2020 IEEE REGION 10 CONFERENCE (TENCON)
Osaka, Japan, November 16-19, 2020

Transfer Learning Based Method for COVID-19 Detection From Chest X-ray Images

Nayeeb Rashid[1], Md Adnan Faisal Hossain[1], Mohammad Ali[1],
Mumtahina Islam Sukanya[1], Tanvir Mahmud[1], and Shaikh Anowarul Fattah[1]
Email: {nayeebrashid, afhossain97, marcoreus.kazmi, islam.sukanya}@gmail.com, {tanvirmahmud, fattah}@eee.buet.ac.bd
Department of Electrical and Electronic Engineering,
Bangladesh University of Engineering and Technology, Dhaka-1205, Bangladesh

***Abstract*—Radiology examination of chest radiography or chest X-ray (CXR), is currently performed manually by radiologists. With the onset of the COVID-19 pandemic, there is now a need to automate this process which is currently one of the key methods of primary detection of the SARS-Cov-2 virus. This will lead to shorter diagnosis time and less human error. In this study, we try to perform three-class image classification on a dataset of chest X-rays of confirmed COVID-19 patients(408 images), confirmed pneumonia patients(4273 images), and chest X-rays of healthy people(1590 images). In total the dataset consists of 6271 people. We aim to use a Convolutional Neural Network(CNN) and transfer learning to perform this image classification task. Our model is based on a pre-trained InceptionV3 network with weights trained on the ImageNet dataset. We fine-tune the layers of the Inception network to train it to our specific task. We try fine-tuning the network to different extents by freezing a different number of layers and then comparing accuracy for each variation of the network. To evaluate the performance of our network we use several metrics which include Classification accuracy, Precision, Sensitivity, and Specificity. Our proposed method achieves an accuracy of 96.33% on a 3-class classification task (Normal, COVID-19, Pneumonia) and an accuracy of 99.39% on a 2-class (COVID and Non-COVID) classification task.**

***Index Terms*—COVID-19, Medical Image Analysis, Chest X-ray, Deep Learning**

I. Introduction

The novel Coronavirus disease 2019 also known as COVID-19 first appeared in Wuhan, Hubei, China in December 2019 [1] and from then on it turned into a global pandemic affecting millions of lives worldwide. At the time of writing this paper more than 10 million people have been infected with this disease and more than half a million people have died from it. As of now, no vaccine or antiviral treatment is available for COVID-19 [2]. As a result, early detection and isolation is currently the best way of mitigating the spread of the disease and saving lives.

An infected COVID-19 patient may start showing symptoms such as fever, cough, fatigue, sputum production, headache, hemoptysis, diarrhea, dyspnoea, and lymphopenia [3] after 5.2 days [4]. As early detection of the disease is a crucial step in fighting the pandemic, the current methods of testing the presence of a virus in a human body include real-time reverse transcription poly- merase chain reaction (rRT-PCR) [5] and antibody testing [6]. But these tests are both costly and difficult to carry out at the rate the virus is spreading in the densely populated areas. In developing countries, the lack of testing kit available has already become an obstacle in fighting the pandemic. So a testing method with ease of availability and cost-effectiveness has become a necessity in these countries.

COVID-19 is a novel severe acute respiratory syndrome coronavirus which mostly affects the lungs in the human body [7]. Researchers have found ground-glass opacities, consolidation, and lower zone predominance [8] in chest scans of COVID-19 patients. Because of these features in the lung scans, it has been shown that chest scans can be used to detect the virus [9] in patients. The two available methods of chest scans are i)CT scan and ii)X-ray. [10], [11] has shown deep learning-based methods for detecting COVID-19 from chest CT scans.

But CT scans are not as available or cost-effective as X-ray. Almost every hospital has X-ray facilities and it is very easy to get an X-ray done. So being able to detect COVID-19 from X-ray images will have much more impact and will be able to reach a lot more people even in the less developed areas.

