CovidVision: Advanced COVID-19 Detection from Lung-X-rays with Machine Learning or Deep Learnings

*A PROJECT REPORT*

Submitted By

**Team Leader :** ANKIT KUMAR

**Team member :** JEHAN MANDAL

**Team member :** ASHOKKUMAR M

**Team member :** DHANUSH A

**Team ID :** NM2023TMID05684

**BHARATHIDASAN ENGINEERING COLLEGE**

**PROJECT CONTENTS**

**TITLE**

1. **INTRODUCTION**

1.1 Project Overview

1.2 Purpose

2. **IDEATION & PROPOSED SOLUTION**

2.1 Problem Statement Definition

2.2 Empathy Map Canvas

2.3 Ideation & Brainstorming

2.4 Proposed Solution

3. **REQUIREMENT ANALYSIS**

3.1 Functional requirement

3.2 Non-Functional requirements

**4. PROJECT DESIGN**

4.1 Data Flow Diagrams

4.2 Solution & Technical Architecture

4.3 User Stories

**5. CODING & SOLUTIONING (Explain the features added in the project along with code)**

5.1 Feature 1

5.2 Feature 2

5.3 Database Schema (if Applicable)

**6. RESULTS**

6.1 Performance Metrics

**7. ADVANTAGES & DISADVANTAGES**

**8. CONCLUSION**

**9. FUTURE SCOPE**

**10. APPENDIX**

Source Code

GitHub & Project Video **Demo** Link

**Introduction**

**Project OverView**

COVID-19 (coronavirus disease 2019) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which is a strain of coronavirus. The disease was officially announced as a pandemic by the World Health Organization(WHO) on 11 March 2020. Given spikes in new COVID-19 cases and the re-opening of daily activities around the world, the demand for curbing the pandemic is to be more emphasized. Medical images and artificial intelligence (AI) have been found useful for rapid assessment to provide treatment of COVID-19 infected patients. The PCR test may take several hours to become available, information revealed from the chest X-ray plays an important role for a rapid clinical assessment. This means if the clinical condition and the chest X-ray are normal, the patient is sent home while awaiting the results of the etiological test. But if the X-ray shows pathological findings, the suspected patient will be admitted to the hospital for close monitoring. Chest X-ray data have been found to be very promising for assessing COVID-19 patients, especially for resolving emergency-department and urgent-care-center overcapacity. Deep-learning (DL) methods in artificial intelligence (AI) play a dominant role as high-performance classifiers in the detection of the disease using chest X-rays.

* One of the biggest challenges following the Covid-19 pandemic is the detection of the disease in patients. To address this challenge we have been using the Deep Learning Algorithm to build an image recognition model that can detect the presence of Covid-19 from an X-Ray or CT-Scan image of a patient's lungs.
* Transfer learning has become one of the most common techniques that has achieved better performance in many areas, especially in medical image analysis and classification. We used Transfer Learning techniques like Inception V3,Resnet50,Xception V3 that are more widely used as a transfer learning method in medical image analysis and they are highly effective.

**Purpose**

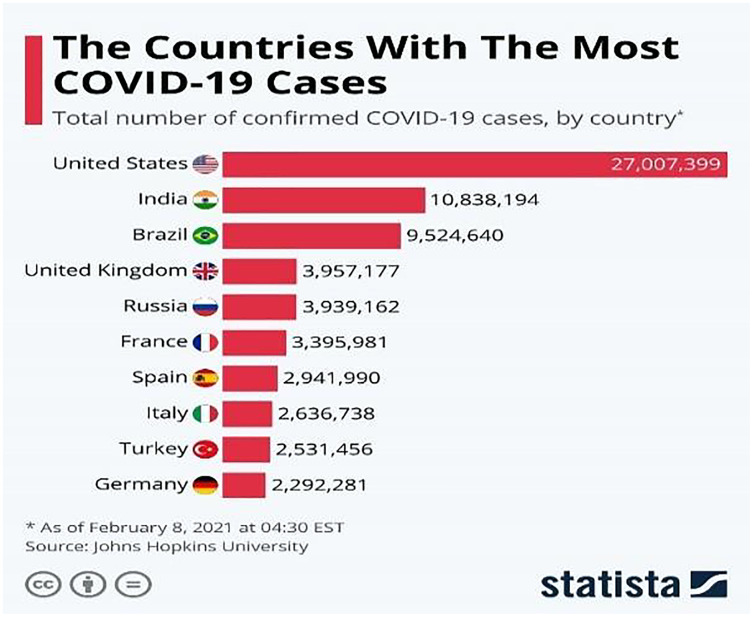
COVID-19 is a widespread deadly virus that directly affects the human lungs. The spread of COVID-19 did not stop at humans but also reached animals, so it was necessary to limit it is spread and diagnose cases quickly by applying a quarantine to the infected people. Recently x-ray lung images are used to determine the infection and from here the idea of this research came to use deep learning techniques to analyze x-ray lung images publicly available on Kaggle to possibly detect COVID-19 infection. In this article, we have proposed a method to possibly detect the COVID-19 by analyzing the X-ray images and applying a number of deep learning pre-trained models such as InceptionV3, DenseNet121, ResNet50, and VGG16, and the results are compared to determine the best performance model and accuracy with the least loss for our dataset. Our evaluation results showed that the best performing model for our dataset is ResNet50 with accuracies of 99.99%, 99.50%, and 99.44% for training, validation, and testing respectively followed by DenseNet121, InceptionV3, and finally VGG16.

