World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(RESEARCH ARTICLE)

Check for updates

Performance analysis of deep learning architectures for chest disease and lymphoma classification

Mustafizur Rahman Shakil [1], Mehedi Hasan [2], Mohammed Imam Hossain Tarek [2], Fakhru Islam Polash [2], and Erin Jahan Meem [3, *]

[1] *Department of Engineering Management, Westcliff University, Irvine, CA 92614, USA.*
[2] *Department of Business Administration, International American University, Los Angeles, CA 90010, USA.*
[3] *Department of Computer Science, Pacific States University, Los Angeles, CA 90010, USA.*

World Journal of Advanced Engineering Technology and Sciences, 2025, 17(01), 233-247

Publication history: Received on 01 September 2025; revised on 06 October 2025; accepted on 08 October 2025

Article DOI: https://doi.org/10.30574/wjaets.2025.17.1.1392

Abstract

Integrating deep learning and federated learning models, the study investigates the detection of cancer and chest diseases. The researchers used a combination of 112,120 chest X-ray photographs annotated by the NIH and 5,400 lymphoma biopsy images from Kaggle to identify three distinct forms of lymphomas. The method's first aims are to improve images, normalize data, and identify features. Afterwards, it evaluates features that rely on contours, such as aspect ratio solidity and intensity fluctuations. The study included seven industry-standard models and federated learning techniques. The InceptionV3, MobileNetV2, DenseNet161, ResNet50, VGG-19, and VGG-16 models are included here. To measure performance, we employ F1-score, recall, accuracy, and computational efficiency. Although InceptionV3 dominated in terms of loss and root-mean-squared error, DenseNet161 had the best accuracy among deep learning models used to detect chest illnesses at 88.01%. for compared to other models, VGG-19 had a superior accuracy rate of 97.5% for classifying lymphomas. The newly integrated models that succeeded were InceptionV3 and VGG-19, which had 95.8% accuracy in diagnosing lymphoma and 97.7% accuracy in detecting chest illnesses. With the use of deep and federated learning algorithms to medical imagery, automated illness detection got increasingly accurate.

Keywords: Deep Learning; Chest Disease Detection; Feature Extraction; Medical Image Analysis; Federated Learning; Transfer Learning

1. Introduction

The thoracic cavity contains various organs which fall under the category of chest diseases that medical science classifies into a wide range of ailments. The human body faces numerous health problems in these regions and additional structures of the thoracic cavity [1]. Since chest diseases present multiple interwoven anatomical components, it becomes essential to use a combination of approaches to manage them effectively. Medical detection of chest diseases becomes possible through three primary diagnostic methods which include CT scans and MRIs along with pulmonary function testing [2], [3].

New technology from artificial intelligence helps deep learning models to increase disease identification accuracy together with operational speed. AI systems using medical image inputs achieve high detection precision to identify diseases at an early stage thus enabling combined diagnoses of pneumonia along with tuberculosis leukaemia and cardiomegaly [4]. Predictive diagnostics receives enhanced effect from neural networks through their analysis of patient information and imaging data which improves healthcare delivery results. The medical diagnostics field experienced a breakthrough with deep learning models that deliver automated picture analysis at very high standards

* Corresponding author: Erin Jahan Meem

of accuracy. The fundamental role of these models particularly CNNs and RNNs enables them to analyse medical imagery effectively [5]. Image segmentation and classification together with anomaly detection form the main use cases for which CNNs have become essential tools in medical diagnostics. The interpretation of X-rays and CT scans by human specialists for chest disease detection through traditional methods shows errors and depends on differing levels of medical expertise. DL models enhance diagnosis accuracy through their ability to find precise patterns and anomalies which results in improved analysis [6], [7]. Rudimentary acceptance of this technology is blocked by the requirement of extensive data and processing capacity as well as navigability issues within these models.

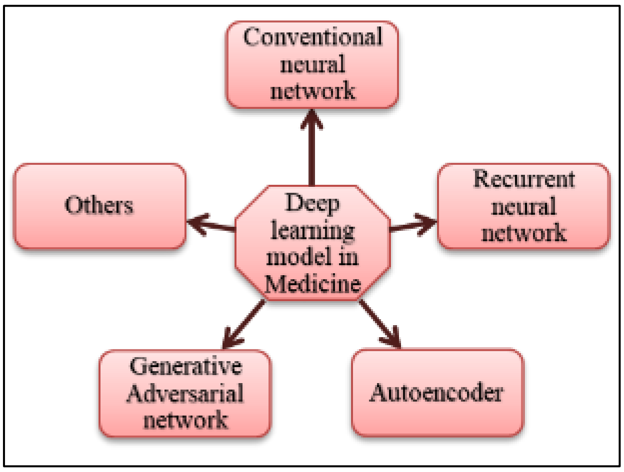

Figure 1 Representation of Deep learning models in medical science

Figure 1 presents the different deep learning model types applied in medical science. Latent-dimensional reduction functions of autoencoders work alongside the image-processing capabilities of CNNs and sequential analysis from RNNs in addition to GAN and hybrid model technology addressing challenges [8]. The worldwide rise of chest diseases results from exposure to pollutants together with exposure to infectious pathogens and tobacco smoke consumption. Multiple chest diseases which affect patients include pneumonia, atelectasis, pleural diseases, consolidation, infiltration, pneumothorax and enema together with emphysema, fibrosis, effusion, cardiomegaly, nodules and masses. As an example, pneumonia produces alveolar inflammation that results in fluid buildup. Atelectasis refers to lung collapse due to airway obstruction or external pressure. The function and elasticity of lungs become adversely affected when patients have pleural diseases and fibrosis which results in serious respiratory complications. It is fundamental to image-based diagnosis since proper understanding of these diseases enables prompt medical treatment.

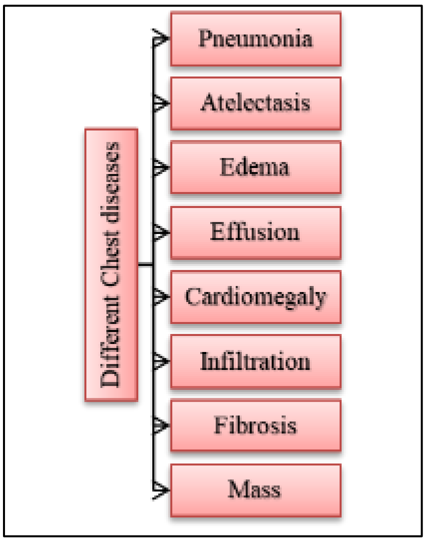

Figure 2 Outline of the types of chest diseases

The diagram in Figure 2 outlines multiple chest diseases by their origins alongside their effects on lung functioning. The identification of chest conditions during early stages leads to decreased mortality statistics. Machine learning (ML) and deep learning algorithms perform advanced detection along with classification of various disorders in a significant manner [9]. Chest X-rays function as the principal diagnostic method because they provide effective results and can be easily accessed. CT scanners and MRI machines enable complex disease analysis which results in detailed descriptions of diseases. AI algorithms evaluate X-ray and CT images to detect tuberculosis, pneumonia and lung cancer by achieving accurate identification [10]. These detection models surpass traditional approaches because they recognize subtle abnormalities which people viewing X-rays normally miss. AI diagnostic tools possess the capability to analyse large data sets with rapid efficiency which enables clinicians to conduct real-time decisions in medical facilities. Modern medical diagnostic tools advanced by AI in medical imaging now perform chest disease detection with high accuracy levels. When artificial intelligence methods employing neural networks and transfer learning techniques are combined, the accuracy of illness detection is significantly raised.

