A Review: Deep Learning Classification Performance of Normal and COVID-19 Chest X-ray Images

M A Markom¹,²*, S Mohd Taha³, A H Adom¹,³, A S Abdull Sukor¹,⁴, A S Abdul Nasir³, H Yazid³, E S M M Tan³, A M Markom⁵

¹ Advanced Sensor Technology, Center of Excellence, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia
² Faculty of Electronic Engineering Technology, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia
³ Faculty of Electric Engineering Technology, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia
⁴ Faculty of Mechanical Engineering Technology, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia
⁵ Universiti Teknologi MARA Cawangan Johor Kampus Pasir Gudang, Jalan Purnama, Bandar Seri Alam 81750 Masai, Johor, Malaysia

*mark@gmail.com

Abstract. COVID19 chest X-ray has been used as supplementary tools to support COVID19 severity level diagnosis. However, there are challenges that required to face by researchers around the world in order to implement these chest X-ray samples to be very helpful to detect the disease. Here, this paper presents a review of COVID19 chest X-ray classification using deep learning approach. This study is conducted to discuss the source of images and deep learning models as well as its performances. At the end of this paper, the challenges and future work on COVID19 chest X-ray are discussed and proposed.

1. Introduction

In early December 2019, the world has been under immense strain due to the outbreak of Coronavirus Disease 2019 (COVID-19). Coronavirus disease 2019 (COVID-19) is a respiratory infection that caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) that was first identified in Wuhan City, Hubei Province, China [1]. The disease was at first called as SARS-CoV-2, and later it was named COVID-19 by World Health Organization (WHO) in February 2020 [2]. This new virus only took only 30 days to spread from Wuhan to other regions of China. COVID-19 symptoms include fever, myalgia, dry cough, nausea, sore throat, and chest pain [3]. In serious cases, the infection can lead to pneumonia, respiratory difficulties, multi-organ failure, and fatal [4-5].

COVID-19 seems to be highly infectious and rapidly spreading worldwide, infecting more than 210 countries. [3]. There is currently no specific medication or treatment available to cure this infection. According to the World Health Organization (WHO), more than five million people have been infected worldwide to date, with approximately three million confirmed death cases [3]. In addition, due to the high and rising growth rate of COVID-19 cases, many developing countries' health services are on the
verge of collapsing. Since the COVID-19 epidemic has become a national crisis, the WHO declared COVID-19 a global pandemic on March 11, 2020 [6].

Since the virus normally infects the lungs, it has been discovered that medical imaging methods such as chest X-ray plays an important role in diagnosing COVID-19 cases. To date, The X-ray image is used as a supplementary or support material in diagnose COVID19 severe level [7]. However, the severity results are still requiring a decision from the experts and other materials, for example, oxygen saturation in blood (SpO2).

For radiologist, they find it difficult to classify the severity in COVID-19 cases because the images of the chest X-rays in each class are nearly same. Another issue that are facing by the radiologist is there are many common features between medical images of COVID-19 and pneumonia. They must identify the typical patterns of the disease that are often shared with other types of viral pneumonia, such as common flu (Influenzas A) which leads to errors in their diagnosis. This similarity makes a differential diagnosis of COVID-19 cases by expert radiologists challenging. These issues need to be highlighted since it can speed up the triage process of non-COVID-19 cases and maximize the allocation of hospital resources to COVID-19 cases.

Furthermore, the study of classification of COVID-19 chest X-ray is still new and hence require more studies to assist it since the number of COVID-19 cases is still increasing. This research therefore introduces the use of a machine learning approach for classification of COVID-19 and normal CXR images. Machine learning (ML) based methods have been widely used in the analysis of medical images. ML-based approaches are scalable, automatable, and easy to implement in clinical settings [8-9]. In addition, this approach can be easily implemented in any future viral outbreak for the rapid classification of CXR images.

Practically, the classification or identification of severity level of COVID-19 only can be conducted by specialist of medical officer through chest X-ray. Therefore, machine learning can be a good tool for assisting medical staff in determining the severity level of the COVID-19. Most of studies present deep learning approach to classify normal and COVID19 chest x-ray images [10-11]. This is due to the fact that deep learning techniques in computer-aided methods add greatly to state-of-the-art medical image analysis and offer good results.

In this study, a few highlighted parameters when conducting classification of deep learning will be discussed. The parameters consist of image or data collection, pre-processing method, deep learning models and their performance. Aside from that, there will be an open discussion about the challenge and potential of the COVID19 classification system, which uses chest X-ray as the primary diagnostic technique.

