RESEARCH ARTICLE

X-ray versus computerized tomography (CT) images for detection of COVID-19 using deep learning [version 1; peer review: awaiting peer review]

Nandish Siddeshappa1, Tejashri Varur2, Krithika Subramani2, Siddhi Puranik2, Niranjana Sampathila2

1Manipal School of Information Sciences, Manipal Academy of Higher Education, Manipal, Karnataka, 576104, India
2Biomedical Engineering, Manipal Institute of Technology, MAHE, Manipal, Karnataka, 576104, India

First published: 17 Dec 2021, 10:1292
https://doi.org/10.12688/f1000research.74839.1
Latest published: 17 Dec 2021, 10:1292
https://doi.org/10.12688/f1000research.74839.1

Abstract

Background: The recent outbreak of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the disease corresponding to it (coronavirus disease 2019; COVID-19) has been declared a pandemic by the World Health Organization. COVID-19 has become a global crisis, shattering health care systems, and weakening economies of most countries. The current methods of testing that are employed include reverse transcription polymerase chain reaction (RT-PCR), rapid antigen testing, and lateral flow testing with RT-PCR being used as the golden standard despite its accuracy being at a mere 63%. It is a manual process which is time consuming, taking about an average of 48 hours to obtain the results. Alternative methods employing deep learning techniques and radiologic images are up and coming.

Methods: In this paper, we used a dataset consisting of COVID-19 and non-COVID-19 folders for both X-Ray and CT images which contained a total number of 17,599 images. This dataset has been used to compare 3 (non-pre-trained) CNN models and 5 pre-trained models and their performances in detecting COVID-19 under various parameters like validation accuracy, training accuracy, validation loss, training loss, prediction accuracy, sensitivity and the training time required, with CT and X-Ray images separately.

Results: Xception provided the highest validation accuracy (88%) when trained with the dataset containing the X-ray images while VGG19 provided the highest validation accuracy (81.2%) when CT images are used for training.

Conclusions: The model, VGG16, showed the most consistent performance, with a validation accuracy of 76.6% for CT images and 87.76% for X-ray images. When comparing the results between the modalities, models trained with the X-ray dataset showed better performances than the same models trained with CT images. Hence, it
can be concluded that X-ray images provide a higher accuracy in detecting COVID-19 making it an effective method for detecting COVID-19 in real life.

**Keywords**
COVID19, Deep Learning, Artificial Intelligence, Classification, Prediction

**Corresponding author:** Niranjana Sampathila (niranjana.s@manipal.edu)

**Author roles:**
- **Siddeshappa N:** Conceptualization, Methodology, Project Administration, Supervision, Validation, Writing – Review & Editing
- **Varur T:** Conceptualization, Data Curation, Formal Analysis, Methodology, Writing – Original Draft Preparation
- **Subramani K:** Formal Analysis, Writing – Original Draft Preparation
- **Puranik S:** Formal Analysis, Methodology, Software, Writing – Original Draft Preparation
- **Sampathila N:** Conceptualization, Formal Analysis, Project Administration, Supervision, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** The author(s) declared that no grants were involved in supporting this work.

**Copyright:** © 2021 Siddeshappa N et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Siddeshappa N, Varur T, Subramani K et al. X-ray versus computerized tomography (CT) images for detection of COVID-19 using deep learning [version 1; peer review: awaiting peer review] F1000Research 2021, 10:1292 https://doi.org/10.12688/f1000research.74839.1

