID-19 patients.

VI. REFERENCES

[1] Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., and Xia, L. (2020). Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology*, *296*(2), E32-E40.

[2] Ouyang, X., Huo, J., Xia, L., Shan, F., Liu, J., Mo, Z., and Shen, D. (2020). Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia. *IEEE Transactions on Medical Imaging*, *39*(8), 2595-2605.

[3] Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., and Xia, J. (2020). Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. *Radiology*, *400*(1), E23-E30.

[4] Wang, S., Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., and Xu, B. (2021). A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). *European Radiology*, *45*, 1-9.

[5] Jin, S., Wang, B., Xu, H., Luo, C., Wei, L., Zhao, W., and Xu, W. (2020). AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system in four weeks. *Applied Soft Computing*, *98*, 106897.

[6] Islam, M. M., Karray, F., Alhajj, R., and Zeng, J. (2021). A review on deep learning techniques for the diagnosis of novel coronavirus (covid-19). *IEEE Access*, *9*, 30551-30572.

[7] Altaf F, Islam S M S, Akhtar N, et al. Going deep in medical image analysis: concepts, methods, challenges, and future directions[J]. IEEE Access, 2019, 7: 99540-99572.

[8] Zhou S K, Greenspan H, Davatzikos C, et al. A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises[J]. Proceedings of the IEEE, 2021.

[9] Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., and Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage: Clinical*, *17*, 16-23.

[10] Liu, S., Liu, S., Cai, W., Pujol, S., Kikinis, R., and Feng, D. (2014). Early diagnosis of Alzheimer's disease with deep learning. *In 2014 IEEE 11th International Symposium on Biomedical Imaging*, 1015-1018.

[11] Albarqouni, S., Baur, C., Achilles, F., Belagiannis, V., Demirci, S., and Navab, N. (2016). Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE Transactions on Medical Imaging*, *35*(5), 1313-1321.

[12] Zech, J. R., Badgeley, M. A., Liu, M., Costa, A. B., Titano, J. J., and Oermann, E. K. (2018). Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. *PLoS Medicine*, *15*, e1002683.

[13] Liao, F., Liang, M., Li, Z., Hu, X., and Song, S. (2019). Evaluate the malignancy of pulmonary nodules using the 3-d deep leaky noisy-or network. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(11), 3484-3495.

[14] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. *In Proceedings of the IEEE International Conference on Computer Vision*, 618-626.

VII. Acknowledgements

As a student living in China in February 2020, I was one of the first people to witness the effects of the COVID-19 pandemic. I watched as hospitals and doctors were overwhelmed in Wuhan by the number of suspected cases and as they struggled to diagnose and quarantine patients effectively. I was extremely aware of the devastating effects of the pandemic. Unfortunately, the virus has spread beyond China, and the number of cases skyrocketed. Research reveals that the current testing methods in many countries, which uses the RT-PCR test as a golden standard, are not extremely effective, and it only yields 59% sensitivity. Given the rapid transmission of COVID-19, this is a severe problem. Thus, I knew that COVID-19 diagnosis was a major issue that needed to be further researched.

In many countries such as China, where I conducted my research, CT scanning is used alongside the RT-PCR test, and it has demonstrated higher sensitivity than the RT-PCR. However, it is time-consuming and labor-intensive for radiologists to analyze, as each CT scans consists of approximately 300 slices, and the signs of illness are often subtle. Thus, the goal of my research was to develop a computer-guided classification method that can guide radiologists and reduce the time needed to diagnose patients with CT.

Throughout the course of this study, Professor Feng Rui and Professor Guo Yi from Fudan University kindly guided me through my research process (both were uncompensated). Prior to conducting my own research, Professor Feng presented me with the opportunity to participate in a project involving diagnosis of COVID-19, segmentation of COVID-19 lung lesions, and the final clinical experimentation of completed models.

Initially, I mainly worked as part of a team that focused on COVID-19 classification. I participated in performing data processing of the CT scans we received from the Shanghai Public Health Clinical Center (SPHCC), and I learned more about the CT manifestations and imaging features of COVID-19. I also focused on experimenting with baseline models such as ResNet18, ResNet34, VGG16, DenseNet121, etc. In addition, I read many technical papers with the team, some directly related to COVID-19 and others that were generally about medical imaging and deep learning.

Afterwards, I started researching individually. I focused exclusively on three-class classification between healthy, CAP, and COVID-19. I included CAP because I realized that differentiating between COVID-19 and CAP is difficult, and it is a common challenge for radiologists. Later, I began conducting experiments and creating models individually, with the general guidance of Professor Guo Yi. She offered feedback on my ideas and helped guide me in performing the ablation studies.

Finally, Professor Feng Rui helped review my work and guided me in the process of writing a paper.