2. COVID19 Chest X-ray Image Dataset

In machine learning, classification performance with high accuracy can be depended on a lot of images. This is because, a variety of images are forced to learn the classification algorithm. A huge number of images will provide more learning samples to the algorithm, which may bring more possible of input to the classification. The beginning of COVID19 X-ray images studies is started in April 2020 [12]. At early stage, many studies have obtained their images from public library, such as, GitHub by Joseph Cohen from Standford University, USA [12], Radiopedia.org and Italian Society of Radiology. In terms of ethical documentation, certain images have ethical, while, some have not [12]. Table 1 shows high citation of deep learning studies with their images. While, figure 1 shows variety of normal and COVID19 chest X-ray images that obtained from the dataset.

In the public library, such as in Github by Dr Cohen [12], there are COVID19 chest X-ray, which is properly kept in one folder. While, in Kaggle repository [25], the chest X-ray images have been separated accordingly, normal, COVID19 and pneumonia. Figure 1 shows various of normal and COVID19 patient chest X-ray images that taken from one of the public libraries. All these images are original. In the other words, no augmentation or enhancing process has taken place.
Table 1. Related studies with their sources of images

| No | Author and reference | Sources of images                                                                 | Number of citations based on Google scholar |
|----|----------------------|----------------------------------------------------------------------------------|---------------------------------------------|
| 1  | Apostolopoulos, I. D. et. All (2020) [13] | 1. Cohen et al (2020) [12]  
2. Radiological Society of North America (RSNA)  
3. Radiopaedia  
4. Italian Society of Medical and Interventional Radiology (SIRM)  
5. Kermany, D. S et al. [23] | 682 |
| 2  | Ozturk, T. et. All (2020) [14] | 1. Cohen et al (2020) [12]  
2. Wang, Y. et al (2020) [24] | 628 |
| 3  | Narin, A. et. Al (2021) [15] | 1. Cohen et al (2020) [12]  
2. Kaggle repository [25] | 567 |
| 4  | Khan, A. I. et all (2020) [16] | 1. Cohen et al (2020) [12]  
2. Radiological Society of North America (RSNA)  
3. Radiopaedia | 267 |
| 5  | Chowdhury, M. E. et al [17] | 1. Cohen et al (2020) [12]  
2. Radiological Society of North America (RSNA)  
3. Italian Society of Medical and Interventional Radiology (SIRM)  
4. Kaggle repository [25] | 257 |
| 6  | Abbas, A. et al (2021) [18] | 1. Cohen et al (2020) [12]  
2. Japanese Society of Radiological Technology (JSRT) | 253 |
| 7  | Oh, Y. (2020) [19] | 1. Cohen et al (2020) [12]  
2. Japanese Society of Radiological Technology (JSRT)  
3. Kaggle repository [25]  
4. U.S. National Library of Medicine (USNLM)  
5. Corona Hack: Chest X-Ray-Dataset [26] | 217 |
| 8  | Minaee, S. et al [20] | 1. Cohen et al (2020) [12] | 183 |
| 9  | Farooq, M. et al (2020) [21] | 1. Wang, L. et al (2020) [27] | 180 |
| 10 | Panwar, H. (2020) [22] | 1. Cohen et al (2020) [12]  
2. Kaggle repository [25]  
3. Radiopaedia | 127 |

Images that have been collected, basically, will be proceed to validation by medical officers or analysed using statistical analysis. After that, most of the images will continue to an augmentation process in order to ensure the images are suitable and perfect to execute in deep learning algorithm. Also, to ensure that the number of images are sufficient to process using deep learning algorithm. There are a few operations to do the augmentation process, which is involved with cropping, resizing, flipping, rotating and padding.

In [28], they show how they applied flipping and rotation to generate more images for their deep learning algorithm. This is one of the methods to get sufficient number of samples for the deep learning algorithm. In this study, they started with flipping operation for 90 original images. Then, the original 90 images have further been rotated by 90°angle to get 90 more images. Then the images are rotated again by 180°angle to get 90 more images. Finally, the original 90 images were further rotated by 270°angle to get more 90 images. These operations resulted in a dataset containing 450 COVID-19 X-ray images. Figure 2 shows the flipping and rotation operation.
3. Deep Learning Models and Its Performances

Deep learning is a sub-field of machine learning that is inspired by the brain's structure. In the field of medical image processing, as in many other fields, deep learning approaches have continued to demonstrate excellent results in recent years. It is hoped that by applying deep learning algorithms to medical data, relevant results might be extracted. The good points of deep learning, it can cater high dimensional and huge number of data, which is very suitable to images. Besides, this approach may reduce a risk of overfitting process and misclassification results [29].

