ACCSCIENCE PUBLISHING

Artificial Intelligence in Health

ORIGINAL RESEARCH ARTICLE

Deep learning on chest X-ray and computed tomography scans for detection of COVID-19 as a part of a network-centric digital health stack for future pandemics

Ajay Kumar Gogineni[1], Madapathi Hitesh[1], Prashant Kumar Jha[2], Soumya Suvashish Sen[3], Shreeja Das[2], and Kisor Kumar Sahu[2,4,5]*

[1]School of Electrical Sciences, Indian Institute of Technology, Bhubaneswar, Odisha, India

Abstract

Developing a reliable rapid screening protocol for highly infectious diseases like COVID-19 is of paramount interest since it facilitates the isolation of infected patients from the rest of the population. Reverse-transcription polymerase chain reaction (RT-PCR) test is presently the most widely accepted gold-standard test to detect COVID-19. In this method, the RNA of the virus is duplicated by a process called reverse transcription to form DNA for facilitating the copying process. Fluorescent dye is attached to the viral genetic material and copied billions of times through the process called polymerase chain reaction. Enhanced fluorescence is used to identify the presence of genetic material of the virus. These tests are time-consuming and have significant false negatives, *i.e.*, a person with COVID-19 might be categorized as not having the virus. Large-scale RT-PCR testing has its own share of problems such as logistics, availability and affordability in underdeveloped nations, and reliability of the test results. Machine learning algorithms can act as a cheaper supplementary/alternative diagnostic tool for the testing process. In the current study, using publicly available chest X-ray image datasets, different convolutional neural network (CNN)-based models were developed for efficient identification of COVID-19 infected patients, and their efficacies were compared. Key innovations in training the CNNs are discussed. Our results indicate that EfficientNet, SeResNext, and ResNet are best at classifying normal, pneumonia and COVID-19 cases, respectively. The ResNet architecture with transfer learning performed best at detecting COVID-19 with an accuracy of 94%, a rate far superior to that in the RT-PCR test, which is typically in the range of 70 – 80%. This is particularly attractive as an additional noninvasive protocol since such technology-augmented detection is likely to help in reducing the psychological refractory period due to COVID-19 infections. Toward the healthy lung initiative in the post-COVID-19 era, we propose close coupling of the present diagnostic protocols with digital approaches to ensure more reliable personal care within the ambit of large-scale pandemic control mechanisms. Such integration with emerging technological tools can create a benchmark for the first line of defense against future global pandemics.

Keywords: COVID-19; Machine learning; Deep learning; EfficientNet; ResNet; SeResNext; Network-centric digital health stack

***Corresponding author:**
Kisor Kumar Sahu
(kisorsahu@iitbbs.ac.in)

Citation: Gogineni AK, Hitesh M, Jha PK, Sen SS, Das S, Sahu KK. Deep learning on chest X-ray and computed tomography scans for detection of COVID-19 as a part of network-centric digital health stack for future pandemics. *Artif Intell Health*. 2025;2(1):29-41. doi: 10.36922/aih.2888

Received: February 5, 2024

1st revised: April 17, 2024

2nd revised: May 15, 2024

3rd revised: July 3, 2024

Accepted: July 17, 2024

Published Online: October 7, 2024

[2]School of Minerals, Metallurgical and Materials Engineering, Indian Institute of Technology, Bhubaneswar, Odisha, India
[3]Department of Pulmonary Medicine, Srirama Chandra Bhanja Medical College and Hospital, Cuttack, Odisha, India
[4]Centre of Excellence for Novel Energy Materials (CENEMA), Indian Institute of Technology, Bhubaneswar, Odisha, India
[5]Virtual and Augmented Reality Centre of Excellence, Indian Institute of Technology, Bhubaneswar, Odisha, India

1. Introduction

In December 2019, clusters of pneumonia-like cases of unknown origin were first reported in the city of Wuhan in China. Upon investigations, it was found that the disease was caused by a new type of single-stranded RNA virus, which was officially named by the World Health Organization (WHO) as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).[1,2] The disease caused by SARS-CoV-2 was subsequently named COVID-19. The rate of spread and severity of COVID-19 all around the world forced the WHO to declare it as a pandemic on March 11, 2020. It had caused unparalleled disruptions in the post-internet modern era of global civilization and, in a way, demonstrated the serious shortcomings and vulnerabilities of traditional approaches for tackling such issues.[2] It was depicted that, although we had the bits and pieces to construct a robust technology-armed first line of defense against global pandemics based upon smart integration of the new age tools, we did not invest enough to construct the requisite technological architecture to realize this. Therefore, various countries attempted to do what they could do best under the prevailing situations. Some implemented nationwide lockdowns, limiting movement of the population, commanding social distancing, and expanding the consciousness of cleanliness and good hygiene as a range of preventive measures against the pandemic; an early analysis of such measures has been reported in the literature.[3]

SARS-CoV-2 can cause symptoms of fever, fatigue, headache, cough, sore throat, myalgia (muscle pain), anosmia (loss of smell), and other respiratory symptoms. Most people recover from the disease without requiring any special treatment but those who have comorbid medical conditions such as chronic kidney disease, autoimmune disease, cancer, heart conditions, obesity, diabetes, and respiratory disease are more prone to develop serious illness.[3] COVID-19 causes serious damage to the lungs, causing oxygen deficiency in the patient, a condition referred to as hypoxia. This can be easily identified by lower oxygen saturation level, typically less than 94%. Infections due to bacterial pneumonia and pulmonary embolism present similar characteristics. Furthermore, other underlying conditions such as chronic obstructive pulmonary disease typically complicate the identification of COVID-19. Atelectasis (indicating a partial or complete collapse of the lungs caused by COVID-19) can be partially reversed by maintaining continuous positive airway pressure, which forces the collapsed alveoli to remain open with oxygen-rich air. A far worse situation demands the use of ventilators, where breathing is assisted with a life-support device. Thus, since it is a common symptom across various respiratory diseases, depleting oxygen saturation cannot be a definitive indicator of COVID-19 infection.

To arrest the spread of the disease, affected countries have adopted reverse-transcription polymerase chain reaction (RT-PCR) tests as the gold-standard diagnostic method to detect COVID-19 infection and isolate the infected individuals as early as possible to limit further transmission of the infection. Despite having many positive aspects, this testing method requires the setting-up of entirely new facilities with specially trained personnel for sample collection and analysis. Such facilities are largely authority-regulated and many underdeveloped countries find it difficult to procure sufficient numbers of test kits. Moreover, it is time-consuming and sometimes gives non-negligible false negative (infected person is tested negative) as well as false positive (uninfected person is tested positive) results.