**First published:** 17 Dec 2021, 10:1292 https://doi.org/10.12688/f1000research.74839.1
Introduction

The recent coronavirus disease 2019 (COVID-19) pandemic instigated by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has become a global catastrophe with 257,469,528 registered cases and 5,158,211 deaths registered worldwide as of 23rd November 2021 [WHO Coronavirus (COVID-19) Dashboard, Accessed on 24th November 2021]. The virus is primarily spread in 3 different ways- inhaling air that has been contaminated with fine droplets and or aerosol particles consisting of the virus, deposition of exhaled droplets that contain the virus onto exposed mucous membranes which may be the mouth, nose or the eyes, and touching mucous membranes after coming in contact with contaminated or soiled surfaces and or objects [CDC, Accessed on 24th November 2021]. Therefore, it is essential to maintain and/or improve hygienic habits such as washing one’s hands for at least 20 seconds, and sanitizing hands with sanitizers that contain a minimum alcohol content of 60%.4 Avoiding extended stays in enclosed spaces with improper ventilation is also crucial to protect oneself [CDC, Accessed on 24th November 2021]. An individual contracting the disease may depict a wide range of symptoms: 83%-98% of the patients have fever, 59%-78% of the patients exhibit dry cough and 3%-29% of the affected patients may require admission into an intensive care unit (ICU) for the management of a myriad of complications, which include extreme pneumonia, hypoxemic respiratory failure, acute respiratory distress syndrome (ARDS), multiple organ failure etc.5 On the other hand, it is observed that a small percentage of the affected patients presented no symptoms.6 Affected individuals that exhibit symptoms are often diagnosed to have COVID-19 using the reverse transcription polymerase chain reaction test, commonly referred to as the RT-PCR test. Samples are collected from the nasal and or pharyngeal passages or the throat. Sputum is also often used as a sample. Genetic material is isolated from the sample and multiplied. Fluorescent indicators are added to the multiplied genetic material that glow in the presence of the SARS-CoV-2’s genetic material [Cleveland Clinic, Accessed on 24th November 2021]. Despite being the gold standard of testing, it provides an average accuracy rate of 63%7–9 as the detection of the virus’s genetic material depends on when the patient takes the test and the amount of the virus present in the sample collected. A recovered patient may continue to get a positive RT-PCR test result due to small amounts of genetic material being present, even though they have recovered and cannot spread the virus. On the other hand, a recently infected patient may receive a negative test result due to smaller amounts of virus present in the sample [Cleveland Clinic, Accessed on 24th November 2021]. To overcome these drawbacks, doctors and scientists alike are researching and coming up with various other methods to diagnose COVID-19 that provide higher accuracy rates. One such method is using radiological images and deep learning techniques.6,7 Often, the imaging modality used for this purpose is computerized tomography (CT)10–11 and/or X-ray.12–15 It is found that COVID-19 patients depict similar features on radiographic images such as bilateral, multifocal ground glass opacities with peripheral or posterior distribution. This is predominantly found in the lower lobes of the lungs in the early stages. In the later stages, these features can be found in areas of pulmonary consolidation.16 Unfortunately, these symptoms overlap with the symptoms of bronchopneumonia, hence, making it difficult for the radiologists to tell COVID-19 apart from other diseases that may affect the lungs.17 Deep learning techniques are used, in hopes of revealing underlying features, that often go unnoticed which may aid in confirming diagnosis.

There have been several studies that have employed deep learning methods to diagnose COVID-19 from radiological images- a total of 8 studies with 4 using CT and 4 using X-ray as their imaging modalities. They have been presented in a consolidated form in Table 1 and Table 2.

Past works have used either X-ray image or CT images to train and test their models but not both. There are no studies conducted, to the best of our knowledge, that compare the performances of the models when they are trained with CT and X-ray images. This paper presents a comparative study using 3 non-pre-trained models and 5 pre-trained models to find which imaging modality provides better accuracies in detecting COVID-19 using deep learning techniques.

Methodology

This section elucidates the dataset used, the kind of network employed and the working of this study that took place from January 2021 to July 2021. The general workflow used in the study is shown in Figure 1. Non-pre-trained models were developed for the study and were rejected due to their low validation accuracies. To compare the results and efficacy and to deduce the best of the two imaging modalities, 5 pre-trained models namely VGG16, VGG19, ResNet50, InceptionV3, and Exception were loaded and fine-tuned according to the needs of this study. The pre-trained models that were selected were chosen primarily based on their accuracy (where higher accuracy yields better results), the speed of the model training and predictions (where faster training is preferred). Ideally, we would use a model that is smaller in size with a higher accuracy and lower time requirement but realistically any model that has deeper layers would make it slower to execute and a smaller model would have much lower accuracy. Hence, we have experimented with the available models to find the best suited option to detect COVID-19.
For the purposes of training our networks, the dataset named ‘Extensive COVID-19 X-Ray and CT Chest Images Dataset’, curated and augmented by Walid El-Shafai and Fathi Abd El-Samie was used. The dataset consists of COVID-19 and Non-COVID-19 folders for both X-ray and CT images. There were 5500 non-COVID-19 and 4044 COVID-19 cases belonging to the X-ray folder. For CT images, there were 2628 non-COVID-19 and 5427 COVID-19 images, which brings the total number of images to 17,599. Sample images from the Extensive COVID-19 X-Ray and CT Chest Images Dataset is shown in Figure 2.
A convolutional neural network (CNN) is one of the many networks that fall under neural networks. It primarily works on the basis that the input will be an image or is made up of images. Therefore, the model is extensively employed in the field of pattern recognition in images, making it an effective method for this study.