Deep Learning-based approaches are being used in the field of medical image analysis and automatic diagnosis for some time now due to the great improvements in their performance and their potential for becoming an adjacent tool for clinicians [12]–[14]. We have seen deep learning being applied for Alzheimer's disease detection from neuroimaging [15], retinal image analysis [16], breast cancer detection [17], skin cancer classification [18] and many other sectors. Deep learning-based methods have seen significant application in chest X-ray image related tasks such as: Nodule classification [19], Tuberculosis detection [20], rib suppression [21], Pneumonia detection [22] and Lung segmentation [23].

Given the success of deep learning-based methods in chest X-ray image-related tasks, it is only natural to use it for classifying COVID-19 from chest X-ray images. A lot of research is being done in this field. [24] suggested COVID-Net, a deep convolutional neural network for classifying COVID-19 in chest X-ray images. It was trained on a dataset containing 3 classes (normal, pneumonia, and COVID) and achieved a 93.3% accuracy across the classes. DarkCovidNet another CNN model for this task developed by [25] was trained on both 3 classes and 2 classes (COVID and Non-COVID) and attained an accuracy of 87.02% and 98.08% respectively. Another CNN model based on the Xception [26] architecture named CoroNet [27] was trained on 4 classes (normal, COVID, bacterial pneumonia and viral pneumonia), 3 classes and 2 classes and its accuracy for each of this case was 89.6%, 95% and 99%. [28] proposed a method of segmenting lungs from a chest X-ray image and using random patches from that segmented image to train a pretrained ResNet-18 [29] to classify COVID-19. [30] used a small dataset of 50 normal and 50 COVID-19 patients images to train InceptionV3, ResNet-

50 and Inception-ResnetV2 models and got an accuracy of 97%, 98% and 87% respectively for 2 classes.

In this study, we propose a deep learning-based approach for detecting COVID-19 from patient's chest X-ray images. We developed a convolutional neural network model with InceptionNetV3 [31] as the backbone and use transfer learning to initialize the model with weights trained on Imagenet [32] dataset. We add our classification layer at the end of this model and do an end to end training for this entire network on a balanced dataset containing 3 class of normal, pneumonia, and COVID chest X-ray image and also train in on a 2 class dataset (COVID and Non-COVID). We were motivated by the efforts of the open-source research community and decided to play our part in this fight against this pandemic through this work.

II. Dataset

As COVID-19 is a new disease, there is a huge scarcity of publicly available chest X-ray images corresponding to COVID-19 patients. So no one specific dataset was available for this study and a combination of several datasets was used to address this problem.

We obtained pneumonia(both viral and bacterial) and normal chest x-ray images from [33], an open sourced dataset released on the Kaggle platform. The dataset contains 5,863 chest x-ray images(4273 pneumonia and 1590 normal) classified into two categories.

For COVID-19 chest x-ray images, we collected our data from Dr. Cohen's [34] open-source Github repository. The repository contains an open database of Covid-19 cases with chest X-ray or CT images and is being updated regularly. Chest x-ray images are largely compiled from websites such as Radiopaedia.org, the Italian Society of Medical and Interventional Radiology, and Figure1.com [34]. At the time of writing this paper, the repository contained 408 COVID-19 chest x-ray images.

Our combined primary dataset consisted of 4273 pneumonia, 1590 normal, and 408 COVID-19 chest x-ray images. As the dataset was unbalanced, a resampling technique called random under-sampling [35] was used to make the dataset balanced. It involved randomly deleting examples from the majority class until the dataset becomes balanced. Thus, our final dataset consists of 408 pneumonia, 408 normal, and 408 COVID-19 chest x-ray images. From this, we randomly distributed the images into train and test sub-folders and generated five different folds for cross-validation. The training set consists of 978 images of three different classes and the test set contains 246 images, also classified into three different classes. We also created a separate dataset for two-class classification which contains 408 COVID-19 and 816 non-COVID-19 chest x-ray images. Also, all the images were resized to 299 × 299 pixels with a resolution of 96 dpi.