Machine learning (ML) has made a significant impact in many disciplines of science and technology.[4-11] An ML tool, based on chest X-ray and/or computed tomography (CT) scan images of COVID-19 suspected individuals, can be very attractive in this scenario, primarily because of the very limited resources required and short diagnosis time. Computer-aided biomedical image analysis might turn out to be an additional useful tool to assist medical practitioners in correct and quick decision-making. For example, Habib *et al.*[7] introduced a novel modified lightweight SqueezeNet (SQN-MF) model (demonstrated in non-medical application) coupled with continuous wavelet transformation for converting acoustic emission signals into two-dimensional (2D) images, achieving 100% classification accuracy (surpassing traditional techniques by 20.8%). The lightweight model (0.5 MB) is suitable for field programmable gate array implementation, enabling real-time monitoring. Their success in achieving high accuracy with a memory-efficient model demonstrates the potential for similar approaches in medical diagnostics that will help to build reliable, rapid screening protocols for infectious diseases like COVID-19. This kind of lightweight

computer-based automated tools can be extremely useful, particularly when faced with a logistics bottleneck or when the entire medical infrastructure is overwhelmed for some reason or other, which can be very handy during the outbreak of some other potent infectious diseases.

The method outlined in the present article could be a very important building block in an emerging integrated digital medical doctrine that can ensure a more reliable, personalized, and targeted medical intervention even within the ambit of a very large-scale pandemic control initiative spanning across prefectures/territories/states/countries as previously outlined.[2] One of the foundational principles of modern-era scientific practices is to decouple critical processes to maximize control over them and create provisions for more efficient resource allocation. In the RT-PCR test, the processes of physical examination of the patient by the doctor, sample collection, and actual evaluation of the testing results are highly integrated and codependent. A decoupling approach in this scenario is difficult to achieve. In the present ML-based method, the physical examination of the patient in the form of taking an X-ray radiograph and the evaluation of the testing results from the ML model can be effectively decoupled, thus offering a modular approach that tremendously enhances flexibility in operational and clinical protocols for pandemic control. This effective decoupling and modular approach will be a crucial component for the network-centric digital 'health stack' for the future pandemic control by effectively predicting its geographical occurrences and this will be dealt with in the second article (in writing progress) in this series. Let us illustrate this point further with a simple example for the sake of brevity (the comprehensive architecture will be outlined in the follow-up article): say an economically weaker country "A" lacks enough resource pool of highly trained medical professionals and/or RT-PCR test kits reserves. However, medical X-ray radiography is one of the oldest and most common diagnostic techniques. As a result, most countries (including "A") are already well-equipped with X-ray devices. The approach proposed in this article requires minimal training of personnel who can supervise X-ray radiography tests satisfactorily (that is, just to take a satisfactory radiograph; in the case it is not satisfactory, it can be automatically flagged by a computer/professional located in another country to re-do the test). However, the difficult part of this approach is to train a large number of medical professionals in analyzing/getting familiar with X-ray images for COVID-19 or future pandemic detection. Here, the great advantage of decoupling the critical process becomes clear. By using internet-enabled technologies, the tests and results can then be transmitted online for further evaluation by automated ML algorithms. Only a few confusing cases need to be cross-examined by trained medical professionals in another country, say "B," who can undertake the post-ML decision-making with the help of well-developed resources natively available in "B." It is important to note that the present protocol is performed by an ML algorithm implemented on a computer, and the actual job of a medical expert is minimal, limited to only overviewing/double-checking the assessment of the computer to eliminate the remote possibility of mistakes (since the machine accuracy is more than 94% even without human intervention, as demonstrated in a later section). Such global alliances can be a game changer in providing equitable diagnostics to underdeveloped regions of the world (Global South), as already demonstrated in the case of vaccination under the COVID-19 Vaccines Global Access, (COVAX).[12] Moreover, free exchange of valuable medical data across the boundaries will have the potential to create huge synergistic effects, for example, in faster identification of newer strains, proper epidemiological analysis and consequent prevention strategy, and most importantly, a reliable prediction about the future trends based on hard data obtained through data-driven approach. Developing an automated analysis system can therefore save valuable time for the medical professionals in the country "A," which could be best utilized to address some other critical conditions. This will also ensure optimal resource utilization and advanced preparation for the pandemic in country "B." As pointed out earlier, the detailed digital architecture and infrastructure planning will be presented in the follow-up article. It is important to note that such a design will not be COVID-19-specific but flexible enough to handle a broad spectrum of other diseases.

COVID-19 radiograph (CORAD) scores from CT scans are considered a definite diagnosis of COVID-19 even in the case of negative RT-PCR results. Chest X-ray or CT scan of a COVID-19-infected patient generally reports abnormal findings, such as ground glass opacities, and coarse horizontal linear opacities scattered throughout the lungs, often with consolidation.[13] They represent tiny air sacs getting filled with fluid. Another finding is called the "crazy paving" pattern, which is caused due to swelling of the interstitial septum along the walls of the lung lobes superimposed on the background of ground glass opacities. The latter finding is observed in the advanced stage of infection.[14] Although RT-PCR remains the gold standard procedure for COVID-19 detection, these findings in X-ray images can help in the initial screening of the suspected patients. There are six patterns indicative of COVID-19 infection that can be observed in a patient's X-ray report. They are: (i) reverse batwing, (ii) multifocal lower lobe predominant consolidation,

(iii) peribronchial rounded consolidations, (iv) multifocal bilateral consolidations, (v) ball pattern or round pneumonia, and (vi) bilateral symmetrical diffuse lung involvement. Out of these patterns, pattern (vi) is the most severe, suggesting acute respiratory distress syndrome in the patient.

Several research groups have used deep learning-based techniques for COVID-19 and pneumonia detection.[15-31] Since most of these techniques are well discussed and debated in the literature, we will very briefly review only a few of them. Wang *et al.*[16] used deep learning techniques on CT images to screen COVID-19 patients with an accuracy, specificity, and sensitivity of 89.5%, 88%, and 87%, respectively. A novel convolutional neural network (CNN), COVID-Net for detecting COVID-19 using chest X-ray images presented by a different research team has an accuracy of 83.5%.[15] Transfer learning offers several benefits, primarily in saving training time, significantly improving the performance of ML models, and requiring smaller data sets for training. For example, Sohaib *et al.*[8] proposed a novel approach (demonstrated in non-medical applications) that used step transfer learning (STL) combined with extreme learning machine (ELM), primarily focusing on improving the generalization power of deep learning models for autonomous inspection. By leveraging STL to extract generalized abstract features from diverse source images and utilizing ELM to overcome optimization limitations of traditional neural networks, their model achieved significant improvements in accuracy (2.5%), recall (4.8%), and precision (0.8%) compared to existing studies. This approach enhanced generalization demonstrating the usefulness of transfer learning techniques for increasing the robustness of the detection models and making the model more generalizable. Transfer learning has found many applications in the biomedical arena.[32] Joaquin[17] used a small dataset of 339 images for training and testing by utilizing the ResNet50-based deep transfer learning technique and obtained a validation accuracy of 96.2%. Ahmmed *et al.*[32] conducted an in-depth analysis of brain tumor classification using transfer learning across multiple classes, utilizing robust frameworks, such as ResNet 50 and Inception V3 for MRI images. Their research meticulously curated paired datasets and incorporated advanced techniques such as Early Stopping, ReduceLROnPlateau, and hyperparameter optimization. These strategies significantly improved model accuracy, achieving exceptional classification rates for various types of brain tumors. Similarly, Podder *et al.*[31] developed a deep learning model using optimized DenseNet architectures to diagnose infectious diseases from chest X-rays, achieving high detection rates for COVID-19 and other conditions. Their modifications to the DenseNet architecture and hyperparameter tuning demonstrated the potential of deep learning models in improving diagnostic accuracy and early disease detection.

