DETECTION AND CLASSIFICATION OF LUNG DISEASES FOR PNEUMONIA AND COVID-19 USING MACHINE AND DEEP LEARNING TECHNIQUES

Prof. Kishor Ugale[1], Bhushan Sandeep Waghmare[2], Kunal Ganesh Bhavsar[3], Om Somnath Shinde[4]

[1]*Professor, Electrical Department, MET's Institute of Engineering, Maharashtra, India*
[2] *Student, Electrical Department, MET's Institute of Engineering, Maharashtra, India*
[3]*Student, Electrical Department, MET's Institute of Engineering, Maharashtra, India*
[4]*Student, Electrical Department, MET's Institute of Engineering, Maharashtra, India*

ABSTRACT

Lung Disease is common throughout the world. These include chronic obstructive pulmonary disease, pneumonia, asthma, tuberculosis, fibrosis, etc. Timely diagnosis of lung disease is essential. Many image processing and machine learning models have been developed for this purpose. The recent development of imaging and sequencing technologies enables systematic advances in the clinical study of lung cancer. Machine learning-based approaches play a critical role in integrating and analyzing these large and complex datasets, which have extensively characterized lung cancer through the use of different perspectives from these accrued data. Machine learning techniques, ranging from traditional algorithms to sophisticated deep learning models, have proven valuable in the analysis of medical images for lung disease detection. The emergence of CNNs, designed to excel in image analysis tasks, has particularly transformed this field, enabling automated detection and classification with unprecedented accuracy. These technologies not only offer the ability to expedite diagnosis but also to reduce the workload on healthcare professionals. The benefits of utilizing machines and deep learning techniques extend beyond accuracy and speed; they hold the potential to create scalable and robust solutions applicable in diverse healthcare settings, including resource-limited environments. This abstract provides a foundation for a comprehensive exploration of these methods, encompassing their methodologies, challenges, and prospects, as they play a pivotal role in the fight against respiratory diseases, ultimately contributing to more efficient and effective healthcare systems

Keyword: - *Lung Disease, Machine Learning, Deep Learning, Image Processing,*

---

1. INTRODUCTION

Human cells, which heavily rely on a constant and sufficient oxygen supply to operate efficiently, will suffer if oxygen inflow is interrupted. The lungs are where the body takes in oxygen and releases extra carbon dioxide, which can be harmful to the body. However, since the lungs are fragile organs that are continuously vulnerable to the outer environment, they are exposed to many diseases known as lung diseases. Many of these diseases include lung cancer, chronic obstructive pulmonary disease, Covid-19, asthma, chronic bronchitis, influenza, lung fibrosis, sarcoidosis, and tuberculosis. The most concentrated lung disease these days is Covid-19, because of the pandemic. To date, at least 106 million cases and at least 2 million reported fatalities have been reported by the COVID-19 pandemic. Early diagnosis of infectious patients and accurate surveillance of critically ill patients are crucial for pandemic management and mortality mitigation due to the high transmittable rate, long incubation time, and comparatively high mortality rate.

Detection and classification of lung diseases, particularly pneumonia and COVID-19, have become critical areas of research and application in the field of medical imaging and healthcare. Machine learning and deep learning techniques have emerged as powerful tools to address this pressing need. In this introduction, we will discuss the significance of these techniques in the context of identifying and distinguishing pneumonia and COVID-19 in lung images. Machine learning methods, such as support vector machines, decision trees, and random forests, have traditionally been used in image analysis for lung disease detection. However, the advent of deep learning, particularly convolutional neural networks (CNNs), has revolutionized this field. CNNs are specifically designed to excel in image analysis tasks, enabling the automated detection and classification of lung diseases with remarkable accuracy. They have demonstrated their potential not only in the diagnosis of pneumonia and COVID-19 but also in distinguishing them from other lung conditions, offering more comprehensive healthcare solutions.