**Keywords:**Deep learning, Supervised learning, Classification, Transfer learning, COVID-19

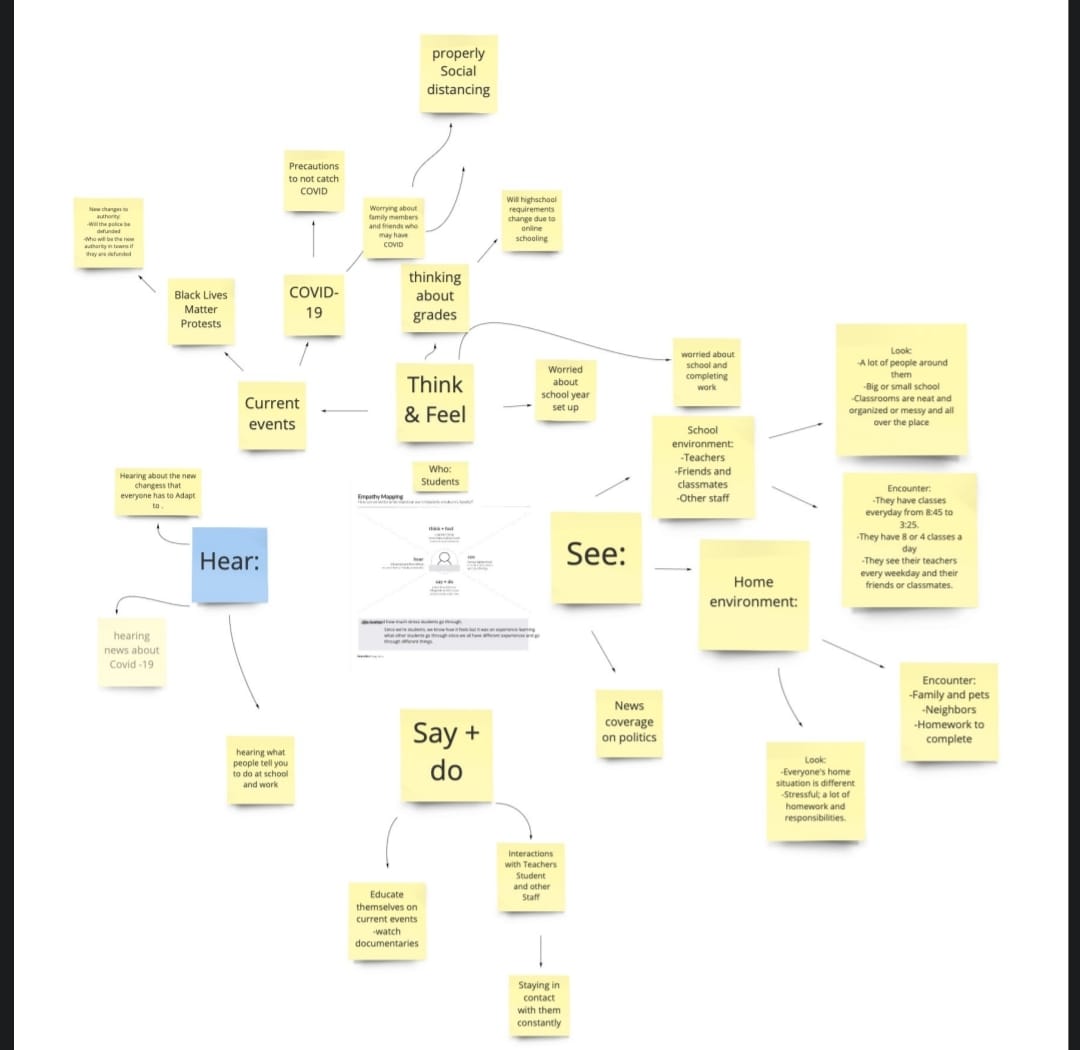
**IDEATION & PROPOSED SOLUTION**

Problem Statement Definition

The COVID-19 ([WHO, 2019](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/#ref-28)) virus is a fast-spreading disease that has appeared all over the world; [Fig. 1](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/figure/fig-1/) ([Armstrong & Richter, 2021](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/#ref-2)) represents the high spreading of COVID-19 in many countries as of February 8, 2021. It affects all aspects of life, economically, educationally, healthily, socially, and, most dangerously, human health, as it has caused the death of millions of people ([Maatuk et al., 2022](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/" \l "ref-13)). There are two ways to control the spread of the virus: one is to take precautions, and the other is to take the vaccinations that are being specifically developed to defend against the disease. Studies have shown that vaccinations could cause allergic responses ([Schumaker et al., 2020](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/" \l "ref-20)), thus some people may not be allowed to take it and would lead to the other way which is the early detection of the disease by chest x-ray to develop a treatment plan which provides a great opportunity for recovery and to prevent its spread



Emphathy Map Canvas

****

Ideation & Brainstorming

**Brainstorm & Idea Prioritization Template:**

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions.

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

Reference: <https://www.mural.co/templates/empathy-map-canvas>

**Step-1: Team Gathering, Collaboration and Select the Problem Statement**

Graphical user interface, application

Description automatically generated

**Step-2: Brainstorm, Idea Listing and Grouping**Graphical user interface, treemap chart

Description automatically generatedProposed Solution

**Proposed Solution Template:**

Project team shall fill the following information in proposed solution template.

