ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# A DEEP LEARNING-BASED HYBRID MODEL FOR AUTOMATED LUNG DISEASE DETECTION: ADDRESSING COVID-19 DIAGNOSTIC CHALLENGES

**NIGHILA ASHOK K , S SIVAKUMARI**
Computer Science and Engineering Department
Avinashlingam Institute of home science and higher education for women
Coimbatore

Computer Science and Engineering Department
Avinashlingam Institute of home science and higher education for women
Coimbatore
21pheop005@avinuty.ac.in, sivakumari_cse@avinuty.ac.in

**ABSTRACT**

Lung abnormalities in the post-COVID era are a common issue, demanding accurate and timely diagnosis. However, detecting these irregularities with deep learning models has its own difficulties like class imbalance, overfitting, and restricted generalizability due to heterogeneous datasets, all of which can impede accurate detection. Here we apply deep learning classification on extensive datasets compiled during and after the COVID-19 pandemic, with an ability for high-performance AI algorithms to peruse normal lungs and lungs affected by pneumonia, cardiomegaly, and COVID-19. We implemented an integrated framework of two convolution neural network architectures, Places365 GoogLeNet and EfficientNetB0. The fusion of these models employed the AdaBoost ensemble method, significantly enhancing classification accuracy. Places 365 GoogLeNet achieved a validation accuracy of 90.49%, while EfficientNetB0 reached 94.70%. By integrating these models, the classification accuracy improved to 97.48%, showcasing the effectiveness of model fusion in achieving superior performance. This framework demonstrates promise for diagnosing complex lung conditions, particularly those related to COVID-19, offering potential as a robust diagnostic tool.

**Keywords**-*COVID-19, Places 365 GoogLeNet, CNN, EfficientNet B0, ResNet, Pneumonia, Cardiomegaly, Ensemble*

## I. INTRODUCTION

### A. Background

The COVID-19 pandemic has emphasized the critical need for effective diagnostic tools to identify and manage respiratory diseases such as pneumonia, cardiomegaly, and COVID-19. These conditions present significant challenges due to their impact on global health and the complexity of their diagnosis. Rapid and accurate identification is essential for timely intervention, treatment, and improving patient outcomes. Among these conditions, cardiomegaly, characterized by the enlargement of the heart, is often observed in chest X-rays and can indicate underlying health issues such as heart failure or complications related to COVID-19 [1][2]. Deep learning, a branch of artificial intelligence, has emerged as a promising solution for medical image analysis, offering significant advantages in automating diagnostic processes. The integration of deep learning models with medical imaging can alleviate the burden on radiologists, reduce human error, and enhance diagnostic precision. However, existing models face challenges such as class imbalance, overfitting, and limited generalizability. Overfitting is a general problem in machine learning but is especially pronounced in COVID-19-related data due to factors such as small sample sizes for certain classes (e.g., rare cases or outcomes), the use of noisy or heterogeneous data from multiple sources, and the rapid emergence of new variants that may not align well with the training data [3][4].

### B. Research gap

The diagnosis of respiratory conditions has seen significant advancements with the adoption of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

deep learning models [5] such as ResNet, VGGNet, and DenseNet, which have demonstrated strong feature extraction capabilities in medical imaging tasks. However, these models often face critical limitations, including overfitting on small, imbalanced datasets, high computational demands, and limited generalizability to real-world clinical settings. Traditional ensemble methods have been explored to address these issues, but they frequently lack the robustness needed for distinguishing subtle variations among conditions like COVID-19, cardiomegaly, and pneumonia. Furthermore, while chest imaging scores provide valuable insights for disease severity, existing algorithms fail to interpret these scores reliably in diverse populations. To address these gaps, this study proposes a novel integration of GoogLeNet and EfficientNetB0, leveraging their respective strengths in multi-scale feature extraction and computational efficiency. The fusion of these architectures with the AdaBoost ensemble method aims to enhance diagnostic accuracy, mitigate class imbalance by assigning higher weights to misclassified samples, combining multiple weak learners as well as Reducing Bias in Feature Learning, and provide a significant advancement over existing state-of-the-art techniques in automated lung disease classification.

## C. Problem Statement

Globally, millions of COVID-19-recovered patients experience long-term lung complications, leading to increased hospitalizations and misdiagnosis risks. Traditional methods often fail to distinguish COVID abnormalities from other conditions like pneumonia and cardiomegaly, necessitating a more robust AI-driven solution.The conventional methods for diagnosing respiratory conditions rely heavily on manual evaluation by radiologists, which can be time-consuming and prone to errors. The increased workload during the pandemic has highlighted the limitations of traditional approaches. Furthermore, diagnosing conditions like cardiomegaly in the context of COVID-19 often involves interpreting "chest imaging scores," which quantify the severity of lung involvement based on abnormalities detected in chest X-rays and CT scans. These scores serve as a valuable tool for assessing disease progression and guiding treatment plans, especially in severe cases [6]. Despite their potential, these imaging scores require robust algorithms to ensure accurate interpretation.

## D. Contributions

This study aims to address the challenges in lung disease diagnosis by developing an integrated deep learning framework. The main contributions of this research are as follows:

- **Development of an Automated Diagnostic System**: The system uses a deep learning framework that classifies lung conditions (normal lungs, pneumonia, cardiomegaly, and COVID-19) for accurate and efficient disease diagnosis.
- **Fusion of Complementary CNN Architectures**: The model combines Places365 GoogLeNet and EfficientNetB0 in an ensemble member selection strategy, obtaining consensus from complementary architectures to improve diagnostic accuracy.
- **Improved Differentiation of COVID-19 from Other Conditions**: The framework differentiates COVID-19 from pneumonia and cardiomegaly well and reduces the possibility of multi-class misclassification of COVID-19 cases in medical imaging.
- **Mitigation of Class Imbalance and Overfitting:** Data augmentation, weighted loss functions, and AdaBoost ensemble learning are developed to mitigate Class imbalance & Overfitting.

Unlike previous works that focus solely on a single CNN model, our research introduces an innovative ensemble approach leveraging GoogLeNet and EfficientNetB0 with AdaBoost, significantly reducing class imbalance effects and improving generalizability.

