MedCheX: An Efficient COVID-19 Detection Model for Clinical Usage

CHI-SHIANG WANG[1*], FANG-YI SU[2*] AND JUNG-HSIEN CHIANG[3]
Department of Computer Science and Information Engineering
National Cheng Kung University
Tainan, 701 Taiwan
E-mail: {hyaline0317[1], fangyi[2]}@iir.csie.ncku.edu.tw, jchiang@mail.ncku.edu.tw

Due to the highly infectious and long incubation period of COVID-19, detecting COVID-19 efficiently and accurately is crucial since the epidemic outbreak. We proposed a new detection model based on U-Net++ and adopted dense blocks as the encoder. The model not only detects and classifies COVID-19 but also segment the lesion area precisely. We also designed a two-phase training strategy along with self-defined groups, especially the retrocardiac lesion to make model robust. We achieved 0.868 precision, 0.920 recall, and 0.893 F1-score on the COVID-19 open dataset. To contribute to this pandemic, we have set up a website with our model. *https://medchex.tech/*

Keywords: *COVID-19, Convolutional Neural Network, Computer Vision, Deep Learning, Medical Image*

1. INTRODUCTION

In December 2019, a cluster of pneumonia-like cases emerged in Wuhan, China. Soon after, it spread rapidly throughout China causing an epidemic, followed by an increasing number of cases in almost every country throughout the world. In February 2020, the World Health Organization (WHO) designated the disease COVID-19, which stands for coronavirus disease in 2019. The pathogenic virus COVID-19 is named SARS-CoV-2 short for severe acute respiratory syndrome coronavirus 2, because of its phylogenetic similarity to the SARS virus, which also caused an epidemic in 2003; previously, it was referred to as 2019-nCoV [1].

SARS-CoV-2 is a member of the coronavirus family that causes severe illness in humans. This virus is very contagious due to the nonspecific symptoms like cough, fever, and mild dyspnea similar to the common cold caused by other respiratory viruses. COVID-19 is an ongoing pandemic. Until August 1, 2020, 17,396,943 cases had been reported in over 200 countries with 675,060 deaths (approximately 3.88 % case fatality rate) [2]. Even though the Reverse Transcription-Polymerase Chain Reaction (RT-PCR) test serves as the gold standard of diagnosing COVID-19, chest X-ray (CXR) is one of the more easily accessible clinical screening tools available with low cost and high efficiency. Furthermore, previous research on detecting COVID-19 on CXR images usually focused on classifying COVID-19, other pneumonia, and normal lung [3]. Results would only show which group they are, but they were not useful clinically. If we want to apply it online to assist

*Equal Contribution

physicians or other frontline staff to screen for COVID-19, the model would be much more explainable. Hence, many studies utilize class activation map (CAM)-based approaches to visualize the significant features of the diseased position [4]. However, we found that the patterns of pneumonia are difficult to detect accurately by solely depending on CAM; likewise, physicians find it hard to discriminate against pneumonia caused by other pathogens and COVID-19. In this paper, we designed a COVID-19 detection model on CXR images and tried to apply it to clinical usage. The model we proposed, MedCheX, not only can classify COVID-19, diseased lung patterns, and normal images but also can segment lesion areas for visualization. Moreover, MedCheX won the COVID-19 Global Hackathon as a highlight project [5].

2. RELATED WORK

2.1 Deep Learning in Medical Images

Convolution Neural Network (CNN) has worked well on computer vision tasks since 2012 [6]. Recently it was proved that CNNs also did great at complicated tasks which human needs professional training for years such as distinguishing diseased images from healthy ones on computed tomography (CT) and CXR or segmenting lesion areas [7]–[9].

At present, the available architectures for medical image segmentation are almost inspired by the well-known fully convolutional neural network (FCN) [10] or U-Net [11]. In FCN, there is no classic fully connected layer, which often is implemented for classification but full of convolutional layers. At the last layer, FCN takes a single step of upsampling to recover the original resolution for dense pixel prediction. In contrast, U-Net joints convolutional layers with pooling layers and upsampling layers to create contractive and expansive paths. Skip connections are implemented to leverage intermediate feature maps and merge contractive and expansive features. Recently, many extensive models have been proposed to solve medical image tasks based on these two network architectures. We built our own model by taking them as references [12], [13].

