**COVID VISION: Advanced COVID-19 Detection from**

**Lung X-rays with Machine Learning or Deep Learnings**

**Team ID : 1**

**BATCH : GCE\_1**

**Team Size: 4**

**Team Members 1.GURU PRAKASH G**

**2.MANOJ S**

**3.GOWTHAM D**

**4.JENISH DEVA RAJ J**

|  |
| --- |
| **The Source Code of the project, the required Github links and Project Demo Link are attached at the bottom of this Project Report** |

**ABSTRACT:**

The COVID-19 pandemic has caused trouble in people’s daily lives and ruined several economies around the world, killing millions of people thus far. It is essential to screen the affected patients in a timely and cost-effective manner in order to ﬁght this disease. This paper presents the prediction of COVID-19 with Chest X-Ray images, and the implementation of an image processing system operated using deep learning and neural networks. In this paper, a Deep Learning, Machine Learning, and Convolutional Neural Network-based approach for pre- dicting Covid-19 positive and normal patients using Chest X-Ray pictures is pro- posed. In this study, machine learning tools such as TensorFlow were used for building and training neural nets. Scikit-learn was used for machine learning from end to end. Various deep learning features are used, such as Conv2D, Dense Net, Dropout, Maxpooling2D for creating the model. The proposed approach had a classiﬁcation accuracy of 96.43 percent and a validation accuracy of 98.33 percent after training and testing the X-Ray pictures. Finally, a web application has been developed for general users, which will detect chest x-ray images either as covid or normal. A GUI application for the Covid prediction framework was run. A chest X-ray image can be browsed and fed into the program by medical personnel or the general public.

Keywords: Covid-19 prediction; covid-19; coronavirus; normal; deep learning; convolutional neural network; image processing; chest x-ray

1. **Introduction**

In recent times, Coronavirus disease has become the biggest health hazard worldwide. Each country is going up against furious events as far as ensuring the wellbeing of its inhabitants because of the boundless thought of the disease and the detachment of medicine or immunization for it. The COVID-19 pandemic has achieved excellent examinations the world over. The effect on exploring ahead of time at the hour of the boundless, the criticality and difﬁculties of ongoing far and wide ask about, and this epic far-reaching fully emphasizes the importance of a pediatrician-researcher labor force. As it examines and goes through and beyond this far-reaching issue, which has the prospect of having a long-term impact on our reality,

This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original



research, and the biomedical inquiry initiative, it is vital to recognize and address openings and procedures, as well as difﬁculties in analyzing and sustaining the pediatrician-researcher labor force [[1](#_bookmark22)]. In the ﬁrst-place instances of what is by and by perceived as SARS-CoV-2, the deﬁlement, named COVID-19, were point by point in Wuhan, China in December 2019 as instances of deadly pneumonia. Then Italy, then the United States, and Russia. After that, it was the U.S., Brazil, and India. Because the locations of the coronavirus epidemic are spreading and growing, the economic prices are also rising. Bangladesh had a mortality toll of 21,397 people, a total incidence of 1,296,093, and a total amount of recoveries of 1,125,045. In just 24 h, the infection rate was 28.54 percent. On Tuesday, August 4th, as the pandemic continues to inﬂict damage on Bangladesh, health ofﬁcials reported another 235 Covid-19 deaths and 15,776 new cases. On Thursday, August 5th, the total number of COVID-19 infections worldwide surpassed 200 million, as per Worldometers [[2](#_bookmark23)–[4](#_bookmark24)].

The impact on the investigations into the case of COVID-19 was immediate, emotional, and, without a doubt, long-lasting. Most scholarly, business, and government fundamental research and clinical

investigations have been decreased, or investigations have already been diverted to COVID-19. The majority of ongoing clinical trials, including those researching life-saving cures, have been postponed, and the majority of those that are still open to modern recruitment have closed. Continuous clinical trials have been altered to allow domestic organizations to give care and virtual monitoring, reduce the risk of COVID-19 contamination, and avoid the diversion of healthcare resources from a widespread reaction [[5](#_bookmark25),[6](#_bookmark26)].

In a medical environment, learning the resulting outcome takes about 6–7 days, and it is also expensive for the general population. Due to these limitations, radiography checks can be used as a stand-in for diagnosing the disease. Chest radiography images can be evaluated to determine the presence of the novel coronavirus or its side effects. Infections are found in this family, according to studies, and show up as crucial symptoms in radiographic images. Furthermore, Polymerase Chain Reaction (PCR) test results are not always accurate. Furthermore, chest X-rays are more tolerant than other radiological examinations, such as Computed Temography (C.T) scans, and are available in almost every clinic. The difﬁculties of locating Covid-19 patients using chest x-rays (CXR) have been demonstrated, with prepared specialists not always being available, especially in the higher ranges [[7](#_bookmark27)]. Furthermore, the radiological indications associated with Covid-19 are novel and unexplored, with many specialists who have no experience with Covid-19 reporting positive, persistent CXRs. It has also been contrasted with a couple of existing benchmark works by different experts [[8](#_bookmark28),[9](#_bookmark29)].

The COVID-19 pandemic has resulted in a global emotional loss of human lives and is a magniﬁcent test of our entire condition. Public mindfulness and “dos and don’ts” programs for COVID-19 are being implemented in public areas. Environmental factors may also aid the coronavirus. However, the loss and recovery rates show that the pandemic is not being well handled. They couldn’t even get the results on time in most cases. As a result, the patient’s condition declines or dies.

Many corona detection systems like CXR and C.T. images use Transfer Learning and Haralick features, using the internet of things and sending alerts, COVID-19 Deep Learning Prediction Model Using Publicly Available Radiologist. Also, in a medical way, like Molecular point-of-care test, Polymerase chain reaction (qrt-PCR) etc. Due to a shortage of diagnostic kits and the incorrect prediction of RT-PCR in Algeria, public and private hospitals employed CT scans as an alternate diagnostic method to detect COVID-19 in patients [[10](#_bookmark30)]. CXR images, on the other hand, are better than any other in the ﬁeld of Covid-19, according to [[10](#_bookmark30),[11](#_bookmark31)],

encouraging outcomes [[12](#_bookmark32)]. With the help of C.T. check images, a convolutional neural (CNN) demonstration was used to differentiate Covid-19 patients in a study. A few more studies using C.T. scans to detect Covid-19 infection in human lungs with low accuracy [[13](#_bookmark33)]. The authors in [[4](#_bookmark24),[14](#_bookmark34)–[16](#_bookmark36)] investigated the performance of many deep learning formulas like VGG16, CNN-AD, Densenet, Resnet50, Custom CNN etc. to diagnose Covid-CXR and Normal-CXR images and their accuracy was in between (85% to 98.9%), which is quite good.

The study’s main goal is to estimate the most accurate result, to save time and money when it comes to Coronavirus tests. It shows a fully programmed framework for differentiating coronavirus-infected lungs from chest CT scan images and other lung disorders. For the frontend, a Graphical User Interface (GUI) system was used instead of Flask and Django. The novelty of this paper is that we used (CNN, Inception and DenseNet) 3 types of models for comparison. The accuracy obtained from the CNN model is

98.3 percent. Also, the GUI software tool will detect the covid patients’ results within a sec using the given x-ray images. In this way, it can be claimed that classiﬁcation using radiographic images, such as a chest X-ray (CXR), can be precise while also being signiﬁcantly faster and less expensive than a PCR test.

Section one provides an introduction. In section two, methods and methodology are described. In section three, mathematical equations and expressions are presented. The results are provided in section four. A performance comparison is shown in section ﬁve. Finally, section six discusses the conclusion.

1. **Method and Methodology**

The proposed system aims to predict Covid-19 from chest X-ray images using Deep Learning and Convolutional Neural Networks. The available dataset [[17](#_bookmark37),[18](#_bookmark38)] was collected from Kaggle and GitHub, which is about normal, pneumonia, and covid chest x-ray images. To begin, the dataset had 6432 samples and metadata.csv front images data which was adjusted using pictures of almost 196 operations on the tests of two classes. At that point, the chest X-ray 60 pictures were resized to make them uniform in terms of their determination. This section contains the methods and materials that are adapted to implement the system’s goal. The ﬁrst subsection discusses the classiﬁcation part and dataset creation techniques of the system; the next section contains the outlines of the system. Afterward, it will highlight the design and experiments of the system.

As a result, the article recommends a simple and effective Convolutional Neural Network (CNN) and Deep Learning-based technique for classifying Covid-19 positive and negative cases using CXR images. This method can saturate a very small, speciﬁc area of Covid-19 positive patients in a matter of seconds. We supplied an apparatus that can be designed to recognize Covid-19 positive patients as part of this study. To be sure, in the absence of a radiologist or if the topic specialists’ opinions conﬂict, this deep learning-based device will continue to offer interpretation without requiring human participation. We used data from free sources to demonstrate the applicability of the proposed gadget in terms of arrangement exactness and affected ability in this paper. [Fig. 1](#_bookmark3) shows the process of the system.

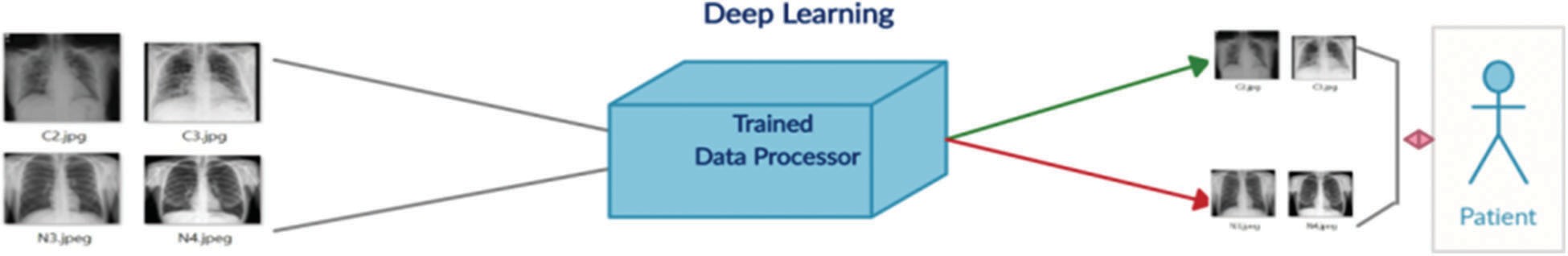


Figure 1: Showing the process

* 1. *Outline of Full System*

[Fig. 2](#_bookmark4) depicts a block diagram of the entire system.

