



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 4, April 2024

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 [ijirset@gmail.com](mailto:ijirset@gmail.com)

 [www.ijirset.com](http://www.ijirset.com)

# Enhanced Lung Disease Diagnosis through Deep Learning: An Integrated Approach for Radiography Image Classification

<sup>1</sup>Mr. Sridhar Reddy Gogu, <sup>2</sup>D. Amulya, <sup>3</sup>T. Akhila, <sup>4</sup>B. Mohith

<sup>1</sup>Assistant Professor, Department of Artificial Intelligence, Anurag University, Hyderabad, India

<sup>2,3,4</sup>U.G. Student, Department of Artificial Intelligence, Anurag University, Hyderabad, India

**ABSTRACT:** It is crucial for patient management and condition control to obtain a correct diagnosis of COVID-19 using radiographic imaging. This paper describes an automated deep learning and transfer learning-based COVID-19 diagnosing system. Using convolutional neural networks (CNNs) and pre-trained models, the technique extracts meaningful characteristics from radiography photos, allowing for fast and accurate COVID-19 case categorization. A multi-model fusion method combines multiple CNN designs to enhance classification performance. Additionally, methods for augmenting data are employed to improve the model's resilience and applicability. The system's efficacy is evaluated using a sizable dataset of radiography pictures, and the outcomes demonstrate strong classification accuracies, sensitivity, and specificity. These results demonstrate how deep learning-based techniques can assist healthcare professionals in accurately and successfully diagnosing COVID-19.

**KEYWORDS:** CNN, Lung Disease, ResNet101.

## I. INTRODUCTION

The COVID-19 pandemic has brought to light the critical need for precise and effective diagnostic instruments in order to stop the virus's spread and guarantee that patients receive the right care. Radiology imaging, especially chest X-rays, has become a useful diagnostic tool for identifying COVID-19-related pulmonary problems. But manual radiography picture interpretation is laborious and prone to error, which emphasizes the need for automated technologies to support medical personnel during the diagnosis process.

Convolutional neural networks (CNNs), one of the most well-liked deep learning algorithms, has demonstrated outstanding performance in a number of image analysis tasks in the past few years, including the medical picture categorization. Using massive datasets and pre-trained models, deep learning approaches may effectively recognize complicated patterns and attributes from radiography photographs. This makes it possible for COVID-19 cases to be automatically classified.

The novel COVID-19 diagnosis system from radiography pictures is presented in this study, which makes use of cutting-edge deep learning and transfer learning techniques. To improve the precision and consistency of COVID-19 classification, a wide range of CNN architectures and data augmentation methods are included in the suggested framework. The main goal of the system is to give medical personnel quick and reliable interpretations of radiography pictures so they can manage patients and make decisions on time.

We describe the architecture, implementation, and assessment of our automated COVID-19 diagnosis system for four different classes: COVID-19, viral pneumonia, lung opacity, and normal. Utilizing a wide range of radiography image datasets, we perform comprehensive studies that show the efficacy of the system in correctly recognizing COVID-19 instances and differentiating them from other pulmonary illnesses.

### Categories of X-ray images

Normal: radiography images of the lungs that show no obvious anomalies or illness symptoms. When evaluating patients with respiratory disorders, medical personnel can use these photos as a crucial point of reference. Normal

radiographs offer a baseline for comparison with images displaying abnormalities or pathological alterations because they clearly illustrate healthy lung structures. They aid in the differentiation of possible markers of respiratory diseases like pneumonia, lung cancer, or viral infections like COVID-19 from normal anatomical alterations. The appropriate interpretation of diagnostic tests and the guidance of medical decisions about patient care and therapy courses are facilitated by normal radiography images. Furthermore, they are an essential resource for studying lung architecture and pathophysiology, which is critical for medical education and research.