TABLE I
Dataset Summary

Disease	No. of Sample Images	
	Primary Unbalanced Dataset	Balanced Dataset
COVID-19	408	408
Normal	1590	408
Pneumonia	4273	408

In table I, the summary of the prepared dataset is shown.

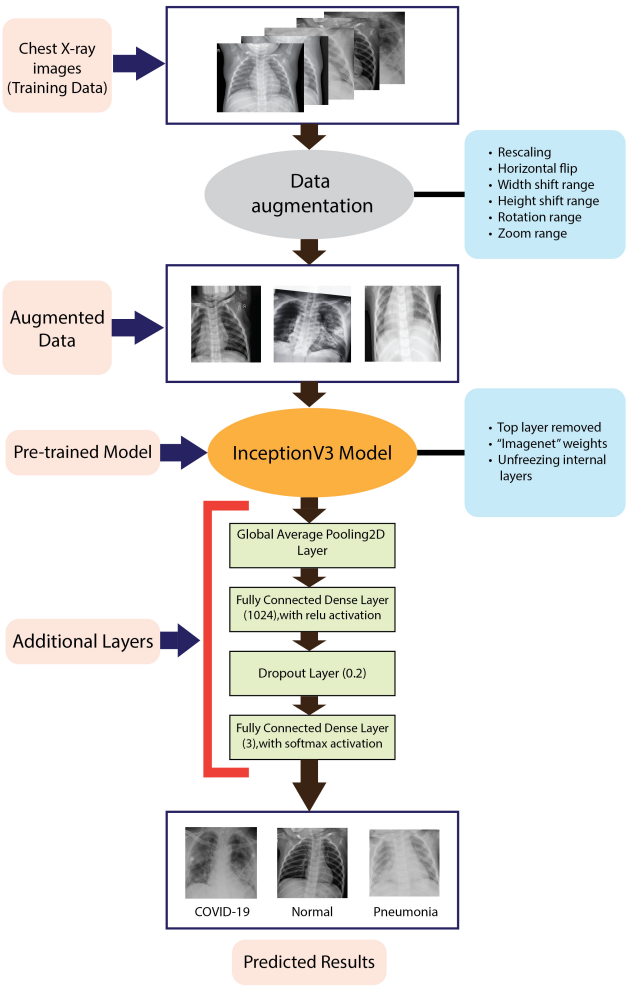

Fig. 1. Model Architecture and Workflow

III. Methodology

To solve our task of classifying COVID19 patients from chest X-ray images we developed a deep learning model and used various training schemes to improve the performance of the system. These are described in the following sections.

A. Data Augmentation

The input image size of the network architecture is 299 x 299. Some data augmentations of the following categories were also introduced to the input data: horizontal flip, translation, rotation, zoom etc. The model was built to classify chest x-ray images into three classes: Covid-19, Normal condition and Pneumonia (viral bacteria) which were labelled as 0, 1 and 2 respectively.

B. Model Architecture

1) Inception Block: Rather than preparing a deep learning model from scratch, it is always a better idea to lay a foundation of a tried and tested model and build a model on top of that. In our proposed model, we considered InceptionV3 to be our base model. In the history of the continuous progression of CNN networks, the Inception network was a revolutionary step forward because of its complex and well-engineered design.

InceptionNet was the first CNN classifier to use meticulous measures to ensure better performance while balancing out speed and accuracy at the same time by making the network wider rather than deeper. This wide network feature was achieved by using varying sizes of filters on the same level and this was implemented in InceptionV1 [31]. Later, in order to upgrade computational speed, a new scheme of factorizing filter size 5 x 5 convolution into two 3 x 3 convolutions was introduced along with representing any 3 x 3 convolution filter with a 3 x 1 convolution filter following a 1 x 3 convolution filter. This upgraded strategy was utilized in InceptionV3 [36] network which is the base model of our proposed architecture. InceptionV3's ability to learn an increasing number of features as a very deep CNN while maintaining high efficiency for the network was the motive behind this choice. It also has a smaller computational cost than other networks such as VGGNet. Although we are working with a comparatively smaller dataset now, eventually when we move on towards extended datasets, InceptionV3 possesses more probability of performing better than other seemingly simple models namely VGG, AlexNet, GoogLeNet etc.