2.2 Deep Learning for Detecting COVID-19

Applied deep learning tools for detecting COVID-19 is pivotal since the epidemic outbreak. Deep learning researchers had approved that the CNN model could classify lung lesion patterns very well [14] Hence, investing resources in researching exerting CNN detection ability to help fight against the pandemic is momentous.

Ali et al. used ten well-known CNN models to classify COVID-19 and non-COVID-19 on CT images and in the methods for the detection of lung abnormalities, particularly in the early stages of the disease [15]. However, with the increasing number of hospitalizations and intensive care unit (ICU) admissions, frequent CT scans are unrealistic.

Therefore, Borghesi et al. were trying to use CXR imaging for monitoring patients with confirmed infection [16]. Even though CXR is considered not sensitive enough for the detection of early-stage lung diseases. During this COVID-19 pandemic, many studies had shown uplifting results on CXR images with deep learning models [3], [4], [7].

2.3 Attention Mechanism

Attention can be interpreted as a mechanism of biased towards the most informative components of a signal, just like humans attend important information while reacting. Attention mechanism has demonstrated its utility across many tasks including text translation, sequence localization, understanding in images, and image captioning. In these applications, it can be implemented in every part of a network you want to attend just like a plug-in device. Some studies proposed spatial or channel attention in image tasks and achieved good performance [17], [18].

3. METHODOLOGY

To achieve our goal of detecting COVID-19, we proposed a new model and well-designed lesion segmentation label groups. In this section, we will introduce our proposed model in detail. The overall architecture is shown in Figure 1. The pneumonia detection model is based on U-Net++ [13] and utilized dense block to be the encoder backbone. We also added the squeeze-and-excitation block [19] into the model to generate more efficient features in the decoder. Skip connections are constructed to feed previous feature maps into the next convolutional block.

