# Development Of Convolutional Neural Network (CNN)-Based Deep Learning Model For Prediction Of Covid-19 Infection

**Henshaw Jumbo**[1]
Department of Computer Engineering Heritage Polytechnic, Eket
Akwa Ibom State Nigeria
henshawjumbo@yahoo.com

**Idara Ufot Akpan**[2]
Department of Computer Science Heritage Polytechnic, Eket
Akwa Ibom State Nigeria
akpan.idara@gmail.com

**Clement, Kingsley Bassey**[3]
Department Of Electrical/Electronic And Computer Engineering,
University of Uyo, Akwa Ibom State Nigeria
kingsleyclement@uniuyo.edu.ng

***Abstract***— **In this paper, development of convolutional neural network (CNN)-based deep learning model for prediction of Covid-19 infection is presented. The CNN-based deep learning technique was based on Covid-19 chest X-ray images dataset from the GitHub repository with 712 images of those persons with Covid-19 and 1583 images of those persons that are normal. The study utilized the chest X-ray image dataset to train, validate and test the model. The dataset images were first pre-processed to suit the application of deep learning techniques after which Google Colab Graphics processing unit (GPU) was used to train the COVID-19 model for 3 hours and 70 epochs. The classification model results show a training loss values of 0.0565 with training accuracy values of 98.34%. Also, the training precision results is 0.9872, training recall values is 0.9891 while the validation loss value is 0.0633. Furthermore, the results shows that the validation accuracy is 98.62%, the validation precision results is 0.9921 while the validation recall value is 0.9881. In addition, the test loss results is 0.0791, the test accuracy results is 96.07% while the test precision values 0.9571, and the test recall values is 0.9842. Altogether, the COVID-19 classification mode accuracy was very high with a value above 98 %. This shows that the model can effectively predict COVID-19 infection by analyzing the chest X-ray images of people suspected to have such infection.**

***Keywords— Deep Learning, COVID-19, chest X-ray images, Google Colab, Graphics Processing Unit, Classification Model, Hyper-Parameters***

## 1. INTRODUCTION

CORONAVIRUS disease denoted as COVID-19 is a form infectious disease that is attributed to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1,2,3,4,5,6,7,8]. It spread rapidly across the globe to such an extent that the WHO (world health organisation) declared it a pandemic [9,10,11,12,13,14,15]. The popular symptoms of COVID-19 include fever, sore throat, respiratory disorder, fatigue, shortness of breath as well as muscular pains. [16,17,18,19,20,21,22,23] Available clinical reports show that early detection of the infection and isolation of infected persons are the most effective ways to stem the spread of the disease. Also, early detection will help to commences treatment to avoid complications of severe infection.

Accordingly, in this paper, a convolutional neural network (CNN)-based deep learning model for prediction of COVID-19 infection is presented [24,25,26,27,28,29,30]. Although, there are some other ways of screening for COVID-19 infection, the approach presented in this paper is based on the use of the dataset of ChestX-ray images which is used to train the CNN-based deep learning model such that the model can be used to effectively predict from the frontal-view ChestX-ray images the likelihood of COVID-19 infection. The effectiveness of such technology lies on the careful pre-processing of the dataset ChestX-ray images, careful selection of appropriate architecture for the COVID-19 classification model and fine tuning of the various parameters and model hyper-parameters. Eventually, the prediction performance of the prediction model is characterized and presented in various useful comprehensive metrics.

## METHODOLOGY

### 2.1 Dataset

The deep learning technique was based on Covid-19 chest X-ray images dataset from the GitHub repository available at https://github.com/education454/datasets.git [31,32,33,34,35]. The dataset consists of 712 images of those persons with Covid-19 (shown in Table 1, Figure 1 and Figure 2) and 1583 images of those persons that are normal (shown in Table 1). The dataset was segmented into training, validation, and test set (which about 20% of the training dataset).

Table 1: Total number of images in COVID dataset

| Category | No of images in the dataset |
|---|---|
| COVID-19 | 712 |
| Normal | 1583 |

