International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Enhancing Lung Disease Diagnosis through a Hybrid Deep Learning Approach Pro Chest X-Ray Image Classification

Dr. Sreeja K. A.[1], Dr. Arshey M.[2], Dr. Gayathry S. Warrier[3], Dr. Arun Pradeep[4]

Submitted: 27/12/2023 **Revised**: 03/02/2024 **Accepted**: 11/02/2024

Abstract: Lung disorders have a wide-reaching influence, resulting in reduced lung function and a variety of complications, such as breathing difficulty, airway blockages, and exhalation problems. Due to limited resources for lab tests and imaging procedures, early diagnosis of illnesses such as pneumonia, fibrosis, etc., remains difficult. The use of chest X-ray pictures for quick disease monitoring, crucial for ICU patients, has gained popularity due to this problem, and image processing & machine learning models have become more popular. Deep learning for lung disease detection entails three critical steps: picture pre-processing, training, and classification. Relevant features indicative of lung disorders are extracted utilizing a range of deep learning methodologies such as CNNs, RNNs, Attention Mechanisms, Transfer Learning, GANs, and VGG architectures after improving the raw quality of X-ray images by optimal filtering techniques. While typical CNNs may struggle with complicated characteristics, potentially compromising lung cancer classification accuracy, a unique method has been developed via hybrid VGG-CNN architecture. This hybrid architecture captures local and global elements; whereas CNNs excel at detail-oriented aspects, VGG networks efficiently capture wider patterns. The effectiveness of this methodology is demonstrated using open datasets that include NIH Chest X-ray data. The classification of the gathered CNN features is so, therefore, performed using Random Forest and Support Vector Machine models. A variety of systems of measurement, such as accuracy, precision, recall, and F-measure, are used to evaluate this method's effectiveness. The normal CNNVGG-SVM model's accuracy of 93.54% is significantly outperformed by the hybrid SVM-RF model's outstanding accuracy of 97.89%, a gain of 4.35%. Similarly, the hybrid RF model achieves an accuracy of 98.99%, outperforming the standard CNNVGG-RF model by 4.32%, or 94.89%. These metrics thoroughly evaluate the methodology's capacity to diagnose various lung illnesses reliably. The effectiveness of the suggested technique is demonstrated by its exceptional accuracy in improving lung disease diagnosis.

Keywords: *Chest X-ray images, deep learning, Visual geometry group, Diagnosis improvement, Performance metrics*

1. Introduction

Lung illnesses cause damage to the airways of the lungs as well as other pulmonary structures. They are also known as respiratory diseases [1, 2]. Lung conditions brought on by various circumstances have caused more deaths in recent years. The new Covid-19 virus causes minimal to moderate side effects in pneumonia patients, including high body temperature, hacking, and difficulty of breath. [2]. However, a few people died from severe pulmonary pneumonic diseases [3-5].

The high chest obstruction (Pneumonia) that affected many Coronavirus cases who died due to the sickness resulted in a large decrease in oxygen levels and eventually cardiovascular failure. Pneumonia, a type of lung disease, irritates the body's small air sacs in the lungs. It can cause you to drink many fluids, making it difficult to relax. Numerous illnesses, such as infections caused by bacteria, viral infections, and ordinary colds, can result in pneumonia [6]. The probability of surviving and recovering are both increased by early identification [7-8]. The probability of an individual living a long life is generally said to rise if a cancer case is discovered early, diagnosed, and successfully treated [9–10]. Skin examination, tests for blood, mucus sample tests as such, chest X-ray exams, and chest computed tomography (CT) analyses have all been used to identify lung illness. Medical specialists are necessary to analyze medical data and diagnose ailments, and because medical images are so complicated, expert opinions usually differ while examining them. In the field of medicine, artificial intelligence is critical. Because they provide cutting-edge solutions for medical applications, ML and DL algorithms have grown popular in recent years for analyzing medical images and detecting ailments [11]. Although research in this sector is ongoing, providing a prediction system that delivers precise diagnoses and classifications remains challenging.

[1] Associate Professor, Department of Electronics and Communication Engineering, SCMS School of Engineering and Technology, Karukutty Ernakulam, Kerala, India.
[2] Assistant Professor, Computer Science And Engineering, Scms School Of Engineering And Technology, arsheym@scmsgroup.org
[3] Assistant Professor
Department of Computer Science
Christ (Deemed to be University), Bangalore, Karnataka.
[4] Associate Professor,
Department of Electronics and Communication Engineering.
Saveetha Engineering College (Autonomous), Chennai, Tamil Nadu-602105.
Corresponding Author: ka.sreeja@gmail.com
** Corresponding Author Email: kavitavce@gmail.com*

The classification of lung diseases using DL and ML methods has attracted much attention in recent years. DL and ML algorithms have been applied in numerous computer vision methodologies. Recently, several research has been put out that predict Covid-19 using DL and ML models. The most recent theories for using chest X-ray pictures to categorize lung diseases are covered in this section.

