ISSN: 2666-2795 Vol. 7 No. 1, 2022, Netherland

International Journal of Applied Engineering Research

Exploring COVID-19 in Chest X-Ray Images Through Deep Learning Analysis

Waqas Khan[1*], Muhammad Bilal[2], Muhammad Junaid Iqbal[2], Muhammad Junaid Hassan[1], Muhammad Faheem Khalil Paracha[2]

waqasdurrani1993@gmail.com , mbilal9k@gmail.com, muhammadjunaid.uos@gmail.com, junaidhassan626@gmail.com, faheemparacha.uettazila@gmail.com

[1]Department of Physical and Numerical Science, Qurtuba University of Science and Information Technology, D.I.Khan, KPK

[2]Institute of Computing and Infromation Technology, Gomal University, D.I.Khan, KPK Pakistan 29050

[1*]Corresponding Author: waqasdurrani1993@gmail.com

AbstractL Timely diagnosis of COVID-19 is crucial for pandemic control, and this study explores the efficacy of Chest X-Ray images using deep learning techniques, particularly ResNet-50, which demonstrated superior accuracy at 98.33%. The investigation includes an array of models such as VGG-16, RestNet-34, InceptionV3, and VGG-19, emphasizing ResNet-50's robust performance. The proposed CoroDet-1 CNN model, tailored for automated COVID-19 detection, accommodates two, three, and four diagnostic classes. Evaluation involved 100 confirmed COVID-19 chest X-ray images and 1431 images from the Chest X-ray 14 dataset, showcasing the model's effectiveness and potential clinical application. This research underscores the pivotal role of advanced technology in achieving accurate and rapid COVID-19 screening.

Keyworks : Diagnostic,Detection,Deep Learning ,COVID-19, X-Ray, Convolutional Neural Network. VGG-16, VGG-19, ResNet-34, ResNet-50.

1. INTRODUCTION:

The Covid-19 pandemic, an unprecedented global health crisis, has defied efforts to find a vaccine, necessitating alternative treatments like plasma therapy and X-ray imaging [1][2]. Initiated in March 2020, research on diverse platforms such as Github and Kaggle aims to develop effective diagnostic tools and understand Covid-19's societal impact [3]. The virus, characterized by symptoms ranging from throat infections to respiratory issues, requires rapid and accurate diagnosis to prevent transmission and manage its economic and human toll [1]. This study focuses on innovative strategies and diagnostic approaches to combat the persistent spread of Covid-19, underlining the urgency for healthcare professionals to adapt and respond effectively to this ongoing challenge [4].

1.2 Background of the Study:

As of May 28, 2021, COVID-19 has affected over 5.5 million people worldwide, resulting in 353,373 deaths [1]. The virus presents a spectrum of symptoms, making efficient patient screening crucial [5]. Traditional genetic detection faces practical challenges, prompting the use of chest radiography, like X-rays or CT scans, to detect SARS-CoV-2 infection [7]. However, differentiating COVID-19 pneumonia from other diseases is difficult for radiologists, affecting computer-assisted diagnosis. To improve accuracy, advanced Convolutional Neural Networks (CNN) are used, but face limitations due to small datasets and simplistic categorizations [6]. These models are vital for quick evaluation and refined screening in COVID-19 cases, guiding treatment decisions.

1.3 Motivation:

The COVID-19 pandemic, with 170 million cases and 3.54 million deaths, has dramatically altered healthcare practices, underscoring the need for efficient testing [8]. Challenges in testing accessibility and distribution persist globally, affecting countries like Bangladesh and the United Kingdom [9][10][11]. To overcome issues like prolonged detection times and testing kit shortages, there's a shift towards using CT and x-ray scan images for rapid, accurate diagnosis [12]. This method, utilizing widely available technology, significantly reduces patient wait times. Concurrently, experiments are integrating radiological imaging with Convolutional Neural Networks (CNN) to enhance the diagnostic effectiveness of these imaging techniques in detecting COVID-19 [5].

1.4 Convolutional Neural Network:

Convolutional Neural Networks (CNNs) are essential in image processing, effectively retaining image properties while minimizing size. For instance, a 35x35x1 image is processed through CNNs to maintain quality with reduced dimensions [14]. Managing larger images, like 100x100 pixels, poses computational challenges, but CNNs use convolutional and pooling layers to efficiently reduce neuron count, thus lowering computational demands [15].

In clinical practice, CNNs are highly effective for medical imaging and disease detection, adept at identifying minute yet critical details, such as differentiating healthy and diseased lung tissues in radiography [18]. The proposed CNN architecture incorporates five convolutional layers, beginning with a 244 x 244 breast image tensor, and the first layer uses 64 kernel filters with a 1 x 1 stride [15].

Table 1.1: Several strengthened limits of the CNN model construction and favored weights are investigated in work.

M1	M2	M3	M4	M5	M6	M7	M8
CNN	224x224	RMS prop	32	10 fold	3e 4	BCE CCE	50

In the CNN architecture, a 244 x 244 breast image tensor undergoes processing via five convolution layers, as per hyperparameters M1 to M8. The first layer uses 64 filters with a 1 x 1 stride, followed by a 2 x 2 stride max-pooling, reducing the size to 112 x 112. Subsequent layers with 5x5x64 filters and 1x1 stride further diminish the size to 56 x 56 and then 28 x 28, incorporating ReLU activation and max-pooling. The fourth layer has 512 filters (5 x 5 x 256 kernel), and the fifth uses 512 filters (14 x 14 x 512 kernel), followed by max-pooling to a 7 x 7 x 512 tensor. This is flattened to 25,088 neurons, reduced via dropout layers and fully connected layers to 64 neurons, aiding in classifying the image as healthy, COVID-19, or pneumonia.

1.4.1 Convolution Layer:

In CNNs, filters applied to input images generate layers, capturing key features and characteristics of the original image.

F*t = (I¬n¬) (f) x (t), illustrating the process. (1.1)

The standard definition of the Adjoint F* of a linear map F ∈ F*t is a linear map that satisfies the following, assuming that t only takes integer values.

$$F(t) = \sum_a x \begin{matrix} l_n * a \\ f \\ (a-t) \end{matrix} \quad (1.2)$$

F(t) as a linear map switches F's domain and codomain, with its adjoint reversing direction, extending to two-dimensional convolutions using (m, n) inputs and f(a, b) kernels.

$$F^*t = \sum_{a\,x} \sum_b x\ l_n x\ (b,a)\ x\ f x(a-m, b-n) \quad (1.3)$$

To avert overfitting, error metrics and regularization of parameters are used. After learning, F* reconstructs new images, inverting the kernel per commutative rules.

$$Fx\ t = \sum_{a\,x} \sum_b x\ l_n\ x\ (m-a, n-b).f(a,b) \quad (1.4)$$

The reconstruction function is a regularized minimization problem, contrasting with the learning approach's parametric function for inverse problems. Neural networks use cross-correlation, similar to convolution, without flipping the kernel.

$$F * t = \sum_a x \sum_b x \, l_n (m + a, n + b).f(a, b) \quad (1.5)$$

Layer values depend on preceding layers, influenced by F (training data input) and t (weight). This layer prevents gradient vanishing and enhances analysis and training by zeroing negative inputs.

$$R(x) = \max 1\,(0, x) \quad (1.6)$$

This overview presents popular CNN models, their features, comparisons, and development history, while exploring their architectural theories. However, detailing their specific mathematical rules, particularly neuron input (x), remains challenging.

