Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 12, 2025

DEEP LEARNING-ENABLED LUNG DISEASE CLASSIFICATION USING A RESOURCE-EFFICIENT VGG ARCHITECTURE

[1]Nasir Khan, [*2]Muhammad Sohail, [3]Muhammad Bilal, [4]Mobashir Husain

[1]*Department of Computing and Technology, Sarhad University, Peshawar, 25000, Khyber Pakhtunkhwa, Pakistan*

[*2]*Department of Computing and Technology, Sarhad University, Peshawar, 25000, Khyber Pakhtunkhwa, Pakistan*

[3]*Forest department Khyber Pakhtunkhwa, Pakistan*

[4]*Department of Computing and Technology, Sarhad University, Peshawar, 25000, Khyber Pakhtunkhwa, Pakistan*

[1]nasirkhan.csit@gmail.com, [*2]sohail.csit@suit.edu.pk [3]engr.bilal219@gmail.com, [4]mobashir70ustb@gmail.com

DOI: https://doi.org/10.5281/zenodo.21821075

Keywords:
Chest X-ray classification, convolutional neural network, deep learning, VGG-16, VGG-12, transfer learning, COVID-19, pneumonia, tuberculosis, medical image processing.

Article History
Received on: 14 Nov, 2025
Accepted on: 24 Dec, 2025
Published on 27 Dec, 2025

Corresponding Author:
Muhammad Sohail
sohail.csit@suit.edu.pk

Abstract

Deep learning has significantly advanced medical diagnostics by enabling accurate and automated detection of lung diseases. Its powerful feature extraction capabilities enhance early diagnosis, aiding in timely and effective treatment. This study focused on the identification of COVID-19, pneumonia, tuberculosis and normal cases through the utilization of an adopted lightweight model named VGG-12. This modified model based on convolutional neural networks provides a resource efficient and inexpensive solution for identification lung ailments at early stages. The model was trained on a large dataset of 14,000 images based on multi-class classification sourced from Kaggle. The adopted VGG-12 model obtained comparatively good results on the test dataset with 96.67% average accuracy with a reduced number of training parameters and training time. To improve the performance of the model and mitigate over fitting, advanced preprocessing and data augmentation techniques were used. The performance of the model was compared with multiple pre-trained networks like MobileNetV2, ResNet-50, and VGG-16. In comparison to these pre-trained models, the modified VGG-12 framework obtained comparatively good results in the classification of pneumonia, COVID-19, tuberculosis, and normal cases. This development will help medical professionals with a precise and quick diagnosis tool, contributing to timely and effective patient treatment. This work sets the stage for future advancements in automated medical diagnosis by demonstrating the potential of lightweight frameworks for efficient medical image processing, alongside the immediate benefits offered by the proposed model.

I. INTRODUCTION

Healthcare faces major challenges in the timely and accurate diagnosis of life-threatening diseases, especially respiratory illnesses such as pneumonia, COVID-19, and tuberculosis. Medical imaging, particularly chest X-rays, plays a vital role in detecting these conditions at an early stage. However, manual interpretation by radiologists is time-consuming, subjective, and often limited in resource-constrained regions. In the modern era of artificial intelligence (AI) based frameworks, the integration of machine learning (ML) and deep learning (DL) has become crucial in the field of medical diagnosis, with the aim of continuously enhancing efficiency. Automated medical diagnosis has improved tremendously, with a significant focus on achieving desired levels of precision. X-ray pictures can now be used to promptly identify hazardous disorders, including pneumonia, COVID-19, tuberculosis, and other chest infections. The use of computer-based picture analysis not only demonstrates the diagnostic procedure, but it also shows to be an affordable alternative [1].

Pneumonia, a lung infection caused by viruses, bacteria, or fungi, is a significant cause of death worldwide. It is especially dangerous for children and has an elevated mortality rate, with 1.2 million children under the age of five infected in 2016 and 880,000 killed [1]. Pneumonia causes swelling inside the air passages of lungs, which prevents normal breathing. The SARS-CoV-2 virus, which caused the COVID-19 declared as a global pandemic, is now regarded as among the most serious health crises in recent memory. Daily fatalities due to its pervasive influence have made it difficult for testing facilities to provide sufficient medical care. The quick spread of the infection, combined with the worrisome volume of newly identified infections, has reduced the time accessible for providing not just well-known but also critical medications. COVID-19 alone has claimed 4.45 million lives [2].

Every year, a large number of people get TB, with a high rate of death. The yearly estimate of TB infections is 8 million, with two to three million fatalities [3], [4]. Mycobacterium The disease-causing bacterium is TB, which typically affects the lungs but can also affect the kidneys, spine and brain. TB has two manifestations: TB illness and underlying infection with TB. TB is harmful when ignored because it spreads via the air via breath or saliva. After being eaten, viruses dwell in the lungs before travelling to the coronary arteries and kidneys via the circulation of blood [5]. Fig. 1 shows automated detection of chest disease using CNN.

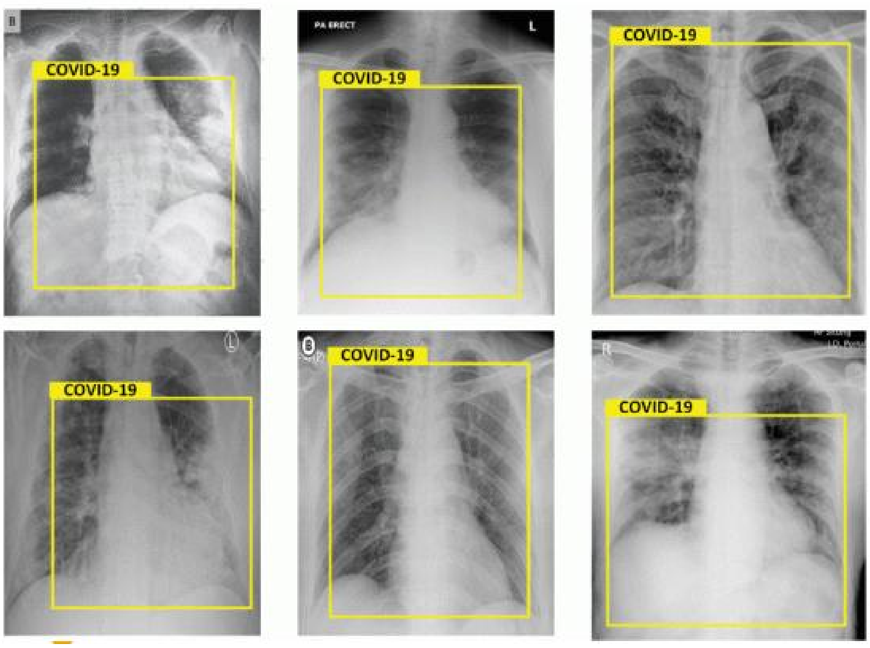

Figure 1. Automated detection of chest diseases.

Detecting disease signs early and initiating treatment on time might minimize death rates caused by chest disorders such as TB, pneumonia, and covid-19. CXR pictures are routinely used to identify signs of chest illnesses. The primary causes of disease detection delays are naked-eye processes for pulmonary illness detection and a lack of area experts [1]. According to the numbers provided above, an alternate method is required to accelerate the timeliness of disease identification and thereby lower the overall effect of lung diseases on the community [2].

CXR (CT-scan and X-ray) pictures are commonly used to detect early indications of lung illnesses such as tuberculosis, COVID-19, and pneumonia [2]. Medical professionals typically utilize a simple visual approach to detect lung problems, which has various drawbacks, including being time-consuming, error-prone, and subjective, especially when a healthcare official lacks expertise. To address these issues, DL is an alternate method for detecting early indications of chest infections. Deep learning frameworks are well-known for their precision and generalization in comparison to the naked eye and traditional approaches, but they require a large number of training images. For example, a CNN model for a certain domain requires several thousand photos as training data. DL algorithms have the potential to improve the accuracy of healthcare diagnostics through automated decision-making [3], [5].

DL is a type of AI that uses multiple hidden layers based on deep neural networks to learn from the training dataset and then makes autonomous decision in the future using this trained models based on supervised learning. It is utilized in every field of computer science, including the identification of voice tune, interpretation of images, natural language processing (NLP), and data based on time series. They have the capability to extract features from input data and use them to make autonomous judgments. Deep learning algorithms replicate the power of the human brains to examine and grab information from training data, which makes them useful for processing complicated and unorganized raw data and demanding less preparation approaches as in contrast to conventional and traditional procedures [6].

CNNs are well-known deep learning networks that are commonly used to interpret picture data. CNNs' capacity to extract complicated features from pictures makes them useful in the interpretation of healthcare images. CNNs can recognize even small alterations in the source data more accurately than peoples. The invention of models based on CNNs have improved both the precision and speed of ailment detection, and experts are still working to introduce new mechanisms into present CNN structures to revolutionize medicine and lessen the burden on medical experts. Recent advances in CNNs, such as networks called U-Net transformers and GANs, have established a solid foundation for automated diagnosis to improve clinical care and treatment. CNNs are made up of two components: the backbone, which extracts relevant information from videos and images, and the head, which makes final decisions based on the knowledge gathered from the data set used for training. CNNs are also utilized for NLP-based tasks such as sentiment assessment, time series data processing, and disease prediction using patient demographics [4]. Fig. 2 shows the general steps to build AI model for automated disease identification.

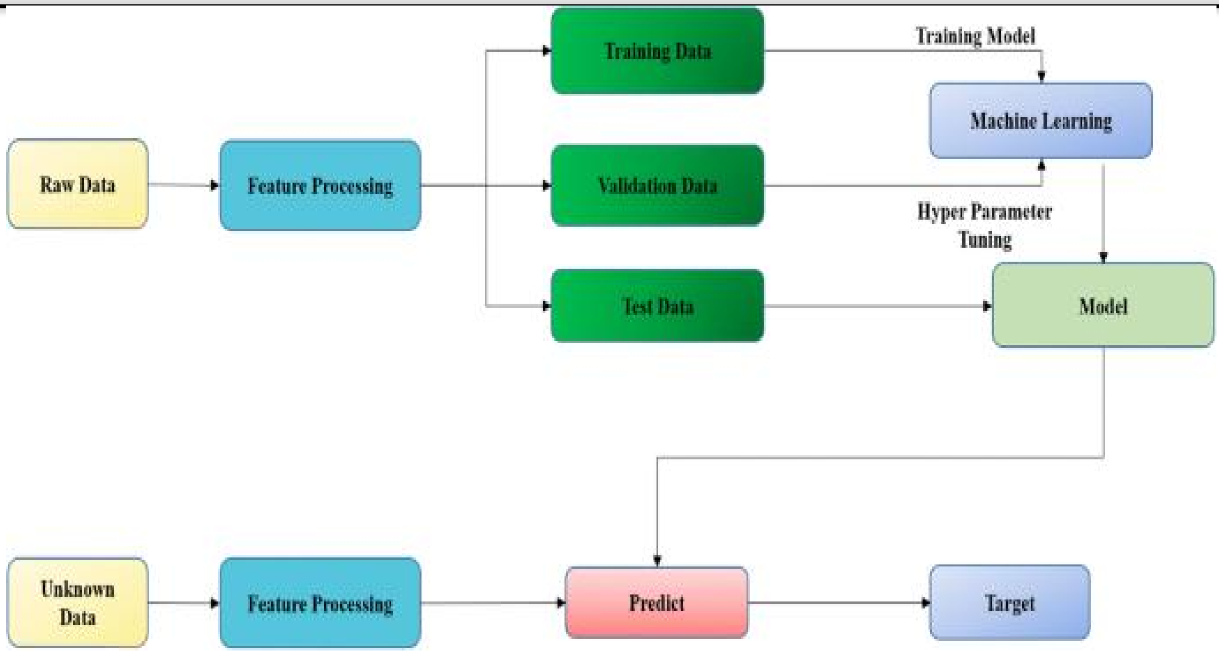

Figure 2. General steps to build an AI model for disease diagnosis.