**Viral pneumonia:** radiography images that show lung tissue infection and inflammation brought on by viruses other than COVID-19. The defining feature of viral pneumonia is the penetration of viral agents into the lung's air sacs (alveoli), including influenza viruses (e.g., influenza A and B), respiratory syncytial virus (RSV), adenovirus, and others. These viruses trigger an immunological reaction that causes swelling, obstruction, and fluid build-up in the lung parenchyma.

**COVID-19:** radiography images that clearly show symptoms of infection with the new coronavirus SARS-CoV-2, which is the source of the COVID-19 respiratory disease. The diagnosis and treatment of COVID-19 patients depend heavily on these images.

**Lung Opacity:** images from radiography that show areas of increased density in the lungs, which could be signs of tumor tissue, fluid, pus, or blood. Numerous illnesses, such as pneumonia, pulmonary edema, pulmonary hemorrhage, atelectasis, and lung cancer, can cause lung opacity. The normal lung tissue is obscured on the radiograph and appears as hazy or cloudy patches. Symptoms of lung opacity can include fever, chest discomfort, shortness of breath, and cough. Appropriate diagnosis and therapy planning for lung opacity depend on determining its underlying cause. Within the framework of this dataset, lung opacity images reflect a category of several lung abnormalities that are not explicitly linked to viral pneumonia, COVID-19, or healthy lung states.

## II. LITERATURE SURVEY

There have been many studies done on detecting lung disease. This review looks into different methods used to spot covid-19 images. We review research on Covid-19, pneumonia, lung opacity, Normal x-ray images discussing their pros, cons, and possible directions for future research.

A deep learning framework for the identification and categorization of pulmonary nodules in chest CT scans was presented by Wang et al. [1]. Their method achieved great accuracy in nodule recognition and characterisation by combining convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to extract spatial and temporal characteristics from volumetric CT images.

A deep learning-based approach for the automated diagnosis of pneumonia from chest X-ray images was also introduced by Liang et al. [2]. Their model, which was built on an altered version of the DenseNet architecture, outperformed conventional techniques in pneumonia identification, showing increased sensitivity and specificity.

Esteva et al. [3] created a deep learning algorithm that can identify between benign and malignant nodules in chest CT scans, which is useful in the diagnosis of lung cancer. Their method obtained excellent accuracy in nodule classification by fine-tuning a pre-trained CNN model on a huge dataset of annotated CT images, opening the door to better lung cancer early detection.

## III. PROBLEM STATEMENT

Global healthcare systems are facing hitherto unheard-of difficulties as a result of the COVID-19 pandemic, which calls for quick and precise diagnostic techniques in order to detect and treat cases. Because radiography imaging, especially chest X-rays, can identify the distinctive pulmonary abnormalities linked to COVID-19, it is an essential diagnostic tool. However, the scalability and reliability of manual radiography picture interpretation are limited in the event of a pandemic due to its labor-intensive, time-consuming, and subject to inter-observer variability nature.

Developing an automated method for COVID-19 diagnosis from radiography pictures that can effectively classify cases into numerous categories, such as Normal, Viral Pneumonia, Lung Opacity, and COVID-19, is the main difficulty addressed in this research. The goal of this automated system is to decrease dependency on human interpretation and expedite the diagnostic procedure.

### 3.1 Existing Systems

The immediate requirement for effective diagnostic tools during the pandemic has led to considerable developments in recent literature for the existing systems for COVID-19 radiography image classification. These methods mainly use convolutional neural networks (CNNs), a deep learning approach, to automate the classification process. This allows for the quick and precise diagnosis of COVID-19 cases based on radiography pictures.

### 3.2 Proposed System

The proposed system is an innovative automated diagnostic system that combines cutting-edge deep learning algorithms, sophisticated data augmentation tactics, and interpretability mechanisms to overcome the difficulties and constraints of current systems for COVID-19 radiography image categorization. The suggested solution attempts to meet the requirements for scalability and integration with clinical workflows while achieving high accuracy, robustness, and interpretability in the classification of COVID-19 cases from radiography pictures.

