

TOWARDS ACCURATE COVID-19 VARIANT DETECTION: A DEEP LEARNING APPROACH ON CHEST CT DATA

Muhammad Hammad u Salam^{*1}, Shujaat Ali Rathore², Nazir Ahmad³, Muhammad Irfan³

^{*1,2}Department of Computer Science & Information Technology, University of Kotli, Azad Jammu and Kashmir.

³Department of Computer Science, NCBA&E, Sub-Campus Multan, 60000, Pakistan

^{*}hammad.salam@uokajk.edu.pk

DOI: <https://doi.org/10.5281/zenodo.17231985>

Keywords

Image Analysis, Deep Learning, CT Chest, COVID 19

Article History

Received: 24 April 2025

Accepted: 30 June 2025

Published: 24 July 2025

Copyright @Author

Corresponding Author: *
Muhammad Hammad u
Salam

Abstract

The global COVID-19 pandemic, which started in late 2019, and its subsequent waves have put an unparalleled burden on healthcare systems worldwide, necessitating the rapid development of diagnostic tools and treatments. The prompt identification and accurate classification of COVID-19 patients is critical to stopping the virus's spread and improving patient outcomes. This paper takes up the challenge and embarks on a groundbreaking journey to provide an innovative and practical answer utilizing deep learning, a subset of artificial intelligence. The proposed study suggests a classification method to classify the COVID-19 and other cases from the chest image dataset. There are three steps of proposed methodology, in first step a dataset of CT scan chest images is downloaded from the Kaggle website. The second step of our proposed model is feature extraction and image classification using different deep learning models MobileNetV2, ResNetV2 and proposed a Custom CNN model for better performance. In third step, proposed methodology is detecting Covid-19, Pneumonia and Normal cases from the given image's dataset. Finally, a comparative analysis is done between the proposed and other well-known techniques on the basis of training and testing accuracy as well as loss. The comparative results show that our proposed model performed well with testing accuracy of 96% percent and give minimum loss.

INTRODUCTION

This section provides a thorough overview of the complex field of COVID-19, covering its history, effects on world health, and the importance of timely detection. It explores the nuances of different detection strategies used in the battle against COVID-19, including conventional diagnostic procedures like RT-PCR and the growing significance of medical imaging in the diagnosis process. This section also clarifies the study's main goal, which is to develop and validate a deep learning-based methodology for COVID-19 case classification [1-2] using chest image analysis. It also

makes clear the critical issue statement and the difficulties with current diagnostic techniques. This chapter establishes the framework for the following sections, which will elaborate on the approaches, findings, discussions, and conclusions of this research project, by offering a solid theoretical foundation.

The COVID-19 pandemic that began in late 2019 and the waves that followed have placed an unprecedented strain on healthcare systems around the world, requiring the quick creation of diagnostic instruments and remedies[3]. One of the most

important factors in halting the virus's propagation and enhancing patient outcomes is the timely identification and precise classification of COVID-19 patients. This paper takes up the challenge and, using deep learning[4], a subset of artificial intelligence, sets out on a revolutionary journey to deliver a creative and effective solution. Chest radiography—that is, chest X-rays[2]—has shown to be a very useful tool in the diagnosis of respiratory conditions, particularly COVID-19. Our goal is to add to the array of diagnostic instruments that medical practitioners have at their disposal by creating a deep learning-based method for classifying COVID-19 cases from chest pictures. This chapter presents the overall objectives of the study, describes the thesis structure, and explains why deep learning is being used to meet the urgent demand for a precise and quick COVID-19 diagnosis. It also establishes the framework for the next section, which explore the study's methodology, findings,

comments, and conclusions. To aid in diagnosis, the covid-19 disease—an infectious disease brought on by significant respiratory syndrome—is divided into two forms: covid-19 and Normal. Dry cough, fever, dysphagia, and dysphagia are among the symptoms of COVID-19, whereas normal refers to symptoms that are not associated with the virus. Covid-19 is more dangerous because it spread quickly and resulted in fatalities. The first case report of COVID-19 was published in China in 2019, and at that point, the virus became global. Many researchers are working to manage this disease and offer the public with the best guidelines for keeping oneself safe. These guidelines include wearing masks, keeping a 6-foot social distance, washing your hands for 20 seconds, and sanitizing your hands after performing any job [5]. Machine learning has great applications in the classification of medical images.

