Automatic Prognosis of COVID-19 from CT Scan using Super-convergence CNN Algorithm

Steve A. ADESHINA
Department of Computer Engineering
Nile University, Abuja, Nigeria.
steve.adeshina@nileuniversity.edu.ng

Adeyinka P. ADEDIGBA
Department of Mechatronics Engineering,
Federal University of Technology, Minna, Nigeria.
adeyinka.adedigba@futminna.edu.ng

Abstract—Due to the high incident rate of COVID-19, the number of suspected patients needing diagnosis presents overwhelming pressure on hospital and health management systems such that the disease outbreak elapsed into a global pandemic. More so, the infected patients present a higher risk of being infected to the health workers because once a patient is positive of the virus, the progress of recovery or deterioration needs to be monitored by medical experts and other health workers, which eventually exposes them to the infection. In this paper, we present an automatic prognosis of COVID-19 from a CT scan using deep CNN. The models were trained using a super-convergence discriminative fine-tuning algorithm, which uses a layer-specific learning rate to fine-tune a deep CNN model; this learning rate is increased or decreased per iteration to avoid the saddle-point problem and achieve the best performance within few training epochs. The best performance results were obtained as 98.57% accuracy, 98.59% precision and 98.55% recall rate.

Index Terms—COVID-19, CT Scan, deep CNN, Hyperparameter optimization, medical image, radiology, super-convergence.

I. INTRODUCTION

The novel SARS-CoV-2 virus, which causes COVID-19, uses widely expressed Angiotensin-converting enzyme-2 (ACE-2) receptor to enter human cells (usually pneumocytes, respiratory tracts, renal and epithelium) through nasal and oral pharyngeal mucosa to viremia [1]. The ACE-2 is an enzyme attached to the membranes of cells located in the lungs, arteries, heart, kidney, and intestines – the vital organs. Further, because the virus enters through nasal and oral tracts, studies show that the respiratory tracts are the primary site of infection and disease morbidity [2], [3]. This explains why COVID-19 patients show difficulty in breathing, and many die within a short period.

Although the virus has a high incident rate and spread rapidly, there is a broad spectrum of disease severity ranging from mild (which accounts for 80%), severe case of about 15% and a critical case of about 5%. Globally, the reported number of COVID-19 cases is over 147.812 million, with about 3.12 million death, representing 2.11% of reported cases. In Nigeria, the total number of reported cases is 164,912, with 2063 reported death representing 1.25%. It could be inferred that although the disease has a high incident rate, the mortality rate is low.

Due to its high incident rate, the number of suspected patients needing diagnosis presents overwhelming pressure on hospital and health management systems such that the disease outbreak elapses to a global pandemic. More so, the infected patients present a higher risk of being infected to the health workers because once a patient is positive of the virus, the progress of recovery or deterioration needs to be monitored by medical experts and other health workers, which eventually exposes them to the infection. Within the first three months, over 33,000 health workers were infected with the virus in China, 5000 in Italy with about 60 death [4]. Currently, over 300,000 health workers have been infected with COVID-19 globally, with over 7,000 death recorded in 79 countries [5]. Hence, the development of a CAD system for the detection of COVID-19 is not sufficient, but a method of automatically monitoring the recovery progress of infected patients is needed to reduce the risk of health workers.
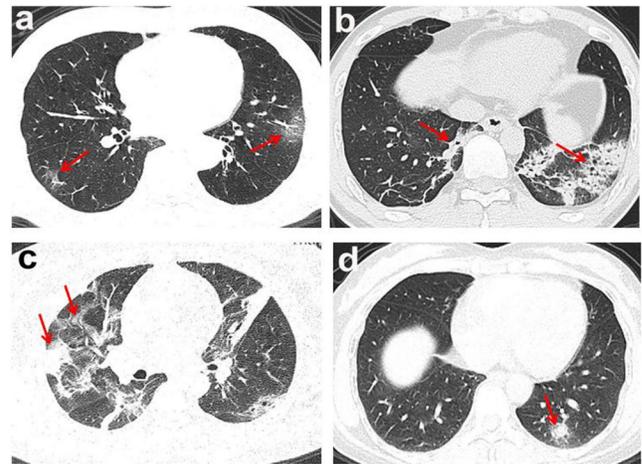
Fig. 1. CT Scan of an infected patient showing morphological observations. Red arrows indicating abnormalities (a) Ground Glass Opacities; (b) consolidation; (c) consolidation with GGO; (d) solid nodules [3]

In Fleischner statement [6], the role of medical imaging in COVID-19 was extensively discuss. This consensus agreement is based on the expert opinion of a panel of fifteen thoracic radiologists, ten pulmonologists and one pathologist, in addition to experts in emergency medicine, infection control and laboratory medicine drawn from nine advanced countries. The CT scan was identified to be more sensitive in the statement, thereby recommended for monitoring disease progression.