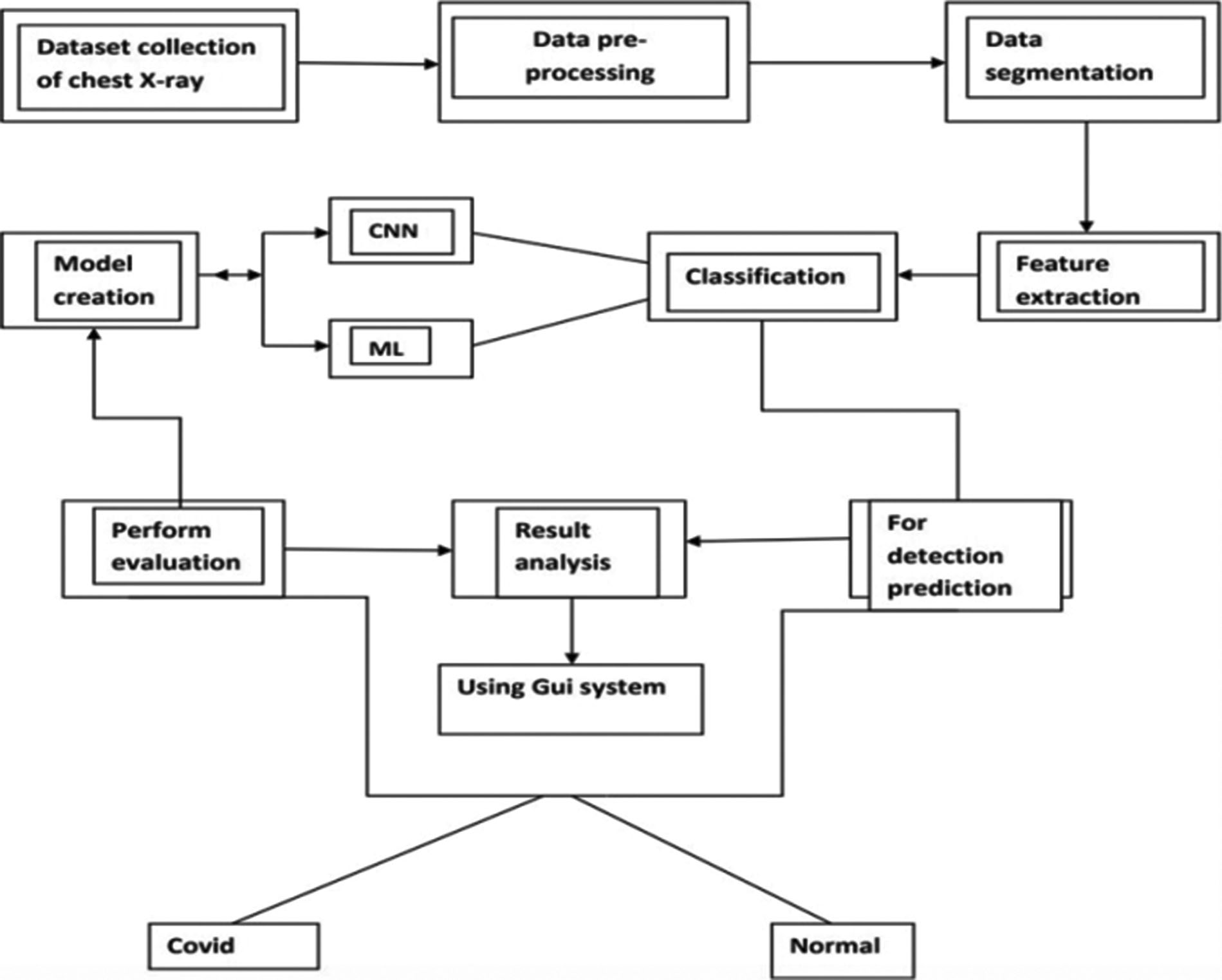


Figure 2: Block ﬂow diagram of the system

The system is made up of a dataset that includes normal and Covid patients’ chest X-ray images [[17](#_bookmark37),[18](#_bookmark38)]. Then, from the X-ray, we employed Artiﬁcial Intelligence to identify the occurrence of the Covid virus. Then, for a prediction model, data augmentation and dataset training using Deep Learning. So, here is the system showing the creating dataset part; data preprocessing may well be a get ready for arranging the rough data and making it suitable for a machine learning illustration. CXR images, in comparison to earlier image classiﬁcation tasks, contain a high level of pair-wise similarity and little intra-subject variability. This type of data can easily cause model deviation and over ﬁtting issues, diminishing the network’s generalization performance and making picture classiﬁcation jobs more complicated [[19](#_bookmark39)]. It is an essential and crucial step in making machine learning illustrated. Then data augmentation, data augmentation methods on CNN utilize Tensorﬂow and Keras. But a few times as of late, any procedure: Images are resized. The most commonly utilized image broadening strategies are code cases and representation of pictures after extension. From here onwards, data will be alluded to like pictures. We are going to be utilizing Tensorﬂow or OpenCV composed in Python in all our cases. After that, the feature was extracted. In machine learning and measurements, classiﬁcation could be an administered learning approach. The

computer program learns from the input information and makes a show that has been utilized to evaluate the performance and result analysis. To predict the result, using the GUI system as the front end. That system will show the detective result of the coronavirus after chest x-ray image dataset training using CNN and then testing of CXR images. Using this mproposed application, Covid-19 prediction will be shown.

* 1. ***Object Detection***

To identify objects for our Covid-19 detection system from chest X-ray images, we used OpenCV. This image processing technique helps with object detection based on the color, size, and shape of images. Different benchmark CNN models have been embraced in our proposed work. They have been trained individually to make independent predictions. Then the models are combined, using the new method of weighted average assembling technique, to predict a class value. We used OpenCV to detect particles in chest X-ray images for our Covid-19 detection system. This image processing technique assists in recognition of objects based on image color, size, and shape. In our proposed work, we used a variety of benchmark CNN models. Individually, they’ve been taught to make predictions on their own. The models are then merged to determine a class value, utilizing a novel way of the weighted average assembly procedure. This modern proposed assembling technique is expected to produce a more powerful expectation. DenseNet, h5, and the Inception model are three pre-trained CNN models in our proposed study. To train models, we applied Keras and TensorFlow with supplied parameters. Then, using the weighted normal gathering of the three models, they run the prepared models on the test images and choose lesson name 0 or 1 based on the results. Whereas partitioning the pictures into preparing and testing guarantees that there’s no persistent cover, i.e., distinctive pictures of the same quiet aren’t displayed in both preparing and testing datasets.

* + 1. ***Densenet Architecture***

Firstly, the DenseNet’s convolution creates a greater number of highlight maps. The number of yields, including maps of a layer, is characterized as the development rate. DenseNet has lower requirements for wide layers since layers are thickly associated with little excess within the learned highlights. [Fig. 3](#_bookmark5) shows the architecture of the CNN model DenseNet [[20](#_bookmark40)].

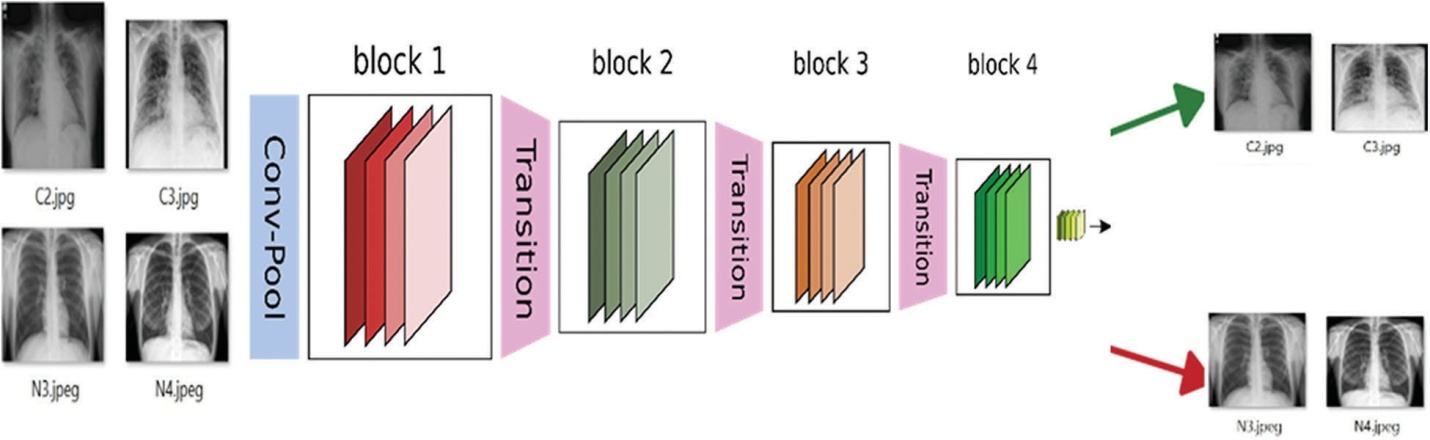


Figure 3: DenseNet architecture

Densely interconnected Convolutional Systems, also known as DenseNets, are the next step in the evolution of profound convolutional systems. After executing a composite of operations, traditional feed- forward neural systems interface the layer to yield it to the next layer. As we’ve seen, this composite typically includes a convolution operation or pooling layers, group normalization, and actuation work. The Dense Nets are divided into Dense Blocks, where the highlight map measurements remain constant

within a piece, but the number of channels varies [[21](#_bookmark41)]. These layers are called Move Layers, and they focus on down sampling while utilizing a variety of normalization techniques, including 1 x 1 convolution and 2 x 2 pooling layers. I can comprehend how this behavior, including 32 times the number of layers, is really, really, performed within the new deeper level, speaking to the primary Thick Layer inside the basic Thick Square. We do a 1 x 1 convolution with 128 ﬁlters to lower the highlight map estimate, and then a more expensive 3 x 3 convolution with this chosen 32 number (keeping in mind adding the cushioning to assure the measurements stay consistent). The input volume is then concatenated with the

results of the two processes (which are the same for each Thick Layer inside each Thick Piece).

* + 1. ***Inception Module***

The paper proposes a modern sort of engineering – Google Net or Initiation v1. It is essentially a convolutional neural network (CNN) that is 27 layers dense. The 1 × 1 Convolutional layer is sometimes applied to another layer, which is primarily utilized for dimensionality reduction. [Fig. 4](#_bookmark6) shows the architecture of the CNN model Inception Module.

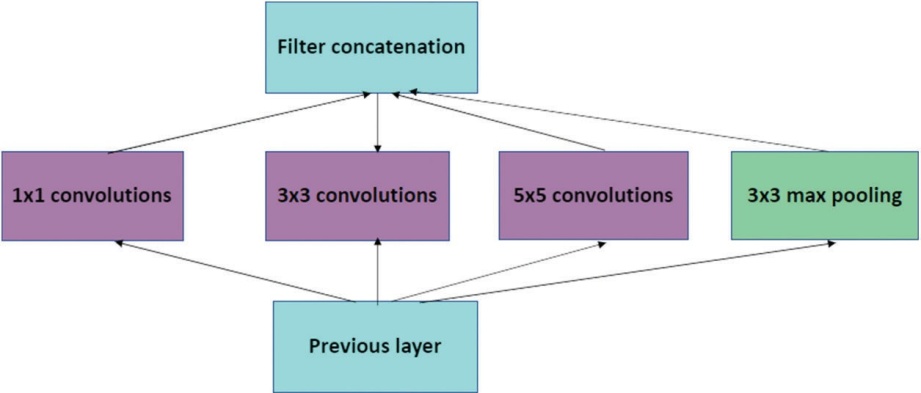


Figure 4: Inception module

Beginning Convolutional Neural Network systems (CNNs) are used to reduce processing costs by combining modules. Because a neural network interacts with an inﬁnite number of pictures, each of which has a wide variety of signiﬁcant components, they must be appropriately described. Convolution is performed on input with not one but three distinct channel sizes in the ﬁrst disentangled adaption of a

starting module (1 x 1, 3 x 3, 5 x 5). Furthermore, maximum pooling is employed. The following yields are concatenated and passed to the next phase at that time. The following yields are concatenated

and passed to the next tier at that time. The arrangement becomes dynamically more extensive, not more profound, by organizing CNN to complete its convolutions at the same level [[22](#_bookmark42)]. So, here this DenseNet and Inception model are used for better results and fewer validation errors. Also, the model ﬁles are saved as h5 model ﬁle names, which we used for the prediction result. It is divided into two-part Covid

(0) and Normal (1).

* + 1. ***Dataset Generation***

Convolutional neural networks were used to detect objects, and two chest X-ray datasets were analyzed. It consists of normal, COVID-19, and Pneumonia patients’ chest X-ray samples [[17](#_bookmark37)]. The second dataset can be accessed from the Github repository [[18](#_bookmark38)]. The other one was downloaded from Kaggle. That dataset basically contains images in three categories: normal, bacterial pneumonia, and viral pneumonia. It has around six thousand images, which are pretty big, so we separated 196 random images of type normal. In order to train a classiﬁer, we need positive and negative samples. Women, men, and children of different ages, patients’ chest X-ray image collections and other information were present in the GitHub folder.

The ﬁrst dataset contains a total of 392 images containing normal and covid samples. 196 images were found and extracted as COVID-19. Some of the X-rays have a front view, and some of the X-rays have a top view or side view. As we needed only the front view of X-ray images, we separated them and found a total of 196 posteroanterior (P.A.) views of X-rays. So, we created 50 percent Covid and 50 percent normal samples.

The dataset images, which are divided into two parts. Those are Covid positive and normal CXR images. The images are then converted into (224,224) forms and normalized. At that point, the pictures are rearranged and split into preparing and testing information. Thus, the training part has 60 images and 2 classes. The same testing part has 60 images and 2 classes. There are many possibilities that the same patients’ CXR images are kept in both the training and testing parts. It can be overlapping, but it’s kind of promising that training of the model, which has been examined by testing and validation checking, deﬁnes the ability of the trained model. Covid-19 positive and negative patients’ chest X-Ray images below in [Figs. 5](#_bookmark7) and [6](#_bookmark8).