In this study, Keras API (version 2.4.3), an open-source software library providing an interface for neural networks, is used to build and train the models. We trained and tested a total of 8 models- 3 models that were built from scratch and 5 pre-trained models. Each model was trained with the CT dataset and the X-ray dataset separately and the obtained results were compared. Description of each of the models are mentioned below. The code is available from GitHub and archived with Zenodo.21

Model 1: Consists of 4 alternate convolution and max pooling layers, 32 filters, 4 × 4 kernels and Relu activation function present in each convolution layer. The model consisted of 5 epochs and has 23,618 total parameters, all of which are trainable.
Model 2: This model consists of a total of 6 convolution layer and 3 max pooling layers along with a dropout layer with varying number of filters (from 32 till 256), $3 \times 3$ kernels in both layers. This model has 5 epochs and 333,154 total parameters and all of them are trainable.

Model 3: This model was an attempt to fine tune Model 2 by varying the number of filters (from 32 till 128) and adding a dropout layer with a dropout rate of 0.3. This model has 462,338 total parameters and all of them are trainable. For experimentation, this model has 10 epochs.

VGG 16: VGG-16 is one of the many pre-trained CNN models that is considered as an excellent vision model and is one of the extensively used image-recognition architectures [Deep Convolution Networks VGG16 for Image Recognition in Keras, Medium, Accessed on 24th November 2021]. This model consists of 3 epochs and has a total of 14,766,866 parameters out of which 50,178 are trainable. Figure 3 shows the model architecture that was used for this study.

VGG 19: The difference between VGG-16 and VGG-19 is the number of layers -16 and 19 layers, respectively. The architecture of this network is shown in Figure 3. Compared to VGG16, VGG19 performs slightly better but it requires more memory due to more layers [Understanding the VGG19 Architecture, Opengenus Foundation, Accessed on 24th November 2021]. The model has 20,074,562 parameters out of which 50,178 are trainable and 3 epochs. Figure 4 shows the model architecture that was used for this study.

ResNet 50: ResNet 50 grants users the provision to train tremendously deep neural networks of about 100 and more layers successfully without facing the problem of vanishing gradients [Understanding and Coding a ResNet in Keras, Towards Data Science, Accessed on 24th November 2021]. The general architecture of the model is presented below in Figure 5. The model has 23,788,418 total parameters, including 200,706 trainable parameters and used 3 epochs.

Inception V3: It is a commonly utilized image recognition model that has been proven to have high accuracies when tested on a large dataset, consisting of asymmetric and symmetric blocks of convolution and max pooling layers, dropouts, and fully connected layers [Advanced Guide to Inception V3 on Cloud TPU, Google Cloud, Accessed on 24th November 2021]. The model consisted of 21,905,18 parameters out of which 102,402 are trainable with 3 epochs. Figure 6 depicts the model’s architecture.

---

Figure 3. Model architecture of pre-trained model - VGG16.
Xception: This model is derived from the Inception V3 model, with the inception modules replaced by depth wise separable convolutions. It slightly outperforms the Inception V3 due to its efficient use of model parameters. Figure 7 shows the architecture of the model. The model has 21,062,186 total parameters of which 200,706 are trained and 3 epochs.
Once the models were trained, parameters such as number of epochs, training accuracy and loss, validation accuracy and loss, true positive and negative, false positive and negative, prediction accuracy, sensitivity, and training times were recorded.

