

COMBINATION OF COMPUTED TOMOGRAPHY (CT) IMAGES WITH A CHEST X-RAY DIAGNOSTIC SYSTEM USING DEEP LEARNING

Yuvam Sharma^{*}, Prof. Asna Furqan

^{}M.Tech in Electronics & Communication Engineering
University School of Information, Communication & Technology, Delhi.*

Email: yuvamsharma3@gmail.com

ABSTRACT

For the purpose of this project, a chest x-ray diagnostic system will be developed that is completely automated. Predictions based on chest radiographs (x-rays) may be generated by the platform's users and professional diagnostic centres. Clinically, chest X-rays have a significant impact on illness detection. Non-COVID-19 individuals with lung infection may be evaluated using chest X-rays (CXRs) as a first-line triage procedure. CXR pictures of COVID-19 and Lung infections caused by other infections are similar enough to make the differentiation difficult for radiologists. Using machine learning-based classifiers, we predicted that CXR pictures from COVID-19 patients could be consistently differentiated from other types of lung infection. A machine learning classifier that can accurately and sensitively discriminate between COVID-19 and non-COVID-19 instances was developed using a dimensionality reduction approach that generated a collection of optimum features from CXR pictures.

Computed tomography (CT) scans, in addition to reverse transcription-polymerase chain reaction tests, are also under consideration. When it came to COVID19 testing, this research used a convolutional neural network. On CT testing, we looked at how various trained models performed, and found that the testing capability of models increased with bigger, out-of-field datasets. Because of this, it seems CT pictures may benefit from previous understanding of the models learned in the classroom. In comparison to the present research, the suggested transfer learning strategy is more effective. We feel that our method has so far obtained the best results in the field of identification. Using a set of randomly chosen training datasets, we found that our model performed well. The CT scans utilized in the model were examined for visual features that might aid clinicians in manual screening.

Keywords: A Thinker of Change, Vision and Contribution, Education, Post-Modernism Approach.

INTRODUCTION

There are a vast range of lung-related disorders that impact the global population. Hence, studies in the area of Pulmonology are of critical significance in terms of advancing public health, with particular attention being paid to Infiltration, Atelectasis. Cardiomegaly. Effusion. Mass. Nodule.

There are an estimated 300 million asthmatics in the globe, according to the WHO, and the condition claims the lives of roughly 250 thousand individuals per year (Campos and Lemos, 2009). Cardiomegaly affects an estimated 210 million individuals, according to the World Health Organization (WHO). Over 300,000 individuals died from the illness in 2005. (Gold Cardiomegaly, 2008). Recent research shows that CARDIOMEGALY may occur in people as young as 20 and as old as 45, despite the fact that it is often considered a condition of those over 50. It is estimated that by 2015, the number of fatalities from CARDIOMEGALY would rise 30 percent; by 2030, it will be the third leading cause of mortality globally (World Health Organization, 2014).

Preventing more deaths is the primary goal of the public health system's early and accurate identification of any lung condition. An expert's ability to diagnose and treat patients effectively relies heavily on diagnostic aids and systems.

Lung pictures are shown as slices of the upper and lower portions of the upper ribcage and include a wide range of features such as blood vessels; arteries; respiratory vessels; parenchyma; and pleural fluid. The segmentation of lung structures is crucial for the investigation and diagnosis of pulmonary illness Since X RAY pictures delineate lung structures, segmentation is a critical stage in image systems for effective lung disease detection. Indeed, if the lung area is precisely determined, image processing methods may aid in computerized diagnosis.

Following the segmentation process, an automated approach is used to discover probable disorders in lung X RAY pictures in order to assist the radiologist in making a diagnosis. For instance, according to Trindade (2009; Haralick et al., 1973), using textural descriptors taken from the grey level concurrence matrix (GLCM), it is possible to distinguish between three different types of disease—nodule, emphysema, and frosted glass—and a healthy pattern. The GLCM texture descriptors used by Shimo et al. (2010) may also be used to assess the health of the lungs. The identification of particular disorders, such as nodules (Ayres et al., 2010; Silva and Oliveira, 2010) or emphysema (Ayres et al., 2010), is also addressed in certain articles (Felix et al., 2007, 2011).

Medical photos can now be reliably analyzed using machine learning (ML) algorithms. Scalability, automation, and ease of implementation are all advantages of ML-based techniques in healthcare contexts. The categorization of photos with strikingly comparable attributes is a typical use of ML-based image analysis. In order to properly classify images, this technique makes use of image segmentation, image feature extraction from the segmented area in the spatial or frequency domain,

and machine learning to choose the best classification method. "Many ML-based approaches to diagnosing COVID-19 medical photos have been suggested in the last few years." By using a deep learning network called DenseNet 121 to categorize CT pictures of COVID-19 imaging tests as positive or negative, Wang et al. achieved an accuracy rate of 81.24%. A small dataset of lung ultrasonography (LUS) pictures was used by Roy et al. to study the applicability of deep learning models to evaluate COVID-19 infections (only 11 patients). ResNet-18 classifier model for COVID-19, Lung infection, and normal lung-lesion segmentation in CT images was suggested by Zhang and colleagues and produced a 92.49% accuracy rate.

In this study, we anticipated that an ML-based classifier could accurately discriminate COVID-19 patients' CXR pictures from those of patients with other types of lung infection. In order to produce an optimal model with a collection of synthetic features that can accurately discriminate COVID-19 photos from non-COVID-19 instances with a precision of 94%, we applied a dimensionality reduction strategy. In our approach, characteristics may be identified and extracted from the whole CXR picture without the need for any segmentation of chest lesions.

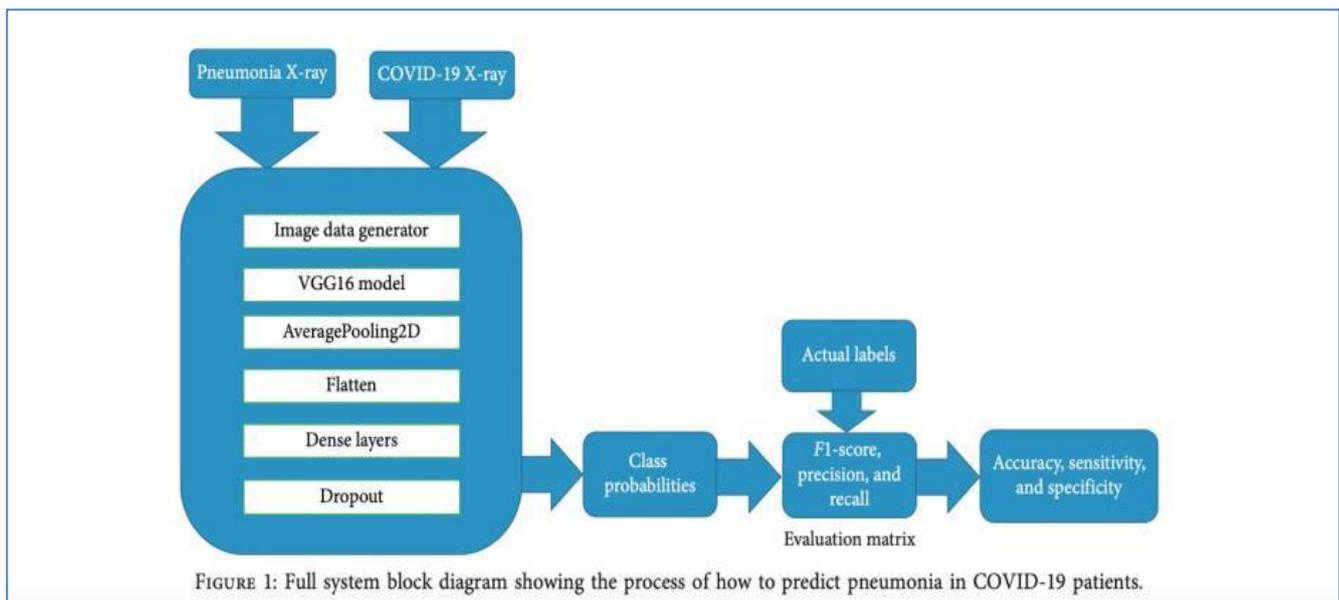
As a result of this new quantitative marker, we are able to eliminate segmentation mistakes and lower the computing cost of our final model. Our work shows that basic ML-based classification may be used as an auxiliary to other tests to aid in the differentiation of CXR pictures from COVID-19 patients. It is also possible that we may use this method in any future viral epidemic to quickly classify the CXR pictures of patients.

CT imaging has been presented as a vital alternative for COVID-19 screening, which is particularly successful as a companion approach to RT-PCR, and was able to conduct quick prediction compared to RT-PCR. CT scans for non-COVID-19 purposes, such as before elective surgeries and nervous system evaluations, are particularly beneficial in testing for COVID-19 infection, according to studies. Patients who have deteriorating respiratory problems or comparable issues and are tested negative for RT-PCR may exhibit clinical evidence that is similar to COVID-19 positivity when utilizing CT imaging. It is possible, according to preliminary study, that CT pictures include a large number of infection-related signs, although this has not yet been confirmed. As a result, radiologists will have certain difficulties when attempting to use CT scans to detect COVID-19 infections. In addition, analyzing CT scans visually takes time, particularly when working with a large number of patients or in a study of this scope. As a consequence of privacy issues, most CT scans used for diagnostic reasons are not freely available, which means that the findings from neural network training on one dataset cannot be repeated or utilized in other institutions. Since COVID-19 CT pictures are not openly available, the development of sophisticated artificial intelligence systems for improved CT detection on COVID-19 testing is severely hindered by this fact. We discussed how transfer learning can improve the performance of convolutional neural networks on COVID-19 testing using CT images in light of the urgent need for solutions to deal with the COVID-19

pandemic and based on recent efforts among researchers to promote open-source data and open access. We found that pre-trained models trained on larger out-of-domain datasets have better performance in COVID-19 detection. Our approach outperforms the architecture revealed by a machine-driven design exploration process utilizing generative synthesis in every assessment criteria.