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Parameter** | **Description** |
|  | Problem Statement (Problem to be solved) | Evaluate the CovidVision: To Advanced COVID-19 Detection from Lung X-rays with Machine Learning or Deep Learnings |
|  | Idea / Solution description | The COVID-19 ([WHO, 2019](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/#ref-28)) virus is a fast-spreading disease that has appeared all over the world; [Armstrong & Richter, 2021](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/#ref-2)represents the high spreading of COVID-19 in many countries as of February 8, 2021. It affects all aspects of life, economically, educationally, healthily, socially, and, most dangerously, human health, as it has caused the death of millions of people ([Maatuk et al., 2022](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/" \l "ref-13)). There are two ways to control the spread of the virus: one is to take precautions, and the other is to take the vaccinations that are being specifically developed to defend against the disease. Studies have shown that vaccinations could cause allergic responses ([Schumaker et al., 2020](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/" \l "ref-20)), thus some people may not be allowed to take it and would lead to the other way which is the early detection of the disease by chest x-ray to develop a treatment plan which provides a great opportunity for recovery and to prevent its spread. |
|  | Novelty / Uniqueness | COVID-19 is a widespread deadly virus that directly affects the human lungs. The spread of COVID-19 did not stop at humans but also reached animals, so it was necessary to limit it is spread and diagnose cases quickly by applying a quarantine to the infected people. |
|  | Social Impact / Customer Satisfaction | Social constraints in the COVID-19 pandemic force individuals to adapt to isolation and increase the prevalence of violence in the family, depression, anxiety, post-traumatic stress disorder. |
|  | Business Model (Revenue Model) | In its simplest form, a business model is how an organisation makes money.  All businesses need a business model to ensure their organisation is viable, and that all the elements that form a profitable business are in place and working cohesively together to make a profit.  A business model also needs to be flexible, to ensure that an organisation can respond to changes in the market and keep making money.  This need for organisational flexibility has been tested to the maximum over the past year, with COVID-19 disrupting the operations of almost every business worldwide. According to [research by McKinsey](https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/how-covid-19-has-pushed-companies-over-the-technology-tipping-point-and-transformed-business-forever), companies responded to a range of COVID-19 related changes much more quickly than they ever thought possible before the crisis.  Most of these coronavirus-induced changes have involved increased digitization. In fact, [more than half of businesses are now investing in technology for competitive advantage](https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/how-covid-19-has-pushed-companies-over-the-technology-tipping-point-and-transformed-business-forever) or refocusing their entire business around digital technologies. |
|  | Scalability of the Solution | When a business owner wants to improve their marketing and data management, they use data solutions to access and utilize data effectively. Businesses can change their marketing plans in an instant, and they can use scalable models to manage their business models.  Companies that were growing before the COVID-19 pandemic must cut back so that they can save money and survive. The tips below show how scalable resources and services are saving businesses.  **Data Management Solutions**  **Managing Customer Data**  **Scalable Material Ordering**  **Scalable Office Employee Work**  **Scaling Back Your Office Management Expenses** |

**REQUIREMENT ANALYSIS**

Functional Requirement

**Functional Requirements:**

Following are the functional requirements of the proposed solution.

|  |  |  |
| --- | --- | --- |
| **FR No.** | **Functional Requirement (Epic)** | **Sub Requirement (Story / Sub-Task)** |
| FR-1 | Scope | This International Standard specifies functional requirements for a self-initiated disease symptom checker app to be used for early screening and control of pandemics such as COVID-19.  There are four functional components specified in the standard: 1) self-registration; 2) symptom checking; 3) guidance on screening stations; and 4) health consultation. |
| FR-2 | Term and Condition ( Confirmed case) | person who has been confirmed to be infected with the infectious disease pathogen according to the diagnostic testing standard, regardless of clinical manifestations |
| FR-3 | Isolation | separates sick people with a contagious disease from people who are not sick |
| FR-4 | Pandemic | worldwide spread of a disease |
| FR-5 | quarantine | separates and restricts the movement of people who were exposed to a contagious disease to see if they become sick. These people may have been exposed to a disease and do not know it, or they may have the disease but do not show symptoms |
| FR-6 | Self-Symptom Checker | ool that enables individuals to record and report pandemic-related symptoms, provide information on screening test sites, and offer consultation channels that link them to health experts for disease spread containment |

**Non-functional Requirements:**

Following are the non-functional requirements of the proposed solution.

|  |  |  |
| --- | --- | --- |
| **FR No.** | **Non-Functional Requirement** | **Description** |
| NFR-1 | Availablity | COVID-19 is a respiratory condition caused by a coronavirus. Some people are infected but don’t notice any symptoms (doctors call that being asymptomatic). Most people will have [mild symptoms](https://www.webmd.com/lung/covid-19-symptoms) and get better on their own. But some will have severe problems, such as [trouble breathing](https://www.webmd.com/lung/breathing-problems-causes-tests-treatments). The odds of more serious symptoms are higher if you’re older or have another health condition like [diabetes](https://www.webmd.com/diabetes/default.htm) or heart disease. |
| NFR-2 | Symptoms | The most common things people who become ill with COVID-19 have include:   * Fever or chills * A dry cough and shortness of breath * Feeling very tired * Muscle or body aches * Headache * A loss of taste or smell * Sore throat * Congestion or runny nose * Nausea or vomiting * Diarrhea |
| NFR-3 | Testing | Get tested for COVID-19 if:   * You’ve had symptoms of the virus * You’ve come into close contact with a person who has COVID-19 (take a test at least 5 days after you last saw the individual) * You’ve been asked to get tested by your school, health care provider, workplace, state, local, tribal, or territorial health department (regardless of your vaccination status) |
| NFR-4 | Prevention | Until you’re vaccinated, be sure to take these steps to prevent COVID-19:   * [Wash your hands](https://www.webmd.com/men/video/dirty-truth-handwashing) often, for at least 20 seconds each time, with soap and [water](https://www.webmd.com/a-to-z-guides/features/wonders-of-water). * Use an alcohol-based sanitizer with at least 60% alcohol if you don't have soap and water handy. * Limit your contact with other people. Stay at least 6 feet away from others if you have to go out. * Wear a well-fitted protective face mask in public places. * [Avoid people who are sick](https://www.webmd.com/lung/coronavirus_quarantine). * Don’t touch your [eyes](https://www.webmd.com/eye-health/ss/slideshow-eye-conditions-overview), nose, or [mouth](https://www.webmd.com/oral-health/rm-quiz-what-do-you-know-about-your-teeth) unless you’ve just washed your hands. * Regularly clean and disinfect surfaces that you touch a lot. |
|  |  |  |
|  |  |  |