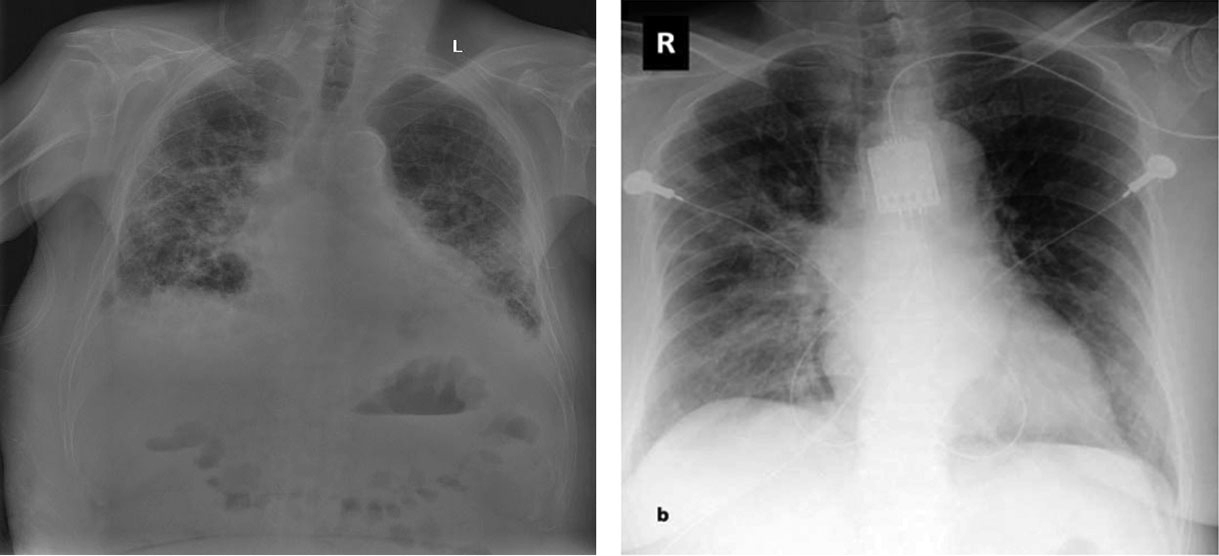


Figure 5: Covid-19 positive patient CXR pictures

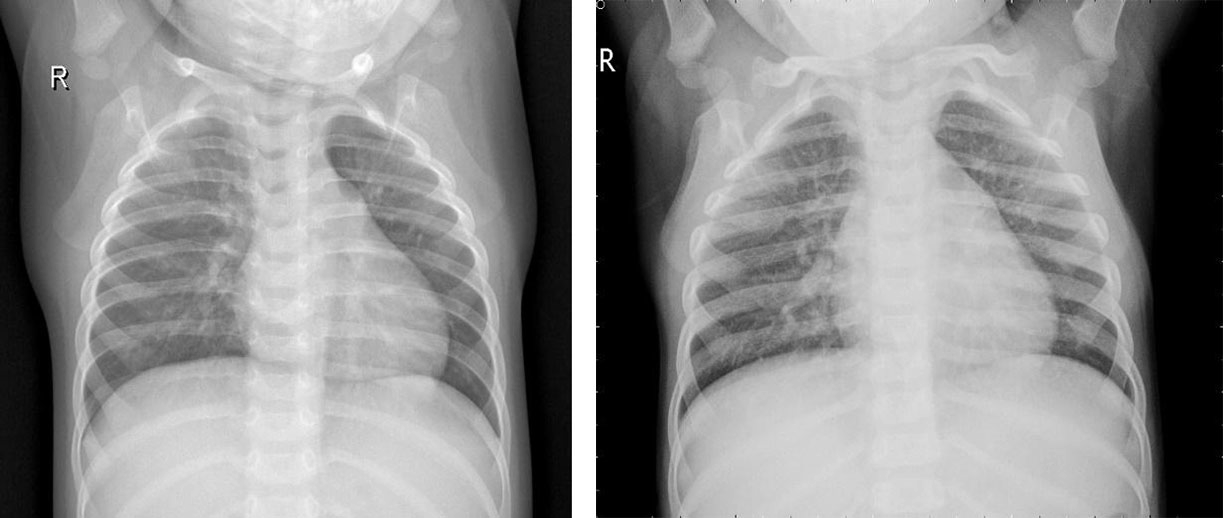


Figure 6: Covid-19 negative patient CXR pictures

Here are 4 images from the dataset which have been taken from the Kaggle dataset. Covid and Normal. The ﬁrst dataset was distributed in 2018, and from that point forward, around 100 exploration articles have 121 been distributed so far, including its example. The benchmark paper by Kermany et al. reports

122 claims a grouping exactness of 92.8 percent while utilizing the Inception V3 architecture (pre- prepared for the 123 ImageNet datasets) to recognize Normal and COVID tests. When it came to recognizing normal, bacterial, and viral pneumonia, they achieved a 90.9 percent accuracy rate. Notwithstanding, 125 ongoing examinations have revealed better order results (for paired groupings) on these 126 datasets. For instance, in 2020, Chouhan et al. detailed an examination portraying a Transfer Learning127 based methodology for Covid identiﬁcation, which brought about a 96.39 percent grouping

exactness 128. Nahid et al. suggested a two-channel CNN-based pneumonia location technique that generated a 132-characterization exactness of 97.92 percent [[23](#_bookmark43)]. They utilized ﬁve diverse pre-prepared CNN structures, including extraction and a 129-troupe procedure for the last grouping. Mittal et al. [[13](#_bookmark33)] used Dynamic Capsule Routing 130 to achieve a maximum categorization accuracy of 95.90 percent using the second dataset [[17](#_bookmark37)]. The creators utilized numerous pictures preparing strategies 133 to deal with the examples prior to playing out the grouping utilizing the profound learning model. Covid-19, a CNN-based design with 21 layers, including standard 135 and the depth of the astute detachable convolution operations, was introduced by R. 134 Siddiqi. Different parallel arrangements of 136 were used in that review to recognize pneumonia, and Covid-19 accomplished a 94.80 percent 137 arrangement exactness on the testing tests. For pneumonia identiﬁcation, Hu et al. presented MD- Conv, a multi-portion, profound, and comprehensive 138 convolution plot. On the popular Chest X-beam

14 dataset, they tested their 139 methods and got a 98.30 percent Area Under the Curve (AUC) 140 score. 141 A little research, such as the 142 suggested approach, used chest X-beam images to perform three-class characterization, including the initial dataset. Mahmud et al. 143, for example, presented CovXNet (a multi-enlargement CNN) for normal and COVID-19 identiﬁcation and 144 Covid- 19 arrangements in 2020. They tried various layouts using subsets of 145 of the two chest radiograph datasets and a variety of deep learning models. In any case, they 146 accomplished pinnacle results by utilizing CovXNet with adaptable multi-open element 147 advancement. Using CNN and Transfer Learning, Jain et al. achieved three-class characterizations to recognize Covid and its 148 types. To fulﬁll their objective, they looked at several avenues using 149 six different CNN-based models, four of which were pre-prepared 150 models, similar to the former inquiry [[23](#_bookmark43)].

* 1. ***Design of Interface***

This proposed work’s most signiﬁcant beneﬁt is that it is very user-friendly, and the system beneﬁts in all sections of socities. Everyone can use it if anyone has a minimum knowledge of browsing and selecting images. This whole system is mainly done in the Python language, and for the overall design, we use the Tkinter library, which is Python’s default graphical user interface toolkit, to generate a standard user interface. A dataset was created using a Jupyter notebook. We train and validate our data in Google Colab. We use Flask, which helps us to import libraries. For the CNN base model, we use Keras for image processing; we also use OpenCV and Tensorﬂow, which resize images and zoom. Also, the standard image size is 224,224. We use scikit-learn to maintain our algorithm-like epochs. To highlight training loss, training accuracy, validation accuracy, and validation loss, we use Matplotlib, from which the library was born, as well as Sklearn. Metrics, which allow us to show true positive, true negative, falsely positive, and falsely negative, actually show how successfully the system’s work is done. To detect the image, we use artiﬁcial intelligence. In the backend part, deep learning is used.

1. Equations and Mathematical Expression

Actual or classiﬁcation accuracy, which we may obtain using some variant of cross-validation data, or sensitivity, which is the ratio of the correctly + classiﬁed by our program to all, are the performance metrics used to measure the success of the proposed system. The ratio of accurately + classiﬁed by our software to all

+ classiﬁed is known as precision. After that, the overall score was calculated.

* 1. ***In-line Style***

Accuracy measured as follows:

Classification accuracy ¼ ð*TP* þ *TN* Þ=ð*TP* þ *FP* þ *FN* þ *TN* Þ (1)

Sensitivity ¼ *TP*=ð*TP* þ *FN* Þ (2)

Precision ¼ *TP*=ð*TP* þ *FP*Þ (3)

F1 Score ¼ 2 x sensitivity x precision=sensitivity þ precision (4)

It’s called a classiﬁcation metric, which, in summary, is the number of correct and incorrect predictions made by a classiﬁer.

* 1. ***Display Style***

For these confusion metrics, we need the Python machine learning library nameScikit-learn. The True Positive/Negative name refers to the anticipated result of a test, whereas the True/False refers to the real result. So, in the event, I anticipated that somebody would be Covid-19 positive, but they weren’t. At that point, that would be a False Positive since the real result was wrong, but the expectation was positive.

* 1. *Tables and Figures*

It is a summary of our proposed solutions and results. The technique is repeated until a speciﬁed meeting model is fulﬁlled, a predetermined number of eras has elapsed, or the arrangement has remained the same for a few successive periods. [Tab. 1](#_bookmark9) shows the data preprocessing.

Table 1: Architecture of the model used for deep feature extraction

|  |  |  |
| --- | --- | --- |
| Layer (type) | Output Shape | Parameters |
| conv2d (Conv2D) | (None, 222, 222, 32) | 896 |
| conv2d\_1 (Conv2D) | (None, 220, 220, 64) | 18496 |
| max\_pooling2d(MaxPooling2D) | (None, 110, 110, 64) | 0 |
| dropout (Dropout) | (None, 110, 110, 64) | 0 |
| conv2d\_2 (Conv2D) | (None, 108, 108, 64) | 36928 |
| max\_pooling2d\_1(MaxPooling2) | (None, 54, 54, 64) | 0 |
| dropout\_1 (Dropout) | (None, 54, 54, 64) | 0 |
| conv2d\_3 (Conv2D) | (None, 52, 52, 128) | 73856 |
| max\_pooling2d\_2(MaxPooling2) | (None, 26, 26, 128) | 0 |
| dropout\_2 (Dropout) | (None, 26, 26, 128) | 0 |
| ﬂatten (Flatten) | (None, 86528) | 0 |
| dense (Dense) | (None, 64) | 5537856 |
| dropout\_3 (Dropout) | (None, 64) | 0 |
| dense\_1 (Dense) | (None, 1) | 65 |

Total params: 5,668,097

Trainable params: 5,668,097

Non-trainable params: 0

The model is sequential. It is also a CNN-based model in Keras. From here, we can see the layer types and our shape number. Total params number, where 5,668,097 is trainable and non-trainable number is 0.

The table summarizes the training and validation loss and accuracy of two ﬁles, where for Covid positive it is 0 and for normal it is 1: for example, [Tab. 2](#_bookmark10).

Table 2: Training and validation loss, accuracy

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| COVID  Positive | Normal | Training accuracy | Training loss | Validation accuracy | Validation accuracy |
| 0 | 1 | 0.9643 | 0.0904 | 0.9833 | 0.0363 |

From [Tab. 2](#_bookmark10), it can be seen that training accuracy is 0.96, training loss is 0.09, validation accuracy is

0.98 and validation loss is 0.03.

[Tab. 3](#_bookmark11) shows the Model evaluate generator for the Training and Validation part of Covid and Normal. Here for Covid it’s 0, and normal it’s 1. After the Model Training it has got the accuracy of training and validation up to 97 percent.

Table 3: Model evaluate generator

|  |  |  |
| --- | --- | --- |
| Covid-19 class indices | Training generator | Validation generator |
| 0 | 0.06231 | 0.03630 |
| 1 | 0.97321 | 0.98333 |

The model evaluates the generator of training and validation is Training accuracy of 0.6231 and loss of 0.97321. Validation accuracy is 0.98333, and loss is 0.03630.

All models are trained for a total of 10 epochs with a step-per-epoch of 7. Model ﬁt is generated here, and the model is saved as an a.h5 ﬁle. Model training takes 9 s for each epoch. The training accuracy of this model is almost 96 percent, and the validation accuracy is almost 98 percent.

1. **Results and Analysis**

The experimental evaluation of the suggested method is described in this section. The goal of the ofﬂine training and testing experiments is to ﬁnd the best machine learning model with the highest output for real- time sentiment polarity prediction. We analyzed the performance using two machine learning models and training on a total of ﬁve thousand CXR images. Where 60 images for testing and validation.

So, starting by describing the confusion matrix. For these confusion metrics, the True Positive/Negative name refers to the anticipated result of a test, whereas the True/False refers to the real result. In [Fig. 7](#_bookmark12), it can be seen in the accuracy table of performance evaluation.

Here, T.P. is 30, TN is 29, F.P. is 1, and F.N. is 0. So, the train generator class index Covid is 0 and normal is 1.