Morphological features of the disease infection observable in CT scan includes the following: ground-glass opacity (GGO), consolidation, hazy paving, and solid nodules see Fig. 1.

Ground-glass opacity (GGO) is a sign of decreased air content in the lungs without totally obliterating the alveoli. This could be due to the lungs air spaces partially filled with fluid, the walls of the alveoli becoming thickening and the space between the lungs parenchyma becoming thickening. GGO is an opacity that does not obscure the underlying vessels; it appears as grey opacity on CT scan images.

That CT scan can be used for monitoring disease progression has been demonstrated in literature [3]. A typical observation by [2] shown in Fig. 2 is a 35 years old woman. The scan in Fig 2A was obtained on the first day. The scan shows multiple GGO on the lower lobe of the right lung. The scan in Fig 2B was obtained on the fifth day; the scan shows an increased consolidation. Fig 2C was obtained on the eleventh day; it shows multiple consolidations with a similar extent as obtained on the fifth day. Lastly, the scan in Fig 2D was obtained on the fifteenth day; the scan shows a smaller extent than the eleventh day. Hence, it is clear that a CT scan can monitor the progress of COVID-19 infected patients. In this paper, we explore the possibility of automating this process without exposing the health worker to further risk.

The rest of this paper is organized as follows: the review of related work is presented in Section II, the methodology adopted is presented in Section III, while the result and discussion of the result are presented in Section IV.

II. REVIEW OF RELATED WORK

Most researchers focused on the classification of CT scan images for the diagnosis of the novel coronavirus. Some of these research works are reviewed herewith.

Authors gathered a comprehensive dataset of CT scan images and chest x-rays of normal and infected patients in [7], then a comparative result of a model trained from scratch and a fine-tuned model was presented. The fine-tuned model of AlexNet achieves 98% accuracy while a model trained from scratch is 94% accurate. Although the result is appreciably good, the dataset on which these models were trained is too scanty, leading to model overfitting. Furthermore, the two models show a wide difference between sensitivity (72%) and specificity (100%), which indicated the case of overfitting, this kind of model should be avoided in medical applications. On the other hand, [8] gathered 104,009 CT scan images of 1,489 patients (unfortunately, the dataset is not publicly available). They build a new model called COVIDNet-CT, a 90 layer deep model using the generative-synthesis technique. The result obtained shows 99.1% accuracy, a comparatively higher result than ResNet50 architecture, which obtained 98.7% on the same dataset. While this result is impressive, such a voluminous dataset and the computational power to train such deep learning models are not evenly distributed, often unavailable to researchers in low- and middle-income countries. For instance, without such a large dataset, [9] trained a 10-layer CNN model

from scratch and obtained 82.1% accuracy on 738 CT scan images.

However, better results were reported when models are trained using the transfer learning approach. For instance, an improved result was obtained in [10] where 828 CT scans were used to fine-tune EfficientNetB0 and accuracy of 96.2% was obtained. On the other hand, [11] trained eight states of the art deep learning models using a fine-tuning approach; they obtained the highest classification accuracy of 89.0% with DenseNet169. The poor result shows that the desired result cannot be obtained when transfer learning is not skillfully done.

Therefore, in this paper, we present a data and computational efficient fine-tuning approach to achieve high accuracy without overfitting the small dataset. The design of this algorithm is presented in Section III.

III. METHODOLOGY

In this section, the methodology adopted is discussed. We present the dataset used in training the deep learning model, followed by the data augmentation method. Then the superconvergence discriminative fine-tuning algorithm is then presented. The algorithm uses layer-specific learning rates to fine-tune a deep CNN; these learning rates were dynamically increased or reduced to avoid saddle-point problem. Lastly, the deep learning architecture used is presented.