2) Classification Layers: In our architecture, there are some other layers following the base model which is InceptionV3 Network. The first layer right after the base model is a global average pooling (GAP) [37] layer. This GAP layer lowers the total number of parameters making the model less prone to overfitting. Next comes a dense layer of 1024 units which uses the rectified linear unit (ReLU) activation function followed by a dropout layer where dropout fraction = 0.2. Finally, the model is completed with a softmax layer producing the outputs. This model has a total of 23,904,035 parameters, out of that 23,869,603 are trainable parameters and the other 34,432 are non-trainable parameters. Architecture details with layer parameters, output shape, etc are given in table II.

TABLE II
LAYER AND PARAMETER DETAILS OF THE PROPOSED ARCHITECTURE

Layer (type)	Output Shape	Parameters
InceptionV3 (Model)	(None, 8, 8, 2048)	21802784
Global_Average_Pooling	(None, 2048)	0
Dense	(None, 1024)	2098176
Dropout(0.2)	(None, 1024)	0
Dense	(None, 3)	3075
Total Parameters: 23,904,035		
Trainable Parameters: 23,869,603		
Non-trainable Parameters: 34,432		

C. Training Scheme

While training our network, we initialized the parameters using parameters from the pre-training done on the ImageNet dataset. This transfer learning scheme was adopted to avoid overfitting since the training dataset was relatively smaller. The hyperparameter values used while training the model are learning rate=0.01, epoch=40, batch size=32. To improve the performance of the model, the value of the loss function was checked every 3 epochs, and if it stayed constant learning rate was reduced to one-tenth of its value while ensuring the minimum learning rate to be 1e-6. The model was trained using the Adam optimization algorithm. The proposed model was implemented with Keras library using TensorFlow 2.0 backend. The entire training and testing process was performed on Google Colaboratory Server.

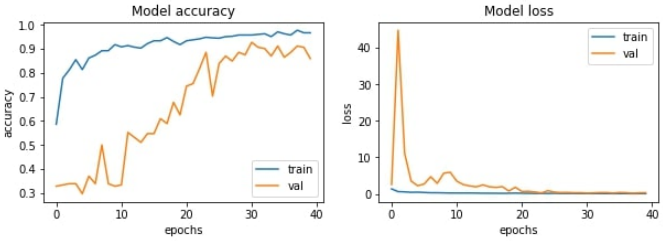

Fig. 2. Accuracy and Loss plot of the training

IV. RESULT

The model was trained for each of the 5 folds of data and it was tested on 3 class test set of each fold. For each of the test set we calculated Precision, Sensitivity, F1-score and Accuracy as our performance metric and it can be seen in table III.

TABLE III
PRECISION, SENSITIVITY, F1-SCORE AND ACCURACY ACROSS ALL 3 CLASSES FOR THE 5 FOLDS OF DATA

Folds	Precision(%)	Sensitivity(%)	F1-score(%)	Accuracy(%)
Fold 1	98.37	98.37	98.37	98.37
Fold 2	94.28	93.90	93.89	93.90
Fold 3	96.38	95.94	95.92	95.94
Fold 4	96.34	95.89	95.89	95.89
Fold 5	97.57	97.53	97.53	97.53
Average	**96.59**	**96.33**	**96.32**	**96.33**

From table III we can see that our model got a highest accuracy of 98.37% from fold 1 and the average accuracy for all the 5 folds is 96.33%. We also generated the same performance metrics in a class-wise basis for all of the folds. The class-wise result for fold-1 can be seen in table IV.

TABLE IV
PRECISION, SENSITIVITY, F1-SCORE AND ACCURACY OF THE 3 CLASSES FOR FOLD 1

Class	Precision(%)	Sensitivity(%)	F1-score(%)	Accuracy(%)
COVID19	98.80	100	99.39	100
Normal	97.56	97.56	97.56	97.56
Pneumonia	98.77	97.56	98.16	97.56

As evident from table IV, our model performance exceptionally well in the COVID19 class getting an accuracy of 100%. While for both the Normal and Pneumonia class it gets an accuracy of 97.56%. This claim are further supported by the confusion matrix we generated for each of the folds. The confusion matrix for fold-1 and fold-3 are presented in figure 3.