**Procedure**

The open-source dataset, Extensive COVID-19 X-Ray and CT Chest Images Dataset, was downloaded. The dataset contained 5500 non-COVID-19 and 4044 COVID-19 cases belonging to the X-ray folder and there were 2628 non-COVID-19 and 5427 COVID-19 images belonging to the CT folder, which brings the total number of images to 17,599. The images were divided into training, testing and validation sets. For each imaging modality, 20 COVID-19 and 20 non-COVID-19 images were set aside for the testing set. About 1860 images of both COVID-19 and non-COVID-19 were initially kept aside for validation for each of the imaging modalities. Upon experimenting with the number of images in the validation sets, it was found that validation sets consisting of 2500 images of both the cases for each modality resulted in the better validation accuracies. The remaining images were used for training the various models.

Models 1 was trained with the CT training set first. The model was trained with the training set five times, hence 5 epochs. Once trained, the validation set was fed into the model to obtain a validation accuracy of 24.89%. Model 1 was then tweaked in terms of the number of convolution layers, number of max pooling layers, adding dropout layers and changing kernel sizes to obtain Model 2. Model 2 was then trained with the training set with the same number of epochs and then validated. A validation accuracy of 20.74% was obtained. To improve the validation accuracy, certain parameters, such as varying the number of filters and adding a dropout layer with a dropout rate of 0.3, were added and changed in Model 2 to give rise to Model 3. Model 3 was then trained with the CT training set with 10 epochs. Increasing the number of epochs increased training time but somehow decreased the validation accuracy to 19.94%.

The above process was repeated with the X-ray training and validation sets. Models 1, 2 and 3 showed validation accuracies of 31.32%, 29.87% and 29.87% respectively.

Due to poor validation accuracies, pre-trained models were used in hoped of obtaining higher validation and testing accuracies. The top layers of the models were dropped from the pre-trained models to make it more suited for the study’s objective. Preprocessing layers were also added to the front of certain models.

Once the models are trained, validated, and tested, parameters such as the number of epochs, training accuracy and loss, validation accuracy and loss, true positive and negative, false positive and negative, prediction accuracy, sensitivity, and training times were all recorded and tabulated.

**Statistical analysis**

There was no statistical analysis conducted on features during the period of the study.

**Results**

The results obtained during our study have been tabulated and categorized in terms of imaging modalities. Table 3 presents the tabulated performance of the CNN models when CT images are used. Similarly, Table 4 presents the performance of the CNN models when X-ray images are used. The number of epochs, various losses and accuracies are mentioned in both Table 1 and Table 2. The training times are also mentioned.

The 3 non-pretrained, CNN models were first employed to gain more exposure to how the network functions and how we can alter it to increase accuracy rates. From Tables 3 and 4, it can be observed that both Models 1 and 2 have 5 epochs while Model 3 has 10. Among Models 1 and 2, it can be noted that Model 2 has a higher training accuracy [79.52% for CT images and 84.63% for X-ray images] that is due to the presence of more convolutional layers and increased filter size.

### Table 3. Performance of the convolutional neural network (CNN) models using computerized tomography (CT) images to detect coronavirus disease 2019 (COVID-19).

<table>
<thead>
<tr>
<th>Model</th>
<th>No. of epochs</th>
<th>Training loss</th>
<th>Training accuracy</th>
<th>Validation loss</th>
<th>Validation accuracy</th>
<th>Training time (Hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL 1</td>
<td>5</td>
<td>0.4279</td>
<td>0.7331</td>
<td>3.5918</td>
<td>0.2489</td>
<td>0.804</td>
</tr>
<tr>
<td>MODEL 2</td>
<td>5</td>
<td>0.3795</td>
<td>0.7952</td>
<td>7.0109</td>
<td>0.2074</td>
<td>0.483</td>
</tr>
<tr>
<td>MODEL 3</td>
<td>10</td>
<td>0.3126</td>
<td>0.8513</td>
<td>7.3482</td>
<td>0.1994</td>
<td>0.981</td>
</tr>
</tbody>
</table>
Model 3 took the most time to train, which is the result of increased number of epochs. However, higher number of epochs which resulted in longer training periods provided higher training accuracies [85.13% for CT images and 85.94% for X-ray images]. Figure 8 is a comparison of the 3 non-pretrained models based on training accuracy. While the CNN models provided high training accuracies, their validation accuracies were not satisfactory. For deep learning techniques to be employed in a clinical setting to detect COVID-19, the validation accuracies must be high. Hence, pre-trained networks were chosen and trained with our dataset. Table 5 and 6 show the performances of the 5 pre-trained models when they are trained and tested using CT images and X-ray images, respectively.