Here, this study presents a few models of deep learning that have been used in COVID19 chest X-ray applications. The common models in typical convolution neural network (CNN) [30]. Table 2 shows a summary of deep learning models and its performance, particularly implemented in COVID19 chest Xray studies. The most common models that have been widely used are VGGNet and ResNet.

VGGNet is a Convolutional Neural Network (CNN) architecture proposed by Karen Simonyan and Andrew Zisserman from the University of Oxford in 2014 [31]. Some states that this VGGNet improves AlexNet performance. The architecture of VGGNet consists of convolution, max pooling and fully connected process. The architecture also includes activation, loss and optimiser function. The keyword ‘Net’ after ‘VGG’ term, defines the number of layers that been included in the model. ResNet, short for Residual Network is a specific type of neural network that was introduced in 2015 by Kaiming He [32]. It is very successful learning algorithm which very well-known to solve complex problems. It does have a sub-function to skip connections, or shortcuts to jump over some layers. This allows the algorithm to have less layers, and perhaps, less learning time. Similar to VGGNet, the keyword ‘Net’ indicates the number of layers in the model structure.
Table 2. Deep learning models and its performances

| DL Model          | Researchers                        | Performance                                                                 | No of citation based on Google scholar till June 2021 | Reference |
|-------------------|------------------------------------|----------------------------------------------------------------------------|-------------------------------------------------------|-----------|
| CNN - VGGNet      | 1. Apostolopoulos, I. D. et. All (2020)  
                     2. Abbas, A. et al (2021)  
                     3. Wang, L. et al (2020) | 1. Accuracy: 98.75%, Sensitivity: 92.85%, Specificity: 98.75%  
                                                               2. Accuracy: 97.35%, Sensitivity: 98.23%, Specificity: 96.34%  
                                                               3. Accuracy: 83%, Sensitivity: 58.7% | 1. 682  
                                                               2. 253  
                                                               3. 877 | 1. [13]  
                                                               2. [18]  
                                                               3. [27] |
| CNN - MobileNet V2| 1. Apostolopoulos, I. D. et. All (2020) | 1. Accuracy: 97.40%, Sensitivity: 99.10%, Specificity: 97.09% | 1. 682 | 1. [13] |
| CNN - ResNet      | 1. Narin, A. et al (2021)  
                     2. Abbas, A. et al (2021)  
                     3. Minaee, S. et al (2020)  
                     4. Farooq, M. et al (2020)  
                     5. Wang, L. et al (2020) | 1. Accuracy: 96.1%  
                                                               2. Accuracy: 95.12%, Sensitivity: 91.87%  
                                                               3. Sensitivity: 98%, Specificity: 90.7%  
                                                               4. Accuracy: 96.23%  
                                                               5. Accuracy: 90.6%, Sensitivity: 83% | 1. 575  
                                                               2. 253  
                                                               3. 183  
                                                               4. 180  
                                                               5. 877 | 1. [15]  
                                                               2. [18]  
                                                               3. [20]  
                                                               4. [21]  
                                                               5. [27] |
| CNN - Inception   | 1. Narin, A. et. Al (2021) | 1. Accuracy: 95.4% | 1. 575 | 1. [15] |
| CNN - CoroNet/CovNet | 1. Khan, A. I. et all  
                     2. Farooq, M. et al (2020)  
                     3. Panwar, H. (2020)  
                     4. Wang, L. et al (2020) | 1. Precision: 90%, Specificity: 96.4%, Accuracy: 89.6%  
                                                               2. Accuracy: 83.5%  
                                                               3. Accuracy: 97.62%  
                                                               4. Accuracy: 93.3%, Sensitivity: 91% | 1. 267  
                                                               2. 180  
                                                               3. 132  
                                                               4. 877 | 1. [16]  
                                                               2. [21]  
                                                               3. [22]  
                                                               4. [27] |
| CNN - GoogleNet   | 1. Abbas, A. et al (2021) | 1. Accuracy: 94.71%, Sensitivity: 97.88%, Specificity: 95.76% | 1. 253 | 1. [18] |
| CNN - SqueezeNet  | 1. Abbas, A. et al (2021)  
                     2. Minaee, S. et al (2020) | 1. Accuracy: 94.90%, Sensitivity: 95.70%, Specificity: 94.71%  
                                                               2. Sensitivity: 98%, Specificity: 92.9% | 1. 253  
                                                               2. 183 | 1. [18]  
                                                               2. [20] |
| Inception-ResNetV2| 1. Narin, A. et. Al (2021) | 1. Accuracy: 94.2 | 1. 575 | 1. [15] |
| ResNet-SVM        | 1. Ismael, A. M. et al (2021) | 1. Accuracy: 95.79%, Sensitivity: 94%, Specificity: 97.73% | 1. 50 | 1. [30] |

Based on Table 2, most studies that implemented deep learning, have been used multiple types of models in order to do their classification. These models are studied to obtain the best performance. And, the performances are always indicated in terms of Accuracy, Specificity and Sensitivity. In this summary, there are same models that performed with high and low accuracy, such as, model of CNN-
VGGNet. Three studies have performed using this model but obtained different results. This is caused by other parameters, for example, CNN structures or architecture, activation function, loss function and etc. The main reason is due to the usage of images. More images, usually, will get better performance compared to a smaller number of images.