2. LITERATURE SURVEY

The arrival of Covid-19 has brought significant threat to human life which started from China in November 2019 and later spread across the world. It has been reported that more than 63.2 million people have already been infected in the world, which includes approximately 1.47 million deaths. The world health organization (WHO) continuously provides the necessary information for nations to protect against Covid-19. Countries like United States, India, Brazil, Russia, France, Italy, China are the highly suffered nations from this threat. The people infected with Covid-19 have moderate or mild symptoms such as fever, cough, and breathe shortness. However, people suffered from severe pneumonic conditions in their lungs that resulted in death as well. Lungs may fill up with a significant amount of fluid, causing difficulty in breathing. Pneumonia may be caused by viral infections (like Covid-19 or flu), common cold, and bacterial infections. Due to the arrival of Covid-19 disease, it is a very challenging task for medical experts to detect lung infections (either viral/bacterial pneumonia or Covid-19 pneumonia) from chest X-ray images . [1]

The usage of chest X-rays can help screen cases of COVID-19. They were already applied to detect the SARS-COV-1 and Middle East Respiratory Syndrome (MERS) coronavirus, as they allow a quick, consistent, and cheaper pneumonia detection. COVID-19 leads to radiological evidence of lower respiratory tract lesions of lung even in patients who do not have clinical pneumonia [8], which favors detection by radiographs of a larger group of contaminated people. To this end, it is critical to differentiate the COVID- 19 pneumonia radiographs from those with other types of pneumonia since they may have some similarities in the lung's affected areas [2]

Machine getting to know (ML) is a utility of AI which suggests flexibility to robotically study and enhance its performance by experience without programming. ML classifiers are very famous for detecting breast and lung cancers. ML algorithms can be differentiated into three different categories namely Supervised, Unsupervised, and Reinforcement Learning. We used an exclusive model for detecting carcinoma in CT test pictures with the aid of using an ensemble classifier that consists of five exclusive ML supervised algorithms like decision tree, KNN, SVM, RF, MLP, logistic regression, etc. to get more accurate results.[3]

Machine learning is an exciting field of research in computer science and engineering. It is considered a branch of artificial intelligence because it enables the extraction of meaningful patterns from examples, which is a component of human intelligence. The appeal of having a computer that performs repetitive and well-defined tasks is clear: computers will perform a given task consistently and tirelessly; however, this is less true for humans. More recently, machines have demonstrated the capability to learn and even master tasks that were thought to be too complex for machines, showing that machine learning algorithms are potentially useful components of computer aided diagnosis and decision support systems. Even more exciting is the finding that in some cases, computers seem to be able to "see" patterns that are beyond human perception. This discovery has led to substantial and increased interest in the field of machine learning— specifically, how it might be applied to medical images.[4]

3. OBJECTIVES

1. Early Detection: Early detection of lung diseases is crucial for improving patient outcomes and reducing mortality rates. Machine learning algorithms can play a vital role in this by analyzing medical imaging data, such as X-rays and CT scans, to identify subtle abnormalities indicative of lung diseases. Research has shown that machine learning models can achieve high accuracy in detecting various lung conditions, including pneumonia, tuberculosis, and lung cancer, at an early stage when treatment is most effective [1]. For instance, a study published in Nature Medicine demonstrated that a deep learning model outperformed radiologists in detecting malignant lung nodules on CT scans, highlighting the potential of machine learning in improving early diagnosis [2].

2. Data Handling: Machine learning excels at handling large volumes of medical data, which is essential for discovering complex patterns and correlations relevant to lung disease diagnosis. With advancements in imaging

technology, healthcare institutions generate vast amounts of imaging data daily. Machine learning algorithms can efficiently process this data, extract meaningful features, and identify disease-specific patterns that may not be apparent to human diagnosticians. Moreover, techniques such as transfer learning allow models to leverage knowledge learned from one dataset to improve performance on another, enhancing the generalizability of diagnostic models [3].
3. Public Health Initiatives: Machine learning has the potential to support public health initiatives aimed at controlling infectious diseases like tuberculosis and COVID-19. For instance, machine learning algorithms can analyze chest X-rays or CT scans to detect signs of tuberculosis infection, facilitating early diagnosis and treatment initiation [4]. In the context of COVID-19, machine learning models have been utilized for mass screening of chest X-rays to identify individuals with suspected COVID-19 pneumonia, aiding in triage and resource allocation during the pandemic [5]. Additionally, machine learning techniques can assist in contact tracing by analyzing demographic and clinical data to identify individuals at higher risk of infection and monitor disease spread in communities [6].
By leveraging machine learning's capabilities in early detection, data handling, and supporting public health initiatives, the overarching aim is to enhance the accuracy, efficiency, and accessibility of lung disease diagnosis and management. Ultimately, these efforts contribute to improving patient care and advancing medical knowledge in the field of pulmonary medicine.