1.4.2 Pooling-Max Layer:

This layer employs a model-based discretization approach, assuming features in binned input sub-regions, thus reducing dimensionality while preserving substantial internal representation invariance. Carpooling uses a 3x3 kernel in max pooling.

$$(I * K)_{ij} = \sum_{m=0}^{k_1-1} \sum_{n=0}^{k_1-1} \sum_{c=1}^{C} K_{m,n,c} \, , I_{i+m,j+n,c} + b \quad (1.7)$$

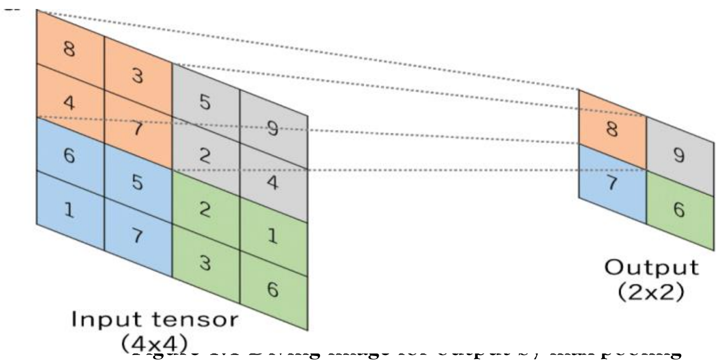
Convolution filters process images by computing dot products, with each filter extracting distinct features. Element-wise multiplication and summation occur between a 2x2 filter and a 5x5 image.

Max Pooling code for separating images in Python

```
import numpy using the np

import Sequential from keras.models

import MaxPooling2D from keras.layers

# Define input image

image = np.array

([[2, 2, 7, 3],

[9, 1, 4, 5],

[8, 5, 2, 6],

[3, 1, 2, 6]])

photograph = photograph.reshape(1, 4, 4, 1)

# Describe a model with a single max pooling layer.

model = Sequential([MaxPooling2D(pool_size=2, strides=2)])

# Produce combined results.

model.predict(image) as output

# Print the resultant picture.
```

np.squeeze(output) = output

take printout (output)



1.4.3 Loss Function:

In the class score vector, 'y' indicates the CNN score for the jth positive class, with 'c' being the total classes per image. SoftMax logarithmically translates this into predictive probabilities.

$$L_i = -\log\left(\frac{e^{\beta y}}{\sum_j^c e^{\beta j}}\right) \tag{1.8}$$

Hyperparameters are tuned to minimize J (average loss) by optimizing weights (Li) and biases (eby), similar to using residuals in statistics for minimizing deviation in regression analysis.

1.4.4 Accuracy of Regularization:

Dropout, as a simple yet effective regularization technique, is chosen for its ease of implementation, with the hyperparameter set at 0.50 for optimal regularization in this study.

1.5 Research Criteria:

i. Has the integration of Deep Learning (DL) significantly enhanced the current standard of care for COVID-19 diagnosis?
ii. What alternative modalities can be employed in conjunction with DL to support COVID-19 identification and diagnosis?
iii. How successful has DL been in addressing the limitations of existing diagnostic modalities for COVID-19?
iv. The comparative effectiveness of various DL forms and topologies in facilitating the diagnosis of COVID-19.

Researchers conducted a comprehensive search in medical and computer science databases, focusing on PubMed, Web of Science, and Scopus, among others, to identify the most pertinent publications for their study. The search, conducted from November 1, 2019, to July 20, 2020, utilized keywords such as "COVID-19," "diagnostic," "detection," and "deep learning." EMBASE and IEEE databases were excluded due to redundant content.

1.6 Data Extraction

Data extraction forms capture research details, including DL algorithms. Two researchers collaborate to ensure accuracy and resolve discrepancies through discussion.

1.7 Use of DL Techniques for Recognition & Verdict of Corona Virus

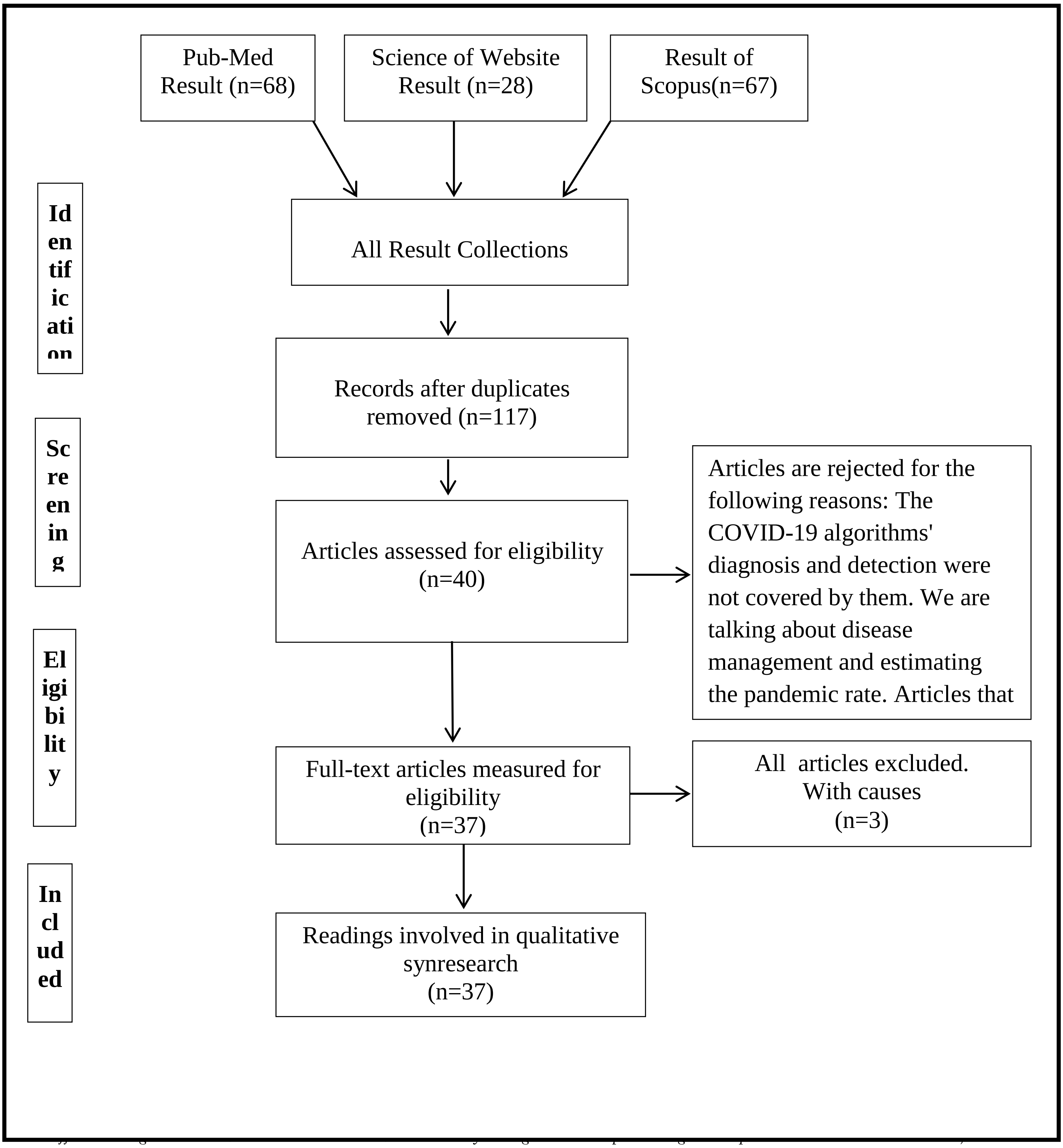
DL techniques effectively detect, diagnose, categorize, predict, and assess COVID-19. Studies employ various datasets, evaluating algorithms for sensitivity, specificity, and accuracy, with some pioneering novel methodologies.