For the pre-trained models, the comparison has mostly been based on validation accuracies. Validation accuracies signify the generalizability of the model to detect COVID-19, in this scenario, from the dataset that was held back from the model while training. Another important parameter used is sensitivity. This gives us the ratio of actual positives that been predicted correctly by the model. The epochs of the training were decided to be set at 3 after experimentation - a higher number of iterations did not result in higher validation accuracies and smaller numbers of iterations result in lower accuracy rates. Figure 9 compares the obtained validation accuracies of the 5 pre-trained models for both the imaging modalities.

**Discussion**

The obtained results are compared to the highest obtained results from previous works for each of the imaging modality. These comparisons are done based on the imaging modality and parameters such as validation accuracy, sensitivity and prediction accuracy are taken into consideration (wherever they are available). Figure 10 presents the comparison of our results with previous works when X-ray images are used, and Figure 11 presents the comparison when CT images are used.

---

**Table 4. Performance of the convolutional neural network (CNN) models using X-ray images to detect coronavirus disease 2019 (COVID-19).**

<table>
<thead>
<tr>
<th>Model</th>
<th>No. of epochs</th>
<th>Training loss</th>
<th>Training accuracy</th>
<th>Validation loss</th>
<th>Validation accuracy</th>
<th>Training time (Hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL 1</td>
<td>5</td>
<td>0.4098</td>
<td>0.8246</td>
<td>1.9988</td>
<td>0.3132</td>
<td>1.273</td>
</tr>
<tr>
<td>MODEL 2</td>
<td>5</td>
<td>0.3779</td>
<td>0.8463</td>
<td>1.7220</td>
<td>0.2987</td>
<td>1.162</td>
</tr>
<tr>
<td>MODEL 3</td>
<td>10</td>
<td>0.3545</td>
<td>0.8594</td>
<td>1.7405</td>
<td>0.2987</td>
<td>2.625</td>
</tr>
</tbody>
</table>

**Figure 8. Comparison of the training accuracies of the 3 non-pretrained, convolutional neural network (CNN) models. CT = computerized tomography.**

Model 3 took the most time to train, which is the result of increased number of epochs. However, higher number of epochs which resulted in longer training periods provided higher training accuracies [85.13% for CT images and 85.94% for X-ray images]. Figure 8 is a comparison of the 3 non-pretrained models based on training accuracy. While the CNN models provided high training accuracies, their validation accuracies were not satisfactory. For deep learning techniques to be employed in a clinical setting to detect COVID-19, the validation accuracies must be high. Hence, pre-trained networks were chosen and trained with our dataset. Table 5 and 6 show the performances of the 5 pre-trained models when they are trained and tested using CT images and X-ray images, respectively.
### Table 5. Performance of various pre-trained models using computerized tomography (CT) images to detect coronavirus disease 2019 (COVID-19). TP = True positive, FP = False positive, TN = True negative, FN = False negative.

<table>
<thead>
<tr>
<th>Model</th>
<th>No. of epochs</th>
<th>Training loss</th>
<th>Training accuracy</th>
<th>Validation loss</th>
<th>Validation accuracy</th>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>Prediction accuracy</th>
<th>Sensitivity</th>
<th>Training time (Hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGG 16</td>
<td>3</td>
<td>0.1114</td>
<td>0.9544</td>
<td>0.8128</td>
<td>0.7664</td>
<td>16</td>
<td>4</td>
<td>20</td>
<td>0</td>
<td>95%</td>
<td>80%</td>
<td>1.98</td>
</tr>
<tr>
<td>VGG 19</td>
<td>3</td>
<td>0.1334</td>
<td>0.9456</td>
<td>0.66</td>
<td>0.812</td>
<td>18</td>
<td>2</td>
<td>18</td>
<td>2</td>
<td>90%</td>
<td>90%</td>
<td>3.19</td>
</tr>
<tr>
<td>Resnet</td>
<td>3</td>
<td>0.4334</td>
<td>0.8454</td>
<td>1.0943</td>
<td>0.7436</td>
<td>14</td>
<td>6</td>
<td>7</td>
<td>13</td>
<td>52.50%</td>
<td>70%</td>
<td></td>
</tr>
<tr>
<td>Inception v3</td>
<td>3</td>
<td>0.4654</td>
<td>0.9277</td>
<td>4.5027</td>
<td>0.6832</td>
<td>18</td>
<td>2</td>
<td>16</td>
<td>4</td>
<td>85%</td>
<td>90%</td>
<td>2.04</td>
</tr>
<tr>
<td>Xception</td>
<td>3</td>
<td>0.6184</td>
<td>0.9277</td>
<td>3.9648</td>
<td>0.7268</td>
<td>19</td>
<td>1</td>
<td>18</td>
<td>2</td>
<td>92.5%</td>
<td>95%</td>
<td>1.53</td>
</tr>
</tbody>
</table>