Here the graph shows that training and validation accuracy is increasing. At epoch 4, the training accuracy was a little bit, but after epoch 04, the training accuracy increased. Here, the highest training accuracy is 0.95 after epoch 8. Here the graph also shows the validation accuracy; it also increases when epoch 0; the validation accuracy is 0.92. The validation accuracy is highest at epoch 1, after epoch 7, it is a little bit decreasing. After epoch 8, the validation accuracy again increased. The training loss is 0.09, and the validation loss is 0.03. Here, when the epoch is 0, the training loss is high. At the highest point, training loss is decreasing. Here the graph shows that in epoch 2, 4, 6, 8, training loss is decreasing, and

after the epoch, the training loss rate is the lowest. But in epoch 4, 7, it slightly increased, and after epoch 08, the training loss was very low. On the other hand, the graph also shows validation loss. Here the highest validation loss is close to 0.58 when epoch 0. When the epoch is 1 to 2, range validation is constant, and after epoch 2, the validation loss also decreases. When epoch is in the range of 3–7, the change in validation loss is more or less the same, but at epoch 7, the validation is a little bit increased, but after epoch 7, the validation loss again decreases. Training loss is higher than validation loss. It can be shown with the help of both graphs below. [Figs. 8](#_bookmark13) and [9](#_bookmark14) show the graph of the accuracy and loss:

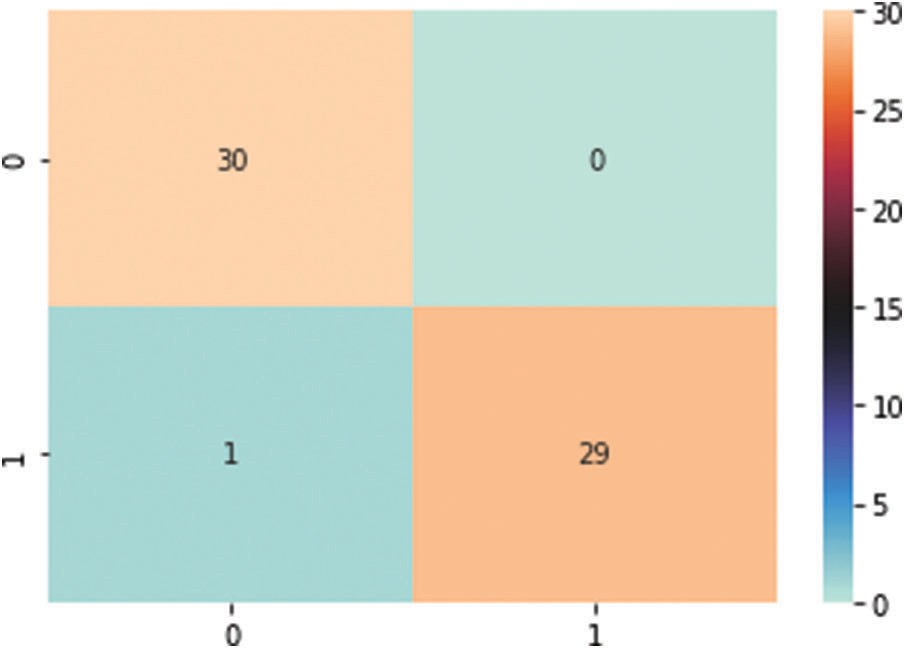


Figure 7: Accuracy table

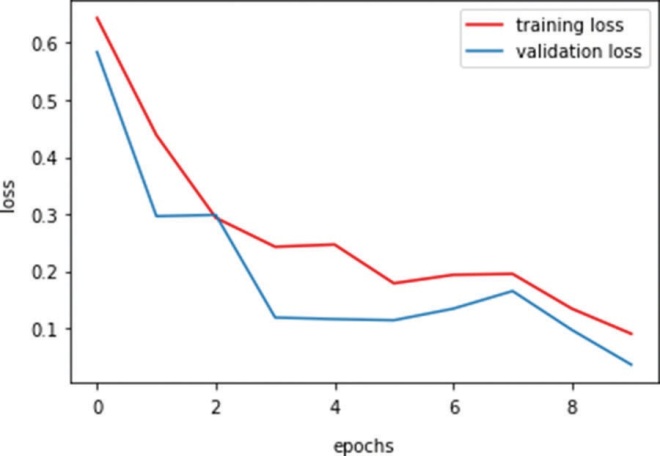


Figure 8: Training and validation loss

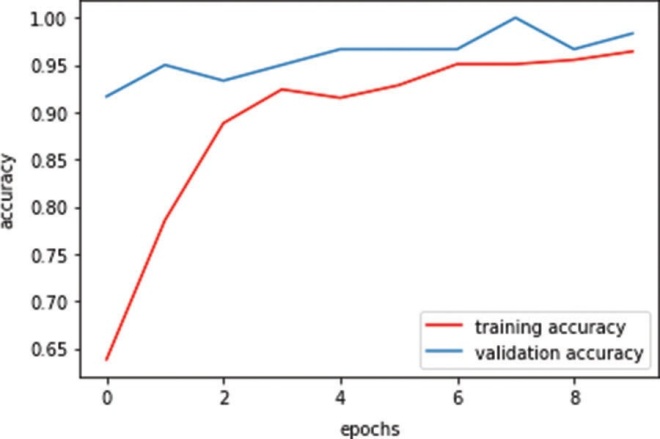


Figure 9: Training and validation accuracy

Overall validation accuracy is higher than training accuracy. It can be shown with the help of both graphs below. For graphing the epochs for training and validation, it has been used by matplotlib and the Sklearn library.

* 1. ***Frontend Tool***

Based on the proposed course of action, a GUI application for the Covid Prediction framework was run. In a way that will minimize Covid-19 positive and negative situations, a direct desktop program was developed, as shown in [Fig. 10](#_bookmark15).

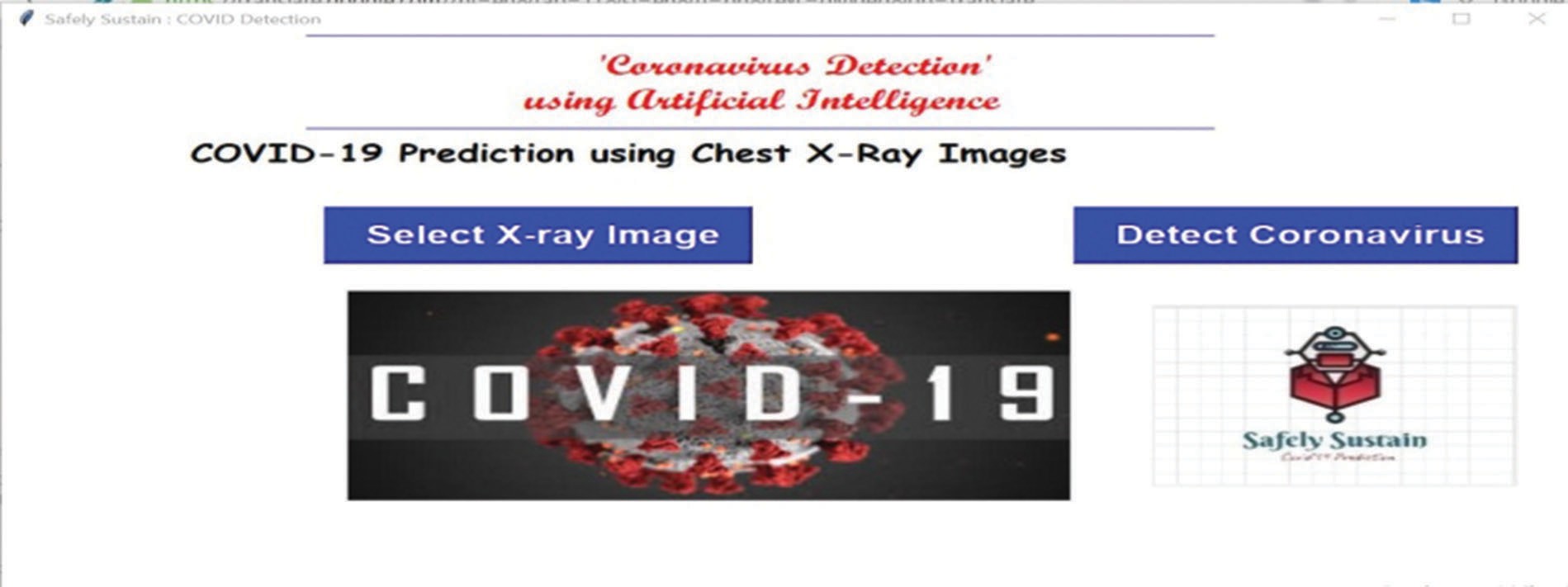


Figure 10: Gui Covid-19 prediction desktop tool

A chest X-ray image can be browsed and fed into the program by medical personnel or the general public. In turn, the application will implement the provided illustration in this paper and assign a title to the given Chest X-Ray image, such as Covid or Normal. As a result, they will identify the Covid +ve and Covid -ve situations in addition to their probabilities, as illustrated in [Fig. 11](#_bookmark16). This might be used on a variety of platforms, including Windows, Mac OS X, and Linux. This interface can be used in any Covid-19 testing center or other medical ofﬁces for speciﬁc areas of the disease. This arrangement to utilize devices alongside the fundamental code for data course of action and appear planning is accessible freely on the GitHub account.

Firstly, to get the desktop tool in Anaconda Prompt, press the tool name gui\_covid.py. Then enter the platform where it can be seen like the one below. In [Fig. 10](#_bookmark15), it shows the Gui Covid-19 prediction desktop tool.After that, press the select X-ray image, open your ﬁle explorer, select your chest x-ray image, and it will look like [Fig. 11](#_bookmark16).

After the selection of the images, it will now press the detect button to predict your result. Then press OK. After that, press the detect the coronavirus button. After a few seconds, the result will appear on the screen, either Covid or Normal. The result can be shown in [Fig. 12](#_bookmark17).

Finally, after image analysis, the user will get the result and it will show users a message. [Figs. 13](#_bookmark18) and [14](#_bookmark19) show the text box of the result. The image is shown below.

When the user crosses the x-ray image, the text message will be shown to the user. If the patient gets a result of Covid-19 positive, it shows Take Care, be alert and Stay Safe. If the patient gets the result of Covid- 19 negative, it shows don’t worry, and you are safe. After that, the user can end their prediction or they can again apply the same procedure with different chest x-ray images.

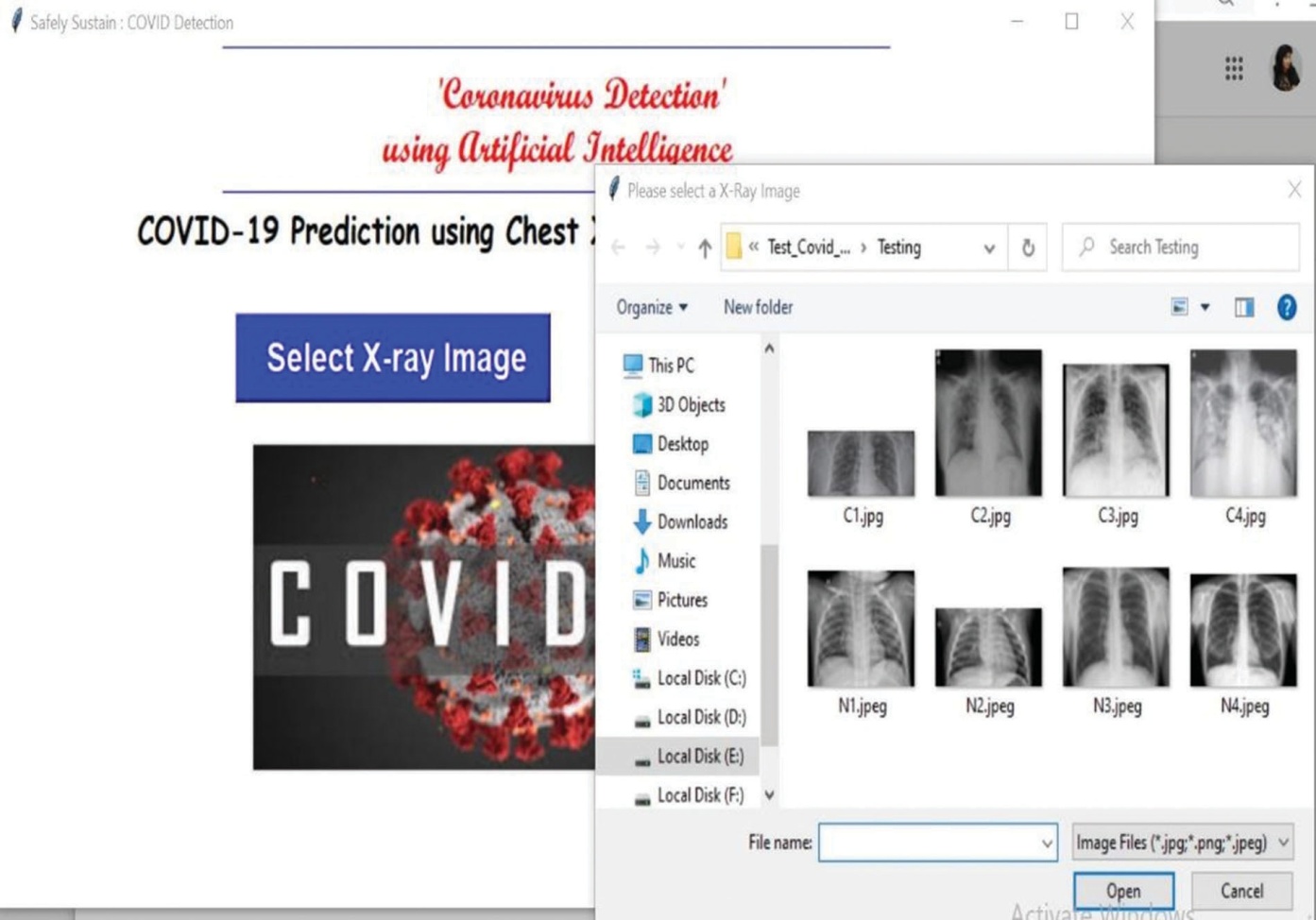


Figure 11: Selection of CXR images

So, developing tools where they have been used in the Gui application for indexing and templates. For the Gui application, we imported the Python library named Tkinter. To predict exact results, we used our training model.h5 ﬁle. So. This is how users can get their Covid-19 results.