Apart from study using different types of models, deep learning in COVID19 chest X-ray study can have more various objectives or purposes. The list below is examples that can be used as objective when conducting deep learning study:

i. To study the performance of same type of deep learning model, but, execute with different number of layers. For example, make use model of VGG11, VGG16 and VGG19.

ii. To study the performance of same type of deep learning model, but, execute with different number of convolution kernel, pooling process and convolution layer.

iii. To study the performance of different activation, loss and optimiser function.

Actually, there are a lot of studies can be considered when it deals with deep learning for COVID19 chest X-ray. Perhaps, for other disease too.

4. Deep Learning Challenge in Diagnosis of COVID19 chest X-ray

In these two years, starting 2020 to 2021, many deep learning studies on X-ray or radiography have been conducted to promote and propose as a primary diagnosis tool for COVID19 disease. These studies have shown their good results and performance. Also, the best deep learning algorithm models. These studies have proven that deep learning algorithms are able to classify or differentiate between normal and COVID19 chest X-ray.

Hence, to keep the study more reliable and to be part of primary diagnosis tool, perhaps, now is the time to move the study to classify the severity level of COVID19. In Malaysia, COVID19 severity of infection is classified into five stages or level, beginning from stage 1 up to stage 5. Stage 1 defines as early stage of infection. This classification decision is made mostly based on the condition of SpO2 level, respiratory frequency and difficulty of breathing symptom. Table 3 shows the severity level and its common symptoms [33]. However, clinical guidelines and clinical studies may have overlapping or differing criteria for each category, and a patient's clinical condition may alter over time.

| Severity Level | Known as                              | Description                                                                 |
|----------------|---------------------------------------|----------------------------------------------------------------------------|
| 1              | Asymptomatic or Presymptomatic Infection| Individuals who have no symptoms that are consistent with COVID-19.         |
| 2              | Mild Illness                          | Individuals who have light symptoms of COVID-19 (e.g., fever, cough, sore throat, malaise, headache, muscle pain, nausea, vomiting, diarrhoea, loss of taste and smell). |
| 3              | Moderate Illness                      | Individuals who show evidence of lower respiratory disease during clinical assessment or imaging and have an oxygen saturation (SpO2) ≥94% on room air at sea level. |
| 4              | Severe Illness                        | Individuals who have SpO2 <94% on room air at sea level, a ratio of arterial partial pressure of oxygen to fraction of inspired oxygen (PaO2/FiO2) <300 mm Hg, respiratory frequency >30 breaths/min, or lung infiltrates >50%. |
| 5              | Critical Illness                      | Individuals who have respiratory failure, septic shock, and/or multiple organ dysfunction. |
Apart from the severity level classification, there is a challenge based on medical reports, stated that, in certain circumstances, the chest X-ray displayed at early stage of infection, but, the SpO2 level of patient is severe, which is can be less than 80%. This end up the patient to be in stage 4 or 5, and required oxygen support or ventilation machine. This kind of cases always happen to senior citizen or individual with high-risk disease. Due to this circumstance, it leads the X-ray results become the supplementary tool. Due to this challenge and difficulties, a huge number of studies must be conducted. Since X-ray really helps and has been used to support the severity level of COVID19, the study can be enhanced to the next level by focusing to severity level classification. Prospective predictions are the future.