In this study, we developed deep transfer learning-based approaches for the automatic detection of COVID-19 using various CNN-based models on chest X-ray images and implemented several innovations in the training protocol. Beyond the immediate applicability of the present method as an additional diagnostic tool for identifying COVID-19 infections, it also might have far-reaching consequences. An earlier big challenge was to identify and detect the infected individual; however, slowly but steadily, the major focus has shifted toward the long-term care of the lungs of the infected persons. All seven human coronaviruses affect the respiratory tracts and lungs. However, three of these virus types, *i.e.*, SARS-CoV, Middle East respiratory syndrome coronavirus, and SARS-CoV-2 are known to severely affect the lungs. Unfortunately, the long-term impact of SARS-CoV-2 infection is neither clearly understood nor widely studied as the viruses are dynamically evolving across the world. Most severe infections are known to cause post-COVID-19 sequelae, *i.e.*, fibrosis of the lungs in the future. A worthwhile initiative toward recording post-COVID-19 symptoms in this direction has been reported in the literature.[33] An ML-based system that relies on the chest X-ray and/or CT scan image analysis might prove to be highly valuable and instrumental in the long-term care of the lungs by appropriately creating a highly scalable and easily integrable digital infrastructure through curating, cataloging, classifying and most importantly, creating an appropriate data repository of the vital information about different stages of the lungs as a function of time (using automated time-stamps), significantly augmenting the healthy lung initiative in the post-COVID-19 era.

2. Data and methods

2.1. Data

The dataset of chest X-ray and CT scan images[34,35] used in this study were sourced from a freely available GitHub repository maintained by Cohen.[36] X-ray images containing pneumonia and normal images were obtained from the Kaggle dataset.[37] We used 1763 images to analyze the performance of the neural networks used for this study and utilized 1260 images to train the model along with 251 and 252 images for validating and testing the model, respectively. The dataset consists of 563, 947, and 223 images that belong to COVID-19, normal, and pneumonia categories, respectively. Figure 1 shows some examples from the dataset.

2.2. Methods: Different CNN architectures

Various CNN models such as ResNet, SeResNext, DenseNet, and EfficientNet were used for classification. Here we briefly

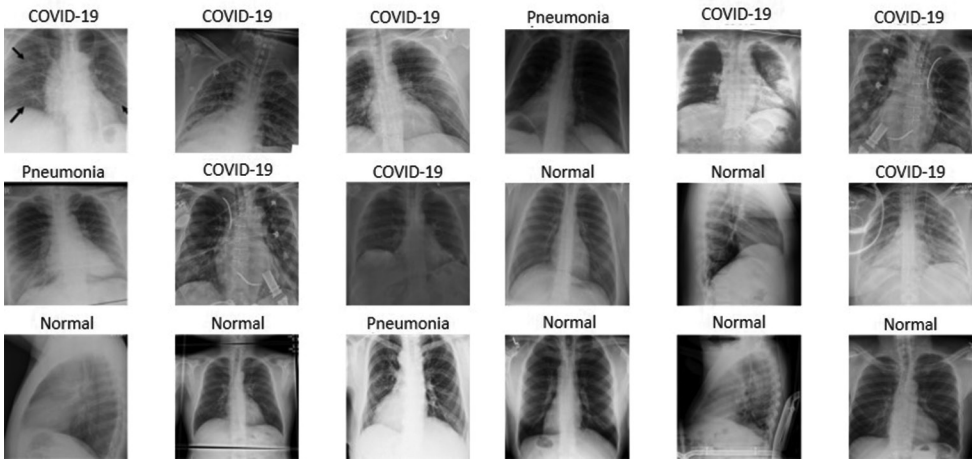

Figure 1. Sample images are taken from the dataset.[34,35,37] The indication above each image corresponds to the associated label. Image created by the author.

sketch their outlines, although complete details can be found in the reference mentioned in the corresponding sections. Fundamental building blocks are schematically depicted in Figure 2A-D.

2.2.1. Method 1: ResNet34

Developing a proper training protocol is a matter of serious concern for the implementation of any deep neural network. One such major issue is the divergence from local minima, leading to improper training. To address this, in ResNet,[38] a modification in the network architecture was introduced by incorporating skip connections, expressed as H(x) + x, between layers. This alteration facilitated quicker and more efficient model training. The smooth loss landscape of ResNet prevents the model from becoming trapped in local minima or saddle points, resulting in improved training speed and accuracy. In our study, we utilized a variant of ResNet, specifically ResNet34, consisting of a total of 34 convolutional layers.

2.2.2. Method 2: SeResNext50

SeNet was originally proposed by Hu *et al.*[39] SeNet differs from conventional neural network designs by emphasizing the exploration of channel-wise features rather than solely focusing on spatial features. In its fundamental structure, the squeeze-and-excitation (SE) block transforms the input, denoted as x, into a feature map U through convolution. This map undergoes a squeeze operation, consolidating feature maps across spatial dimensions to produce channel descriptors. These descriptors encapsulate the global distribution of channel-wise feature responses. Following this, an excitation operation converts the descriptors into per-channel modulation weights. These weights are then applied to the feature map U to yield the output of the SE