A. Dataset

The dataset for this work was collected from Hospitals in Sao Paulo, Brazil and made publicly available by [12]. The dataset contains 2484 CT scan images, of which 1252 were positive of COVID-19 while 1230 were negative. Although the data is small for training a deep learning model, data augmentation is performed to obtain a more comprehensive distribution and curate the dataset from defective scans. The data augmentation is presented in the following subsection.

B. Data Augmentation

To obtain a good generalization with deep learning models, they must be trained on a vast dataset in hundreds of millions. Meanwhile, the medical dataset is very scanty and hard to get due to privacy issues. The COVID-19 dataset, like every medical dataset, is very scanty, available in the order of tens of hundreds. Therefore, data augmentation is an essential step in the design of a good model.

Unlike general computer vision task, medical images must be carefully augmented as augmented images may contribute nuisance images which confuse the model and prevent it from learning the key desired features from the image [13]. For instance, image flipping and mirroring are well-known augmentation techniques in computer vision; however, medical images contain anatomical features with chirality, whereas these augmentation techniques do not preserve this feature. Hence, these augmentation techniques and similar techniques that do not preserve anatomical features of interest must be avoided.
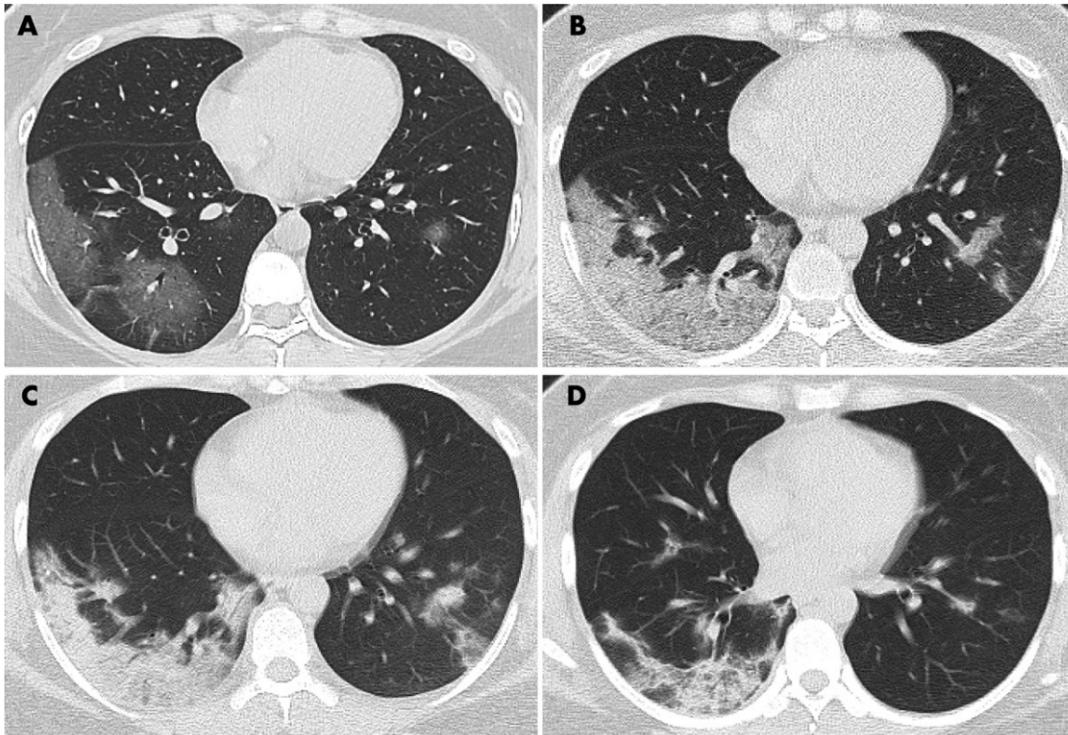
Fig. 2. CT Scan of a 35yrs old patient infected with COVID-19

More so, in our previous work [13], we showed that augmentation could be done either at run-time or at the pre-processing stage. Data augmentation at run-time is commonly implemented in popular deep learning frameworks such as PyTorch, TensorFlow, Theano, etc. Data augmentation at the pre-processing stage was presented in [13]; the result obtained shows that this is the right way to go, especially for medical images. We recommended that data augmentation for medical images be done at the pre-processing stage so that the augmented images can be curated and inappropriate images are removed before training the model with the dataset.