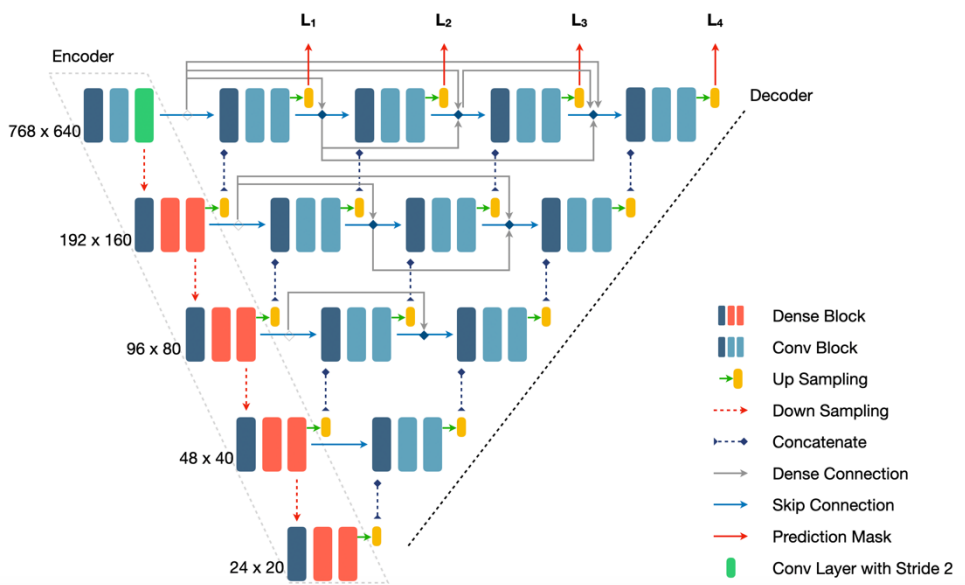

Fig. 1. Overview of MedCheX model architecture

3.1 Overview of Detection Model

We proposed a new model, MedCheX, for detecting COVID-19 by distinguishing patterns from other lung diseases and normal CXR images. As mentioned in section 2, FCN and U-Net are well-known architecture for designing medical image segmentation models. However, they could lose significant information during the downsampling and upsampling processes even though skip connections could compensate for the drawbacks. Consequently, we designed our model on U-Net++ which improved feature extraction and reusability on the skip connections. Furthermore, we adopted dense blocks as encoder backbone. The dense blocks as shown in Figure 2 are adopted from DenseNet [20] and were also used by Jegou et al in Tiramisu [21] which directly connects every layer within the block in a feed-forward fashion. Even then the network will be more accurate. In addition, the convolution blocks as shown in Figure 1 are constructed by two 3x3 convolutional layers. Each layer passed through group normalization [22] and then activated by ReLU functions [23].

To make a model which utilizes feature maps efficiently and accurately. Hu et al. proposed Squeeze-and-Excitation (SE) blocks [19]. It's an attention architectural unit designed to improve the representational power of a network by enabling it to perform dynamic channel-wise feature recalibration. As shown in Figure 2, the squeezing part leveraged the weights to do the weighted sum, then produce new feature maps. SE blocks can be applied to our model directly and are also computationally lightweight and impose only a slight increase in model complexity and computational burden. Therefore, we added SE blocks to MedCheX.

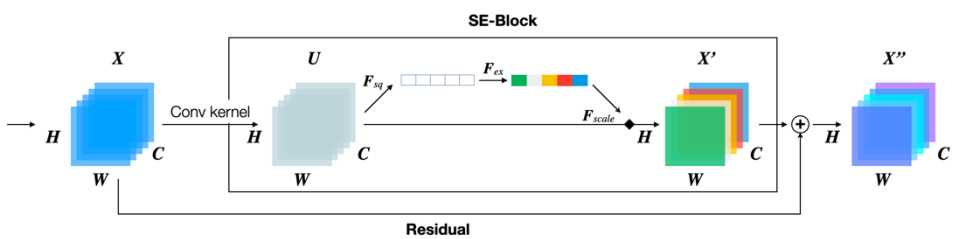

Fig. 2. This figure shows how to add the SE-block into the residual convolutional block. After obtaining the U, we utilized the global average pooling in F_{sq} to get the global information from each feature map. We adopted two fully connected layers with ReLU and sigmoid function to apply the non-linear transformation to attention information in F_{ex}. We then multiplied the attention weight, produced by F_{ex}, to feature map U to obtain the more efficient features X' in F_{scale}.

3.2 Training/Test Protocol

We designed a two-phase training strategy. During phase 1, we used pneumonia CXR images collected from quarantine station at National Cheng Kung University Hospital in Taiwan to train a pneumonia classifier model. We used Adam [24] optimizer with an initial

learning rate 0.0001 along with 0.1 decay every 20 epochs. Our goal was to train a model that could classify pneumonia accurately. Moreover, according to the concept of big success in fine-tuning on transferring a pre-trained model for downstream tasks [25]. In phase 2, We then transferred our pre-trained pneumonia classifier to the COVID-19 dataset for fine-tuning to distinguish COVID-19 from pneumonia with a lower learning rate along with cosine annealing. Further, we implemented data augmentation while training in two phases by computer vision methods.

3.3 Model Predictions

The model, MedCheX, will predict two different kinds of patterns including pneumonia and retrocardiac. Thus, we set two 5x5 convolutional kernels with sigmoid activation function at the last layer in the model for predicting the probability of groups and creating masks for three groups respectively. We trained our model by using the Binary Cross Entropy (BCE) and then minimized the loss function as followed:

$$\mathcal{L}_{model} = -\frac{1}{N \times M} \sum_{i=0}^{N} \sum_{j=0}^{M} y_{i,j} \times log\left(\hat{y}_{i,j}\right) + \left(1 - y_{i,j}\right) \times log\left(1 - \hat{y}_{i,j}\right) \quad (1)$$

where the N and M denote the height and width of the input image and the $y_{i,j}$ presents the target information and $\hat{y}_{ij}$ shows the predicted value of this pixel. Moreover, we adopted deep supervision [26] to adjust the features start from the early step by utilizing the loss L_1 to L_4 shown in Figure 1. All of the loss values followed (1) to minimize the loss function, after that we applied the widely-used optimization method, Adam [24], to optimize the model. Thereafter, we detected pneumonia and normal lung by the strategy as shown in (2) when we obtained the prediction mask x' from the model, where the θ is a threshold to divide the normal and diseased lung, and we set it to 0.5 in this study.