2. METHODOLOGY USED

In the first few paragraphs of this part, we offer a block diagram of the system and a flowchart that summarizes this study effort. The next section focuses on several deep learning characteristics and models used to correctly identify Lung infection in COVID-19 patients. Figure 1 depicts the model design in its entirety. It includes the VGG16 pre-trained model, an image data generator, AveragePooling2D, flatten, dense, and dropout layer layers. Before identifying Lung infection, we additionally preprocessed the data and did data augmentation. F1-score, precision, and recall were evaluated in a classification matrix to determine accuracy and sensitivity as well as specificity.



Working Steps of Lung infection Prediction. As a result of using a deep learning algorithm, the model is able to identify Lung infection with more accuracy. The flowchart and the subsequent working phases offer a thorough understanding of this research activity.

- Dataset collection: Kaggle’s dataset was used for this research project.
- Data processing and augmentation: after collecting images from the dataset, the noise in the X-ray images is removed and cleaned. After that, the data are resized.
- Feature extraction: VGG16 model was used to build a Lung infection prediction model.

- Data split in training and test set: the data were split into 80% training and 20% testing data. Then the data are fed into the VGG16 model for training.
- Data test: after training the data into the model, the test data were used for prediction.

Finally, the results and conclusions were utilised to compute accuracy, develop a classification matrix, and establish sensitivity and specificity from the results.

2.1 The Proposed CNN Model

Feature extractors and classifiers make up the CNN model employed in this research. As a consequence, CNN models offer the highest accurate results in picture classification, with the fewest mistakes. Fig. 3 depicts the overall CNN architecture used in the current investigation.

A CNN architecture known as VGG16 is used to classify images from a big dataset. Weights are loaded into ImageNet, however the fully connected layer (FC) head is not linked. Following a sequence of convolutional layers, the FC layer contains three completely connected layers. The first two tiers had 4096 channels each. IISVRC classification is performed 1000 ways by the third layer's 1000 channels. The output layer has the same number of nodes as the softmax layer, the last layer. To solve multiclass classification issues, this layer is often utilized.

In deep learning, an epoch is a complete iteration of the data. In machine learning, the epoch parameter specifies how many times a model is applied to the training dataset. In each epoch of training, the internal model parameters are updated by a sample from the dataset. There may be more than one batch in an era. It is known as the batch gradient descent learning method when one batch is used [22]. There were 25 epochs in this study's coding phase.

To ensure that the model's parameters are constantly being improved, the hyperparameter batch size specifies how many samples must be sent through before the parameters may be updated. One or more samples may be used to create predictions based on the batch size. After that, the predicted results are compared to the actual ones obtained at the conclusion of the batch, and the difference between the two is determined. For example, the current model moves down the error gradient [22] to improve itself from this mistake. A total of 16 batches were used in this investigation, with an initial learning rate of $1e^{-3}$. Using flattening, multidimensional data is transformed into one-dimensional data so that it may be inserted into the next layer. The output of the convolutional layers into a one-dimensional feature vector was flattened in this work. It is subsequently sent to the categorization layer [23] for further processing. A pooling layer [24] with an average pool size of was also used (4, 4).

When it comes to directing medical choices and giving information that may help us better understand the patterns of infection by COVID-19, access to first-hand CT imaging and clinical data is crucial. This is especially true when it comes to the worldwide spread of the pandemic. To make the worldwide battle against COVID-19 more efficient, a comprehensive database of CT scans and associated clinical signs should be created. According to the section on related work, a number of datasets for COVID-19-related research have been created and are available for use by researchers, clinicians, and data scientists alike. There is currently a lack of patient demographic variety in the COVIDx-CT dataset, despite the fact that this dataset is clearly bigger than many other CT datasets utilized in the literature in the field of COVID-19 testing. However, as only data from China's provinces are accessible in this CT image collection (COVIDx-CT), symptoms of COVID-19 in the CT scans may not be applicable for patients outside of this country. Deep neural networks will become more diverse and comprehensive as the quantity and diversity of patients increases, allowing for more generalizability and applicability in a variety of clinical settings throughout the globe. Previous researchers created the COVIDx CT-2A and COVIDx CT-2B datasets by meticulously processing and categorizing CT pictures of patients from a variety of CT equipment, solutions, and validation capabilities. COVIDx CT-2A includes 194,922 pictures from 3,745 patients ranging in age from 0 to 93 years, with a median age of 51 years. The number of CT slides in a patient's CT scan varies widely. It's an image issue since the CT slides are used as input pictures for COVID-19 detection.

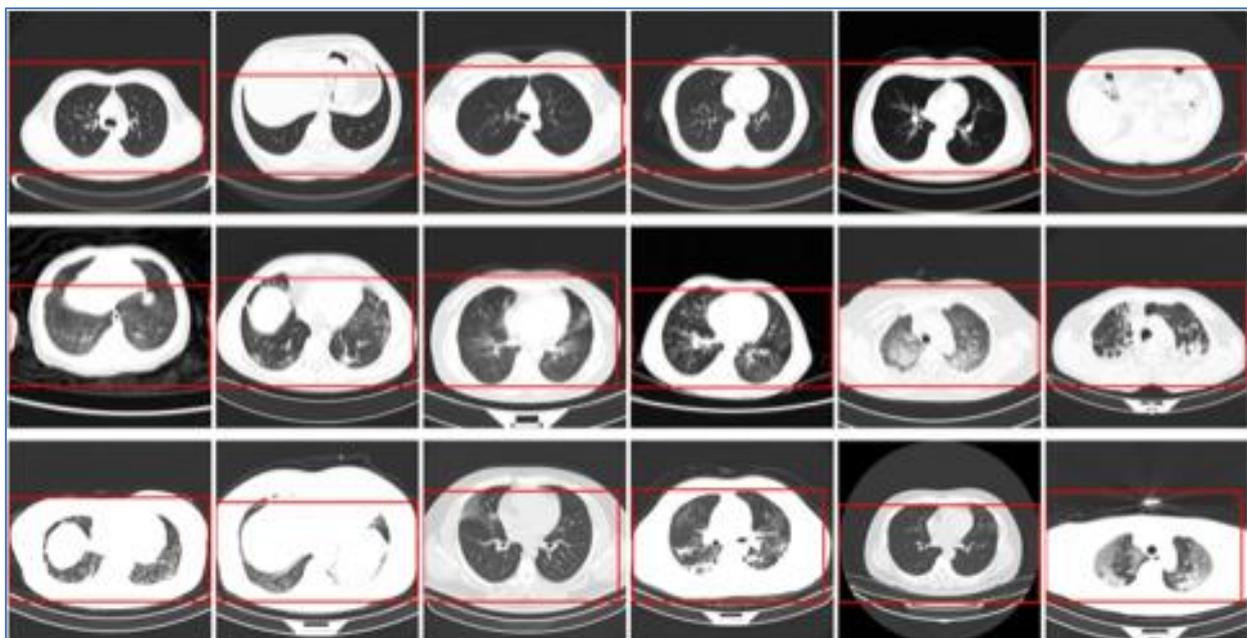


Figure 2. Relevant examples of CT images in COVIDx CT-2. The red frames refer to the marker frames of CT images in the dataset. The first line shows the normal controls, the second line shows the cases with CP, and the third line shows the cases with NCP infected by SARS-CoV-2.

Classification problem. The CT images are provided as 512×512 pixels. The sources of input for the images in COVIDx CT-2A are as follows:

- National Center for Bioinformation (CNCB) (China)
- National Institutes of Health Intramural Targeted Anti-COVID-19 (ITAC) Program (countries unknown)
- Negin Radiology Medical Center (Iran)
- Liyuan Hospital and Wuhan Union Hospital of Tongji Medical College of Huazhong University of Science and Technology (China)
- COVID-19 CT lung and infection segmentation project, annotated and verified by Nanjing Drum Tower Hospital (China)
- Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI) (countries unknown)
- Open access online collaborative radiology resource (Radiopaedia) (countries unknown)

The Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Department of Health Care of Moscow, added weak validation (MosMed) to COVIDx CT-2B, building on COVIDx CT-2A. (Russian Federation). To see whether adding weak validation (i.e., results without employing RT-PCR and radiological testing) training data will improve the model's performance, this validation set was created. Dataset breadth and variety may be increased by doing this validation. We used COVIDx CT-2A for COVID-19 testing in the pre-sent research because of the comparison to earlier working models and the availability of data. COVIDx CT-2A dataset includes CT scans for three kinds of lung infections: new coronavirus lung infection (NCP) infected by SARS-CoV-2, common lung infection (CP), and normal controls. To make our models easier, we tweaked photographs from the database. With this in mind, we used an automated cropping method that removes the backdrop from photographs in order to eliminate any possible biases from the models (as shown by the red frames in Fig. 2). We were able to identify the ground glass opacity (GGO), lung consolidation, and even the existence of white lung infection in the groups of CP and NCP by comparing different kinds. Even for radiologists, it may be difficult to tell the difference between those with a typical lung infection and those infected with SARS-CoV-2 since the visual distinctions between the two infections are so minor. Figure 2 shows the distribution of various kinds of infections and photos in training, test, and validation sets.

Model Selection. Previous researchers created the deep convolutional neural network COVID-Net-CT for COVID-19 testing based on CT images using the design exploration mode formed by machine-driven creation. This architecture was subsequently used to create the COVID-Net CT-2 that followed. We used ResNet-v2, which is a modified version of ResNet, in our experiment. Next, we made a substitution.