Hence, the parameters for the augmentation pre-processes are presented in Table I.

TABLE I
TABLE OF PARAMETERS FOR THE AUGMENTATION PRE-PROCESSES

Data Augmentation	Parameter	Value(s)
Rotation	Rotation angle	$\pm$ [5, 10, 15]
Gaussian Blurring	Kernel size	3
Random Zoom	Scale	1.3
Random Lighting	Intensity	1.4

C. Discriminative Fine-tuning and Mixed-precision training

To obtain a good generalization, deep learning models must be trained on a large, well-labelled training dataset, using a high specification computer with Graphics Processing Units for a very long time. Hence, the state of the art models in computer vision were trained on ImageNet Large Scale Visual Recognition Challenge (ILSVRC) data, consisting of hundreds of millions of well-labelled training data. When this kind of huge training data is not available or computational power is limited, the usual practice is to perform transfer learning.

Transfer learning involves retraining a previously trained model (base model) on a new dataset from the current problem (target) domain. Depending on the similarity of the target domain and the domain where the base model is trained (usually called source domain), transfer learning could be feature extraction or fine-tuning. Feature extraction is usually done when the dataset in the target domain is scanty and is similar to the source domain. This is done by replacing the last fully connected layer of the base model architecture with a new layer corresponding to the target output, initializing the other layers with the weights from the previous training scenario and retraining only the newly added layer. Fine-tuning, on the other hand, is done either when the dataset is scanty or when the problem domains are different. This is done by replacing the last layer of the base model with a new layer corresponding to the target output, initializing the other layers with weights from the previous training scenario and training the entire network all over again.

Because fine-tuning involves training the entire network all over again, the performance of the model on the current problem depends on how well the training is conducted besides the high demand for computational power. One of the main challenges facing fine-tuning is called overfitting, where a model performed very well on training dataset but poorly on test or validation data; such model performed woefully when deployed and should be avoided, especially in medi-

cal applications. Methods of overcoming overfitting include training with an extensive training set, data augmentation and regularization. What we mean by an extensive training set is quite explicit, and data augmentation has been introduced in subsection III-B. Regularization refers to techniques that make slight modifications to the learning algorithm such that the model generalizes better on the unseen dataset. Regularization can be achieved by optimal selection of hyperparameters such as learning rates, weight decay, batch size, and dropout. In our previous work [14] we introduced discriminative fine-tuning, where we assigned different learning rate and momentum to each layer of the network. The idea is that we found that each layer of the network is learning different features and, as such, has different objectives. Hence, it would be good to tune each layer with different learning rates and momentum to facilitate the learning process without getting stuck in the local minimum or saddle point. This idea has been found to produce a better result and a good generalization. The complete algorithm is listed in Algorithm 1.

D. CNN Architecture

To validate the performance of the algorithm, we trained two deep state of the art architectures – Densenet [15] and Resnet models [16]. The justification is presented herewith.

Resnet 152 is a 152-layer deep network that surpasses human-level performance in the 2015 ILSVRC with a 3.57% top-5 error rate; since then, it has become a state of the art model in the deep learning community. Deeper networks like this have been shown to perform substantially better than shallower counterparts. However, deeper networks are more prone to vanishing gradient problems, making them hard to train [16]. This problem was addressed by the implementation of *Residual Block* in ResNet. The Residual block modelled in equation (1) creates a connection between the output of a convolutional layer and the earlier input to the layer using identity mapping [16]. Thus, the activation of a Residual block is given as:

$$\mathbf{a}_l = \mathcal{H}(\mathbf{a}_{l-1}) + \mathbf{a}_{l-1}, \tag{1}$$

where $\mathbf{a}_l$ is the activation of layer l, $\mathcal{H}(\cdot)$ is a nonlinear convolutional transformation of the layer and $\mathbf{a}_{l-1}$ is the activation of previous layer $l-1$. The skip connection of (1) enables more layers to be stacked on each other resulting in a remarkably deep network.