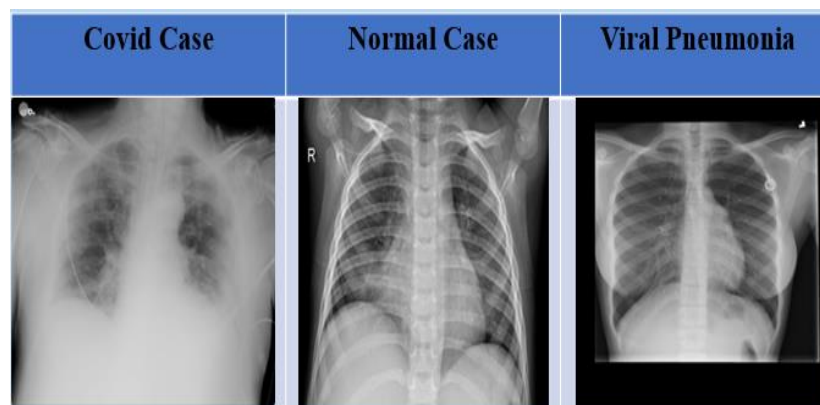


Figure 1FIGURE 1.

Figure 2 Classes of Images

Few researchers study this illness and offer the medical community improved methods for accurately diagnosing[6] it. Machine learning techniques are enhanced by deep learning [7] characteristics as well. Since deep learning relies on machine intelligence, it is related to the more advanced and specialized field of machine learning. The effectiveness of medical image classification models is more sufficiently impacted by deep learning techniques. Several techniques, including RT-PCR, X-rays, ultrasounds, and CT scans, are being used to identify COVID-19 disease. Reverse

transcription polymerase chain reaction, or RT_PCR, is employed in PCR reports to categorize the covid-19 results. Additionally, instances of false negatives in PCR samples have been made due to insufficient cellular content, a laborious process, or insufficient detection in cases where positive radiological collection was present. Chest X-rays (CXRs) are more accurate in diagnosing COVID-19 infections due to the complex anatomical lung pattern, which might change over time in terms of size and visual aspect. This diagnosis is mostly dependent on radiological competence [8-9]. Higher accuracy in identifying sample patterns may expedite

the screening procedure and, once suspected or confirmed, the diagnosis and treatment of COVID-19-positive individuals[10]. In recent years, lung ultrasonography (LUS) has advanced to both its theoretical and practical aspects. The ability of LUS to detect variations in the tissue-to-air ratio in the apparent lung surface is one of its unique qualities. Covid-19 detection also makes advantage of the classification of computed tomography (CT) images[11]. CT scan of the chest, a non-invasive technique called a scan picture yields a detailed view of the patient's chest.

Compared to a standard x-ray, the advanced x-ray technology it uses gives more detailed images of the chest. More thorough chest pictures reveal details about the patient's organs, fat, and bones, among other things. Better hints are given to doctors by this detailed information, which is also beneficial for precise diagnosis. In order to progress the current COVID-19 uncovering utilizing a limited quantity of publically accessible CXR dataset, we develop and gadget a CXR based COVID-19 sickness recognition and classification pipeline by means of a amended VGG-16 architecture in this study [12]. This study uses a multi class labeled data set with x-ray pictures

classified as normal and COVID and other. In order to enhance the COVID-19 class that is currently being offered, we train adversarial architecture to produce synthetic COVID-19 images during our trials. Our choice was primarily influenced by the lightweight nature of convolutional backbones[13-14] as well as their execution qualities, which include accuracy, precision, and recall for consistently identifying COVID-19.

LITERATURE REVIEW

In order to learn about cutting-edge methods for automated COVID-19 identification using chest images [15] and to pinpoint flaws in these studies, we conducted a thorough literature review in this chapter. We examined some of the most recent classifications that have been reported. Computer techniques for COVID-19 detection. These methods are generally broken down into three categories: 1) Machine learning classifiers 2) Classifiers using DL 3) A combined strategy. A quick description of these categories and their several research studies on COVID-19 detection have used subcategories.