Several studies have suggested techniques for diagnosing lung conditions from chest X-ray pictures. [12] developed and assessed DL-based methods for CNN-based bronchial illness diagnosis. When paired with delicate registration procedures, the authors demonstrated considerable training and testing time effects of the CNN model. Using computer vision methods, the authors of [13] occurred able to identify pneumonia in chest X-ray pictures. They employed lung segmentation and neural groupings to identify sites of pneumonia. A similar strategy was employed in [14], which used a CNN model for lung disease prediction focusing on pneumonia detection. [15] demonstrated pneumonia classification utilizing chest X-ray images using a CNN model called VGG16, with an accuracy of 90.54%. [16] used CNNs to distinguish between pneumonic and normal X-ray images, whereas [17] provided an innovative automated method for detecting pneumonia in chest X-ray images, obtaining excellent classification accuracy. A CNN model was utilized to diagnose pneumonia cases in [18]. To recognize pneumonia symptoms, the authors fine-tuned the last layers of a loaded model (VGG16). [19] presented a weighted classifier that combined DL models and achieved good accuracy on a pneumonia dataset. To diagnose pneumonia, [20] suggested combining the deep CNN model CheXNet with the VGG-19 algorithm. For correctly detecting COVID-19 chest X-ray pictures, DeTraC was developed [21]. CNNs were also employed by [22] to increase pneumonia classification training and validation accuracy. In [23], the effectiveness of CNN frameworks for COVID-19 detection was investigated. [24] suggested COVID DetectioNet, which would use CNN-enhanced AlexNet to detect COVID-19. [25] used transfer learning on chest X-ray datasets to distinguish between bacterial pneumonia, viral pneumonia, and COVID-19. CheXGCN, which uses Graph Convolutional Networks, was launched in [26] to categorize chest X-rays. [27] the Convolutional Support Estimation Network was proposed to address difficulties with execution speed and space. [28] employed transfer learning-based algorithms to diagnose pulmonary illness. [29] recommended using CAD to categorize lung- and health-related disorders. Chest X-ray scans were used in several investigations to check for COVID-19. [30–31] identified COVID-19 patients using the VGG19 and ResNet50 models. [32] achieved outstanding accuracy with a dataset of 6505 images, whereas [33] reached 92.9% accuracy with 5941 images. Applying DL and ML approaches in chest X-ray images for lung disease diagnosis shows potential for assisting radiologists and boosting diagnostic accuracy. The dearth of large and well-annotated datasets is a fundamental obstacle to precise lung illness categorization utilizing machine learning and deep learning techniques. Model development is hampered by the scarcity of medical imaging data for specific lung illnesses. A lack of representation and diversity in datasets might hamper the generalization and usefulness of existing models. Complex categorization structures and algorithms may produce difficult-to-interpret findings, particularly for healthcare personnel lacking data science skills. It is critical to ensure model interpretability in order to acquire trust and acceptance in therapeutic applications. Transferring models from certain datasets to diverse patient groups and imaging techniques can be difficult. Variability in the presentation of lung illness, imaging modalities, and patient variables might restrict model performance and induce bias. Efforts are needed to improve the adaptability of models across varied demographics and imaging differences. Using machine learning and deep learning models in clinical contexts brings practical obstacles such as user-friendly interfaces, regulatory compliance, and interaction with existing medical systems. Bridging the research-to-clinical-use gap necessitates collaboration among researchers, physicians, and technology specialists. The use of AI to classify lung diseases raises ethical questions about privacy, consent, and potential biases. Biases in data collection, annotation, or model training might result in discrepancies in diagnosis and treatment. These ethical challenges and ensuring fairness and inclusivity in model building are significant research topics. Despite advances, further study is required in a variety of areas. This includes creating robust models that can handle class imbalances and rare diseases, improving model interpretability, leveraging multimodal data for more accurate predictions, and conducting extensive validation studies in various clinical settings to assess the models' real-world effectiveness and impact.

The study outlines a thorough method representing classifying chest X-ray pictures utilizing a hybrid VGG-CNN framework to improve the diagnosis of lung diseases. The topic of accurate lung disease classification in chest X-ray pictures is discussed after the introduction. The suggested methodology enhances the extraction and classification of features using a hybrid VGG-CNN model. The report then thoroughly examines the chest X-ray picture dataset before presenting the findings and analyzing their ramifications. Finally, the research emphasizes the importance of the hybrid approach in advancing lung disease diagnostics.

2. Problem Statement

The global burden of lung illnesses is a major healthcare concern, affecting millions globally. COVID-19, pneumonia, fibrosis, and tuberculosis affect pulmonary function, leading to serious problems such as breathing difficulties, airway blockages, and impaired exhalation. Despite advances in medical technology, early and reliable detection of lung illnesses remains difficult, owing mostly to limited laboratory resources and imaging tools. Current diagnostic methods sometimes rely on time-consuming, subjective approaches requiring specific skills. The demand for a more efficient and trustworthy diagnosis process has prompted researchers to investigate the capabilities of modern technologies, notably image processing and machine learning. Chest X-ray images, widely used as diagnostic tools, provide information about lung problems by visualizing lung health. This wealth of data has inspired the development of automated algorithms for analyzing photos and detecting indications of various lung illnesses. Despite these developments, it is difficult to diagnose lung illness from chest X-ray pictures automatically. CNNs, a well-known deep learning method, have excelled in tests of picture categorization.

The complexity of lung disorders, combined with the inherent variances in medical imaging data, can, however, limit the efficiency of traditional CNN models. These difficulties are exacerbated when dealing with complicated traits suggestive of lung malignancies, making precise diagnosis difficult. In order to overcome these limitations, it is vitally important to offer a sophisticated and trustworthy framework that can overcome the difficulties of lung disease diagnosis via chest X-ray pictures. In particular, CNNs, Recurrent Neural Networks (RNNs), Attention Mechanisms, Transfer Learning, Generative Adversarial Networks (GANs), and the capabilities of the Visual Geometry Group (VGG) should all be utilized to their full potential within this framework. Considering the quickness with which lung disease progresses, especially in critical care settings, including ICUs, the suggested architecture must also allow real-time or virtually real-time interpretation of chest X-ray pictures. This unique strategy attempts to improve accuracy, increase sensitivity, and decrease false negatives in the detection of lung illnesses by merging various deep-learning approaches. Creating such a comprehensive solution might significantly affect patient care, ease the strain on healthcare systems, and even save lives.