DenseNet 169 [15] is a 169-layer network with 14 million parameters, which can easily overfit on small data. This model is deeper; hence, it achieves higher performance than the ResNet152 on ImageNet dataset due to its *dense block*. The dense block implements a connection that allows a layer to be connected to all layers before it within the network [15]. That is, layer l receives feature activations from all its preceding $l-1$ layers as follows:

$$\mathbf{a}_l = \mathcal{T}([\mathbf{a_0}, \mathbf{a_1}, \dots \mathbf{a_{(l-1)}}]), \tag{2}$$

Algorithm 1 Discriminative Fine-tuning Algorithm

1: **procedure** DFT
2:     **Input:** (α_{min} : minimum learning rate,
3:              α_{max} : maximum learning rate,
4:              m_{min}: minimum momentum,
5:              m_{max}: maximum momentum,
6:              $\mathcal{N}$: size of dataset, batch size)
7:     **Output:** (θ: Network parameters)
8:
9:     $t \leftarrow \frac{\mathcal{N}}{batch\ size}$
10:    $pct \leftarrow$ random number between 0.5 and 1
11:    *// pct determines how rapidly the learning rate increases or reduces*
12:
13: **while** $t \leq pct \times t$ **do**:
14:    $\alpha_t \leftarrow \alpha_{min} + \left(\frac{\alpha_{max}-\alpha_{min}}{\alpha_{max}}\right)t$ *// Increase learning rate for this iteration*
15:    $m_t \leftarrow m_{min} + \left(\frac{m_{max}-m_{min}}{m_{min}}\right)t$ *// Increase momentum for this iteration*
16:    **for** l in each layer **do**:
17:       $\alpha_t^l \leftarrow \alpha_{min} + \left(\frac{\alpha_t-\alpha_{min}}{\alpha_{min}}\right)l$ *// Increasing the learning rate per layer*
18:       $m_t^l \leftarrow m_{min} + \left(\frac{m_t-m_{min}}{m_{min}}\right)l$ *// Increasing the momentum per layer*
19:       $v_t^l \leftarrow m_t^l v_t^l - \alpha_t^l \frac{dJ(\theta_t^l)}{\theta_t^l}$
20:       $\theta_t^l \leftarrow \theta_t^l + v_t^l$ *// Update the layer parameters*
21:    end **for**
22:    $t \leftarrow t + 1$
23: end **while**
24:
25: **while** $pct \times t < t < t_{max}$ **do**:
26:    $\alpha_t \leftarrow \alpha_{max} - \left(\frac{\alpha_{max}-\alpha_{min}}{\alpha_{max}}\right)t$ *// Reduce learning rate for this iteration*
27:    $m_t \leftarrow m_{max} - \left(\frac{m_{max}-m_{min}}{m_{min}}\right)t$ *// Reduce momentum for this iteration*
28:    **for** l in each layer **do**:
29:       $\alpha_t^l \leftarrow \alpha_{min} + \left(\frac{\alpha_t-\alpha_{min}}{\alpha_{min}}\right)l$
30:       $m_t^l \leftarrow m_{min} + \left(\frac{m_t-m_{min}}{m_{min}}\right)l$
31:       $v_t^l \leftarrow m_t^l v_t^l - \alpha_t^l \frac{dJ(\theta_t^l)}{\theta_t^l}$
32:       $\theta_t^l \leftarrow \theta_t^l + v_t^l$
33:    end **for**
34:    $t \leftarrow t + 1$
35: end **while**

---

where $\mathbf{a}$ is the activation of the lth layer, $[a_0, a_1, \dots a_{(l-1)}]$ is a concatenation of all the previous layer activations, which can be seen as a form of collective information gathered by the network up to that layer $l-1$. $\mathcal{T}(\cdot)$ is a nonlinear transformation function that maps the concatenated activation to the activation of layer l.

E. Overview of Experiment

A hp G8 server with 128GB RAM and 3.5GHz Intel Xenon processor was the computer hardware for the experiments

described herewith. The data augmentation experiment was carried out using Python OpenCV library as follows: the data was divided into training and validation set using ration 70:30; then the data augmentation procedure was carried out on each set, saving each new images to disk; the new images were then curated, and invalid images were deleted.