## E. Rationale for Choosing GoogLeNet and EfficientNet

GoogLeNet was selected for its ability to capture multi-scale image features through its Inception modules, which are particularly effective in identifying complex patterns in chest X-ray images. EfficientNetB0, on the other hand, offers a balance between computational efficiency and accuracy, making it well-suited for medical image classification tasks [7]. The combination of these models leverages their respective strengths—GoogLeNet's robust feature extraction and EfficientNet's scalability—to create a more reliable and accurate diagnostic tool. The fusion of these models using the AdaBoost ensemble method further improves classification performance by aggregating the predictions of both networks [8], [9], [10].

Journal of Theoretical and Applied Information Technology
30th June 2025. Vol.103. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

F. Judtifuing adaboast and Addressing Model Recency

The AdaBoost algorithm in this study is utilized as an ensemble technique to combine the predictions of two independently trained models: GoogLeNet and EfficientNetB0. These models, while sophisticated deep learning architectures, are not weak learners in the traditional sense defined by AdaBoost. Instead, they act as "base learners" in an ensemble setting, where their outputs are weighted and combined to create a more robust classifier. AdaBoost [11] adjusts the weights of misclassified samples iteratively and combines the predictions of these base learners to improve classification performance.

The use of AdaBoost in this context deviates from its conventional application with weak learners but is justified by its ability to reduce overfitting and enhance model generalizability through weighted combination. The rationale for not concatenating the two models into a single neural network stems from the need to preserve the unique feature extraction capabilities of each architecture. Concatenation may lead to challenges in training, increased computational complexity, and potential redundancy in feature representation. By leveraging AdaBoost, the complementary strengths of GoogLeNet [12] and EfficientNetB0 [13] are retained while achieving an aggregated improvement in classification accuracy. Although GoogLeNet and EfficientNetB0 are established models, their selection over more recent alternatives is intentional. The computational efficiency of these models ensures their applicability in real-world clinical settings, where resource constraints often limit the deployment of more complex architectures. Additionally, the complementary strengths of these models—GoogLeNet's feature extraction and EfficientNetB0's scalability—offer a balanced approach that aligns well with the study's objectives. The proposed ensemble method mitigates potential limitations of individual models, demonstrating performance that rivals or exceeds state-of-the-art techniques.

The remainder of this manuscript is structured as follows: The section II examines existing deep learning-based approaches for lung disease classification by evaluating their research gaps while documenting their limitations and benefits. The Dataset Description discussed in section III outlines the origin of utilized datasets while describing their class distribution and describing preprocessing procedures. Section IV discuss the proposed framework that merges Places365 GoogLeNet and EfficientNetB0 along with AdaBoost ensemble learning and strategies for class imbalance and overfitting solutions. Also, the Experimental Setup and Implementation segment in this section describe the training procedures and includes hyperparameter optimization followed by evaluation methods and computational infrastructure descriptions. Model performance analysis in Results and Discussion section includes accuracy levels as well as confusion matrices and model comparisons against existing approaches. Finally, Future Work along with Conclusion reviews major research outcomes while preparing clinical applications and suggesting research pathways.

## 2. RELATED WORKS

Numerous studies have explored deep learning models for detecting and classifying respiratory diseases, including COVID-19, pneumonia, and cardiomegaly. Key advancements in this domain are summarized below:

- A model was developed to classify chest X-rays into four categories: pneumonia, tuberculosis (TB), COVID-19, and normal. This model further categorized COVID-19 cases based on severity using VGG16, DenseNet161, and ResNet18. It achieved test accuracies of 95.9% and 98% [14][15].
- Another study utilized CNN features with a dataset of 600 radiographs, achieving 97.19% DSC and 94.13% DSC for lung and infection segmentation, respectively [16].
- A VGG model with 16 convolutional layers classified chest X-ray images into pneumonia, viral pneumonia, and normal cases, demonstrating generalization on a dataset of over 9,000 images [17].
- A comparative study evaluated 11 CNN models (e.g., VGG16, ResNet50V2, DenseNet169) and showed accuracy in detecting COVID-19, bacterial, and viral pneumonia [18]. Pandit et al. applied VGG19 with transfer learning to process 1,428 chest X-ray images resulting in 92.53% COVID-19 classification accuracy while dealing with restricted scaling concerns. With CNN and machine learning algorithms Sekeroglu and Ozsahin [19] operated on 1,808 of 6,100 images to reach a 96.51% ROC result but their approach encountered high computational complexity issues. Khan and colleagues [20] designed CoroNet using Xception architecture

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to analyze 3,000 chest X-ray images at 95% accuracy while managing multiple classification projections yet requiring significant computational power.

- Another work achieved 99.83% accuracy using CNN with PCA and ELM for binary pneumonia classification [21].
- Yamac et al. developed a sparse support estimator-based model, achieving 87.07% accuracy for four-class classification and 95.90% for COVID-19 detection [22].

Table 1 shows the comparative study of existing methods and its advantages and disadvantages.

Studies such as Chandra et al. [23] utilized a majority voting-based ensemble classifier for COVID-19 detection, while Nahiduzzaman et al. [24] proposed a CNN-based approach for multi-disease classification, but both approaches lacked a dedicated strategy for addressing class imbalance and overfitting. Serte and Serener [25] employed a trained ResNet model to classify pleural effusions (PE) derived from TB, pneumonia, and COVID-19 cases. Their work achieved detection accuracies of 99%, 75%, and 100% for PE related to pneumonia, TB, and COVID 19 respectively. For classification tasks, their average accuracy reached 83%. Sahlol et al. [26] on the other hand utilized a trained MobileNet model to extract approximately 50,000 features, from CXR images for TB detection.

An optimization algorithm based on ecosystems was utilized to select features surpassing the performance of current methods, with accuracy rates of 90.20% and 94.10% for Shenzhen (SZ) and Dataset 2 respectively [26]. Chandra et al. developed a computer-aided diagnosis (CAD) system for TB detection from CXR images.