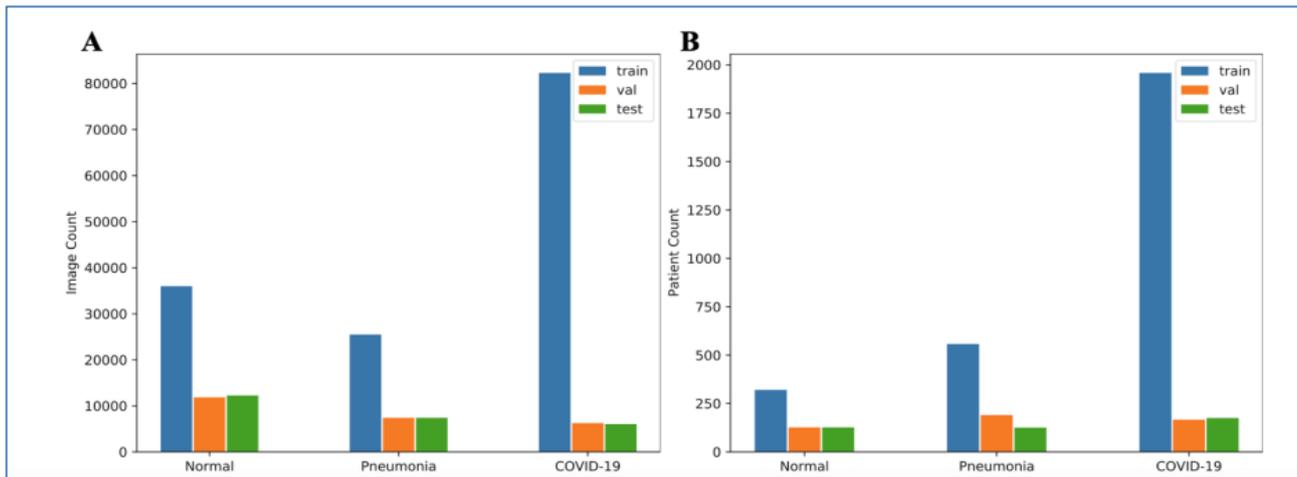


Figure 3. COVIDx CT-2A Dataset Training, Testing and Validation Dataset.
(A) CT Image Distribution. (B) Patient Distribution.

3. OBJECTIVES

This project exploits the convergence of imaging research and system development at the NLM and NIH policy objectives in global health. The following are project objectives:

1. Improve the state-of-the-art in CXR image analysis using automated software. Automatically identify the presence of pulmonary illnesses such as tuberculosis (TB) and other pertinent diseases in digital CXRs, resulting in appropriate screening discrimination and a level of confidence in the finding.
2. Create deployable screening software to assist field clinical officers in making decisions at the point of treatment and to help radiologists organise their workload.
3. Recognizing the seriousness of lung illnesses and the scarcity of radiological facilities in western Kenya, AMPATH has created software that can be installed on a self-powered mobile X-ray vehicle that may be used in rural regions. Their team uses NLM-developed software to check for the existence of lung illnesses and other disorders by taking chest x-rays of the general public.
4. To investigate the influence of these hyperparameters on the outcomes, we employed different training steps, resolutions with and without mixup, and observed that a higher resolution and a suitable number of training steps are useful in improving model performance. Because the model already produces outstanding results when the data is adequate, adopting mixup has no influence on the outcomes.

5. We investigated the influence of initialised parameters on model performance using five alternative parameter initialization procedures in the models. Our findings show that various pre-training factors have an impact on the ultimate performance of fine-tuned models. The model can be more successfully generalised by using a bigger out-of-field dataset for pre-training.

6. We show that our models based on transfer learning are superior than those based on structural design by comparing our findings to those from earlier research, and that our models attain state-of-the-art performance. Furthermore, we tested our model's performance in the absence of downstream data and found that it was still capable of recognising COVID-19.

7. We used visualisation to study the process underlying the COVID-19 testing model in order to improve clinical decision-making.

3.1 Research Gaps

Until far, research on this issue has been confined to one illness, but here we have used two training models to identify numerous diseases in one go, resulting in many detections with better accuracy owing to the availability of a large dataset.

3.2 Problem Definition and Scope

The main issue is that specialist physicians are unable to diagnose a patient's problem using just their x-ray since there is no clear visibility in them, which may lead to ignorance of the patient's impending problem. So, the present project's goal is to achieve that visibility by using image processing methods, comparing them to the given data collection, and generating conclusions that can subsequently be checked by professionals. This saves time and eliminates the possibility of human mistake.

4. PERFORMANCE ANALYSIS

Ist Part

We evaluated the performance of the proposed model based on different metrics: accuracy, recall, sensitivity, specificity, and precision. The metrics are evaluated by various parameters in the confusion matrix, such as true positive (TP),

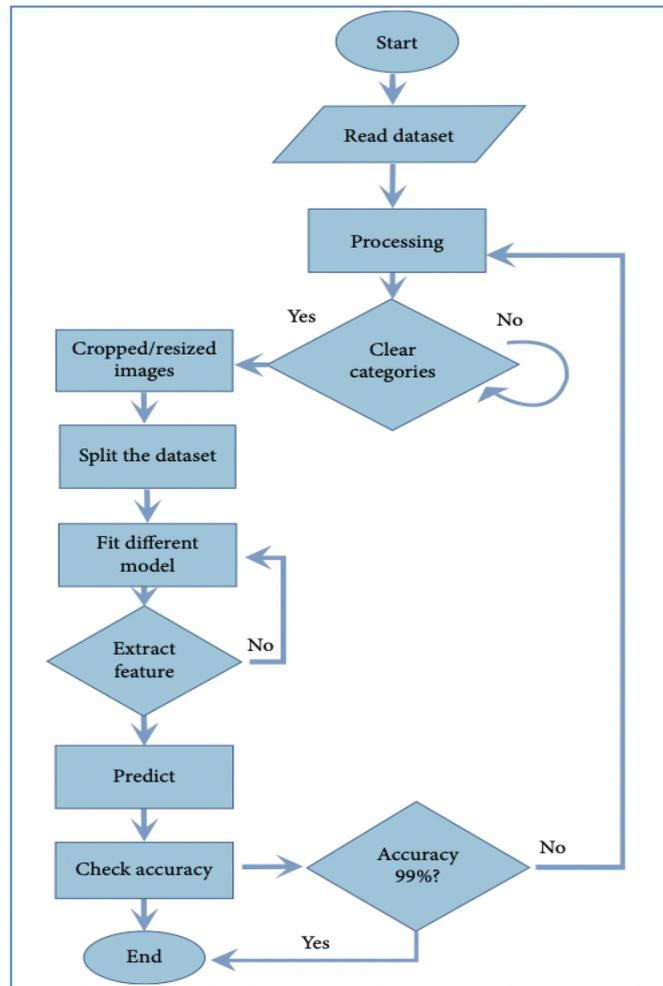


Figure 4: Flowchart of The Working Steps to Predict Lung Infection in COVID-19 Patients.

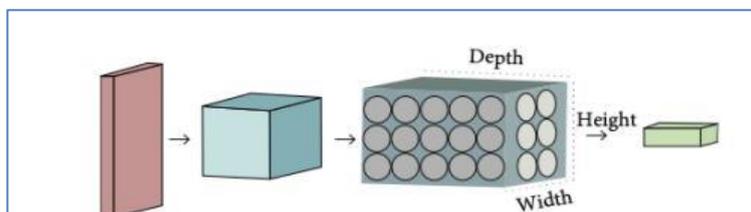


Figure 5: A general 3-Dimensional CNN Structure

True negative (TN), false positive (FP), and false negative (FN). The metrics are defined as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

It is the fraction of positive instances that can be reliably anticipated by sensitivity that is most important. This statistic assesses the model's ability to anticipate. For the calculation of sensitivities, use the following formula:

$$\frac{\text{Sensitivity}}{\text{Recall}} = \frac{TP}{TP + FN} \quad (2)$$

Specificity was employed to clarify the fraction of real negative situations, which was accurately anticipated. The capacity of a model to predict true-negative instances of a particular category is measured by its specificity. To understand the results, these metrics were applied to each classification model. The following is the formula for determining specificity:

$$\text{specificity} = \frac{TN}{FP + TN} \quad (3)$$

Precision demonstrates the performance of the model on the test data. It shows the number of models predicted correctly from all positive classes. This should be as high as possible:

$$\text{precision} = \frac{TP}{TP + FP} \quad (4)$$

The whole code for predicting lung infection in COVID-19 patients was written in Anaconda's Jupyter notebook. An open-source platform, the libraries required to perform this study may be accessible via Jupyter notebook.

IInd Part

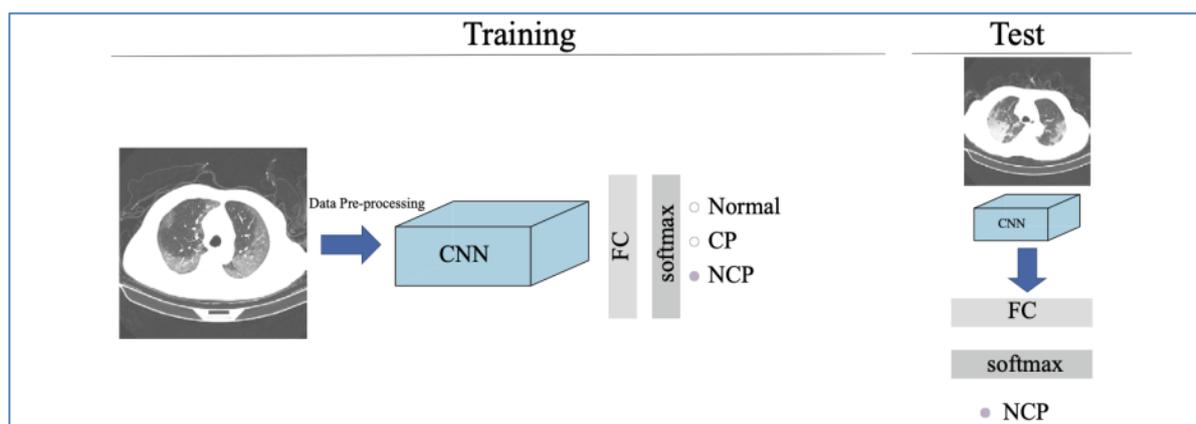


Figure 6: A General Flowchart of Deep Learning Based COVID-19 Diagnosis System.

Hyperparameter Settings For Training. Fig. 1 depicts the overall flowchart of the COVID-19 deep learning diagnostic system. There are two main parts to the system. The training data are utilised to update the model parameters, and the performance of the created model is evaluated using test data in the training portion. It is possible to extract a feature and then identify class labels based on that feature in the test section by utilising the model. As a final step, the model's performance is evaluated using a variety of criteria such as accuracy and sensitivity.