1.8 Short coming:

Past studies encountered challenges like complex structures, high costs, and low accuracy. Outdated models contributed to accuracy limitations, emphasizing the need for technological upgrades to improve accuracy.

1.9 Statement of the Problem:

COVID-19, originating in Wuhan, China in December 2019, has caused widespread pneumonia with varying effects. It has strained healthcare, economies, education, and commerce worldwide. Timely detection is vital, but testing kit shortages, particularly in developing countries, are challenging. Recent research employs radiology, specifically X-ray images, for COVID-19 diagnosis. Advanced techniques like Convolutional Neural Network (CNN) applied to lung X-rays enable swift and accurate disease identification, offering a valuable solution amidst testing kit shortages.



RestNet-34, InceptionV3, and VGG-19 are assessed, with ResNet-50 demonstrating superior performance. The potential for automated COVID-19 diagnosis from chest X-rays is highlighted, offering a swift and efficient screening tool for healthcare professionals.

1.10 Aims and Objectives:

1. Examine chest X-ray images from the poster anterior (PA) perspective in both healthy persons and Covid-19 patients.

2. Evaluate the performance of DL-based CNN models after data cleaning and application of data augmentation.
3. Examine and contrast the relative accuracy of the ResNeXt, Xception, and Inception V3 models.

1.11 Researcher Questions:

1. Analyze how much the use of deep learning has enhanced conventional methods for COVID-19 diagnosis.
2. Examine the ways in which COVID-19 detection and diagnosis can be aided by deep learning.
3. Assess whether deep learning has been successful in compensating for the flaws in diagnostic modalities.
4. Compare the efficiency of different forms of deep learning and their architectures in boosting COVID-19 diagnosis relative to each other.

1.12 The Pros & Addition

1. Integration of machine learning with DL improves COVID-19 diagnostic accuracy.
2. The research enhances DL by applying clustering to the dataset for COVID-19 recognition.
3. Examining dataset examples aids in identifying distinctions and supporting learning.
4. Radiological imaging showcases chest X-ray efficacy for COVID-19, pneumonia, and negative cases.
5. Limitation: The approach relies solely on chest X-ray data, while COVID-19 diagnosis could involve various medical datasets.

2. METHODOLOGY:

In my research, three crucial experiments are conducted, each assessing the effectiveness of the models and examining the impact of various process steps. Each experiment involves a unique dataset, utilizing identical images for all COVID-19 positive cases. For negative scenarios, three different datasets are employed. Experiments 1 and 2 sequentially compare positive and negative case datasets, while Experiment 3 incorporates images captured with the COVID-19 imaging system (images from 2019-2020).

2.1. Datasets

The proposed method validates using Chest X-ray images from two sources: COVID-19 X-ray images from Cohen JP's dataset [16] and normal chest X-ray images from GitHub repository 1 [17]. Data split is 71% training and 29% testing.

2.1.1. Filtering for Image Projection

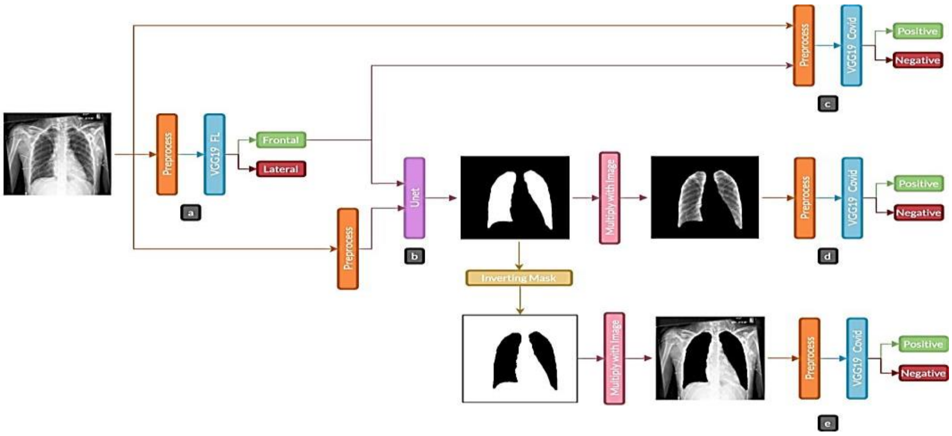
The technique prioritizes frontal COVID-19 datasets, emphasizing poster anterior and anteroposterior projections, to mitigate mismatches and enhance model performance while preserving valuable information.

2.1.2 Lung Segmentation

Despite the apparent data scarcity, the ample volume and variety of photos provided sufficient information to develop an effective segmentation model for classification purposes.

2.2 Image Separation

The classification task is divided into three sections to prevent biases and overfitting. Frontal images have 1,140, 723, and 609 pixels in the training, testing, and validation sets, while lateral images have 375, 236, and 204 pictures in each set. Positive COVID-19 cases comprise 6,475 training images, 3,454 testing images, and 2,873 validation images. Negative BIMCVCOVID-19 data has 2,342 training, 1,228 testing, and 1,040 validation images. The pre-COVID-19 dataset includes 2,803 training images, 1,401 testing images, and 1,265 validation images. The COVID-19 cases dataset has 286 training, 96 testing, and 96 validation images for comparison, denoted as tasks 'a' to 'e' in the experiment diagram.



Figurer 2.1: Converting x-ray to positive and negative result

Figure 2.1 illustrates the transformation of raw x-ray data into positive and negative results, providing information images that determine the final outcomes.