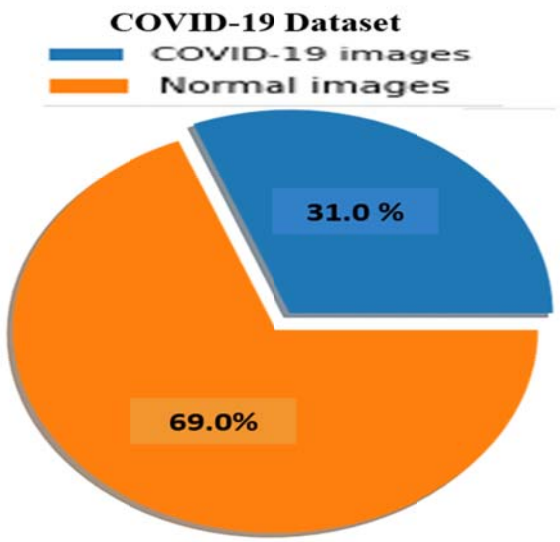


Figure 1: Pie Chart of COVID-19 dataset

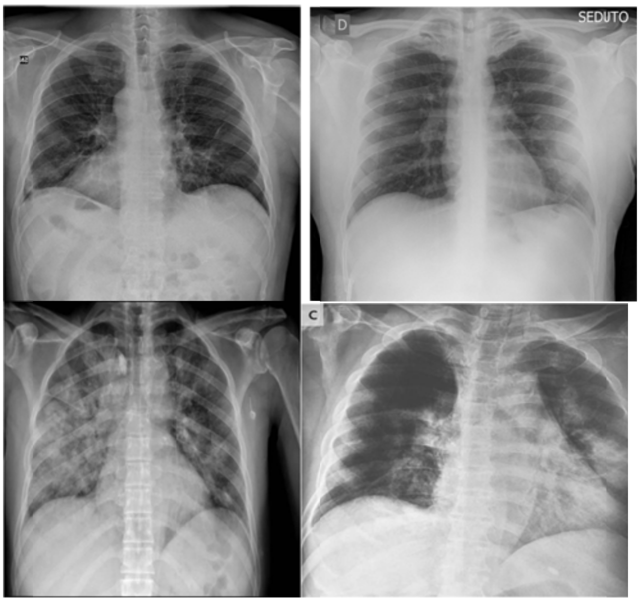

Figure 2: COVID-19 Chest X-ray Scan

### 2.2 Data Pre-processing

Overfitting problem was addressed by applying data augmentation which at the same time was used to boost the number of images that are used in the model training. Specifically, horizontal flip and zoom data augmentations strategies were employed along with the settings shown in Table 2. Furthermore, all the training, validation and testing data are normalized. The dataset images were resized to 200 x 200 with batch size of 32 and binary class mode.

**Table 2: Image augmentation settings**

| Method | Setting |
|---|---|
| Horizontal Flip | True |
| Zoom range | 0.2 |
| Rescale | 1/255 |

### 2.3 Model Architecture for COVID-19 classification model

The model architecture consists of a convolutional block with a convolutional layer that has 32 filters along with 5 stride. The padding parameter was set to *'same'* and the activation function selected was *relu*. Also, max-pooling layer was added and it has a 2 as its pool size setting. Furthermore, a dropout layer was added. The second block consists of a convolutional layer with 64 filters with a stride of 5. Padding was set to 'same' and relu activation was used. A max-pooling layer of pool size 2 and a dropout layer were added and then a flattening layer was also added. Eventually, a fully connected layer which has 256 nodes was added along with a dropout layer. The output layer is made up of a node with the sigmoid activation function. The architecture for COVID-19 CNN classification model is presented in Figure 3 and the summary is presented in Figure 4. Furthermore, Adam optimizer, binary cross-entropy loss function, and accuracy metric were used to compile the COVID-19 CNN classification model and compilation hyper-parameters are presented in Table 4.

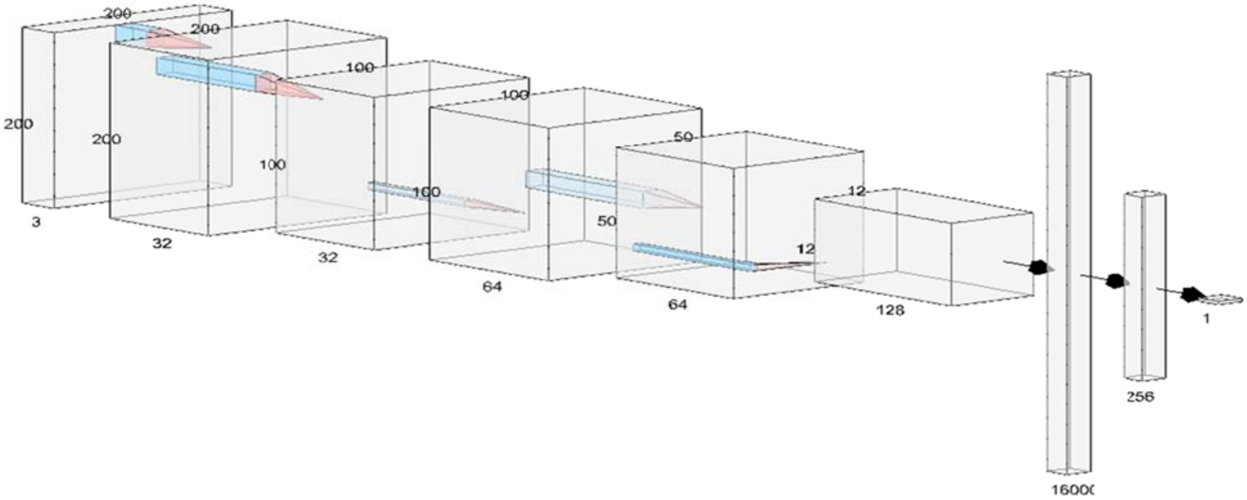


**Figure 3: CNN Model Architecture for COVID-19 classification I**

Model: "sequential_1"