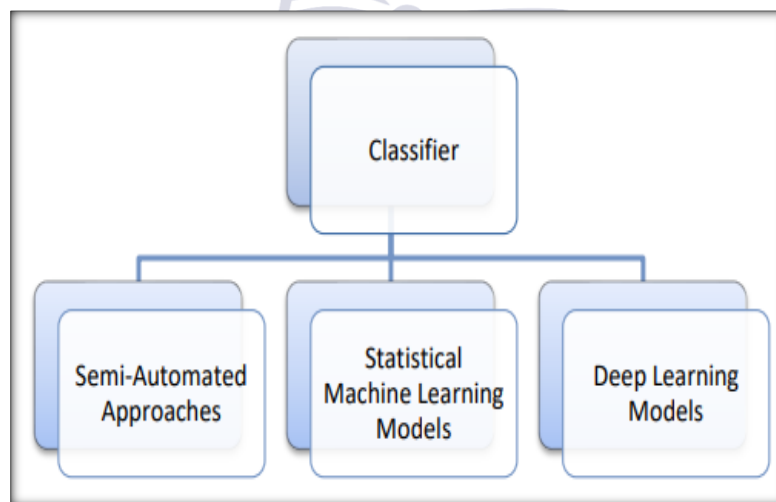


Figure 3
Figure 45

Figure 6 CLASSIFICATION TECHNIQUES

The World Health Organization (WHO) reported that a new infectious virus recognized as COVID-19 has been diffusion rapidly in a number of nations worldwide during the past several years. Typically,

COVID-19 diagnosis is connected to pneumonia symptoms, which can be identified by imaging and genetic tests. The COVID-19 timely identification helped to prevent the sickness from spreading [15]. Image testing can help quickly identify COVID-19

and, as a result, help to stop the disease's spread. The imaging approaches of computed tomography (CT) and chest X-ray (CXR) are crucial in the diagnosis of COVID-19 [16] disease. One of the most well-liked and successful methods for diagnosing COVID-19 from digital photos is the convolutional neural network (CNN). To highlight current developments in COVID-19 detection, several reviews have been conducted. For instance, a CNN based model was used to find COVID-19 illness in CT. A new study provides a modified version of the ResNet-50 pre-trained network that can categorize CT scans into three groups: COVID-19-positive, bacterial pneumonia, and healthy [17]. To identify the COVID-19 cases, those attributes were loaded into the SVM model as image classifier. The effectiveness of cutting-edge CNN architectures previously labelled for medical image categorization was assessed CXR images from patients with pneumonia, proven COVID-19 illness, and normal occurrences. According to the study, transfer learning can extract important COVID-19 disease-related parameters.

In 2021 conducted a deep learning study to detect the Covid19 using deep neural network. In this study they explained that the first imaging method that is crucial in the diagnosis of COVID-19 disease is the chest X-ray. Convolutional neural networks (CNNs) have been used with remarkable success for picture recognition and classification since large-scale annotated image datasets are widely available. The classification of medical images, however, continues to be the major issue in medical diagnosis because there aren't many annotated medical images available. By reassigning information from general object identification tasks to targeted tasks, transfer learning is an efficient process that can offer a potential solution. To identify the COVID 19 symptoms, used a deep CNN dubbed Decompose, Transfer, and Compose (DeTraC) [1]. A research work was proposed by in which Covid-19 variants are classified from genomic dataset. In this work Sarwan and Murray stated that more information about this virus, including its genetic sequence data, is becoming available as a result of COVID-19's quick global dissemination. There are presently several million genetic sequences that are publically accessible on websites like GISAID, and that number is growing daily. Researchers now have a