1. ***Performance Comparison***

We compared the model’s execution to comparable strategies in [Tab. 4](#_bookmark20) to put it in context. To execute two-class classiﬁcations, the ﬁrst two questions on the table require distinct deep-learning algorithms. There are also ensembles. Regardless, we must emphasize that none of the cited papers [[24](#_bookmark44)] address the speciﬁc dataset in question. They took the corresponding lesson exams from the parent dataset in the majority of cases.

In summary, the success of our proposed solution, It’s better than models of individuals. The closest execution is that of DenseNet201. In [Tab. 5](#_bookmark21), we compared the model’s performance to that of similar methods to put it in context. References [[14](#_bookmark34),[16](#_bookmark36)] use numerous deep-learning algorithms to produce four- class classiﬁcations. In one case, the described model’s performance was practically identical to this proposed method (98.3 percent). In every case, the strategy outperforms the presented lead in terms of classiﬁcation accuracy in those studies. However, none of the referenced papers are relevant to the speciﬁc dataset used in this analysis. They took the associated class samples from the parent dataset in the majority of situations.

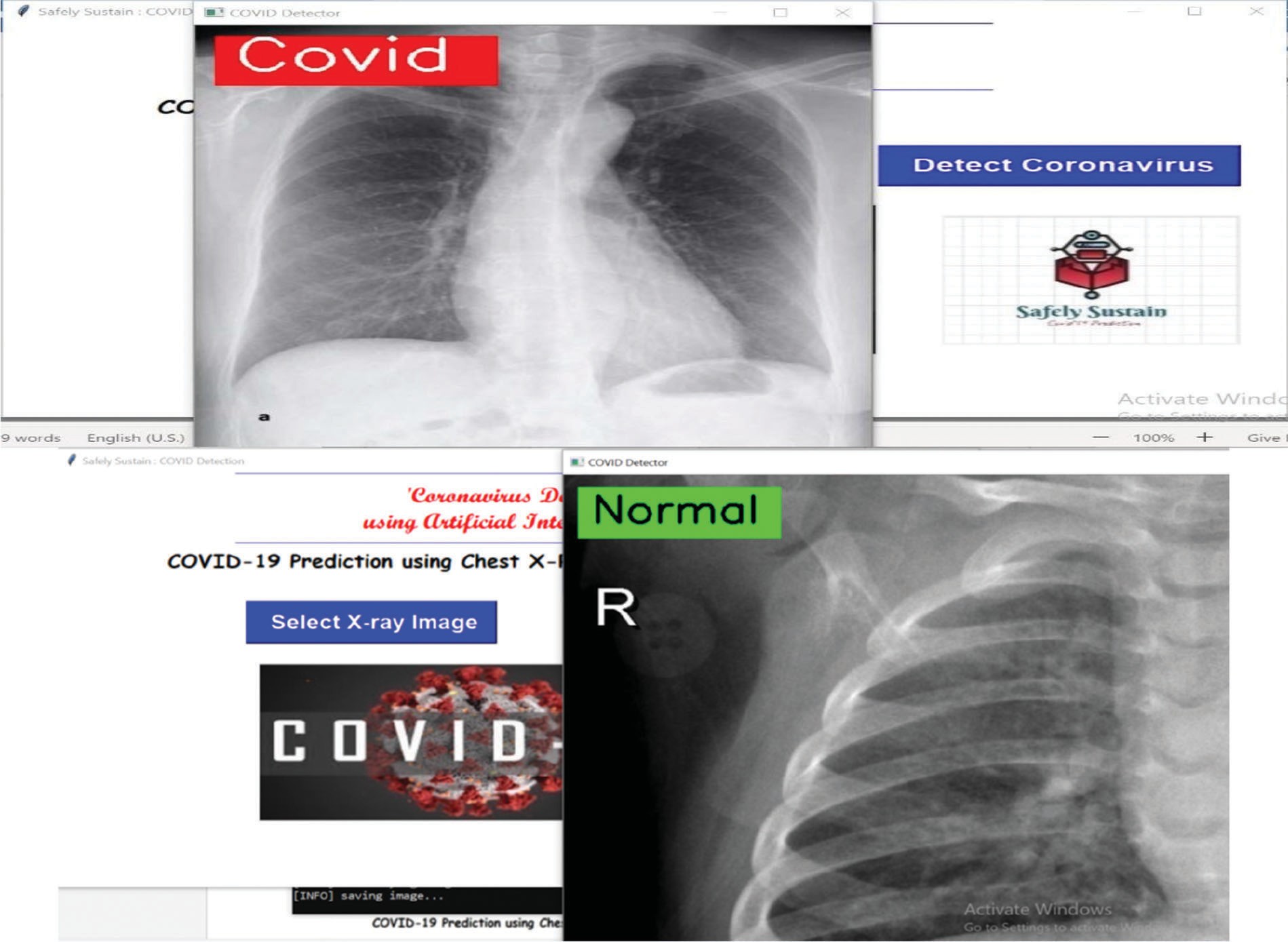


Figure 12: Detect coronavirus result

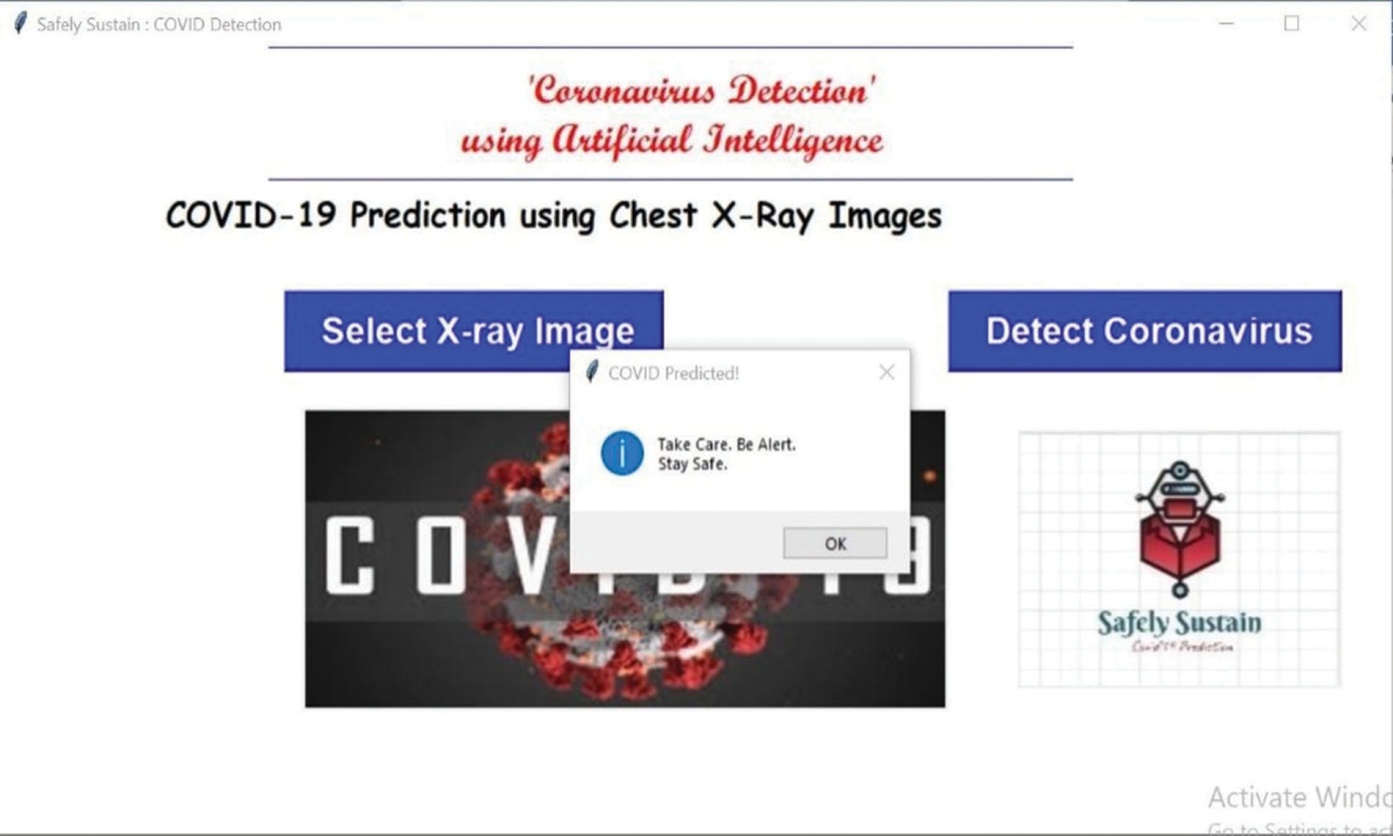


Figure 13: Showing the text for Covid report



Figure 14: Showing the text for normal report

Table 4: Summary of performance

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Models | | Parameters | | Validation Accuracy (%) | Precision (%) | Sensitivity (%) | F1-Score (%) |
| Individual Networks | |  | |  |  |  |  |
| DenseNet201 | | 5,668,097 | | 96.8 | 96 | 98 | 97.4 |
| Inception v3 | | 5,634,029 | | 95 | 91 | 92 | 95.8 |
| Ensembled networks | |  | |  |  |  |  |
| Unweighted average | |  | | 94.5 | 93 | 95 | 95.1 |
| Weighted average (accuracy) [[23](#_bookmark43)] | |  | | 94.5 | 94 | 95 | 95.1 |
| Weighted average (rank) [[23](#_bookmark43)] | |  | | 95.3 | 95 | 97 | 95.8 |
| Proposed Approach | |  | | 97.6 | 96 | 98 | 98.3 |
| Table 5: Compares the performance of the method to that of other similar methods | | | | | | | |
| Reference | Identiﬁed Classes | | Number of Samples | | Classiﬁcation Model | | Accuracy (%) |
| [[4](#_bookmark24)] | COVID-CXR | | 21456 | | CovXNet | | 95.00 |
| [[14](#_bookmark34)] | COVID-CXR | | 5349 | | CNN-AD | | 85.00 |
| [[14](#_bookmark34)] | NORMAL-CXR | | 5349 | | CNN-AD | | 85.00 |
| [[14](#_bookmark34)] | COVID-CXR | | 5349 | | CNN-SA | | 95.00 |
| [[14](#_bookmark34)] | NORMAL-CXR | | 5349 | | CNN-SA | | 95.00 |

(Continued )

Table 5 (continued ).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Reference | Identiﬁed Classes | Number of Samples | Classiﬁcation Model | Accuracy (%) |
| [[15](#_bookmark35)] | COVID-CXR | 6432 | VGG16 | 91.69 |
| [[16](#_bookmark36)] | COVID-CXR | 3000 | RF | 89.41 |
| [[16](#_bookmark36)] | COVID-CXR | 3000 | LR | 88.36 |
| [[16](#_bookmark36)] | COVID-CXR | 3000 | KNN | 69.25 |
| [[16](#_bookmark36)] | COVID-CXR | 3000 | DenseNet | 85.73 |
| Proposed | COVID-CXR | 5668 | CNN | 98.3 |

“–” denotes that the information is not mentioned in the associated paper.