Algorithm 1 provides the pseudocode for fine-tuning the Convolutional Neural Network (CNN) and achieving accuracy. Each iteration, we used b CT scans to compute the gradient and updated the network parameters. In contrast to the prior conventional training approach, we restricted the training steps rather than the period of repetition. A stochastic gradient descent (SGD) algorithm was utilised, with a learning rate of 0.003, momentum of 0.9, and a batch size of 64 as parameters. The final input to the suggested model was a $512 \times 512 \times 3$ picture that had been reordered in the RGB colour space. According to the annotated cropping frame, the training set photos were first customised according to the annotated cropping frame, and then they were adjusted to the dimensions of the annotated cropping frame, randomly segmented to 480×480 pixels, and then normalised. After cropping the photographs according to the annotation, we simply scaled them to fit on the screen. Experiments with 10,000 training steps yielded the best results. A warm-up of the learning rate was done before we lowered it three times at a rate of 10x throughout training to polish the model. In the Parameter Sensitivity section, you'll find further information. Finally, we augmented our data using mixup (Eq. (1)).

$$\begin{cases} \tilde{x} = \lambda x_i + (1 - \lambda)x_j, \\ \tilde{y} = \lambda y_i + (1 - \lambda)y_j, \end{cases}$$

Here, x_i and x_j are the initial input vectors, while y_i and y_j are the labels. Through mixup, we obtained new vectors and labels. With strong convergence qualities for convex optimization problems, we choose to utilise cross entropy as the loss function for calculating loss using cross entropy.

$$\begin{aligned} \text{loss}(x, \text{class}) &= -\log \left(\frac{\exp(x[\text{class}])}{\sum_j \exp(x[j])} \right) \\ &= -x[\text{class}] + \log \left(\sum_j \exp(x[j]) \right) \end{aligned}$$

where $x \in \mathbb{R}^{N \times C}$ is the output of the model, $\text{class} \in \mathbb{R}^N$ is the label of the CT imaging and $0 \leq \text{class}[i] \leq C-1$.

Algorithm 1: Fine-tuning the neural network and obtaining the detection accuracy

```

Input : Initialize(net)
Output : Test Acc

1 while  $l$  do
2   for  $batch = 1, 2, \dots, \#images/b$  do
3     images  $\leftarrow$  uniformly random sample  $b$  images;
4      $x, y \leftarrow$  progress(images);
5      $lr \leftarrow$  get_lr( $lr$ , step);
6      $x, y_a, y_b \leftarrow$  mixup_data( $x, y$ );
7      $z \leftarrow$  forward( $net, x$ );
8      $\ell \leftarrow$  loss( $z, y_a, y_b$ );
9     grad  $\leftarrow$  backward( $z, y_a, y_b$ );
10    update( $net, grad$ );
11    step++;
12    if  $lr = None$  then
13      break ;
14    end
15  end
16 end
17 Acc = Eval( $net, test_k$ )

```

5. RESULT

For predicting Lung infection in COVID-19 patients, the findings of the suggested approach are discussed above. We began by enhancing the photos using the Keras image data generator. Rotation was maintained at 15 degrees and fill mode was set to closest. Label Binarizer () is then used to encode the labels one at a time. Data were then divided into 80 percent training and 20 percent testing data, as shown in Table 1 Base and head models were constructed; the latter was modified using AveragePooling2D [25], flattening, thick dropout layers and eventually developing the entire model. "Following this, the Adam optimizer was used to construct the entire model, and the testing phase was forecast." After analysing chest X-ray pictures, 91.69% of COVID-19 patients were correctly diagnosed with Lung infection. Deep learning models employ a variety of parameters in the training phase, when they increase their accuracy with each epoch. Training loss and validation loss reduced throughout the course of seven epochs, resulting in the system achieving its maximum level of accuracy. Each epoch, the model learns more about the world as it gathers input from numerous



sources. It is therefore possible to improve the accuracy of the VGG16 model with less loss in training and validation.

It is shown in Table 3 that the suggested model begins running epoch after epoch after training all of the pictures in the training dataset. In spite of the inclusion of 25 epochs, the model had to halt after 7 epochs since it had become overfitted. At epoch 1, training accuracy was 89.17%, loss.

TABLE 1: Dataset split for the model’s training and testing.

Dataset	Number of images
Training	5144
Validation	1288
Testing	1288

TABLE 2: Model training and validation parameters.

Parameters	Value
Initial learning rate	$1e - 3$
Batch size	16
Shuffling	Each epoch
Optimizer	Adam
Max epochs	25
Execution environment	GPU

TABLE 3: Training loss, training accuracy, validation loss, and validation accuracy on different epochs.

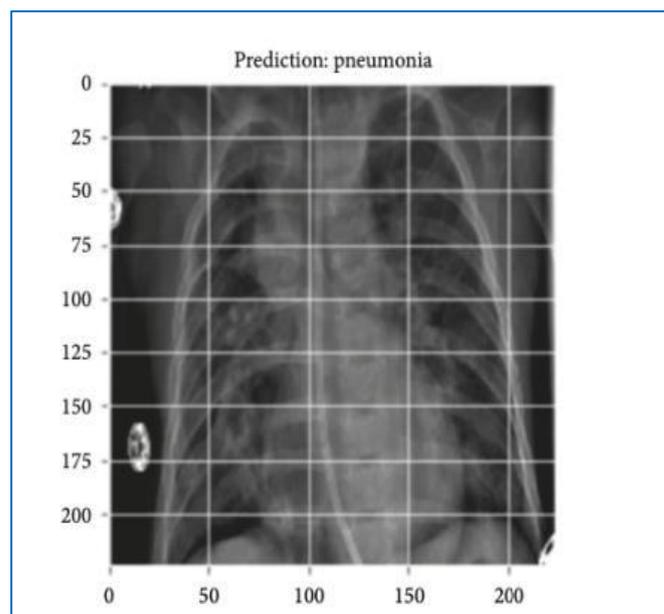
	Loss	Accuracy	VAL_Loss	VAL_Accuracy
0	0.286430	0.891719	0.232570	0.911491
1	0.269631	0.900272	0.223635	0.920031
2	0.255112	0.903188	0.219480	0.917702
3	0.261968	0.905132	0.221931	0.920807
4	0.250308	0.904743	0.202796	0.923137
5	0.245776	0.905327	0.212749	0.922360
6	0.236960	0.910964	0.216345	0.916925

28.64%, validation accuracy 91.14%, and validation loss 23.25%. After seven epochs, training accuracy was 91.1%, loss 23.69%, validation accuracy 91.69%, and validation loss 21.63%.

X-ray pictures predicted to be Lung infection-infected after fitting the suggested model for testing are shown in Figure 4. It's important to note that the testing process has been redesigned. There are three separate classes of chest X-ray pictures predicted by the model, and the predicted and the actual classes are identical. A model developed in this paper was able to correctly classify all of the classes it tested.

There is a categorization report included in Tables 4 and 5. As can be seen, the precision of COVID-19 prediction is 99 percent, recall is 81 percent, and F1-score is 89 percent, whereas the precision of normal case detection is 83 percent, recall is 91 percent, F1-score is 87 percent, and the precision of Lung infection prediction is 95 percent, recall is 93 percent, and F1-score is 94 percent. This means that the suggested method has a lower risk of incorrectly identifying Lung infection compared to normal and COVID-19 instances when it comes to Lung infection prediction.

In order to make this conclusion clearer, the figure presents TP and TN Lung infection prediction cases, as well as FP and FN Lung infection prediction cases. 798 patients were discovered to have a lung infection, although only 387 of them had pneumonia in the traditional sense. There were 57 incorrect diagnoses of lung infection and 46 false diagnoses of non-lung illness as a result of the model's faulty assumptions, which were both produced simultaneously. To address this, this model's sensitivity and



Proposed Model Prediction on A Test Image.

Specificity values are calculated, which are 95.92% and 100%, respectively. In the categorization of COVID-19-induced Lung infection, the suggested technique has a 91.69 percent accuracy rate after completing epochs.

Figures illustrate the model's training and validation accuracy and loss values. The graphs illustrate the relationship between training and validation losses and improvements in training and validation precision. Epochs on the x-axis and precision on the y-axis are shown in this graph. Training and validation accuracies were low when the epoch started, and the loss was substantial. There were also decreases in training and validation loss as epochs grew, however this was not the case for the first epoch. For example, after 7 iterations, the training and validation accuracies are near to 90%; training losses and validation losses are close to 2% each.

Using the suggested concept, patients require just a chest X-ray and a mobile app to diagnose Lung infection, rather than seeing a doctor and spending enormous amounts of money on clinical testing and exams. As a consequence, the findings of this research will have a significant impact on the disadvantaged and rural populations. After using the app to diagnose a Lung infection, patients may simply follow their doctor's instructions and take the prescribed medication.

Deep learning-based approaches for COVID-19 classification using chest X-rays have made great progress. Researcher's primary emphasis was on COVID-19, normal, and lung infection detections. The lack of huge datasets is a major issue in the assessment of the suggested models. "Transfer learning approaches have been used to address the issue of tiny datasets." The models are trained on the ImageNet dataset before they are used. Additionally, COVID-19 detection use ensemble learning approaches, which integrate predictions from many models to generate more accurate findings. This reduces the model's generalisation error and variance, which improves the model's prediction accuracy.