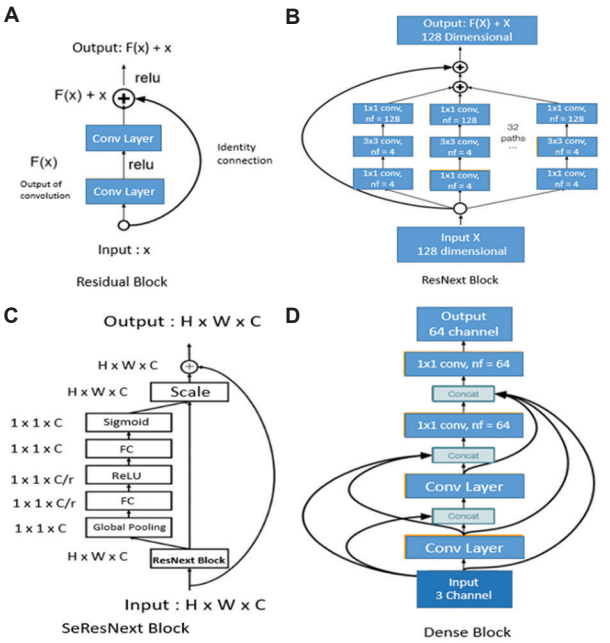

Figure 2. Building blocks of different convolutional neural network-based architectures: (A) ResNet, (B) ResNext, (C) SeResNext, and (D) DenseNet. Copyright © 2020 Springer International Publishing. Reprinted with permission of Springer International Publishing.

block. In the SeResNext model,[40,41] SE block is integrated into every non-identity branch of the ResNext block—a variant of the ResNet block characterized by multiple convolution layers and skip connections. A ResNext block consists of several convolution layers, each having a distinct set of filter sizes and dimensions, and incorporates a single skip connection for mitigating the vanishing gradient problem during training. A SeResNext block is shown in Figure 2C.

2.2.3. Method 3: DenseNet

DenseNet architecture was proposed by Xie *et al.*[40] Its architecture is very similar to that of ResNet, where there are feed-forward connections from each layer to the next layer.[38] In DenseNet, feature maps of one layer are concatenated with feature maps of all the following layers. This approach offers a benefit by leveraging features extracted from early layers for subsequent layers. Convolution blocks are sequentially stacked, and interspersed with basic convolution layers to preserve dimensionality across the network's depth. It consists of various "dense blocks." A simple "dense block" is depicted in Figure 2D.

2.2.4. Method 4: EfficientNet

Most of the architectures such as ResNet, VGG Net, and Inception Net, are created manually by researchers where they specify the complete network architecture upfront, for example, number of layers, filter size, and number of channels based on previous experiments/experience. EfficientNets[42] are created using neural architecture search where the complete model is built algorithmically by keeping a constraint on the number of parameters. It uses ResNet as a baseline model and modifies the number of layers, number of channels, and input image dimensions in the baseline model to create the desired model. The smallest model, *i.e.*, with a minimum number of parameters in EfficientNet is called b0, and seven other models are generated by changing the constraints, such as the number of parameters and the number of Floating Point Operations Per Second. EfficientNet b7 is the largest model among EfficientNets. EfficientNets have very less inference time when compared to other models with a similar number of parameters. As the image size increases, larger EfficientNet models are preferred since they have a greater number of layers and the channel size also increases. This helps in obtaining useful features from the larger image.

Overall, the rational for choosing this architecture is as follows: ResNet was selected for its skip connection architecture, which facilitates stable learning during training; DenseNet for its dense connectivity facilitating feature reuse across layers; SeResNext for its integration of SE modules for enhanced feature recalibration; and EfficientNet for its efficient model scaling strategy, which collectively provides a diverse range of architectural innovations for achieving accurate and reliable image classification. It should be noted that it is always possible (and sometimes more desirable) to make an ensemble of these models for further improving the overall prediction. Since the present article focuses on the fundamental aspects of the implementation of these models, the ensemble strategy is not explored in this work.

Transfer learning uses the initial weights of the neural network pre-trained on a different database, which is ImageNet[43] for the present study. Transfer learning plays a crucial role in improving the model performance when working on a dataset where the number of images are limited. The amount of labeled data available in the biomedical domain is limited mainly due to the time taken to annotate the dataset. Initializing the model weights in this manner helps the model to capture important information from the images. The initial layers of the network capture very generic information from an image such as horizontal and vertical edges, whereas the later layers of the network capture patterns in an image that are very specific to the dataset of study as previously described.[41] This pre-training on a large dataset provides a solid foundation, allowing the model to start with a better understanding of general visual patterns. The fine-tuning process adapts the model to the unique features and characteristics of the target dataset, allowing it to specialize in recognizing patterns relevant to the biomedical images at hand.

2.3. Training

We employed a learning rate scheduler[44] to determine the most effective learning rate for our specific dataset. During this process, the learning rate is cautiously increased after each mini-batch, with the corresponding loss recorded at each increment. Subsequently, we plotted the loss versus learning rate, as illustrated in Figure 3, revealing how different learning rates impacted the model's performance. Notably, for a very low learning rate, the loss diminishes at a slower pace. As the learning rate increased, the loss showed a rapid decline, indicating the optimal range. Beyond this point, further increases in the learning rate caused the loss to rise sharply, suggesting overshooting. By identifying the point of the steepest loss decline (0.002 in our case), we determined the optimal learning rate for

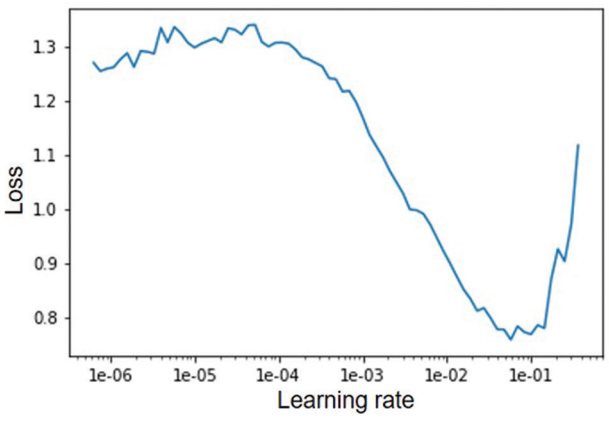

Figure 3. Scheduling the learning rate by investigating its impact on the loss function. Image created by the authors.

training. It is crucial to avoid exceeding this optimal value, as it may lead to overshooting the global minima of the loss function. This balanced approach ensures efficient training and better model generalization by avoiding both slow convergence and the risk of overshooting the global minima.

Adjusting the learning rate plays a critical role in training neural networks. A smaller value of the learning rate leads to gradual changes in the loss function but can prolong convergence due to small gradients. Conversely, a larger value of the learning rate may cause overshooting of the global minimum. Hence, striking a balance between the two extremes is essential for efficient training and better generalization. Furthermore, employing a fixed learning rate can lead to challenges such as becoming trapped in local minima or saddle points where gradients are insufficient for optimization. To mitigate these issues, we set upper and lower bounds on the learning rate, typically differing by a factor of ten. For instance, in our experiments, the upper bound is set at 10^{-4} and the lower bound at 10^{-5}. This range allows the model to explore the loss landscape effectively, avoiding stagnation in local minima or saddle points. In addition, in lieu of employing a uniform learning rate throughout the entire network during the training process, this study adopts a strategy known as discriminative learning rates. Here, learning rates are tailored for different layers of the classifier, typically ranging between 0.0001 and 0.01. This approach acknowledges that various network layers capture distinct types of information, thus warranting diverse learning rates. Initial layers receive lower learning rates compared to later layers, reflecting their differing roles in feature extraction and abstraction.