* *Real-Time Reverse Transcription Polymerase Chain Reaction*: the viral gene detection by Real-Time Reverse Transcription Polymerase Chain Reaction (RT-PCR) is very sensitive because it can detect a copy of a specific genomic sequence, which has lead to the development of many commercial technologies that use nasal or nasopharyngeal swabs along with RT-PCR for COVID-19 detection [[Yüce et al.,](#_bookmark16) [2021].](#_bookmark16)

However, the detection of COVID-19 using RT-PCR is complex, the materials are sometimes slow to deliver, and the complete process can only be performed by qualified clinical laboratory personnel, which take over 24 hours from taking the sample to getting the analysis results. It is also expensive due to one kit can cost over 100 USD and setting a lab costs more than 15,000 USD. Additionally, many factors like storage, collection, processing, and genomic mutations, can lead to incorrect results [[Afzal,](#_bookmark17) [2020].](#_bookmark17) Other drawbacks include low availability in some countries [[Aziz et al.,](#_bookmark18) [2020],](#_bookmark18) and high false-negative rates [F[an et al.,](#_bookmark19) [2020].](#_bookmark19)

* *Computed Tomography scan*: this is an advanced technique that allows to generates detailed 3D images of organs and soft tissues [[Ohata et al.,](#_bookmark20) [2021].](#_bookmark20) Unlike RT-PCR, a CT scan is fast to obtain and it is relatively easy to perform. It has been recently reported that this technique shows typical features of COVID-19 like ground-glass opacities and multifocal patchy consolidation, even in patients with negative PCR but clinical symptoms [[Ai et al.,](#_bookmark21) [2020].](#_bookmark21) However, decontaminate CT equipment after scanning COVID-19 patients may damage it. Thus, to minimize the risk of cross-infection, it is suggested to use portable devices like chest radiography, which is already a triage tool in many hospitals [Wong et al.](#_bookmark22) [[2020].](#_bookmark22)
* *Chest X-ray*: X-ray refers to a medical imaging technique that uses radiation to generate an image of internal structures of the human body. The main elements that are assessed using X-ray are bones, which appear white on the image; soft tissues, which appear as light gray; fat, which appears gray; and gas, which appears black [[Anis et al.,](#_bookmark14) [2020].](#_bookmark14) In particular, a chest X-ray image allows a doctor to evaluate multiple organs, structures and conditions [Breiding](#_bookmark23) [[2009].](#_bookmark23) It is one of the most used methods to diagnose pneumonia worldwide [[Jaiswal](#_bookmark24) [et al.,](#_bookmark24) [2019].](#_bookmark24)

Chest X-ray devices can be portable, are affordable, fast, and gives the patient a lower radiation dose than CT. It has been reported that common CT findings can also be detected on chest X-ray images, even in patients with initial negative RT-PCR for COVID-19. However, the diagnosis of COVID-19 using chest X-ray images is more difficult than using CT or other imaging modalities and can only be performed by specialist physicians, which scarce [[Narin et al.,](#_bookmark25) [2021,](#_bookmark25) [Wong et al.,](#_bookmark22) [2020].](#_bookmark22)

There is a trade-off between quality and accessibility when choosing the imaging technique to use. CT produces a higher quality image but requires a much more complex device, not always available in many institutions. On the contrary, X-ray devices are much more affordable, can be portable, and are less harmful, given that a single CT scan can deliver a median effective radiation dose as high as 442 chest X-ray series [[Breiding,](#_bookmark23) [2009].](#_bookmark23)

Concerning the radiological findings on chest X-ray and CT associated with COVID-19 pneumonia, the most common are ground glass opacities. These marks are usually bilateral, meaning they affect both lungs and are more likely to be located in the periphery and lower areas of the lungs [[Kaufman et al.,](#_bookmark26) [2020,](#_bookmark26) [Yasin and Gouda,](#_bookmark27) [2020].](#_bookmark27) They can be seen in chest X-ray images, or CT images as regions of increased whiteness due to the augmented density [[Cleverley et al.,](#_bookmark28) [2020],](#_bookmark28) which do not cover blood vessels and airway walls completely [[Rousan et al.,](#_bookmark29) [2020].](#_bookmark29) As the disease progresses, this finding becomes denser and covers blood vessels and airway walls on the image, becoming consolidations. Fig. [1](#_bookmark0) presents a comparison of the chest X-ray images for a COVID-19 negative subject and a COVID-19 positive subject; additionally, for the COVID-19 case, the image on the right shows the masks over the regions of ground glass opacities (yellow) and consolidations (purple).

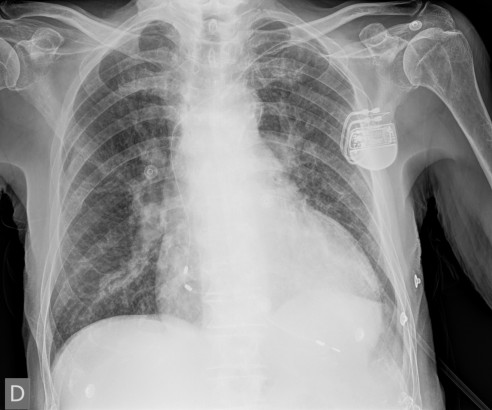
The WHO solidarity consortium from February 2021 presents a study to find a drug against COVID-19, however, it was found that the mortality, initiation of ventilation, and hospitalization duration were not definitely reduced by any trial drug. Until now, no specific drug has been found against COVID-19 [WHO Solidarity Trial Consortium](#_bookmark30) [[2021].](#_bookmark30) Currently, there are 153 vaccine candidates, 476 vaccine trials ongoing, 23 vaccines approved by at least one country, and 7 vaccines approved for use by the WHO against COVID-19. The AstraZeneca vaccine is the one approved in the largest number of countries, followed by Pfizer/BioNTech, Moderna, and Janssen [[WHO,](#_bookmark31) [2021].](#_bookmark31)

Artificial Intelligence (AI) techniques, including Machine Learning (ML), can be used for COVID-19 diagnosis from chest X-ray images and set foundations for automatic decision-making support systems [[Arteaga-Arteaga et al.,](#_bookmark32) [2022].](#_bookmark32) AI refers to the process of providing computer features from human intelligence. ML is a subset of AI that holds the mathematical models used to achieve this task, whereas Deep Learning (DL) is a subset of ML itself and relates the models and algorithms based on neural networks [[Goodfellow et al.,](#_bookmark33) [2016].](#_bookmark33) In general, ML and DL techniques are designed to extract features and find relationships between data samples. Thereby, these approaches are well-suited for tasks relying on the human experience [[Orozco-Arias et al.,](#_bookmark34) [2019,](#_bookmark34) [Tabares-Soto et al.,](#_bookmark35) [2019,](#_bookmark35) [Bravo Ortíz et al.,](#_bookmark36) [2021,](#_bookmark36) [Reinel et al.,](#_bookmark37) [2021,](#_bookmark37) [Arteaga-Arteaga et al.,](#_bookmark38) [2021]](#_bookmark38) such as classifying a chest X-ray image as positive or negative for COVID-19. Besides decision-making support systems in the medicine and healthcare field, AI has been used to perform tasks from managing medical data and analyzing health plans to drug development and health monitoring [[Amisha](#_bookmark39) [et al.,](#_bookmark39) [2019].](#_bookmark39) AI applications in medicine aim to improve diagnostic performance and offer a better quality of service [[Ahuja,](#_bookmark40) [2019].](#_bookmark40)

Regarding the detection of COVID-19 in chest X-ray images using AI techniques, most research papers propose a transfer learning approach using Convolutional Neural Networks (CNNs) such as VGG19, Inception, and MobileNet [[Pham,](#_bookmark41) [2021].](#_bookmark41) A different approach creates novel CNNs to classify chest X-ray images as positive or negative for COVID-19 [[Hussain et al.,](#_bookmark42) [2021].](#_bookmark42) Different traditional ML approaches have been proposed involving a manual feature extraction stage employing texture or morphological descriptors of the images [[Hussain et al.,](#_bookmark43) [2020,](#_bookmark43) [Pereira et al.,](#_bookmark44) [2020].](#_bookmark44) The availability of data is probably the biggest limitation when designing AI systems to detect COVID-19, although nowadays there are several public image databases, the quality of images and information is highly variable, which makes it difficult for researchers to evaluate their systems on appropriate conditions. Furthermore, there is not a standard benchmark to evaluate and compare the different proposals, which in combination with data variability, makes the reported results difficult to compare with each other.

The paper is organized as follows: Section Survey Methodology explains the criteria used to perform the literature review; Section Development Of The Subject presents the results and the state of the art; and Section Conclusions and Future Work.

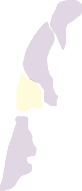
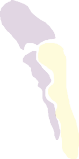
**Without COVID-19**



**With Covid 19**



**With COVID-19 & ROI**



# Survey Methodology

A systematic review of scientific papers was conducted, which explains the design and implementation of CNN, ML algorithms, or segmentation methods, for COVID-19 classification from chest X-ray images.

## Identification of the need for a review

Given the need to develop more efficient and effective diagnostics tools for the COVID-19 disease, many research papers have been written since the beginning of the pandemic. One estimate suggests that more than 200,000 research papers have been published in journals and preprints repositories only in 2020 [[Else,](#_bookmark45) [2020].](#_bookmark45) Therefore there is a need for a new state-of-the-art review.

State-of-the-art works up to March 21, 2021, are summarized in Table [1,](#_bookmark1) where is shown if they specify the corresponding preprocessing techniques and datasets used in the selected works, the date of search, whether or not it is systematic, the number of search databases used, the data modalities included in the work, the number of X-ray related articles analyzed and the number of X-ray related databases described. Table [1](#_bookmark1) shows that some of the available state-of-the-art works do not follow a systematic approach, and only include works up to July 2020 or before, unlike our work that covers from January 1, 2020, to March 21, 2021. Additionally, our work uses the largest number of search databases and some of the available state-of-the-art include a limited amount of papers that identify COVID-19 using X-ray images

The description of the available datasets presented in this work also fills a gap in the literature, since the available works barely describe some of them. We also present the preprocessing techniques and the specific datasets used on each work selected for this literature review, what is not included in any of the available works, along with the models, tasks and results described help researchers to build a complete panorama of the actual strategies used for the detection of COVID-19 using X-ray images. Our work is a complement to many of the available works since they invest more effort in discussing the risk of bias, recommendations, and deficiencies than in the specific methodologies details as we do.

Therefore, our bibliographic review contributes relevant and up-to-date information about the development of AI-based systems to detect COVID-19 from chest X-ray images. It will also offer a baseline for researchers regarding methods, architectures, databases, and current limitations.

## Research questions

In order to describe the state-of-the-art approaches for COVID-19 detection, this paper aims to answer the following questions:

* Which are the different architectures and novel components of CNNs used to detect COVID-19 on chest X-ray images?
* What are the detection performances of COVID-19 on chest X-ray images using CNNs?
* Which digital image databases are the most used for COVID-19 detection on chest X-ray images?
* Which segmentation methods are applied on chest X-ray images for the automatic detection of COVID-19?

## Bibliographic search

The key words chosen for this search are:

* COVID-19.
* Deep Learning.
* Machine Learning.
* Classification.
* Segmentation.
* Chest X-ray.

After defining the search terms, the search string was built with logical operators. Due to COVID-19 being a disease that emerged in late 2019, the search is limited between 2020 - Present, only in the English language. The general search string is: *Covid-19 AND (Machine Learning OR Deep Learning) AND (Classification OR Segmentation) AND X-ray*. Table [2](#_bookmark2) shows the databases and search strings used for the review. The gray literature search included papers with novel COVID-19 classification methods from chest X-ray images.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Article** | **Database** | | **Preprocessing** | **Date of search** | **Date of** | **Systematic** | **Search** | **Articles** | **Databases** |
|  | **specification** | | **specification** |  | **publication** |  | **databases** | **analyzed** | **described** |
| [Rahman](#_bookmark46) | [et al.](#_bookmark46) | * X | | - March 02, 2021 | | X | - | 23 | 12 |
| [[2021]](#_bookmark46) |  |  | |  | |  |  |  |  |
| [Nayak](#_bookmark47) | [et al.](#_bookmark47) | X X | | - March 20, 2021 | | ✓ | 4 | 41 | 1 |
| [[2021a]](#_bookmark47) |  |  | |  | |  |  |  |  |
| [[Wynants et al.,](#_bookmark48) | | X | X | July 01, 2020 | April 07, 2021 | ✓ | 5 | 22 | - |
| [2020]](#_bookmark48) | |  |  |  |  |  |  |  |  |
| [Shi et al.](#_bookmark49) [[2021]](#_bookmark49) | | X | X | March 31, 2020 | April 16, 2020 | X | - | 4 | - |
| [Swapnarekha](#_bookmark50) | | X | X | May 03, 2020 | May 26, 2020 | X | - | 12 | - |
| [et al.](#_bookmark50) [[2020]](#_bookmark50) | |  |  |  |  |  |  |  |  |
| [Waleed Salehi](#_bookmark51) | | X | X | - | June 23, 2020 | X | - | 5 | - |
| [et al.](#_bookmark51) [[2020]](#_bookmark51)  [Albahri et al.](#_bookmark52) | | X | X | May 15, 2020 | June 25, 2020 | ✓ | 4 | 11 | - |
| [[2020]](#_bookmark52) | |  |  |  |  |  |  |  |  |
| [Bansal et al.](#_bookmark53) | | X | X | - | August 01, 2020 | X | - | - | - |
| [[2020]](#_bookmark53) | |  |  |  |  |  |  |  |  |
| [Syeda et al.](#_bookmark54) | | X | X | June 27, 2020 | January 11, 2021 | ✓ | 3 | 22 | 23 |
| [[2021]](#_bookmark54) | |  |  |  |  |  |  |  |  |
| [Roberts et al.](#_bookmark55) | | X | X | - | March 15, 2021 | ✓ | - | 22 | - |
| [[2021]](#_bookmark55) | |  |  |  |  |  |  |  |  |
| Our work | | ✓ | ✓ | March 21, 2021 | - | ✓ | 6 | 23 | 18 |

Table 2: *Databases and search strings for literature review*

|  |  |
| --- | --- |
| **Name of the**  **search database** | **Search string** |
| Scopus | TITLE-ABS-KEY (Covid-19 AND (Machine Learning OR Deep Learning)  AND (Classification OR Segmentation) AND X-ray) |
| Web of Science | (SUBJECT OR TITLE: Covid-19 AND (Machine Learning OR Deep Learning)  AND (Classification OR Segmentation) AND X-ray) |
| SpringerLink | Covid-19 AND (Machine Learning OR Deep Learning)  AND (Classification OR Segmentation) AND X-ray. |
| PubMed | Covid-19 AND (Machine Learning OR Deep Learning)  AND (Classification OR Segmentation) AND X-ray. |
| IEEE Xplore | Covid-19 AND (Machine Learning OR Deep Learning)  AND (Classification OR Segmentation) AND X-ray |
| Google Scholar | Covid-19 AND (Machine Learning OR Deep Learning)  AND (Classification OR Segmentation) AND X-ray. |

## Inclusion and exclusion criteria

The inclusion criteria taken into account are:

* Papers published in Journals.
* Papers written in English.
* Papers found in the databases in Table [2.](#_bookmark2)
* Papers that use DL o ML to detect COVID-19 from chest X-ray images.
* Papers using novel methods for COVID-19 detection. The exclusion criteria taken into account are:
* COVID-19 classification or segmentation papers without application of DL or ML methods.
* COVID-19 classification papers that do not use chest X-ray.
* Papers with methods that include X-ray and CT simultaneously in the methods training.