| Layer (type) | Output Shape | Param # |
|---|---|---|
| conv2d_2 (Conv2D) | (None, 200, 200, 32) | 2432 |
| max_pooling2d_1 (MaxPooling2 | (None, 100, 100, 32) | 0 |
| dropout_1 (Dropout) | (None, 100, 100, 32) | 0 |
| conv2d_3 (Conv2D) | (None, 100, 100, 64) | 51264 |
| max_pooling2d_2 (MaxPooling2 | (None, 50, 50, 64) | 0 |
| dropout_2 (Dropout) | (None, 50, 50, 64) | 0 |
| flatten (Flatten) | (None, 160000) | 0 |
| dense (Dense) | (None, 256) | 40960256 |
| dropout_3 (Dropout) | (None, 256) | 0 |
| dense_1 (Dense) | (None, 1) | 257 |

Total params: 41,014,209
Trainable params: 41,014,209
Non-trainable params: 0

**Figure 4: CNN Model Summary for COVID-19 classification**

**Table 3: Hyper-parameters for COVID-19 model compilation**

| | |
|---|---|
| **Optimizer** | ***Adam*** |
| **Loss function** | ***Binary cross-entropy*** |
| **Metric** | ***Accuracy, Precision, Recall*** |
| **Learning Rate** | ***0.001*** |

## 3. RESULTS AND DISCUSSION

Google Colab Graphics processing unit (GPU) was used to train the COVID-19 model for 3 hours and 70 epochs. In the course of the training, fine-tuning of the parameters as well as the tuning of the hyper-parameters were done so as to improve the model's performance.

### 3.1 Results of the Training and Validation of COVID-19 classification Model

The COVID-19 classification model results are shown in Table 4 which shows a training loss values of 0.0565 with training accuracy values of 98.34%. Also, the training precision results is 0.9872, training recall values is 0.9891 while the validation loss value is 0.0633. Furthermore, the results shows that the validation accuracy is 98.62%, the validation precision results is 0.9921 while the validation recall value is 0.9881. In addition, the test loss results is 0.0791, the test accuracy results is 96.07% while the test precision values 0.9571, and the test recall values is 0.9842. The graphs for the different performance parameters of the COVID-19 classification model are given in Figure 5 to Figure 10. Altogether, the COVID-19 classification model accuracy was very high with a value 98.34 %. This shows that the model can effectively predict COVID-19 infection by analyzing the chest X-ray images of people suspected to have such infection.

**Table 4: The COVID-19 classification model Results**

| Final Result | COVID-19 Model |
|---|---|
| Training Accuracy | 98.34% |
| Validation Accuracy | 98.62% |
| Training Loss | 0.0565 |
| Validation Loss | 0.0633 |
| Training Precision | 0.9872 |
| Validation Precision | 0.9921 |
| Training Recall | 0.9891 |
| Validation Recall | 0.9881 |