3. Proposed Methodology

Figure 1 depicts the study's methodology, which combines state-of-the-art image processing and deep learning methods to boost the accuracy of lung disease diagnosis using chest X-rays

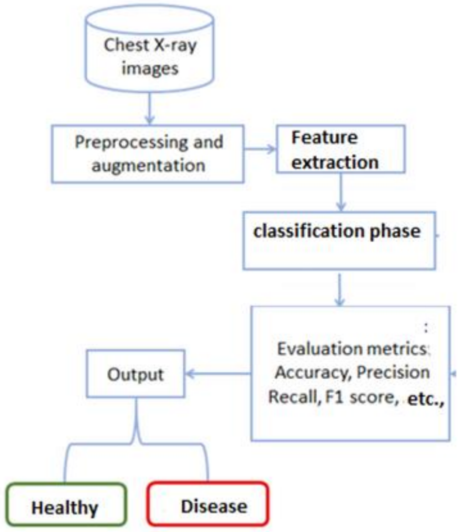


Fig 1 Proposed Methodology

Considering the constraints caused by resource shortages and the need to diagnose critically ill patients in ICUs, the goal is to address the issues of early diagnosis of lung illnesses such as COVID-19, pneumonia, fibrosis, and tuberculosis. The suggested methodology entails several phases, including image pre-processing, model training, and disease categorization. The technique identifies regional and global features in chest X-ray images, offering a thorough analysis for precise diagnosis.

3.1. Data set Description

X-rays of the chest are frequently used in medical imaging because they are affordable and efficient. This technique, frequently in great demand in clinical settings, is a vital tool for identifying lung-related disorders. However, compared to diagnosing the chest using computed tomography (CT) imaging, diagnosing the lungs with chest X-rays may present certain complications. Chest X-rays can make clinically meaningful diagnoses, but only some comprehensive and resource-rich public datasets are available. This presents a substantial difficulty when applying computer-aided detection across multiple medical institutions. Properly labeling many images presents a substantial challenge when creating large-scale chest X-ray datasets. The largest available dataset before the emergence of the current datasets was Openi, which had 4143 chest or lung X-ray images on websites like Kaggle. A sample of the full dataset [34], consisting of a randomly chosen 5% of the dataset, is provided as a subset. Each image in this subset has a resolution of 1024 x 1024 pixels, making up 5606. Patient-specific data and associated class labels have been collected into a (.csv) file format to simplify data administration and enable efficient classification. The dataset includes a wide range of 15 classes and various medical cases. One of these classes denotes "No findings," while the other 14 classes cover a spectrum of different disorders. These cover various illnesses and disorders that doctors regularly diagnose from X-ray records. Healthcare professionals may efficiently diagnose, monitor, and interpret a variety of lung disorders using the rich insights gained from X-ray chest images. Due to the dataset's multidimensionality, intelligent robots can work with doctors to improve diagnostic capabilities and the accuracy of health assessments.

This work uses three representative X-ray images from the extensive dataset [34], as shown in Figure 2.

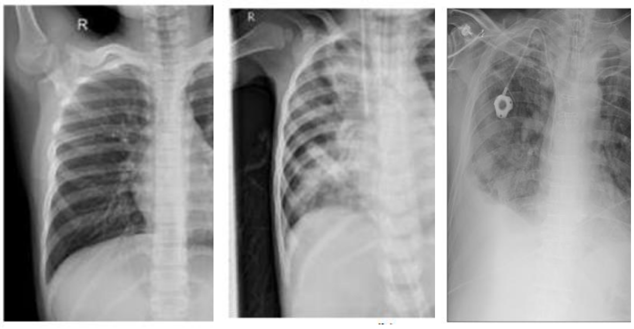
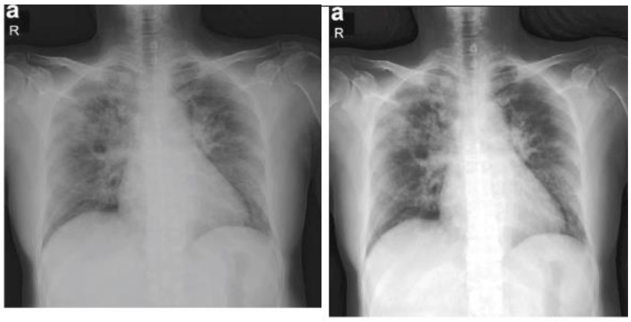


Fig 2 Sample X-ray Images

3.2 Image pre-processing

Great image quality is essential across various image and video processing applications to prevent skewed results and potential consequences. Advanced biometric scanning technologies in the medical field result in lower-quality medical images like X-rays. Unfortunately, current automatic Computer-Aided Diagnosis (CAD) systems for lung disorders frequently ignore these quality issues related to X-ray images. This oversight reduces the accuracy of CAD models, especially in situations involving real-time patient monitoring. Our method attempts to address this problem by improving the quality of input photographs using sensible and simple strategies. The first step in standardizing each input chest image is to shrink it to 512 X 512 pixels & convert it to a grayscale format. The 2D grayscale X-ray image 'x' is processed with contrast enhancement, Wiener filtering, and histogram equalization to strike a compromise between improved image quality and reduced data loss. The procedure begins with a contrast adjustment to improve the weaker regions of the input image. The adjust (.) function enhances the image's contrast and adds noise and artifacts. In order to fix this, the filtered image is created by applying a 2D Wiener filtering operation on the corrected image, "x1." Wiener filtering is the most efficient method of filtering when quality measures like Peak to Signal Noise Ratio, Structural Similarity Index Matrix, and Root Mean Square Error are considered. Since Wiener filtering has adaptive noise reduction capabilities, it produces better results. In the Wiener filtering procedure, the default neighborhood size 'N' is consistent with the strategy described in reference [35]. This thorough enhancement procedure successfully resolves input X-ray picture quality concerns, enhancing the precision and dependability of future studies. Figure 3 depicts an example of a pre-processed image.