The training methodology incorporates a one-cycle training policy, characterized by a dynamic adjustment of learning rates across epochs. Initially, a higher learning rate is applied, gradually decreasing toward the final epoch. This technique promotes improved model performance and stability, facilitating parameter updates at an appropriate pace. By mitigating the risk of local minima entrapment, the model's generalizability is enhanced.

Implementation details involve Python programming utilizing the fastai library,[45] a PyTorch-based open-source platform tailored for deep learning model development. Execution takes place on Google Collaboratory, leveraging its provision of a free K-80 GPU with 12GB RAM, ideal for executing ML algorithms. The source code and neural network weights utilized in this study are publicly available for reference.[46] In the following, we summarize some of the key novelties in the training strategy in the present study. We used training strategies to optimize the performance of EfficientNet, ResNet, and SeResNext for COVID-19 detection using X-ray images. Key innovations include the use of a dynamic learning rate scheduler to identify the optimal learning rate, setting adaptive boundaries (10^{-4} – 10^{-5}) to avoid local minima, and employing discriminative learning rates tailored for different network layers (ranging from 0.0001 to 0.01) to enhance feature extraction and abstraction. In addition, we implemented a one-cycle training policy, dynamically adjusting learning rates across epochs to improve model performance and stability. These training methodologies significantly enhance the model's robustness and accuracy, providing a distinctive contribution to the field. Altogether, our proposed model introduces several key innovations that enhance the robustness and accuracy of these models, making our work distinct from prior studies.

3. Results

A total of 1763 images were used to build the ML model, of which 1260 images were used to train the model and 251 and 252 images were employed for validating and testing the model, respectively. After the screening, the 563 images indicating COVID-19 were split into train (450 images), valid (71 images), and test datasets (72 images).

Figure 4A-D shows the confusion matrix for various CNN architectures. ResNet and DenseNet obtained the best accuracy, at 94.09%. The confusion matrix corresponds to the predictions of the model on the test dataset. The model is able to distinguish clearly among various classes. The confusion matrices indicate that EfficientNet is best at classifying normal images and SeResNext is best at classifying pneumonia. ResNet performs best for classifying images pertaining to COVID-19.

Figure 5 shows the predicted class, actual class, loss, and probability of actual class for a set of misclassified images for each model in the format "Prediction/Actual/Loss/Probability" on the top of each image. Each image illustrates the target class results for the model, highlighting areas where the model's predictions did not align with the true labels.

Figure 6 shows the sensitivity as well as specificity of the CNN models used in this work. The sensitivity of a class measures the proportion of images belonging to a particular class that are correctly classified by the model. The specificity of a class measures the proportion of images that do not belong to the class of interest and are correctly classified by the model.

4. Discussion

EfficientNet has the highest specificity toward images of class COVID-19 and ResNet has the highest sensitivity

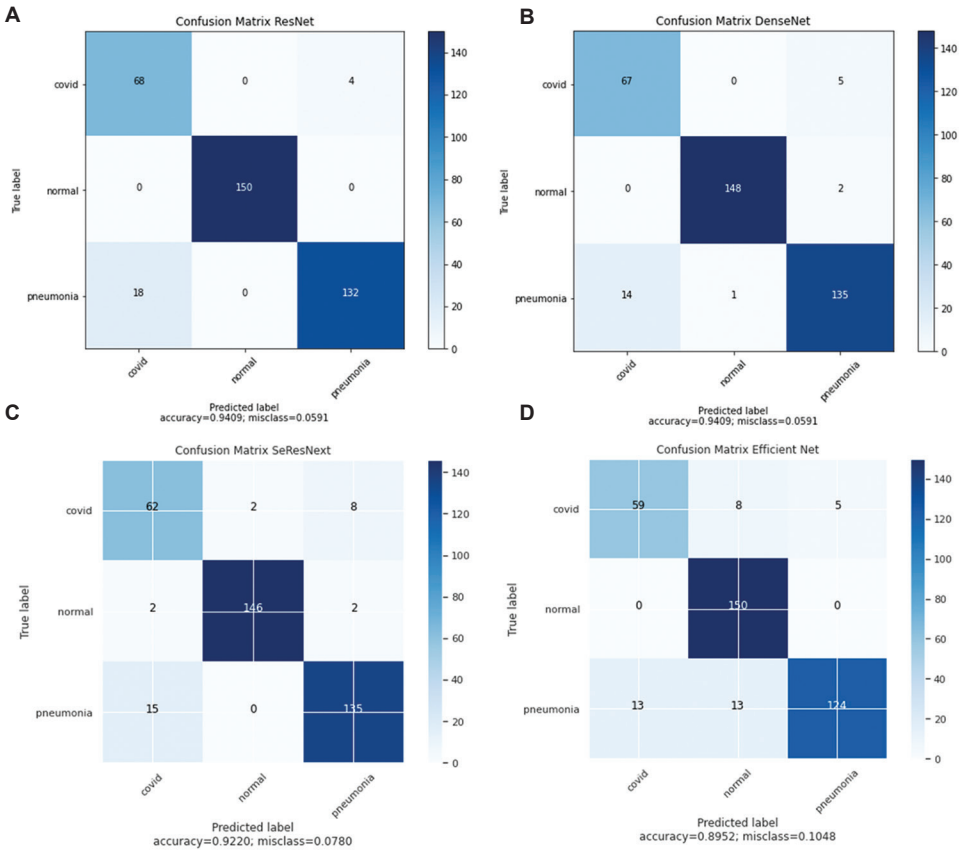

Figure 4. Confusion matrices for (A) ResNet (34), (B) DenseNet, (C) SeResNext, and (D) EfficientNet
Note: The number 34 indicates the number of convolution layers).

toward images of class COVID-19. The sensitivity and specificity values in Figure 6A and B, respectively, were computed by evaluating the performance of each CNN model (ResNet, DenseNet, SeResNext, EfficientNet) on a labeled test set consisting of three classes: COVID-19, pneumonia, and normal X-ray images. Sensitivity (true positive rate) was calculated as the ratio of correctly identified positive cases to the total actual positive cases for each class. Specificity (true negative rate) was determined by the ratio of correctly identified negative cases to the total actual negative cases for each class. For each model, predictions were compared against the ground truth labels to calculate these metrics, providing a comprehensive assessment of each model's ability to correctly identify and distinguish between the three classes.