The deep learning models were trained using PyTorch deep learning framework. The inputs to the Discriminative fine-tuning (see Algorithm 1) are mini batch-size, minimum and maximum momentum (m_{min} and m_{max}); and minimum and maximum learning rate (α_{min} and α_{max}). A batch size of 64 was used for both models; the minimum and maximum momentum for ResNet was selected as 0.8 and 0.99, respectively, while 0.79 and 0.9 were chosen for DenseNet, respectively. The learning rates α_{min} and α_{max} were carefully selected because the values could greatly slow down or hasten the training process. This learning rate choice was accomplished by running a single epoch trial experiment on each model using different learning rates and observing how the loss function increases or decreases during this epoch. The best learning rate is selected within the range where the slope of the loss function reduces sharply, which indicates a large derivative.

Throughout the experiment, Adam optimizer was used for the optimization of the backpropagation with a constant L2 norm weight-decay of 0.01. The results of a single epoch and the training experiment is presented in the next section.
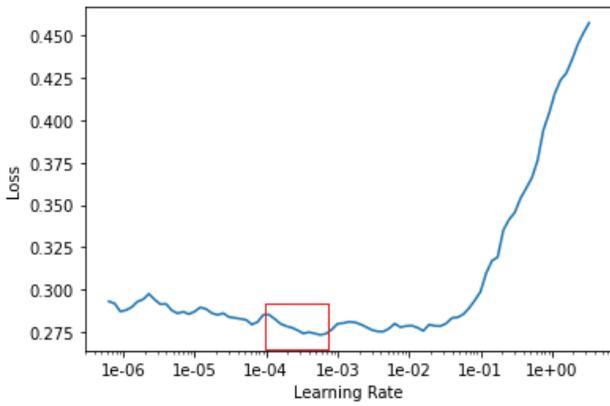
Fig. 3. Finding the optimum learning rate that best optimises the loss function. The Graph shows the variation of training loss with learning rate.

IV. RESULTS AND DISCUSSION OF RESULTS

The result of the experiment is presented in this section. We give more consideration to the performance of our model in terms of precision, recall, sensitivity and specificity. This performance was compared with those reported in the literature, especially on the same dataset as ours.

The learning rate choice for the DFT algorithm was accomplished by running a single epoch trial experiment using different learning rates and observing how the loss function increases or decreases during this epoch. Fig. 3 presents the result of this trial experiment obtained for ResNet; a similar

graph is obtained for DenseNet. The learning rate selected is within the range where the slope of the loss function reduces sharply. From Fig. 3, this range is taken from $1e^{-4}$ to $1e^{-3}$; hence α_{max} is $1e^{-3}$ while α_{min} is $1e^{-4}$. A similar experiment was conducted for DenseNet and the α_{max} was selected to be $1e^{-2}$, α_{min} was $1e^{-4}$.

The performance of ResNet and DenseNet trained using DFT fine-tuning approach is presented in Table II. In addition, the confusion matrix of the two model is displayed in Fig. 4 and Fig. 5 respectively. From the figures, it can be seen that both models show a similar trend because their precision and recall is similar. Hence, the sensitivity and specificity of the models are 98.81% and 98.73%, respectively.

TABLE II
VALIDATION RESULTS OF DISCRIMINATIVE FINE-TUNING.

	Accuracy (%)	Precision (%)	Recall (%)
DenseNet	98.57	98.59	98.55
ResNet	98.56	98.57	98.56