*Table 1 Comparative Summary Of The Existing Methods*

| Study | Methodology | Dataset Size | Accuracy | Advantages | Disadvantages |
|---|---|---|---|---|---|
| **Model [13][14]** | VGG16, DenseNet161, ResNet18 | 64 × 64 images | 95.9% - 98% | Multi-class classification | Limited resolution and generalizability |
| **Segmentation Study [15]** | ED-CNNs, UNet, FPN | 600 radiographs | 97.19% (DSC) | High segmentation precision | Complex implementation |
| **VGG Model [16]** | VGG with 16 convolution layers | 9,000+ images | Not specified | Strong feature generalization | Potential overfitting |
| **Comparative Study [17]** | VGG16, ResNet50V2, DenseNet169 | X-ray dataset | Not specified | Comprehensive performance evaluation | No detailed generalization metrics |
| **Pandit et al. [18]** | VGG19 Transfer Learning | 1,428 images | 92.53% | Strong COVID-19 classification | Limited scalability |
| **Sekeroglu and Ozsahin [19]** | CNN and ML Algorithms | 1,808/6,100 images | 96.51% (ROC) | High multi-class classification | Computational complexity |
| **Khan et al. [20]** | CoroNet (Xception-based) | 3,000+ images | 95% | Handles multi-class classification | Computationally intensive |
| **PCA-ELM [21]** | CNN with PCA and ELM | 600 images | 99.83% | Excellent binary classification | Limited to two classes |
| **Yamac et al. [22]** | Sparse Support Estimator Network | 6,200 images | 87.07% | Novel sparse estimation | Lower accuracy for complex datasets |

Their approach involved reducing image noise identifying lung boundaries and extracting characteristics followed by SVM classification. They achieved accuracies of 95.60% and 99.40% for the Montgomery (MT) and SZ datasets respectively [27]. Tawsifur et al. [28] employed nine transfer learning (TL) models along with two Net models to detect TB using a database containing 7,000 CXR images. After applying augmentation techniques, they achieved accuracies of 96.7% and 98.6% using

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ChexNet and DenseNet201 respectively. Additionally, they incorporated t distributed stochastic neighbor embedding as a means of visualizing the data [29-32]. Muhammad et al., on the other hand, employed three transfer learning models to extract features from a set of 7,000 CXR images and used eXtreme Gradient Boosting for TB detection [33, 34]. Chest Computed Tomography (CT) and X-ray imaging are widely used to evaluate the condition of the lungs and chest in individuals affected by COVID-19. Radiologists and healthcare professionals often assign scores to these images to measure the severity of lung involvement. These scores can indicate the extent and characteristics of lung abnormalities, including ground glass opacities, consolidation, and other pulmonary findings. A comprehensive report on COVID-19 patients in Japan [35] presents findings regarding the imaging characteristics observed upon admission and their correlation with disease severity. The main discoveries are as follows; (1) The study found that older age and male gender were associated with cases of COVID-19, which aligns with previous research. Notably, in cases chest abnormalities seen on X-rays and CT scans were frequently detected in the central regions of both lungs whereas milder cases showed these abnormalities predominantly in the peripheral and basal regions. (2) The severity of COVID 19 when evaluated using a scoring system based on chest X-ray and CT findings demonstrated a relationship that depended on the dose received. A specific score of 9 points on a chest x-ray was identified as a threshold, for predicting disease.

Interestingly, the study found that an enlarged heart, known as cardiomegaly, was strongly linked to disease severity on chest X-rays. This correlation held true in cases where there were no existing heart conditions. The study emphasized the value of using chest X-rays to detect cardiomegaly and assess the severity of the disease in situations where performing CT scans might be difficult. Additionally, the study highlighted that a higher score for opacity and density in the middle lung areas was indicative of severe disease.

2.1. Challenges:

Though other research in lung disease shows promise in classification results, factors including class imbalance, computational inefficiency, overfitting, and a dearth of generalizability remain. Most current models depend on one CNN structure and aren't versatile on various datasets. To this end, this work presents an ensemble strategy utilizing Places365 GoogLeNet and EfficientNetB0 with AdaBoost to indeed achieve a superior classification accuracy, generalizability, and a lessened impact of class imbalance and overfitting.

## 3. DATASET

Researchers from Qatar University, Doha, Qatar have collaborated with their counterparts from the University of Dhaka, Bangladesh, and partners from Pakistan and Malaysia to establish a significant chest X-ray image database [36]. It is a major resource for healthcare providers and researchers who are dealing with COVID-19. It first had 219 COVID-19, 1341 normal, and 1345 viral pneumonia chest X–rays, which have been regularly augmented for added value since then. The second modification grew the database to comprise 3616 COVID-19-positive cases, 10 thousand 192 routine cases, 6 thousand 12 cases of Lung density (non-COVID lung infection) and 1345 Viral Pneumonia. The continuous efforts make updated diagnostic, as well as treated data available for use by health professionals and researchers. To boost the use of artificial intelligence in cancer diagnosis, the National Institute for Health (NIH) Clinical Center has released 100,000 de-identified chest X-ray pictures and data. The dataset, comprising 30,000 patients, puts the privacy of patients first [37]. Though reading chest X-rays is complicated, it can be considered training an AI to read multiple scans and support radiologist findings including unexplored aspects. This is a step that may help in medical imaging and diagnostics which will benefit people around the world. The diseases COVID-19 and Viral Pneumonia images are shown in Table 2

Table 2: Dataset details

| Diseases | Number of Images |
|---|---|
| Cardiomegaly | 1094 |
| COVID | 3616 |
| Normal | 5552 |
| Viral Pneumonia | 3145 |

a. Dataset Sources and Inclusion Criteria

The dataset used in this study was compiled from publicly available repositories, including the Qatar University database and the National Institute for Health (NIH) Clinical Center's dataset. The inclusion criteria focused on high-quality chest X-ray images representing four categories: normal, COVID-19, cardiomegaly, and viral pneumonia. Images with significant artifacts, low resolution, or incomplete annotations were excluded to ensure the reliability of the dataset. The final dataset consisted

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of over 10,000 labeled images, distributed across the four categories.

b. The inclusion criteria

The inclusion criteria focused on high-quality chest X-ray images representing four categories: normal, COVID-19, cardiomegaly, and viral pneumonia. Images with significant artifacts, low resolution, or incomplete annotations were excluded to ensure the reliability of the dataset. The final dataset consisted of over 10,000 labeled images, distributed across the four categories.