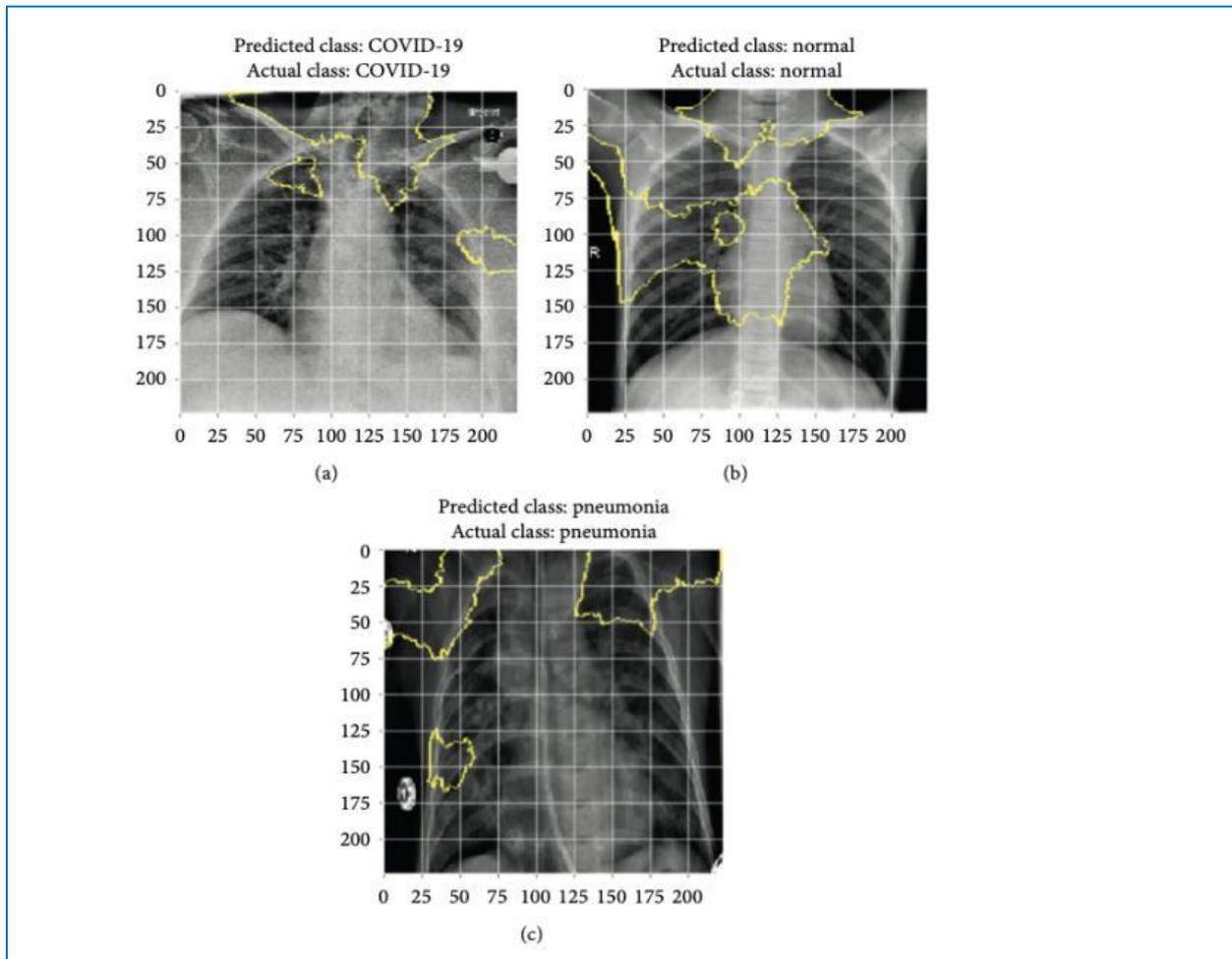


TABLE 4: Classification report of the pneumonia prediction model.

	Precision	Recall	F1-score	Support
0	0.99	0.81	0.89	116
1	0.83	0.91	0.87	317
2	0.95	0.93	0.94	855
Accuracy			0.92	1288
Macro avg	0.92	0.89	0.90	1288
Weighted avg	0.92	0.92	0.92	1288

TABLE 5: Performance score on test data.

Disease	Precision (%)	Recall (%)	F1-score (%)
COVID-19	99	81	89
Normal	83	91	87
Pneumonia	95	93	94

Chest X-ray pictures have also been utilised to identify COVID-19 patients via the application of domain adaptation. By adapting feature adversarial and implementing a new classifier strategy, Zhang et al. used this technique to handle data. The method showed significantly better

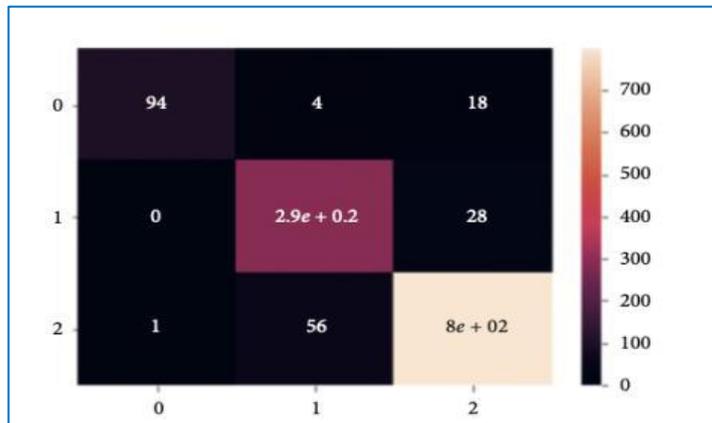
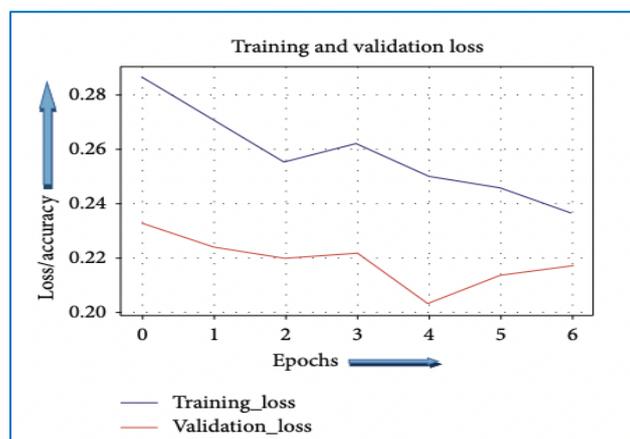
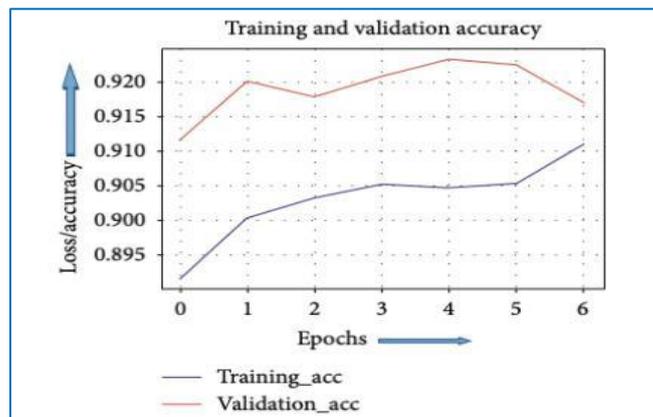


Figure: Confusion Matrix of The Lung Infection Prediction Model.



Second Part

In this section, we investigate the model performance in testing for COVID-19. Specifically, we endeavor to address the following questions:

- How are different hyperparameters, including various resolutions, training steps, and mixup, used to affect the model performance?
- How do different weight initializations trained from the different datasets affect result?
- Can we obtain a satisfactory result with the proposed model with limited CT images?
- How can we understand the decisions made by the deep convolutional neural network to assist in clinical?

Test Performance. The models were trained using the Hyperparameter settings for training section training setup provided in that section. Table 1 summarises the findings and compares them to the most sophisticated techniques currently available. A random initialization model and pre-training techniques based on ILSVRC-2012 and ImageNet-21k will be introduced in the Impact of parameter initialization section under the names Random, Bit-S, and Bit-M. A more sophisticated COVID-Net CT-2 L model was compared to ours. It is shown in Table 1 that the accuracy of our Bit-S and Bit-M models relying on transfer learning increased by 1.12 percentage points compared to COVID-Net CT-2 L model. Random initialization's accuracy was 3.60 percent greater than COVID-Net CT-1's, demonstrating that when compared to models utilising structure space search, our model also has outstanding performance. The distribution of CT image representations after dimensionality reduction can be seen in the figure, which demonstrates how clearly the various groups are separated. In the confusion matrix in Fig., we demonstrate that even though radiologists may sometimes fail to distinguish between CP and NCP,

Table Accuracy of The COVIDx CT-2A Benchmark Datasets.

Model	Methods	Accuracy(%)
COVID-Net CT-1	Structure design	94.5
COVID-Net CT-2 L	Structure design	98.1
COVID-Net CT-2 S	Structure design	97.9
Random (ours)	Transfer learning	97.9
Bit-S (ours)	Transfer learning	98.8
Bit-M (ours)	Transfer learning	99.2

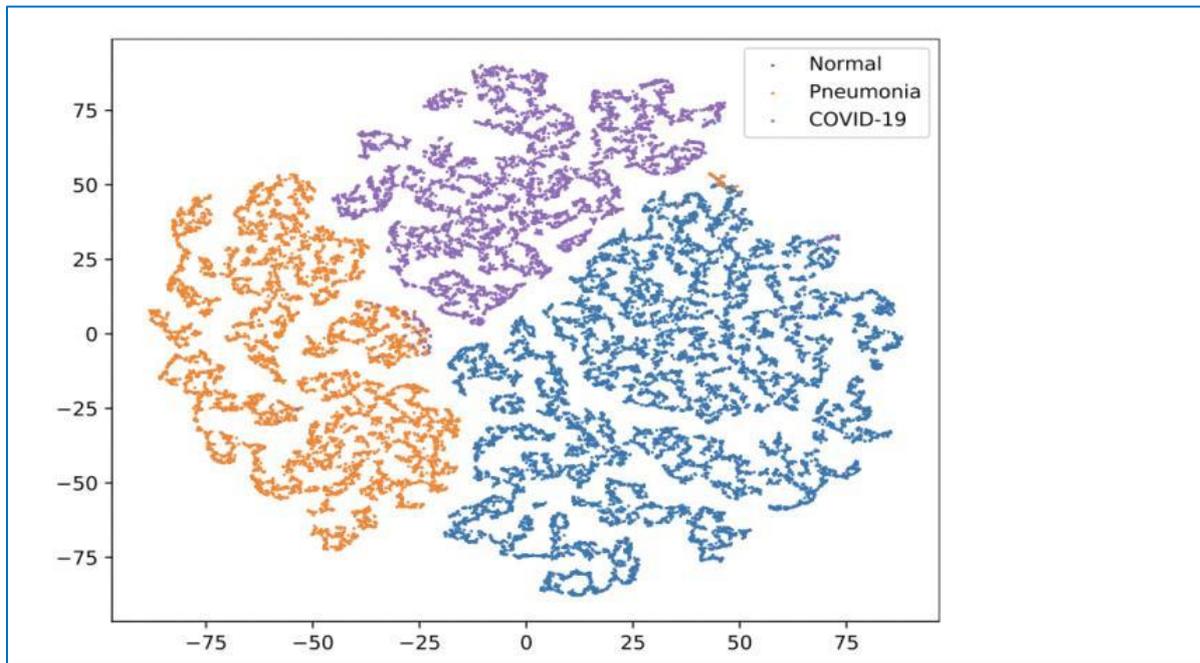


Figure: Distribution of Characteristics of CT Images After Dimensionality Reduction With T-SNE. Each Node Refers to A Different CT Image, The Colour Reflects the Information on Categories, And the Meaning of The Colour Is Defined in The Legend.