$$classified(x'; \theta) = \begin{cases} 1, if\ max(x) \geq \theta \\ 0, otherwise \end{cases} \quad (2)$$

As mentioned previously, for applying our model to detect the COVID-19, we took advantage of transfer learning to extend our model to predict one more mask for COVID-19. If the model predicted pneumonia and COVID-19 masks with a confidence score greater than 0.5 simultaneously, we would consider this result as the COVID-19 group. But for describing clearly, we show all masks on our website platform.

4. EXPERIMENTS

In this section, we will describe how we construct our dataset and experiment results.

4.1 Data Collection

To protect the patients' privacy, we anonymize all the private information of patients during dataset construction. For our task, we collected overall 2497 chest X-ray images at National Cheng Kung University Hospital in Taiwan. All of our data are interpreted by professional radiologists who helped us to identify the lesion area and segment it on images

as shown in Figure 3. We divided the segmented labels into 3 groups: (a) Diseased lung group: We clustered ground-glass opacity (GGO), and consolidation patterns that were not infected by COVID-19. (b) Retrocardiac group: Physicians labeled specific patterns which behind the cardiac area. (c) COVID-19 group: This group data was an open-source at GitHub*, which collected 237 COVID-19 infection confirmed images.

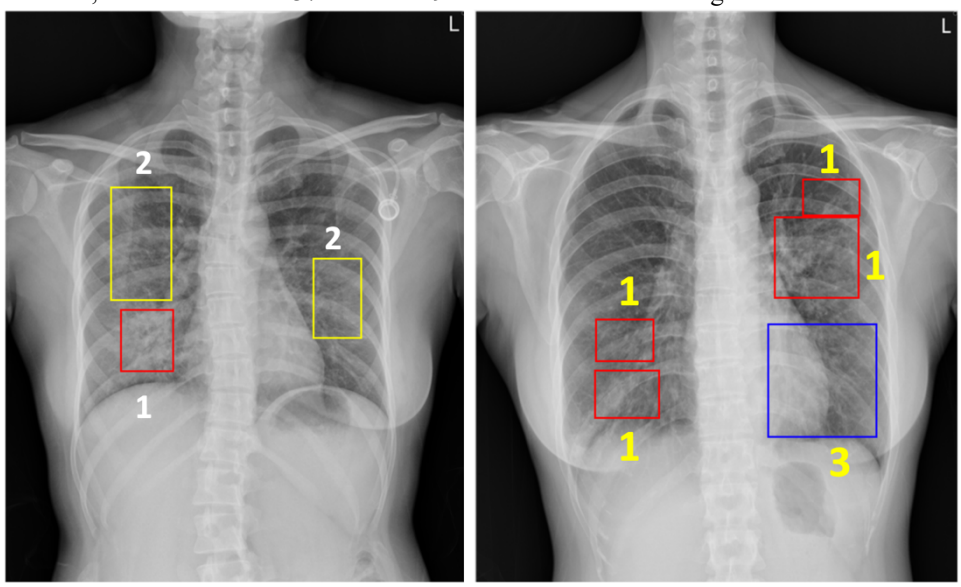

Fig. 3. This figure shows different label groups: (1) Consolidation (2) Ground-glass opacity. (3) Retrocardiac.