## 4. METHODOLOGY

Figure 1 shows the proposed method which involves data preprocessing with translation, reflection, and scaling, while model enhancements include BatchNorm, Leaky ReLU, and EfficientNetB0. An ensemble approach with AdaBoost combines model predictions to boost accuracy, creating a robust framework for scene recognition.

a. Preprocessing

Data preprocessing was a critical step in this study to standardize the input data and improve model performance. The preprocessing steps included:

- **Image Resizing:** All images were resized to a uniform dimension of 224 × 224 pixels to ensure compatibility with the CNN architectures.
- **Normalization:** Pixel values were scaled to the range [0, 1] to enhance computational efficiency and stabilize the training process.
- **Noise Reduction:** Gaussian filters were applied to reduce noise and improve image clarity.
- **Label Verification:** Each image's label was cross-verified to eliminate potential misclassifications.

B. Challenges and Mitigation Strategies

One significant challenge was class imbalance, where certain categories, such as COVID-19 and cardiomegaly, had fewer samples compared to the normal category. To address this, the following strategies were implemented:

- **Data Augmentation:** Techniques like translation, reflection, and scaling were applied to increase the diversity of the minority class samples.
- **Weighted Loss Functions:** The loss function was modified to assign higher weights to underrepresented classes, ensuring balanced learning across all categories.

C. Data Augmentation Techniques

- Translation:

The RandXTranslation and RandYTranslation operations are responsible for randomly translating (shifting) the images in both the $X$ and $Y$ directions within the specified pixel range. Let's denote the original image as $I$, and the translated image as $I'$. The translation operation can be mathematically expressed as shown in equation (1) and (2) for both X and Y directions:

$$I'(x,y) = I(x - \Delta x, y) \quad (1)$$

$$I'(x,y) = I(x, y - \Delta y) \quad (2)$$

$\Delta x$ and $\Delta y$ represent the random translations within the specified pixel range.

- Reflection:

The RandXReflection operation is responsible for randomly reflecting (flipping) the images horizontally as shown in equation (3).

$$I'(x,y) = I\,(width - x, y) \quad (3)$$

Here, I is the original image, I' is the reflected image, and 'width' is the width of the image.

- Scaling:

The RandXScale and RandYScale operations perform random scaling in both the X and Y directions within the specified scale range. Let's denote the original image as I, the scaled image as $I'$, and the scaling factors as $S_x$ and $S_y$. The scaling operation can be mathematically expressed as follows for both X and Y directions: Equation 4 shows the X-Direction:

Journal of Theoretical and Applied Information Technology
30th June 2025. Vol.103. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

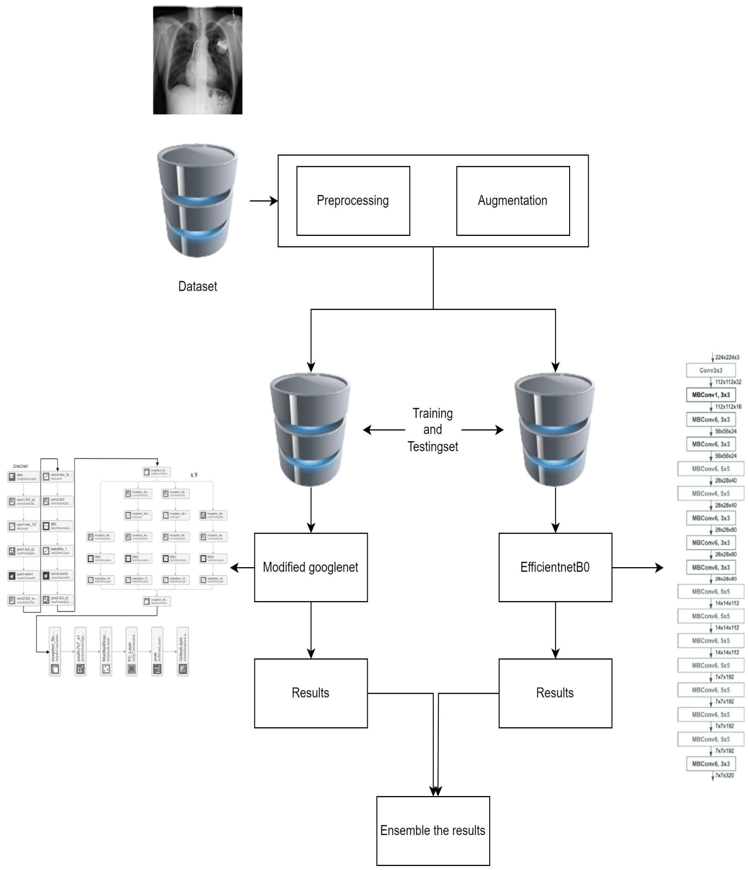



*Figure 1 The Generalized Diagram Of The Proposed Method*

$$I'(x, y) = I(\frac{x}{S_x}, y) \quad (4)$$

Equation 5 shows the for Y-Direction:

$$\text{I}'(\text{x, y}) = \text{I}(\text{x}, \frac{y}{S_y}) \quad (5)$$

Sx and Sy represent the random scaling factors within the specified scale Range.

These operations are applied to the images in the dataset, and the resulting images are used for training, validation, and testing. The specific values of $\Delta x$, $\Delta y$, $S_x$, $S_y$, and other parameters are determined by the data augmentation settings and randomization. These operations help introduce variations into the dataset to enhance the model's robustness and ability to handle different variations in input data.

D. Places 365 GoogLeNet

The main breakthrough is in including convolutional filters of different sizes within inception modules. This lets the network catch multi-scale features needed for complex pattern recognition. Also, auxiliary classifiers are added to the middle layers. These help with gradient flow during training and boosting the model's overall performance. While training, it was clear the model had problems. It had plateauing loss and reduced accuracy. To fix this, batch normalization was added between the Convolutional and Leaky ReLU layers. This change was meant to speed up training, decrease sensitivity to network starting setup, and improve convergence. We've made a key improvement involving adding batch normalization (BatchNorm) layers to the network structure. BatchNorm works by standardizing layer actions inside every mini-batch. The batch average gets deducted from each channel, and this output is divided by the batch's standard deviation. Then, adjustable settings come into play to scale and move the standardized activities. This lets the network pick the best scale and shift for each layer. The modified GoogleNet Network diagram is shown in figure 2.