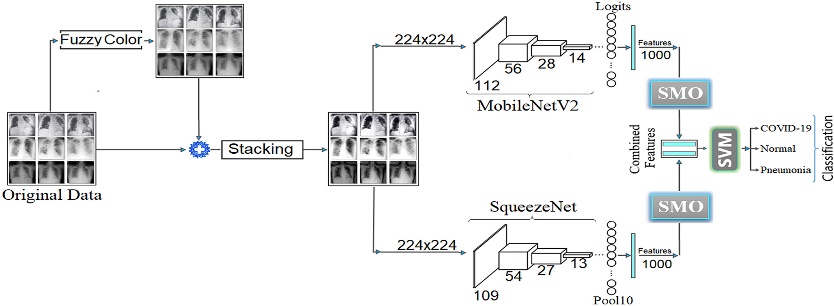
**PROJECT DESIGN**

Data Flow Diagram

A Data Flow Diagram (DFD) is a traditional visual representation of

the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

Example(Simplified)

****

# Example: Deep Learning System for Detection of COVID-19 Diagnosis Aid Using X-ray

Solution & Technical Architecture

**Solution Architecture:**

-Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions.

-The system aims to improve transfer learning and model integration by categorizing Lungs X-ray pictures into three categories:

-Normal, COVID-19, and infectious pneumonia. Choose the models ResNet-101 and ResNet-152 with best outcomes for fusion based on accuracy and loss value, and apathetically enhance its weight percentage during the training procedure.

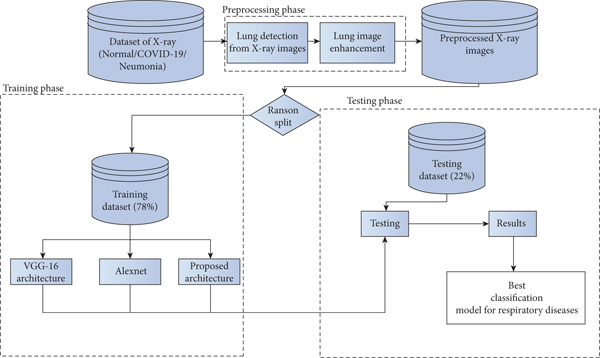
-The system can obtain 97 percent accuracy in Lungs X-Ray pictures post training. In the diagnosing of lung nodules, this technique offers a higher accuracy than radiologist.

-It can assist radiologists enhance their efficiency levels and diagnostic performance as an additional diagnostic technique.

-Keywords: COVID-19, pneumonia, x-ray, convolutional neural networks, Coronavirus.

**Example - Solution Architecture Diagram**

# *Figure 1: Architecture and dataflow diagram for Covid-19 using deep learning on lungs X-ray*



USER STORIES

| **User Type** | **Functional Requirement (Epic)** | **User Story Number** | **User Story / Task** | **Acceptance criteria** | **Priority** | **Team Member** |
| --- | --- | --- | --- | --- | --- | --- |
| Patient (Offline & Online) | Registration | USN-1 | I can fill the form on spot or through online for registering after detection of covid-19 | I can fill the form for admitting the patient | High | Jehan |
|  | Paper Work | USN-2 | Confirm that all details are written correctly in form | I can confirm details while admitting | High | Jehan |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Follow Up | USN-3 | Ward is allocated, shift the patient to specific room safely by keeping distance | Shift the patient to allocated ward | High | Ankit |
|  | Review Medication | USN-4 | Provide proper equipment and medicine | I can provide medical equipment | Medium | Ankit |
| Patient (Billing) | Payment | USN-5 | To pay bills through online or offline mode | Billing for the treatment | Medium | Ashok |
|  | Discharge | USN-6 | Discharge the patient after condition of health become better | Make sure last paperwork for discharge is completed | Medium | Dhanush |

**CODING & SOLUTIONING (Explain the features)**

### Dataset

Two publicly available datasets on Kaggle were used and merged in this research. The COVID-19 Radiography Dataset consists of 21,178 images of lung x-rays distributed over four classes that are normal, viral pneumonia, lung opacity, and COVID-19. The second dataset was called COVID-19 X-ray, which has 2,133 categorized into three categories which are normal, pneumonia, and COVID-19. Because only COVID-19 disease is discussed in our work, the images of the other classes are discarded. [Table 1](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/table/table-1/) shows the distribution of the data over all the classes of the two datasets and [Table 2](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/table/table-2/) shows the distributions after discarding irrelevant classes.

### Table 1

**Original dataset distribution.**

| **Class** | **Dataset 1** | **Dataset 2** |
| --- | --- | --- |
| Normal | 10,192 | 711 |
| Lung opacity | 6,012 | – |
| Pneumonia | 1,345 | 711 |
| COVID | 3,616 | 711 |
| Total | 21,178 | 2,133 |

### Table 2

**Distribution of data after filtering.**

| **Class** | **Dataset 1** | **Dataset 2** | **Full dataset** |
| --- | --- | --- | --- |
| Normal | 10,192 | 711 | 10,903 |
| COVID | 3,616 | 711 | 4,327 |
| Total | 13,808 | 1,422 | 15,230 |

### Pre-processing and splitting

Several steps are taken into consideration when data preparation for the model is required. In this article, four main steps are applied to the acquired dataset to prepare them for training. The first step was to resize all the acquired images into (224 × 224 × 1) images. The value 224 was chosen for the number of pixels as all the pre-trained networks that were used in this article have been trained on the ImageNet dataset that has the size of 224 × 224 and because the images are x-rayed images, then changing the number of channels in the images to one making them in greyscale is necessary.

After resizing the images, the images were renamed because some did not have meaningful names. For that reason, all the images inside a specific folder were renamed to the category that they belong to. It is known in deep learning that the model trains on a set of data, then it is validated on another set and finally is tested on totally new data, thus data splitting comes in handy to help to achieve this flow of training process of deep learning models.