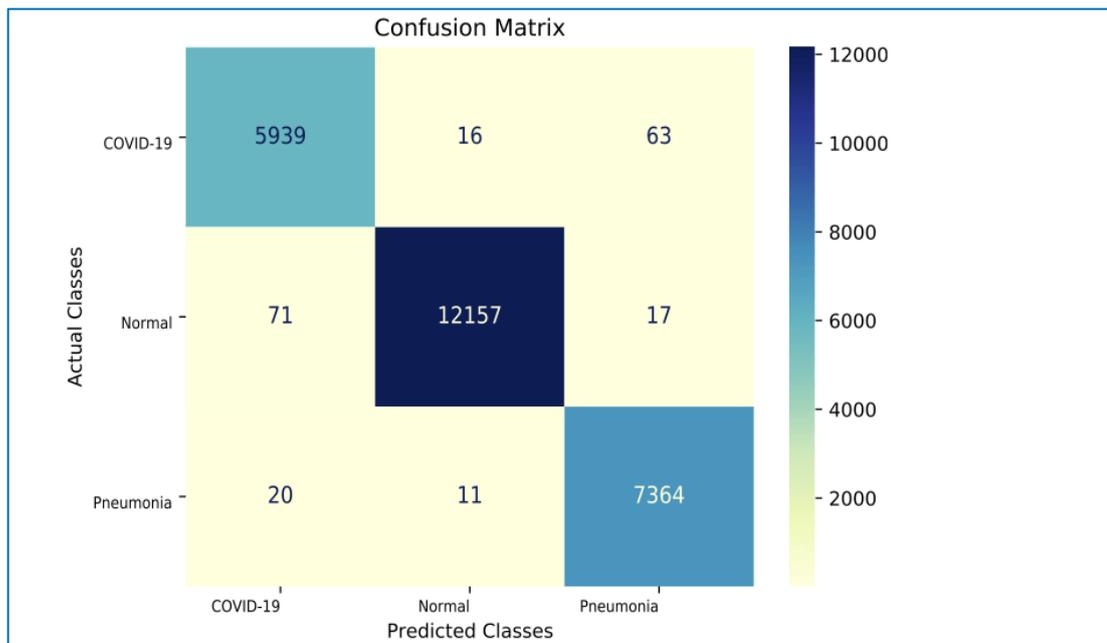


Figure: Confusion Matrix for COVID-19 Testing Using Bit-M Fine-Tuned Model. The Color Bar Indicates the Intensity of Normalization.

Our model provides accurate classifications. It was decided to include sensitivity (Sn) and specificity (Sp) as well as positive and negative predictive values in the models for better quantitative analysis. When it came to COVID-19 sensitivity (98.7 percent), positive predictive value (98.5 percent), specificity (99.5 percent), negative predictive value (98.7 percent) we found that the BiT-M model based on transfer learning produced the best results in the industry (99.6 percent). Pre-training the model on a bigger dataset from outside of the domain allows the model to gain more general information, which improves its performance. High sensitivity ensures that patients with COVID-19 infection are not overlooked because of false negatives, and high PPV guarantees that false positives do not burden the health care system. Our Bit-M model's high specificity and NPV guarantee that in the great majority of instances, COVID-19 negative predictions are in fact genuine negatives, and the prediction results for COVID-19 negative patients are real and credible. The problem of treating false.

Table Sensitivity, PPV, Specificity, and NVP of The Test Data in COVIDx CT-2A Benchmark Datasets.

Network	Sensitivity (%)			PPV (%)			Specificity (%)			NPV (%)		
	Normal	CP	NCP	Normal	CP	NCP	Normal	CP	NCP	Normal	CP	NCP
COVID-Net CT-1	98.8	99.0	80.2	96.1	90.2	97.6	96.3	95.7	99.4	98.9	99.6	94.2
COVID-Net CT-2 L	99.0	98.2	96.2	99.4	97.2	96.7	99.5	98.8	99.0	99.1	99.3	98.8
COVID-Net CT-2 S	98.9	98.1	95.7	99.3	97.0	96.4	99.3	98.8	98.9	99.0	99.2	98.7
Random(ours)	97.9	98.6	96.9	98.9	97.5	96.4	99.0	99.0	98.9	98.1	99.4	99.1
Bit-S(ours)	99.0	99.3	97.9	99.6	97.9	98.4	99.6	99.1	99.5	99.1	99.7	99.4
Bit-M(ours)	99.3	99.6	98.7	99.8	98.9	98.5	99.8	99.6	99.5	99.3	99.8	99.6

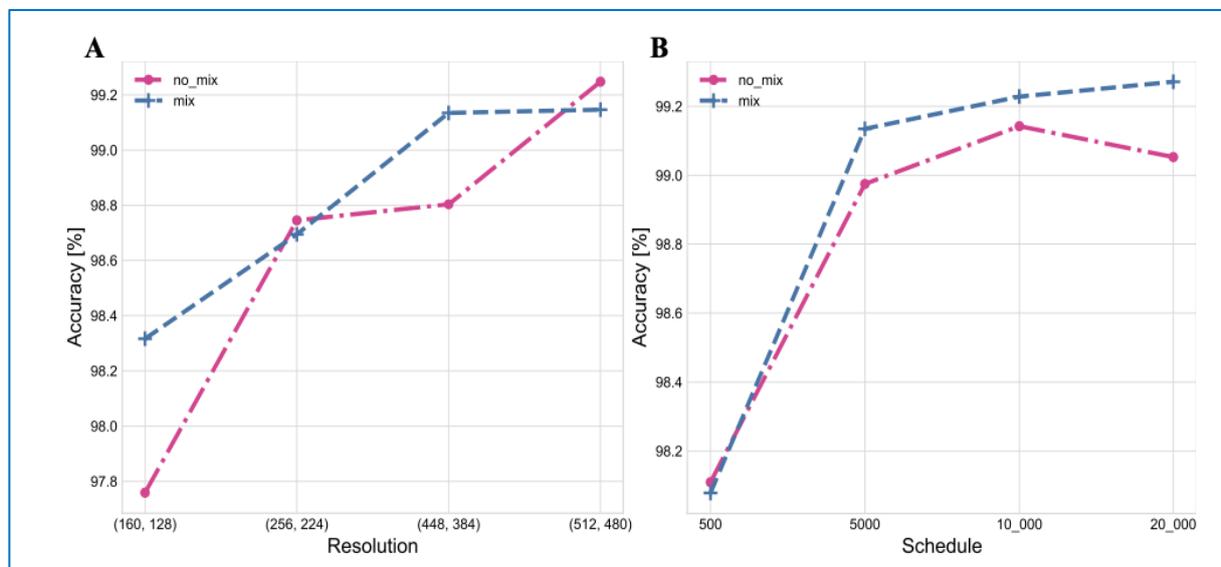


Figure Test Accuracy of COVIDx CT-2 with Various Hyperparameters. (A) Resolution. (B) Schedule.

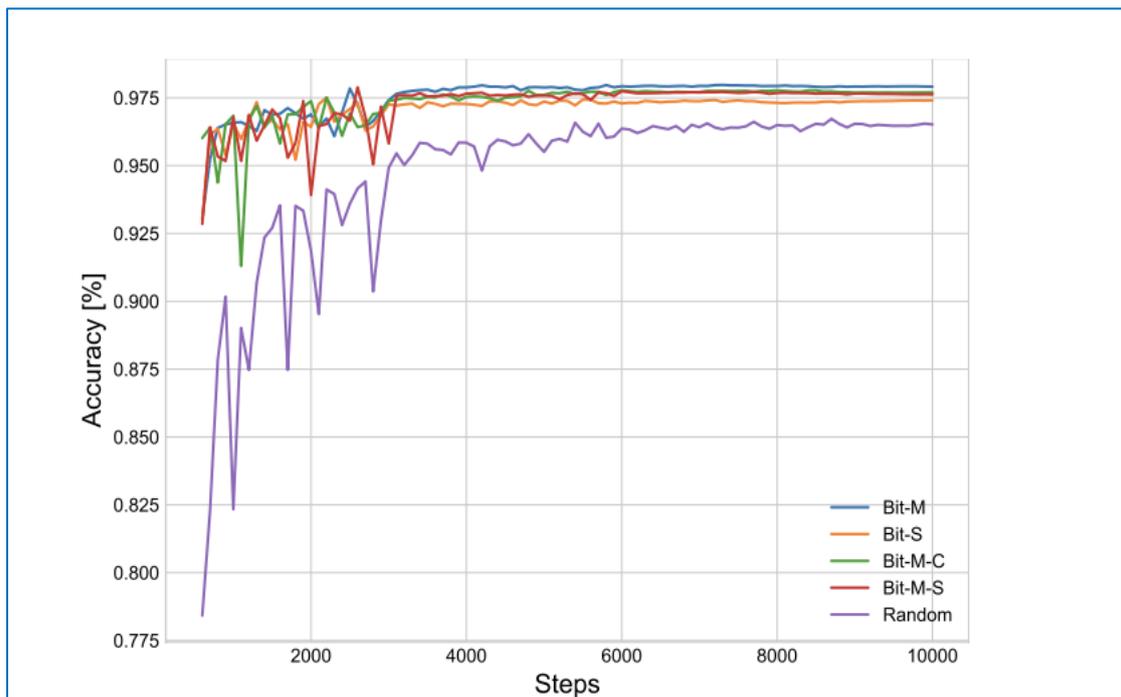
positives and false negatives is equivalent, specifically, To classify a COVID-19 positive patient as negative would put the patient at risk of re-entering the community thinking he or she is free of COVID-19 and so leading to the spread of the illness throughout that population. Over diagnosing COVID-19 negatives as positives puts an additional strain on the healthcare system and creates a state of fear among the general people. If a negative person is diagnosed as positive, they may experience psychological distress.