Federated learning provides a new approach that safeguards privacy during the collaborative AI model training process between multiple institutions [11]. Multiple hospitals and research centres can contribute to AI progress through this decentralized approach as it ensures they do not need to share protected patient data. Deep learning technology has automated medical image analysis which enables fewer need for human interpreters to work on these diagnostic tasks. VGG-16 ResNet50 along with DenseNet161 represent several pretrained neural networks that detect diseases widely in various applications [12], [13]. Transfer learning in the models enhances both the diagnostic accuracy and decreases the computational needs. Deep learning operates unlike traditional ML approaches because it needs little or no feature engineering to function successfully [14]. The networks acquire their hierarchical data representations from the data they process directly and produce very effective image pattern identification within medical contexts. Advantages from self-contained design led to better diagnostic tasks and efficient healthcare outcomes.

Hospitals and institutions now use distributed model training through federated learning to revolutionize how they deploy artificial intelligence systems for healthcare. Traditional centralized models differ from FL because data stays locally while this system permits only model updates to be exchanged for patient protection. A central server within federated learning distributes the first version of a model to participating clients who apply their individual datasets for local training. Multiple institutions update their local models to refine the global model allowing improved predictive capability and keeping data private in the process. The method provides exceptional benefits to chest disease detection by allowing access to various datasets which enhances model robustness and generalizability properties. This research unites deep learning with federated learning to develop an AI-driven medical diagnostic method which presents scalable protection of patient data during chest disease detection processes.

2. Related Works

Recent breakthroughs, such as regenerative wound healing [15], hybrid nanoconjugates for glioblastoma [16], and molecular erasers for cancer immunity [17], highlight the power of advanced technologies in healthcare. In parallel, artificial intelligence has shown transformative potential across domains—improving precision agriculture through self-driving systems [18], advancing road segmentation via attention models [19], enabling Alzheimer's diagnosis with hybrid CNN-SVM methods [20], predicting machine failures with multimodal learning [21], and enhancing plant disease detection through deep stacking models [22]. Together, these innovations illustrate how AI and biomedical research converge to address complex challenges in critical sectors.

AI applications for chest disease detection now receive wide acceptance due to their dual capability to cut down misdiagnosis and provide support for radiologists in their diagnostic choices. CRIS models that utilize large databases of medical images exhibit extraordinary precision when detecting different kinds of chest infections and abnormalities [10]. Different AI technology approaches have been studied with deep neural networks, transfer learning as well as hybrid AI model use. CNN technology has become essential for medical imaging analysis since it excellently identifies complex radiographic patterns. Studies have shown that selected models provide excellent diagnostic results which validate artificial intelligence capabilities in medical settings.

Applications of deep learning and related machine learning techniques have revolutionized the solution of medical image problems [23]. Deep neural systems enable automatic process of extracting helpful information without requiring human interpretation. The training of deep learning models linked with big chest X-ray datasets enables them to successfully detect various pulmonary diseases with high accuracy levels [24]. The field of health research applies transfer learning by adjusting pre-trained CNN networks including ResNet, VGG16 and InceptionV3 for better diagnosis outcomes. Machine learning plays an essential role in respiratory medicine because certain models demonstrate high accuracy levels. The analysis includes support vector machines (SVMs) and long short-term memory (LSTM) networks in addition to CNNs for ML exploration. Several techniques have found their applications throughout multiple aspects of detecting chest diseases particularly through lung sound evaluation and airflow tracking and disease development forecasting [25]. Multiple computers learning methods can achieve better predictive competencies through integrative model combinations according to roundtable analysis.

Deep learning techniques have proved essential for increasing the precision of chest disease diagnosis through their applications. Modern neural network systems analyze chest radiographs to detect pneumonia and fibrosis and tumors inside the body. Multiple studies have applied convolutional neural networks and autoencoders and graph convolutional networks for imaging analysis and chest classification purposes [26]. Research groups have created specific CNN designs that deliver maximum accuracy results when conducting multi-class and binary classifications. Deep learning technology achieved success in chest disease diagnosis by creating new neural network designs DC-ChestNet and VT-ChestNet. The models have received training through large-scale dataset samples to identify normal patients from those with lung diseases or heart ailments [27]. The VT-ChestNet model stands out because it demonstrates excellent performance by reaching an area under the curve (AUC) score higher than 95% in multi-class classification tasks.

The global COVID-19 pandemic caused healthcare institutions to fast-track their implementation of AI solutions in medical diagnostics. The detection of COVID-19 pneumonia through chest X-ray and CT scan images has been accomplished using intelligent AI-based models [28]. The development of AI-driven COVID-19 detection systems heavily depends on CNNs together with transfer learning strategies. These models show great accuracy results which enable quick screening and early diagnosis processes. According to research deep learning systems training on COVID-19 datasets reveal the ability to discover COVID-19 at high sensitivity together with specificity rates in almost real-time. Artificial intelligence diagnosis systems using patient clinical data from oxygen saturation tests and lung function analysis achieve higher accuracy in medical diagnoses [29]. Federated learning serves as a research area to enable AI training in multiple healthcare institutions without compromising patient privacy of data.

AI applications extend beyond medical image analysis by being used for lung sound analysis in respiratory disease diagnosis. Deep learning algorithms receive training from audio data of lung sounds to detect asthma as well as bronchitis and COPD among other conditions [30]. These diagnostic systems examine airborne sound patterns to detect abnormal patterns which supports non-invasive assessment of chest diseases. Studies reviewed 160 publications which focused on deep-learning-based lung sound analysis. Research results confirm that artificial intelligence models achieve accurate classification of lung sounds for diagnosing pulmonary conditions at an early stage [31]. New advancements include the coupling of artificial intelligence with electronic stethoscopes which enhances respiratory sound analysis efficiency thus making it suitable for use in telemedicine.