new opportunity to thoroughly explore this pathogen thanks to the availability of such Big Data. With all of the dynamics of the COVID-19 variations that are circulating and emerging, this is especially crucial. With the ultimate goal of reducing or completely eradicating such risks, this rich data collection provide us with knowledge on the most effective strategies to conduct genomic surveillance for current and potential future pandemic threats. It is a difficult undertaking to analyze and process the several million genetic sequences. Despite being effective, conventional approaches for classifying sequences are not intended to handle these particular kinds of genomic sequences. Additionally, the majority of the currently used approaches struggle with scalability. Previous research that was focused on coronavirus genomic data suggested employing spike sequences, which correspond to a portion of the genome, rather of the entire sequence, to carry out various machine learning (ML) tasks like classification and grouping[18]. However, scalability problems plague those methods. In this study, they offer a method termed Spike2Vec, which represents each spike sequence as a scalable and effective feature vector for use in subsequent machine learning tasks. Through studies, was demonstrated that Spike2Vec beats baseline models in terms of prediction accuracy, F1 score, and other metrics. It is also saleable on millions of spike sequences. Since this kind of study on such a large number of spike sequences hasn't been done before, the study give academics new opportunities to use this data and carry out various tasks to uncover previously undiscovered information. In order to determine the significance of each amino acid in the spike sequence, scientists additionally apply information gain (IG). The amino acids with higher IG values frequently match those published by the CDC (Centers for Disease Control and Prevention), a US-based organization, for various variations.

The World Health Organization (WHO) has designated the coronavirus, also known as COVID-19, a global pandemic. Globally, there are 19.18 million COVID-19 confirmed cases as of 8 August 2020, and there have been 0.716 million deaths. It is crucial to identify patients who have the coronavirus

to stop the virus from spreading. A CNN model is suggested on this site to identify COVID-19 patients in chest X-ray pictures. With comparative analytical analysis, two additional CNN models with various convolution layer counts and three further models built on the pretrained ResNet50, VGG-16, and VGG-19 are assessed. Datasets 1 and 2 [19] are used to train and validate each of the six models. In contrast to Dataset 2, which is significantly larger and contains 659 normal and 295 COVID-19 chest X-ray pictures, Dataset 1 contains 201 normal and 201 COVID-19 chest X-rays. With Dataset 2, the suggested model performs with an accuracy of 98.3% and a precision of 96.72%. With Dataset 2, this model yields a Receiver Operating Characteristic (ROC) curve area of 0.983 and an F1-score of 98.3. Additionally, this paper provides a comparative analysis of how the size of the dataset and the number of convolutional layers effect classification results.

Joy Long-Zong Chen and colleagues discovered in 2021 that COVID-19 viral detection using chest image categorization is effective. The histogram-oriented gradients (HOG) CNN algorithm is used in this study to extract features and identify COVID-19 patterns. The greatest results for classifying COVID-19 have come from this tertiary classification system. Disease using a data set of medical X-ray pictures from an online library. The loudness is as the number of classes rises, the tertiary classification's precision declines for the intended methodology. In

2021, Samira Chamyani and colleagues investigate the Naive Bayes, Random Forest, and KNN ML models. The author utilized the fundamental deep learning technique for chest X-ray images. Three pooling layers, two FC levels, and five convolution layers make up the Alex net. The newest deep learning method that takes advantage of batch normalization and shortcuts is Res-Net. Analysis of X-ray images has expanded in breadth thanks to deep learning [13]. The results of the experiments demonstrate that the suggested technique performs better than past approaches. The precise and accurate examination of Covid-19 illness will be aided by this research.

PROPOSED METHODOLOGY

Our suggested deep learning model for classification of covid-19 and other cases contains three tiers (The first tier, the second tier and the output tier). In the first tier of proposed model we input our CT scan images and performs different image processing techniques such as, images resizing, data augmentation, data normalization and data splitting into training and testing to improve the quality of images as well as evaluation of proposed Convolutional Neural Network (CNN) Algorithm. The second tier of our proposed model is Feature extraction and image classification using different deep learning models MobileNetV2, ResNetV2 and proposed a Custom CNN model for better performance.