The mathematical representation of BatchNorm (X) for a given layer as shown in equation (6):

$$X = (\gamma(X - \mu))/\sigma + \beta \quad (6)$$

Here, X represents the layer's input, μ is the batch mean, σ is the batch standard deviation, γ is a learnable scale parameter, and β is a learnable shift parameter. These parameters are fine-tuned during training.

In addition to BatchNorm, the traditional ReLU (Rectified Linear Unit) activation function was replaced with the Leaky ReLU (Leaky Rectified Linear Unit) function. The special feature of the Leaky ReLU is that it gives a tiny slant to the negative- side to help counter the "vanishing gradient" problem. The mathematical representation of the Leaky ReLU is as as shown in equation (7) and (8):

$$LeakyReLU(x) = x, if\ x > 0; \quad (7)$$

$$LeakyReLU(x) = \alpha * x, if\ x \leq 0, \quad (8)$$

where α is a small positive constant.

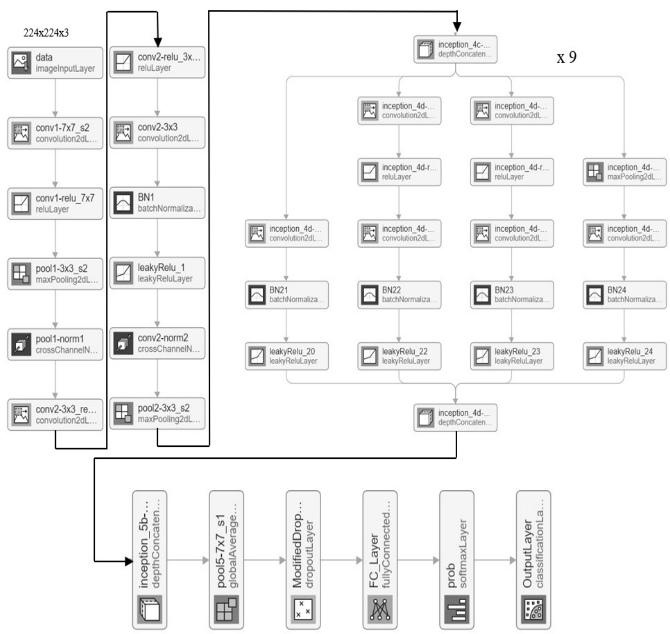


Figure 2 Network diagram of Modified Places 365 GoogleNet

E. Hyperparameter Tuning:

The optimization process involves fine-tuning several key hyperparameters to improve the model's performance. These hyperparameters include:

**Batch Size:** The number of images processed in each training iteration. An optimal balance must be struck to ensure efficient convergence. In this work, a minimum batch size of 32 was utilized.

**Pooling Operation**: Determines the specific aggregation function applied, with average pooling being chosen in this instance.

**Learning Rate:** Governs the speed at which the model learns optimal weights during training. A learning rate of 0.001 was selected.

**Optimizer:** The method employed to adjust neural network parameters during training. Among the three optimizers considered, 'sgdm' was chosen as the most suitable for this case. These hyperparameters collectively influence the model's training process, convergence, and ultimately, the accuracy of the GoogLeNet Places365.

F. Efficientnet B0:

EfficientNetB0 is a CNN network design that achieves a balance, between model size, computational efficiency, and accuracy in computer vision tasks. The compound scaling approach optimizes the network's depth (D) width (W) and resolution (R) as depicted in [38]. Efficientnet B0 Network diagram is shown in figure 3. The network depth is determined by a scaling parameter called "d," which controls the number of blocks within the architecture. These blocks contain repeated layers of depth convolutions as shown in equation (9). The width of EfficientNetB0 is controlled by another scaling parameter, "w," which scales the number of channels in each layer as shown in equation (10). Additionally, the input image resolution is scaled using a factor "r" that represents an aspect ratio between width and height as shown in equation (11). To optimize this architecture, we need to solve an optimization problem. The major goal is to minimize a cost function that combines accuracy and computational efficiency metrics while

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

respecting constraints, on depth, width and resolution scaling factors. This careful parameterization ensures that EfficientNetB0 performs across tasks while remaining computationally efficient as discussed in [38].

$$d = \alpha^{\varphi} \quad (9)$$
$$w = \beta^{\varphi} \quad (10)$$
$$r = \gamma^{\varphi} \quad (11)$$

Subject to the following constraints:

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$$
$$\alpha \geq 1, \beta \geq 1, \gamma \geq 1$$


*Figure 3. Network Diagram Of Efficientnetb0*

G. Justification of selecting Places 365 GoogLeNet and EfficientnetB0

The selection of GoogLeNet and EfficientNet for ensemble modeling is grounded in their complementary architectures and demonstrated effectiveness in image classification tasks. GoogLeNet, with its inception modules, employs multi-scale feature extraction within a single layer, making it adept at capturing features across varying resolutions. On the other hand, EfficientNet utilizes a compound scaling strategy that optimally balances the depth, width, and resolution of the network, achieving high accuracy with fewer computational resources. These architectures, though distinct, offer synergistic strengths: GoogLeNet's capacity for detailed local feature extraction complements EfficientNet's ability to generalize across diverse datasets, resulting in a robust combination. Furthermore, both models are computationally efficient, with GoogLeNet reducing parameter count through inception modules and EfficientNet excelling in scalability without significant computational overhead. This efficiency makes them ideal candidates for ensemble modeling, where the fusion of their predictions can mitigate individual model biases and improve overall performance. Together, these characteristics ensure that the ensemble benefits from diverse feature extraction strategies and resource-efficient computation, making it well-suited for complex classification tasks.