2.3 Preprocessing

Preprocessing addresses data variability from diverse image sources like the Montgomery, National Institutes of Health segmentation collections (from Pakistan and worldwide), and JSRT collection (from the Japanese government), which have different X-ray equipment, technologies, and resolutions. The steps include:

1. Resize all images to 224x224 pixels in one channel.

2. Resize each image to 224x224 pixels in a separate channel.

3. Resize every grayscale image to 224x224 pixels in a third channel.

Normalization is represented as Eq. (2.1), where x is the original image, and N is the normalized image. Standardization uses Eq. (2.2), with Z for standardized view and N for normalized image in each standardization case.

$$N_i = \frac{x_i - \min x . x}{\max(x) - \min . x} \tag{2.1}$$

$$Z_i = \frac{N_i - \text{mean} . N}{\text{std} . N} \tag{2.2}$$

U-Net architecture segments chest X-ray images, producing masks with 1s in the reconstruction region and 0s elsewhere. Filter numbers are optimized using three layer numbers on CNN, calculated by Eqs. (2.3) and (2.4) [References not included].

$$\#Filters_{cont} = F_0 * 2^{i-1} \tag{2.3}$$

$$\#\text{Filters}_{expan} = \frac{F_f}{2_i} \tag{2.4}$$

In production, the model takes an X-ray chest image as input and produces a predicted mask of the chest image as output.

2.4 Hyper Parameters

In the convolutional layers, standard kernel initialization and padding use 3 x 3 kernel sizes. The first two blocks of Expansion and Contraction have a pool size of 2 x 2, and dropout rates vary from 0.1 to 0.3 in subsequent blocks. No dropout is applied in the first two blocks. Transposed convolutional layers have a kernel size of 2, strides of 2, and padding similar to the original convolutional layers. The final convolutional layer uses a single filter with a 1 x 1 kernel size.

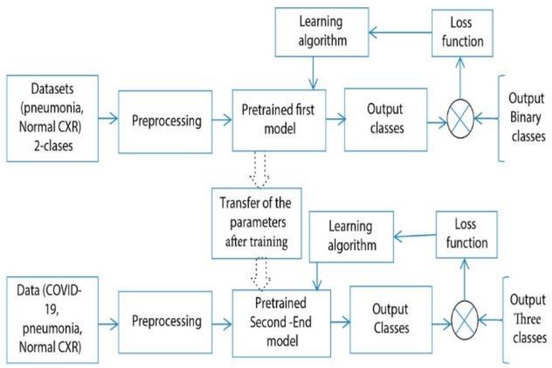
2.5 The Method

The method is thoroughly discussed below, and each step is explained in detail.

2.5.1 System Architecture.

1. Goal: Develop a system to accurately detect viral pneumonia and COVID-19 in X-ray images.
2. Dataset: Training, validation, and testing sets consisting of pneumonia and normal X-ray images.
3. Preprocessing: Apply scaling, data augmentation, and resampling techniques to the images.
4. Model: Construct a CNN structure to calculate the model's output using chest X-ray images.
5. Training: Adjust the CNN parameters based on the loss function by comparing results with target classes.
6. Transfer Learning: Utilize a pre-trained model to train a second model using transfer learning methods.
7. Iteration: Repeat steps 3-6 for each dataset and epoch in the analysis.
8. Dataset Expansion: Create a research dataset with COVID-19, pneumonia, and normal X-ray images.
9. Preprocessing: Apply data augmentation, resampling, and scaling to the research dataset.
10. Model Update: Update the parameters of the pre-trained model using the training method.

F



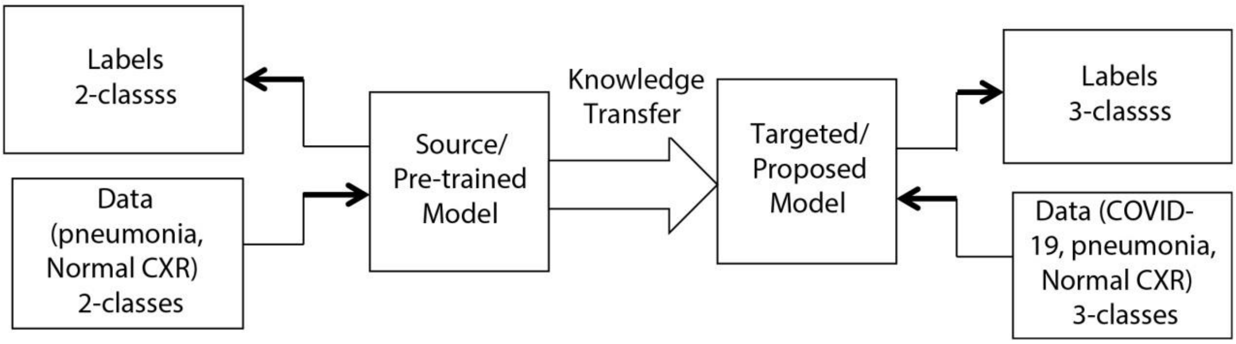


Figure 2.8: Transfer learning structure.

Figures 2.7 and 2.8 demonstrate transfer learning's use in building the model. Two datasets are used: one with normal and pneumonia patients, and another with COVID-19, pneumonia, and healthy individuals. The first model is trained on the first dataset, and the second model incorporates knowledge from the first model to handle all three conditions. This shows how transfer learning impacts the final model architecture.

.

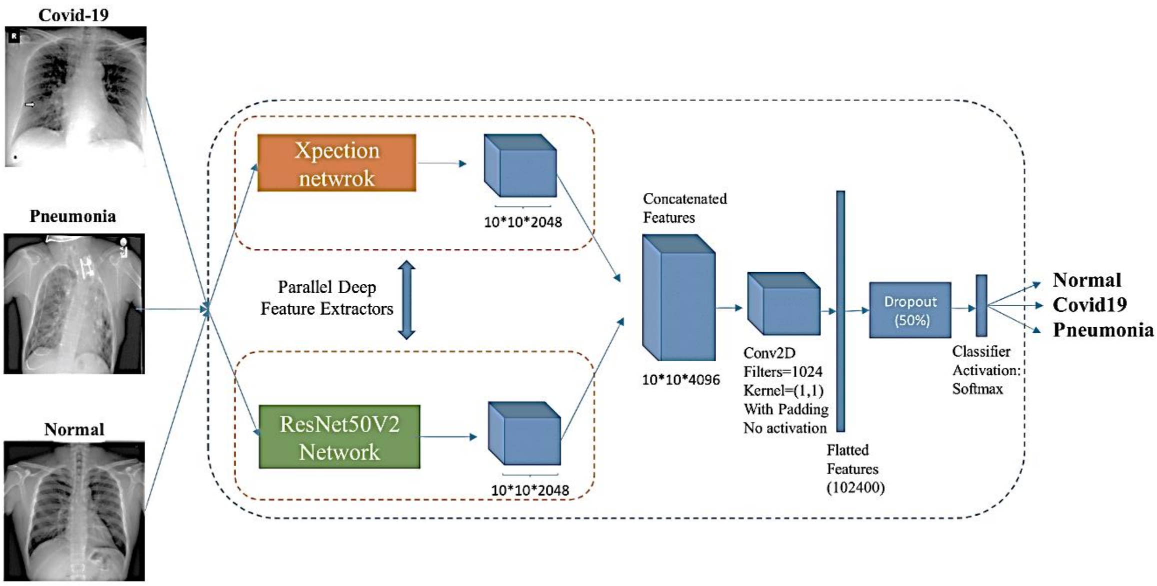


Figure 2.9: Proposed model

2.6 Proposed Model

Accuracy(ACC) = $\frac{T \times P + T \times}{P + N}$ x 100 (2.5)