TABLE III
PERFORMANCE COMPARISON OF OUR RESULT WITH THOSE REPORTED IN THE LITERATURE

Reference	Accuracy (%)	Precision (%)	Recall (%)
[12]	97.38	99.16	95.53
[17]	95.60	99.00	95.00
[18]	98.70	99.67	84.17
[19]	90.83	95.75	85.89
[20]	94.49	96.55	93.52
Our Model	**98.57**	**98.59**	**98.55**
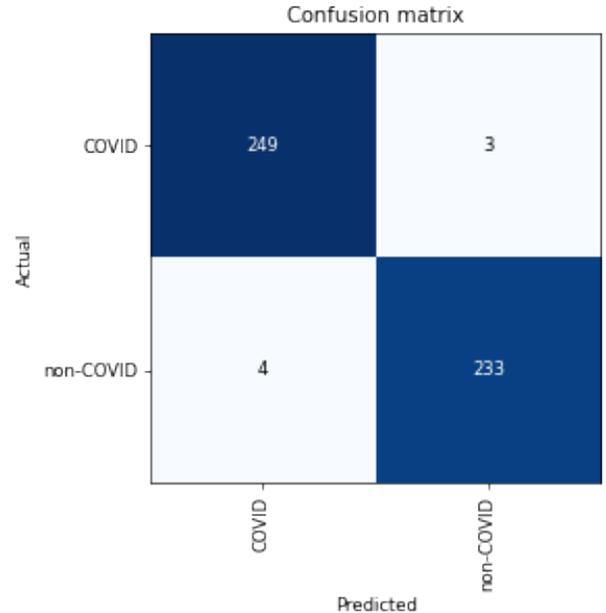
Fig. 4. Confusion Matrix of DenseNet model on Validation Data.

Lastly, we compare our result with those obtained in the literature; previous works are done on the dataset we used in
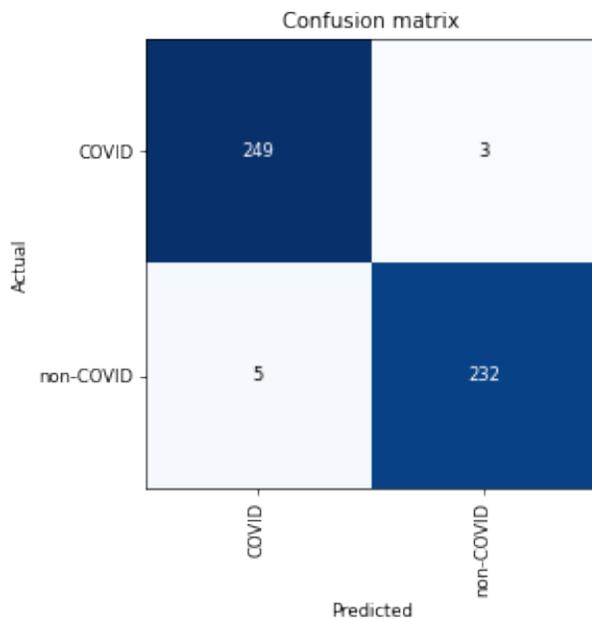
Fig. 5. Confusion Matrix of ResNet model on Validation Data.

this work. The result is presented in Table III. It could be noted that the models reported in the literature have a low recall rate compared to precision. The recall is the ability of the model to identify all relevant positive cases within a dataset, although most of these model shows near-perfect precision, low recall rate in the case of medical diagnosis during a pandemic would not be acceptable. Compared with these result, our model and approach achieve superiority.

V. CONCLUSION

The automatic prognosis of COVID-19 from CT scan images using deep convolution neural networks is presented in this paper. The models were trained using discriminative fine-tuning, which uses a layer-specific learning rate to fine-tune a deep CNN model; these learning rates were dynamically increased or decreased per iteration to avoid the saddle-point problem and achieve the best performance within few training epochs. The best performance results were obtained as 98.57% accuracy, 98.59% precision and 98.55% recall rate.

REFERENCES

[1] H. P. Jia, D. C. Look, L. Shi, M. Hickey, L. Pewe, J. Netland, M. Farzan, C. Wohlford-Lenane, S. Perlman, and P. B. McCray, "Ace2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia," *Journal of virology*, vol. 79, no. 23, pp. 14 614–14 621, 2005.

[2] Y. Wang, C. Dong, Y. Hu, C. Li, Q. Ren, X. Zhang, H. Shi, and M. Zhou, "Temporal Changes of CT Findings in 90 Patients with COVID-19 Pneumonia: A Longitudinal Study," *Radiology*, vol. 296, no. 2, pp. E55–E64, 2020.

[3] X. Li, W. Zeng, X. Li, H. Chen, L. Shi, X. Li, H. Xiang, Y. Cao, H. Chen, C. Liu, and J. Wang, "CT imaging changes of corona virus disease 2019(COVID-19): A multi-center study in Southwest China," *Journal of Translational Medicine*, vol. 18, no. 1, pp. 4–11, 2020. [Online]. Available: https://doi.org/10.1186/s12967-020-02324-w

[4] Z. F. Udwadia and R. S. Raju, "How to protect the protectors: 10 lessons to learn for doctors fighting the covid-19 coronavirus," *Medical Journal, Armed Forces India*, 2020.