4.2 Metrics

Sensitivity, specificity, and accuracy were employed to evaluate the performance of the pneumonia classification model in phase 1. Besides, due to a lack of negative samples, we adopted precision, recall, and F1-score as metrics for assessing the ability to distinguish COVID-19 from pneumonia in phase 2. The formula of metrics is listed below:
(TP: True Positive; TN: True Negative; FP: False Positive; FN: False Negative)

$$Accuracy = \frac{No.\ of\ images\ correctly\ classified}{Total\ no.of\ images} \tag{3}$$

$$Sensitivity = Recall = \frac{TP}{TP+FN} \tag{4}$$

$$Specificity = \frac{TN}{TN+FP} \tag{5}$$

$$Precision = \frac{TP}{TP+FP} \tag{6}$$

$$F1-score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{7}$$

* https://github.com/ieee8023/covid-chestxray-dataset

4.3 Results

At phase 1 training, we got 94.1% sensitivity, 95.1% specificity, and 94.6% accuracy on training set; 83.6% sensitivity, 82.2% specificity, and 82.5% accuracy on the test set for classifying pneumonia and non-pneumonia lung as shown in Table 1a and 1b. We find that our model can classify pneumonia and normal lung accurately. During phase 2, we transferred our model to finetune on the COVID-19 dataset. We achieved 0.868 precision, 0.920 recall, and 0.893 F1-score as shown in Table 2. Due to a lack of negative samples, we did not calculate specificity and choose other metrics for evaluation.

Table 1a. This table shows the phase1 training results.

Phase 1 Training Data Acc=0.946					
	Present	**Num.**	**Absent**	**Num.**	**Total**
Positive	True Positive	685	False Positive	34	719
Negative	False Negative	43	True Negative	658	701
Total		728		692	1,420
Sensitivity	0.941		**Specificity**	0.951	

Table 1b. This table shows the phase 1 test results.

Phase 1 Test Data Acc=0.825					
	Present	**Num.**	**Absent**	**Num.**	**Total**
Positive	True Positive	46	False Positive	182	228
Negative	False Negative	9	True Negative	840	849
Total		55		1,022	1,077
Sensitivity	0.836		**Specificity**	0.822	

Table 2. This table shows the phase 2 evaluation results after fine-tuning.

Phase 2 Evaluation Results F1-score=0.893					
	Present	**Num.**	**Absent**	**Num.**	**Total**
Positive	True Positive	46	False Positive	7	53
Negative	False Negative	4	True Negative	0	4
Total		50		7	57
Precision	0.868		**Recall**	0.920	

In figure 4, we show the predictions by phase 1 pneumonia classifier. From left to right are (a) original image, (b) pneumonia lesion including GGO and consolidation, and (c) retrocardiac lesion. We use different colors to display predictive probability.

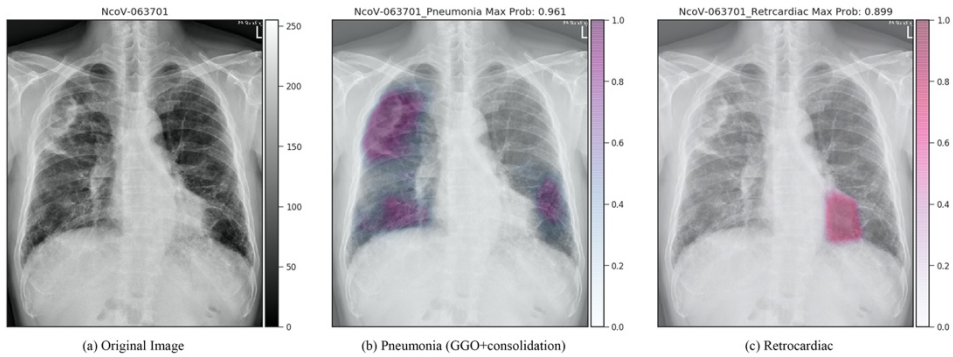

Fig. 4. This figure shows pneumonia classifier predictions in phase 1.

In Figure 5, we show two set of predictions by phase 2 COVID-19 detector. From left to right are (a) original image, (b) pneumonia lesion including GGO and consolidation, (c) retrocardiac lesion, and (d) COVID-19. We use different colors to display predictive probability the same as Figure 4.