H. Ensemble method

By combining the outcomes of two network models, namely Places365 GoogLeNet and EfficientNetB0 through the utilization of the AdaBoost ensemble method we can create a robust approach, for recognizing or categorizing scenes. This ensemble technique enhances the capabilities of both models by assigning weights to their predictions and aggregating them into a final more precise prediction. In this research paper, the aim is to enhance classification accuracy by employing AdaBoost, a known method that merges the predictions, from multiple weak learners. In this case neural network models. The process involves adjusting sample weights and merging predictions to build a learner. Figure 4 shows the fusion process of GoogleNet and Efficient Net model.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

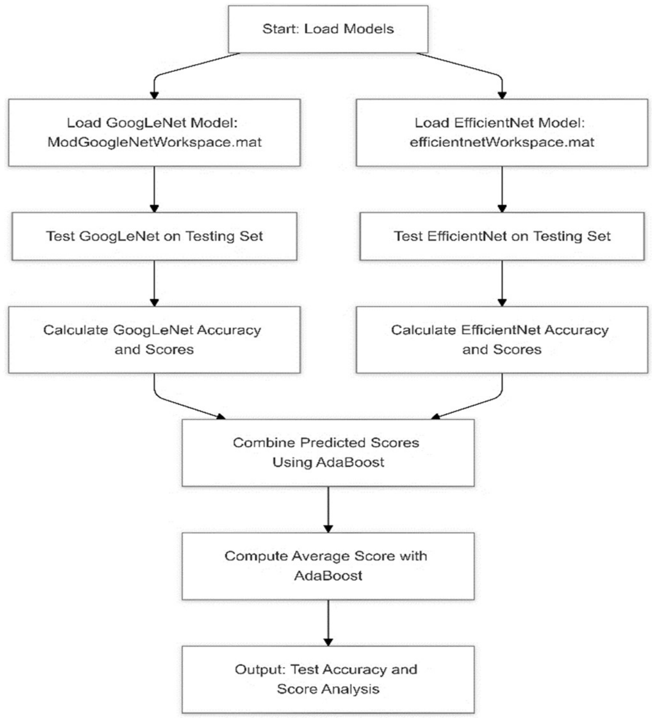



*Figure 4. Fusion Of Places 365 Googlenet And Efficient Net Model*

I. AdaBoost Algorithm

All data instances get an equal weight (D). These weights are a measure of the importance of each of the samples to the ensemble. In specific, one of the trained versions of the neural networks is applied to a training set with GoogLeNet and EfficientNetB0. In training a Weak learner ($M_k$), there are sample weights referred to as $D_k$ for misclassified samples. Its weight ($\alpha_k$) takes into account how well it has classified the training data. They assign greater weights to a high-precision weak learner. The sample weights (D) are increased on mis-classified examples and decreased on correctly classified ones using a weak classifier. Then, the existing weak learner combines the prediction by assigning each weak prediction with a weight ($\alpha_k$). To get the final output, F(x) is derived by averaging out the weak learner's prediction on the whole. These weak learner predictions are typically combined in a weighted sum. The final prediction (F(x)) can be expressed in equation (12):

$$F(x) = \sum_k \alpha_k \cdot M_k(x) \qquad (12)$$

Here, F(x) represents the final prediction for input data x. $\alpha_k$ is the weight assigned to the k-th weak learner, and $M_k(x)$ is the prediction of the k-th weak learner for input x. The training process shown in table 3 involves using the EfficientNet-B0 model and modified googlenet for the target dataset. The dataset is split into 40% for training, 20% for validation, and 40% for testing, with the first 10 layers of the model frozen for pre-trained feature retention, and fine-tuning is applied to the remaining layers. Hyperparameters include a mini-batch size of 32, an initial learning rate of $1\times10^{-4}$, L2 regularization of $1\times10^{-5}$, and a maximum of 15 epochs. The Adam optimizer is used with an 'l2norm' gradient threshold method to prevent gradient explosion.

Journal of Theoretical and Applied Information Technology
30th June 2025. Vol.103. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3 training process and parameters for EfficientNet B0 and Modified googlenet

| Aspect | Details |
|---|---|
| Training Process | - Model Selection: EfficientNet-B0, and Modified googlenet |
| | - Dataset Preparation: Data split into training (40%), validation (20%), and testing (40%). |
| | - Training Workflow: First 10 layers frozen; fine-tuning performed on remaining layers using augmented training data. |
| Hyperparameters | - Mini-Batch Size: 32 |
| | - Initial Learning Rate: $1 \times 10^{-4}$ |
| | - L2 Regularization: $1 \times 10^{-5}$ |
| | - Maximum Epochs: 15 |
| | - $Validation\ Frequency \left\lfloor \frac{Number\ of\ training\ set\ Images}{Batch\ Size} \right\rfloor$ |
| Optimization | - Optimizer: Adam (adaptive learning rates with momentum). |
| | - Gradient Threshold Method: 'l2norm' (to avoid gradient explosion). |
| Evaluation Metrics | - Accuracy: Mean of correct predictions; visualized in a confusion matrix. |
| | - ROC Curve and AUC: Trade-off between sensitivity and specificity, with AUC as a performance measure. |
| | - Per-Class Performance: Confusion matrix provides insights into individual class performance. |
| Regularization | - Data Augmentation: Reflections, translations, and scaling applied to increase training data variability. |
| | - L2 Regularization: Penalizes large weight values to reduce overfitting. |
| | - Freezing Layers: First 10 layers of EfficientNet are frozen to retain pre-trained features and avoid overfitting. |
| | - Early Stopping: Monitors validation performance to halt training if overfitting is detected. |

Evaluation metrics include accuracy, the ROC curve with AUC for performance, and per-class performance through a confusion matrix. Regularization strategies involve data augmentation, L2 regularization, freezing layers, early stopping, and dropout layers to reduce overfitting.