The detection of chest-related illness signs is primarily dependent on CT scans and X-ray pictures, which present various obstacles utilizing naked-eye methods, such as inaccurate interpretation, subjectivity, and overlooking tiny alterations, as well as increasing diagnosis time [7]. Deep learning-based evaluation is a substitute that is more accurate, free of subjectivity, capable of detecting minor differences, and produces fast findings. Deep learning algorithms, especially CNNs, have dramatically improved accuracy and generalization, leading to more effective patient care and treatment [8]. Incorporating CNN-based algorithms into healthcare imaging is a significant advancement that opens up new possibilities for improved precision, efficiency, and early detection of illness symptoms, such as pneumonia, tuberculosis, COVID-19, and lung carcinoma.

The advancement in healthcare imaging has proven to be a positive step towards correct diagnosis. Yet, differentiating between the early signs of various chest ailments can be difficult, requiring technical knowledge and expertise in the sector. Deep learning, particularly CNNs, has demonstrated significant strength in mitigating these issues by identifying unique signs of chest illnesses and offering automatic evaluation of images for better results. CNN models continue to confront challenges such as inconsistency in data used for training and questions regarding the utility of trained models. However, new advances in medical imaging may have long-term benefits. These models are expected to improve and grow smart tools for physicians, assisting in precise diagnoses and ultimately improving patient care and treatment in lung disease, offering precise and swift treatment outcomes, and increasing the overall efficacy of chest disease recognition [9].

The inspiration for this work stems from recent advances in smart diagnostics. DL-based systems are currently being used in every aspect of the health sector, including automated diagnosis using medical imaging and patients' demographic information. Earlier computation techniques employing traditional and machine learning methods were less precise and generalizable. Current models based on deep learning have renewed the medical field, while scholars are continually working to enhance these approaches. The proposed effort also aims to develop a novel CNN-based model for automatically predicting early signs of lung abnormalities such as tuberculosis, COVID-19, and pneumonia. The study's goal is to create a CNN model with fewer training parameters that may be deployed on resource-constrained platforms while maintaining accuracy and generalization.

DL is playing an important role in transforming healthcare, including automatic identification of diseases from medical images. Among the famous tools in healthcare, medical image processing stands out as a useful area where AI researchers work continuously to build models that assist healthcare professionals in making rapid and accurate diagnoses. The integration of AI and clinical images has considerable potential, but further efforts are needed to enhance the performance of these approaches.

This study aims to improve the performance of VGG-Net for small datasets, specifically focusing on the common lung diseases. Its aim is to focus on reducing the parameters while improving the accuracy regarding CXR images like TB, pneumonia, Covid-19, and normal cases. The main aim of the study is to develop a resource-efficient CNN approach that is prepared for effective training, inference and deployment, improving resources utilization with the goal of enhancing medical diagnosis capabilities in healthcare.

The objectives of this study are as follows:

- We aim to build a novel DL-based approach to classify lung diseases efficiently.
- To gather a comprehensive set of CXR images, including COVID-19, normal cases, and pneumonia, followed by image labeling/annotation, preprocessing and data augmentation approaches.
- To modify and fine-tune the structure of VGG-16 to successfully identify lung diseases with high accuracy and generalization.
- To compare the proposed approach against previous state-of-the-art (SOTA) approaches.

The aim of this study is to develop a resource-efficient model utilizing CNN for detection of early signs of lung diseases from CT-Scan and X-ray images. After a lot of experiments where different design features were changed, the study created a new and strong DL system designed to diagnose several lung diseases using CXR images. The proposed VGG-12 framework is a modified version of the famous pre-trained network called VGG-16. VGG-12 has 10 convolutional layers, removing the final group of 3 convolutional layers of the original VGG-16 model. The proposed model has two FC layers with 1024 and 512 neurons, respectively. After each convolutional block, batch normalization (BN) is used, followed by dropout layers with a dropout rate of 30%. Due to the multi-disease classification environment, softmax is used as the final layer. The model was trained using 14000 pictures from four distinct groups (TB, COVID-19, pneumonia, and normal). Advanced preprocessing and image augmentation strategies are used to improve model training and reduce over fitting.

The main contributions of this study are described below:

- A large dataset of training images was collected from multiple publically available datasets.
- A lightweight version of VGG-16 model named VGG-12 has been built to identify multiple disease types of lungs using CXR photos.
- Using several kinds of preprocessing approaches, as well as offline and online image augmentation, to improve model performance and avoid bias towards the majority classes.
- To evaluate the performance of the modified VGG-12 network, the results were compared with state-of-the-art (SOTA) models focusing on similar lung ailments.

The proposed study tried to develop a lightweight and accurate CNN model for early detection of lung diseases such as COVID-19, pneumonia, and TB using CXR images. It focused on improving diagnostic accuracy in low-resource healthcare settings by proposing a resource-efficient approach. The study included the collection and use of publicly available datasets, targeting early and accurate identification of lung diseases. It aims to support automated diagnostic workflows in clinical environments to enhance the diagnostics process with accurate and rapid results. The research contributes to ongoing efforts in AI-driven medical imaging for timely disease detection and healthcare enhancement.

The subsequent sections of the dissertation are structured as follows: The 2nd chapter contains an in-depth assessment of previous latest works, exhibiting the crucial role of DL in healthcare imaging, which helps to make advancements in this sector. The third chapter contained the workflow of the proposed study, providing readers with a comprehensive understanding of its plan of action. It includes a comprehensive description of all necessary steps of the proposed approach, including image preprocessing, data augmentation, and system architecture. Chapter 4 presents the study's findings, as well as a comparative analysis that includes facts, figures, and comparisons. The thesis winds up with the final section, which provides final reflections and suggests new topics for further research on this unusual subject.

II. LITERATURE REVIEW

The rapid expansion of computational biology has positioned healthcare and agriculture image analysis at the forefront of innovation, driving a concerted drive to automate these disciplines. Computational biology has risen in prominence as AI, machine learning, and CNNs have emerged, particularly in medical picture categorization. Recent breakthroughs in this discipline include the use of transfer learning, ensemble learning, and other approaches. Such approaches increase diagnosis accuracy while optimizing healthcare processes on a daily basis. As edge computing advances real-time diagnosis, AI collaborates with competent medical practitioners to build a future in which technology is the single most effective tool for addressing the issues of human health. This section discusses the most recent advances in image categorization, focusing on major medical imaging projects in general as well as those focused particularly on lung disease diagnosis.

Image recognition is a well-known machine vision and digital imaging technology that has a wide range of applications, including forensics and safety, agriculture, medicine, monitoring, and transport. CNN models based on DL have significantly accelerated training and allowed models to successfully learn complicated patterns and attributes from input images. Following very accurate experimental findings and forecasts, these models can be simply deployed across a variety of systems (desktop, drone cameras, web-based, edge devices, mobile, etc.). Furthermore, with ongoing model improvements and multidisciplinary collaboration, they may be utilized more responsibly across a wide range of sectors to solve numerous issues and improve many parts of our lives [10]. Ullah N et al. [11] developed DeepPlantNet, an innovative deep learning-based technique for identifying plant illnesses. DeepPlantNet consisted of 28 layers, comprising 25 convolution layers and 3 FC layers. The model utilized Leaky ReLU, BN, fire modules, and a combination of 5×5 and 1×1 kernels in each layer of the convolution block. This network categorized plant diseases into ten, eight, and three classes, respectively. It attained an average accuracy of 98.49% and 99.85%, respectively. This method overcomes the constraints of its conventional alternatives, like shallow CNN models, by offering an automated approach for illness identification issues and recognizing plant illnesses prior to their spread, thereby improving agricultural production and preventing substantial losses in yield because of multiple kinds of pathogens.

Shekar et al. [12] employed CNN combined SVM as the last layer to recognize signatures on electronically stored information such as pictures and PDF documents. Zhu et al. [13] suggested a model for segmenting food items in useful manners from visual inputs, addressing the challenge of getting pixel-level annotation. They suggested a single-component classification model that was trained on a recently constructed picture dataset with only one class per photo and a standardized biological-based nested material layout. The feature representations of this algorithm were employed in a multi-ingredient categorization system yielding fairly acceptable outcomes on the FoodSeg103 data. The suggested CNN design, named AttNet, consisted of convolution layers started with an activation function called Relu, which was followed by a max pool layer to obtain characteristic maps from image inputs, FC layers for identification, and the last layer for output creation. The suggested architecture was evaluated against SOTA already trained networks, including ResNet50 and EfficientNetB1, to demonstrate its robustness and uniqueness. The paper claimed to have contributed by offering a thorough approach and outcomes for substance identification that did not depend on pixel-level labeling.

Akash et al. [14] employed the YOLOv5 and Inception-v3 networks to identify aggression on the streets and track irregularities in public settings. Both of these models received training using transfer learning. Hassan et al. [15] presented a unique deeper CNN approach for early identification of plant illnesses using leaf pictures. They used depth-wise separable convolutions in the Inception model and InceptionResNetV3 to reduce parameter intensity while maintaining the precision of identification. They utilized Mobile-Net and EfficientNetB2 as pre-trained networks, achieving accurate classification rates of 98.97% and 99.58%, respectively, exceeding typical machine-learning methods. CNN algorithms built using transfer learning that incorporate regularization strategies (dropout layers and BN) have demonstrated relatively acceptable precision as well as training time. The algorithms' integration with smartphones and tablets, notably MobileNetV2, demonstrates the opportunity for real-time use in agriculture, which has accelerated crop illness diagnosis.