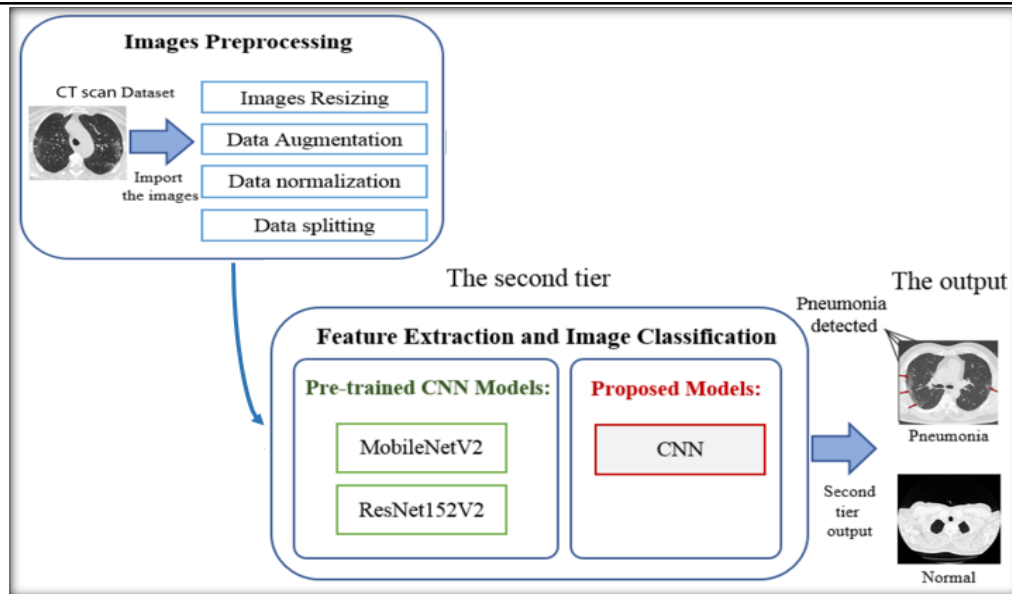


Figure 7
Figure 8 PROPOSED METHODOLOGY

The last and the final tier of our proposed methodology is uncovering the Covid-19, Pneumonia and Normal belongings from the COVID 19 image's dataset.

FIRST TIER

We divide first tier into five different steps image data input, image resizing, data augmentation, data normalization and data splitting. Following are the descriptions of how these steps are work and

improve the performance of the conducted study: The input dataset is taken from the online Kaggle dataset repository and the name of dataset is "COVID19 Chest X-Ray Images dataset". The selected dataset contains three classes (Covid 19, Normal and Viral Pneumonia). The total of 6432 images are divided as 576 images as Covid 19 cases, 1583 images form Normal cases and 4273 images from Pneumonia cases.

TABLE 1. Table 1 DATASET CLASS DISTRIBUTION

Class	Number of images
COVID - 19	576
Normal-Cases	1583
Viral Pneumonia	4273
Total	6432

SECOND TIER

The second tier of our proposed methodology involves the extraction of features and classification of images into different classes.

FEATURE EXTRACTION

In image processing, feature extraction is the process of turning raw pixel data into a collection of representative features that effectively capture pertinent details about an image's content. The visual content is characterized and described by these qualities, which improves its suitability for computer

vision or machine learning applications. The following are a few standard methods for image processing feature extraction:

Raw Pixel Values: Using the raw pixel values as features is the most basic method of feature extraction. Basic image processing jobs frequently make use of this.

Color histograms: Usually broken down into color channels, these show how colors are distributed throughout a picture (e.g., RGB histograms).

Three operators are used to identify edges in an image: Laplacian, Prewitt, and Sobel. Features can be derived from the generated edge maps.

Canny Edge Detector: An advanced edge detection method that uses several image processing stages to accurately identify edges.

We extract features by using different deep learning models such as MobileNet, ResNet and CNN.

CLASSIFICATION

We use a custom CNN classification model to classify the images by using different layers such as convolutional layer, maxpooling layer, flatten layer and dense layer. Output tier contains the classified images after every earlier tier.