[5] A. F. Goddard and M. Patel, "The changing face of medical professionalism and the impact of covid-19," *The Lancet*, vol. 397, no. 10278, pp. 950–952, 2021.

[6] G. D. Rubin, C. J. Ryerson, L. B. Haramati, N. Sverzellati, J. P. Kanne, S. Raoof, N. W. Schluger, A. Volpi, J. J. Yim, I. B. Martin, D. J. Anderson, C. Kong, T. Altes, A. Bush, S. R. Desai, O. Goldin, J. Mo Goo, M. Humbert, Y. Inoue, H. U. Kauczor, F. Luo, P. J. Mazzone, M. Prokop, M. Remy-Jardin, L. Richeldi, C. M. Schaefer-Prokop, N. Tomiyama, A. U. Wells, and A. N. Leung, "The role of chest imaging in patient management during the covid-19 pandemic: A multinational consensus statement from the fleischner society," *Radiology*, vol. 296, no. 1, pp. 172–180, 2020.

[7] H. S. Maghdid, A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq, and M. K. Khan, "Diagnosing COVID-19 pneumonia from x-ray and ct images using deep learning and transfer learning algorithms," pp. 1–8, 2020.

[8] H. Gunraj, L. Wang, and A. Wong, "COVIDNet-CT: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases From Chest CT Images," *Frontiers in Medicine*, vol. 7, no. December, pp. 1–11, 2020.

[9] V. Shah, R. Keniya, A. Shridharani, M. Punjabi, J. Shah, and N. Mehendale, "Diagnosis of COVID-19 using CT scan images and deep learning techniques," *Emergency Radiology*, 2021.

[10] L. A. Privor-dumm, G. A. Poland, J. Barratt, D. N. Durrheim, M. Deloria, P. Vasudevan, M. Jit, P. E. Bonvehí, and P. Bonanni, "Automatic distinction between COVID-19 and common pneumonia using multi-scale convolutional neural network on chest CT scans," *Chaos, Solitons and Fractals*, vol. 140, no. June, 2020.

[11] C. M. Do and L. Vu, "An approach for recognizing COVID-19 cases using convolutional neural network applied to CT scan images," in *Applications of Digital Image Processing XLIII*, no. August. nternational Society for Optics and Photonics, 2020, p. 108.

[12] P. Angelov and E. Almeida Soares, "Sars-cov-2 ct-scan dataset: A large dataset of real patients ct scans for sars-cov-2 identification," *medRxiv*, 2020.

[13] A. P. Adedigba, S. A. Adeshina, and A. M. Aibinu, "Deep learning-based mammogram classification using small dataset," in *2019 15th International Conference on Electronics, Computer and Computation (ICECCO)*. IEEE, 2019, pp. 1–6.

[14] A. P. Adedigba, S. A. Adeshina, O. E. Aina, and A. M. Aibinu, "Optimal hyperparameter selection of deep learning models for covid-19 chest x-ray classification," *Intelligence-Based Medicine*, p. 100034, 2021.

[15] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[16] 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.

[17] H. Panwar, P. Gupta, M. K. Siddiqui, R. Morales-Menendez, P. Bhardwaj, and V. Singh, "A deep learning and grad-cam based color visualization approach for fast detection of covid-19 cases using chest x-ray and ct-scan images," *Chaos, Solitons & Fractals*, vol. 140, p. 110190, 2020.

[18] M. Rahimzadeh, A. Attar, and S. M. Sakhaei, "A fully automated deep learning-based network for detecting covid-19 from a new and large lung ct scan dataset," *Biomedical Signal Processing and Control*, vol. 68, p. 102588, 2021.

[19] Z. Wang, Q. Liu, and Q. Dou, "Contrastive cross-site learning with redesigned net for covid-19 ct classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2806–2813, 2020.

[20] D. Konar, B. K. Panigrahi, S. Bhattacharyya, N. Dey, and R. Jiang, "Auto-diagnosis of covid-19 using lung ct images with semi-supervised shallow learning network," *IEEE Access*, vol. 9, pp. 28 716–28 728, 2021.