The work in [16] offered an AI-based approach to address the considerable impact of illnesses on rice output, emphasizing the importance of early disease detection to improve productivity. They designed and used a CNN algorithm using InceptionResNetV3 weights that were already trained, known as transfer learning. The researchers claimed an impressive 95.67% precision for determining rice ailments of the leaves. The dataset included 5200 photos from 4 classes, such as leaf blasts, brown spots, bacterial blights, and normal leaf images. They employed image pre-processing, which included rescaling and resizing as well as augmentation procedures such as rotation and flipping, to expand the dataset. The InceptionResNetV3 algorithm has been fine-tuned with hyper parameters, revealing the possibility of automatic identification of illnesses in rice crops, which farmers may employ to increase precision farming.

As image identification improves quickly, automatic diagnosis in healthcare imaging is going to grow more effective [16]. Additionally, it is easily accessible due to its easy-to-use architecture, which significantly improves medical diagnostics for better results. Z. Wu et al. [17] explored the efficiency of CNNs, including VGG-16, ResNet50, Inception-V3, EfficientNet-B0, DenseNet, Xception, and Inception-ResNet, for the classification of face skin diseases using medical images. The study used China's largest clinical image archive to construct a dataset of 2656 face photographs encompassing six common skin diseases. Following careful analysis, the frameworks produced promising results, particularly when TL was applied. In the test set of 388 facial pictures, the best model showed the efficacy of the proposed approach by obtaining impressive recall rates for specific conditions. A new CNN-based technique named Tumour-ResNet was employed in this research [18] to detect brain tumors from MRI images. Three FC layers were placed after 20 convolutional layers in the framework. The project aimed to develop an inexpensive CNN algorithm that could identify brain tumors with good generalization and precision. The research's detection rate was 99.33%.

A skin disease classification approach using DL that combines Long Short-Term Memory (LSTM) and MobileNet was proposed by Srinivasu et al. [19]. MobileNetV2 handles portable gadgets like cellphones well, whereas LSTM enhances the performance of the model. The model was fine-tuned and trained on a public dataset and has an accuracy of about 85%. Quick diagnosis of diseases was made possible by the algorithm's smartphone app, which helped both physicians and patients. By using texture-based data evaluation, the Co-occurrence Matrix with Grey Level enhanced the evaluation of skin conditions. In [20], the authors proposed an enhanced CNN model and assessed the use of deep learning for breast tumor detection. This new method simplifies the laborious process by using a CNN for processing instead of distinct extraction of features and compression steps as in the earlier methods. An artificial fish school approach is used in the suggested optimization method. It uses a dataset of breast tumor images and preprocessing methods like Wiener filtering for classification. Through effective training and evaluation, the integrated CNN surpasses previous methods in terms of sensitivity, precision, specificity, F1 measure, and recall.

In order to identify breast tumors, Ullah et al. [21] developed and trained an innovative CNN approach using VGGNet. Because of the deep, layered design, the VGG-16 exhibits over fitting on small and moderate datasets and necessitates a huge collection of training photos. In order to solve this problem, the authors reduced the VGG-16's layer depth and created a simpler version called VGG-12, which has 12 layers. The performance of the VGG-12 network was assessed using fresh breast tumor test images. The adopted VGG-12 method outperformed other SOTA methods, according to the authors, with an identification rate of 95.6%.

According to Nigat et al. [22], CNN algorithms can effectively classify skin conditions caused by fungi. They utilized a training set of 407 pictures and improved the training of the model through the use of preprocessing strategies such as image standardization and the process of augmentation. The suggested sequential CNN framework outperformed MobileNet and ResNet in terms of accuracy, achieving 93.3%. The study highlighted the possible application of CNN algorithms in automating dermatology screening based on pictures, especially in environments with limited assets. A deep learning network was constructed by Eweje et al. [23] using 1,065 MRI images of bone lesions. The model, which was based on the EfficientNetB1 pre-trained network, used both patient demographic information and picture-based features to identify the patient's abnormality severity. This hybrid approach achieved high accuracy when compared to professional radiologists. An improved AUC ROC of 0.80 from evaluation on the test data demonstrated the network's accuracy in differentiating between malignant and benign bone disorders. The method is a useful tool for evaluations at local hospitals and has the potential to reduce needless admissions and investigations. Sampath et al. [24] employed a DL-based model followed by data augmentation and image processing to distinguish images that were healthy and those that were afflicted by tumors. A pre-trained network, AlexNet, was fine-tuned using transfer learning, which worked well, with testing accuracies of 98%. The researchers used a dataset of 1140 CT scan images to train the model. Before breaking down the images, the researchers changed them to greyscale and used median filtering to clean them up, then applied smart edge detection and K-means clustering. According to the results of the assessment, Alex-Net is preferable since it offers accurate identification with a high chance of a timely tumors diagnosis and a shorter processing period.

Computer vision innovations have permeated the medical domain, opening the door to autonomous diagnosis of conditions such as pneumonia, COVID-19 and TB, a lung-related illness that requires algorithms like CNN to readily analyze pictures from X-ray and CT scans. In addition to offering quick and precise recognition, this integration expedites the diagnosis procedure, showcasing the ability of the technology to enhance patient care and healthcare imaging.

Mahbub et al. [2] suggested an inexpensive CNN model to examine CXRs for lung disorders linked to viral illnesses including Covid-19, pneumonia, and tuberculosis. The network featured 6 convolution layers subsequent to a dropout layer, a dense layer followed by a dropout layer, and a final layer based on an activation function called 'sigmoid' used for binary classification. Three chest diseases such as TB, COVID-19, and pneumonia were used as training data for the model. With a 99.63% accuracy rate, the approach was additionally optimized for binary identification using training datasets of COVID-19 and healthy pictures. Additionally, they reported that their multi-classification examination of test pictures had identification rates of 99.55%, 99.76% and 99.87% for TB, COVID-19, and pneumonia, respectively. An improved hybrid method based on conventional feature design with CNNs was proposed in the article in [25] for using CXR pictures to detect 10 pulmonary disorders, such as COVID-19. The framework includes a workflow with standard point extraction strategies (SURF and ORB), Deep CNN with transfer learning (VGG-19), and ML models for illness recognition, as well as Info-MGAN for segmenting lung photographs. The method supported automatic diagnosis of pulmonary conditions, such as COVID-19, with a respectable identification rate of 97.98%. The training set included several datasets related to lung diseases, each with preprocessing tools to improve the photos used for training and data augmentation tools to lessen the likelihood of over fitting. The research provides guidelines for the suggested technique, image sets, and outcomes that will be helpful to healthcare professionals in the present. Researchers in the future may find the thorough overview useful in training CNN algorithms for a variety of purposes.

In order to automatically detect illnesses such as COVID-19, pneumonia, heart disease, lung opaqueness, and thoracic problems from digital CXR images, Nahiduzzaman et al. [1] built an inexpensive CNN algorithm named ChestX-ray6. The algorithm achieved an accuracy of 80% in identifying these illnesses after its training on a collection of 9,514 pictures. The study demonstrated the growing significance of machine learning for healthcare assessment, namely in assisting radiologists with X-ray image analysis. When it came to distinguishing between normal individuals and pneumonia situations, ChestX-ray6 worked better than SOTA designs, with an accurate diagnosis percentage of 97.94%. The authors claim to have combined many datasets, proposed an inexpensive CNN framework, improved efficiency through image augmentation, and compared the results using various TL approaches. In order to improve the CNN framework's stability and accuracy in its recognition of healthcare photos, Aytaç et al. [26] suggested a novel CNN approach with an adjustable rate of learning. No parametric tuning was required because the acquisition pace was in harmony with the corrections performed using the margin of error for each epoch. Three distinct medical picture collections, such as NIH-Chest X-ray, COVID-19 CT-Scan, and REMBRANDT-Brain Tumor, were used to empirically determine the methodology. For a variety of medical picture classification applications, the study offered an enhanced reverse propagation technique.

Using CXR images, Alshmrani et al. [27] proposed a DL model for multi-class diagnosis of common chest illnesses, such as COVID-19, TB, lung cancer, pneumonia, and lung opacity. Three convolutional layers for obtaining data and 3 dense layers for identification are included in the framework after a previously trained VGG-19 model. Upon being trained and validated on an image set of 80,000 lung images, the proposed model outperformed earlier studies with an accuracy of 96.48% and an AUC of 99.82%. The study highlights the possible importance of DL in the accurate and efficient detection of lung problems in light of COVID-19. In order to analyze B-lines in pulmonary images, the researchers in [28] developed an algorithm known as CNN using ultrasound pictures. The method has demonstrated strong agreement with skilled readers for the existence or absence (kappa 0.87) and shown exceptional performance for identifying B-lines (93% efficiency). The approach shows potential for lowering variation and enhancing B-line evaluation, which could result in better diagnosis and illness evaluation, even though differentiating intensity became more difficult (kappa 0.65).

The application of DL approaches for CT scan-based TB detection was covered in the study in [29]. The focus was on categorizing TB types and identifying resistance to multiple drugs. Picture structures, mask structures, and age/gender data were among the multi-channel sources employed by the networks. CT scan characteristics were extracted using the pre-trained networks such as ResNet and VGG-19. While the algorithm for TB category identification scored 53% efficiency and a Kappa score of 34%, the best-performing strategy for resistance to multiple drugs achieved 74.13% efficiency and a total of 64% AUC. The study highlights the importance of prompt identification for appropriate therapy and the possibility of DL to improve the diagnosis process of diseases, including TB. In order to identify multiclass lung abnormalities using CXR pictures, Sharmat et al. [30] proposed an algorithm based on transfer learning and CNN that targets diseases such as pneumonia, COVID-19, and TB. The pre-processed, class-balanced, and augmented training image collection was gathered from multiple web collections. The model outperformed other previously trained models with an accuracy rate of 98.89%. Ens-Net was used to remove image annotations, and an auto encoder was used to de-noise the images. Using a large image set of 85,105 images of the chest X-ray, the research illustrated the outcome of the proposed approach using a number of analyses, such as ROC plots and AUC figures.

Due to a shortage of successful antifungal therapies, Duong et al. [31] addressed the critical need for prompt identification and surveillance of COVID-19. By using state-of-the-art machine learning techniques, such as MixNet and EfficientNet to CXR imaging and lung CT pictures, the researchers offered a workable approach. Training effectiveness was improved by three TL methods. In 5-fold validation on 4 datasets from the real world, the suggested approach outperformed previous research with an accuracy of over 95%. The study presented a smartphone app design for real-world application and underlined the importance of precise diagnosis in the fight against the pandemic. Neural network performance, TL techniques, identification speed, and matching to predetermined previous levels were the main areas of assessment. Seelwan et al. [32] used a particular CXR image set to investigate the generalizability of a deep CNN approach created for TB diagnosis. The model was trained based on binary classification of TB and healthy images. When developing a framework, it is crucial to take into account technological requirements, severity of disease distribution, and data features. The proposed CNN model demonstrated excellent precision in the Shenzhen image set, but its performance drastically declined in the ChestX-ray8 data. The research emphasized the demand for careful analysis prior to practical use and the difficulties in applying machine learning (ML) techniques for TB detection across varied settings.