The implementation of AI technologies for chest disease diagnosis continues to meet various obstacles even with existing advancements. The main restriction in AI deployment arises from the insufficient use of diverse standardized datasets [32]. AI systems can only deliver accurate diagnoses when trained on specific datasets even though they become less effective when used with different groups of patients. The model performance suffers whenever training datasets have an unbalanced distribution of diseases. AI models currently present difficulties when it comes to their ability to show interpretable outputs. Deep learning systems operate as opaque mechanisms which prevent healthcare providers from understanding their determination logic [33]. Healthcare professionals need transparent models for clinical acceptance since interpretability allows them to validate diagnostic results. The rendition of AI algorithms sometimes contains systemic biases which might endanger medically just healthcare [34]. Research reveals that AI systems receiving input data which contains bias create mistakes for population segments that lack proper representation in training samples. Fair and reliable AI-based diagnosis needs ethical standards for handling biases and collection of datasets that represent diverse patient populations [35].

The integration of AI into medical imaging through machine learning makes chest disease diagnosis more accurate and efficient along with being capable of large-scale implementation. Lung disease classification from radiographic images succeeds at the highest level through deep learning models whose main component are convolutional neural networks [36]. The breast cancer detection using AI has reinforced the diagnostic capabilities of artificial intelligence for medical purposes [37]. AI deployment in healthcare requires solving problems with biased data and lack of interpretability and privacy limitations to secure responsible implementation. AI research progress gives reason to expect positive changes in chest disease diagnosis and custom treatment plans for the near future [38].

Deep learning and federated learning provide substantial potential to improve healthcare results during chest disease detection processes. This research aims to

- A deep learning system should be created to recognize chest diseases successfully within X-ray and CT imaging data.
- Netted training happens by using federated learning together with data privacy protection features.
- The implementation of advanced AI techniques should focus on improving diagnostic efficiency together with precision of detection.

The study works to solve major medical imaging problems which supports the development of advanced diagnostic systems that are both more secure and efficient.

3. Methodology

The quick progress of artificial intelligence (AI) technology has reshaped healthcare through its substantial influence on medical image examination. Medical detection of pneumonia along with lung fibrosis and lymphoma needs fast and precise detection methods to achieve better patient results. Medical diagnostic procedures require significant time and show subjective evaluation results. A deep learning framework utilizing federated learning and different pre-trained models and feature extraction approaches proposes to detect chest diseases together with lymphoma. The Figure 3 shows the illustration of the proposed approach for detecting the chest diseases.

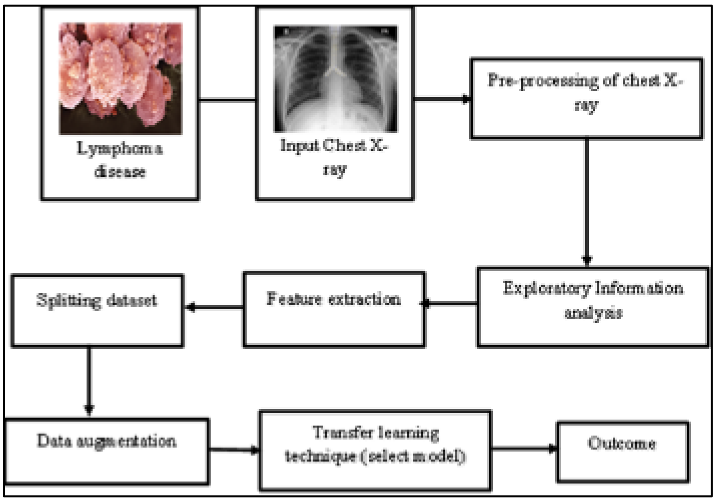

Figure 3 Illustration of the proposed approach

3.1. Dataset Description

The research uses the NIH Chest X-ray dataset as well as the Malignant Lymphoma Classification dataset from Kaggle which serves as the primary sources. This NIH dataset includes 112,120 chest X-ray images with 14 different disease annotations ranging from pneumonia to pneumothorax and fibrosis to atelectasis to edema. A structured labeling system based on Natural Language Processing technology extracts medical report information regarding diseases. All patient files contain metadata links that include demographic information about age, gender, diagnostic type, and imaging protocol details. The set provides bounding box annotations which allow viewers to precisely identify regions of disease appearance in order to enhance area localization.

Histopathological images from many pathologies centers total 5,400 samples which medical professionals grouped into the three lymphoma types Chronic Lymphocytic Leukemia (CLL), Follicular Lymphoma (FL), and Mantle Cell Lymphoma (MCL). The staining with Hematoxylin and Eosin (HandE) technique helps enhance image contrast so cellular structures become easier to view. The RGB format with high resolution provides the dataset for enhanced features extraction capabilities. The extensive dataset enables deep learning models to learn many patterns which relate to chest diseases alongside lymphoma characteristics. The research depends on two different datasets to build diagnostic capabilities for disease identification in X-ray and biopsy images because deep learning proves its potential for medical imaging data sets.

3.2. Preprocessing Techniques

The deep learning models undergo extensive preprocessing steps that both improve image quality as well as normalize inconsistent data in the dataset before training. Applying a standardized resolution to every image constitutes the first part of preprocessing before deep learning model training. The procedure begins by resizing Chest X-ray pictures to 224 × 224 pixels and setting lymphoma images at 1388 × 1400 pixels. The next stage applies normalization to convert all pixel values between zero and one so that every image possesses equal intensity ranges. The training process reaches convergence speedily because of this technique.

Gaussian filtering smooths medical images without distorting essential edges in order to eliminate background noise. The Gaussian function exists as follows:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

The standard deviation σ defines the width of the Gaussian kernel in this formula. The pre-processing contains a filtering function to remove small variations together with imaging artifacts which preserves the quality of feature extraction.

The process of image enhancement through contrast enhancement plays an essential role when working with chest X-ray images because it helps reveal important disease characteristics which may otherwise stay hidden owing to weak contrast levels. The process of equalizing pixel intensities with histogram equalization improves image contrast in the data. The processing of missing metadata values within the dataset occurs either through substitution techniques or by taking out incomplete records. Last but not least in the preprocessing scheme are categorical variables that receive numerical encoding which enables their compatibility with deep learning frameworks. By implementing preprocessing steps to the dataset its quality improves which results in better model performance and generalization during classification.

3.3. Data Augmentation

The medical image datasets show substantial variations regarding disease samples distribution where specific conditions possess significantly less image data than others. The data augmentation technique creates artificial images through modifications which apply to existing dataset images. The augmentation techniques generate new image versions while the learning process focuses on identifying features which remain consistent throughout. This leads to better generalization together with improved robustness.

The applied data augmentation strategies include both rotation and translation methods as well as horizontal flipping and scaling operations and zoom functions. Images get rotated by small angular values (±5∘) as a method to correct slight positioning differences between X-ray and biopsy images. The translational movements apply to both x- and y-axes to keep disease characteristics visible while maintaining small repositioned images. The system applies horizontal image flipping at random to X-ray images for simulating patient position differences. The image features are refined by scaling and zooming operations while maintaining slightly modified dimensions because this method allows the model to develop mastery over various levels of image magnification.