We also trained our model on a 2-class dataset derived from the 3-class dataset where the Normal and Pneumonia classes were labeled as Non-Covid19. We used the Precision, Sensitivity, Specificity, F1-score and Accuracy as the performance metric for this task. This detailed result is presented in table V.

As we can see from table V, our model performance increases significantly for the 2 class dataset and it obtained an accuracy of 98.78% for the COVID19 class. As for the Non-COVID19 class it was able to detect all the images correctly. We ran this setup for all of the folds and the confusion matrix

TABLE V
PRECISION, SENSITIVITY, SPECIFICITY, F1-SCORE AND ACCURACY OF THE 2 CLASSES FOR FOLD 2

Class	Precision(%)	Sensitivity(%)	Specificity(%)	F1-score(%)	Accuracy(%)
COVID19	100	98.78	100	99.39	98.78
Non-Covid19	99.39	100	98.78	99.67	100
Average	**99.70**	**99.39**	**99**	**99.54**	**99.39**

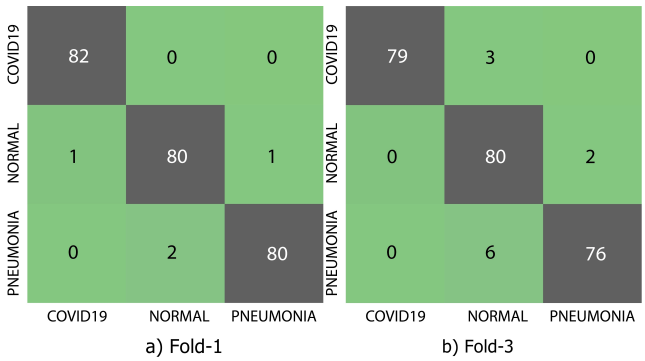

Fig. 3. Confusion Matrix of the Test Set for 3-class Dataset

for fold-1 and fold-2 are presented in Figure 4. From the confusion matrices we can see that the model was able to classify most of the test set images correctly for both the folds.

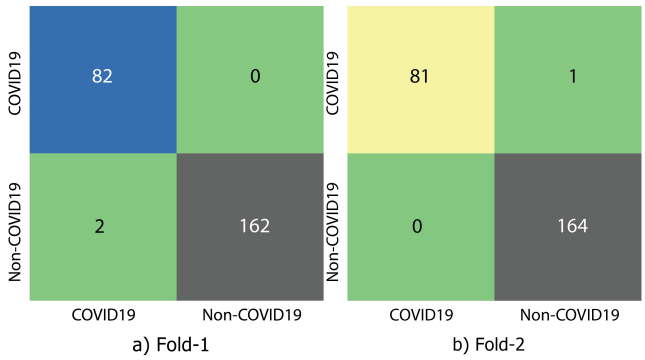

Fig. 4. Confusion Matrix of the Test Set for 2-class Dataset

As stated in section 3, we used data augmentation in our training scheme. We also tried to train model without using any of the data augmentations and compared their performance in table VI.

TABLE VI
COMPARISON OF PERFORMANCE FOR DATA AUGMENTATION

Data Type	Precision(%)	Sensitivity(%)	F1-score(%)	Accuracy(%)
With Augment	96.59	96.33	96.32	96.33
Without Augment	94.61	94.31	94.23	94.31

From table VI we can observe that the model performed better when using data augmentation techniques compared to when we trained without using any data augmentation. We got an accuracy of 94.31% without using data augmentation which is lower than our average accuracy of 96.33%.

While using Imagenet weights as initialization, we had the option of freezing some of the layer of InceptionV3 model with that weight and training rest of the layers. We tried freezing different number of layers and compared their accuracy in Figure 5.