RT-PCR's sensitivity and specificity are typically in the range of 70 – 80% and 99 – 100%, respectively. Therefore, the 94% sensitivity achieved in the present study with very limited numbers of training images is an indication of good performance, which is expected to get better with more training images. Therefore, our method is significantly more accurate. Subsequently, doctors may be consulted for a confirmed diagnosis. Some examples of COVID-19 scans are described in the literature.[47] In addition to the tremendous promise, the efficacy of the present method can be significantly enhanced, if it is supplemented with blood lymphocyte count (as lymphopenia—a lower count of lymphocytes—is mostly associated with COVID-19 and indicates severe form) and RT-PCR test data from nasopharyngeal samples collected through swabs. The efficacy of the present method will significantly improve, even when not assisted by other methods, with more and more usage (as the method progressively learns and gets better at the job), as is the case for any ML method.

It is imperative to mention that collecting data from diverse geographical regions can significantly improve the performance of our model by introducing a wider variety of image characteristics and potential variations in COVID-19 presentation. Different regions may have variations in imaging equipment, patient demographics, and prevalence of comorbidities, all of which can influence the appearance of X-ray images. By incorporating a more diverse dataset, our model can learn to generalize better across different populations and imaging conditions,

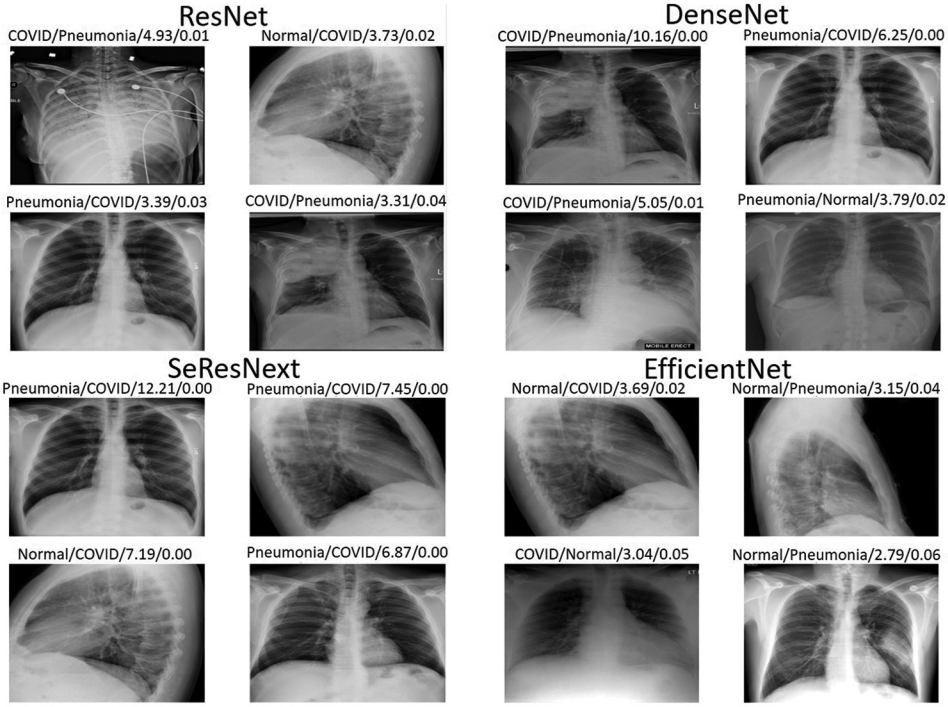

Figure 5. Misclassified X-ray images with prediction details using ResNet, DenseNet, SeResNext, and EfficientNet models. On the top of each subplot, "Prediction/Actual/Loss/Probability" details for each individual image are shown.

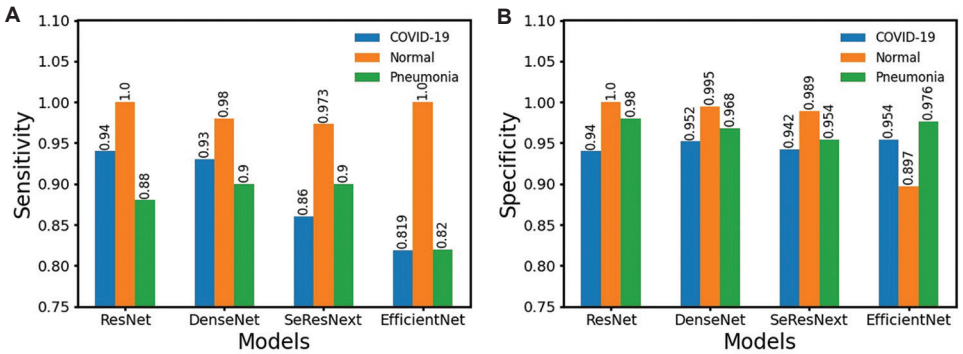

Figure 6. Sensitivity (A) and specificity (B) of ResNet (34), DenseNet, SeResNext, and EfficientNet for COVID-19, normal, and pneumonia class detection
Note: The number 34 indicates the number of convolution layers).

leading to more robust and accurate predictions. This enhanced diversity helps mitigate biases and ensures that the model performs well in real-world, heterogeneous environments, ultimately improving its reliability and effectiveness in detecting COVID-19, pneumonia, and normal cases at the global scale and for future pandemics, as well.

5. Conclusion

In this work, we implemented various CNN models with a transfer learning-based approach to classify COVID-19 and pneumonia from normal patients through chest X-rays. This research article introduces several key innovations in the training methodology that distinguish it from prior works:

(1) *Dynamic learning rate optimization.* We developed a novel dynamic learning rate scheduler that adaptively identifies the optimal learning rate within carefully set boundaries (10^{-4} – 10^{-5}). This approach mitigates the risk of getting trapped in local minima, a common challenge in neural network training.

(2) *Layer-specific discriminative learning.* This study implemented a discriminative learning rate strategy, applying different learning rates (ranging from 0.0001

to 0.01) to various network layers. This nuanced approach enhances both the low-level feature extraction and high-level abstraction capabilities of the models.

(3) *One-cycle policy implementation.* We integrated a one-cycle training policy that dynamically adjusts learning rates across epochs. This method has been shown to significantly improve model convergence, stability, and overall performance.

(4) *Comparative analysis of advanced architectures.* While individual CNN architectures have previously been applied to COVID-19 detection, our study provides a comprehensive comparison of EfficientNet, ResNet, and SeResNext using these advanced training strategies. This comparison offers valuable insights into the relative strengths of these architectures for this specific task.