(a)Actual image (b)Pre-processing image

Fig 3 Sample of actual and pre-processing Image

3.3 Feature extraction

A hybrid approach utilizes the benefits of both CNNs and VGG networks to detect lung diseases effectively. While CNNs are excellent at collecting the minute details evident in the photographs, VGG networks are better at recognizing larger patterns and traits. The pre-processed chest X-ray pictures extract relevant properties for feature extraction using CNN. Algorithms for classifying data subsequently include these aspects. They record crucial visual cues indicative of different lung diseases.

One of the most well-known deep learning algorithms is the Convolutional Neural Network (CNN), which has proven particularly effective in recognizing image patterns. CNNs are neurons with adjustable weights and biases, just like the neural networks in your brain. Each neuron processes the weighted sum of the various inputs it receives. The activation function then uses this weighted sum to create an output. Convolution layers, which are a feature of CNNs and are seen in Figure 4, make them different.

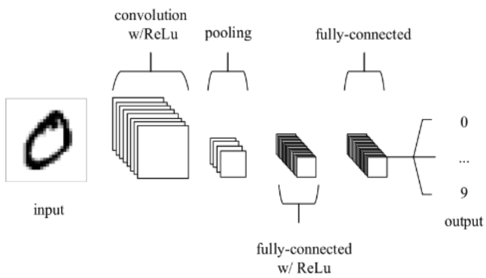


Fig 4 CNN architecture

The three main types of layers in a CNN are convolutional, pooling, and fully connected. In the convolutional layer, the input data is multiplied by a set of weights, a technique known as convolution. This collection of weights is referred to as a kernel or filter. A dot product is produced when a segment of the input information with filter-sized dimensions is multiplied by the filter since the input data's dimensions are greater than the filter's. These dot products are combined to create one value. However, the pooling layer gradually reduces the representation's spatial dimensions, which places a cap on network parameters & computations. This curbing lessens overfitting. The CNN is given a rectified linear unit, which introduces nonlinearity by acting as an element-wise activation function utilized by the output of the preceding layer. For in-depth information on CNNs, see references [36] and [37]. Feature extraction & classification are the two crucial stages of the learning process for CNNs. A feature map is produced due to the convolution of the input data using a filter or kernel during feature extraction. The CNN calculates the likelihood that a picture will be allocated to a particular class or label during the classification step. One of CNNs' outstanding abilities is its ability to recognize and categorize images, deftly combining characteristics without requiring manual feature extraction [38]. Additionally, CNNs can be recycled and altered for usage in various domains through transfer learning [39]. According to [40], this transfer learning approach has led to better categorization outcomes.

The deep convolutional neural network architecture known as VGG, or Visual Geometry Group, has attracted much attention for its effectiveness and simplicity in picture categorization tasks. Its creation was specifically attributed to the Visual Geometry Group at the University of Oxford, hence the name. VGG16, which has 16 weight layers, and VGG19, which has 19 weight layers, are the two most well-known and extensively explored variations of the VGG design and are both shown in Figure 5. The fundamental idea of the VGG architecture is its consistent design philosophy, which uses a series of small convolutional filters (3x3) that have a stride of 1 and a fixed padding of 1, then max-pooling layers (2x2) having a stride of 2. The network may acquire a hierarchical set of features from the input image by continually utilizing this structure.

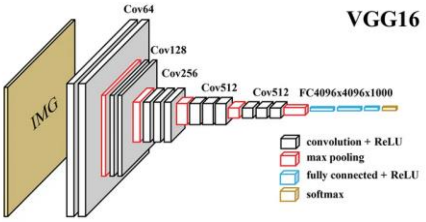


Fig 5 Original architecture of VGG16

Here is a step-by-step breakdown of the VGG architecture, focusing on VGG16:

- Input Layer: The input is an image with predetermined dimensions, such as RGB pictures, which are 224x224x3.
- Convolutional Layers: The input layer is followed by several convolutional layers. The input data is processed using a set of 3x3 filters, commonly referred to as kernels, in each convolutional layer. The filters are trained to identify characteristics like edges, textures, and more intricate patterns.
- Max-Pooling Layers: A max-pooling layer is inserted after each set of convolutional layers. Max-pooling shrinks the feature maps' spatial dimensions, which lowers the computational burden and limits overfitting.
- Fully Connected Layers: After extracting hierarchical features, the convolution and pooling layers combine to form a stack of fully linked layers. These layers incorporate the features discovered from the preceding layers to produce predictions about the image's class.
- Output Layer: Many nodes are present in the final layer as classes are in the classification process. Depending on the issue, the output layer's activation function is chosen. A sigmoid activation is utilized for binary classification, and a softmax activation is used for multi-class situations.

Key characteristics of VGG architecture:

- The consistent use of 3x3 filters and 2x2 max-pooling windows throughout the architecture simplifies the design and implementation.
- The repeated stacking of convolutional and max-pooling layers allows the network to learn increasingly complex features from the input image.
- Despite its simplicity, VGG architectures tend to have many parameters due to their deep stacking. This can lead to longer training times and the risk of overfitting on smaller datasets.