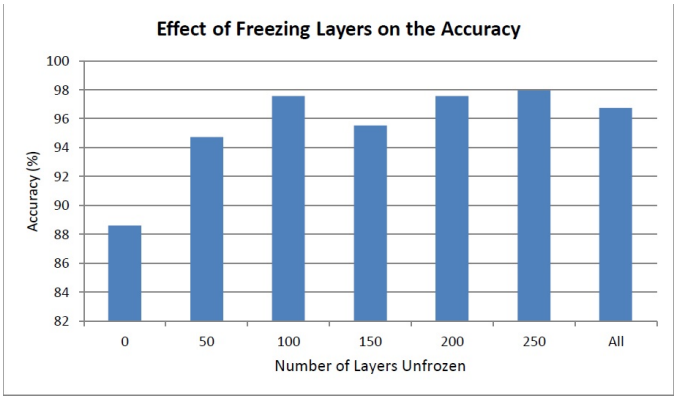

Fig. 5. Accuracy vs Number of Unfrozen Layers plot for our proposed model

As we can see from Figure 5 when all the layers were frozen the model performance was relatively poor but as we increased the number of unfrozen layers the accuracy went up. But there is very little difference in performance from 200 unfrozen layers onward.

As we mentioned in section 1, a lot of research work is currently being done on classifying COVID19 patients from chest X-ray images. These studies are being conduct in both 3-class and 2-class datasets with variation in number of images in the dataset and the model architecture. A comparison of our proposed system with the existing literature is presented in table VII.

V. DISCUSSION

Our proposed method of using InceptionV3 based architecture with pretrained Imagenet weights and employing data augmentation in the training scheme has resulted in an overall good accuracy for both the 3-class and 2-class dataset. As we can observe in table III the model had accuracy in the range of 93.90% to 98.37% for all of the folds of data. Even the lowest accuracy of 93.90% is still quite high. The model's performance did not vary that much across the folds meaning the system is robust and reliable.

The system's performance improves dramatically when applied in the 2-class dataset as should be the case because the task becomes much easier for the learning algorithm. Another thing to notice in table III and table V is the high precision and the sensitivity of the model. High sensitivity means a low number of false-negative cases and that our model misses less number of COVID19 cases. This holds for both the 3-class and 2-class setup and makes us hopeful of the system's potential.

Table VI shows that our intuition of using data augmentation and pre-trained weights for the network does lead to better performance. As in these cases using these techniques yields a much higher accuracy. Using data augmentation makes the

TABLE VII
COMPARISON OF OUR PROPOSED METHOD WITH THE EXISTING LITERATURE

Study	Architecture	Accuracy 3-class(%)	Accuracy 2-class(%)
Ioannis et al. [38]	VGG19	93.48	98.75
Wang and Wong [24]	Covid-Net	N/A	92.4
Sethy and Behra [39]	Resnet-50	N/A	95.38
Hemdan et al [40]	VGG19	N/A	90
Narin et al [30]	Resnet-50	N/A	98
Ozturk et al [25]	DarkCovidNet	87.02	98.08
Khan et al [27]	CoroNet	89.6	99
Proposed Method	**InceptionV3**	**96.33**	**xxxx**

model more generalized and robust which makes it able to deal with much more variation in test set distribution and helps these data-hungry deep learning algorithms learn better representation [41]. As for using pre-trained weights rather than training from the scratch, it was reported in [42] that when limited training data is available pre-trained networks tend to outperform networks that are trained from scratch. And our experimentation also led to the same conclusion.

Table VII compares our proposed method with the existing literature. It can be seen here that our model manages to perform better than the other methods presented in this table. We get an average accuracy of 96.33% for the 3-class setup, while for the 2-class setup it is 99.39%. Another point to note here is that we use more COVID19 class data in our study compared to the other studies mentioned. [38] acquired an accuracy of 93.48% for the 3-class setup but used 224 COVID, 700 Pneumonia, and 504 normal class images. DarkCovidNet [25] used 224 COVID, 500 Pneumonia and 500 normal class images that resulted in an accuracy of 87.02% in the 3-class setup. CoroNet [27] used 310 normal, 330 pneumonia-bacterial and 327 Pneumonia-viral X-ray images for their 3-class setup and got an accuracy of 89.6%. So using more COVID19 data and improving the performance makes our system a lot more reliable.