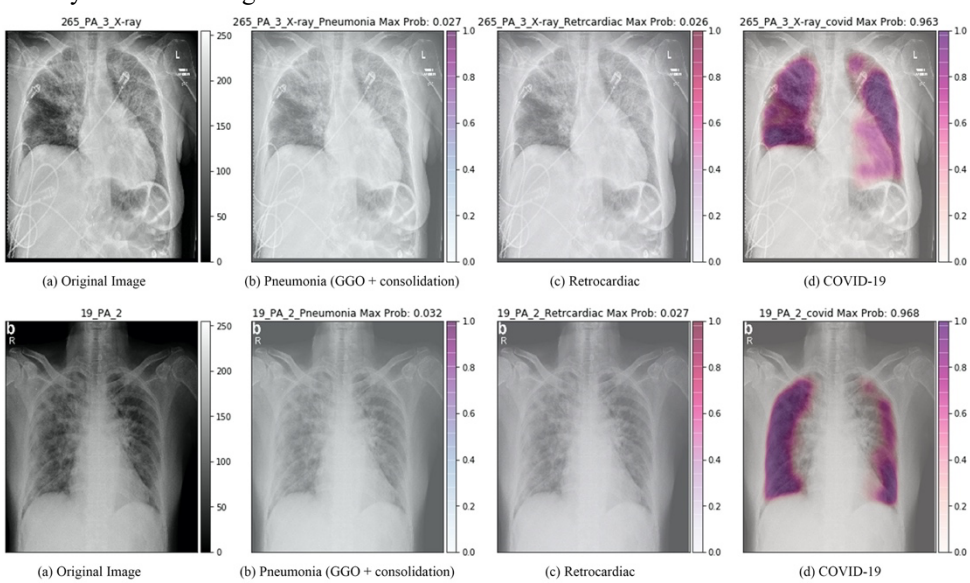

Fig. 5. This figure show COVID-19 detector predictions in phase 2.

As shown in Figure 4 and 5, our model can predict pneumonia and COVID-19 respectively and accurately, even though COVID-19 patterns were to some degree the same similar to pneumonia caused by other pathogens.

4.4 Comparison

For proving the robustness of our model, we also conducted some experiments to compare our proposed model to other baseline models on the same phase 2 evaluation COVID-19 images. We trained another two models, which utilize respectively original U-

Net++ and U-Net++ with SE block as the backbone. The results are shown below in Table 3:

Table 3. This table shows the comparison results with other models.

Models	TP	FP	FN	Precision	Recall	F1-score
Original U-Net++	39	11	7	0.78	0.848	0.813
U-Net++ / **SE block**	44	8	5	0.846	0.898	0.871
Ours (U-Net++/ **SE /Dense block**)	**46**	**7**	**4**	**0.868**	**0.920**	**0.893**

As shown in Table 3, our model achieved the best results in all metrics evaluations. Although there was little performance gain from the Dense block, we still glad to see the increase in recall as we focus that during screenings.

5. DISCUSSION

There were still some limitations in our model. First, the small sample size warrants further investigation. More enrollment will likely improve the robustness of our model. Second, artifacts such as blurred lesions or ground-glass opacities would influence model performance. Moreover, from the professional aspect of physicians, although not much differences between pneumonia and COVID-19, we could moderate how deep learning model works by applying CAM heatmaps. There is still limited evidence for clinical usage.

6. CONCLUSIONS

Early-stage diagnosis and treatment of COVID-19 are essential during this pandemic. In this paper, we proposed an efficient COVID-19 detection model based on U-Net++, which distinguishes COVID-19 patterns from pneumonia caused by other pathogens accurately. Due to the limitations mentioned in section 5, we will collect more data for further training. Also, we have plans on modifying our model to a two-stage classifier and segmentation mask generator to make it more robust.

REFERENCES

[1] "WHO Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020." https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020.

[2] "Coronavirus disease (COVID-19) Situation Report – 194." https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200801-covid-19-sitrep-194.pdf?sfvrsn=401287f3_2.