The VGG16 & VGG19 architectures have been used as benchmarks to assess the performance of progressively complex architectures. The VGG may not be the most effective parameter utilization or computing cost compared to more contemporary architectures like ResNet or Inception. However, it is still a significant historical turning point in the evolution of convolutional neural networks. This emphasizes that deeper structures are crucial for improving feature representation and picture classification precision. Regarding the 2014 ILSVR (ImageNet) competition, the VGG-16 won first place. It is often cited as the state-of-the-art in vision model architecture. For VGG-16's training, we turned to the ImageNet database. VGG-16 refers to the number of weighted layers, which is 16. VGG-16's comprehensive training allows it to achieve high accuracy even when presented with relatively tiny image datasets. The VGG-16 image classification model can classify 1000 images into 1000 distinct categories and has an object identification accuracy of 92.7%. This approach to classifying images uses transfer learning and follows the standard format. Learning times for neural networks could be reduced, and model reliability could be increased with the help of batch normalization and additional layers. Overfitting is avoided with the help of dropout layers, the ReLu activation, and the Sigmoid activation in VGG-16.

3.4 Classification

The proposed methodology is tested using two categorization techniques: Support Vector Machine (SVM) and Random Forest. These algorithms identify the retrieved attributes based on the observed patterns and provide a prognosis for each chest X-ray image.

SVM is a classification algorithm [42–43] that focuses on locating the hyperplane in feature space that optimally divides data points from various classes. SVM looks for a hyperplane that minimizes classification errors while maximizing the margin between classes. The data points nearest to the hyperplane, known as support vectors, are crucial in determining this hyperplane. Using the kernel approach, SVM can handle both linearly separable and non-linear data. The original feature space is changed into a higher-dimensional space where data points are easier to distinguish. Polynomial, Gaussian, and linear kernel functions are frequently used (RBF). Using feature vectors derived from chest X-ray images and their accompanying class labels, SVM trains on a labeled dataset. Regulates the trade-off between minimizing the classification error and increasing the margin. Larger values prioritize accurate classification, while smaller values prioritize a broader margin but tolerate some misclassification. Non-linear data separation requires careful consideration when choosing a kernel function. The same extraction procedure used in training is applied to each fresh chest X-ray image by SVM to transform its features into a feature vector. SVM predicts the class label based on which side of the learned hyperplane the feature vector lies.

An ensemble learning system called Random Forest[44] consists of several decision trees, each making a unique prediction. RF creates an ensemble of decision trees by employing a subset of the training data and a subset of features. The ultimate prediction determines which individual tree projections received the most votes. Similar to SVM, RF trains on a labeled dataset. However, it uses arbitrary differences in the data and features they observe to train numerous decision trees. Establishes the forest's number of decision trees. Up until a certain point, increasing this value can improve performance, but after that, it may cause overfitting. Limits each decision tree's depth to avoid overfitting. RF processes the features of a fresh chest X-ray image to produce a feature vector. The class of an image is separately predicted by each decision tree in the Random Forest. The class label with the most support among the decision trees serves as the foundation for the final forecast. Two classification methods, SVM and RF, are used to evaluate the effectiveness of the proposed methodology. These algorithms identify the retrieved attributes based on the observed patterns and provide a prognosis for each chest X-ray image.

3.5 Evaluation metrics

Several publicly accessible lung illness datasets, including the National Institute of Health Chest X-ray Dataset, are used in the methodology's evaluation process. The main objective is to evaluate the method's effectiveness and accuracy in identifying cases of lung illness. When the model applies classification to the data, it generates four possible results, each of which is symbolized differently:

- **True Positive (TP):** This outcome is denoted as TP. It signifies instances where the model accurately predicts positive cases of lung disease. In the evaluation context, these are cases that the model correctly identifies as having a lung disease.
- **True Negative (TN):** The notation for true negatives is TN. This result pertains to instances where the model correctly classifies negative cases, accurately identifying cases without lung disease as negative.
- **False Positive (FP):** Represented as FP, this outcome occurs when the model incorrectly classifies a negative case as positive. In other words, the model predicts the presence of lung disease where there is none.
- **False Negative (FN):** Denoted as FN, this result arises when the model inaccurately classifies a positive case as negative. In this situation, the model fails to identify the presence of a lung disease that is present.

The assessment of the methodology's performance involves the use of various performance metrics, which collectively provide a comprehensive evaluation of its ability to detect different lung diseases accurately:

• Accuracy: This metric, sometimes represented by the "Acc," assesses how well the model generally predicts. The percentage of TP and TN incidences that were accurately predicted is determined.

• Precision: Denoted by "P" or "Prec," precision measures the reliability of the model's affirmative predictions. It is determined by dividing the number of TP occurrences by the total number of TP and FP occurrences.

Recall (sometimes referred to as sensitivity or true positive rate) is symbolized by the letter "R." It measures how well the model can find good examples. It is determined by dividing the number of TP occurrences by the total number of TP and FN occurrences.

• F-measure: Also written as "F1" or "F-measure," this metric considers accuracy and how well the model can remember previous data. The harmonic mean considers false positives and false negatives and is the mean of accuracy and recall.

These performance metrics collectively evaluate the methodology's effectiveness in accurately identifying various lung diseases using the provided datasets. By examining these measures, researchers can learn about the model's true positive, false positive, and false negative rates, as well as other elements of classification accuracy.

4. Results and Discussion

It employed state-of-the-art methods to build the necessary model in MATLAB. The answer relied on 8 GB of RAM and a 2.10 GHz Intel® Xeon® Gold 6130 processor's integrated central processing unit. The system's Graphical Processing Unit also included a Titan RTX graphics card. The suggested High-Dimensional Latent Analysis method was applied to both datasets using several different classifiers. The effectiveness of various solutions was then evaluated using crucial performance metrics like accuracy, recall, precision, and F1-score.

The dataset[34] contains the distribution of images among various classes, each representing a distinct medical condition. It presents the number of images available for each class, including Atelectasis with 508 images, Pneumonia with 62 images, Hernia with 13 images, Edema with 118 images, Emphysema with 127 images, Cardiomegaly with 141 images, Fibrosis with 84 images, Pneumothorax with 271 images, Consolidation with 226 images, Pleural Thickening with 176 images, Mass with 284 images, Effusion with 644 images, Infiltration with 967 images, Nodule with 313 images, and No Finding with the highest count of 3044 images.