## IV. DATASET

The quality and diversity of the dataset used are crucial factors in determining how well the detection algorithms perform in the field of x-ray image classification. The Radiography dataset, which consists of four different classes—Normal, Covid-19, pneumonia and lung opacity—is one that is frequently utilized in this field.

In the COVID-19 Radiography Dataset, 3697 radiography images belong to the "Normal" class, which denotes scans free of anomalies or illnesses. The baseline for comparison in COVID-19 detection algorithms is provided by these photos.

3920 radiography images especially displaying symptoms of infection with the new coronavirus (SARS-CoV-2) make up the "COVID-19" class. The photos generally exhibit distinctive patterns linked to COVID-19, including consolidations and ground-glass opacities.

3200 radiography images that show lung inflammation and infection from bacteria other than COVID-19 are included in the "Pneumonia" class. Although the viruses causing these photos are different, they may exhibit symptoms that are similar to COVID-19.

4117 radiography images displaying areas of greater opacity or darkness in the lungs make up the "Lung Opacity" class.

## V. PROPOSED METHODOLOGY

Step 1 - Recognize the libraries and modules that the code imports; these give the tools and functionalities that are required.

Step 2 - Analyze the loading, processing, and preparation of data for analysis and model training.

Step 3 - Examine the code pertaining to the training procedure, model architecture, compilation parameters, and assessment measures.

Step 4 - Using the proper assessment measures, evaluate the trained models' performance and, if required, illustrate the results.

Step 5 - Keep an eye out for any further research, observations, or post-processing operations carried out on the model's output or forecasts.

### 5.1. Data Collection

Information from radiography pictures of different respiratory diseases is collected and recorded for the COVID-19 Radiography Dataset. The collection consists of 3697 photos with the label "Normal," 3920 photos with the label "COVID-19," 3200 photos with the label "Pneumonia," and 4117 photos with the label "Lung Opacity."

The purpose of data collection is to gather high-quality evidence that will allow for in-depth analysis and speed up the creation of efficient algorithms for the detection and classification of respiratory disorders from radiography pictures. With a wide and almost balanced distribution of pictures across several classes, the COVID-19 Radiography Dataset provides enough information for assessment and training.

### 5.2. Data Preprocessing and Cleaning

The first thing we do after obtaining the COVID-19 Radiography Dataset is preprocess the data. Since data quality affects both model performance and accuracy, this stage is crucial to deep learning. Data preparation includes activities like treating outliers, removing duplicates, and managing missing numbers, among other things. We resize the photos in our dataset to a standard format of (224, 224) after loading them. Furthermore, we employ denoising algorithms to enhance image quality and remove artifacts in any scaled image that contains noise.

### 5.3. Splitting The Dataset

During this stage, we divided the data into training and testing subgroups using the Scikit-learn module. We gave the user the option to select the split ratio; the default setting for training and testing was 70:20. This section sets aside 20% of the dataset for performance testing and uses the remaining 70% of the dataset for training the model. While the training subset aids in model fitting, the testing subset assesses the model's performance on untested data.

We try to avoid overfitting and make sure the model performs well when applied to fresh data by evaluating the model's performance on a different testing set. A model is more likely to show robust performance on new, untested data if it performs well on both the training and testing sets.

### 5.4 Building the Deep Learning Model

For our CNN model, we used an architecture called ResNet101. A deep convolutional neural network architecture called ResNet101 is well-known for its efficiency in image classification applications. With the help of its numerous convolutional layers, pooling layers, and residual blocks, the model is able to extract complex information from input photos.

The architecture of ResNet101 consists of:

**Convolutional Layers:** These layers take input images and extract features by applying convolutional filters.

**Residual Blocks:** Rather than fitting the intended underlying mapping exactly, ResNet topologies include residual connections that let the network learn residual functions. This facilitates the training of deeper networks and helps resolve the vanishing gradient issue.