The mathematical definition of augmentation transformation appears as follows

$$I' = SRTI \quad (2)$$

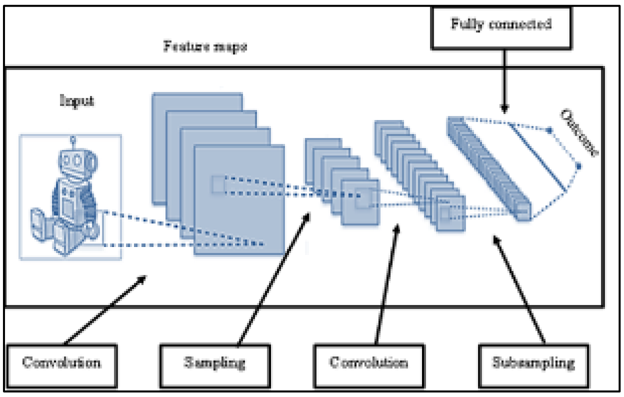

Figure 4 Deep Learning Architecture for Chest Disease Classification

The image transformation produces I′ with scaling factor S and rotation matrix R and translation T. The images used for lymphoma contain color jittering because this method randomly alters brightness levels and contrast together with saturation conditions to replicate different staining variances noticed in biopsy examples. The augmented data elements enlarge the dataset diversity to reduce overfitting while boosting model performance during real medical image evaluations.

Figure 4 illustrates a Convolutional Neural Network (CNN) design that applies convolutional layers for feature extraction after which it moves on to pooling and fully connected layers for classification. Medical image processing within CNNs reveals their ability to detect chest disorders as well as distinguish different lymphoma categories.

3.4. Feature Extraction

Medical image classification highly depends on feature extraction because it enables deep learning models to recognize important image patterns. The extraction process of morphological and texture-based features helps improve disease characteristic understanding. The shape characteristics and dimensional information that concern disease-affected areas form part of the morphological features extracted from the images. These include

Area of the disease region

$$Area = height \times width \qquad (3)$$

The boundary length of disease regions is captured as the measurement known as perimeter.

$$Per = \sum\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \qquad (4)$$

In Aspect Ratio, healthcare practitioners track the stretch length of medical conditions at the target site.

$$Aspect\ Ratio = \frac{width}{height} \qquad (5)$$

The compactness of a lesion is determined by filling a syringe with its material while measuring the volume.

$$Solidity = \frac{Area}{Convex\ Hull\ Area} \qquad (6)$$

Texture-Based Features: The features describing pixel intensity variations that exist inside disease regions fall under the category of texture features. This measure evaluates the standard measurement of grayscale value across an area.

$$I_{mean} = \frac{1}{N}\sum_{i=1}^{N} I_i \qquad (7)$$

Histogram equalization enhances contrast:

$$H(I) = \frac{(I - I_{min})}{(I_{max} - I_{min})} \times 255 \qquad (8)$$

The features of CNNs enable them to concentrate on disease-relevant patterns which enhances their ability to classify medical images effectively.

3.5. Model Selection and Training

This research uses deep convolutional neural networks (CNNs) since they demonstrate high efficiency in image classification tasks. Each CNN organizes into three essential layers that consist of convolutional layers and pooling layers and fully connected layers. The convolution procedure involves the following definition

$$O(i, j) = \sum_m \sum_n I(i + m, j + n) \cdot K(m, n) \qquad (9)$$

The process uses three variables namely input image I(i,j), filter kernel K(m,n) and output feature map O(i,j). SGD and backpropagation methods enable training the models to reduce classification errors.

3.6. Transfer Learning Models

Deep learning utilizes transfer learning as a strong method that adapts a large dataset-trained model for applications in related but different contexts. Deep neural networks receive better performance after pre-trained ImageNet models are adjusted to analyze medical pictures. The training method results in better performance together with reduced training duration and successful resolution of the low-number of labeled medical data challenge.

This research utilized six state-of-the-art transfer learning models namely VGG-16, VGG-19, ResNet-50, DenseNet-161, MobileNetV2, and InceptionV3 for detecting chest diseases and lymphomas. Each architecture within these models provides features that aid extraction and classification processes.

VGG-16 (Visual Geometry Group-16): The Visual Geometry Group (VGG) at the University of Oxford has created VGG-16 as their deep CNN model. This network contains sixteen layers starting with thirteen convolutional layers before ending in three fully connected layers. Feature extraction at the model occurs through the combination of 3×3 convolutional filters and small receptive field elements and max pooling layers.

A REL activation function follows each convolutional layer for the purpose of creating non-linear behavior.

$$f(x) = \max(0, x) \qquad (10)$$

The last classification layer includes a softmax activation function for its operation.

$$P(y = j|x) = \frac{e^{z_j}}{\sum_k e^{z_k}} \qquad (11)$$

VGG-19: VGG-19 expands upon VGG-16 by adding more layers to reach a total of 19 between convolutional and fully connected components. The model includes extra depth to its features since it employs the same 3×3 convolutional filters and max pooling layers.

ResNet-50 (Residual Network-50): ResNet-50 introduced residual learning for deep network training (50 layers) through the solution of the vanishing gradient problem. Skip connections are the fundamental concept of ResNet because they direct gradient propagation across layers for faster network convergence.

Res Net trains itself to solve a residual function instead of designing a function H(x).

$$F(x) = H(x) - x \qquad (12)$$

The output of a residual block is:

$$y = F(x) + x \qquad (13)$$

DenseNet-161 (Densely Connected Convolutional Networks): DenseNet-161 represents an advanced CNN design which builds feature propagation through sequential layer connections between each element during model forward processing. The DenseNet architecture implements feature map concatenation instead of residual connection methods like ResNet does to improve efficiency in feature reuse.

For layer l, the output is

$$x_l = H_l([x_0, x_1, \cdots x_{l-1}]) \qquad (14)$$

MobileNetV2: The solution creators designed MobileNetV2 specifically for efficient deep learning models at low power because this model works well for medical imaging tasks on limited-capacity systems. Depthwise separable convolutions form a key part of MobileNetV2 by decreasing computational requirements without sacrificing performance quality.

Instead of a standard convolution

$$Y = X * K \qquad (15)$$

MobileNetV2 factorizes it into

$$Y_{depthwise} = X * K_d \quad (16)$$

$$Y_{pointwise} = Y_{depthwise} * K_p \quad (17)$$

InceptionV3: The efficiency of CNN depends on InceptionV3 because it conducts parallel convolution layers and multi-scale feature extraction. The network employs parallel 1×1, 3×3 and 5×5 convolution operations to obtain pattern recognition at varying scales as an alternative to using standard filter size.