Specificity = $\frac{T \times N}{N}$ x 100 (2.6)

Precision = $\frac{T \times P}{TP + FP}$ x 100 (2.7)

Recall = $\frac{TP}{P}$ x 100 (2.8)

F_1 – Measure = $\frac{2 \times precision \times recall}{precision + recall}$ x 100 (2.9)

Area under curve (AUC) = $\frac{1}{2}\left(\frac{TP}{P} + \frac{TP}{N}\right)$ (2.10)

Mathews Correlation Coefficient (MCC) = $\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP) \times P \times N \times (TN+FN)}}$ (2.11)

The *z*-test and Friedman average ranking are two statistical techniques that are used to thoroughly validate the acquired data. Furthermore, to guarantee the validity of the results, post-hoc multiple comparison techniques created by Shaffer (Shaffer, 1986) and Holm (Holm, 1979) are used.

2.7 Simulation Discussion

MATLAB in 2021 facilitated the task with its user-friendly interface for technical computing. It integrates computation, visualization, and programming, making it popular in prototyping, simulation, modeling, and data analysis. MATLAB's array-centric approach, extensive toolboxes, and contributed user functions have made it a cornerstone in scientific and engineering fields.

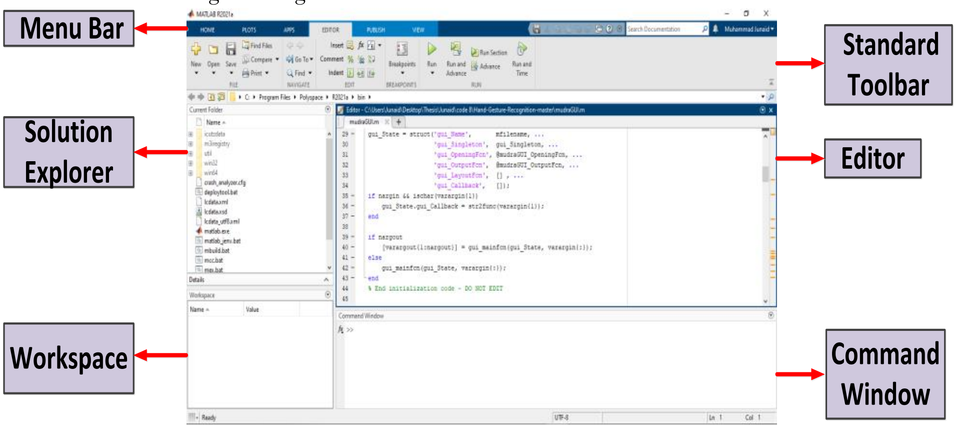


Figure 2.10: MATLAB main window preview

3. Results and Discussion

This article explores using X-ray images and Deep Learning (DL) to detect Covid-19. It covers the suggested model structure, theoretical assessment, literature review, simulation model, and machine learning analysis of the results.

3.1. Dataset For Testing

The dataset used in this study was obtained from the Kaggle repository and consisted of average chest X-ray images of Covid-19 patients. The dataset, containing a total of 6432 images, was divided into three categories: Covid set, training (5467 images), and validation (965 images). During validation, the study included 641 pneumonia patients, 86 Covid cases, and 238 normal cases. For effective model training, 576 posterior-anterior (PA) scans of Covid-19 patients were collected and resized to 128x128. Table 1 provides details on the data distribution used for testing and training.

3.2 Training and Data Formulation

The research involves multiple phases: data acquisition from GitHub and Kaggle repositories, splitting the dataset into training and validation sets, pre-processing, and training. Due to the small dataset size, online augmentation was performed for each epoch. Transfer learning using the ResNet50 model was employed and fine-tuned. The ResNet50 model demonstrated excellent performance with its number of layers. Figure 3.1 depicts the validation accuracy of the trained model, achieving 98.7% accuracy. Positive labels were assigned to infected chest X-ray images, while negative labels were assigned to non-infected images.

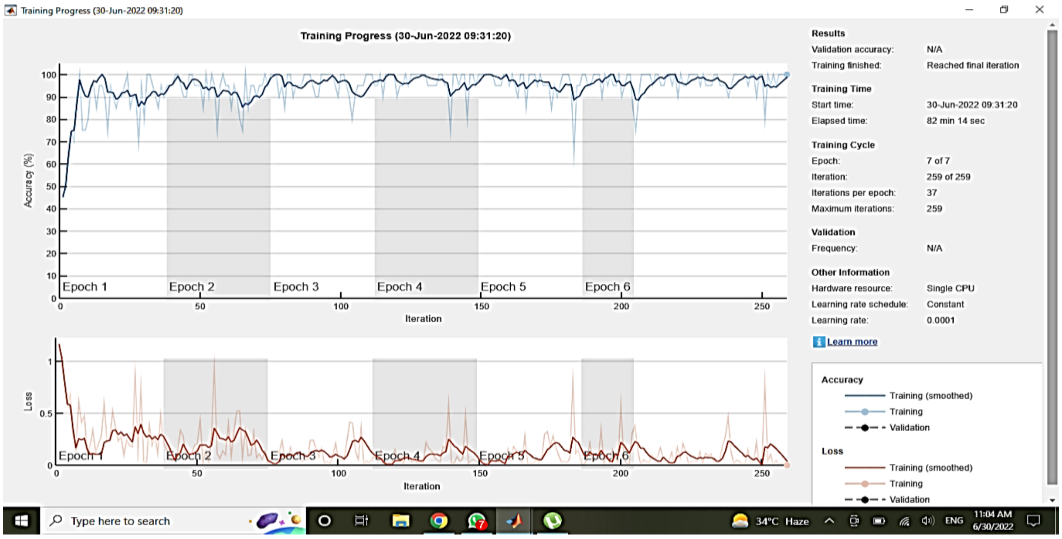


Figure 3.1: Validation accuracy of trained network model

A MATLAB GUI was developed for individual X-ray detection, featuring separate interfaces for COVID-19 positive and negative cases (Figure 3.2 and 3.3, respectively).