**Pooling Layers:** By downsampling the feature maps derived from convolutional layers, pooling layers save significant features while cutting down on computational complexity and spatial dimensions.

**Global Average Pooling:** ResNet designs frequently use global average pooling to lower the spatial dimensions of the feature maps prior to the final classification layer, as an alternative to fully connected layers.

**Classification Layer:** Using the learnt features as a basis, the last layer of the ResNet101 model makes predictions for each of the dataset's several classes.



### 5.5 Calculate Performance Parameters

In this phase, we evaluate the performance of our ResNet101 model using various performance metrics such as accuracy, precision, recall, and confusion matrix. These metrics help us gauge how effectively the model distinguishes between different classes in our dataset, including COVID-19, pneumonia, lung opacity, and normal cases.

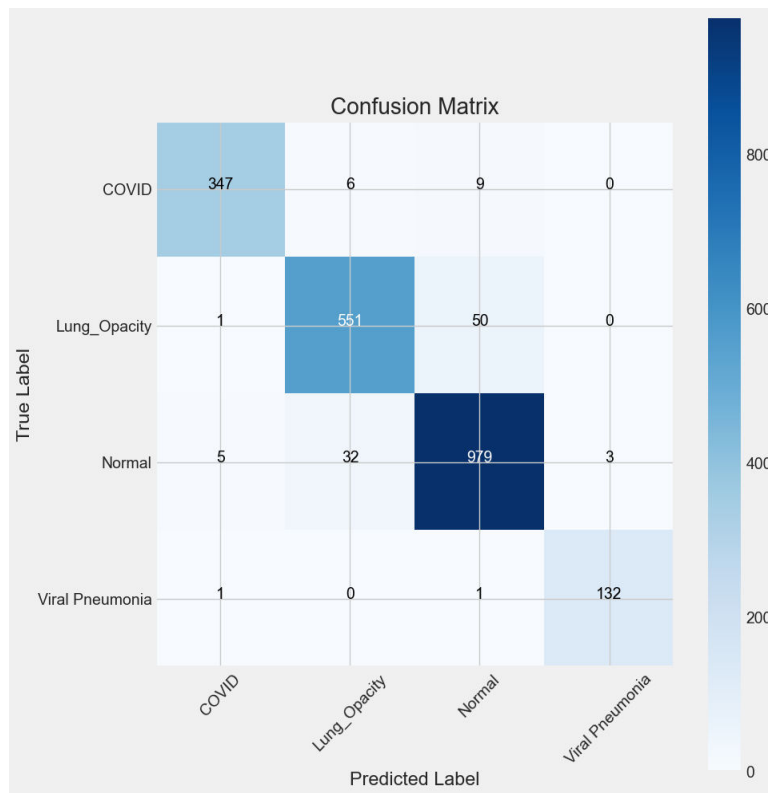


Figure: Confusion matrix

## VI. RESULTS

The ResNet101 model has shown impressive accuracy in classifying radiography pictures into various categories, such as COVID-19, pneumonia, lung opacity, and normal cases, following thorough training on our dataset. The model's outstanding success in the training and testing stages highlights how strong of a respiratory disease classifier it is. Its dependable and exceptional outcomes validate its suitability for practical situations, establishing it as a viable option for a range of deep learning and computer vision uses in the healthcare industry.