Given an input X, the output is computed as

$$Y = \sum_{i=1}^{N}(X * K_i) \quad (18)$$

Here Ki = filters, Kd = depthwise filter, Kp = pointwise filter, xl = outcome, Hl = transformation, $[x_0, x_1, \cdots x_{l-1}]$ = concatenation of previous feature, x = input, F(x) = learned residual mapping, y = outcome,

3.7. Federated Learning for Privacy Preservation

The training process in Federated learning takes place on various devices locally without revealing actual data to protect privacy. During the aggregating process through Federated Averaging (FedAvg) the algorithm unites model updates sent by multiple device clients.

$$w_{t+1} = \sum_{i=1}^{k} \frac{n_i}{n} w_i \quad (19)$$

Each client provides their model updates wi while also reporting their sample count ni. Hospitals can train models together through this approach which stands as a guarantee for patient data confidentiality.

The combination of deep learning along with federated learning enhances disease identification by improving medical diagnostic security through controlling personal information access.

The research contains medical image classification segments using deep learning with federated learning techniques which enhance privacy and disease detection precision. A prediction enhancement model utilizes CNNs together with transfer learning and morphological feature extraction through the proposed approach.

4. Result Analysis

The results section provides an extensive analysis of the proposed deep transfer learning and federated learning models which classify chest diseases and lymphomas. Analysis of effectiveness measures accuracy, loss, precision, recall, F1-score, and AUC provides an assessment of model performance. The analysis determines medical diagnostic optimal models by measuring different architectural efficiency. Visual model performance evaluation for model effectiveness is enabled using AUC curves and loss-accuracy plots. The research investigates multiple threshold parameters together with dataset variations to validate the proposed method through accuracy validation. Accuracy demonstrates the number of correctly classified instances among all available instances in a particular study. It is given by

$$Acc = \frac{Correct\ positive + Correct\ negative}{Total\ samples} \quad (20)$$

Precision evaluates the number of truly positive cases among all those cases that the model predicts positively. The system needs to perform well in situations which require a low rate of false positive results.

$$Pre = \frac{Correct\ positive}{correct\ positive + Incorrect\ positive} \quad (21)$$

Recall demonstrates which actual positive cases exist and correctly get detected. Medical diagnosis requires immediate detection of all positive cases because missing a single one constitutes a false negative.

$$Rec = \frac{Correct\ positive}{correct\ positive + Incorrect\ negative} \quad (22)$$

The F1-score calculates precision and recall with an equation that computes their harmonic mean for balanced evaluation. The metric finds its best application during work with data that contains imbalanced distributions.

$$FS = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (23)$$

AUC-ROC provides insight into model discrimination ability through the representation of True Positive Rate vs False Positive Rate at different threshold choices.

$$TPR = \frac{Correct\ positive}{Correct\ positive + Incorrect\ negative} \quad (24)$$

$$FPR = \frac{Incorrect\ positive}{Incorrect\ positive + Correct\ negative} \quad (25)$$

The evaluation of predicted and actual value divergence is accomplished through loss functions. Two common loss functions are

$$MSE = \frac{1}{m}\sum_{k=1}^{m}(actual\ value\ -\ predicted\ value) \quad (26)$$

$$EL = -\sum_{i=1}^{m} b_i \log(\hat{b}_i) \quad (27)$$

The error quantity known as RMSE functions as another metric to calculate actual-predicted value discrepancies. The loss function assigns greater weight to large deviations above and below the true value than it does to MSE.

$$RMSE = \sqrt{\frac{1}{m}\sum_{k=1}^{m}\left(b_i - \hat{b}_i\right)^2} \quad (28)$$

MCC maintains balanced measurements for datasets with any imbalance level. TMCC presents a full performance evaluation system by using TP and TN together with FP and FN.

$$MCC = \frac{(CP \times CN) - (IP \times IN)}{(CP+IP)(CP+IN)(CN+IP)(CN+IN)} \quad (29)$$

Here CP = correct positive, CN = correct negative, IP = incorrect positive, IN = incorrect negative, m = total number of information points, bi = actual value, $\hat{b}_i$ = Predicted value, m = total samples.

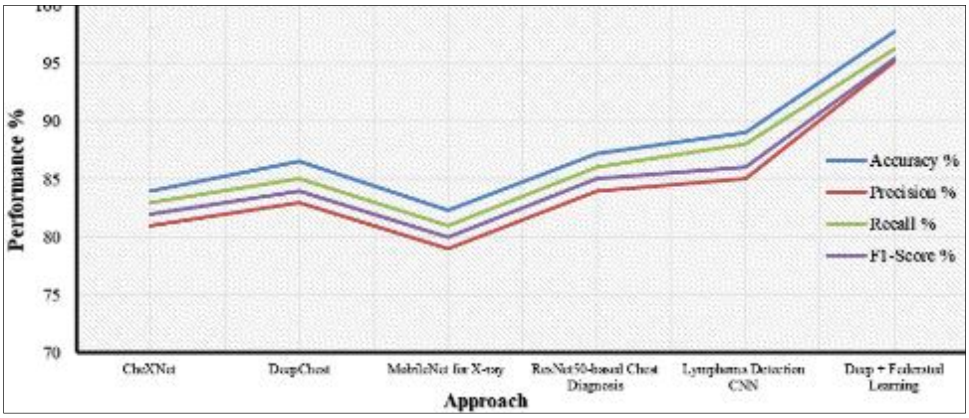

Figure 5 Visualization of compared performance of existing approach with suggested approach

The performance metrics display that the Deep + Federated Learning method ranks higher than traditional methods for classifying chest diseases and lymphomas is shown in Figure 5. X-ray models using CheXNet, Deep Chest, and Mobile Net for X-ray achieve diagnostic accuracy from 82.3% to 87.2% whereas the Lymphoma Detection CNN stands at 89%. The proposed method demonstrates superior diagnostic capability through its 97.7% accuracy which surpasses other systems. Deep + Federated Learning produces outstanding results through its precision of 95.2% and recall of 96.2%

and F1-score of 95.4% which surpasses the Lymphoma Detection CNN's performance indicators (85%, 88% and 86% respectively). The proposed detection system exhibits 96.2% recall which maintains maximum sensitivity toward disease diagnosis thus making it ideal for medical practice. Using deep learning in conjunction with federated learning allows the proposed model to safeguard patient data privacy while boosting classification accuracy which leads to becoming an efficient advanced instrument for detecting diseases in real medical environments.