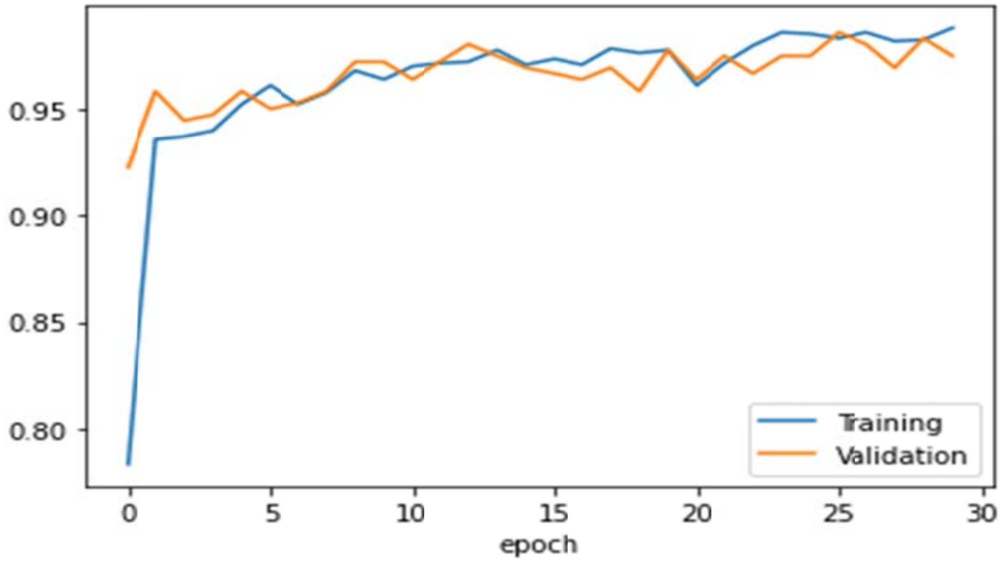


**Figure 5: Training and Validation Accuracy Results for the COVID-19 Classification Model**

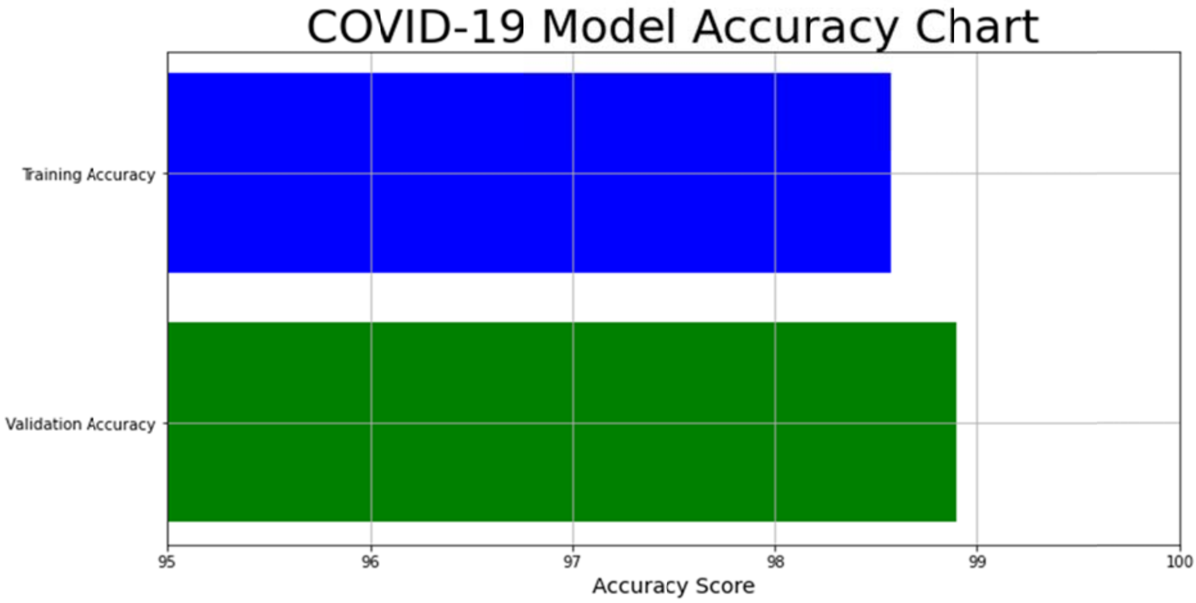


**Figure 6: COVID-19 Classification Model Accuracy Chart**

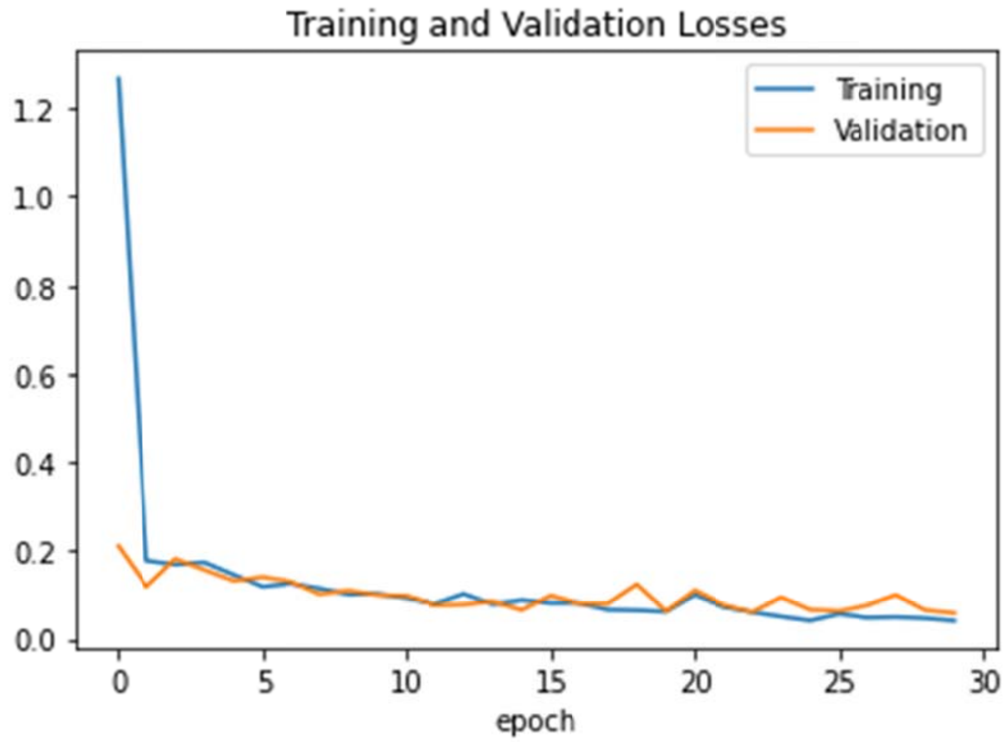


**Figure 7: Training and Validation Losses Results for the COVID-19 Classification Model**

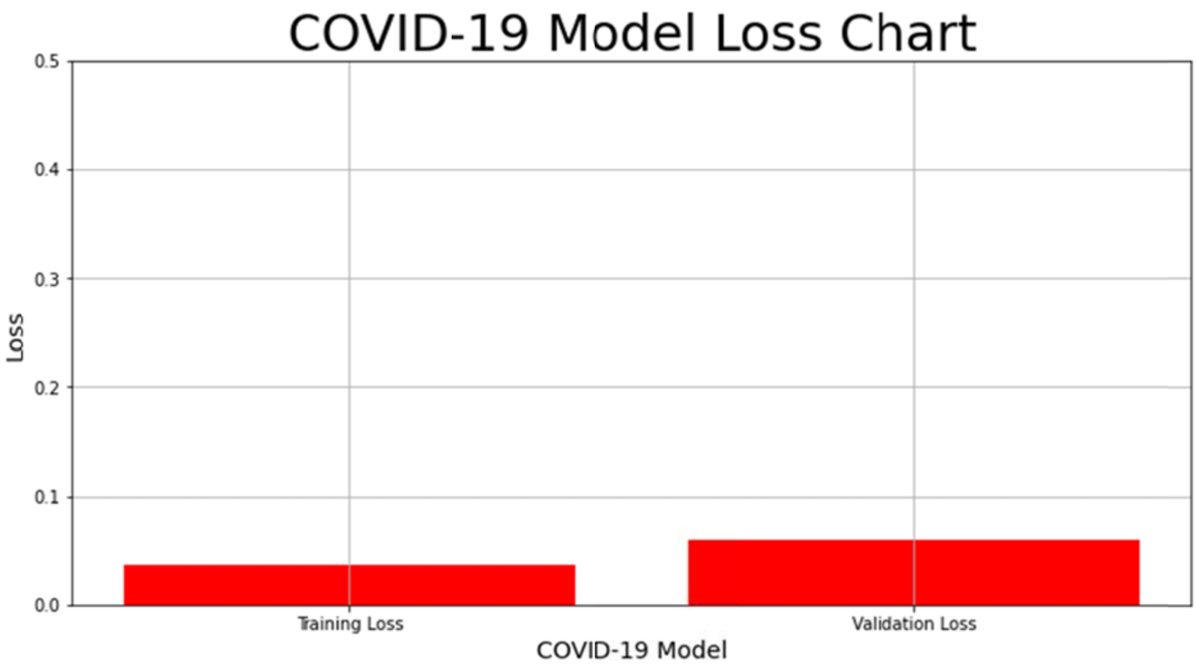


**Figure 8: COVID-19 Classification Model Loss Chart**

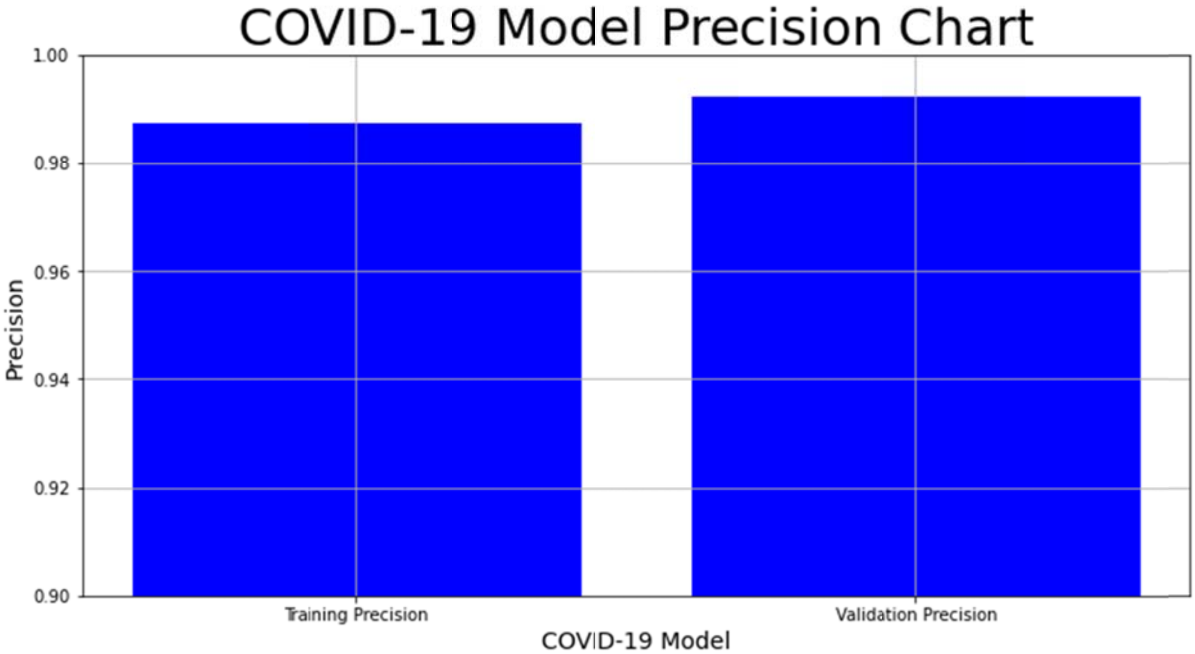