(5) *Reproducibility and benchmarking.* By using a publicly available dataset and clearly documenting our methodologies, we provide a robust benchmark for future studies in medical image analysis, which extends beyond the scope of COVID-19 detection.

These methodological innovations collectively enhance the robustness, accuracy, and generalizability of CNN models for medical image analysis. While the immediate application to COVID-19 may seem less pressing now, the techniques we developed have broader implications for improving deep learning approaches in medical imaging across various conditions that might be very useful in our fight against future pandemics.

Among the models implemented in the present study, ResNet and DenseNet have achieved more than 94% accuracy. This is far superior to the typical sensitivity of 70 – 80% for RT-PCR. Our results indicate that EfficientNet is best at classifying normal images, and SeResNext is best at classifying pneumonia. ResNet performs best for classifying images pertaining to COVID-19. While the accuracy of the present method is expected to get better with increasing usage, which is an inherent feature of artificial intelligence, there is no such chance for RT-PCR, since this traditional method is not a smart protocol. The model is able to learn the inherent features of pneumonia and COVID-19 from a relatively small dataset. The performance of the model can be improved further by collecting data from diverse geographical regions. This will also improve the generalizability of the model.

We strongly believe that this ML-aided diagnostic protocol can help in detecting individuals suspected of carrying infections with greater speed and accuracy, and more importantly, it charts out the blueprint to rapidly develop a new med-tech protocol for quick screening of future pandemics. It is pertinent to point out here that the decoupling of physical examination of the patient and analytical pathology leads to an effective and modular approach. This is likely to significantly enhance detection speed, accuracy, and sensitivity, expected to form the fundamental cornerstone that will be pivotal for an extensive digital architecture to safeguard against many future pandemics (to be elaborated in the follow-up article). Furthermore, these models help in the early screening of suspects in remote places in countries where the health care providers as well as resources (such as RT-PCR kits and CT scan machines) are limited.

Acknowledgments

The authors acknowledge the sincere help from the authorities of VARCoE, IIT BBS, and SCBMCH for the research support and encouragement.

Funding

None.

Conflict of interest

The authors declare no conflicts of interest.

Author contributions

Conceptualization: Ajay Kumar Gogineni, Kisor Kumar Sahu
Formal analysis: All authors
Investigation: Ajay Kumar Gogineni, Madapathi Hitesh
Methodology: Ajay Kumar Gogineni, Kisor Kumar Sahu
Writing – original draft: All authors
Writing – review & editing: All authors

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Data are available at the following resources:

(1) Cohen JP, Morrison P, Dao L. COVID-19 Image Data Collection. arXiv.org. Accessed April 9, 2021. https://doi.org/10.48550/arXiv.2003.11597

(2) Yang X, He X, Zhao J, Zhang Y, Zhang S, Xie P. COVID-CT-Dataset: A CT Scan Dataset about COVID-19. *arXiv:200313865 [cs, eess, stat]*. Published online June 17, 2020. http://arxiv.org/abs/2003.13865

(3) Cohen JP. ieee8023/covid-chestxray-dataset. GitHub. Published June 10, 2020. https://github.com/ieee8023/covid-chestxray-dataset

(4) Chest X-Ray Images (Pneumonia). www.kaggle.

com. Accessed April 9, 2021. https://kaggle.com/paultimothymooney/chest-xray-pneumonia

(5) Gogineni A. AjayKumarGogineni777/covid_cnn. GitHub. Published October 16, 2020. Accessed April 10, 2021. https://github.com/AjayKumarGogineni777/covid_cnn

References

1. World Health Organization. *Determinants of Health*; 2023. Available from: https://www.who.int [Last accessed on 2024 Sep 16].
2. Pathak AD, Saran D, Mishra S, Hitesh M, Bathula S, Sahu KK. Smart War on COVID-19 and Global Pandemics. *In: Computational Modeling and Data Analysis in COVID-19 Research*. CRC Press; 2021:67-94.

   doi: 10.1201/9781003137481-5
3. Kishore R, Jha PK, Das S, *et al*. A kinetic model for qualitative understanding and analysis of the effect of complete lockdown imposed by India for controlling the COVID-19 disease spread by the SARS-CoV-2 virus. *arXiv*. Preprint posted online 2020.

   doi: 10.48550/arXiv.2004.05684
4. Varadi M, Anyango S, Deshpande M, *et al*. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high accuracy models. *Nucleic Acids Res*. 2021;50(D1):D439-D444.

   doi: 10.1093/nar/gkab1061
5. Chowdhery A, Narang S, Devlin J, *et al*. PaLM: Scaling Language Modeling with Pathways. *arXiv*. Preprint posted online 2022.

   doi: 10.48550/arXiv.2204.02311
6. Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A ConvNet for the 2020s. *arXiv*. Preprint posted online 2022.

   doi: 10.48550/arXiv.2201.03545
7. Habib A, Hasan J, Kim J. A Lightweight deep learning-based approach for concrete crack characterization using acoustic emission signals. *IEEE Access*. 2021;9:104029-50.

   doi: 10.1109/access.2021.3099124
8. Sohaib M, Hasan MJ, Chen J, Zheng Z. Generalizing infrastructure inspection: Step transfer learning aided extreme learning machine for automated crack detection in concrete structures. *Meas Sci Technol*. 2024;35:055402.

   doi: 10.1088/1361-6501/ad296c
9. Liu Z, Lin Y, Cao Y, *et al*. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. *arXiv*. Preprint posted online 2021.

   doi: 10.48550/arXiv.2103.14030
10. Dosovitskiy A, Beyer L, Kolesnikov A, *et al*. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv*. Preprint posted online 2020.

    doi: 10.48550/arXiv.2010.11929
11. Radford A, Kim JW, Hallacy C, *et al*. Learning Transferable Visual Models From Natural Language Supervision. *arXiv*. Preprint posted online 2021.

    doi: 10.48550/arXiv.2103.00020
12. WHO. *COVAX: Working for Global Equitable Access to COVID-19 Vaccines*. World Health Organization; 2020. Available from: https://www.who.int/initiatives/act-accelerator/covax [Last accessed on 2021 Apr 09].
13. Cleverley J, Piper J, Jones MM. The role of chest radiography in confirming covid-19 pneumonia. *BMJ*. 2020;370:m2426.

    doi: 10.1136/bmj.m2426
14. Zu ZY, Jiang MD, Xu PP, *et al*. Coronavirus disease 2019 (COVID-19): A perspective from China. *Radiology*. 2020;296(2):E15-E25.

    doi: 10.1148/radiol.2020200490
15. Wang L, Lin ZQ, Wong A. COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Sci Rep.* 2020;10(1):19549.

    doi: 10.1038/s41598-020-76550-z
16. Wang S, Kang B, Ma J, *et al*. A deep learning algorithm using CT images to screen for Corona virus disease (COVID-19). *Eur Radiol*. 2021;31:6096-6104.