According to the literature mentioned above, CNNs have the potential to be a useful tool for the early diagnosis and detection of lung conditions. However, problems still exist, including restrictions on the quantity and quality of lung-related image collections that are now accessible. A further obstacle is the computational expense of CNN deployment and training. Notwithstanding these challenges, CNNs exhibit promise in the fields of medical imaging and early disease symptom detection. To increase the precision and effectiveness of CNNs for healthcare image categorization, continuous studies are essential. Table 1 shows some previous work based on identifying lung diseases leveraging deep learning.

Table 1: ***Recent Work On Automated Identification Of Lung Diseases***

Ref#	Year	Authors	Method	Dataset (Classes)	Accuracy
[33]	2025	Alqaissi, et al.	Lightweight CNN	COVID-19, Pneumonia, TB	96.53%
[34]	2025	Saddique, et al.	Transfer Learning (EfficientNet + Transformer)	Normal, Pneumonia, COVID-19	95.50%
[35]	2023	Ullah, N. et al.	Deep CNN (20 layers)	TB, COVID-19, Lung opacity, Normal, BP, VP	95.57%
[36]	2023	Sharma S. et al.	Transfer Learning (VGG-16)	Pneumonia,	94.54%
[37]	2022	Jain D. K., et al.	YOLO-V3	COVID-19	92.15%
[38]	2024	Asnake N. W. et al.	Transfer Learning (Xception)	Pneumonia	96.00%
[39]	2023	Goyal S., et al.	CNN-RNN	Kaggle's Chest X-ray Dataset	95.04%
[40]	2022	Taher F. et al.	CNN	Kaggle's Chest X-ray Dataset	94.00%
[41]	2024	Bhosale R. D., et al.	Deep CNN	Normal, Pneumonia, COVID-19	96.00%
[42]	2024	Ko J. et al.	Vision-Transformer	COVID-19, Pneumonia, TB	95.87%

III. RESEARCH METHODOLOGY

The goal of this effort was to develop a resource efficient CNN model that could accurately and broadly identify early indicators of lung illnesses from CXR scans. Therefore, an extensive hybrid collection of CXR pictures was used to train an innovative CNN approach called VGG-12. The training set, picture processing, image augmentation, structure of VGG-16, and the improved version of VGG-16 known as modified

VGG-12 were all covered thoroughly in this part. The main procedure for the proposed work is depicted in Fig. 3.

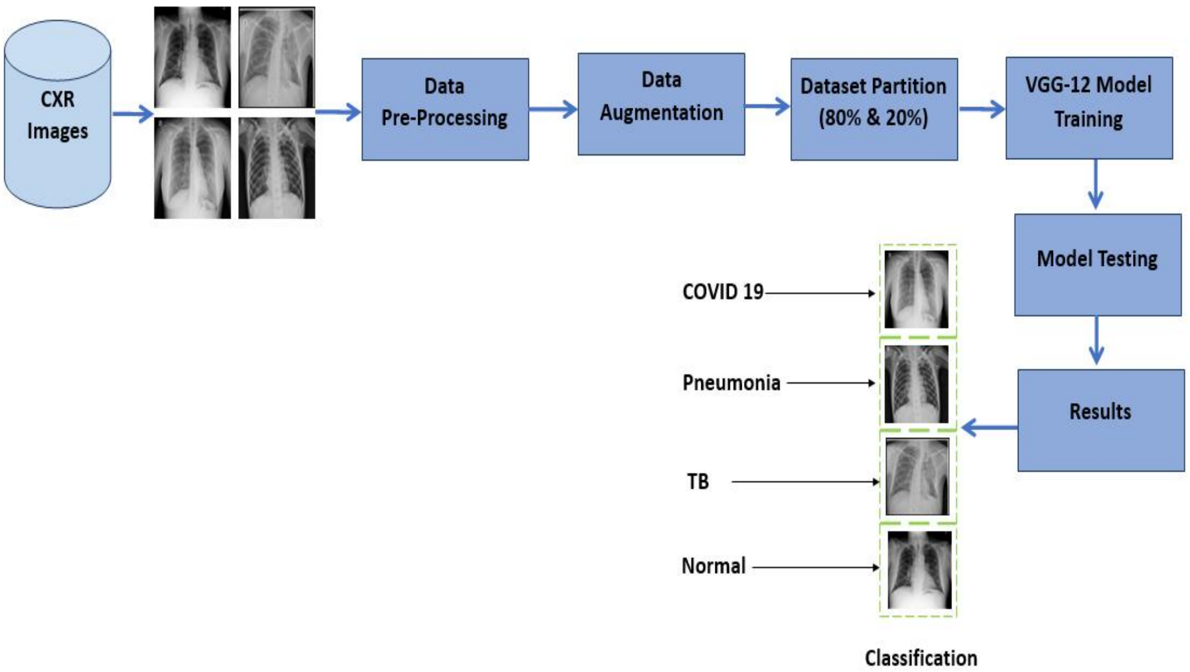

Figure 3. General Flow Of The Proposed Method.

A. Training Dataset

In the field of supervised learning, data is an essential and fundamental element. The well-known proverb "garbage in, garbage out" emphasizes how crucially important the standard of input data is in influencing how computing operations turn out. This is particularly relevant in the realm of deep learning, where a thorough grasp of the specific domain tasks and a significant quantity of outstanding data are essential for successfully training networks. The combination of deep learning, transfer learning, and analysis of medical images has received a lot of attention currently. Our focus in the current research is multiclass classification, and CNN is the tool we have selected to do this. Because CNNs are capable of acquiring structural characteristics from inputs, they are especially appropriate for tasks involving images. We obtained a substantial set of images from Kaggle, a repository recognized for its extensive and broad data, to aid in our research. CXR scans from various Kaggle locations, containing normal instances, COVID-19, TB and pneumonia, are included in the training data. The dataset employed in this research consists of 1419 photographs for TB and 4000 images for pneumonia, COVID-19 and the normal category. Particularly, offline augmentation was applied to the TB photographs, expanding the number of photos to 2000 pictures. The training data set was divided into the ratio of 0.80 and 0.20 for training and validation, respectively. Another set of photographs, referred to as the testing dataset, containing 200 photographs per category, was employed to assess the model's efficiency on new training images. Fig. 4 shows sample images from training dataset.

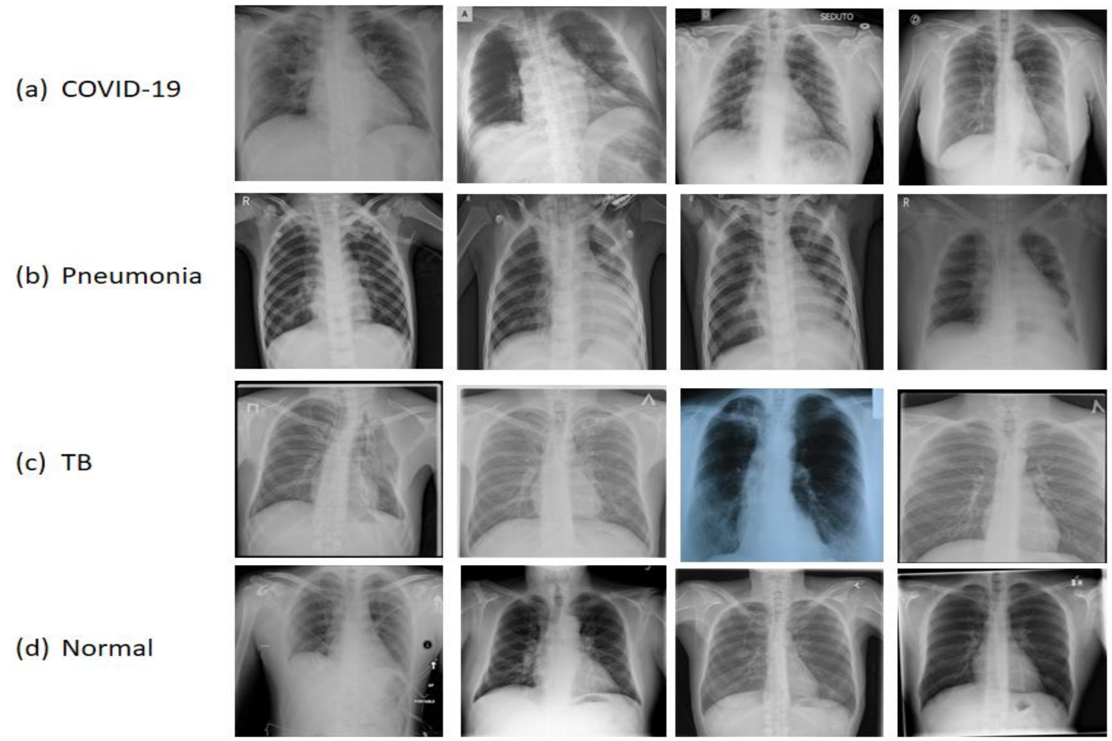

Figure 4. Samples From The Training Set.

b. Pre-Processing

One crucial and typical stage in CNN model development is the preprocessing strategies used for image enhancement. A comprehensive hybrid training dataset was collected for this investigation from multiple Kaggle sources. In the presence of healthcare professionals, the photos were manually reviewed, and any that were deemed insufficient were eliminated from the training data. During training, CNN simulates consistency in the input photos it receives. As a result, every image was resized to 256 x 256 pixels. Rescaling or normalization is another crucial image processing step that experts perform during training CNN algorithms. Additionally, it is a crucial phase that promotes the convergence of models. All of the supplied images were transformed to the 0–1 scale in the proposed method. We developed the framework with improved accuracy in validation and training, and extremely low over fitting was observed during training, exhibiting improved generalization of the model.

C. Image Augmentation

Image augmentation is a common technique used to train deep learning models with improved generalization and reduced over fitting. Image augmentation serves two functions in CNN model training: it increases the quantity of the training dataset and adds changes to the training image set to guarantee the algorithm's generalization in real-world situations. Both online and offline augmentation techniques were employed in the suggested investigation. After increasing the number of TB photos from 1419 to 2000 using offline augmentation techniques, the entire training data was subjected to online augmentation. Rotation, shear, flipping, brightness, and zooming are some of the famous methods employed for data augmentation. Fig. 5 shows several strategies used in this research's augmentation procedure. The stability of our suggested CNN framework in the categorization of CXR photos is strengthened by this multidimensional augmentation method, which also improves its ability to generalize to fresh, unseen images and strengthens its ability to resist over fitting.