## 5. RESULTS AND DISCUSSION

The proposed menthod is built on MATLAB 2023a and the simulations are performed using NVIDIA RTX 3050 with 16 GB RAM.

a. Modified Places 365 GoogLeNet

The primary objective of this study was to develop an effective model for classifying chest X-ray images and to evaluate its performance. After 15 epochs of training, the model achieved a validation accuracy of 90.49%, as shown in Figure 5. Notably, the model's performance improved significantly during training, with validation accuracy rising from an initial 35.43% to a peak of 89.48%. The model demonstrated strong potential for accurately classifying chest X-ray images, particularly in distinguishing between COVID-19, normal, and viral pneumonia cases—an inherently challenging task due to the subtle differences in these conditions. The achieved accuracy, especially in identifying COVID-19 cases, is promising and highlights the model's potential to assist radiologists in diagnosing respiratory conditions. The use of GoogLeNet, a deep convolutional neural network, proved effective in extracting relevant features from the dataset. A rigorous de-identification process was applied to ensure patient privacy. This process involved anonymizing all personally identifiable information (PII), such as patient names, IDs, and other sensitive details, in compliance with ethical standards and data protection regulations. These measures ensured that the dataset used for training and evaluation maintained patient confidentiality while providing valuable data for research purposes.

The results indicate that the trained GoogLeNet model has the potential to assist healthcare professionals in classifying chest X-ray images, particularly for COVID-19 diagnosis. However, further research on larger and more diverse datasets is essential to confirm the model's clinical utility and generalizability. The confusion matrix and ROC curve, presented in Figures 6a and 6b, further illustrate the model's classification performance.

The model's validation accuracy reached its peak at 92.43% during epoch 11 but declined to

Journal of Theoretical and Applied Information Technology
30th June 2025. Vol.103. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

88.51% by epoch 14, signaling potential overfitting. Initially, training and validation errors decreased steadily through epochs 1–8, but validation performance began to diverge from training accuracy beyond epoch 12, indicating overfitting. Training accuracy remained stable between 93% and 97% by epoch 5, while validation accuracy fluctuated, widening the disparity. This overfitting issue could be mitigated through techniques such as regularization, data augmentation, or hyperparameter tuning. Furthermore, adopting a declining learning rate schedule could enhance the model's validation performance and reduce overfitting.

Overall, while the model performed well on the validation set up to epoch 11, additional modifications are necessary to improve its generalizability and prevent overfitting in subsequent training epochs.

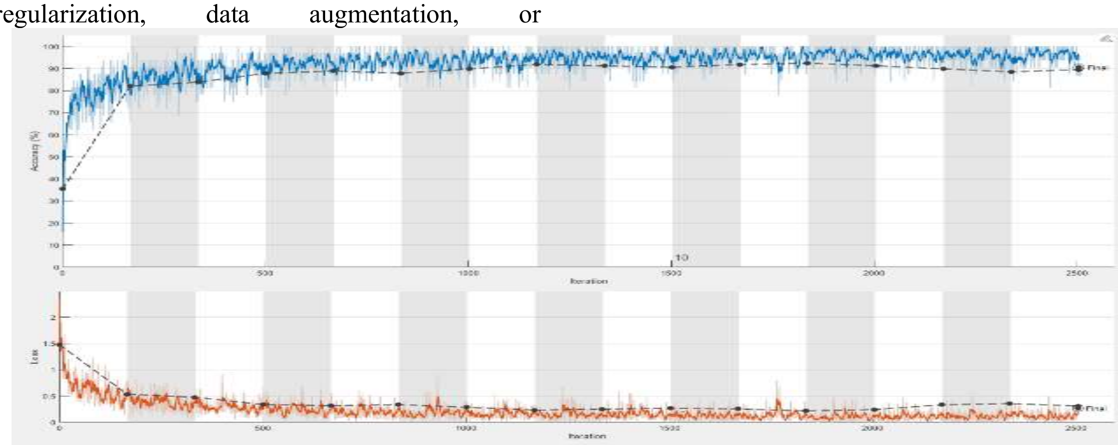
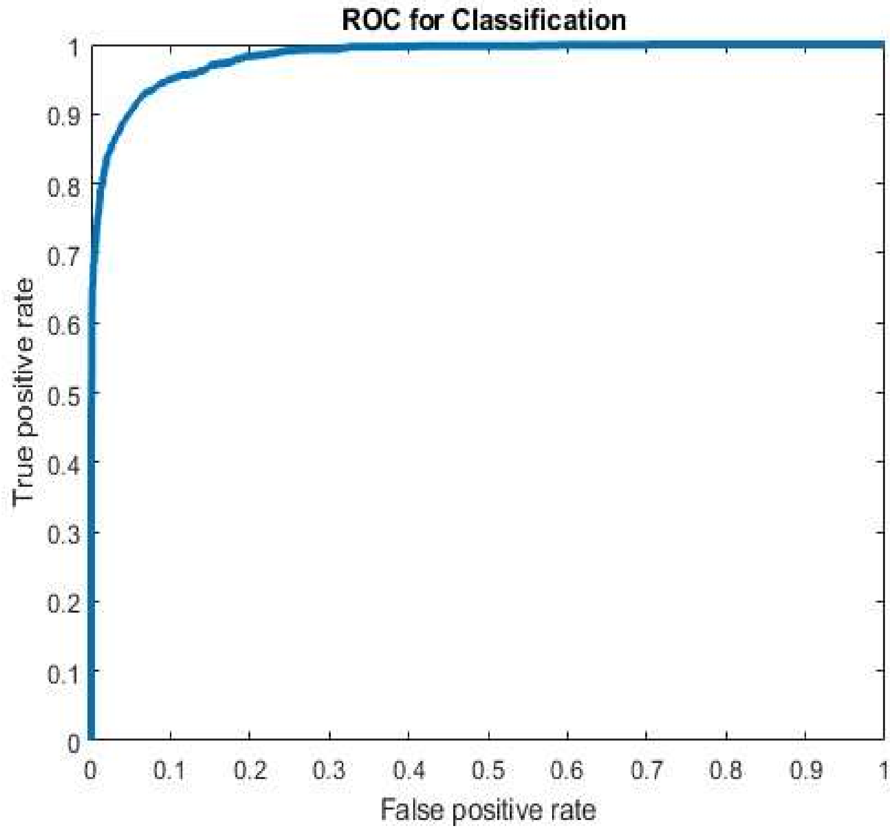


*Figure 5. Training Progress Of Places 365 Googlenet*


*(A)*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195




(B)