VI. CONCLUSION

Despite the success of our method when compared against other existing literature we do have certain limitations, especially when it comes to deploying this sort of a model for large scale practical use. The most important limitation is the lack of verified datasets of chest X-rays of Covid-19 patients available for training our model to make it practically deployable. Currently, we have only trained the model for 3-classes, but training it with more classes such as SARS, bacterial and viral pneumonia, etc. would certainly make it more robust. Finally, we still do not know the human-level accuracy for detecting Covid-19 from chest X-rays, and it would be interesting to compare our model against that.

It has been more than 6 months since the advent of the Covid-19 pandemic, and still, the number of people being infected by the virus is increasing every day. Using chest radiology is one of the most common methods used to detect infected patients. However, this process is done manually by radiologists and as such, there is a glaring shortage of radiologists around the world in this critical period. At the same time manually checking every patient's X-ray is a time-consuming task and is prone to human errors. Thus it is vital to develop a fast, reliable, and automated process for detection of the virus in chest X-rays. Our method of automatic COVID 19 detection from Chest X-ray images is one such process. Despite not being production-ready, we believe that based on the ideas of our literature a large-scale model can be built and deployed. Through that, we can achieve a faster diagnosis, faster isolation of infected patients, and less human contact during diagnosis, all of which will lead to much stronger mitigation of the Covid-19 pandemic.

REFERENCES

[1] Z. Wu and J. M. McGoogan, "Characteristics of and important lessons from the coronavirus disease 2019 (covid-19) outbreak in china: summary of a report of 72 314 cases from the chinese center for disease control and prevention," *Jama*, vol. 323, no. 13, pp. 1239–1242, 2020.

[2] J. M. Sanders, M. L. Monogue, T. Z. Jodlowski, and J. B. Cutrell, "Pharmacologic treatments for coronavirus disease 2019 (covid-19): a review," *Jama*, vol. 323, no. 18, pp. 1824–1836, 2020.

[3] H. A. Rothan and S. N. Byrareddy, "The epidemiology and pathogenesis of coronavirus disease (covid-19) outbreak," *Journal of autoimmunity*, p. 102433, 2020.

[4] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong *et al.*, "Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia," *New England Journal of Medicine*, 2020.

[5] V. M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D. K. Chu, T. Bleicker, S. Brünink, J. Schneider, M. L. Schmidt *et al.*, "Detection of 2019 novel coronavirus (2019-ncov) by real-time rt-pcr," *Eurosurveillance*, vol. 25, no. 3, p. 2000045, 2020.

[6] L. Guo, L. Ren, S. Yang, M. Xiao, D. Chang, F. Yang, C. S. Dela Cruz, Y. Wang, C. Wu, Y. Xiao *et al.*, "Profiling early humoral response to diagnose novel coronavirus disease (covid-19)," *Clinical Infectious Diseases*, 2020.

[7] P. Verdecchia, C. Cavallini, A. Spanevello, and F. Angeli, "The pivotal link between ace2 deficiency and sars-cov-2 infection," *European Journal of Internal Medicine*, 2020.

[8] M.-Y. Ng, E. Y. Lee, J. Yang, F. Yang, X. Li, H. Wang, M. M.-s. Lui, C. S.-Y. Lo, B. Leung, P.-L. Khong *et al.*, "Imaging profile of the covid-19 infection: radiologic findings and literature review," *Radiology: Cardiothoracic Imaging*, vol. 2, no. 1, p. e200034, 2020.

[9] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia, "Correlation of chest ct and rt-pcr testing in coronavirus disease 2019 (covid-19) in china: a report of 1014 cases," *Radiology*, p. 200642, 2020.