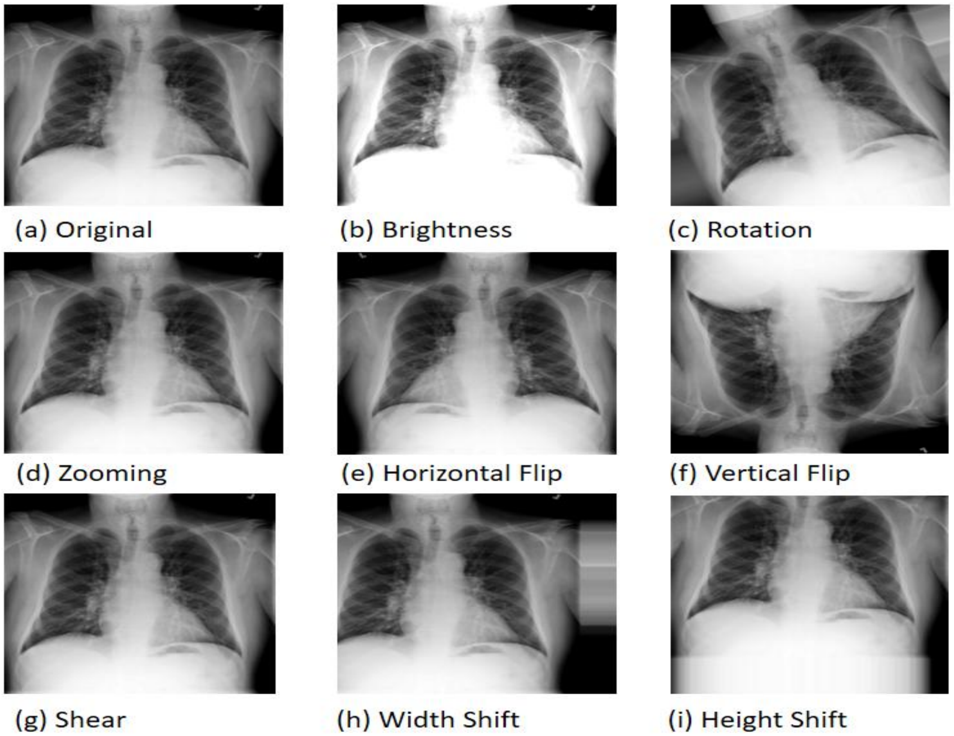

Figure 5. Data Augmentation Methods Utilized In This Work.

D. VGG-16

VGG-16, a groundbreaking CNN pre-trained model containing sixteen layers structure, Thirteen convolution layers (3 x 3 kernel size), and Three dense layers, was developed by the "Visual Geometry Group (VGG) at the University of Oxford" [43]. Five blocks of convolution layers make up the structure of the VGG net, where the first block's first two layers have 64 maps of features, while the second block's two layers have 128 maps of features. The third block contained three layers, each of which has 256 feature maps, and the final two blocks, each of which has three layers with 512 feature maps per layer. Each convolutional block is followed by a max-pool layer. The ReLU activation function is used by both convolutional and dense layers to add non-linearity to the model. VGG-16 is a standard for computer vision because of the perfect balance between width and depth. It is a component of a number of major initiatives in the fields of object identification and classification of images. As a pre-trained model known as transfer learning, it is also frequently utilized. Fig. 6 shows the general structure of VGG-16.

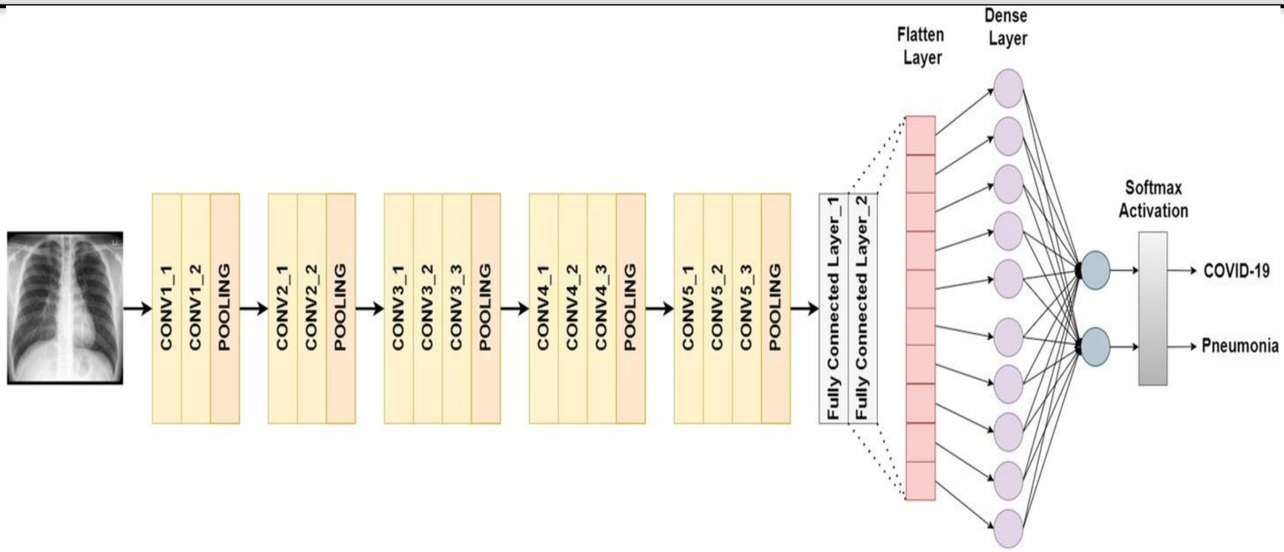

Figure 6. Architecture Of Vgg-16.

E. Proposed Model

The suggested new modified VGG-12 architecture is a modified version of VGG16 and having reduced number of training parameters compared to VGG-16 framework. To accommodate the suggested dataset, a number of layers were eliminated from the previous VGG-16 design in this research. This version of the VGG-16 design is inexpensive and resource-efficient, which can be adopted on small training datasets. After eliminating the last block of 3 convolution layers seen in the VGG-16 approach, it has 10 convolution layers overall. Two dense layers with neuron counts of 512 and 256, respectively, were used in the design. The output layer contained softmax activation mapping for multiclass classification with 4 neurons representing the different kinds of lung diseases. When it came to identifying the three types of lung illnesses, the model performed rather well. Fig. 7 illustrates the modified VGG-12 structure. Both the convolutional and dense layers make use of ReLU as an activation function. To avoid over fitting and improve the VGG-12 algorithm's resilience and generalization, a max pool is applied after every convolutional block, and dropout with a 30% dropout ratio is used after every dense layer.

CNN frameworks are built around convolution layers. Convolution procedures are mostly used to extract low-level features from images, such as borders, contours, lines, and object boundaries. The definition of the convolutional process in mathematics is:

$$S(i,j) = (I * K)(i,j) = \Sigma\Sigma\, I(m,n) \cdot K(i-m,\, j-n) \quad (1)$$

Where S(i,j) denotes the element at position (i, j) in the output map of features, I depicts the input picture matrix, K signifies the kernel or filter matrix while "m" and "n" are indices used for summation. The input picture is first seen as a numeric array in convolution processing, with every component representing a pixel amplitude level. The picture matrix is subjected to a methodical application of a collection of kernels, sometimes referred to as 'feature detectors'. Such filters usually have tiny sizes, like three-by-three or five-by-five, and have obtainable weights. Each of the filters in the suggested approach was three by three in dimension. Each of the filters is moved over the picture as part of the convolutional procedure, which then multiplies the filter by the correct picture area element-by-element and adds up the outcomes. Each kernel's convolutional results in a unique feature pattern that displays the filter's activations throughout the image. Feature maps highlight areas where the kernel or filter has identified pertinent features.

ReLU was utilized in both the dense and convolution layers of the suggested framework. In artificial neural networks, ReLU is a frequently used activation function. By adding non-linear behavior, it enables the algorithm to learn complex information and improve its decision-making. It functions as:

$$\text{ReLU}(x) = \max(0, x) \quad (2)$$

It requires only one input value, usually a weighted total of a neuron's input data. ReLU essentially "kills" or deactivates the neuron by simply producing the data value itself if it is either zero or positive, and 0 if it is negative. Compared to other functions for activation like tanh and sigmoid, it is less computationally expensive due to its simple computation. Pooling layers, like average pooling or max pooling, are commonly used by convolution layers to reduce the geographic extent of maps of features. Max pool was employed for this study following each convolution block. CNNs commonly use a max pool layer, a down sampling technique, to reduce the input data volume's dimension of space. Dividing the input data into independent regions and extracting the maximum result from every one is the fundamental idea underlying the max pool layer. This method reduces the amount of space while preserving the most important elements. This lowers the computational cost and enables more effective processing. The max pool procedure has the following mathematical definition:

Let A be an input matrix with size $m \times n$ and a pooling window (filter) of size $p \times q$, the resultant matrix B is obtained by selecting the maximum value from each non overlapping region:

$$b(i,j) = \max\{ a(i,j), a(i,j+1), \ldots, a(i+p-1, j+q-1) \} \quad (3)$$

Convolution layers gradually extract more complex features from the source image via this series of actions, which eventually allows the algorithm to extract complex features and generate accurate forecasts using visual information.

The flatten layer transforms data with multiple dimensions into one longer array that can be fed into the dense layers in order to apply the maps of features to them. The high-degree data gathered from the maps of features created by the convolutional layers is extracted by dense layers and is then utilized by the final layer for the final ranking. Every neuron in a layer of the dense layers connects to every other neuron in the layer below it. The input for the first dense layer is the result from the preceding layer. Softmax is used as an activation function in the CNN's last layer for multi-class recognition, and it is an essential component of the categorization procedure. Softmax is a quantitative function that creates an expected distribution from an array of numbers that are real. For multi-class recognition, CNNs use it as an outcome layer. After converting every component of the vector's input to an exponent value, and the softmax mapping normalizes the resultant values to produce likelihood. Here's the formula for the Softmax function applied to an input vector z with N elements:

$$\text{Softmax}(z_i) = e^{z_i} / \Sigma(j=1 \rightarrow n)\, e^{z_j} \quad (4)$$

Where Softmax(zi) is the i-th element of the output layer, e is the base of the natural logarithm (Euler's number), zi is the i-th element of the input vector z and the denominator is the sum of the exponential values of all elements in the input vector. The Softmax function essentially transforms the raw, un-normalized scores (logits) into probabilities. The exponentiation ensures that the resulting values are positive, and the normalization ensures that the sum of all probabilities is equal to 1.