Nevertheless, the imbalanced data problem still exists and needs to be dealt with. The other approach to solve this challenge is to select a specific number of samples for each class where the number of samples may differ, but they should be close to each other. Several splitting combinations are tried to overcome the imbalance issue in the data set like 1,500, 2,000, 3,000, and 3,500 for the training set and the rest for validation and testing. The best split size that received good results is 3,500 for COVID and 5,000 for the normal group for the training and 200 for validation and 626 for testing and the rest is discarded. So, in this research, the results that are provided are done on these splitting sizes which are done randomly, and [Table 3](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/table/table-3/) shows the distribution of the images across the three types of datasets.

### Table 3

**Distribution of data after splitting.**

| **Class** | **Dataset 1** | **Dataset 2** | **Full dataset** |
| --- | --- | --- | --- |
| Normal | 10,192 | 711 | 10,903 |
| Class | Training | Validation | Testing |
| Normal | 5,000 | 200 | 626 |
| COVID | 3,500 | 200 | 626 |
| Total | 8,500 | 400 | 1,252 |

### Transfer learning

As mentioned in the introduction section, transfer refers to using pre-trained networks on a large dataset as the foundation for creating a new network architecture that may be fine-tuned and used on a new dataset for a specific task. This allows the use of weights that were obtained during the training process on the large dataset and requires fewer data to be trained on for the new task, which makes training the network easier and more efficient. Transfer learning models are widely used with pretrained models for large and complex picture classification competitions, such as the ImageNet competition.

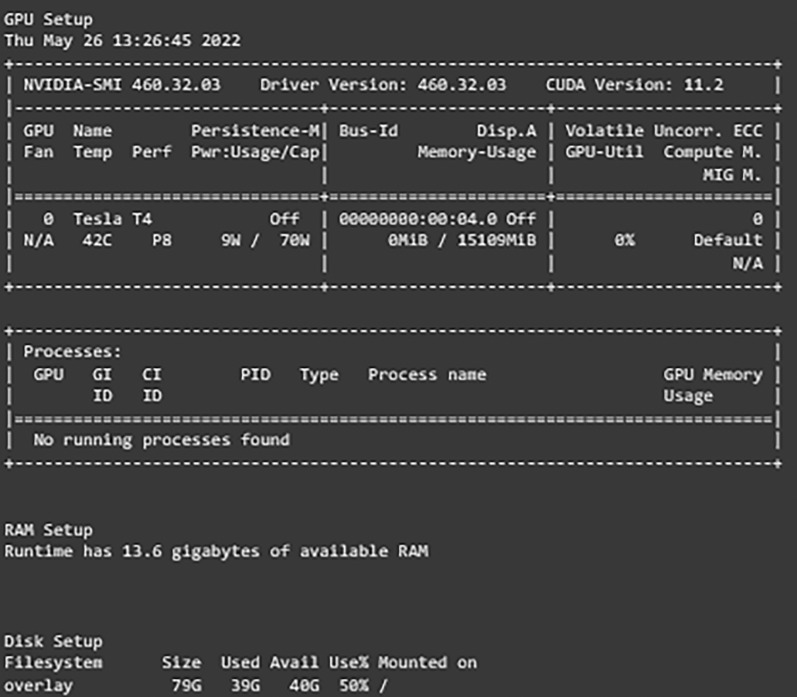
### Experiments and evaluation

After setting and implementing the architecture for the networks, it is time to run experiments on them. Many hyperparameters can affect the learning process of the network like the optimization function, activation function, learning rate, batch size, dropout percentage, and others. In our work, experiments are run with different values for learning rate, dropout percentage, batch sizes, and decay for learning rate. [Table 4](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/table/table-4/) shows the settings of the experiment. The first experiment included running each model with both learning rates and then choosing the one with better results. After that, another experiment with the decay values and freezing the learning rate to the one chosen earlier and comparing the results.

### Table 4

**Expirement settings.**

| **Parameter** | **Value** |
| --- | --- |
| Learning rate | 0.001/0.0001 |
| Decay | Yes/No |
| Dropout percentage | 0.5/0.2 |
| Batch size | 128/64/32 |



**RESULTS**

In this section, the results of the experiments are presented and compared to leverage one of the networks to be the best-suited network for our dataset and use case. The comparisons are based on the training, validation, and testing loss. The network with the lowest testing loss and close to the validation and training losses is considered to be the best one as it is proof that the network does not overfit or underfit. Furthermore, in case of two networks have the same losses, comparisons are expanded to include reports such as classification report data and confusion matrix evaluation to determine the best one.

### Results of transfer learning networks

The results of experiments using different hyperparameters values are shown and the best settings for each one are selected. The format of the model column of the provided tables is as follows: Model Name, Learning Rate, With decay or No decay, Dropout Percentage, and Batch Size.

The first network, “VGG16”, is discussed and shown in [Table 5](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/table/table-5/). From our experiments, it is noticed that decreasing the learning rate while applying the decay rate yields better results. Furthermore, choosing a higher dropout percentage reduces the differences between the validation loss and testing loss which reduces the effects of overfitting. Finally, increasing or decreasing the batch size caused the network to misclassify some of the images. In conclusion, the best set of hyperparameters for the VGG16 network is found to be using 0.0001 as the learning rate with applying decay on it, 0.5 dropout percentage, and a batch size of 64.