4. SYSTEM ARCHITECTURE:

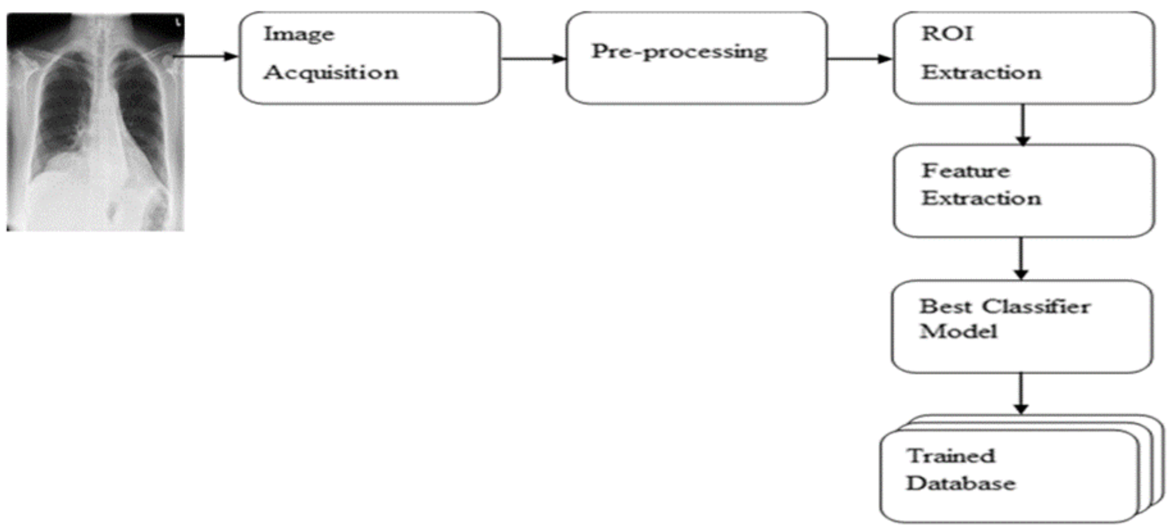

Fig -1: System Architecture

4.1 Data Acquisition and Preprocessing: Chest X-ray or CT scan images are acquired from diverse sources such as medical databases or healthcare institutions. Preprocessing is essential to standardize images, adjust contrast, and resize them to a consistent format, ensuring data quality and consistency [1]. Techniques like histogram equalization and normalization are commonly used for preprocessing medical images [2].
4.2 Image Input: Preprocessed images serve as input data for machines and deep learning models. These images capture vital visual cues related to lung conditions like pneumonia or COVID-19, enabling accurate disease detection [3].
4.3 Feature Extraction: Feature extraction involves identifying relevant patterns and features from image data for classification. Deep learning models, particularly Convolutional Neural Networks (CNNs), automatically learn hierarchical representations of features from raw image data [4]. Traditional machine learning methods often require manual feature extraction, where domain knowledge is used to engineer informative features from the images [5].
4.4 Machine Learning and Deep Learning Models: This step represents the core of the system, where various machine learning and deep learning models are employed for disease detection and classification. Machine learning models such as support vector machines, decision trees, random forests, or logistic regression offer interpretability and efficiency in classification tasks .On the other hand, deep learning models, especially CNNs, excel in learning complex patterns directly from raw image data, leading to state-of-the-art performance in medical image analysis .