Hyperparameters Sensitivity. A variety of hyperparameters, such as training steps, picture resolutions, and whether or not to utilise mixup, are examined in this section. We experimented with four different combinations of training phases and input resolutions. Each doublet has a first value that represents the adjustment scale during training and a second value that determines the size of random cropping during both training and testing for CT pictures (160, 128), (256, 224), (448, 384), and (512, 480). As for the length of the training project's duration, we utilised the following ranges: [100; 200; 300; 400; 500; 500; 1,500; 3,500; 4,500; 5,000]. To begin, the warmup step has a number of steps and an end step. The remainder of the step nodes have a learning rate that decays by ten times. With and without mixup, the test accuracy is shown in the following figure. According to the study findings, a better resolution may improve identification, which indicates that crisper CT pictures carry more clinical information. It is possible to enhance accuracy by increasing the training step, however this is less effective when it surpasses 10,000. For (512, 480) and 10,000 training steps between Figs. A and B, the findings show that the accuracy rates are completely opposite (The hyperparameter settings for the experiments are the same). Random sampling is to blame for this occurrence. It suggests that the model's performance is not boosted by the mixup since the data is already sufficiently rich.

Impact of Parameter Initialization. ResNet50x1 models that have already been trained to a high level were utilised to study the effect that upstream pre-training may have on fine-tuning performance. Random denotes that the model's parameters were set at random. ImageNet-21k, a publicly available dataset of 14,200,000 pictures and 21,000 categories, was used to train BIT-M. There's a chance that the photographs include more than one label. BIT-S was pre-trained on the ILSVRC-2012 variant from ImageNet.

Categorization Accuracy of Test and Validation Sets with Different Weight Initialization.

Weight initialization	Val	Test
Random	96.5	97.7
Bit-S	97.4	98.8
Bit-M	97.9	99.3
Bit-M-C	97.7	99.2
Bit-M-S	97.6	98.9



Validation Accuracy Curves of Various Initialized Models.

Which include 1,280,000 images and 1,000 categories. BIT-M-S was first pre-trained on the ImageNet-21k dataset and then fine-tuned on ILSVRC-2012. CIFAR-10, an image database that comprises 60,000 pictures (32 by 32 pixels) from 10 different categories, was used to train BIT-M-C. A prior research used out-of-domain data to train the weight initialization. Setting the training step to 10,000 and using mixup allowed us to do a fair comparison, while the other variables remained the same. Table depicts the effect of weighting initialization. Due to random sampling and model parameter initialization, the experiment outcomes were somewhat different from the first time around. We discovered that the parameter trained on ImageNet-21k performed better in generalisation than that trained on ILSVRC-2012. Out-of-field datasets will have no effect on this performance, though. Finally, we computed the test results for every 100 steps, as shown in the figure (fig.). Pre-trained models on ImageNet-21k (BIT-M, BIT-M-S, and BIT-M-C) outperformed the ILSVRC-2012 initialised weighting in the assessment with the test set at subsequent stages (BIT-S). Training with a bigger dataset resulted in better generalizability, as shown by this study.

Influence Of the Size of Labeled Training Data on Model Performance. Randomly picked photos from each category were utilised to assess the models' performance on the tiny downstream datasets that would be used in a real-world setting. As a test of the trained model's ability to correctly identify test sets, we randomly selected 50, 100, 500, and 1,000 samples for training. In Fig., the results of the testing were shown. Predictions based on the whole training set, including CT-2L, CT-2S, and CT-1,

are shown in the right-hand histogram. We discovered that a smaller number of labelled photos improved the test accuracy for BIT-M. It was already more accurate (94.8 percent) than the experimental result using CT-1 when 100 photos from each category were picked (94.5 percent). The accuracy (98.0%) was on par with CT-2S when a sample of 1,000 pictures was used (97.9 percent). That our transfer learning models can still work effectively with a short sample provides credence to their enormous potential. Pre-training on big, non-field datasets may nevertheless guarantee great performance in the situation of limited training data, according to this study.

Qualitative Analysis of Covid - 19 Testing of The Model. The network's decision-making behaviour cannot be explained by performance measures, which are useful for model assessment. "Specifically, we used the Grad-CAM visualisation method in COVID-19 testing in order to better understand which properties of CT images are crucial for diagnostic accuracy and hence improve clinical decision-making." As demonstrated in Fig., we first cropped the images using the detection frame (introduced in Hyperparameter settings for training Section), enlarged them to 480×480 pixels, and used

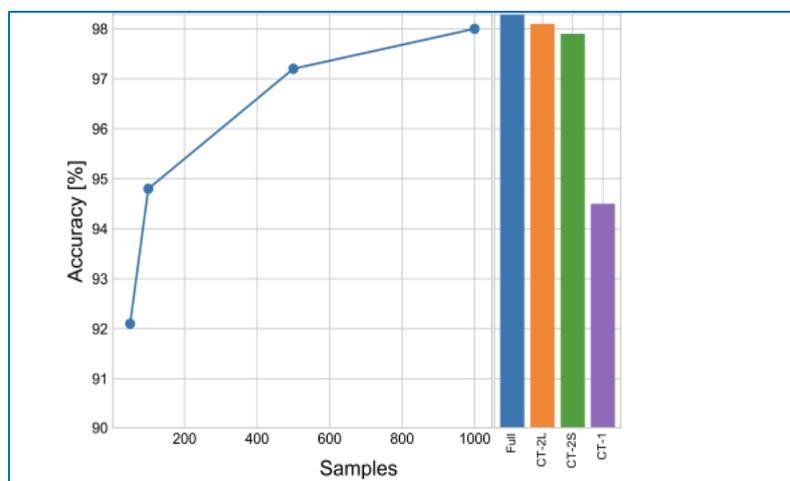


Figure: Impact of The Number of Each Category of Images in The Training Set on The Performance of The Model on The Test Set.

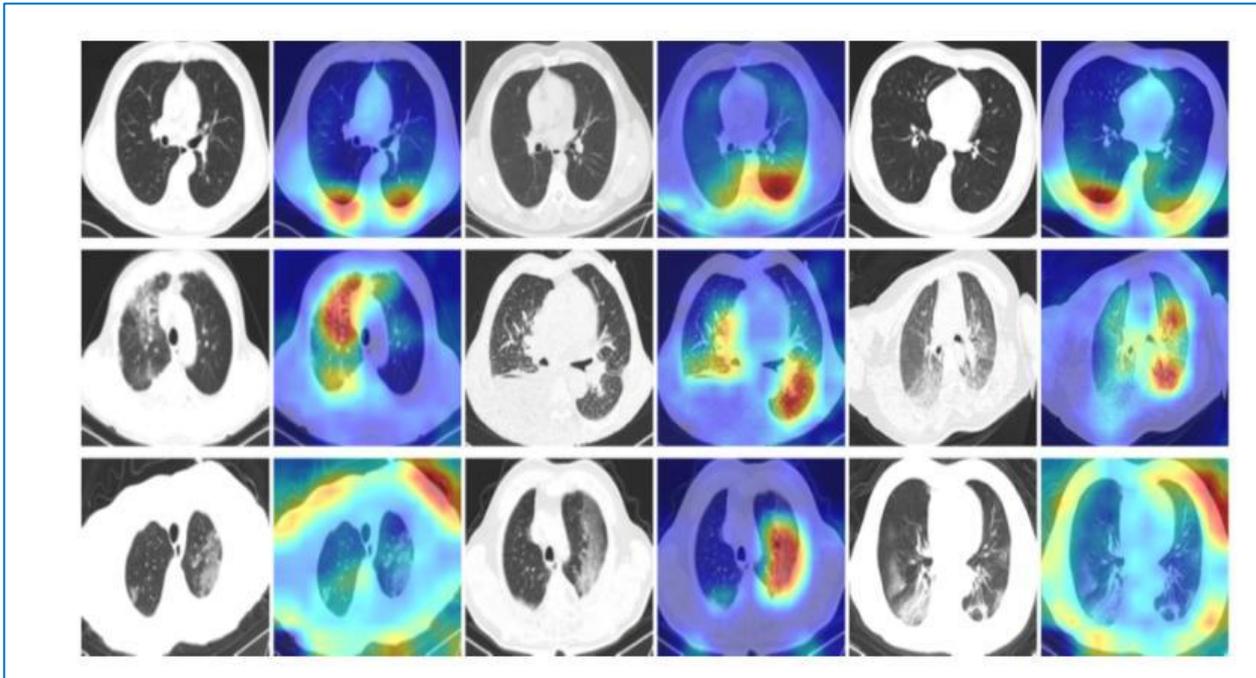


Figure: Grad-CAM Visualization of Bit-M. The First Line Depicts the Normal Case, The Second Line the Case Of CP, And the Third Line the Case of COVID-19 (NCP) Due to Infection by SARS-CoV-2. The Model Using CT Images as Input Source Have Yielded Accurate Predictions.