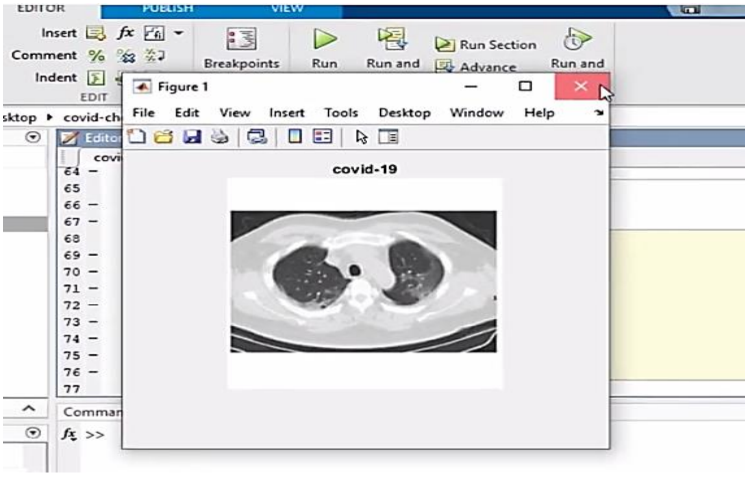


Figure 3.2. Individual result for covid-19 positive case

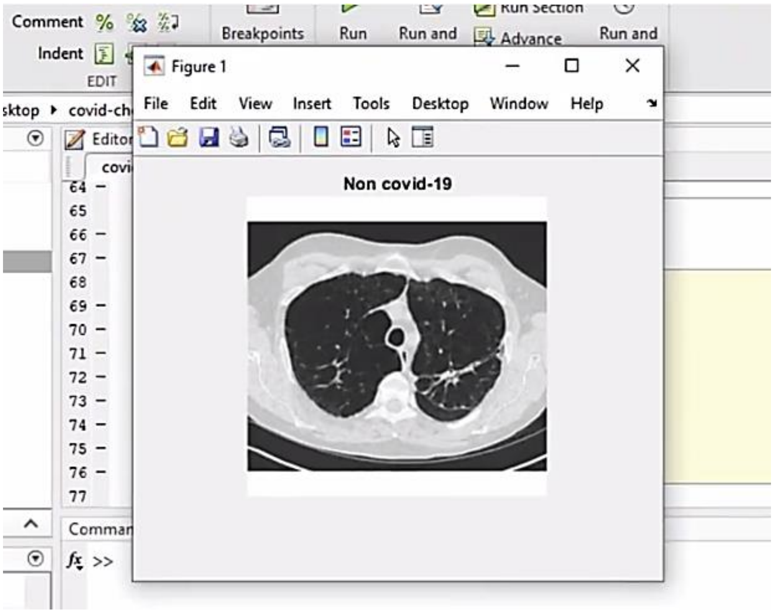


Figure 3.3. Individual result for covid-19 negative case

3.3 ResNet50 Spooling X-Ray Images

Figure 3.4 illustrates the process of processing chest X-rays, segregating infected and non-infected images through the implementation of the ResNet50 algorithm.

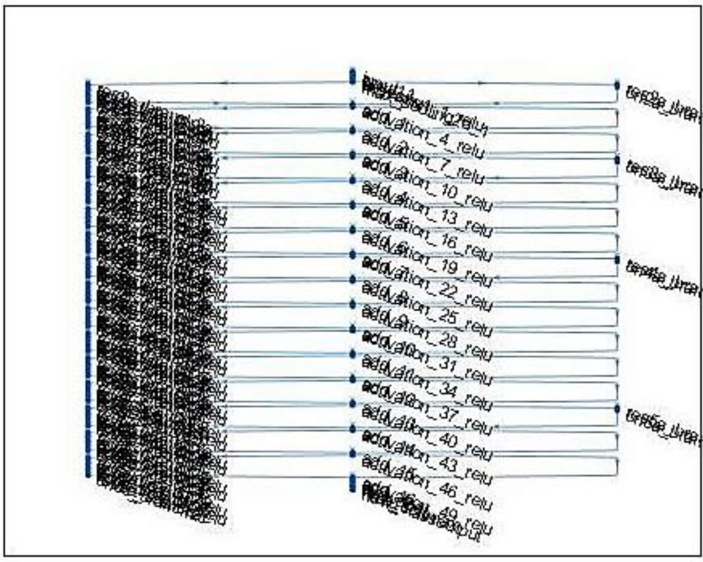


Figure 3.4 ResNet50 spooling x-ray images

3.3.1 Spooling image of ResNet50

Figure 3.5 demonstrates the application of the ResNet50 algorithm to classify chest X-rays,distinguishing between infected and non-infected images.

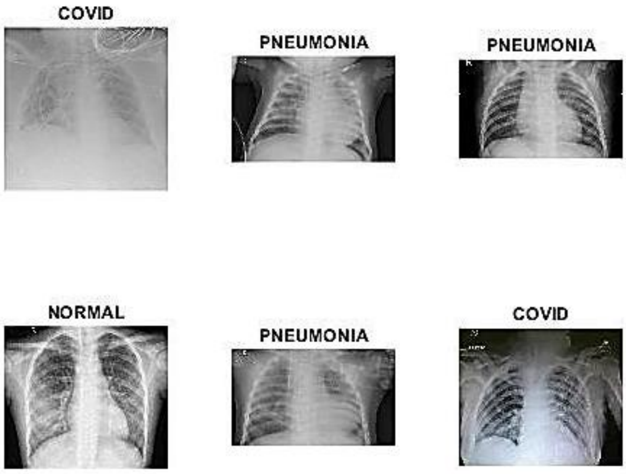


Figure 3.5 Spooling image of ResNet50

3.3.2 Final Training Progress Epoch 5 out of 5

Figure 3.13 displays the final outcome of the MATLAB code run, revealing a reduced loss rate and an improved accuracy rate compared to previous results, meeting the paper objectives.

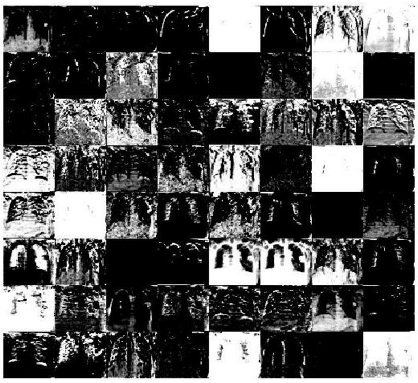


Figure 3.6 Final training progress epoch 5 out of 5

3.3.3 Final Result ResNet50

Figure 3.7 presents the final outcome of chest X-ray images, depicting infected areas as the darkest regions and non-infected areas as the lightest regions.

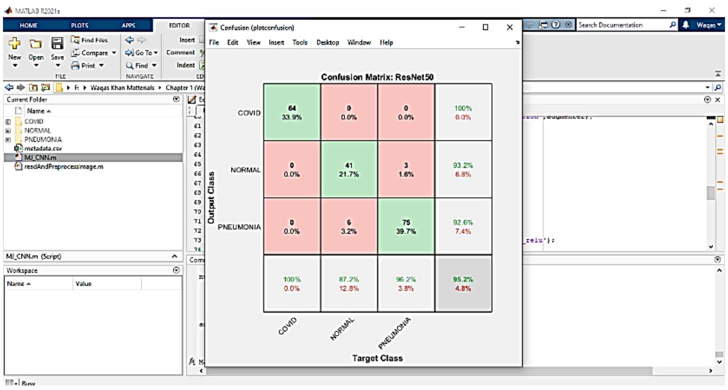


Figure 3.7 Final result ResNet50

Table 3.1 Classification results Comparison of all eight CNN models.