4.5 Training: Models are trained using a labeled dataset of chest images, where each image is associated with disease labels (e.g., pneumonia, COVID-19, or healthy) for supervised learning. During training, models adjust their internal parameters to minimize classification errors and improve their ability to differentiate between different lung conditions.
4.6 Validation and Testing: Trained models are evaluated using separate validation and testing datasets. The validation dataset aids in fine-tuning hyper parameters and preventing over fitting, while the testing dataset provides an independent assessment of the model's performance.
4.7 Classification Output: Once trained and tested, models are used to classify new, unseen chest images. The classification output indicates the predicted lung disease, aiding in early diagnosis and treatment planning.
4.8 Post-processing and Visualization: Post-processing steps involve generating visual reports, providing confidence scores for predictions, and creating heat maps to highlight regions of interest in the images. These visualizations aid healthcare professionals in interpreting model predictions and making informed decisions.
4.9 Decision Support: The system provides decision support for healthcare practitioners by presenting model classifications and relevant information, assisting them in making accurate diagnoses and treatment decisions [12].
4.10 Reporting and Archiving: Results and predictions are stored in databases or healthcare information systems for future reference, research, and documentation, contributing to the advancement of medical knowledge in pulmonary medicine.
These steps outline a comprehensive framework for leveraging machine learning and deep learning techniques in lung disease detection and classification, aiming to enhance patient care and advance medical research.