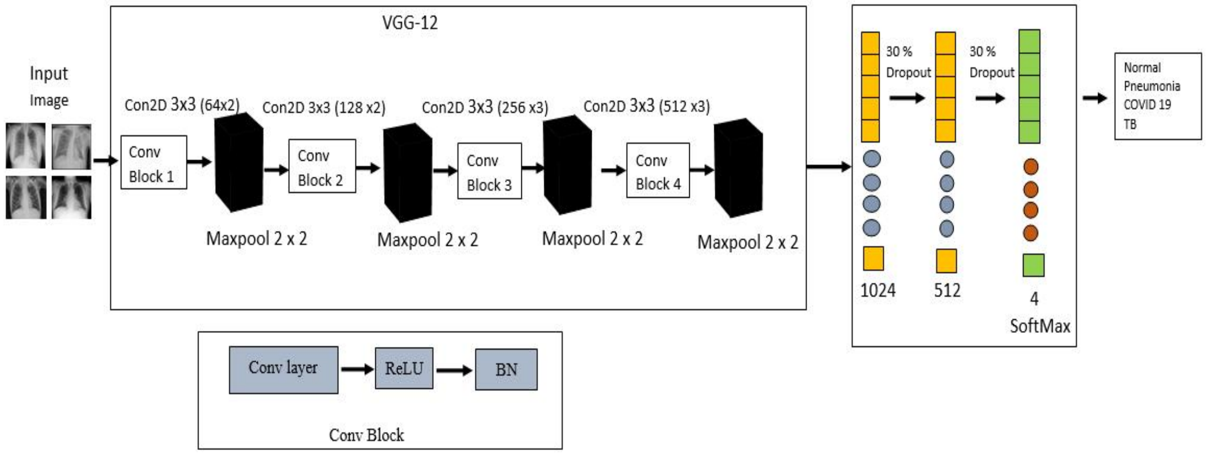

Figure 7. Adopted VGG-12 Architecture.

F. Need for Modifying the Original VGG-16 Architecture

The original VGG-16 architecture was designed for very large-scale datasets (e.g., ImageNet) and contains a deep stack of high-capacity convolutional layers that can lead to over fitting when applied to comparatively smaller datasets. Although 14,000 images were used in this study, this size remains limited relative to the depth and strength the parameters used in the original VGG-16 model, particularly for low-texture chest X-ray images based on fine-grained classification. Therefore, the last three 512-filter convolutional layers were removed to reduce model complexity and the number of trainable parameters. This architectural simplification improved generalization by mitigating over fitting while preserving the essential feature-extraction capability required for discriminating COVID-19, tuberculosis, pneumonia, and normal cases. The reduced 10-layer convolutional backbone focused on learning mid-level and clinically relevant patterns rather than highly abstract features that may not be necessary for this task. Two fully connected layers with a dropout rate of 0.3 were employed to further regularize the model and stabilize training. An iterative design strategy, systematically adding and removing layers and tuning hyper parameters, was adopted to empirically identify an optimal balance between model depth, performance, and computational efficiency. This modification ensures that the network architecture is well aligned with the dataset characteristics and the medical classification objective, leading to improved robustness and reliable performance on unseen data.

IV. RESULTS AND DISCUSSION

A. Training Environment

The enhancement in AI models like machine learning and deep learning to automate various fields, including precision agriculture and smart diagnosis. But it also necessitated the use of powerful hardware, including graphical processing units (GPUs) for visual data instead of central processing units (CPUs). It also needs a large size of memory and storage spaces to hold the data during model training. In the proposed study, Google Colab, a cloud-based platform owned by Google, was used to train the model. It provides powerful GPUs with large memory. It also provides Jupyter Notebooks with pre-installed libraries and frameworks such as TensorFlow, Keras, Matplotlib, NumPy, etc.

B. Model Measurement Metrics

After finalizing the training procedure, model evaluation is the most important step. In this step, the model is tested on new unseen images, and various formulas are used to measure the performance of the model thoroughly. In this study, the most common metrics, such as precision, accuracy, recall and F1 score, were utilized. Accuracy is used to gauge the overall performance of the model. Precision evaluates the proportion of correctly predicted positive instances (also called true positives (TPs)) out of all predicted positive instances. Recall measures the proportion of correctly predicted TPs out of all actual positives. The F1 score is the harmonic mean of the precision and recall, guaranteeing balanced evaluation of the model. All the metrics are shown below in mathematical form.

$$\text{Accuracy} = (TP + TN)/(TP + FP + TN + FN) \quad (5)$$

$$\text{Precision} = TP/(TP + FP) \quad (6)$$

$$\text{Recall} = TP/(TP + FN) \quad (7)$$

$$\text{F1-score} = 2 \cdot (\text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall}) \quad (8)$$

Where FN stands for false negative, TN stands for true negative and FP stands for false negative.

C. Results and Analysis

In this study, a modified approach based on CNNs named adopted VGG-12, a lightweight version of VGG-net, is used to identify the early symptoms of lung diseases based on multi-class classification. The purpose of the study was to develop a highly accurate model with reduced training time and parameters compared previous studies including the famous pre-trained models. Table 2 contains the layered structure of modified VGG-12 while the list of common parameters is shown in Table 3.

Table 2: Structure of the Adopted VGG-12 Framework

Layers	Filters/Neurons	Kernal-Size
Block-1		
Cov2D	64	3x3
Cov2D	64	3x3
Max pooling		2x2
Block-2		
Cov2D	128	3x3
Cov2D	128	3x3
Max pooling		2x2
Block-3		
Cov2D	256	3x3
Cov2D	256	3x3
Cov2D	256	3x3
Max pool		2x2
Block-4		
Cov2D	512	3x3
Cov2D	512	3x3
Cov2D	512	3x3
Max pool		2x2
Flatten-layer		
Dense (Dropout (0.3))	512	
Dense (Dropout (0.3))	256	
Dense (Softmax)	4	

Table 3: List of Common Parameters

Parameter	Value
Input	256 x 256 x 3
Optimizer	SGD with Momentum (0.9)
Loss-Function	Categorical-Cross Entropy
Batch-Size	32
Train-Set	80%
Validation-Set	20%
Learning-rate	0.0001
Epochs	62
Shuffle	Every-Epoch
Activation	ReLU

A big hybrid training dataset of 14,000 photos using multi-class identification was used to train the algorithm. Three categories of lung disorders were included in the dataset, while the fourth category included healthy lung X-ray photos. The model was trained using up to 60 epochs with an early stopping mechanism. The adopted VGG-12 details are displayed in Tables 2 and 3. The training growth of VGG-12, incorporating training and validation accuracy and loss, is depicted in Fig. 8.

The testing dataset, which consists of 600 fresh, unseen photos, was used to evaluate the performance of the model. Multiple evaluation metrics, including accuracy and F1-score, were employed as assessment measures to assess the results of the proposed framework. Table 4 displays the test results and ensures that this model can be used in real-world situations to assist medical professionals in enhancing the overall diagnostic effectiveness of disorders related to the lung, according to the outcomes produced by adopted VGG-12.

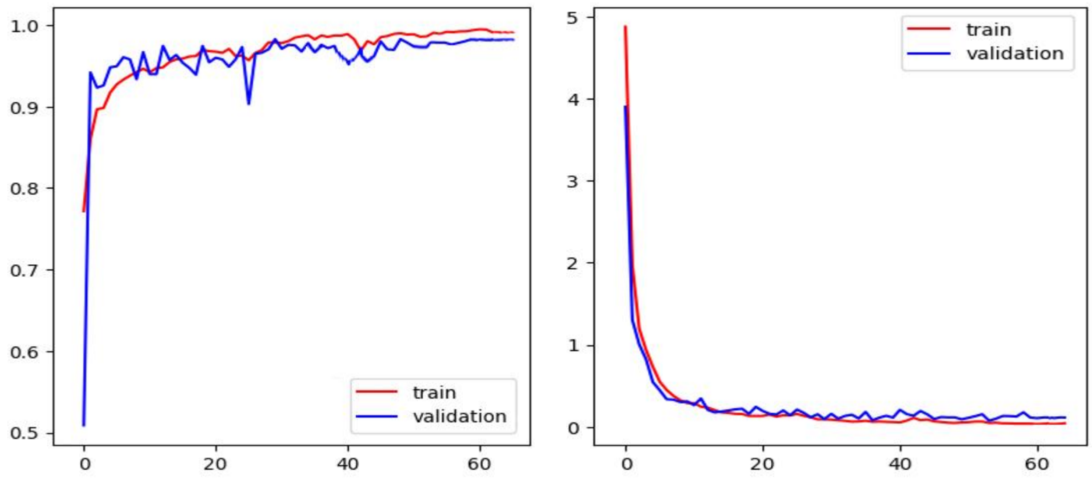

Figure 8. Training And Validation Progress Of The Proposed Model.

The training graph at left side in Fig. 8 illustrates the training and validation accuracy over multiple epochs. The horizontal axis represents the number of training epochs, while the vertical axis denotes classification accuracy. Both training and validation accuracy increase rapidly during the initial epochs and then gradually stabilize near a high value, indicating effective learning. The close alignment of the two curves suggests good generalization with minimal over fitting. The model training was terminated at the 63rd epoch using an early stopping mechanism, indicating convergence of the learning process. At this point, the model achieved high and stable performance, with the training and validation accuracy reaching approximately 98–99%.

The training graph at right side in Fig. 8 shows the training and validation loss across epochs. Loss is a quantitative measure of the error between the model's predicted outputs and the true class labels during training. It is important because minimizing the loss guides the optimization process, enabling the model to learn accurate and generalizable feature representations. The horizontal axis corresponds to the number of epochs, and the vertical axis represents the loss value. A sharp decrease in loss is observed during early training, followed by convergence to a low, stable value. The similarity between training and validation loss curves demonstrates stable optimization and confirms that the model learns robust features without significant over fitting. The training process was stopped at the 63rd epoch using the early stopping criterion, where the loss curves had converged. At this stage, both training and validation loss reached low and stable values (approximately 0.1 or lower).

Table 4: ***Performance Of Modified Vgg-12 On Test Data***

Disease	**Accuracy** (%)	**Precision** (%)	**Recall** (%)	**F1-Score** (%)
Pneumonia	96.68	96.66	96.69	96.70
COVID-19	97.30	96.00	97.30	96.65
TB	94.70	98.00	94.68	95.30
Healthy	98.10	96.05	97.90	97.05
Average	96.70	96.69	96.70	96.40

We conducted iterative trials using a shallow model based on CNN as the baseline to finalize the modified VGG-12 as the suggested framework. In order to achieve the study's objective, we progressively appended layers to the network and repeatedly and methodically adjusted different training parameters to raise the accuracy of the model. Although we employed a variety of sequential CNN-based structural configurations, the VGG-16 framework offered the best chance of achieving the study's objective. Due to its data-greedy character, the VGG-16 algorithm displayed a reasonable rate of over fitting. Throughout the course of the experimental process, we employed the early stopping strategy. ES is a method for preventing over fitting, which occurs when an algorithm produces extremely precise results but performs poorly on the test images. The opposite of generalization, this state is also known as 'memorization'. Early stopping, which ends the model training before the framework reaches a state of over fitting, assists in ensuring that any DL algorithm will perform adequately with fresh unseen images, which is its ultimate goal. Compared to earlier tests, VGG-16 produced favorable outcomes and showed less over fitting.