**Figure 9: COVID-19 Classification Model Precision Chart**

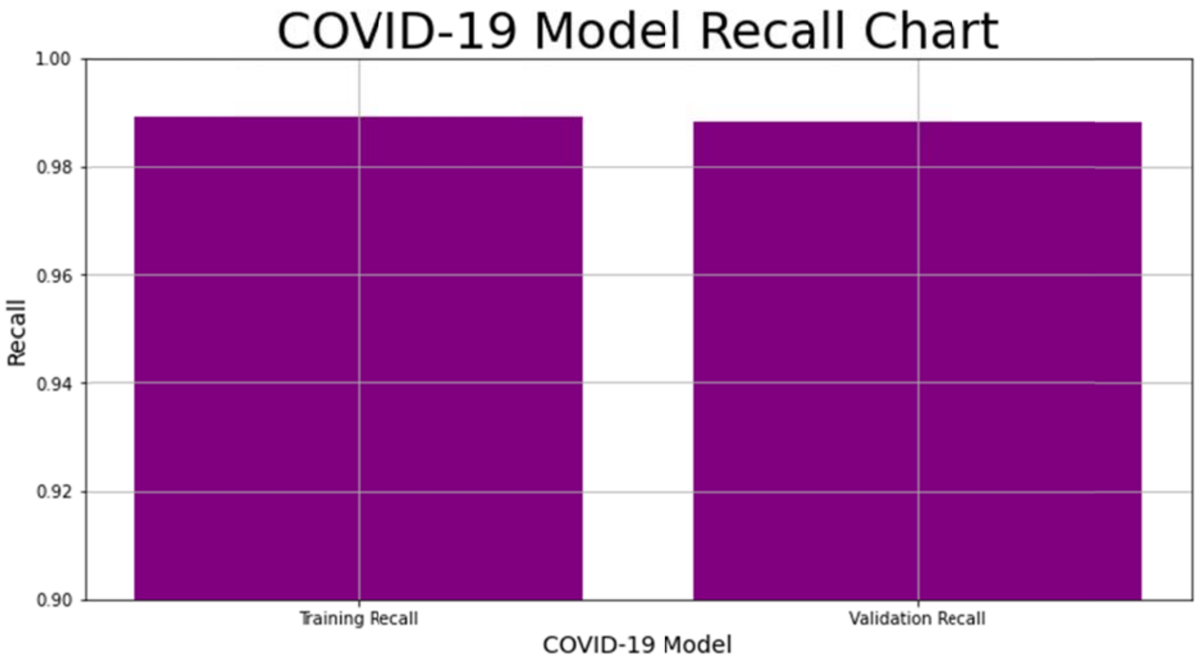


**Figure 10: COVID-19 Classification Model Recall Chart**

## 4. CONCLUSION

A convolutional neural network (CNN)-based deep learning model for predicting COVID-19 infection is presented. The study utilized chest X-ray image dataset to train, validate and test the model. The dataset images are first pre-processed to suit the application of deep learning techniques after which Google Colab Graphics processing unit (GPU) was used to train the COVID-19 model for some hours and epochs. In the course of the training, fine-tuning of the parameters as well as the tuning of the hyper-parameters were done so as to improve the model's performance. Altogether, the COVID-19 classification model accuracy was very high with a value above 98 %. This shows that the model can effectively predict COVID-19 infection by analyzing the chest X-ray images of people suspected to have such infection.