Here, we give a high-level summary of the simulation results performed on the dataset [34], which is split into two groups, "normal" and "Disease." The fundamental purpose of these simulations is to investigate and comprehend how various methods affect detection accuracy, precision, recall, and F1-score performance measures. The visual representation of these findings is depicted in Figures 6,7,8 and 9, where each figure corresponds to a specific performance metric. These figures allow for a clear visualization of the outcomes and trends in the context of the techniques employed. During the investigation, we explored two primary approaches for feature extraction: conventional CNN feature extractors and hybrid VGG CNN features.

Furthermore, we applied two different classifiers—SVM and RF. This study's core focus was to assess the impact of implementing a feature scaling technique on automatically extracting CNN features. The rationale behind this investigation is rooted in the recognition that feature scaling, as a preprocessing step, can significantly affect the performance of machine learning algorithms. By applying this technique to extracting CNN features, we aimed to gauge how it influences the overall effectiveness of the classifiers.

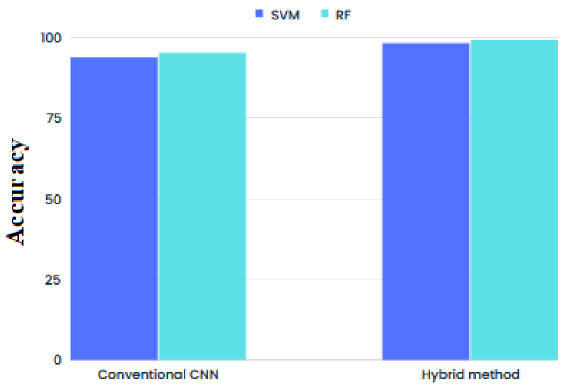


Fig 6 Accuracy analysis

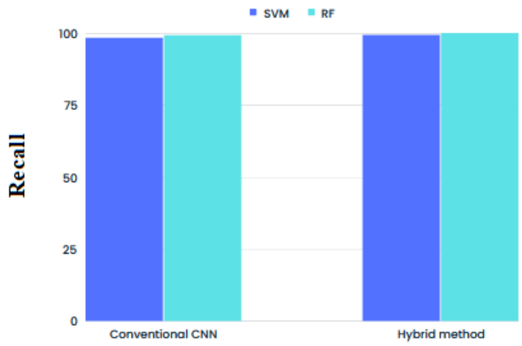


Fig 7 Recall analysis

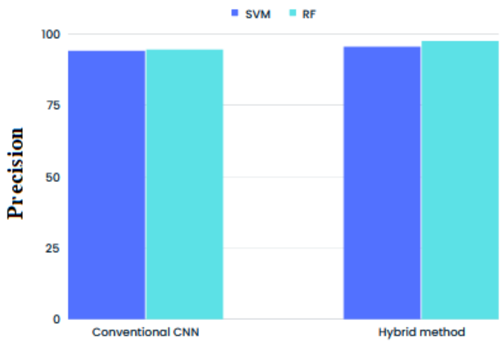


Fig 8 Precision analysis

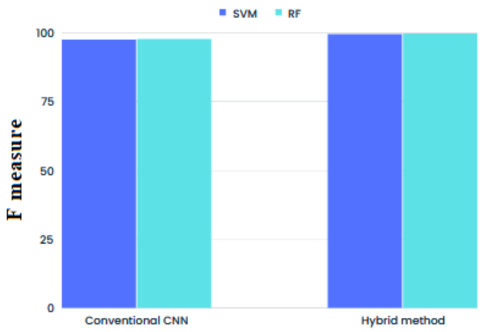


Fig 9 F Measure analysis

This exploration is valuable as it provides insights into the potential improvement of classifier performance by integrating appropriate pre-processing methods. The outcomes of these simulations offer valuable information for refining and optimizing the methodology, contributing to a more accurate and reliable classification of the "normal" and "Disease" classes within the dataset.

Table 1 Comparative Analysis of Performance Metrics for Conventional and Hybrid Methods

Performance metrics	Conventional CNN		Hybrid method	
	SVM	RF	SVM	RF
Accuracy	93.54	94.89	97.89	98.99
Recall	98.1	99.0	99.09	99.8
Precision	93.76	94.25	95.24	97.23
F measure	97.06	97.25	99.08	99.28

The performance metrics for conventional and hybrid methods are compared in Table 1, emphasizing how Support Vector Machines (SVM) and Random Forest (RF) are used in each method. Accuracy, Recall, Precision, and F-measure are among the measures assessed. The SVM and RF models achieved respective Accuracy scores of 93.54% and 94.89% in the case of the traditional Convolutional Neural Network (CNN), according to the results, proving their proficiency in classification tasks. The Recall values of 98.1% for SVM and 99.0% for RF underscore their ability to identify true positive instances effectively. Precision scores of 93.76% and 94.25% highlight the models' proficiency in minimizing false positive classifications. The F-measure, which harmonizes both Precision and Recall, yields values of 97.06% and 97.25%, signifying a balanced trade-off between precision and recall in the conventional approach. Comparatively, the hybrid method, which combines elements of the Convolutional Neural Network with SVM and RF models, showcases enhanced performance. The hybrid SVM-RF model attains an Accuracy score of 97.89%, while the hybrid RF model achieves a remarkable 98.99%, illustrating the superior classification accuracy of these hybrid approaches. Furthermore, the Recall values of 99.09% and 99.8% emphasize the hybrid models' exceptional ability to capture most positive instances. Precision metrics of 95.24% for the hybrid SVM-RF and 97.23% for hybrid RF illustrate their proficiency in making accurate positive predictions.