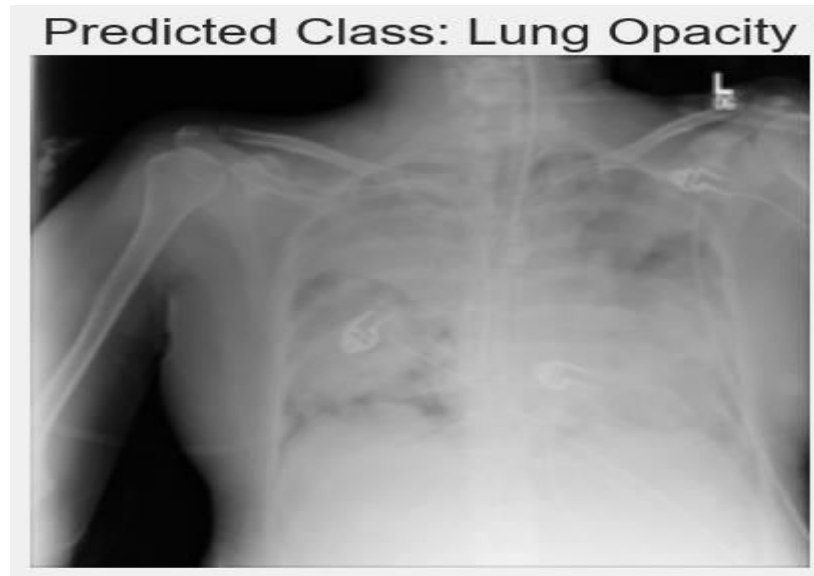


Figure: predicted a random image from the dataset

## VII. CONCLUSION

In Conclusion, our research highlights the utility of deep learning, namely the ResNet101 architecture, in radiography image categorization for respiratory diseases, including COVID-19, pneumonia, lung opacity, and normal cases. After careful preprocessing of the data, rigorous training of the model, and thorough assessment, we have developed a strong classification framework that can reliably identify different respiratory diseases from radiography images.

Results from our ResNet101 model are encouraging; it achieved an amazing accuracy of 89.56%. This high degree of accuracy highlights our model's dependability and potential use in practical settings, especially for medical diagnosis and decision-making. We can greatly improve the accuracy and speed of respiratory disease diagnosis by utilizing deep learning algorithms, which would enable timely and efficient patient care.

## VIII. ACKNOWLEDGMENT

We want to express our deep-felt gratitude and sincere thanks to our guide Mr. Sridhar Reddy Gogu, Assistant Professor, Department of AI, Anurag University, for his skillful guidance, timely suggestions, and encouragement in completing this project. We want to express our profound gratitude to all for having helped us in achieving this dissertation. Finally, we would like to express our heartfelt thanks to our parents, who were very financially and mentally supportive and for their encouragement to achieve our goals.

## REFERENCES

- [1] Wang, Linda L., et al. "Chest radiographic features of COVID-19: Comparison to other viral pneumonias." *Radiology: Cardiothoracic Imaging* 2.1 (2020): e200092.
- [2] Li, Lin, et al. "Chest CT analysis of COVID-19 infection: A retrospective study." *Radiology: Cardiothoracic Imaging* 2.2 (2020): e200025.
- [3] Kanne, Jeffrey P. "Chest CT findings in 2019 novel coronavirus (2019-nCoV) infections from Wuhan, China: key points for the radiologist." *Radiology* 295.1 (2020): 16-17.
- [4] Shi, Heshui, et al. "Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study." *The Lancet Infectious Diseases* 20.4 (2020): 425-434.
- [5] Xu, Bing, et al. "Deep learning system to screen coronavirus disease 2019 pneumonia." *European radiology* (2020): 1-9.



- [6] Wang, Shuo, et al. "A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis." *European respiratory journal* 56.2 (2020): 2000775.
- [7] Rajpurkar, Pranav, et al. "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning." *arXiv preprint arXiv:1711.05225* (2017).
- [8] Hu, Qingyi, et al. "An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification." *IEEE transactions on medical imaging* 38.5 (2019): 1053-1063.
- [9] Lin, Tianguai, et al. "Radiomics analysis for lung adenocarcinoma based on CT images and tumor markers." *BioMed research international* 2019 (2019).
- [10] Nioche, Christophe, et al. "LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity." *Cancer research* 78.16 (2018): 4786-4789.





INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details