### Table 5

**VGG resutls.**

| **Model** | **Training accuracy (%)** | **Training loss** | **Validation accuracy (%)** | **Validation loss** | **Testing accuracy (%)** | **Testing loss** |
| --- | --- | --- | --- | --- | --- | --- |
| VGG16\_0.001\_WD\_0.5\_64 | 98.93 | 0.03 | 98 | 0.05 | 97.28 | 0.09 |
| VGG16\_0.001\_WD\_0.2\_64 | 98.28 | 0.05 | 98.75 | 0.05 | 96.88 | 0.1 |
| VGG16\_0.001\_ND\_0.5\_64 | 97.27 | 0.09 | 97.25 | 0.09 | 96.65 | 0.1 |
| VGG16\_0.001\_WD\_0.5\_64 | 98.66 | 0.04 | 98.00 | 0.04 | 97.60 | 0.09 |
| VGG16\_0.0001\_ND\_0.2\_64 | 99.05 | 0.04 | 97.75 | 0.08 | 97.28 | 0.09 |
| VGG16\_0.0001\_WD\_0.2\_64 | 99.16 | 0.04 | 97.50 | 0.07 | 97.04 | 0.11 |
| VGG16\_0.0001\_ND\_0.5\_64 | 99.75 | 0.01 | 97.50 | 0.08 | 97.52 | 0.09 |
| VGG16\_0.0001\_WD\_0.5\_64 | 99.06 | 0.03 | 97.25 | 0.06 | 98.00 | 0.07 |
| VGG16\_0.0001\_WD\_0.5\_32 | 98.21 | 0.06 | 97.75 | 0.09 | 97.20 | 0.09 |

The next networks that are taken into consideration are ResNet50 shown in [Table 6](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/table/table-6/), and InceptionV3 presented in [Table 7](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/table/table-7/). These two networks are discussed together as they have the same behavior and best set of hyperparameters. The same behavior of adjusting the learning rate and decay is observed. On the other hand, in ResNet50 choosing a lower dropout percentage and batch size improved the performance of the network significantly. Thus, the best set of hyperparameters for the ResNet50 and InceptionV3 networks is found to be using 0.0001 as the learning rate with applying decay on it, 0.2 dropout percentage, and batch size of 32.

### Table 6

**ResNet50 results.**

| **Model** | **Training accuracy (%)** | **Training loss** | **Validation accuracy (%)** | **Validation loss** | **Testing accuracy (%)** | **Testing loss** |
| --- | --- | --- | --- | --- | --- | --- |
| ResNet\_0.001\_ND\_0.2\_64 | 99.62 | 0.01 | 99 | 0.04 | 98.16 | 0.11 |
| ResNet\_0.001\_WD\_0.2\_64 | 99.16 | 0.02 | 99.25 | 0.03 | 98.40 | 0.09 |
| ResNet\_0.001\_ND\_0.5\_64 | 99.42 | 0.02 | 98.75 | 0.03 | 98.80 | 0.08 |
| ResNet\_0.001\_WD\_0.5\_64 | 99.99 | 0 | 99.50 | 0.01 | 98.88 | 0.06 |
| ResNet\_0.0001\_ND\_0.2\_64 | 99.76 | 0.01 | 100 | 0.01 | 98.88 | 0.03 |
| ResNet\_0.0001\_WD\_0.2\_64 | 99.80 | 0.01 | 99.50 | 0.02 | 99.12 | 0.03 |
| ResNet\_0.0001\_ND\_0.5\_64 | 99.56 | 0.01 | 99.00 | 0.03 | 99.12 | 0.03 |
| ResNet\_0.0001\_WD\_0.5\_64 | 99.95 | 0 | 99.75 | 0.01 | 98.96 | 0.04 |
| ResNet\_0.0001\_WD\_0.2\_32 | 99.99 | 0 | 99.50 | 0.02 | 99.44 | 0.02 |

### Table 7

**InceptionV3 results.**

| **Model** | **Training accuracy (%)** | **Training loss** | **Validation accuracy (%)** | **Validation loss** | **Testing accuracy (%)** | **Testing loss** |
| --- | --- | --- | --- | --- | --- | --- |
| InceptionV3\_0.001\_ND\_0.2\_64 | 99.34 | 0.02 | 99 | 0.05 | 98.56 | 0.04 |
| InceptionV3\_0.001\_WD\_0.2\_64 | 99.12 | 0.02 | 99.25 | 0.03 | 98.16 | 0.07 |
| InceptionV3\_0.001\_ND\_0.5\_64 | 98.81 | 0.04 | 97.50 | 0.08 | 94.33 | 0.2 |
| InceptionV3\_0.001\_WD\_0.5\_64 | 99.59 | 0.01 | 98 | 0.05 | 98.40 | 0.05 |
| InceptionV3\_0.0001\_ND\_0.2\_64 | 99.96 | 0 | 99.25 | 0.03 | 98.16 | 0.06 |
| InceptionV3\_0.0001\_WD\_0.2\_64 | 99.86 | 0 | 99 | 0.03 | 98.24 | 0.06 |
| InceptionV3\_0.0001\_ND\_0.5\_64 | 99.86 | 0.01 | 98.50 | 0.03 | 98 | 0.06 |
| InceptionV3\_0.0001\_WD\_0.5\_64 | 99.89 | 0.01 | 99.25 | 0.03 | 98.56 | 0.05 |
| InceptionV3\_0.0001\_WD\_0.5\_32 | 99.34 | 0.02 | 99.50 | 0.02 | 98.96 | 0.04 |

The final network to be examined is DenseNet121 displayed in [Table 8](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/table/table-8/). Two sets of hyperparameters are found to give the same results (DenseNet121-0.001-ND-0.5-64 and DenseNet121-0.0001-WD-0.2-64), this observation is proof of the need to study and experiment with different values as modifying some of them results in networks to act in the same manner. That being said, only one set is chosen to be the best set. Even though the two sets yield the same results, when looking at the learning curves of the networks it can be noticed that DenseNet121-0.0001-WD-0.2-64 had more stable learning whereas DenseNet121-0.001-ND-0.5-64 had a lot of ups and drops in the loss during learning.