VGG-Nets are well-known for their straightforward design and precise outcomes, but they also require a lot of data. In this case, the 14,000-picture dataset did not satisfy the data requirements of the VGG-16 model. As a consequence, an ongoing decrease in the amount of layers was tried. Improved outcomes were shown along with less over fitting after using 12 layers, 10 convolutional and 2 dense layers, along with a decrease in neuron width from 4096 to 512 and 256, respectively. The addition of BN and dropouts improved the efficiency of the framework even further. This improvement brought the framework to an acceptable level and made it consistent with the goals of the suggested investigation. After the arrangement of layers was finalized, the framework was trained several times using various optimizers like RMSProp, Adam, and SGD. However, SGD with momentum (0.9) turned out to be the best option, even if it took a little longer to reach convergence.

Three well-known pre-trained networks, such as MobileNetV2, ResNet50, and VGG-16, were trained and tested on the same data, and their results were compared to the suggested approach. These models were trained using transfer learning with the identical dense layers and training parameters. Each of these models responded more poorly than adopted VGG-12, although achieving positive outcomes, as seen in Table 5.

Table 5: *Results Comparison of Adopted VGG-12 with Well-Known Pre-Trained Models*

Network	Class	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ResNet50	COVID-19	95.33	96.62	95.33	95.78
	Pneumonia	96.00	94.11	96.00	94.97
	TB	94.00	95.91	94.00	94.78
	Healthy	96.67	95.39	96.67	96.03
	Average	95.50	95.51	95.50	95.39
MobileNet-v2	COVID-19	94.67	94.67	94.67	94.77
	Pneumonia	94.00	95.27	94.00	94.66
	TB	94.67	94.03	94.67	94.55
	Healthy	95.33	94.70	95.33	94.79
	Average	94.67	94.67	94.67	94.69
VGG16	COVID-19	95.33	94.70	95.33	94.79
	Pneumonia	94.67	95.30	95.03	95.14
	TB	93.33	93.33	93.33	93.44
	Healthy	94.00	94.00	94.00	94.34
	Average	94.33	94.33	94.42	94.43
VGG12	Average	96.67	96.68	96.67	96.42

This study demonstrated the effectiveness of deep learning structures, particularly the customized and modified VGG-12 model, and marked a substantial advancement in the field of autonomous medical diagnostics. This study is at the core of efforts to transform the identification of lung disorders because of the careful experimental design and thorough analysis of the outcomes. A visual comparison of multiple approaches used in this study is presented in Fig. 9.

D. Performance Evaluation of Adopted VGG-12 Network on a Public Image Set

With the primary goal of differentiating between healthy and unhealthy samples, the proposed model was thoroughly trained and validated utilizing a publicly accessible data set of apple leaves. The 3171 photographs in the collection were divided into four categories: healthy (1644 images), scab (631 images), and black-rot (620 images), and rust (276 images). The categorizing effort reduced the set of images to two categories, which were "normal" and "infected". Each of the infected leaf images was categorized as diseased leaf pictures within the "diseased" category. The modified VGG-12 framework underwent only minor changes, namely switching from categorical to binary cross-entropy for the loss function and substituting a sigmoid for the Softmax activation function. Table 6 summarizes the outcomes of the test set, which demonstrate the algorithm's effective operation with an astounding success rate of 97.56%. Additionally, the model showed effectiveness in both the classification of binary challenges and the classification of multiple classes. This twofold ability confirms the algorithm's adaptability in multiple situations.

Table 6: *Performance of Adopted VGG-12 on Public Dataset*

Dataset	Source	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Apple	Plant-Village	97.66	96.51	98.80	97.64

E. Comparison of Adopted VGG-12 with Other SOTA Methods

A thorough comparison with previous research was done to strengthen the significance of the suggested lung disease identification method. Techniques employing the same diseases but with different category counts, as shown in Table 7, made up the comparative evaluation. This thorough analysis aims to demonstrate the improvements and efficiency of the proposed approach in relation to the identification of lung illnesses.

Table 7: ***Comparison Of Adopted Vgg-12 With Previous Sota Methods***

Study	Disease	Method	Images	Accuracy (%)
[44]	COVID-19	CNN (TL)	CT + CXR	93.00
[45]	COVID-19	CNN (TL)	CT	94.00
[46]	COVID-19	ResNet-50	CXR	96.10
[47]	Pneumonia	CNN-LSTM	CXR	95.04
[48]	Pneumonia	CNN (TL)	CXR	95.09
[49]	Pneumonia	CNN (TL)	CXR	95.09
[50]	TB	CNN	CXR	98.45
[51]	TB	CNN (TL)	CT	85.29
[52]	TB	Ensemble CNN (TL)	CXR	88.24
Proposed	COVID-19	VGG-12 (CNN)	CXR	97.33
Proposed	Pneumonia	VGG-12 (CNN)	CXR	96.67
Proposed	TB	VGG-12 (CNN)	CXR	94.66

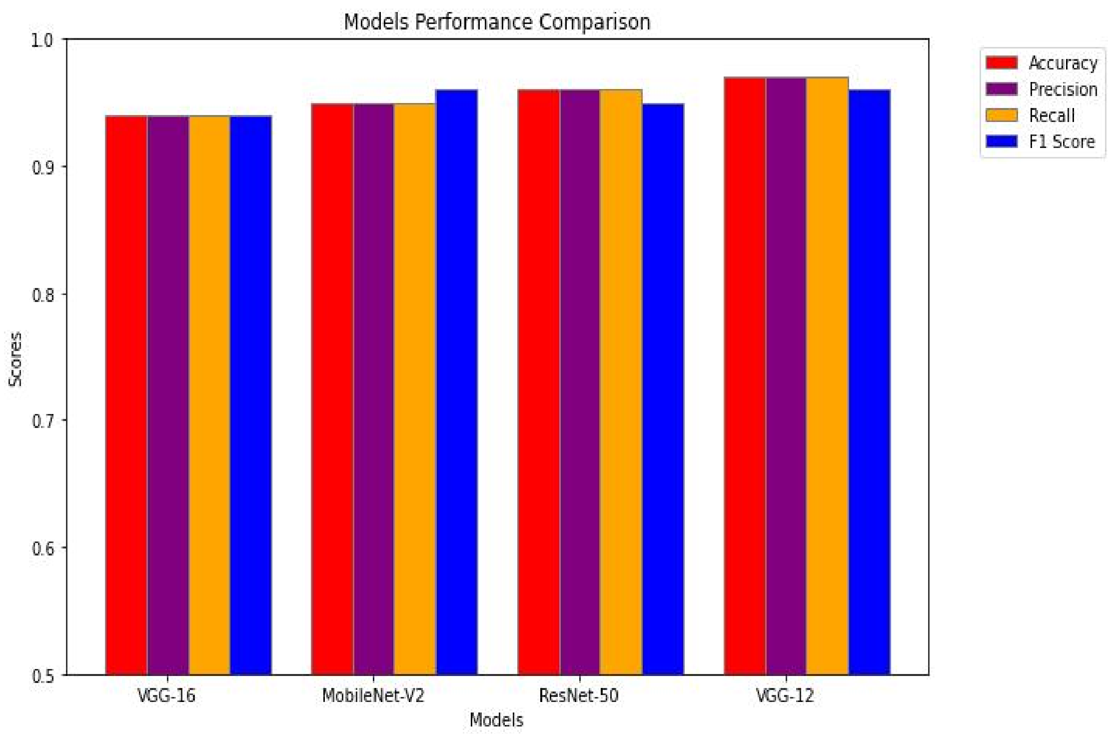

FIGURE 9. Pictorial comparison of multiple models with modified VGG-12.

V. CONCLUSION AND FUTURE WORK

A. Conclusion

The aim of the research was to build an accurate and efficient CNN framework to identify lung ailments from CXR photos. In order to successfully fit the suggested dataset while reducing the amount of over fitting, we strategically reduced the number of layers and training parameters in the well-known VGG-16 model, which we named Adopted VGG-12. CNN approaches require a lot of photos as a training dataset; therefore, a large set of images was collected from multiple sources to train and test the adopted VGG-12 model. The training photos were subjected to augmentation and preprocessing strategies before being fed into the neural network. With an average accuracy of 96.42%, the framework showed strong results on multi-class data when it came to the evaluation of CXR photographs for the identification of prevalent lung illnesses such as pneumonia, TB, and COVID-19 in addition to healthy images. Additionally, its versatility was demonstrated via binary classification, where it achieved an astounding 97.66% success score on a dataset of photographs of apple leaves. This work sets the stage for future advancements in automated medical diagnosis by demonstrating the potential of lightweight frameworks for efficient medical image processing, alongside the immediate benefits offered by the proposed model.

B. Future Work

Our future goals include expanding the training and testing datasets and adding data from surrounding hospitals to the evaluation set. Incorporating more illness classes related to anomalies of the chest is another goal. In the future, we will compare the outcomes of more sophisticated CNN-based approaches, such as attention mechanisms, fire components, and capsule models, with the current adopted VGG-12 and select the most precise and generalized algorithm with the fewest training weights. In order to guarantee accurate results across a range of healthcare imaging issues, we also seek to maintain dataset balance and strengthen the algorithm's abilities for complex diagnosis settings.

VI REFERENCES

[1] Nahiduzzaman, M., Islam, M. R., & Hassan, R. (2023). ChestX-Ray6: Prediction of multiple diseases including COVID-19 from chest X-ray images using convolutional neural network. Expert Systems with Applications, 211, 118576.

[2] Mahbub, M. K., Biswas, M., Gaur, L., Alenezi, F., & Santosh, K. C. (2022). Deep features to detect pulmonary abnormalities in chest X-rays due to infectious diseaseX: Covid-19, pneumonia, and tuberculosis. Information Sciences, 592, 389-401.

[3] Orme I, Secrist J, Anathan S, Kwong C, Maddry J, Reynolds R, et al. Search for new drugs for treatment of tuberculosis. Antimicrob Agents Chemother 2001; 45(7):1943–6.

[4] Xu G, Liu H, Jia X, Wang X, Xu P. Mechanisms and detection methods of Mycobacterium tuberculosis rifampicin resistance: The phenomenon of drug resistance is complex. Tuberculosis 2021;128.

[5] Brooks JV, Orme IM. Evaluation of once-weekly therapy for tuberculosis using isoniazid plus rifamycins in the mouse aerosol infection model. Antimicrob Agents Chemother 1998;42(11):3047–8.

[6] Taye, M. M. (2023). Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. Computers, 12(5), 91.