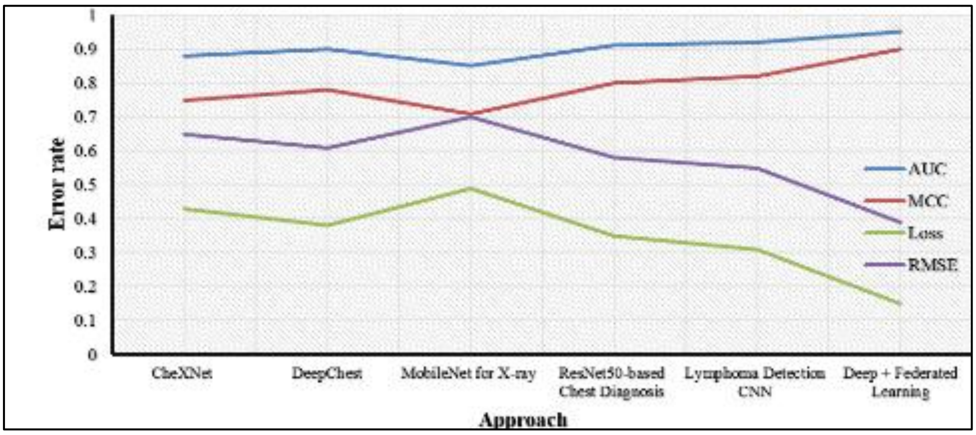

Figure 6 Visualization of compared error rate of existing approach with suggested approach

The combination of Deep with Federated Learning produces enhanced performance results across each essential assessment metric better than other existing techniques is shown in Figure 6. The AUC value of 0.95 stands as the highest among all present methods which demonstrates outstanding discrimination power for separating diseased from non-diseased cases. The MCC (0.90) produces superior results than all other tested models because it demonstrates robust performance in maintaining a strong balance for true positive and true negative identifications within imbalanced datasets. The proposed model performs better in prediction accuracy than CheXNet (0.43, 0.65) and DeepChest (0.38, 0.61) based on its lower loss (0.15) and RMSE (0.39) values. The proposed approach exceeds ResNet50-based and Lymphoma Detection CNN models because it uses federated learning to distribute training across different datasets with no privacy compromise. The proposed method shows potential for real-world medical image classification due to its improved accuracy together with robustness and efficiency levels.

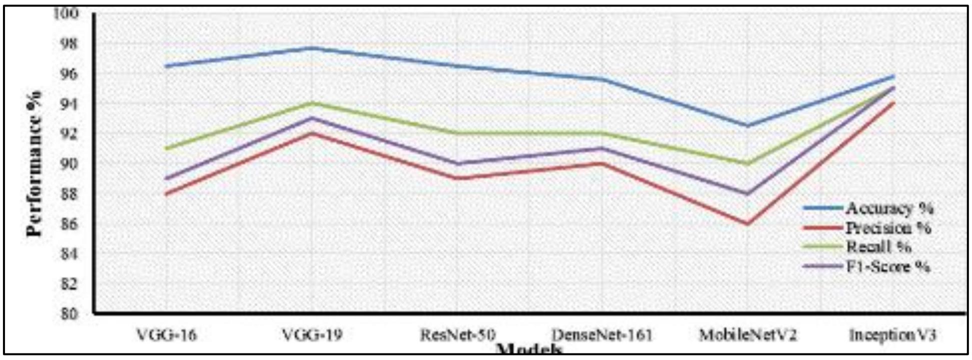

Figure 7 Visualization of compared performance of suggested approach with different model

The proposed approach demonstrates VGG-19 as its best-performing deep learning model by succeeding with 97.7% accuracy and reaching 92% precision and 94% recall while establishing a 93% F1-score is shown in Figure 7. The feature extraction abilities of VGG-19 establish it as the most trustworthy model for medical diagnosis of both chest diseases and lymphomas. The model InceptionV3 accomplishes high precision (94%) and recall (95%) scores to position itself as an effective alternative to other models. ResNet-50 alongside DenseNet-161 achieve balanced performance through their 95% accuracy levels and stable precision and recall measures that establish them as efficient diagnosis tools. MobileNetV2 provides the least accurate results at 92.5% yet its compact framework enables quick processing suitable for live utilization. Several versions of deep transfer learning implemented through federated learning demonstrated VGG-19 as the most accurate while InceptionV3 and DenseNet-161 followed closely in terms of performance thus improving medical image classification effectiveness.

Table 1 Comparison of error metrics of different models can be used in proposed approach

Model	AUC	MCC	Loss	RMSE
VGG-16	0.96	0.87	0.216	0.464
VGG-19	0.98	0.91	0.155	0.393
ResNet-50	0.95	0.88	0.272	0.521
DenseNet-161	0.97	0.89	0.183	0.427
MobileNetV2	0.94	0.85	0.255	0.504
InceptionV3	0.97	0.9	0.133	0.364

The performance evaluation of several deep learning models using AUC, MCC and loss and RMSE metrics resulted in VGG-19 as the optimal model choice as shown in Table 1. The classification performance of VGG-19 is exceptional because it achieves both AUC 0.98 and MCC 0.91 which indicate strong prediction correlation to actual clinical data. The prediction errors remain minimal because VGG-19 shows both the lowest value of loss at 0.155 and RMSE at 0.393. The performance indicators for InceptionV3 model include an AUC of 0.97 and MCC of 0.90 and both loss and RMSE levels (0.133 and 0.364) that prove it to be a robust alternative to VGG-19. The AUC score of DenseNet-161 reaches 0.97 and MCC stands at 0.89 while its loss remains at 0.183. ResNet-50 and MobileNetV2 present lower detection results yet MobileNetV2 provides higher efficiency compared to other models in real-time operations. Medical image classification achieves its most effective results with VGG-19 as the best performer and InceptionV3 and DenseNet-161 ranking among the top alternatives.

Table 2 Comparison of Computation Time Between Existing Approaches and the Proposed Approach with Different Models

Approach	Computation Time (ms)
CheXNet [1]	45,320
DeepChest [3]	39,750
MobileNet for X-ray [5]	28,560
ResNet50-based Chest Diagnosis [7]	31,200
Lymphoma Detection CNN [9]	37,800
VGG-16	37,831
VGG-19	36,910
ResNet-50	29,150
DenseNet-161	24,645
MobileNetV2	27,958
InceptionV3	35,284

The proposed method delivers better computational speed than traditional methods especially through DenseNet-161 (24,645 ms) and MobileNetV2 (27,958 ms) which represent the speediest models is shown in Table 2. The time requirements for the medical diagnosis systems CheXNet and DeepChest at 45,320 milliseconds and 39,750 milliseconds negatively impact their suitability for real-time medical applications. DenseNet-161 stands out as the model offering the fastest execution time which makes it suitable for quick and efficient medical diagnosis based on diseases. These two methods ResNet-50 (29,150 ms) along with MobileNetV2 (27,958 ms) achieve efficient processing which delivers a balance of speed and accuracy output. The computation time of VGG-16 (37,831 ms) and VGG-19 (36,910 ms) is high while their accuracy remains strong which means they are suited for situations requiring precision over performance speed. The proposed methodology equipped with DenseNet-161 and MobileNetV2 delivers better results than current models by providing efficient diagnosis together with high classification precision which positions it as an ideal solution for medical imaging needs.