### Table 8

**DenseNet121 results.**

| **Model** | **Training accuracy (%)** | **Training loss** | **Validation accuracy (%)** | **Validation loss** | **Testing accuracy (%)** | **Testing loss** |
| --- | --- | --- | --- | --- | --- | --- |
| DenseNet\_0.001\_ND\_0.2\_64 | 99.92 | 0 | 99.50 | 0.01 | 99.12 | 0.03 |
| DenseNet\_0.001\_WD\_0.2\_64 | 99.54 | 0.01 | 99.50 | 0.04 | 98.24 | 0.05 |
| DenseNet\_0.001\_ND\_0.5\_64 | 99.99 | 0 | 99.50 | 0.02 | 99.44 | 0.02 |
| DenseNet\_0.001\_WD\_0.5\_64 | 99.22 | 0.02 | 99.25 | 0.04 | 98.80 | 0.05 |
| DenseNet\_0.001\_ND\_0.5\_32 | 99.24 | 0.02 | 99.50 | 0.01 | 98.56 | 0.09 |
| DenseNet\_0.001\_ND\_0.5\_128 | 99.69 | 0.01 | 99.50 | 0.02 | 98.72 | 0.1 |
| DenseNet\_0.0001\_ND\_0.2\_64 | 99.94 | 0 | 99.75 | 0 | 99.28 | 0.03 |
| DenseNet\_0.0001\_WD\_0.2\_64 | 99.93 | 0 | 99.50 | 0.01 | 99.36 | 0.02 |

Thus, we conclude that the best set of hyperparameters for the DenseNet121 network is found to be using 0.0001 as the learning rate with applying decay on it, 0.2 dropout percentage, and a batch size of 64 is reached. Further explanations are found in [Table 9](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/table/table-9/), which represents the best set of values for hyperparameters of each network is provided.

### Table 9

**Best set of hyper parameters.**

| **Model** | **Learning rate** | **Decay (W/N)** | **Dropout** | **Batch size** |
| --- | --- | --- | --- | --- |
| VGG16 | 0.0001 | W | 0.5 | 64 |
| ResNet50 | 0.0001 | W | 0.5 | 32 |
| InceptionV3 | 0.0001 | W | 0.5 | 32 |
| DenseNet121 | 0.0001 | W | 0.2 | 64 |

### Comparisons of different networks

As shown above, the best-performing networks are compared to nominate the best network based on our dataset and experiments to be used in other applications. [Table 10](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/table/table-10/) summarizes the results of all the experiments and presents the best-performing network of each type

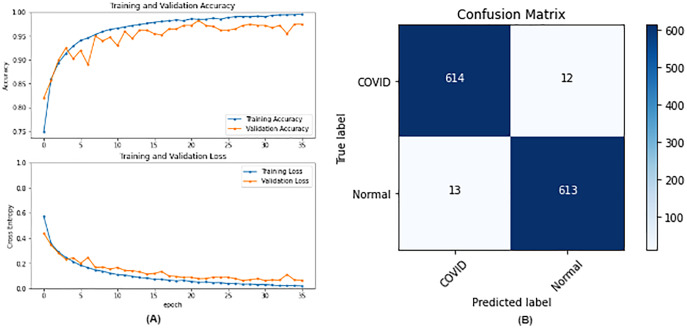
### Table 10

**Comparison of best performance models.**

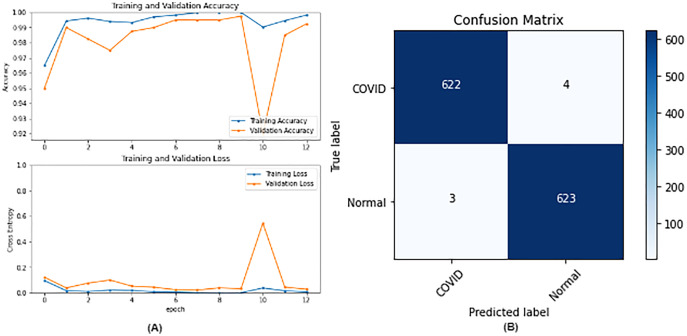
| **Model** | **Training accuracy (%)** | **Training loss** | **Validation accuracy (%)** | **Validation loss** | **Testing accuracy (%)** | **Testing loss** |
| --- | --- | --- | --- | --- | --- | --- |
| VGG16\_0.0001\_WD\_0.5\_64 | 99.06 | 0.03 | 97.25 | 0.06 | 98.00 | 0.07 |
| ResNet\_0.0001\_WD\_0.2\_32 | 99.99 | 0 | 99.50 | 0.02 | 99.44 | 0.02 |
| InceptionV3\_0.0001\_WD\_0.5\_32 | 99.34 | 0.02 | 99.50 | 0.02 | 98.96 | 0.04 |
| DenseNet\_0.0001\_WD\_0.2\_64 | 99.93 | 0 | 99.50 | 0.01 | 99.36 | 0.02 |

**PERFORMANCE METRICS**

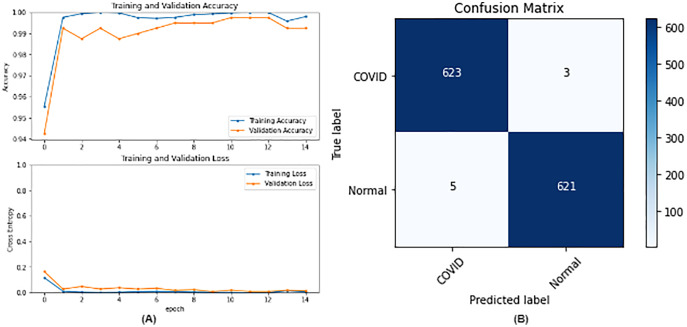
rom the table, it is noticed that ResNet50 and DenseNet121 have very close results and outperform VGG16 and InceptionV3. The two networks have impressive results, and both can be used in other applications. But in our work, only one is chosen to be the best-fitted network. If confusion matrix data is considered, ResNet50 misclassifies seven images whereas DenseNet121 misclassifies eight images. Thus, ResNet50 is nominated to be the most suitable network after conducting the discussed experiments and comparisons of the figures that depict the (a) learning curves and (b) confusion matrix for each one of the best-performing models. For example, [Fig. 4](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/figure/fig-4/) represents VGG16 results, whereas [Figs. 5](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/figure/fig-5/)–[7](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/figure/fig-7/) represent the results for ResNet50, InceptionV3, and DenseNet121 respectively.