[7] Kim, H.E., Cosa-Linan, A., Santhanam, N. et al. Transfer learning for medical image classification: a literature review. BMC Med Imaging 22, 69 (2022).

[8] Zhan X, Liu J, Long H, Zhu J, Tang H, Gou F, Wu J. An Intelligent Auxiliary Framework for Bone Malignant Tumor Lesion Segmentation in Medical Image Analysis. Diagnostics. 2023; 13(2):223.

[9] Pinto-Coelho, L. How artificial intelligence is shaping medical imaging technology: a survey of innovations and applications, Bioengineering (Basel) 10 (12)(2023)

[10] Manakitsa N, MaraslidisGS, Moysis L, Fragulis GF. A Review of Machine Learning and Deep Learning for Object Detection, Semantic Segmentation, and Human Action Recognition in Machine and Robotic Vision. Technologies. 2024; 12(2):15.

[11] Ullah N, Khan JA, Almakdi S, Alshehri MS, Al Qathrady M, El-Rashidy N, El-Sappagh S and Ali F (2023) An effective approach for plant leaf diseases classification based on a novel DeepPlantNet deep learning model. Front. Plant Sci. 14:1212747.

[12] Shekar, B. H., Abraham, W., & Pilar, B. (2022, December). Offline Signature verification using CNN and SVM classifier. In 2022 IEEE 7th International Conference on Recent Advances and Innovations in Engineering (ICRAIE) (Vol. 7, pp. 304-307). IEEE.

[13] Zhu Z, Dai Y. A New CNN-Based Single-Ingredient Classification Model and Its Application in Food Image Segmentation. Journal of Imaging. 2023; 9(10):205.

[14] Akash, S. A., Moorthy, R. S. S., Esha, K., & Nathiya, N. (2022, August). Human violence detection using deep learning techniques. In Journal of Physics: Conference Series (Vol. 2318, No. 1, p. 012003). IOP Publishing.

[15] Hassan, S. M., Maji, A. K., Jasińki, M., Leonowicz, Z., &Jasińska, E. (2021). Identification of plant-leaf diseases using CNN and transfer-learning approach. Electronics, 10(12), 1388.

[16] Krishnamoorthy, N., Prasad, L. N., Kumar, C. P., Subedi, B., Abraha, H. B., &Sathishkumar, V. E. (2021). Rice leaf diseases prediction using deep neural networks with transfer learning. Environmental Research, 198, 111275.

[17] Wu, Z. H. E., Zhao, S., Peng, Y., He, X., Zhao, X., Huang, K., ... & Li, Y. (2019). Studies on different CNN algorithms for face skin disease classification based on clinical images. IEEE Access, 7, 66505-66511.

[18] Shamrat, F. J. M., Azam, S., Karim, A., Islam, R., Tasnim, Z., Ghosh, P., & De Boer, F. (2022). LungNet22: a fine-tuned model for multiclass classification and prediction of lung disease using X-ray images. Journal of Personalized Medicine, 12(5), 680.

[19] Srinivasu, P. N., SivaSai, J. G., Ijaz, M. F., Bhoi, A. K., Kim, W., & Kang, J. J. (2021). Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. Sensors, 21(8), 2852.

[20] Ullah, N., Khan, M. S., Khan, J. A., Choi, A., & Anwar, M. S. (2022). A robust end-to-end deep learning-based approach for effective and reliable BTD using MR images. Sensors, 22(19), 7575.

[21] Khan, A., Khan, A., Ullah, M., Alam, M. M., Bangash, J. I., & Suud, M. M. (2022). A computational classification method of breast cancer images using the VGGNet model. Frontiers in Computational Neuroscience, 16, 1001803.

[22] Nigat, T. D., Sitote, T. M., & Gedefaw, B. M. (2023). Fungal Skin Disease Classification Using the Convolutional Neural Network. Journal of Healthcare Engineering, 2023.

[23] Aytaç, U.C., Güneş, A. &Ajlouni, N. A novel adaptive momentum method for medical image classification using convolutional neural network. BMC Med Imaging 22, 34 (2022).

[24] Eweje, F. R., Bao, B., Wu, J., Dalal, D., Liao, W. H., He, Y., ... & Bai, H. X. (2021). Deep learning for classification of bone lesions on routine MRI. EBioMedicine, 68.

[25] Thilagaraj, M., Arunkumar, N., &Govindan, P. (2022). Classification of breast cancer images by implementing improved dcnn with artificial fish school model. Computational Intelligence and Neuroscience, 2022.

[26] Malik, H., Anees, T., Chaudhry, M. U., Gono, R., Jasiński, M., Leonowicz, Z., &Bernat, P. (2023). A Novel Fusion Model of Hand-Crafted Features with Deep Convolutional Neural Networks for Classification of Several Chest Diseases using X-ray Images. IEEE Access.

[27] Alshmrani, G. M. M., Ni, Q., Jiang, R., Pervaiz, H., & Elshennawy, N. M. (2023). A deep

learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images. Alexandria Engineering Journal, 64, 923-935.

[28] Pinto-Coelho L. How Artificial Intelligence Is Shaping Medical Imaging Technology: A Survey of Innovations and Applications. Bioengineering (Basel). 2023 Dec 18; 10(12):1435.

[29] Duwairi, Rehab & Melhem, Abdullah. (2023). A deep learning-based framework for automatic detection of drug resistance in tuberculosis patients. Egyptian Informatics Journal. 24. 10.1016/j.eij.2023.01.002.

[30] C. Baloescu et al., "Automated Lung Ultrasound B-Line Assessment Using a Deep Learning Algorithm," in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 67, no. 11, pp. 2312-2320, Nov. 2020.

[31] Duong, L. T., Nguyen, P. T., Iovino, L., & Flammini, M. (2023). Automatic detection of Covid-19 from chest X-ray and lung computed tomography images using deep neural networks and transfer learning. Applied Soft Computing, 132, 109851.

[32] Sathitratanacheewin, S., Sunanta, P., & Pongpirul, K. (2020). Deep learning for automated classification of tuberculosis-related chest X-Ray: dataset distribution shift limits diagnostic performance generalizability. Heliyon, 6(8).

[33] Alqaissi, E. (2025). A novel and ultralight convolutional neural network model for real time detection of infectious lung diseases. Digital Health, 11, 20552076251318155.

[34] Saddique, A., Manan, A., Ali, M., Siddiqui, S., & Rehan, M. (2025). hybrid deep learning models for multi-class classification of chest x-ray images: normal, pneumonia, and covid-19. Spectrum of Engineering Sciences, 3(7), 21-33.

[35] Ullah, N., Marzougui, M., Ahmad, I., & Chelloug, S. A. (2023). DeepLungNet: an effective DL-based approach for lung disease classification using CRIs. Electronics, 12(8), 1860.

[36] Sharma, S., & Guleria, K. (2023). A deep learning-based model for the detection of pneumonia from chest X-ray images using VGG-16 and neural networks. Procedia Computer Science, 218, 357–366. https://doi.org/10.1016/j.procs.2023.01.018

[37] Jain, D. K., Singh, T., Saurabh, P., et al. (2022). Deep learning-aided automated pneumonia detection and classification using CXR scans. Computational Intelligence and Neuroscience, 2022, 7474304. https://doi.org/10.1155/2022/7474304

[38] Asnake, N. W., Salau, A. O., & Ayalew, A. M. (2024). X-ray image-based pneumonia detection and classification using deep learning. Multimedia Tools and Applications, 83(21), 60789–60807.

[39] Goyal, S., & Singh, R. (2023). Detection and classification of lung diseases for pneumonia and COVID-19 using machine and deep learning techniques. Journal of Ambient Intelligence and Humanized Computing, 14(4), 3239–3259.

[40] Taher, F., Haweel, R. T., al Bastaki, U. M. H., Abdelwahed, E., Rehman, T., & Haweel, T. I. (2022). Fast COVID-19 detection from chest X-ray images using DCT compression. Applied Computational Intelligence and Soft Computing, 2022, 7. https://doi.org/10.1155/2022/6507594 (verify DOI if available)

[41] Bhosale, R. D., & Yadav, D. M. (2024). Customized convolutional neural network for pulmonary multi-disease classification using chest X-ray images. Multimedia Tools and Applications, 83(6), 18537–18571.

[42] Ko, J., Park, S., & Woo, H. G. (2024). Optimization of vision transformer-based detection of lung diseases from chest X-ray images. BMC Medical Informatics and Decision Making, 24(1), 191.

[43] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[44] V. Perumal, V. Narayanan, S.J.S. Rajasekar, Detection of COVID-19 using CXR and CT images using Transfer Learning and Haralick features, Appl. Intell. 51 (1) (2021) 341–358.

[45] L. Li, Z. Xu, et al, Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT, Radiology 296 (2) (2020) E65–E71.
[46] A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks, Pattern Anal. Appl. 24 (3) (2021) 1207–1220.
[47] S. Goyal, R. Singh, Detection and classification of lung diseases for pneumonia and Covid-19 using machine and deep learning techniques, J. Ambient Intell. Humaniz. Comput. (0123456789) (2021).
[48] K. El Asnaoui, Design ensemble deep learning model for pneumonia disease classification, Int. J. Multimed. Inf. Retr. 10 (1) (2021) 55–68.
[49] J.E. Luja´ n-Garcı´a, M.A. Moreno-Ibarra, Y. Villuendas-Rey, C. Ya´ n˜ ez-Ma´ rquez, Fast COVID-19 and pneumonia classification using chest X-ray images, Mathematics 8 (2020) 9.
[50] S. Sathitratanacheewin, P. Sunanta, K. Pongpirul, Deep learning for automated classification of tuberculosis-related chest X-Ray: dataset distribution shift limits diagnostic performance generalizability, Heliyon 6 (8) (2020) e04614.
[51] X.W. Gao, C. James-Reynolds, E. Currie, Analysis of tuberculosis severity levels from CT pulmonary images based on enhanced residual deep learning architecture, Neurocomputing 392 (2020) 233–244.
[52] R. Hooda, A. Mittal, S. Sofat, Automated TB classification using ensemble of deep architectures", Multimed. Tools Appl. 78 (22) (2019) 31515–31532.
[53] Z. N. K. Swati, A. Ghulam, M. Sohail, J. U. Arshed, R. Sikander, M. S. Malik, and N. Khan, "XGboost-Ampy: Identification of AMPylation protein function prediction using machine learning," VAWKUM Transactions on Computer Sciences, vol. 10, no. 2, pp. 83–95, 2022.

[54] M. Sohail and A. Malook, "Exploring IIoT: Wireless control systems with ESP8266 as a web-server controller basic experimentation," VAWKUM Transactions on Computer Sciences, vol. 13, no. 1, pp. 258–277, 2025.https://doi.org/10.21015/vtcs.v13i1.1902