5. METHODOLOGY

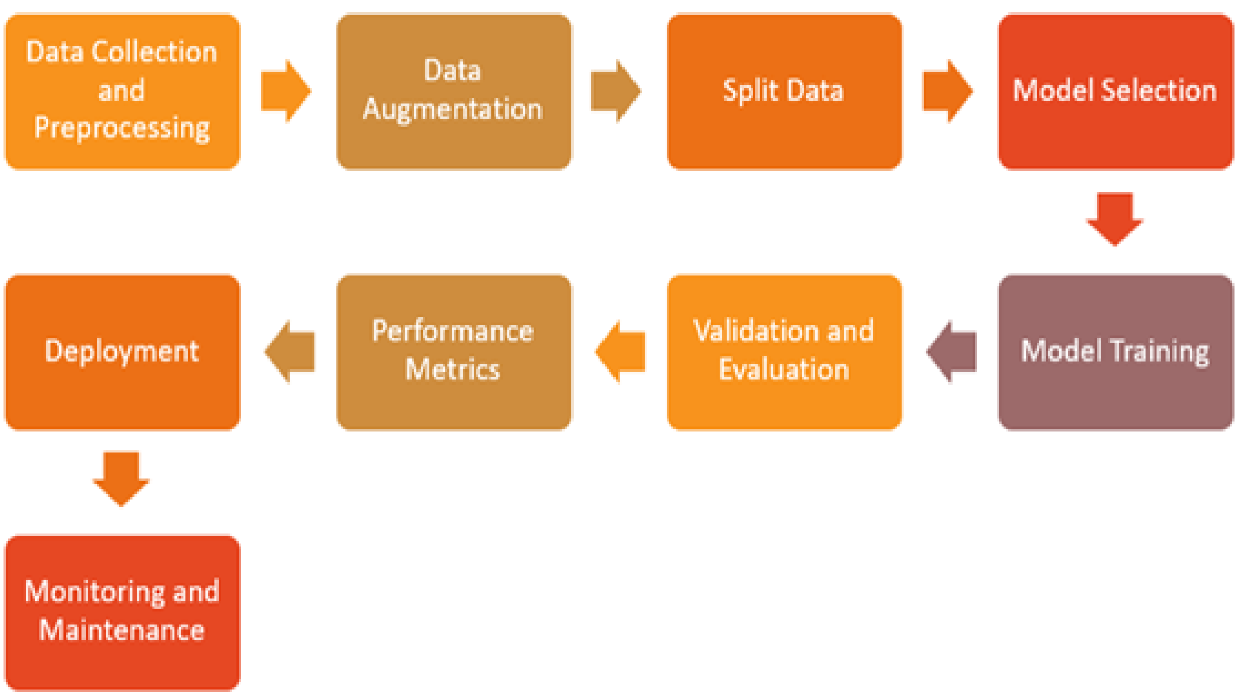

Fig -2: System Methodology

6. RESULTS

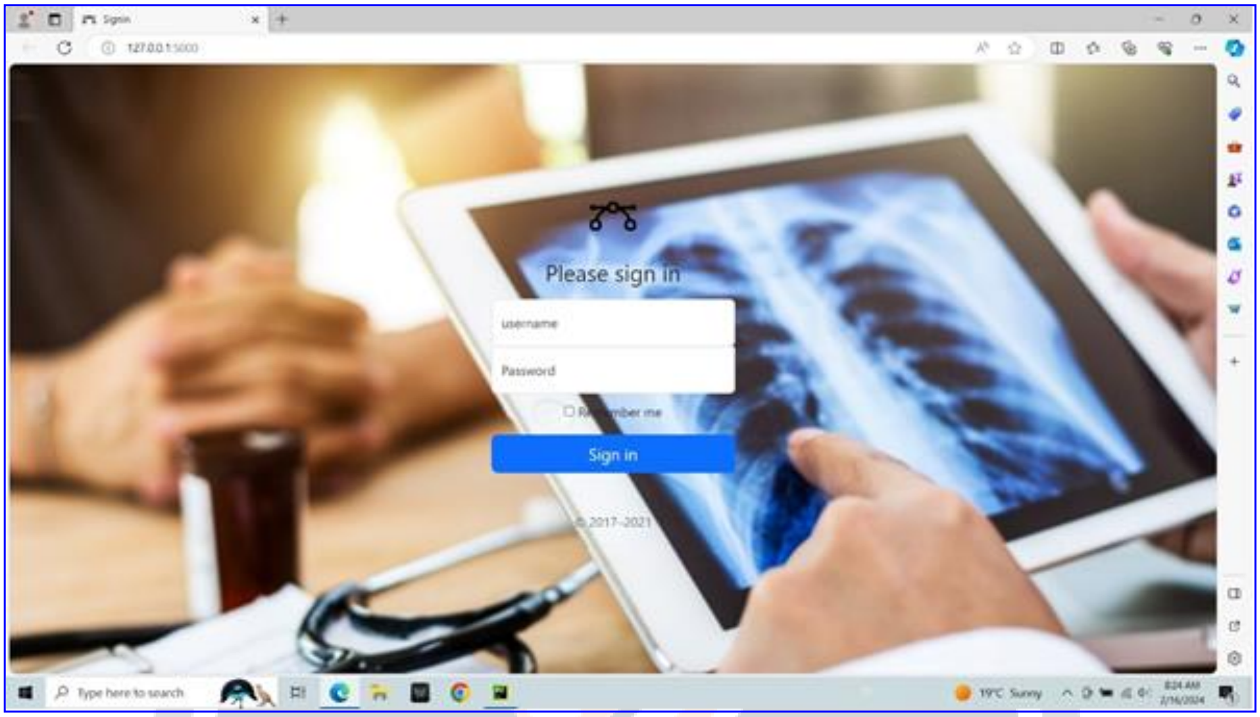

Fig -3: Sign in Page

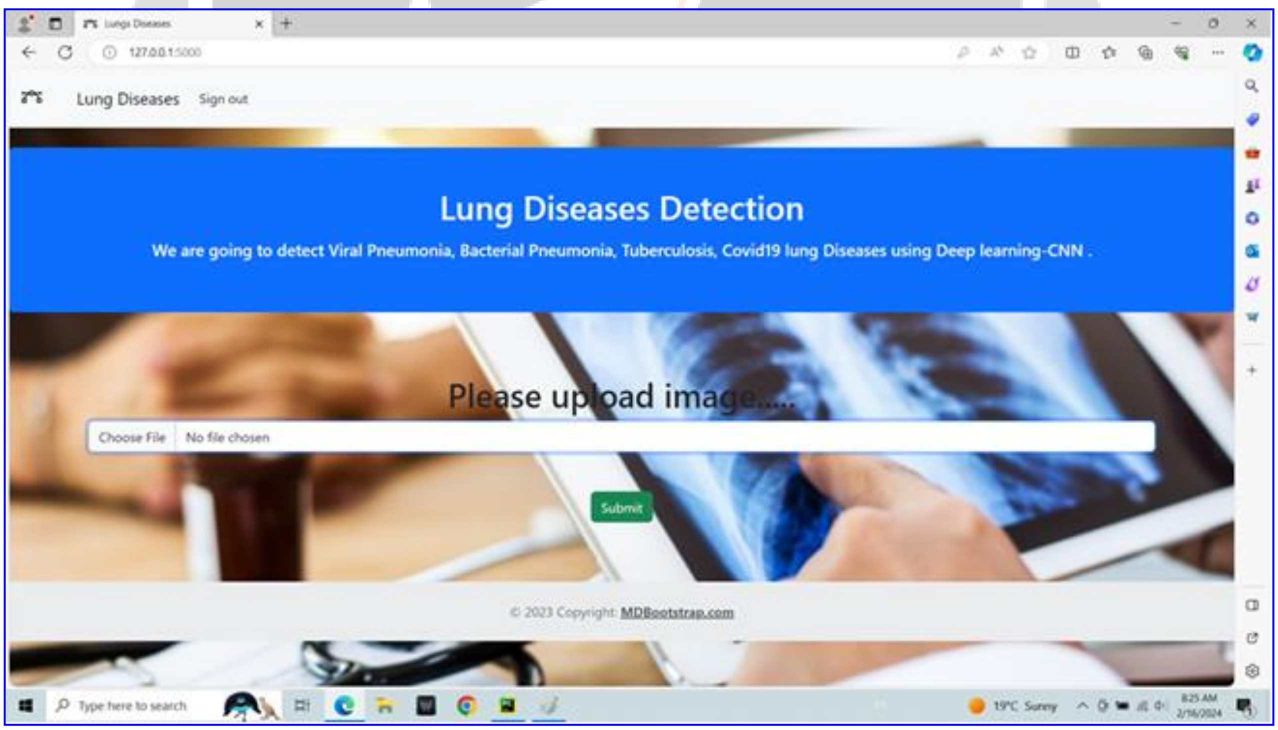

Fig -4: Detection Page

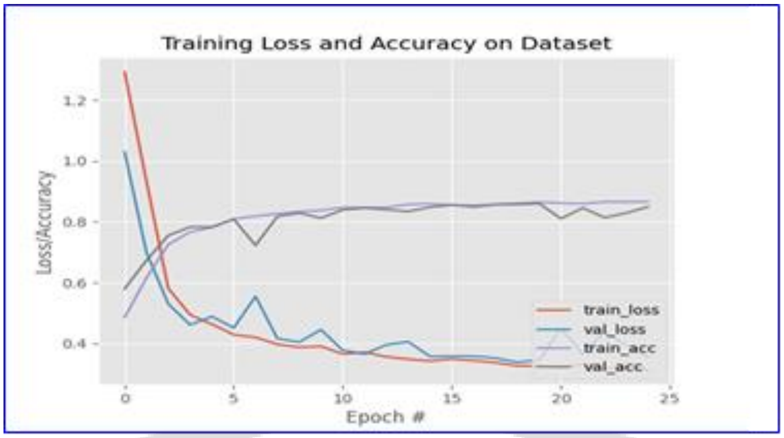

Fig -5: Trained data output

classification report

```
In [17]:  class_report = classification_report(yTest, ypred)
          print(class_report)
```

	precision	recall	f1-score	support
Bacterial Pneumonia	0.75	0.72	0.73	502
Corona Virus Disease	0.99	0.93	0.96	508
Normal	0.91	0.94	0.92	503
Tuberculosis	0.97	0.99	0.98	509
Viral Pneumonia	0.69	0.72	0.70	502
accuracy			0.86	2524
macro avg	0.86	0.86	0.86	2524
weighted avg	0.86	0.86	0.86	2524

Fig -6: Classification Report

7. CONCLUSIONS

The project's commitment to adhering to healthcare regulations and data privacy standards underscores its dedication to maintaining the confidentiality and security of patient information throughout the diagnostic process. By prioritizing patient privacy, the project ensures that individuals can trust in the confidentiality of their medical data, fostering a sense of confidence and trust in the healthcare system.