### Table 6. Performance of the pre-trained models using X-ray images to detect coronavirus disease 2019 (COVID-19). TP = True positive, FP = False positive, TN = True negative, FN = False negative.

<table>
<thead>
<tr>
<th>Model</th>
<th>No. of epochs</th>
<th>Training loss</th>
<th>Training accuracy</th>
<th>Validation loss</th>
<th>Validation accuracy</th>
<th>TP</th>
<th>FP</th>
<th>TN</th>
<th>FN</th>
<th>Prediction accuracy</th>
<th>Sensitivity</th>
<th>Training time (Hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGG</td>
<td>3</td>
<td>0.2651</td>
<td>0.8937</td>
<td>0.2279</td>
<td>0.8776</td>
<td>13</td>
<td>7</td>
<td>19</td>
<td>1</td>
<td>80%</td>
<td>65%</td>
<td>2.80</td>
</tr>
<tr>
<td>VGG</td>
<td>3</td>
<td>0.2943</td>
<td>0.8748</td>
<td>0.4024</td>
<td>0.8232</td>
<td>19</td>
<td>1</td>
<td>12</td>
<td>8</td>
<td>77.50%</td>
<td>95%</td>
<td>2.61</td>
</tr>
<tr>
<td>Res</td>
<td>3</td>
<td>0.504</td>
<td>0.7787</td>
<td>0.5298</td>
<td>0.756</td>
<td>19</td>
<td>1</td>
<td>5</td>
<td>15</td>
<td>60%</td>
<td>95%</td>
<td>1.53</td>
</tr>
<tr>
<td>Inception v3</td>
<td>3</td>
<td>1.0933</td>
<td>0.8518</td>
<td>4.4457</td>
<td>0.648</td>
<td>20</td>
<td>0</td>
<td>4</td>
<td>16</td>
<td>60%</td>
<td>100%</td>
<td>0.64</td>
</tr>
<tr>
<td>Xception</td>
<td>3</td>
<td>0.9504</td>
<td>0.8788</td>
<td>1.1236</td>
<td>0.88</td>
<td>12</td>
<td>8</td>
<td>15</td>
<td>5</td>
<td>67.50%</td>
<td>60%</td>
<td>1.92</td>
</tr>
</tbody>
</table>
The following tables summarize and compare the accuracies obtained by previous works and our study for each image modality. They also mention the key points of the methodology in each of the referred works. Table 7 compares and summarizes previous works based on X-ray together with our results that were obtained when the models were trained with X-ray images. Similarly, Table 8 shows the previous works and our work when trained with CT images.

It is seen that our methods have comparatively lower results in terms of accuracy. This can be due to several factors—lower number of epochs, using models with lesser layers etc. While the initial objective of the study was to employ deep learning techniques in COVID-19 detection that work with adequate accuracies, the main aim was to find which imaging modality provided higher accuracies when used to detect COVID-19. An observable trend is that the performances of the model change with the change in the imaging modality. Hence, comparing CT and X-ray imaging becomes vital as it will help develop an in-depth understanding of how a particular pre-trained network interprets the details provided by the radiological modalities to detect COVID-19.

**Figure 9. Comparison of validation accuracies of the 5 pre-trained models.** CT = computerized tomography.

**Figure 10. Comparison of results with previous works-X-ray.**

The following tables summarize and compare the accuracies obtained by previous works and our study for each image modality. They also mention the key points of the methodology in each of the referred works. Table 7 compares and summarizes previous works based on X-ray together with our results that were obtained when the models were trained with X-ray images. Similarly, Table 8 shows the previous works and our work when trained with CT images.