Model	Pre (%)	Sen (%)	Spe (%)	F1-score	Acc (%)	AUC
ResNet-50	96.77	100.00	96.67	0.9836	98.33	0.9836
ResNet-34	95.24	100.00	95.00	0.9756	97.50	0.9731
VGG-16	95.08	96.67	95.00	0.9587	95.83	0.9487
VGG-19	98.24	93.33	98.33	0.9573	95.83	0.9506
Inception-V3	96.36	88.33	96.67	0.9217	92.50	0.9342

Table 3.16 presents the categorization and comparison report with accuracy rates for different models: Inception V3 - 92.50%, VGG-19 - 95.83%, VGG-16 - 95.83%, ResNet-34 - 97.60%, and ResNet-50 (top performance) - 98.33%.

3.4 Confirmed, Death and Recovered Cases

Fig 3.5 showing the simple scatter of recovered mean of confirmed death cases and recovered cases.

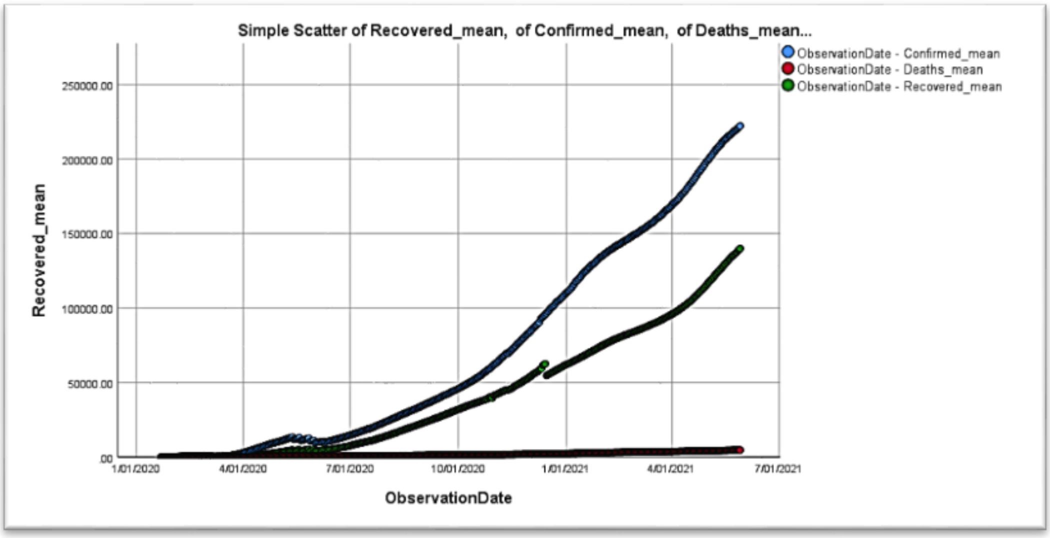


Figure 3.8 confirmed, death and recovered cases

Table 3.2 Sequence Plot

Model Description		
Name of Model		MOD_1
Sequence or Series	1	Confirmed_mean
	2	Deaths_mean
	3	Recovered_mean

Changes	None
Non-Seasonal Variabilities	0
Seasonal Variabilities	0
Duration of the Seasonal Cycle	No periodicity
Labels for Horizontal Axis	ObservationDate
Intervention Commences	None
For Every Observation	Values not joined

Using the MOD_1 model's specs

Table 3.3. Case Processing Summary

Case Processing Summary				
		Confirmed_mean	Deaths_mean	Recovered_mean
Length of Series or Sequence		494	494	494
Plot's total number of missing values	User Error	0	0	0
	Missing System	0	0	0

Cases processing summary

3.5 Observation Date and Chart

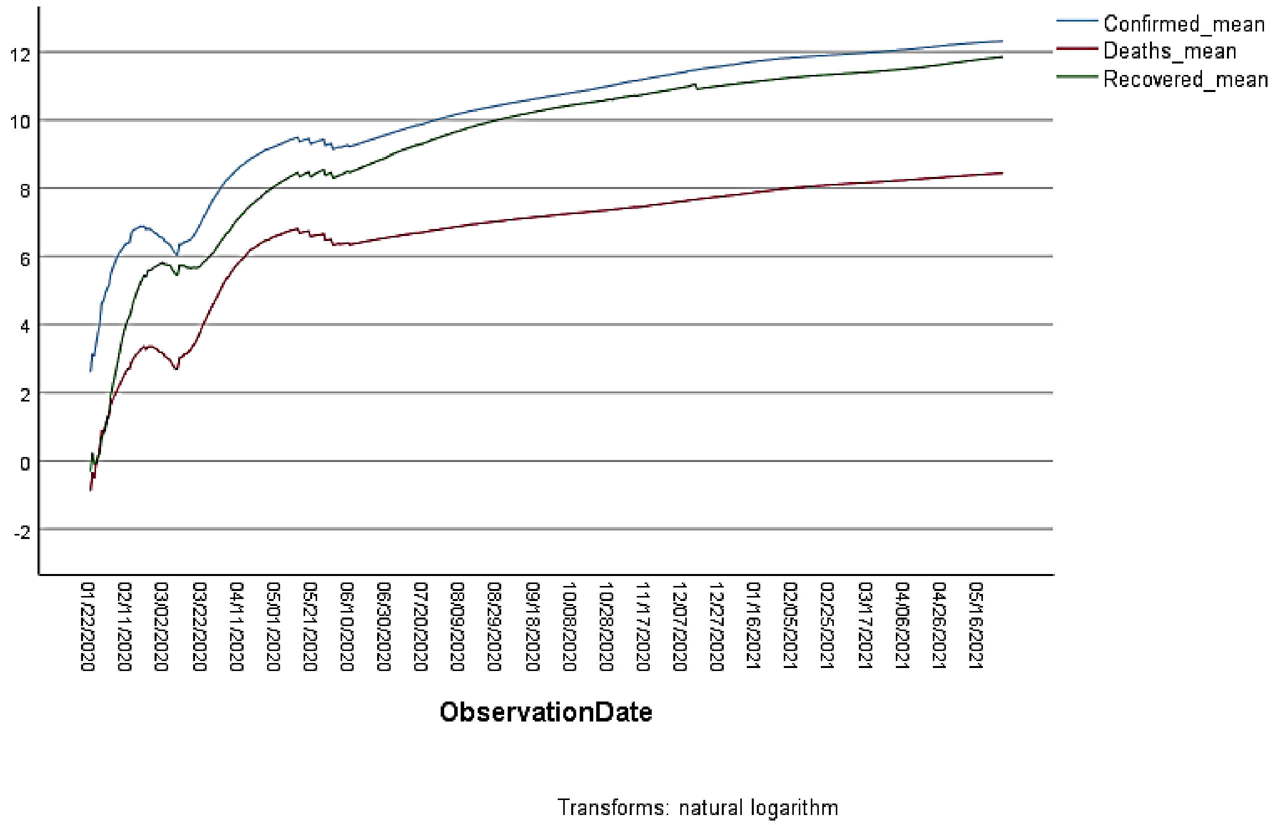


Figure 3.9 Confirmed cases on date wise

Table 3.4 Model Summary

normal_count					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.22	1	.1	6.3	6.3
	2.57	1	.1	6.3	12.5
	2.85	1	.1	6.3	18.8
	2.88	1	.1	6.3	25.0
	3.13	1	.1	6.3	31.3
	3.15	1	.1	6.3	37.5
	3.98	1	.1	6.3	43.8
	3.99	1	.1	6.3	50.0
	5.50	1	.1	6.3	56.3
	6.37	1	.1	6.3	62.5
	6.40	1	.1	6.3	68.8
	6.84	1	.1	6.3	75.0
	6.91	1	.1	6.3	81.3
	7.00	1	.1	6.3	87.5
	7.40	1	.1	6.3	93.8
	11.20	1	.1	6.3	100.0
	Total	16	1.7	100.0	
Missing	System	934	98.3		
Total		950	100.0		