References
[1] Chauhan S 2020 Comprehensive review of coronavirus disease 2019 (COVID-19) Biomedical journal 43(4) 334-340.
[2] WHO 2020 Coronavirus disease 2019 (COVID-19) – Situation report 2 in February 2020.
[3] Lapostolle F, Schneider E, Viana I, Dollet G, Roche B, Berdah J and Adnet F 2020 Clinical features of 1487 COVID-19 patients with outpatient management in the Greater Paris: the COVID-calling study Internal and emergency medicine 15 813-817.
[4] Zaim S, Chong JH, Sankaranarayanan V and Harky A 2020 COVID-19 and multiorgan failure: A narrative review on potential mechanisms Journal of Molecular Histology 51 613–628.
[5] Cucinotta D and Vanelli M 2020 WHO declares COVID-19 a pandemic Acta Bio Medica: Atenei Parmensis 91(1) 157.
[6] Cleverley J, Piper J, Jones MM 2020 The role of chest radiography in confirming covid-19 pneumonia BMJ 2020; 370 :m2426 doi:10.1136/bmj.m2426
[7] Syeda HB, Syed M, Sexton KW, Syed S, Begum S, Syed F, and Yu Jr F 2021 Role of machine learning techniques to tackle the COVID-19 crisis: Systematic review JMIR medical informatics 9(1) e23811.
[8] Kwekha-Rashid AS, Abduljabbar HN and Alhayani B 2021 Coronavirus disease (COVID-19) cases analysis using machine-learning applications Applied Nanoscience 4639 1-13.
[9] Islam MM, 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.
[10] Hariri W and Narin A 2020 Deep Neural Networks for COVID-19 Detection and Diagnosis using Images and Acoustic-based Techniques: A Recent Review Soft computing 1–18.
[11] Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, and Ghassemi M. 2020 Covid-19 image data collection: Prospective predictions are the future. arXiv preprint arXiv:2006.11988.
[12] Apostolopoulos ID and Mpesiana TA 2020 Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Physical and Engineering Sciences in Medicine 43(2) 635-640.
[13] Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O and Acharya UR 2020 Automated detection of COVID-19 cases using deep neural networks with X-ray images Computers in biology and medicine 121 103792.
[14] Narin A, Kaya C and Pamuk Z 2021 Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks Pattern Analysis and Applications 24 1207-1220.
[15] Khan AI, Shah JL and Bhat MM 2020 CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images Computer Methods and Programs in Biomedicine 196 105581.
[17] Chowdhury ME, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub ZB and Islam MT 2020 Can AI help in screening viral and COVID-19 pneumonia? IEEE Access 8 132665-132676.

[18] Abbas A, Abdelsamea MM and Gaber MM 2021 Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network Applied Intelligence 51(2) 854-864.

[19] Oh Y, Park S and Ye J C 2020 Deep learning covid-19 features on cxr using limited training data sets IEEE Transactions on Medical Imaging 39(8) 2688-2700.

[20] Minaee S, Kafieh R, Sonka M, Yazdani S and Soufi GJ 2020 Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning Medical image analysis 65 101794.

[21] Farooq M and Hafeez A 2020 Covid-resnet: A deep learning framework for screening of covid19 from radiographs arXiv preprint arXiv:2003.14395.

[22] Panwar H, Gupta PK, Siddiqui MK, Morales-Menendez R and Singh V 2020 Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet Chaos, Solitons & Fractals 138 109944.

[23] Kermany DS, Goldbaum M, Cai W, Valentim CC, Liang H, Baxter SL and Zhang K 2018 Identifying medical diagnoses and treatable diseases by image-based deep learning Cell, 172(5) 1122-1131.

[24] Wang X, Peng Y, Lu L, Lu Z and Bagheri M, Summers RM 2017 Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2097-2106.

[25] Covid-19 chest x-ray database, c. https://www.kaggle.com/tawsifurrahman/covid19-radiography-database/ Accessed April 9, 2020.

[26] Corona Hack: Chest X-Ray-Dataset. Accessed: Mar. 21, 2020. [Online]. Available: https://www.kaggle.com/praveengovi/coronahackchest-xray dataset

[27] Wang L, Lin ZQ and Wong A 2020 Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images Scientific Reports 10(1) 19549.

[28] Reshi AA, Rustam F, Mehmood A, Alhossan A, Alrabiah Z, Ahmad A and Choi GS 2021 An Efficient CNN Model for COVID-19 Disease Detection Based on X-Ray Image Classification Complexity 2021(6621607) 1-12.

[29] Fakoor R, Ladhak F, Nazi A and Huber M 2013 Using deep learning to enhance cancer diagnosis and classification Proceedings of the international conference on machine learning 28 3937-3949.

[30] Ismael AM and Şengür A 2021 Deep learning approaches for COVID-19 detection based on chest X-ray images Expert Systems with Applications 164 114054.

[31] Simonyan K and Zisserman A 2015 Very deep convolutional networks for large-scale image recognition CoRR. abs/1409.1556. preprint arXiv:1409.1556.

[32] He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition Proceedings of the IEEE conference on computer vision and pattern recognition 770-778.

[33] National Institution of Health 2021 Clinical Spectrum of SARS-CoV-2 Infection. COVID-19 Treatment Guidelines. www.covid19treatmentguidelines.nih.gov/ overview/ clinical-spectrum/. A reference.