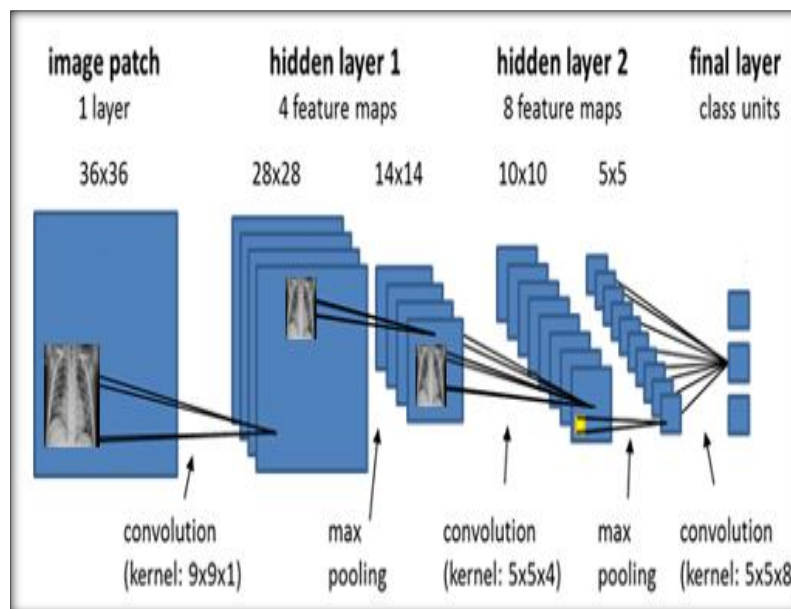


Figure 9FIGURE 4.

Figure 10 Custom CNN ARCHITECTURE
OUTPUT TIER

RESULT

This part of paper presents the empirical assessment of the suggested paradigm, delving into the fascinating world of experiments and outcomes. Here, we provide a thorough examination of the system's effectiveness and performance across a range

of scenarios and benchmarks. Our goal is to offer insights into the model's potential applications, constraints, and capabilities through thorough experimentation. Evaluating the model's performance and suitability for the current tasks is the main objective of this paper.

ARCHITECTURE OF PROPOSED MODEL

TABLE 2. Table 2 PROPOSED CNN MODEL DESCRIPTION

Parameter	Detail
Layers	6
Epochs	30

Loss	Multi class Entropy
Classes	3
Batch size	32
Activation Function	ReLU, SoftMax and Sigmoid

RESULTS MEASURES

The results of our proposed model are calculated on the basis of following parameters: Accuracy and Loss; at training and testing.

RESULTS OF TRAINING & VALIDATION ACCURACY

Table 4.3 show the training and testing/validation accuracy of the proposed CNN. According to the calculated results proposed CNN model give 96% percent accuracy on the training dataset and give 94% percent accuracy on the testing dataset.

TABLE 3. Table 3 ACCURACY

Dataset	Accuracy
Training	96%
Testing	94%

Graph in fig. 5 depict the accuracy results of model proposed model on training and testing dataset. X-axis of the contains epochs values from 0 to 30 because we run over model on 30 epochs. Y-axis of the graph labeled as accuracy it clearly

indicates the values of training and testing accuracy. There are two colors of line in the graph blue line depicts the training accuracy and orange line depicts the testing accuracy.