[10] C. Shen, N. Yu, S. Cai, J. Zhou, J. Sheng, K. Liu, H. Zhou, Y. Guo, and G. Niu, "Quantitative computed tomography analysis for stratifying the severity of coronavirus disease 2019," *Journal of Pharmaceutical Analysis*, 2020.

[11] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song *et al.*, "Artificial intelligence distinguishes covid-19 from community acquired pneumonia on chest ct," *Radiology*, p. 200905, 2020.

[12] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[13] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *Ieee Access*, vol. 6, pp. 9375–9389, 2017.

[14] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual review of biomedical engineering*, vol. 19, pp. 221–248, 2017.

[15] J. Shi, X. Zheng, Y. Li, Q. Zhang, and S. Ying, "Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of alzheimer's disease," *IEEE journal of biomedical and health informatics*, vol. 22, no. 1, pp. 173–183, 2017.

[16] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images," *IEEE transactions on medical imaging*, vol. 35, no. 1, pp. 109–118, 2015.

[17] A. Cruz-Roa, A. Basavanhally, F. González, H. Gilmore, M. Feldman, S. Ganesan, N. Shih, J. Tomaszewski, and A. Madabhushi, "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks," in *Medical Imaging 2014: Digital Pathology*, vol. 9041. International Society for Optics and Photonics, 2014, p. 904103.

[18] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[19] C. Wang, A. Elazab, J. Wu, and Q. Hu, "Lung nodule classification using deep feature fusion in chest radiography," *Computerized Medical Imaging and Graphics*, vol. 57, pp. 10–18, 2017.

[20] H. Kim and S. Hwang, "Scale-invariant feature learning using deconvolutional neural networks for weakly-supervised semantic segmentation," *arXiv preprint arXiv:1602.04984*, 2016.

[21] E. Soleymanpour, H. R. Pourreza *et al.*, "Fully automatic lung segmentation and rib suppression methods to improve nodule detection in chest radiographs," *Journal of medical signals and sensors*, vol. 1, no. 3, p. 191, 2011.

[22] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya *et al.*, "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning," *arXiv preprint arXiv:1711.05225*, 2017.

[23] S. Candemir, S. Jaeger, K. Palaniappan, J. P. Musco, R. K. Singh, Z. Xue, A. Karargyris, S. Antani, G. Thoma, and C. J. McDonald, "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," *IEEE transactions on medical imaging*, vol. 33, no. 2, pp. 577–590, 2013.

[24] L. Wang and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images," *arXiv preprint arXiv:2003.09871*, 2020.

[25] 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," *Computers in Biology and Medicine*, p. 103792, 2020.

[26] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[27] A. I. Khan, J. L. Shah, and M. M. Bhat, "Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images," *Computer Methods and Programs in Biomedicine*, p. 105581, 2020.

[28] Y. Oh, S. Park, and J. C. Ye, "Deep learning covid-19 features on cxr using limited training data sets," *IEEE Transactions on Medical Imaging*, 2020.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[30] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," *arXiv preprint arXiv:2003.10849*, 2020.

[31] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[32] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.

[33] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan *et al.*, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.

[34] J. P. Cohen, P. Morrison, and L. Dao, "Covid-19 image data collection," *arXiv preprint arXiv:2003.11597*, 2020.

[35] M. A. Tahir, J. Kittler, and F. Yan, "Inverse random under sampling for class imbalance problem and its application to multi-label classification," *Pattern Recognition*, vol. 45, no. 10, pp. 3738–3750, 2012.

[36] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[37] M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv preprint arXiv:1312.4400*, 2013.

[38] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, p. 1, 2020.

[39] P. K. Sethy and S. K. Behera, "Detection of coronavirus disease (covid-19) based on deep features," *Preprints*, vol. 2020030300, p. 2020, 2020.

[40] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images," *arXiv preprint arXiv:2003.11055*, 2020.

[41] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *2018 international interdisciplinary PhD workshop (IIPhDW)*. IEEE, 2018, pp. 117–122.

[42] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1299–1312, 2016.