Notably, the F-measure scores of 99.08% and 99.28% demonstrate the balanced and high-quality nature of the hybrid models' classification outcomes. In summary, the comparative analysis of the performance metrics underscores the advantages of the hybrid approach, where combining Convolutional Neural Networks with SVM and RF models results in notably improved classification performance. Such hybrid models outperform their traditional counterparts on measures of accuracy and resilience for classification tasks, including Accuracy, Recall, Precision, and the F-measure.

5. Conclusion

The research presented here reveals the tremendous promise of hybrid VGG-CNN models for improving the precision with which chest X-ray pictures are used to diagnose lung illness. The hybrid technique has improved performance metrics noticeably by merging CNNs and VGG with Support Vector Machines and Random Forest classifiers. The hybrid SVM-RF model achieved an accuracy of 97.89%, showcasing a 4.35% enhancement over the conventional CNN-SVM accuracy of 93.54%. Similarly, the hybrid RF model achieved an accuracy of 98.99%, surpassing the conventional CNN-RF accuracy of 94.89% by 4.32%. These advancements are further emphasized by the hybrid models' significantly improved recall, precision, and F-measure scores. In terms of future scope, expanding this approach to larger and more diverse datasets, integrating clinical information, and exploring emerging techniques could amplify its impact on precise lung disease diagnosis and contribute to the evolution of global healthcare practices.

References

[1] Bousquet, J. Global Surveillance, “Prevention and Control of Chronic Respiratory Diseases”; World Health Organization: Geneva, Switzerland, 2007; pp. 12–36.

[2] Smith DS, Richey EA, Brunetto WL (2020). "A symptom-based rule for diagnosis of COVID-19". SN Compr Clin Med 2:1947–1954.

[3] Elibol E (2020),“Otolaryngological symptoms in COVID-1”9. Eur Arch Otorhinolaryngol 278:1233–1236.

[4] Padda I, Khehra N, Jaferi U, Parmar MS (2020). "The neurological complexities and prognosis of COVID-19”. SN Compr Clin Med 2:2025–2036.

[5] Sharma R, Agarwal M, Gupta M, Somendra S, Saxena SK (2020), “Clinical characteristics and differential clinical diagnosis of novel coronavirus disease 2019 (COVID-19). In: Saxena S. (eds) coronavirus disease 2019 (COVID-19)”. Medical virology: from pathogenesis to disease control. Springer, Singapore.

[6] Jadhav SP, Singh H, Hussain S, Gilhotra R, Mishra A, Prasher P, Krishnan A, Gupta G (2021),"Introduction to lung diseases. In Dua K, Löbenberg R, Malheiros Luzo ÂC, Shukla S, Satija S (eds) Targeting cellular Signalling pathways in lung diseases”. Springer, Singapore.

[7] Yahiaoui, A.; Er, O.; Yumusak, N.,”A new method of automatic recognition for tuberculosis disease diagnosis using support vector machines”. Biomed. Res. 2017, 28, 4208–4212.

[8] Hu, Z.; Tang, J.; Wang, Z.; Zhang, K.; Zhang, L.; Sun, Q.,”Deep learning for image-based cancer detection