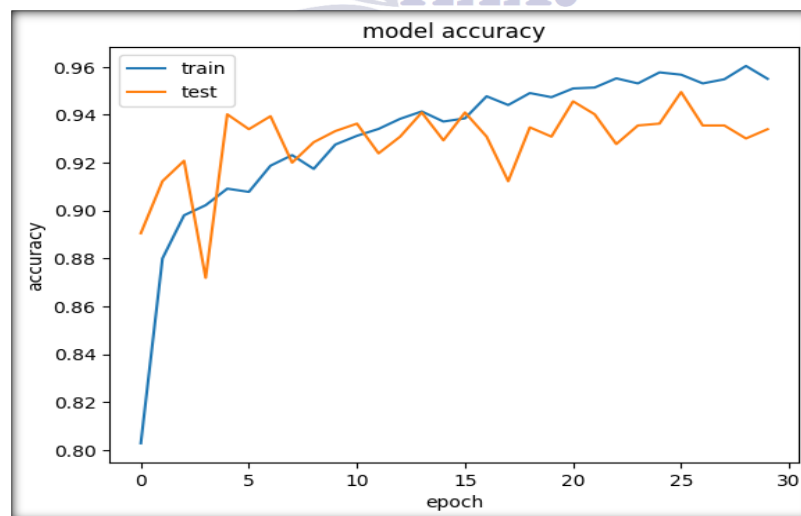


FIGURE 5. Figure 11 TRAINING & TESTING ACCURACY

Results of Training & Testing Loss

Table 4 show the training and testing/validation loss of the proposed CNN. According to the calculated

results proposed CNN model give 0.10 percent loss on the training dataset and give 0.13 percent loss on the testing dataset.

TABLE 4. Table 4 LOSS

Dataset	Loss
Training	0.10%
Testing	0.13%

Graph in fig. 6 depict the loss results of model proposed model on training and testing dataset. X-axis of the contains epochs values from 0 to 30 because we run over model on 30 epochs. Y-axis of

the graph labeled as loss it clearly indicates the values of training and testing loss. There are two colors of line in the graph blue line depicts the training loss and orange line depicts the testing loss.

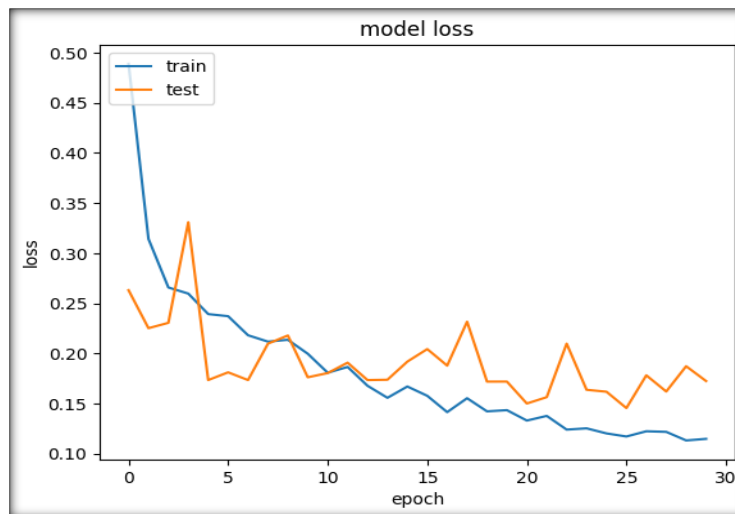


FIGURE 6. Figure 12 TRAINING & TESTING LOSS

Institute for Excellence in Education & Research

CONCLUSION

A classification strategy for COVID-19 and other instances from the chest imaging dataset was proposed in the proposed study. The suggested process comprised three steps. The initial step involved downloading a dataset of chest CT scan pictures from the Goggle website. In order to improve the quality of the images, the chosen dataset was processed using various image processing techniques, including image resizing, data augmentation, data normalization, and data splitting into training and testing. Additionally, the suggested Custom Convolutional Neural Network (CNN) Algorithm was evaluated. Feature extraction and picture classification utilizing several deep learning models, including MobileNetV2, ResNetV2, and a suggested Custom CNN model for improved performance, comprised the second phase of our suggested model. Using the provided picture collection, the suggested methodology identified

COVID-19, pneumonia, and normal patients in the third stage. Lastly, a comparison study was conducted comparing the suggested method with other cutting-edge approaches based on accuracy in testing and training as well as loss. The comparison findings demonstrated that our suggested model gave the least amount of loss and performed well, with a testing accuracy of 96% percent.

REFERENCES

- Khan, T. A., Abbas, S., Ditta, A., Khan, M. A., Alquhayz, H., Fatima, A., & Khan, M. F. (2020). IoMT-Based Smart Monitoring Hierarchical Fuzzy Inference System for Diagnosis of COVID-19. *Computers, Materials & Continua*, 65(3).