It is seen that our methods have comparatively lower results in terms of accuracy. This can be due to several factors—lower number of epochs, using models with lesser layers etc. While the initial objective of the study was to employ deep learning techniques in COVID-19 detection that work with adequate accuracies, the main aim was to find which imaging modality provided higher accuracies when used to detect COVID-19. An observable trend is that the performances of the model change with the change in the imaging modality. Hence, comparing CT and X-ray imaging becomes vital as it will help develop an in-depth understanding of how a particular pre-trained network interprets the details provided by the radiological modalities to detect COVID-19.
When trained with CT and X-ray images, VGG19 and Xception have the best validation accuracies, but their associated prediction accuracies are the lowest. Validation accuracies depict the generalizability of the network while prediction accuracies are calculated using the obtained true and false positive and negative values. Oftentimes, a larger difference between these values indicates overfitting. In terms of prediction accuracies, Xception and VGG16 show promising results when trained with CT and X-ray images, respectively. Their accuracies stand at 95% for Xception and 80% for VGG16.

![Figure 11. Comparison of results with previous works- computerized tomography (CT).](image)

**Table 7.** Comparison and summarized methodology and results of previous work and presented work- X-ray. COVID-19 = coronavirus disease 2019; CNN = convolutional neural network; CT = computerized tomography.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Methods</th>
<th>Measure of performance - accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>L. Wang et al. (2020)</td>
<td>Uses X-ray Images; Proposed model named COVID-net pretrained using both ImageNet and COVIDx dataset; Compared with 2 other pre-trained models- VGG-19 and ResNet-50.</td>
<td>COVID-net provided highest positive predictive value and sensitivity at 98.9% and 91% respectively.</td>
</tr>
<tr>
<td>S. Minaee et al. (2020)</td>
<td>Predicts COVID-19 from chest X-ray images without the need of feature extraction. Dataset names COVID-X-ray-5k was used on 4 pre-trained CNNs- ResNet18, ResNet50, SqueezeNet and, DenseNet-121. Model accuracy was based on sensitivity, ROC Curve and the precision recall curve.</td>
<td>SqueezeNet provided the highest accuracy of about 98%.</td>
</tr>
<tr>
<td>T. Ozturk et al. (2020)</td>
<td>DarkCovidNet detects COVID-19 using X-ray images of the chest. The model is used for both the two-classifier and the three-classifier problem.</td>
<td>The two-classifier problem provided higher accuracy at 98%.</td>
</tr>
<tr>
<td>M. E. Chowdhury et al. (2020)</td>
<td>The presented method uses collected chest X-ray images to train pre-trained models to detect COVID-19 pneumonia. 8 shallow networks and 5 deep networks were employed. Two different experiments- two class image, and three class image classification were done.</td>
<td>The two-class image experiment provided a higher accuracy at 99.7%.</td>
</tr>
<tr>
<td>Presented Method</td>
<td>3 non-pre-trained and 5 pre-trained models were trained and tested. Non-pre-trained models were rejected due to their low validation accuracies. The pre-trained models that were used are VGG16, VGG19, ResNet50, Inception V3, and Xception. All 8 models were trained and tested with X-ray and CT images separately.</td>
<td>Xception provided the highest validation accuracy at 88% when the model was trained with X-ray images.</td>
</tr>
</tbody>
</table>
The obtained accuracies can be further improved by increasing the number of epochs during training, however, increasing epochs would also mean increasing the training time. Another possible improvement would be to add on to the already present dataset for better training.

**Conclusion**

This study was conducted to compare the effectiveness of radiological imaging modalities in the detection of COVID-19, with the help of deep learning techniques. A total of 8 unique models were used- 3 non-pretrained, CNN models and 5 pre-trained models. From the obtained results, it can be noted that the performance of the presented model is compared with 6 other pre-trained models- AlexNet, GoogleNet, VGG16, MobileNetV2, ResNet18, and Inception V3. It was observed that the model, MKs-ELM-DNN, proved to be the most efficient in classifying the expanded CT dataset at 98.36%.