## References

1. Menachery, V. D., Yount, B. L., Debbink, K., Agnihothram, S., Gralinski, L. E., Plante, J. A., ... & Baric, R. S. (2015). A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nature medicine*, *21*(12), 1508-1513.

2. Wu, Z., Yang, L., Ren, X., Zhang, J., Yang, F., Zhang, S., & Jin, Q. (2016). ORF8-related genetic evidence for Chinese horseshoe bats as the source of human severe acute respiratory syndrome coronavirus. *The Journal of infectious diseases*, *213*(4), 579-583.
3. Lau, S. K., Feng, Y., Chen, H., Luk, H. K., Yang, W. H., Li, K. S., ... & Woo, P. C. (2015). Severe acute respiratory syndrome (SARS) coronavirus ORF8 protein is acquired from SARS-related coronavirus from greater horseshoe bats through recombination. *Journal of virology*, *89*(20), 10532-10547.
4. Muth, D., Corman, V. M., Roth, H., Binger, T., Dijkman, R., Gottula, L. T., ... & Drosten, C. (2018). Attenuation of replication by a 29 nucleotide deletion in SARS-coronavirus acquired during the early stages of human-to-human transmission. *Scientific reports*, *8*(1), 1-11.
5. Zhou, J., Chu, H., Li, C., Wong, B. H. Y., Cheng, Z. S., Poon, V. K. M., ... & Yuen, K. Y. (2014). Active replication of Middle East respiratory syndrome coronavirus and aberrant induction of inflammatory cytokines and chemokines in human macrophages: implications for pathogenesis. *The Journal of infectious diseases*, *209*(9), 1331-1342.
6. Wang, Q., Zhang, L., Kuwahara, K., Li, L., Liu, Z., Li, T., ... & Liu, G. (2016). Immunodominant SARS coronavirus epitopes in humans elicited both enhancing and neutralizing effects on infection in non-human primates. *ACS infectious diseases*, *2*(5), 361-376.
7. Wright, W. F., Townsend, J., & Johnson, E. N. (2018). SEVERE ACUTE RESPIRATORY SYNDROME CORONAVIRUS 2 (SARS-COV-2), ALSO KNOWN AS CORONAVIRUS DISEASE 2019. *Essentials of Clinical Infectious Diseases*, 443.
8. Memish, Z. A., Al-Tawfiq, J. A., Makhdoom, H. Q., Assiri, A., Alhakeem, R. F., Albarrak, A., ... & Zumla, A. (2014). Respiratory tract samples, viral load, and genome fraction yield in patients with Middle East respiratory syndrome. *The Journal of infectious diseases*, *210*(10), 1590-1594.
9. Daga, M. K., Kumar, N., Aarthi, J., Mawari, G., Garg, S., & Rohatgi, I. (2019). From SARS-CoV to coronavirus disease 2019 (COVID-19)-A brief review. *Journal of Advanced Research in Medicine (E-ISSN: 2349-7181 & P-ISSN: 2394-7047)*, *6*(4), 1-9.
10. Gupta, M., Gupta, V. K., Kaur, N., Hari, P. K., & Goyal, K. (2015). Prevalence of Anxiety and Depression in India among Medicos and Non-Medicos during Covid 19: A Survey.
11. Archibong, C. S. (2019). LACK OF ADEQUATE CITIZENS'INFORMATION AND THE CHALLENGE OF RESPONDING TO COVID 19 PANDEMIC IN NIGERIA. *JOS JOURNAL OF RELIGION AND PHILOSOPHY*, *1*(2).
12. Parfin, A., Wdowiak, K., Furtak-Niczyporuk, M., & Herda, J. (2019). An influence of social isolation on the level of physical activity as well as on well-being and mental state of people during the coronavirus COVID-19 pandemic. *Polish Journal of Public Health*, *129*(4), 127-131.
13. Begum, N. N. F. (2019). COVID-19 Pandemic. *Journal of Armed Forces Medical College, Bangladesh*, *15*(2), 125-126.
14. AYANDOKUN, A. A., & AKUBUE, U. U. (2019). PUBLIC LIBRARIES AND THE DISSEMINATION OF HEALTH INFORMATION: THE IMPERATIVES FOR COMBATING COVID-19 PANDEMIC IN NIGERIA. *EBONYI JOURNAL OF LIBRARY AND INFORMATION SCIENCE*, 384.
15. Divya, N. S., Vanishree, H. R., & Sharat Kumar, B. (2019). Impact of COVID 19 pandemic on blood transfusion services at a rural based district Hospital Blood-Bank, India. *Blood*, *310*.
16. Daga, M. K., Kumar, N., Aarthi, J., Mawari, G., Garg, S., & Rohatgi, I. (2019). From SARS-CoV to coronavirus disease 2019 (COVID-19)-A brief review. *Journal of Advanced Research in Medicine (E-ISSN: 2349-7181 & P-ISSN: 2394-7047)*, *6*(4), 1-9.
17. Qin, E., Bi, J., Zhao, M., Wang, Y., Guo, T., Yan, T., ... & Zhong, Y. (2015). Clinical features of patients with Ebola virus disease in Sierra Leone. *Clinical infectious diseases*, *61*(4), 491-495.
18. Entesar, A., Almajdoub, O., & Bshaena, A. (2015). SARS-CoV-2 pandemic and its management in AL-Zawiyah city, Libya. 2021; 15 (2).
19. Hossain, K. A., Hossain, M. A., & Haque, M. O. (2019). Scholars Journal of Medical Case Reports.
20. White, D. (2019). Preparing for COVID-19: The lessons from SARS 2003 in Canada. *Asia Pacific Family Medicine*, *18*(1).
21. Danno, K., Cognet-Dementhon, B., Thevenard, G., Duru, G., Allaert, F. A., & Bordet, M. F. (2014). Management of the early symptoms of influenza-like illnesses and ear, nose and throat (ENT) disorders by pharmacists. *Homeopathy*, *103*(04), 239-249.
22. Parekh, U., Gupta, S., & Johnson, A. (2019). Review of clinical syndromes in covid-19 in global perspective: A step towards infection prevention control in forensic practices. *Journal of Forensic Medicine and Toxicology*, *36*(2), 58-63.
23. Fairley, J. K., Kozarsky, P. E., Kraft, C. S., Guarner, J., Steinberg, J. P., Anderson, E., ... & Wu, H. M. (2016, January). Ebola or not? Evaluating the ill traveler from Ebola-affected countries in West Africa. In *Open Forum Infectious Diseases* (Vol. 3, No. 1, p. ofw005). Oxford University Press.
24. Smith, K. P., Kang, A. D., & Kirby, J. E. (2018). Automated interpretation of blood culture gram stains by use of a deep convolutional neural network. *Journal of Clinical Microbiology*, *56*(3), e01521-17.
25. Winkler, J. K., Fink, C., Toberer, F., Enk, A., Deinlein, T., Hofmann-Wellenhof, R., ... & Haenssle, H. A. (2019). Association between

surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition. *JAMA dermatology*, *155*(10), 1135-1141.

26. Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, *284*(2), 574-582.
27. Sirazitdinov, I., Kholiavchenko, M., Mustafaev, T., Yixuan, Y., Kuleev, R., & Ibragimov, B. (2019). Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database. *Computers & electrical engineering*, *78*, 388-399.
28. Soffer, S., Ben-Cohen, A., Shimon, O., Amitai, M. M., Greenspan, H., & Klang, E. (2019). Convolutional neural networks for radiologic images: a radiologist's guide. *Radiology*, *290*(3), 590-606.
29. Park, A., Chute, C., Rajpurkar, P., Lou, J., Ball, R. L., Shpanskaya, K., ... & Yeom, K. W. (2019). Deep learning–assisted diagnosis of cerebral aneurysms using the HeadXNet model. *JAMA network open*, *2*(6), e195600-e195600.
30. Lu, M. T., Ivanov, A., Mayrhofer, T., Hosny, A., Aerts, H. J., & Hoffmann, U. (2019). Deep learning to assess long-term mortality from chest radiographs. *JAMA network open*, *2*(7), e197416-e197416.
31. Sirazitdinov, I., Kholiavchenko, M., Mustafaev, T., Yixuan, Y., Kuleev, R., & Ibragimov, B. (2019). Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database. *Computers & electrical engineering*, *78*, 388-399.
32. Antin, B., Kravitz, J., & Martayan, E. (2017). Detecting pneumonia in chest X-Rays with supervised learning. *Semanticscholar. org*.
33. Islam, S. R., Maity, S. P., Ray, A. K., & Mandal, M. (2019, May). Automatic detection of pneumonia on compressed sensing images using deep learning. In *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)* (pp. 1-4). IEEE.
34. Wibisono, A., Adibah, J., Priatmadji, F. S., Viderisa, N. Z., Husna, A., & Mursanto, P. (2019, July). Segmentation-based knowledge extraction from chest X-ray images. In *2019 4th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS)* (pp. 225-230). IEEE.
35. Jaiswal, A. K., Tiwari, P., Kumar, S., Gupta, D., Khanna, A., & Rodrigues, J. J. (2019). Identifying pneumonia in chest X-rays: a deep learning approach. *Measurement*, *145*, 511-518.