Furthermore, the Detection and Classification of Lung Diseases for Pneumonia and COVID-19 represents a significant advancement in healthcare services. By harnessing the capabilities of machine learning and deep learning, the project empowers healthcare professionals with a powerful tool to aid in the early detection, differentiation, and classification of lung diseases. This technological advancement not only improves diagnostic accuracy but also enables timely intervention and treatment, ultimately leading to better patient outcomes and reduced mortality rates.

Moreover, the impact of this project extends far beyond the realm of medical research. By providing more efficient, accurate, and scalable solutions for healthcare systems worldwide, it contributes to the ongoing global efforts to combat respiratory diseases. In particular, amidst the ongoing COVID-19 pandemic, the project's ability to detect and differentiate COVID-19 from other lung conditions is invaluable in controlling the spread of the virus and managing patient care effectively.

As the field of artificial intelligence and healthcare continues to evolve, the potential for improving patient outcomes and public health remains substantial. By embracing innovative technologies and leveraging data-driven approaches, healthcare systems can enhance their capabilities in disease detection, diagnosis, and treatment, ultimately advancing the quality of care provided to patients worldwide.

8. REFERENCES

[1] Detection and classification of lung diseases for pneumonia and Covid-19 using machine and deep learning techniques Shimpy Goyal1 • Rajiv Singh1[2021]
[2] Detecting pulmonary diseases using deep features in X-ray images Pablo Vieiraa, Orrana Sousaa , Deborah Magalhães b , Ricardo Rabêloc , Romuere Silva[2021]
[3] Prediction of Lung Cancer using Ensemble Classifiers G Ashwin Shanbhag1 , K Anurag Prabhu2 , N V Subba Reddy, B Ashwath Rao[2021]
[4] Machine Learning for Medical Imaging Bradley J. Erickson, MD, PhD Panagiotis Korfiatis, PhD Zeynettin Akkus, PhD Timothy L. Kline, PhD[2017]
[5] A contemporary technique for lung disease prediction using deep learning Er. Rajeshwari P. Suryawanshi*1, Rashmi Tiwari*2, Tejal Bhoyar*3, Janhvi Kawale*4, Anagha Ambekar*5, Ankit Bhujade*6[2023]
[6] A survey on Image Data Augmentation for Deep Learning Connor Shorten* and Taghi M. Khoshgoftaar[2016]
[7] Using Machine Learning via Deep Learning Algorithms to Diagnose the Lung Disease Based on Chest Imaging: A Survey Shaymaa Taha Ahmed1() , Suhad Malallah Kadhem2 1 University of Technology, Baghdad, Iraq 2 University of Diyala, Diyala, Iraq[2021]
[8] Detection of Lung Diseases using Deep Learning Mrs. Rasika Naik Mr. Tejas Wani Ms. Sakshi Bajaj Mr. Shiva Ahir Mr. Atharva Joshi[2016]
[9] Hybrid deep learning for detecting lung diseases from X-ray images Subrato Bharati * , Prajoy Podder, M. Rubaiyat Hossain Mondal Institute of Information and Communication Technology, Bangladesh University of Engineering and Technology, Dhaka, 1205, Bangladesh[2020]
[10] lung disease detection using deep learning Siddhanth Tripathi, Sinchana Shetty, Somil Jain, Vanshika Sharma[2021]
[11]Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network Marios Anthimopoulos, Member, IEEE, Stergios Christodoulidis, Member, IEEE, Lukas Ebner, Andreas Christe, and Stavroula Mougiakakou*[2016]
[12]Lung Diseases Detection Using Various Deep Learning Algorithms M. Jasmine Pemeena Priyadarsini,1 Ketan kotecha , 2,3 G. K. Rajini,4 K. Hariharan,1 K. Utkarsh Raj,1 K. Bhargav Ram,1 V. Indragandhi , 4 V. Subramaniyaswamy , 5 and Sharnil Pandya[2023]