and diagnosis-A survey". Pattern Recognit. 2018, 83, 134–149.
[9] American Thoracic Society. "Diagnostic Standards and Classification of Tuberculosis in Adults and Children." Am. J. Respir. Crit. Care Med. 2000, 161, 1376–1395.
[10] Setio, A.A.A.; Traverso, A.; de Bel, T.; Berens, M.S.; van den Bogaard, C.; Cerello, P.; Chen, H.; Dou, Q.;Fantacci, M.E.; Geurts, B.; et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. Med. Image Anal. 2017, 42, 1–13
[11] Razzak MI, Naz S, Zaib A. "Deep learning for medical image processing: Overview, challenges and the future". Lect Notes Comput Vis Biomech. 2018;26:323–50.
[12] Abiyev RH, Ma'aitah MKS (2018) "Deep convolutional neural networks for chest diseases detection". J Healthc Eng 2018:1–11.
[13] Varela-Santos S, Melin P (2020) "Classification of X-ray images for pneumonia detection using texture features and neural networks". In: Castillo O, Melin P, Kacprzyk J (eds) Intuitionistic and type-2 fuzzy logic enhancements in neural and optimization algorithms: theory and applications. Studies in computational intelligence, vol 862. Springer, Cham.
[14] Wang K, Zhang X, Huang S, Chen F (2019) "Automatic detection of pneumonia in chest X-ray images using cooperative convolutional neural networks". In: Lin Z et al. (eds) Pattern Recognition and Computer Vision. PRCV 2019. Lecture notes in computer science, vol 11858. Springer, Cham
[15] Thakur S, Goplani Y, Arora S, Upadhyay R, Sharma G (2021) "Chest X-ray images based automated detection of pneumonia using transfer learning and CNN". In: Bansal P, Tushir M, Balas V, Srivastava R (eds) Proceedings of international conference on artificial intelligence and applications. Advances in intelligent systems and computing, vol 1164. Springer, Singapore.
[16] Angeline R, Mrithika M, Raman A, Warrier P (2020) "Pneumonia detection and classification using chest X-ray images with convolutional neural network. In: Smys S, Iliyasu AM, Bestak R, Shi F (eds) New trends in computational vision and bio-inspired computing. ICCVBIC 2018. Springer, Cham.
[17] Sarkar R, Hazra A, Sadhu K, Ghosh P (2020) "A novel method for pneumonia diagnosis from chest X-ray images using deep residual learning with separable convolutional networks". In: Gupta M, Konar D, Bhattacharyya S, Biswas S (eds) Computer vision and machine intelligence in medical image analysis. Advances in intelligent systems and computing, vol 992. Springer, Singapore.
[18] Nath M, Choudhury C (2020) "Automatic detection of pneumonia from chest X-rays using deep learning". In:Bhattacharjee A, Borgohain S, Soni B, Verma G, Gao XZ (eds) Machine learning, image processing, network security and data sciences. MIND 2020. Communications in computer and information science, vol1240. Springer, Singapore.
[19] Hashmi MF, Katiyar S, Keskar AG, Bokde ND, Geem ZW (2020) "Efficient pneumonia detection in chest X-ray images using deep transfer learning ."Diagnostics 10(6):417.
[20] Habib N, Hasan MM, Reza MM, Rahman MM (2020) "Ensemble of CheXNet and VGG-19 feature extractor with random Forest classifier for pediatric pneumonia detection". SN Comput Sci 1:359.
[21] Abbas A, Abdelsamea MM, Gaber MM (2020) "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network". Appl Intell 51:854–864.
[22] Dansana D, Kumar R, Bhattacharjee A, Hemanth DJ, Gupta D, Khanna A, Castillo O (2020) "Early diagnosis of COVID-19-affected patients based on X-ray and computed tomography images using deep learning algorithm". Soft Comput 27:2635–2643.
[23] Apostolopoulos ID, Mpesiana TA (2020) "Covid-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks". Phys Eng Sci Med 43:635–640.
[24] Turkoglu M (2020) "COVIDetectioNet: COVID-19 diagnosis system based on X-ray images using features selected from pre-learned deep features ensemble". Appl Intell 51:1213–1226.
[25] Hira S, Bai A, Hira S (2020) "An automatic approach based on CNN architecture to detect Covid-19 disease from chest X-ray images". Appl Intell 51:2864–2889.
[26] Chen B, Li J, Lu G, Yu H, Zhang D (2020) "Label co-occurrence learning with graph convolutional networks for multi-label chest X-ray image classification". IEEE J Biomed Health Inf 1–1.
[27] Yamac M, Ahishali M, Degerli A, Kiranyaz S, Chowdhury MEH, Gabbouj M (2021) "Convolutional sparse support estimator-based COVID-19 recognition from X-ray images". IEEE Trans Neural Netw Learn Syst 32(5):1810–1820.
[28] Ohata EF, Bezerra GM, das Chagas JVS, Lira Neto AV, Albuquerque AB, de Albuquerque VHC, Reboucas Filho PP (2020) "Automatic detection of COVID-19 infection using chest X-ray images through transfer learning". IEEE/CAA J Autom Sin 8:1–10.
[29] Nishio M, Noguchi S, Matsuo H, Murakami T (2020) "Automatic classification between COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy on chest X-ray image: a combination of data augmentation methods". Sci Rep 10(1):17532.
[30] Hemdan, E.E.D.; Shouman, M.A.; Karar, M.E. "Covidx-net: A framework of deep learning classifiers

to diagnose COVID-19 in X-ray images". arXiv 2020, arXiv:2003.11055.
[31] Narin, A.; Kaya, C.; Pamuk, Z. "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks". Pattern Anal. Appl. 2021, 24, 1207–1220.
[32] Brunese, L.; Mercaldo, F.; Reginelli, A.; Santone, "A. Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays". Comput. Methods Programs Biomed. 2020, 196, 105608.
[33] Ghoshal, B.; Tucker, A. "Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection". arXiv 2020, arXiv:2003.10769.
[34] NIH sample Chest X-rays dataset. https://www.kaggle.com/nih-chest-xrays/sample.
[35] Padda I, Khehra N, Jaferi U, Parmar MS (2020) The neurological complexities and prognosis of COVID-19. SN Compr Clin Med 2:2025–2036.
[36] O'Shea, K.; Nash, R. An Introduction to Convolutional Neural Networks. arXiv 2015, arXiv:1511.08458v2.
[37] Ker, J.; Wang, L. Deep Learning Applications in Medical Image Analysis. IEEE Access 2018, 6, 9375–9389.
[38] Al-Ajlan, A.; Allali, A.E. CNN—MGP: Convolutional Neural Networks for Metagenomics Gene Prediction. Interdiscip. Sci. Comput. Life Sci. 2019, 11, 628–635.
[39] Pan, S.J.; Yang, Q. A Survey on Transfer Learning. IEEE Trans. Knowl. Data Eng. 2010, 22, 1345–1359.
[40] Lopes, U.K.; Valiati, J.F. Pre-trained convolutional neural networks as feature extractors for tuberculosis detection. Comput. Biol. Med. 2017, 89, 135–143
[41] Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.
[42] A. M. Alqudah, and M. A. Alqudah, "Sliding Window Based Support Vector Machine System for Classification of Breast Cancer Using Histopathological Microscopic Images," IETE Journal of Research, vol. 7, pp. 1-9, 2019, doi: 10.1080/03772063.2019.1583610.
[43] A. M. Alqudah, "Ovarian Cancer Classification Using Serum Proteomic Profiling and Wavelet Features A Comparison of Machine Learning and Features Selection Algorithms," Journal of Clinical Engineering, vol. 44, no. 4, pp. 165-173, 2019, doi: 10.1097/JCE.0000000000000359.
[44] M. A. Alqudah, A. Albadarneh, I. Abu-Qasmieh, and H. Alquran, "Developing of robust and high accurate ECG beat classification by combining Gaussian mixtures and wavelets features," Australasian physical & engineering sciences in medicine, vol. 42, no. 1, pp. 149-157, 2019, doi: 10.1007/s13246-019-00722-z