- Janjua, J. I., Khan, T. A., & Nadeem, M. (2022, January). Chest x-ray anomalous object detection and classification framework for medical diagnosis. In 2022 International Conference on Information Networking (ICOIN) (pp. 158-163). IEEE.
- Ihnaini, B., Khan, M. A., Khan, T. A., Abbas, S., Daoud, M. S., Ahmad, M., & Khan, M. A. (2021). A smart healthcare recommendation system for multidisciplinary diabetes patients with data fusion based on deep ensemble learning. *Computational Intelligence and Neuroscience*, 2021(1), 4243700.
- Khan, T. A., Fatima, A., Shahzad, T., Alissa, K., Ghazal, T. M., Al-Sakhnini, M. M., ... & Ahmed, A. (2023). Secure IoMT for disease prediction empowered with transfer learning in healthcare 5.0, the concept and case study. *IEEE Access*, 11, 39418-39430.
- Munnoli, P.M., Nabapure, S. & Yeshavanth, G. Post-COVID-19 precautions based on lessons learned from past pandemics: a review. *J Public Health (Berl.)* 30, 973-981 (2022). <https://doi.org/10.1007/s10389-020-01371-3>
- Abbas, T., Khan, A. H., Kanwal, K., Daud, A., Irfan, M., Bukhari, A., & Alharbey, R. (2024). IoMT-Based Healthcare Systems: A Review. *Computer Systems Science & Engineering*, 48(4).
- Javed, R., Abbas, T., Khan, A. H., Daud, A., Bukhari, A., & Alharbey, R. (2024). Deep learning for lungs cancer detection: a review. *Artificial Intelligence Review*, 57(8), 197.
- J. DeGrave, J. D. Janizek, and S. I. Lee, "AI for radiographic COVID-19 detection selects shortcuts over signal," *Nat. Mach. Intell.*, vol. 3, no. 7, pp. 610-619, 2021.
- Abbas, A., Abdelsamea, M. M., & Gaber, M. M. (2021). Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Applied Intelligence*, 50(10), 3670-3680.
- Apostolopoulos, I. D., & Mpesiana, T. A. (2021). Covid-19: a deep learning approach for detection and diagnosis. *IEEE Transactions on Intelligent Systems*, 1-8.
- Luz et al., "Towards an effective and efficient deep learning model for COVID-19 patterns detection in X-ray images," *Res. Biomed. Eng.*, vol. 38, no. 1, pp. 149-162, 2022.
- Hemdan, E. E., Shouman, M. A., & Karar, M. E. (2020). COVIDX-Net: a framework of deep learning classifiers to diagnose COVID-19 in X-ray images. *Artificial Intelligence in Medicine*, 110, 101-108.
- Khan, M. A., Abbas, S., Khan, K. M., Al Ghamdi, M. A., & Rehman, A. (2020). Intelligent forecasting model of COVID-19 novel coronavirus outbreak empowered with deep extreme learning machine. *Computers, Materials & Continua*, 64(3), 1329-1342.
- Khan, M. A., Abbas, S., Atta, A., Ditta, A., Alquhayz, H., Khan, M. F., & Naqvi, R. A. (2020). Intelligent Cloud Based Heart Disease Prediction System Empowered with Supervised Machine Learning. *Computers, Materials & Continua*, 65(1).
- Arias-Garzón et al., "COVID-19 detection in X-ray images using convolutional neural networks," *Mach. Learn. with Appl.*, vol. 6, no. April, p. 100138, 2021.
- Farooq, M., Hafeez, A., & Anwar, W. (2021). COVID-ResNet: A hybrid deep learning framework for the detection of COVID-19 cases from chest X-ray images. *Biocybernetics and Biomedical Engineering*, 41(2), 608-624.
- Narin, A., Kaya, C., & Pamuk, Z. (2021). Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Analysis and Applications*, 24, 1207-1220.

- Apostolopoulos, I. D., & Mpesiana, T. A. (2020). COVID-19: automatic detection from X-ray images utilising transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*, 43(2), 635–640
- Alqudah AM Qazan S (2020) Augmented COVID-19 X-ray images dataset. Mendeley Data, V4. 10.17632/2fxz4px6d8.4