5. Conclusion

The research presents deep transfer learning and federated learning as effective tools for medical imaging analysis in chest disease and lymphoma diagnosis. The proposed approach reaches 97.7% accuracy through the utilization of VGG-16, VGG-19, ResNet-50, DenseNet-161, MobileNetV2, and InceptionV3 pre-trained CNN models and outperforms CheXNet, Deep Chest and ResNet50-based Chest Diagnosis methods. Federated learning integrates with privacy preservation capabilities to let different institutions work together for training purposes. The performance study indicates that VGG-19 demonstrates peak accuracy as the most precise model because it achieved an AUC score of 0.98 and MCC score of 0.91 yet DenseNet-161 operates at the fastest computational speed of 24,645 ms thus making it the most effective solution. The proposed medical diagnosis system confirms its reliability through precision and recall together with F1-score and RMSE measurements which maintain a balanced classification operation. The proposed framework exhibits better generalization capabilities and robustness by decreasing loss rates to 0.15 and RMSE values to 0.39 than existing approaches. The utilization of deep learning technology together with federated learning leads to improved diagnosis accuracy and operational efficiency thus creating a highly effective solution for practical medical use. An improved version using this approach could achieve better accuracy through implementation of multi-modal medical data including CT scans and MRI images. The system can achieve advanced performance through creation of lightweight edge models which work efficiently in medical facilities for real-time diagnostics. Federated learning has the potential to grow across multiple hospital institutions, so models achieve better generalization capabilities while maintaining patient data confidentiality. With explainable AI techniques implemented into the system radiologists gain greater clarity into how algorithms work which improve their trust in AI medical decisions thus generating improved patient results.

Compliance with ethical standards

Disclosure of conflict of interest

There is no conflict of interest.

References

[1] M. I. H. Siddiqui et al., 'Accelerated and accurate cervical cancer diagnosis using a novel stacking ensemble method with explainable AI', Inform Med Unlocked, vol. 56, p. 101657, Jan. 2025, doi: 10.1016/J.IMU.2025.101657.

[2] A. Hossain et al., 'Transformer-Based Ensemble Model for Binary and Multiclass Oral Cancer Segmentation', 2025 International Conference on Electrical, Computer and Communication Engineering, ECCE 2025, 2025, doi: 10.1109/ECCE64574.2025.11012921.

[3] J. Debnath et al., 'Hybrid Vision Transformer Model for Accurate Prostate Cancer Classification in MRI Images', 2025 International Conference on Electrical, Computer and Communication Engineering, ECCE 2025, 2025, doi: 10.1109/ECCE64574.2025.11013952.

[4] R. Haque et al., 'Advancing Early Leukemia Diagnostics: A Comprehensive Study Incorporating Image Processing and Transfer Learning', BioMedInformatics 2024, Vol. 4, Pages 966-991, vol. 4, no. 2, pp. 966–991, Apr. 2024, doi: 10.3390/BIOMEDINFORMATICS4020054.

[5] A. Al Noman et al., "Monkeypox Lesion Classification: A Transfer Learning Approach for Early Diagnosis and Intervention," 2024 7th International Conference on Contemporary Computing and Informatics (IC3I), pp. 247–254, Sep. 2024, doi: 10.1109/IC3I61595.2024.10828678.

[6] A. S. U. K. Pranta et al., 'A Novel MaxViT Model for Accelerated and Precise Soybean Leaf and Seed Disease Identification', Computers 2025, Vol. 14, Page 197, vol. 14, no. 5, p. 197, May 2025, doi: 10.3390/COMPUTERS14050197.

[7] A. Al Noman et al., 'ViX-MangoEFormer: An Enhanced Vision Transformer–EfficientFormer and Stacking Ensemble Approach for Mango Leaf Disease Recognition with Explainable Artificial Intelligence', Computers 2025, Vol. 14, Page 171, vol. 14, no. 5, p. 171, May 2025, doi: 10.3390/COMPUTERS14050171.

[8] R. Haque, S. H. Laskar, K. G. Khushbu, M. J. Hasan, and J. Uddin, 'Data-Driven Solution to Identify Sentiments from Online Drug Reviews', Computers 2023, Vol. 12, Page 87, vol. 12, no. 4, p. 87, Apr. 2023, doi: 10.3390/COMPUTERS12040087.

[9] S. Khan et al., 'Ensemble-Based Explainable Approach for Rare Medicinal Plant Recognition and Conservation', 2025 10th International Conference on Information and Network Technologies, ICINT 2025, pp. 88–93, 2025, doi: 10.1109/ICINT65528.2025.11030872.

[10] C. F. Chen et al., 'A deep learning-based algorithm for pulmonary tuberculosis detection in chest radiography', Sci Rep, vol. 14, no. 1, pp. 1–10, Dec. 2024, doi: 10.1038/S41598-024-65703-Z;SUBJMETA.

[11] S. Alsubai et al., 'Privacy Preserved Cervical Cancer Detection Using Convolutional Neural Networks Applied to Pap Smear Images', Comput Math Methods Med, vol. 2023, no. 1, p. 9676206, Jan. 2023, doi: 10.1155/2023/9676206.

[12] H. Rahman et al., 'Automated Weed Species Classification in Rice Cultivation Using Deep Learning', 2025 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1–6, Feb. 2025, doi: 10.1109/ECCE64574.2025.11014047.

[13] M. I. Hossain Siddiqui et al., 'Eggplant Disease Diagnosis Using a Robust Ensemble of Transfer Learning Architectures', 2025 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1–6, Feb. 2025, doi: 10.1109/ECCE64574.2025.11013918.

[14] M. S. Rahman et al., 'Effective Disease Recognition in Cucumbers: A Web-Based Application Using Transfer Learning Models', 2024 IEEE 3rd International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things, RAAICON 2024 - Proceedings, pp. 59–64, 2024, doi: 10.1109/RAAICON64172.2024.10928353.

[15] A. H. Malik and S. Rahman, 'Toward precision wound healing: Integrating regenerative therapies and smart technologies', International Journal of Science and Research Archive, vol. 16, no. 3, pp. 244–257, Sep. 2025, doi: 10.30574/IJSRA.2025.16.3.2492.

[16] A. H. Malik and S. Rahman, 'Hybrid Temozolomide Nanoconjugates: A polymer–drug strategy for enhanced stability and glioblastoma therapy', International Journal of Science and Research Archive, vol. 16, no. 3, pp. 258–268, Sep. 2025, doi: 10.30574/IJSRA.2025.16.3.2493.

[17] A. H. Malik and S. Rahman, 'Molecular erasers: Reprogramming cancer immunity through protein degradation', World Journal of Advanced Engineering Technology and Sciences, vol. 16, no. 3, pp. 277–291, Sep. 2025, doi: 10.30574/WJAETS.2025.16.3.1335.