## Data extraction and synthesis

We conduct a systematic literature review by applying the preferred reporting items for systematic reviews and meta- analyses guidelines (the PRIMA statement), the Fig. [2](#_bookmark3) presents the number of articles obtained on each step of the guidelines [[Moher,](#_bookmark56) [2009].](#_bookmark56)

Based on the search string provided in Table [2,](#_bookmark2) approximately 720 papers are found and 120 papers per database are analyzed, from which 20 relevant papers per database are pre-selected. The pre-selection is done using the number of citations and the novelty of the COVID-19 detection method proposed. It is relevant to clarify that we found 14 repeated papers that are rejected; therefore, the pre-selection of the papers leads to 106 documents. The remaining 106 articles are filtered according to:

* **Title:** After reading the title, 40 are accepted and 66 papers are rejected.
* **Abstract**: After reading the abstract, 28 are accepted and 12 papers are rejected.
* **Full text**: After reading the entire text, 23 papers are accepted and 5 papers are rejected.

Figure [3](#_bookmark4) presents the percentage distribution of the selected papers in the databases.

# Development of the subject

As mentioned in the previous section, a total of 23 research papers are selected for the systematic review. This paper aims to cover novel approaches and offer a general overview of how AI has been applied to COVID-19 diagnosis.

Scopus

30.44%

Google Scholar

26.09%

Scopus

Google Scholar Web of Science IEEE Xplore PubMed Springer

4.34% Springer

8.70%

PubMed

Figure 3: Percentage of papers selected for bibliographic review in different databases.

## X-ray images databases

Table [3](#_bookmark5) presents the most relevant databases used in the selected articles, they are described along with the total number of images, and the classes. This table includes not only databases with COVID-19 samples but also other chest diseases, due to some of the works published regarding the detection of COVID-19, which intend to classify multiple chest conditions. The most often classified illnesses along with COVID-19 are pneumonia viral, pneumonia bacterial, and tuberculosis. The COVID-19 Image Data Collection published by [Paul Cohen et al.](#_bookmark57) [[2020]](#_bookmark57) is the most used database, it was one of the first ones to be released and more images are included over time. It also provides prospective metadata like survival, ICU stay, intubation events, blood tests, location, and freeform clinical notes. The most regularly used databases in this regard with X-ray samples of other chest diseases are ChestX-ray8 (CRX8), CheXpert, Chest X-ray Images (Pneumonia), and Tuberculosis chest X-ray.

Table 3: Image databases for COVID-19 research

|  |  |  |  |
| --- | --- | --- | --- |
| Item | Dataset name | # Images | Classes |
| A  B C  D  E F G H I  J  K L  M  N  O P Q R | HM Hospitales [[HM Hospitals,](#_bookmark58) [2020]](#_bookmark58)  BIMCV-COVID19 [[Vayá et al.,](#_bookmark59) [2021]](#_bookmark59)  Actualmed COVID-19 chest X-rays [[Ac-](#_bookmark60) [tualmed et al.,](#_bookmark60) [2020]](#_bookmark60)  ChinaSet - The ShenzhenSet [[Jaeger et al.,](#_bookmark61) [2014]](#_bookmark61)  Montgomery [[Jaeger et al.,](#_bookmark61) [2014]](#_bookmark61) ChestX-ray8 (CRX8) [[Wang et al.,](#_bookmark62) [2017]](#_bookmark62) CheXpert [[Irvin et al.,](#_bookmark63) [2019]](#_bookmark63)  MIMIC-CXR [[Johnson et al.,](#_bookmark64) [2019]](#_bookmark64) COVID-19 Image Data Collection [P[aul Co-](#_bookmark57) [hen et al.,](#_bookmark57) [2020]](#_bookmark57)  Kermany et al. [[Kermany et al.,](#_bookmark65) [2018]](#_bookmark65)  RSNA, Radiopedia and SIRM [[Dadario,](#_bookmark66) [2020]](#_bookmark66)  RYDLS-20 [[Pereira et al.,](#_bookmark44) [2020]](#_bookmark44)  COVID-19 Radiography Database (Qatar university) [[Rahman et al.,](#_bookmark67) [2020]](#_bookmark67)  NIH Chest X-ray [[Wang et al.,](#_bookmark62) [2017]](#_bookmark62)  Chest X-ray Images (Pneumonia) [Mooney](#_bookmark68) [[2018]](#_bookmark68)  COVID-19 dataset [[Societa Italiana di Ra-](#_bookmark69) [diologia Medica e Interventistica,](#_bookmark69) [2020]](#_bookmark69) CHUAC dataset [[De Moura et al.,](#_bookmark70) [2020]](#_bookmark70)  COVID-19 X rays [[Dadario,](#_bookmark66) [2020]](#_bookmark66) | 5,560  3,013  188  662  138  61,790  4,623  16,399  481  5,840  73  1,144  21,165  108,948  5,863  115  1,616  79 | COVID-19  COVID-19  COVID-19; Normal Pneumonia; Normal  Pneumonia; Normal Pneumonia; Normal Pneumonia Pneumonia  Viral (COVID-19; SARS;  MERS; among others); Bac- terial; Others  Normal (1,575); Pneumo- nia Bacterial (2,771); non- COVID-19 lung infection (1,494)  COVID-19  Normal (1,000); Pneumonia  (144): (MERS (10); COVID-  19 (90); Pneumocystis (11);  SARS (11); Streptococcus  (12); Varicella (10))  COVID-19 (3,616); Nor-  mal (10,192); Viral Pneumo- nia (1,345); Non-COVID-19  lung infection (6,012) Atelectasis; Mass; Car- diomegaly; Nodule; Effu- sion; Normal; Infiltration; Pneumonia  Pneumonia (bacterial and vi- ral); Normal  COVID-19  Normal (728); Pathological  (648); COVID-19 (240) COVID-19 |

## Approaches for the automatic detection of COVID-19 using X-ray images

Table [4](#_bookmark6) presents relevant information regarding the methodology applied in each selected paper, the best results achieved, the AI models used, the image databases involved, the classes, and the preprocessing operations. The

state-of-the-art networks most used are the VGG and ResNet family and the preprocessing steps more prevalent are resizing, normalization and cropping. Most of the models reported have high performances, however, the experimental setups are not always clear due to most of the authors combine multiple databases and then split the resulting set in the training, validation, and test partitions without doing any further clarification. In many cases, the testing of the model is done using very few images and none of the analyzed works perform clinical validation of the methods.

17.39%

Web of Science

Table 4: Implementation details and results of the selected papers for the detection of COVID-19 using chest X-ray images

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Article | Models | Database | Classes | Preprocessing | Best Results |
| [[Apostolopoulos](#_bookmark71)  [and Mpesiana,](#_bookmark71) [2020]](#_bookmark71)  [[Ucar and Kork-](#_bookmark72) [maz,](#_bookmark72) [2020]](#_bookmark72)  [[Ozturk et al.,](#_bookmark73) [2020]](#_bookmark73)  [[Tog˘açar et al.,](#_bookmark74) [2020]](#_bookmark74)  [[Pereira et al.,](#_bookmark44) [2020]](#_bookmark44)  [[Apostolopoulos](#_bookmark75) [et al.,](#_bookmark75) [2020]](#_bookmark75) | VGG16; Mo-  bileNetV2; In- ception; Xcep- tion; Inception- ResNet-V2  COVID  Diagnosis-  Net (based on SqueezeNet)  Dark CovidNet (based on Darknet-19 model)  MobileNetV2; SqueezeNet; Social Mimic optimization method; SVM  SVM; MLP; DT; RF; Hi-  erarchical Clus-HMC  MobileNetV2 | I, J  I, O  F, I  I, M  I, N, L  I, P | COVID-19;  Pneumonia; Normal  COVID-19;  Pneumonia; Normal  (COVID-19;  Normal); (COVID-19;  Pneumonia; Normal) COVID-19;  Pneumonia; Normal  Normal; COVID-  19; MERS; SARS; Vari- cella; Strep- tococcus; Pneumocystis  COVID-19;  Edema; Effu- sion; Emphys; Fibrosis; Pneumonia; Normal | Resize, 200 x  266  Resize, 227 x  227; Shuffled; Normalization with the mean subtracting operation; Con- version to RGB with 8-bit depth N/A  Convertion to jpg format;  Fuzzy Color technique  Manual crop  Resize, 200 x  200 | MobileNetV2:  Accuracy, 94.7%;  Sensitivity, 98.7%;  Specificity, 96.5%  Accuracy, 98.3%;  Specificity, 99.1%;  F1, 98.3%  Binary Accu- racy, 98.1%  Overall accu- racy, 98.3%; COVID-19  Sensitivity, 99.3%; COVID-19  Specificity, 99.4%  Multiclass: COVID-19  class F1 score, 83.3%  Hierarchical: COVID-19  class F1 score, 88.8%  Accuracy between the seven classes of 87.7% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| [[Waheed et al.,](#_bookmark76)  [2020]](#_bookmark76)  [[Khan et al.,](#_bookmark77) [2020]](#_bookmark77)  [[Das et al.,](#_bookmark78) [2020]](#_bookmark78)  [[Toraman et al.,](#_bookmark79) [2020]](#_bookmark79)  [[Blain et al.,](#_bookmark80) [2021]](#_bookmark80)  [[Horry et al.,](#_bookmark81) [2020]](#_bookmark81) | ACGAN;  VGG16  CoroNet  Truncated Inception Net  CapsNet  U-Net; DenseNet121  VGG19 | I, M, R  I, O  I, O, S  I, N  K  I, N | COVID-19;  Normal  (COVID-19;  Normal; Pneumonia bacterial; Pneumo-  nia viral); (COVID-19;  Pneumonia; Normal) COVID-19;  Pneumonia; Tuberculosis; Normal  (COVID-19;  Normal); (COVID-19;  Normal; Pneumonia)  N/A  COVID-19;  Normal; Pneumonia | Resize, 112 x  112 x 3; Normal- ization  Resize, 224 x  224  Resize, 224 x  224 x 3  Resize, 128 x  128; Data aug- mentation  Lung segmenta- tion  Resize, 224 x  224; Histogram equalization us- ing N-CLAHE | Using actual  data: Sensitivity, 69.0%;  Specificity, 95.0%;  Accuracy, 85.0%  Including syn- thetic images: Sensitivity, 95.0%;  Specificity, 90.0%;  Accuracy, 97.0%  Four classes: Accuracy, 93.0%;  Three classes: Accuracy, 95.0%  COVID-19  positive cases: Accuracy, 99.96%;  AUC, 100%  Binary class: Accuracy, 97.2%  Multi-class: Accuracy, 84.2%  Diagnosing alveolar opaci- ties:  Accuracy, 78.5%  Diagnosing interstitial opacities: Accuracy, 90.7%  X-ray: Accuracy, 86.0%  Ultrasound: Accuracy, 100%  CT:  Accuracy, 84.0% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| [[King et al.,](#_bookmark82) [2020]](#_bookmark82)  [[Karar et al.,](#_bookmark83) [2020]](#_bookmark83)  [[Ohata et al.,](#_bookmark20) [2021]](#_bookmark20)  [[Shorfuzzaman](#_bookmark84) [and Hossain,](#_bookmark84) [2020]](#_bookmark84)  [[De Moura et al.,](#_bookmark70) [2020]](#_bookmark70)  [[Nayak et al.,](#_bookmark85) [2021b]](#_bookmark85)  [[Karakanis and](#_bookmark86) [Leontidis,](#_bookmark86) [2021]](#_bookmark86)  [[Albahli and Yar,](#_bookmark87) [2021]](#_bookmark87)  [[Singh et al.,](#_bookmark88) [2021]](#_bookmark88) | Self-  Organizing Feature Map  VGG16;  ResNet50V2; DenseNet169  MobileNet; DenseNet121; Inception- ResNet-V2; Bayes; RF;  MLP; KNN; SVM  Siamese Net- work  DenseNet161  AlexNet; VGG16;  GoogleNet; MobileNetV2; SqueezeNet; ResNet34; ResNet50; InceptionV3 CNN Designed  NasNetLarge; Xception; InceptionV3; Inception- ResNetV2; ResNet50  VGG19; VGG16;  ResNet50; DenseNet161; DenseNet169; Naive Bayes | I  I  I, R, O, N  I, O  Q  F, I  I, O  I, N  Q, R, M, N, G | COVID-19;  Normal  Normal; COVID-19;  Viral Pneumo- nia; Bacterial Pneumonia COVID-19;  Normal  COVID-19;  Normal; Pneumonia  COVID-19;  Pathologi- cal; Normal; Combinations  COVID-19;  Normal  (COVID-19;  Normal); (COVID-19;  Normal; Bacterial Pneumonia)  COVID-19;  normal; 14 other chest diseases  COVID-19;  Pneumonia; Normal | Resize  Resize, 150 x  150  Resize, (224 x  224; 299 x 299;  331 x 331)  Resize; Nor- malization; Histogram- equalization N/A  Normalization  Resize, 224 x  224  Histogram equalization; Lung and heart segmentation  Histogram equalization (CLAHE);  Dynamic im- age filtering (NLMD) | Euclidean  distance of 1.1 between 1st and 2nd winning neurons Accuracy, 99.9%  MobileNet + SVM (Linear): Accuracy, 98.6%;  F1-score, 98.5%  Accuracy, 95.6%;  AUC, 98.9%  Accuracy, 90.3% in  (Normal & Pathological) vs. COVID-19  ResNet34, 98.3%  Binary classifi- cation: Accuracy, 98.7%;  Sensitivity, 100%;  Specificity, 98.3%  First classifier: Accuracy, 96.3%  Second classi- fier:  Accuracy, 87.8%  Accuracy, 98.7% |