Model summary for model summary

3.6 Overall Result

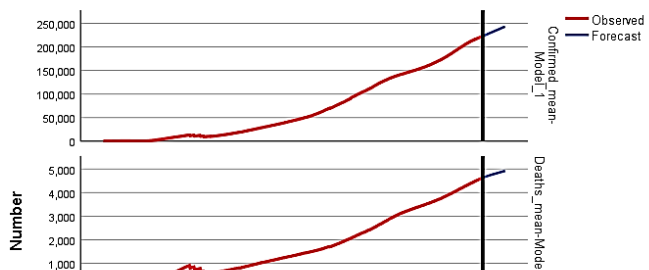


Figure 3.10 Confirmed, death and recovered cases

Table 3.5 Correlation

Correlations

		Confirmed_mean	Deaths_mean	Recovered_mean	GDPpercapita
Confirmed_mean	Pearson Correlation	1	.889**	.986**	.234**
	Sig. (2-tailed)		.000	.000	.004
	N	150	150	150	150
Deaths_mean	Pearson Correlation	.889**	1	.861**	.187*
	Sig. (2-tailed)	.000		.000	.022
	N	150	150	150	150
Recovered_mean	Pearson Correlation	.986**	.861**	1	.198*
	Sig. (2-tailed)	.000	.000		.015
	N	150	150	150	150
GDPpercapita	Pearson Correlation	.234**	.187*	.198*	1
	Sig. (2-tailed)	.004	.022	.015	
	N	150	150	150	150

**At the 2-tailed 0.01 significance level, the correlation is significant.

* The correlation is significant at the two-tailed 0.05 level.

Overall correlation report

3.7 Summary

The dataset, sourced from Kaggle, consisted of a total of 5467 datasets for training and 965 for validation, encompassing both regular and COVID-19 cases. A Deep Learning model was created and tested on individual and combined datasets. The training process utilized 70% of the data, taking approximately 2 hours. The model achieved a remarkable score of 98.7%, with potential for improvement through code updates to reach 98%.

4. Conclusion & Discussion

4.1. Discussion

Numerous studies have emerged since Cohen's dataset release [16], aiming to predict COVID-19 infections with chest X-ray images. Despite the prevalence of CNN-based networks incorporating the concept of time lag (TL) over the years, earlier strategies suffered from limited data points and inaccuracies. This study assesses the performance of AlexNet, VGG-16, MobileNet-V2, SqueezeNet, ResNet-34, and Inception-V3 on X-ray images for COVID-19 prediction. Rigorous testing on a balanced dataset from two sources ([16] and [17]) reveals ResNet models as superior. Notably, ResNet-34 achieves 98.33% accuracy and 100.00% sensitivity, aiding radiologists without additional costs. The study's limitation lies in its evaluation on a small number of COVID-19 samples due to the ongoing pandemic's early stages. Future endeavors involve validating the technique on larger datasets as they become available.

4.2. Conclusion:

This study introduces a DL-based automated method for effectively distinguishing COVID-19-infected patients from healthy individuals through the analysis of chest X-ray images. Various pre-trained CNN architectures utilizing TL are evaluated on publicly available X-ray data, with ResNet-34 outperforming others with a remarkable 98.33% accuracy. This model has potential applications for radiologists, serving as a valuable tool for cross-verifying screening results and saving time. The study opens avenues for the development of more effective CNN models, particularly for multi-class classification challenges, and future research aims to explore optimization techniques to enhance model reliability.

References:

[1] S. R. Nayak, "Application of deep learning techniques for detection of COVID-19 cases," *Biomed. Signal Process. Control*, vol. [add volume if available], pp. [add page range], Feb. 2021.

[2] R. H. Abiyev, "COVID-19 and pneumonia diagnosis in X-ray images using deep learning," *Math. Problems Eng.*, vol. 2021, pp. [1–14], 2021.

[3] C. Huang et al., "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *Lancet*, vol. 395, pp. 497–506, 2020.

[4] K. Elasnaoui and Y. Chawki, "Using X-ray images and deep learning for automated detection of COVID-19," *J. Biomol. Struct. Dyn.*, pp. 1–22, 2020.

[5] F. Shi et al., "Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19," *IEEE Rev. Biomed. Eng.*, vol. [add volume], pp. [add pages], 2020.

[6] Y. Oh, S. Park, and J. C. Ye, "Deep learning COVID-19 features on CXR using limited training data sets," *IEEE Trans. Med. Imaging*, vol. 39, no. [issue], pp. [add pages], 2020.

[7] M. Chung et al., "CT imaging features of 2019 novel coronavirus (2019-nCoV)," *Radiology*, vol. 295, no. 1, pp. [add pages], 2020.

[8] H.-C. Shin et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics, and transfer learning," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.

[9] G. Litjens et al., "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.

[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[11] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[12] X. Gu et al., "Classification of bacterial and viral childhood pneumonia using deep learning in chest radiography," *Proc. 3rd Int. Conf. Multimedia Image Process.*, pp. 88–93, 2018.

[13] V. Chouhan et al., "A novel transfer learning-based approach for pneumonia detection in chest X-ray images," *Appl. Sci.*, vol. 10, no. 2, p. 559, 2020.

[14] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A framework of deep learning classifiers to diagnose COVID-19 in X-ray images," *[Journal/Preprint]*, 2020.

[15] L. Wang and A. Wong, "COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *arXiv:2003.09871*, 2020.

[16] L. Li et al., "Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT," *Radiology*, vol. 296, no. 2, pp. E65–E71, 2020.

[17] E. Hussain et al., "CoroDet: A deep learning-based classification for COVID-19 detection using chest X-ray images," *Chaos, Solitons Fractals*, vol. 142, 2021.

[18] A. Wong et al., "COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Sci. Rep.*, vol. 10, pp. 1–12, 2020.

[19] R. Jain, M. Gupta, S. Taneja, and D. J. Hemanth, "Deep learning-based detection and analysis of COVID-19 on chest X-ray images," *[Journal/Conference]*, 2021.

[20] S. Tabik et al., "COVIDGR dataset and COVID-SDNet methodology for predicting COVID-19 based on chest X-ray images," *[Journal/Conference]*, 2020.

[21] C. Sitaula and M. B. Hossain, "Attention-based VGG-16 model for COVID-19 chest X-ray image classification," *[Journal/Conference]*, 2021.