[3] T. Mahmud, M. A. Rahman, and S. A. Fattah, "CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization," *Comput. Biol. Med.*, vol. 122, no. June, p. 103869, 2020, doi: 10.1016/j.compbiomed.2020.103869.

[4] M. Farooq and A. Hafeez, "COVID-ResNet: A Deep Learning Framework for

Screening of COVID19 from Radiographs," 2020, [Online]. Available: http://arxiv.org/abs/2003.14395.
[5] "COVID-19 Global Hackathon." https://covid-global-hackathon.devpost.com/.
[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
[7] S. Chatterjee *et al.*, "Exploration of Interpretability Techniques for Deep COVID-19 Classification using Chest X-ray Images," no. June, 2020, [Online]. Available: http://arxiv.org/abs/2006.02570.
[8] J. Chen *et al.*, "Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study," *medRxiv*, pp. 1–27, 2020, doi: 10.1101/2020.02.25.20021568.
[9] S. Ying *et al.*, "Deep learning Enables Accurate Diagnosis of Novel Coronavirus (COVID-19) with CT images," *medRxiv*, pp. 1–10, 2020, doi: 10.1101/2020.02.23.20026930.
[10] E. Shelhamer, J. Long, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, 2017, doi: 10.1109/TPAMI.2016.2572683.
[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9351, pp. 234–241, 2015, doi: 10.1007/978-3-319-24574-4_28.
[12] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Comput. Biol. Med.*, p. 103792, 2020.
[13] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11045 LNCS, pp. 3–11, 2018, doi: 10.1007/978-3-030-00889-5_1.
[14] P. Rajpurkar *et al.*, "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," pp. 3–9, 2017, [Online]. Available: http://arxiv.org/abs/1711.05225.
[15] A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, "Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks," *Comput. Biol. Med.*, vol. 121, no. January, p. 103795, Jun. 2020, doi: 10.1016/j.compbiomed.2020.103795.
[16] A. Borghesi and R. Maroldi, "COVID-19 outbreak in Italy: experimental chest X-ray scoring system for quantifying and monitoring disease progression," *Radiol. Medica*, vol. 125, no. 5, pp. 509–513, 2020, doi: 10.1007/s11547-020-01200-3.
[17] F. Wang *et al.*, "Residual attention network for image classification," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, no. 1, pp. 6450–6458, 2017, doi: 10.1109/CVPR.2017.683.
[18] P. Anderson *et al.*, "Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering," *Proc. IEEE Comput. Soc. Conf. Comput. Vis.*

Pattern Recognit., pp. 6077–6086, 2018, doi: 10.1109/CVPR.2018.00636.

[19] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-Excitation Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 8, pp. 2011–2023, 2020, doi: 10.1109/TPAMI.2019.2913372.

[20] F. Iandola, M. Moskewicz, S. Karayev, R. Girshick, T. Darrell, and K. Keutzer, "DenseNet: Implementing Efficient ConvNet Descriptor Pyramids," pp. 1–11, 2014, [Online]. Available: http://arxiv.org/abs/1404.1869.

[21] S. Jegou, M. Drozdzal, D. Vazquez, A. Romero, and Y. Bengio, "The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2017-July, pp. 1175–1183, 2017, doi: 10.1109/CVPRW.2017.156.

[22] Y. Wu and K. He, "Group Normalization," *Int. J. Comput. Vis.*, vol. 128, no. 3, pp. 742–755, 2020, doi: 10.1007/s11263-019-01198-w.

[23] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," 2010.

[24] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv Prepr. arXiv1412.6980*, 2014.

[25] A. Kolesnikov *et al.*, "Big Transfer (BiT): General Visual Representation Learning," 2019, [Online]. Available: http://arxiv.org/abs/1912.11370.

[26] C.-Y. Lee, S. Xie, P. Gallagher, Z. Zhang, and Z. Tu, "Deeply-supervised nets," in *Artificial intelligence and statistics*, 2015, pp. 562–570.