*Precision* + *Recall*

## Interpretability and CNN benchmarking

[Nayak et al.](#_bookmark85) [[2021b],](#_bookmark85) analyze the performance of eight state-of-the-art CNNs, they tune the number of trainable layers of the network, nodes, epochs, layers of the classifier placed at the top of the network, the batch size, the learning rate, and the optimizer algorithm, they find that the ResNet family of architectures has the highest classification accuracy; a similar analysis is implemented by [[Majeed et al.,](#_bookmark95) [2020]](#_bookmark95) using transfer learning and 12 state-of-the-art CNNs architectures, a critical analysis that includes the needed time to train each network is also presented, in the binary scenario the best results were obtained using the networks Xception, Inception-ResnetV2, and SqueezeNet; likewise, [Apostolopoulos and Mpesiana](#_bookmark71) [[2020]](#_bookmark71) implement transfer learning to compare 5 state-of-the-art CNNs, the best result are obtained using the VGG19.

[Pham](#_bookmark41) [[2021],](#_bookmark41) compare the performance of state-of-the-art fine-tuned CNNs and recently developed networks like CovidGAN, CoroNet, and DarkCovidNet, using an experimental setup as similar as possible to the original studies. Despite exact comparisons are not possible due to databases updates, it is concluded that similar performances are obtained using relatively small CNNs like AlexNet or SqueeNet and the new and sometimes more complex architectures. [Tartaglione et al.](#_bookmark96) [[2020],](#_bookmark96) provide insights and also raise warnings regarding the generality of the results of COVID-19 classification using deep learning and chest X-ray images. Alternately [Horry et al.](#_bookmark81) [[2020]](#_bookmark81) compare the results achieved training the VGG16 network using CT, X-ray, and Ultrasound chest images to classify COVID-19, pneumonia and healthy subjects. The best result is achieved using Ultrasound images.

One of the main drawbacks of deep learning is the lack of interpretability, which has imitated its application in some areas [[Singh et al.,](#_bookmark107) [2020].](#_bookmark107) That is why, some authors have included visualization methods that could provide credibility and increase trust in users. Particular focus has been given to heat maps. In [[Ozturk et al.,](#_bookmark73) [2020]](#_bookmark73) is proposed a network that provides a heat map along with the classifications, results are evaluated by a radiologist who concludes that ‘The heat map showed a greater concentration area in the X-rays of patients with COVID-19 than the area in which the disease is not seen’. Besides heat maps, t-SNE is also used to improve explainability [[Arias-Londono et al.,](#_bookmark92) [2020],](#_bookmark92) as well as class activation maps [[Ucar and Korkmaz,](#_bookmark72) [2020].](#_bookmark72)

Finally, a widely used and accepted visualization method is the grand cam, which is used in multiple works [[Singh](#_bookmark88) [et al.,](#_bookmark88) [2021,](#_bookmark88) [Liang et al.,](#_bookmark108) [2021].](#_bookmark108)

## Metrics used in the evaluation of algorithms

Most of the metrics defined below are expressed for binary classification tasks in terms of the numbers of True Positive (TP) predictions, True Negatives (TN), False Positives (FP), and False Negatives (FN). TP refers to the positive instances correctly classified as positive; TN refers to the negative instances correctly classified as negative; FP refers to the negative instances incorrectly classified as positive; FN refers to the positive instances incorrectly classified as negative.

* Accuracy: proportion of correct predictions, usually presented as a percentage or as a number from 0 to 1.

*TP* + *TN*

*Accuracy* =

*TP* + *TN* + *FP* + *FN*

(1)

* Specificity: it measures the ability of the classifier to correctly identify the negative instances of a class.

*TN*

*Specificity* =

*TN* + *FP*

(2)

* Recall/Sensitivity: also known as recall, it measures a classifier’s ability to correctly identify the positive instances of a class.

*Recall* =

*TP*

*TP* + *FN*

(3)

* Precision/Positive Predictive Value: also known as precision, it represents the proportion of positive cases among the instances classified as positive.

Given the worldwide impact of the COVID-19 pandemic, it is a priority to develop alternative diagnostic tools available for everyone and offer reliable results. This paper presents a systematic literature review of AI applied to COVID-19 diagnosis using chest X-ray imaging. It includes the different preprocessing techniques, classification methods using ML algorithms, strategies to increase the interpretability of the models, and the articles that perform a critical analysis of the state-of-the-art and the new architecture designed to perform this task.

The main limitation researchers have faced when developing these systems is the quality and availability of data. To overcome this situation, the use of preprocessing techniques, such as histogram equalization, lung segmentation, data augmentation using rotation

or cropping operations, and synthetic data generation using GANs, have been implemented to improve the detection performances of the models. The most frequent approach for the classification of COVID-19 from X-ray images is transfer learning using pre-trained CNNs architectures, such as VGG16, DenseNet121, and ResNet50. Other proposals involve state-of-the-art architectures as feature extractors and traditional ML methods as classifiers. There are also multiple novel CNNs explicitly designed for this task which allows more flexibility and potentially smaller networks by narrowing the scope of the classification task. Overall, literature methods have exceptional results with classification accuracies over 95% and even 98%, however, the test set and the quality of the data, are usually unclear.

Regarding the questions set at the start of the paper, we have shown that (1) conventional image classification CNNs with pre-trained weights using ImageNet, or more complex approaches as Capsule or Siamese networks have been used to diagnose COVID-19 from chest X-ray images; (2) current detection percentages are over 98% accuracy in binary classification (COVID-19 and Normal). However, no clinical trials have been performed in none of these models and the experimental setups are usually unclear; (3) this paper compiles an active set of databases for training and evaluating AI models, despite the relatively high number of available databases of chest X-rays, there is a limited amount of labeled COVID-19 cases, which leads researchers to combine various databases; and in (4) most of the papers that use lung segmentation as a preprocessing step, do so using a U-Net architecture.

Based on the present literature review, we identify possible research opportunities as follows:

* Construct or contribute to databases of chest X-ray images aiming to create a representation of the different characteristics of real-world images, allowing proper benchmarking and future model proposals.
* Develop new CNNs for image segmentation, focusing on the segmentation of lungs and radiological findings associated with COVID-19 disease.
* Broaden the classification scope to detect factors such as severity, or disease progression.
* Design new preprocessing operations or pipelines, taking into account, for example, the removal of artifacts or medical devices such as necklaces, tubes, or ECG lead wires.
* Detect COVID-19, and its outcome considering other clinical variables such as the patient’s history.
* Perform transfer learning based on networks that have been trained for other lung diseases.
* Design architectures or computational elements of CNNs for the detection of COVID-19 from validated

# Appendix

Table [5](#_bookmark8) shows some acronyms and their meanings.

Table 5: Acronyms

## Acronym Description

ACGAN Auxiliary Classifier Generative Adversarial Network AI Artificial Intelligence

CNNs Convolutional Neural Networks

CT Computer tomography

DL Deep Learning

FN False Negatives

FP False Positives

GANs Generative Adversarial Network

K Kappa statistics

ML Machine Learning

RT-PCR Real-time reverse transcription polymerase chain reaction SARS-CoV-2 Severe Acute Respiratory Syndrome coronavirus 2

SVM Support Vector Machine

TN True Negatives

TP True Positive

WHO World Health Organization

**Source code:**

import os

import pathlib

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import matplotlib.image as mpimg

import random

import cv2

from sklearn.model\_selection import train\_test\_split

from tensorflow.keras.models import Sequential

from tensorflow.keras.models import Model

from tensorflow.keras.layers import Dense, Activation, Conv2D, MaxPool2D, Flatten, Dropout, BatchNormalization

from tensorflow.keras.callbacks import EarlyStopping

import tensorflow as tf

from google.colab import files

from sklearn.metrics import classification\_report,confusion\_matrix

In [ ]:

files.upload()

import os

os.environ["KAGGLE\_CONFIG\_DIR"] = "/content"

!kaggle datasets download -d sid321axn/covid-cxr-image-dataset-research

!unzip \\*.zip

for dirpath,dirnames,filenames in os.walk("/content/COVID\_IEEE"):

print(f"there are {len(dirnames)} directories and {len(filenames)} images in '{dirpath}'.")

data\_dir = pathlib.Path("/content/COVID\_IEEE")

class\_names = np.array(sorted([item.name for item in data\_dir.glob("\*")]))

class\_names

def view\_image(target\_dir, target\_class):

target\_folder = target\_dir+target\_class

random\_image = random.sample(os.listdir(target\_folder),1)

print(random\_image)

img = mpimg.imread(target\_folder+"/"+ random\_image[0])

plt.imshow(img, cmap ="gray")

plt.title(target\_class)

plt.axis("off")

print(f"image shape {img.shape}")

return img

img = view\_image("/content/COVID\_IEEE/","virus")

img = view\_image("/content/COVID\_IEEE/","normal")

img = view\_image("/content/COVID\_IEEE/","covid")

data=[]

labels=[]

covid=os.listdir("/content/COVID\_IEEE/covid/")

for a in covid:

image = cv2.imread("/content/COVID\_IEEE/covid/"+a,)

image = cv2.resize(image, (224, 224))

data.append(image)

labels.append(0)

normal=os.listdir("/content/COVID\_IEEE/normal/")

for a in normal:

image = cv2.imread("/content/COVID\_IEEE/normal/"+a,)

image = cv2.resize(image, (224, 224))

data.append(image)

labels.append(1)

virus=os.listdir("/content/COVID\_IEEE/virus/")

for a in virus:

image = cv2.imread("/content/COVID\_IEEE/virus/"+a,)

image = cv2.resize(image, (224, 224))

data.append(image)

labels.append(2)

data = np.array(data) / 255.0

img\_labels = np.array(labels)

X\_train, X\_test, y\_train, y\_test = train\_test\_split(data, img\_labels, test\_size=0.20, random\_state=42)

y\_train = tf.keras.utils.to\_categorical(y\_train, num\_classes=3)

y\_test = tf.keras.utils.to\_categorical(y\_test, num\_classes=3)

model = Sequential()

*#Block Number 1*

model.add(Conv2D(input\_shape = (224,224,3), filters=32,padding="same", kernel\_size= (3,3)))

model.add(Activation("relu"))

model.add(Conv2D(filters=32,padding="same", kernel\_size= (3,3)))

model.add(Activation("relu"))

model.add(MaxPool2D((2,2)))

*#Block Number 2*

model.add(Conv2D(filters=64,padding="same", kernel\_size= (3,3)))

model.add(Activation("relu"))

model.add(Conv2D(filters=64,padding="same", kernel\_size= (3,3)))

model.add(Activation("relu"))

model.add(MaxPool2D((2,2)))

*#Block Number 3*

model.add(Conv2D(filters=128,padding="same", kernel\_size= (3,3)))

model.add(Activation("relu"))

model.add(Conv2D(filters=128,padding="same", kernel\_size= (3,3)))

model.add(Activation("relu"))

model.add(MaxPool2D((2,2)))

model.add(MaxPool2D((2,2)))

model.summary()

**Github link**:

**<https://github.com/naanmudhalvan-SI/PBL-NT-GP--22030-1683723449.git>**

**Project demo link:**

[**https://drive.google.com/file/d/1hFqArjerHLxOffC32s06IhuLDOPHhzqp/view?usp=drive\_link**](https://drive.google.com/file/d/1hFqArjerHLxOffC32s06IhuLDOPHhzqp/view?usp=drive_link)