    doi: 10.1007/s00330-021-07715-1
17. Joaquin AS. *Using Deep Learning to Detect NCOV-19 from X-Ray Images*. Medium; 2020. Available from: https://medium.com/data-science/using-deep-learning-to-detect-ncov-19-from-x-ray-images-1a89701d1acd [Last accessed 2020 Jun 27].
18. Ismael AM, Şengür A. Deep learning approaches for COVID-19 detection based on chest X-ray images. *Expert Syst Appl*. 2021;164:114054.

    doi: 10.1016/j.eswa.2020.114054
19. Zhang J, Xie Y, Pang G, *et al*. Viral pneumonia screening on chest x-rays using confidence-aware anomaly detection. *IEEE Trans Med Imaging*. 2021;40(3):879-890.

    doi: 10.1109/tmi.2020.3040950
20. Hemdan EED, Shouman MA, Karar ME. COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images. *arXiv*. Preprint posted online 2020.

    doi: 10.48550/arXiv.2003.11055
21. Jain R, Gupta M, Taneja S, Hemanth DJ. Deep learning based detection and analysis of COVID-19 on chest X-ray images. *Appl Intell*. 2020;51:1690-1700.

doi: 10.1007/s10489-020-01902-1

22. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra Acharya U. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med*. 2020;121:103792.

doi: 10.1016/j.compbiomed.2020.103792

23. Apostolopoulos ID, Aznaouridis SI, Tzani MA. Extracting possibly representative COVID-19 biomarkers from X-ray images with deep learning approach and image data related to pulmonary diseases. *J Med Biol Eng*. 2020;40(3):462-469.

doi: 10.1007/s40846-020-00529-4

24. Rahaman MM, Li C, Yao Y, *et al*. Identification of COVID-19 samples from chest X-Ray images using deep learning: A comparison of transfer learning approaches. *J Xray Sci Technol*. 2020;28(5):821-839.

doi: 10.3233/xst-200715

25. Ouchicha C, Ammor O, Meknassi M. CVDNet: A novel deep learning architecture for detection of coronavirus (Covid-19) from chest x-ray images. *Chaos Solitons Fractals*. 2020;140:110245.

doi: 10.1016/j.chaos.2020.110245

26. Ayan E, Ünver HM. Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning. Istanbul, Turkey: Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science (EBBT); 2019. p. 1-5.

doi: 10.1109/EBBT.2019.8741582

27. Asnaoui KE, Chawki Y, Idri A. Automated Methods for Detection and Classification Pneumonia based on X-Ray Images Using Deep Learning. arXiv. Preprint posted online 2020.

doi: 10.48550/arXiv.2003.14363

28. Elshennawy NM, Ibrahim DM. Deep-pneumonia framework using deep learning models based on chest X-ray images. *Diagnostics*. 2020;10(9):649.

doi: 10.3390/diagnostics10090649

29. Ibrahim AU, Ozsoz M, Serte S, Al-Turjman F, Yakoi PS. Pneumonia classification using deep learning from chest X-ray images during COVID-19. *Cogn Comput*. 2021;16:1589-1601.

doi: 10.1007/s12559-020-09787-5

30. Kundu R, Das R, Geem ZW, Han GT, Sarkar R. Pneumonia detection in chest X-ray images using an ensemble of deep learning models. *PLoS One*. 2021;16(9):e0256630.

doi: 10.1371/journal.pone.0256630

31. Podder P, Alam FB, Mondal MR, Hasan MJ, Rohan A, Bharati S. Rethinking densely connected convolutional networks for diagnosing infectious diseases. *Computers*. 2023;12(5):95.

doi: 10.3390/computers12050095

32. Ahmmed S, Podder P, Mondal M, *et al*. Enhancing brain tumor classification with transfer learning across multiple classes: An in-depth analysis. *BioMedInformatics*. 2023;3(4):1124-1144.

doi: 10.3390/biomedinformatics3040068

33. Kostka K, Roel E, Trinh NT, *et al*. The burden of post-acute COVID-19 symptoms in a multinational network cohort analysis. *Nat Commun*. 2023;14:7449.

doi: 10.1038/s41467-023-42726-0

34. Cohen JP, Morrison P, Dao L. COVID-19 Image Data Collection. *arXiv*. Preprint posted online 2020.

doi: 10.48550/arXiv.2003.11597

35. Yang X, He X, Zhao J, Zhang Y, Zhang S, Xie P. COVID-CT-Dataset: A CT Scan Dataset about COVID-19. *arXiv*. Preprint posted online 2020.

doi: 10.48550/arXiv.2003.13865

36. Cohen JP. Covid chest X-ray dataset. *GitHub*; 2020. Available from: https://github.com/ieee8023/covid-chestxray-dataset [Last accessed on 2021 Apr 09].

37. Mooney P. Chest X-Ray Images (Pneumonia). *Kaggle*. Available from: https://kaggle.com/paultimothymooney/chest-xray-pneumonia [Last accessed on Apr 09].

38. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. *In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE; 2016:770-778.

doi: 10.1109/cvpr.2016.90

39. Hu J, Shen L, Albanie S, Sun G, Wu E. Squeeze-and-Excitation Networks. *arXiv*. Preprint posted online 2017.

doi: 10.48550/arXiv.1709.01507

40. Xie S, Girshick R, Dollar P, Tu Z, He K. Aggregated Residual Transformations for Deep Neural Networks. *In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE; 2017:5987-5995.

doi: 10.1109/cvpr.2017.634

41. Simonyan K, Vedaldi A, Zisserman A. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. *arXiv*. Preprint posted online 2013.

doi: 10.48550/arXiv.1312.6034

42. Tan M, Le QV. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *arXiv*. Preprint posted online 2019.

doi: 10.48550/arXiv.1905.11946

43. Simon M, Rodner E, Denzler J. ImageNet pre-trained models with batch normalization. *arXiv*. Preprint posted online 2016.

doi: 10.48550/arXiv.1612.01452

44. Smith LN. Cyclical Learning Rates for Training Neural Networks. *In: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE; 2017:464-472.

    doi: 10.1109/wacv.2017.58

45. Text Learner. *Fastai*. Available from: https://docs.fast.ai/text.learner.html [Last accessed on 2021 Apr 09].

46. Gogineni A. COVID CNN. *GitHub*; 2020. Available from: https://github.com/ajaykumargogineni777/covid_cnn [Last accessed on 2021 Apr 10].

47. Yasin R, Gouda W. Chest X-ray findings monitoring COVID-19 disease course and severity. *Egypt J Radiol Nucl Med*. 2020;51(1):193.

    doi: 10.1186/s43055-020-00296-x