Grad-CAM for visual explanation. All of the model's predictions based on the CT images in Fig. match the actual detection outcomes. It's not uncommon for the model's performance to be comparable to that of typical human visual cognition. This is especially true in the case of CP, when the model is able to accurately depict the diseased portions of the lungs. To make a more informed and confident judgement, the radiologist might use the Grad-CAM colour visualisation technique. The bottom portion of the model receives greater attention in the norm situation. Although the first and third CT pictures (third row in Fig.) were able to identify NCP owing to SARS-CoV-2, the model was more interested in the texture around the edges. In order to better understand how the model detects COVID-19 and which traits they deem most diagnostic, such a visual heuristic that is distinct from human visual perception warrants additional investigation. Finding these characteristics will help explain the power of the model in COVID-19 testing and help clinicians uncover novel visual indications for COVID-19 infections that may be used in manual screening using CT scans.

7. CONCLUSION

COVID-19-induced pneumonia may be identified utilizing a two-stage deep residual learning algorithm using lung X-ray images. Using the VGG16 model, the model was able to distinguish between COVID-19 patients and those with COVID-19-induced pneumonia. The model predicted pneumonia with an average accuracy of 91.69 percent, a sensitivity of 95.92 percent, and a specificity of 100 percent. As a result, training loss is reduced while precision is bolstered. To limit the transmission of infection to frontline workers and establish primary diagnoses for COVID-19 patients, parallel testing may be employed in the present circumstance. To put it another way, the suggested approach may be utilized to identify pneumonia patients. By altering the hyperparameters and transfer learning combinations, CNN architecture may be improved in the future by researchers. Pneumonia and COVID-19 may be better modelled using an enhanced, more complicated network topology. A model trained on ImageNet21k was shown to be highly generalizable to CT pictures in our work by using transfer learning to COVID-19 testing using CT images. We also explored the effects of different startup values on the findings. The suggested methodology has a 99.2 percent success rate in identifying COVID-19 incidents. According to the metrics that we've outlined, our model outperforms the neural architecture search model. Patients who test negative for COVID-19 are more likely to be diagnosed as negative, reducing the number of patients who are misdiagnosed as positive and easing the strain on the healthcare system. Additionally, we tested the model's performance with little data and found that it performed well. In the real world, where huge and varied datasets may not be easily accessible, this indicates that our methodology is still useful with limited data. Finally, we used the Grad-CAM visualization approach to better understand and describe the suggested deep learning model for COVID-19 testing. Using a radiologist's interpretation of the data, the model validates its performance in a way compatible with the model's performance. It is possible to discover new visual indications for clinical professionals to use in subsequent manual screening by analyzing normal and NCP CT pictures. Our models were shown to be successful in COVID-19 testing by conducting trials. The severity of COVID-19 will be closely monitored in the future, and we'll use CT scans to gather more data that may help us fight the pandemic, as well. After doing explanatory analyses on the models, we will uncover critical CT image properties and ease screening by clinical physicians by identifying COVID-19's detection process. Although the system performs well on publicly available datasets, the work is still in the theoretical research stage and the models have not been tested in clinical practice. Because of this, we will conduct clinical trials of our system and contact with doctors to learn more about how they use it and what they think of the models. Consequently, it will allow us to develop the models even more in future research.

REFERENCE

1. Liu, J.; Liao, X.; Qian, S.; Yuan, J.; Wang, F.; Liu, Y.; Wang, Z.; Wang, F.S.; Liu, L.; Zhang, Z. Community transmission of severe acute respiratory syndrome coronavirus 2, Shenzhen, China, 2020. *Emerg. Infect. Dis.* **2020**, *26*, 1320–1323. [CrossRef] [PubMed].
2. Ghinai, I.; McPherson, T.D.; Hunter, J.C.; Kirking, H.L.; Christiansen, D.; Joshi, K.; Rubin, R.; Morales-Estrada, S.; Black, S.R.; Pacilli, M.; et al. First known person-to-person transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in the USA. *Lancet* **2020**, *395*, 1137–1144. [CrossRef].
3. Chen, N.; Zhou, M.; Dong, X.; Qu, J.; Gong, F.; Han, Y.; Qiu, Y.; Wang, J.; Liu, Y.; Wei, Y.; et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study. *Lancet* **2020**, *395*, 507–513. [CrossRef].
4. Long, C.; Xu, H.; Shen, Q.; Zhang, X.; Fan, B.; Wang, C.; Zeng, B.; Li, Z.; Li, X.; Li, H. Diagnosis of the Coronavirus disease (COVID-19): RRT-PCR or CT? *Eur. J. Radiol.* **2020**, *126*, 108961. [CrossRef] [PubMed].
5. Fang, Y.; Zhang, H.; Xie, J.; Lin, M.; Ying, L.; Pang, P.; Ji, W. Sensitivity of chest CT for COVID-19: Comparison to RT-PCR. *Radiology* **2020**, *296*, 200432. [CrossRef].
6. Ai, T.; Yang, Z.; Hou, H.; Zhan, C.; Chen, C.; Lv, W.; Tao, Q.; Sun, Z.; Xia, L. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases. *Radiology* **2020**, *296*, E32–E40. [CrossRef].
7. Kanne, J. Chest CT Findings in 2019 Novel Coronavirus (2019-nCoV) Infections from Wuhan, China: Key Points for the Radiologist. *Radiology* **2020**, *295*, 16–17. [CrossRef].
8. Santosh, K. AI-driven tools for coronavirus outbreak: Need of active learning and cross-population train/test models on multitudinal/multimodal data. *J. Med. Syst.* **2020**, *44*. [CrossRef].
9. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [CrossRef].
10. Sufian, A.; Ghosh, A.; Sadiq, A.S.; Smarandache, F. A survey on deep transfer learning to edge computing for mitigating the COVID-19 pandemic. *J. Syst. Archit.* **2020**, *108*, 101830. [CrossRef].
11. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. ImageNet: A Large-scale Hierarchical Image Database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Miami, FL, USA, 20–25 June 2009; pp. 248–255.

12. Alshazly, H.; Linse, C.; Barth, E.; Martinetz, T. Deep Convolutional Neural Networks for Unconstrained Ear Recognition. *IEEE Access* **2020**, *8*, 170295–170310. [CrossRef]
13. You, Y.; Li, J.; Reddi, S.; Hseu, J.; Kumar, S.; Bhojanapalli, S.; Song, X.; Demmel, J.; Hsieh, C.J. Large batch optimization for deep learning: Training BERT in 76 minutes. In Proceedings of the International Conference on Learning Representations (ICLR), Addis Ababa, Ethiopia, 26 April–1 May 2020.
14. Maaten, L.v.d.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
15. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 618–626.
16. Shin, H.C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [CrossRef] [PubMed].
17. Rajpurkar, P.; Irvin, J.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.; Shpanskaya, K.; et al. CheXNet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv* **2017**, arXiv:1711.05225.
18. Li, L.; Qin, L.; Xu, Z.; Yin, Y.; Wang, X.; Kong, B.; Bai, J.; Lu, Y.; Fang, Z.; Song, Q.; et al. Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy. *Radiology* **2020**, *296*, E65–E71. [CrossRef] [PubMed].
19. Chen, J.; Wu, L.; Zhang, J.; Zhang, L.; Gong, D.; Zhao, Y.; Hu, S.; Wang, Y.; Hu, X.; Zheng, B.; et al. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: A prospective study. *Sci. Rep.* **2020**, *10*, 19196. [CrossRef].
20. Xu, X.; Jiang, X.; Ma, C.; Du, P.; Li, X.; Lv, S.; Yu, L.; Ni, Q.; Chen, Y.; Su, J.; et al. A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering* **2020**. [CrossRef].
21. Silva, P.; Luz, E.; Silva, G.; Moreira, G.; Silva, R.; Lucio, D.; Menotti, D. COVID-19 detection in CT images with deep learning: A voting-based scheme and cross-datasets analysis. *Inform. Med. Unlocked* **2020**, *20*, 100427. [CrossRef].
22. Apostolopoulos, I.D.; Mpesiana, T.A. COVID-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Phys. Eng. Sci. Med.* **2020**, *43*, 635–640. [CrossRef].

23. Elaziz, M.A.; Hosny, K.M.; Salah, A.; Darwish, M.M.; Lu, S.; Sahlol, A.T. New machine learning method for image-based diagnosis of COVID-19. *PLoS ONE* **2020**, *15*, e0235187. [CrossRef].
24. Wu, X.; Hui, H.; Niu, M.; Li, L.; Wang, L.; He, B.; Yang, X.; Li, L.; Li, H.; Tian, J.; et al. Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: A multicentre study. *Eur. J. Radiol.* **2020**, *128*, 109041. [CrossRef] [PubMed].
25. Gianchandani, N.; Jaiswal, A.; Singh, D.; Kumar, V.; Kaur, M. Rapid COVID-19 diagnosis using ensemble deep transfer learning models from chest radiographic images. *J. Ambient Intell. Humaniz. Comput.* **2020**. [CrossRef] [PubMed].
26. Wang, L.; Lin, Z.Q.; 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*, 19549. [CrossRef] [PubMed]
27. Hasan, K.; Alam, A.; Elahi, T.E.; Roy, S.; Wahid, S.R. CVR-Net: A deep convolutional neural network for coronavirus recognition from chest radiography images. *arXiv* **2020**, arXiv:2007.11993.
28. Wang, Z.; Liu, Q.; Dou, Q. Contrastive Cross-site Learning with Redesigned Net for COVID-19 CT Classification. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 2806–2813. [CrossRef] [PubMed].
29. Mukherjee, H.; Ghosh, S.; Dhar, A.; Obaidullah, S.M.; Santosh, K.; Roy, K. Deep neural network to detect COVID-19: One architecture for both CT scans and Chest X-rays. *Appl. Intell.* **2020**. [CrossRef].
30. Farooq, M.; Hafeez, A. COVID-ResNet: A deep learning framework for screening of covid19 from radiographs. *arXiv* **2020**, arXiv:2003.14395.