The presented model, EfficientCovidNet, is a high-quality deep learning model that can be used for screening COVID-19 from Chest CT Images. The performance of the proposed model was evaluated by testing it in 3 different scenarios- Random, Slices and Voting, using a 5-fold-cross-validation. The paper also presented a cross-dataset analysis to mimic real-life situation and unveil drawbacks. Testing the model in the ‘Random’ situation yielded better results. It was found that the initial accuracy of 87.68% of the model dropped to 56.16% during the cross-dataset analysis.

Discusses the performance of various deep learning models using CT Images. Detection uses decision fusion. The decision fusion model outperformed the other pre-trained models with an accuracy of 88.34%.

Presented model, named CTnet-10, was compared with 5 pre-trained models- VGG16, DenseNet161, ResNet50, Inception V3 and VGG19. It was observed that the CTnet-10 had an accuracy of 82.1%, which was lower than the other pre-trained models. Despite the lower accuracy, the presented model took the least time to train, test and execute.

Proposed method

<table>
<thead>
<tr>
<th>Authors</th>
<th>Methods</th>
<th>Measure of performance - accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>T.M (2020)</td>
<td>Discusses a new Multiple Kernels-Extreme-Learning Machine based Deep Neural Network (MKs-ELM-DNN) system to detect COVID-19 using Chest CT Images. The performance of the presented model is compared with 6 other pre-trained models- AlexNet, GoogleNet, VGG16, MobileNetv2, ResNet18, and Inception V3.</td>
<td>It was observed that the model, MKs-ELM-DNN, proved to be the most efficient in classifying the expanded CT dataset at 98.36%.</td>
</tr>
<tr>
<td>P. Silva et al. (2020)</td>
<td>The presented model, EfficientCovidNet, is a high-quality deep learning model that can be used for screening COVID-19 from Chest CT Images. The performance of the proposed model was evaluated by testing it in 3 different scenarios- Random, Slices and Voting, using a 5-fold-cross-validation. The paper also presented a cross-dataset analysis to mimic real-life situation and unveil drawbacks.</td>
<td>Testing the model in the ‘Random’ situation yielded better results. It was found that the initial accuracy of 87.68% of the model dropped to 56.16% during the cross-dataset analysis.</td>
</tr>
<tr>
<td>A. K. Mishra et al. (2020)</td>
<td>Discusses the performance of various deep learning models using CT Images. Detection uses decision fusion.</td>
<td>The decision fusion model outperformed the other pre-trained models with an accuracy of 88.34%.</td>
</tr>
<tr>
<td>V. Shah et al. (2020)</td>
<td>Presented model, named CTnet-10, was compared with 5 pre-trained models- VGG16, DenseNet161, ResNet50, Inception V3 and VGG19.</td>
<td>It was observed that the CTnet-10 had an accuracy of 82.1%, which was lower than the other pre-trained models. Despite the lower accuracy, the presented model took the least time to train, test and execute.</td>
</tr>
<tr>
<td>Proposed method</td>
<td>3 Non-Pre-Trained and 5 Pre-Trained models were trained and tested. Non-Pre-Trained models were rejected due to their low validation accuracies. The pre-trained models that were used are VGG16, VGG19, ResNet50, Inception V3, and Xception. All 8 models were trained and tested with X-ray and CT images separately.</td>
<td>VGG19 provided the highest validation accuracy at 81.2% when the model was trained with CT images.</td>
</tr>
</tbody>
</table>

The obtained accuracies can be further improved by increasing the number of epochs during training, however, increasing epochs would also mean increasing the training time. Another possible improvement would be to add on to the already present dataset for better training.

**Data availability**

**Source data**

Mendeley Data: Extensive COVID-19 X-Ray and CT Chest Images Dataset. [https://doi.org/10.17632/8h65ywd2jr.3](https://doi.org/10.17632/8h65ywd2jr.3)

The dataset is curated and augmented by Walid El-Shafai and Fathi Abd El-Samie. The dataset consists of COVID-19 and Non-COVID-19 folders for both X-ray and CT images. There were 5500 non-COVID-19 and 4044 COVID-19 cases belonging to the X-ray folder. For CT images, there were 2628 non-COVID-19 and 5427 COVID-19 images, which brings the total number of images to 17599.
Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Software availability
- Archived source code at time of publication: https://doi.org/10.5281/zenodo.5741317
- License: Creative Commons Zero “No rights reserved” data waiver (CC0 1.0 Public domain dedication).

References

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com