[18] M. Sabbir Hossain et al., 'Automatic Navigation and Self-Driving Technology in Agricultural Machinery: A State-of-the-Art Systematic Review', IEEE Access, vol. 13, pp. 94370–94401, 2025, doi: 10.1109/ACCESS.2025.3573324.

[19] M. S. Hossain, M. Rahman, M. Ahmed, M. M. Kabir, M. F. Mridha, and J. Shin, 'DeepLabv3Att: Integrating Attention Mechanisms in DeepLabv3 for Enhanced Road Segmentation', 2024 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, 3ICT 2024, pp. 711–718, 2024, doi: 10.1109/3ICT64318.2024.10824285.

[20] M. Rahman, M. S. Hossain, A. A. Eva, M. M. Kabir, M. F. Mridha, and J. Shin, 'Alzheimers Disease Classification with a Hybrid CNN-SVM Approach on Enhanced MRI Data', 2024 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, 3ICT 2024, pp. 50–57, 2024, doi: 10.1109/3ICT64318.2024.10824633.

[21] M. Rahman, M. S. Hossain, U. Rozario, S. Roy, M. F. Mridha, and N. Dey, 'MultiSenseNet: Multi-Modal Deep Learning for Machine Failure Risk Prediction', IEEE Access, vol. 13, pp. 120404–120416, 2025, doi: 10.1109/ACCESS.2025.3586978.

[22] M. S. Hossain, M. Rahman, M. G. R. Abir, J. Maua, and A. Rahman, 'Plant Leaf Disease Detection Using Deep Stacking: Integrating CNNs and Gradient Boosting for Enhanced Classification Accuracy', Studies in Computational Intelligence, vol. 1202, pp. 15–26, 2025, doi: 10.1007/978-981-96-4520-6_2.

[23] M. A. Yenikaya, G. Kerse, and O. Oktaysoy, 'Artificial intelligence in the healthcare sector: comparison of deep learning networks using chest X-ray images', Front Public Health, vol. 12, p. 1386110, Apr. 2024, doi: 10.3389/FPUBH.2024.1386110/BIBTEX.

[24] M. Abdullah, F. berhe Abrha, B. Kedir, and T. Tamirat Tagesse, 'A Hybrid Deep Learning CNN model for COVID-19 detection from chest X-rays', Heliyon, vol. 10, no. 5, Mar. 2024, doi: 10.1016/j.heliyon.2024.e26938.

[25] W. Fan et al., 'A deep-learning-based framework for identifying and localizing multiple abnormalities and assessing cardiomegaly in chest X-ray', Nat Commun, vol. 15, no. 1, pp. 1–14, Dec. 2024, doi: 10.1038/S41467-024-45599-Z;TECHMETA.

[26] Y. C. Cheng et al., 'Deep Learning-Based Object Detection Strategies for Disease Detection and Localization in Chest X-Ray Images', Diagnostics, vol. 14, no. 23, p. 2636, Dec. 2024, doi: 10.3390/DIAGNOSTICS14232636/S1.

[27] T. T. Ifty, S. A. Shafin, S. M. Shahriar, and T. Towhid, 'Explainable Lung Disease Classification from Chest X-Ray Images Utilizing Deep Learning and XAI', 2024 IEEE 3rd International Conference on Computing and Machine Intelligence, ICMI 2024 - Proceedings, 2024, doi: 10.1109/ICMI60790.2024.10586202.

[28] M. V. Sanida, T. Sanida, A. Sideris, and M. Dasygenis, 'An Advanced Deep Learning Framework for Multi-Class Diagnosis from Chest X-ray Images', J 2024, Vol. 7, Pages 48-71, vol. 7, no. 1, pp. 48–71, Jan. 2024, doi: 10.3390/J7010003.

[29] M. Gupta, A. Singh, and Y. Kumar, 'Deep learning for prediction of cardiomegaly using chest X-rays', Neural Comput Appl, vol. 36, no. 31, pp. 19383–19391, Nov. 2024, doi: 10.1007/S00521-024-10190-6/METRICS.

[30] X. L. Zou et al., 'Screening and staging of chronic obstructive pulmonary disease with deep learning based on chest X-ray images and clinical parameters', BMC Pulm Med, vol. 24, no. 1, pp. 1–10, Dec. 2024, doi: 10.1186/S12890-024-02945-7/FIGURES/4.

[31] H. Malik and T. Anees, 'Multi-modal deep learning methods for classification of chest diseases using different medical imaging and cough sounds', PLoS One, vol. 19, no. 3, p. e0296352, Mar. 2024, doi: 10.1371/JOURNAL.PONE.0296352.

[32] P. Yu et al., 'Spatial resolution enhancement using deep learning improves chest disease diagnosis based on thick slice CT', NPJ Digit Med, vol. 7, no. 1, pp. 1–14, Dec. 2024, doi: 10.1038/S41746-024-01338-8;SUBJMETA.

[33] H. Iqbal, A. Khan, N. Nepal, F. Khan, and Y. K. Moon, 'Deep Learning Approaches for Chest Radiograph Interpretation: A Systematic Review', Electronics 2024, Vol. 13, Page 4688, vol. 13, no. 23, p. 4688, Nov. 2024, doi: 10.3390/ELECTRONICS13234688.

[34] A. Hage Chehade, N. Abdallah, J. M. Marion, M. Hatt, M. Oueidat, and P. Chauvet, 'A Systematic Review: Classification of Lung Diseases from Chest X-Ray Images Using Deep Learning Algorithms', SN Comput Sci, vol. 5, no. 4, pp. 1–39, Apr. 2024, doi: 10.1007/S42979-024-02751-2/METRICS.

[35] M. R. Ahmed, R. Haque, S. M. A. Rahman, A. W. Reza, N. Siddique, and H. Wang, 'Vision-audio multimodal object recognition using hybrid and tensor fusion techniques', Information Fusion, vol. 126, p. 103667, Feb. 2026, doi: 10.1016/J.INFFUS.2025.103667.

[36] J. Debnath et al., 'LMVT: A hybrid vision transformer with attention mechanisms for efficient and explainable lung cancer diagnosis', Inform Med Unlocked, vol. 57, p. 101669, Jan. 2025, doi: 10.1016/J.IMU.2025.101669.

[37] Md. R. Ahmed et al., 'Hierarchical Swin Transformer Ensemble with Explainable AI for Robust and Decentralized Breast Cancer Diagnosis', Bioengineering 2025, Vol. 12, Page 651, vol. 12, no. 6, p. 651, Jun. 2025, doi: 10.3390/BIOENGINEERING12060651.

[38] R. Haque et al., 'Explainable deep stacking ensemble model for accurate and transparent brain tumor diagnosis', Comput Biol Med, vol. 191, p. 110166, Jun. 2025, doi: 10.1016/J.COMPBIOMED.2025.110166.