****

**VGG16 results where sub-figure (A) represents the learning curve of the model, while sub-figure (B) illustrates the confusion matrix on the testing data.**

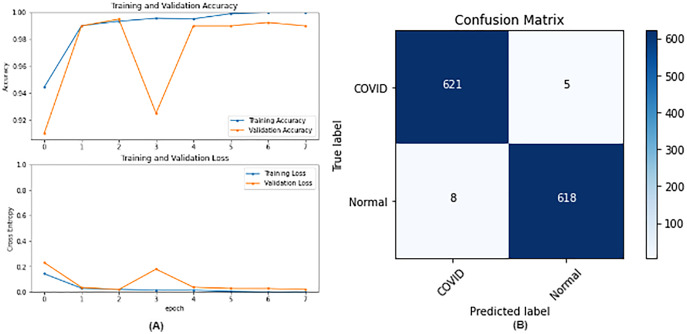


**ResNet50 results where sub-figure (A) represents the learning curve of the model, while sub-figure (B) illustrates the confusion matrix on the testing data.**

****

**DenseNet121 results where sub-figure (A) represents the learning curve of the model, while sub-figure (B) illustrates the confusion matrix on the testing data.**

**IncetionV3 results where sub-figure (A) represents the learning curve of the model, while sub-figure (B) illustrates the confusion matrix on the testing data.**



In our case, the false positives are represented by the top right corner of the confusion matrix figure where they refer to patients that do not have COVID but are positively classified. This could lead to a problem of shortage of medical supplies if they are hospitalized. As for the false negatives that are represented in the bottom left corner, they refer to patients with COVID that are classified as non-COVID. This could lead to a huge problem if they are not quarantined and are allowed to walk freely with crowds. Thus, the best-performing model is the one with the minimum number of FP and FN. In order to verify the validity of our findings, the results of our work is compared with several previous works. As shown in [Table 11](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/table/table-11/), the results of our models outperform all previously developed models included in the literature of this article.

### Table 11

**Comparison with previous work.**

| **Paper** | **Dataset** | **Models** | **Best result (Accuracy)** |
| --- | --- | --- | --- |
| [Chakraborty, Murali & Mitra (2022)](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/#ref-5) | COVID-19 and Pneumonia x-rays | ResNet18 | DenseNet (96.43%) |
| AlexNet |
| DenseNet |
| VGG16 |
| [Yadlapalli et al. (2022)](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/#ref-29) | COVID19-CT | VGG16 | VGG16 (89%) |
| ResNet50 |
| InceptionV3 |
| [Padma & Kumari (2020)](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/#ref-16) | COVID chest Xray | 2D CNN | Training ACC: 99% |
| Validation ACC: 98.3% |
| [Khan et al. (2022)](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/#ref-11) | BIMCV- COVID19+ | EfficientNetB1 | EfficientNetB1 (96.3%) |
| NasNetMobile |
| MobileNetV2 |
| [Aggarwal et al. (2022)](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/#ref-1) | Two datasets | MobileNetV2 |  |
| Xception |  |

**ADVANTAGES & DISADVANTAGES**

Further explanations and discussions are presented. Prior research has shown that DenseNet121 is a good choice for x-ray images as noticed in our work. Furthermore, they stated that InceptionV3 is a good choice but neglected the performance of ResNet50 even though it is used in their work. This might be due to not experimenting with enough hyperparameters to reveal the quality of its performance which has been shown in our work. VGG16 got the lowest results which are normal because of its shallow architecture when compared to the other architectures.

On the other hand, InceptionV3 has 48 layers close to the number of layers of ResNet50 which is 50. Thus, our experiments have shown that the residual block technique is more useful than the inception module technique which is introduced by InceptionV3.

Finally, DenseNet121 which has 121 layers has achieved close to ResNet50 which has far fewer layers. This could be due to the overparameterization problem which means that increasing the number of layers does not necessarily improve the results. Furthermore, training 50 layers take far less time compared with training all 121 layers of the DenseNet.

For the limitations, Google Collaboratory was used in this research which offers free usage of GPU, which was needed in this research to enhance the speed of execution, but with limited time per day and a limited number of sessions in the free model. Furthermore, some configurations could not be executed due to “ResourceExhaustedError” in [Fig. 8](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/figure/fig-8/). The runtime was long even with the GPU and because our research included running a lot of experiments it as difficult to do. Even though data augmentation has shown improvements with image classification but in our case, it adversely affected the results. This could be attributed to the fact that general image augmentation techniques may not be suitable for medical images. This is an interesting research question that is worth investigating in the future.

**CONCLUSION AND FUTURE SCOPE**

The purpose of this research was to train different networks on the lung X-rays dataset to classify whether the patient has COVID-19 or not, and conduct comparisons between them to help the medical field during the faced pandemic. Future work includes exploring segmentation ([Zaitoun & Aqel, 2015](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/" \l "ref-30)) techniques to train the model only on the interesting and important parts of the images, which could improve the results. Furthermore, using hybrid approaches through implementing recurrent neural networks (RNN) ([Sherstinsky, 2020](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/" \l "ref-21)) and transformers along with attention techniques ([Vaswani et al., 2017](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9575860/#ref-26)) should be taken into consideration as it is considered a rising topic these days. Different approaches used by the researchers are noticed, but mainly most of them based their work on transfer learning. This is understandable as building a network from scratch requires a lot of data and is time-consuming. Nevertheless, with enough data, building a model from scratch is applicable and might lead to better results as it is built and trained for a specific task. Future work may include building a network from scratch and conducting a comparison between the results of transfer learning networks and the network that is built from scratch. Finally, further experiments could also be done to find the best parameters for data augmentation and assess its role in medical images.