*Figure 6. A. Confusion Matrix And B. Roc Curve Of Places 365 Googlenet*

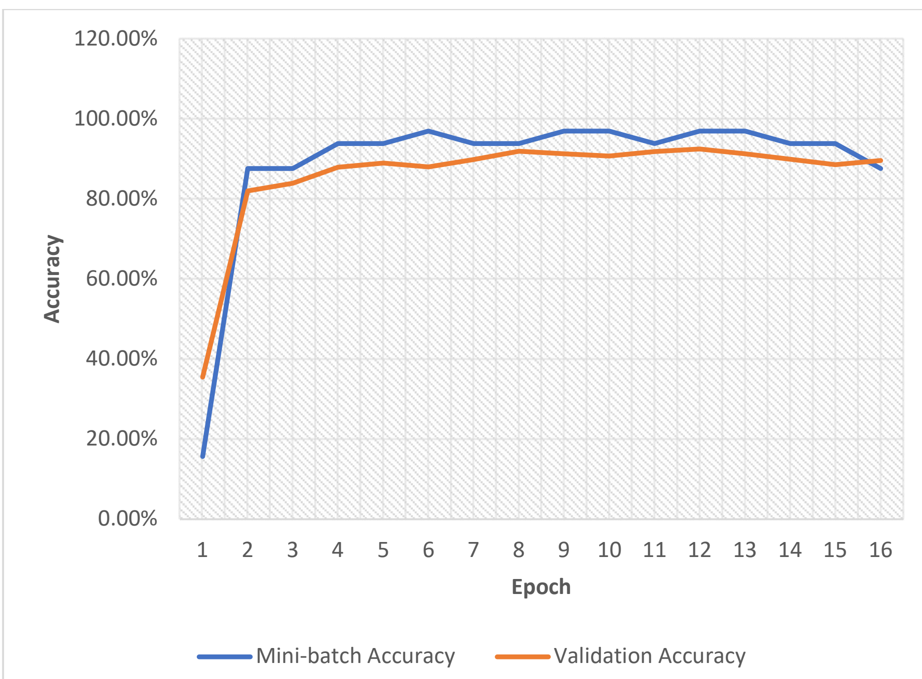



*Figure 7. Accuracy Vs Epoch*

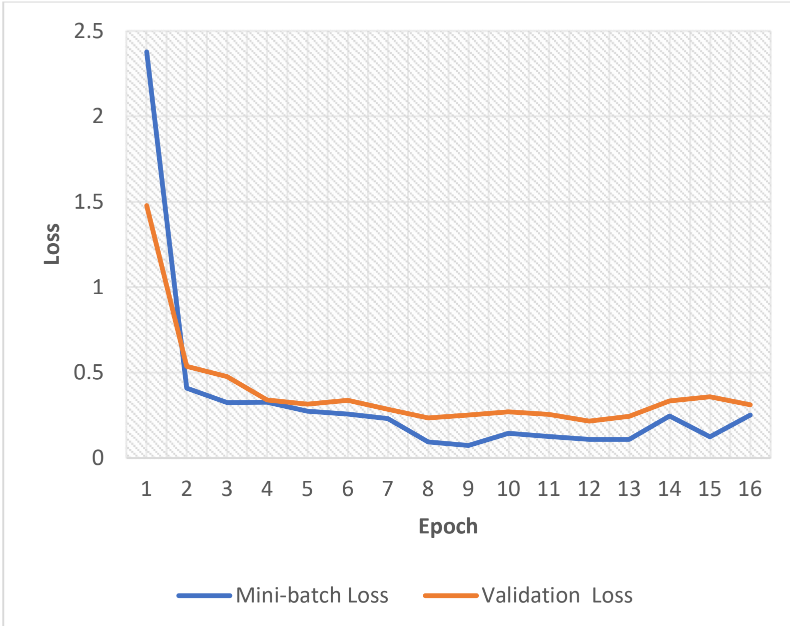



*Figure 8. Loss Vs Epoch*

*Table 4 Detailed Metrics Of Places 365 Googlenet*

| Class | Precision (%) | Recall (%) | F1-Score (%) | Accuracy (%) |
|---|---|---|---|---|
| COVID | 91.57 | 85.54 | 88.45 | 93.56 |
| Cardiomegaly | 81.35 | 89.83 | 85.38 | 97.69 |
| Normal | 91.13 | 83.95 | 87.39 | 89.13 |
| Viral Pneumonia | 77.82 | 97.12 | 86.41 | 94.27 |

The performance metrics for Places 365 GoogLeNet shown in table 4 reveal strong classification capabilities across most classes, with **COVID** and **Normal** achieving high precision (91.57% and 91.13%, respectively), though their slightly lower recall values (85.54% and 83.95%) indicate some misclassified true cases. **Cardiomegaly** demonstrates a balanced performance with an F1-score of 85.38%, reflecting reliable detection with minimal false positives and negatives. The model excels in identifying **Viral Pneumonia**, achieving the highest recall (97.12%), but its lower precision (77.82%) highlights challenges with misclassification into other categories. Overall, the results indicate robust classification with specific areas for improvement, particularly in optimizing recall for underrepresented conditions.

b. EfficientNet B0

EfficientNetB0 was trained on all samples with equal weights to emphasize their significance for the ensemble. After 15 epochs, the model achieved an impressive validation accuracy of 94.70%. The training lasted approximately 332 minutes and 39 seconds and utilized a constant learning rate of 0.0001. The model's progress was characterized by a consistent increase in both mini-batch and validation accuracies, demonstrating successful convergence during training. It achieved 100% accuracy on the mini-batches and 94.07% on the validation set, as shown in Figure 6. While the training began with low accuracy in the first epoch, dramatic improvements were observed, leading to high accuracy measures. A uniform decrease in the loss curves across the training process further supported the model's successful learning.

Despite its high overall accuracy, the model displayed signs of overfitting, with a slight deviation between mini-batch and validation accuracies. Figures 10a and 10b illustrate the confusion matrix and the ROC curve, which reflect the model's classification performance.

As depicted in Figures 11 and 12, validation accuracy steadily increased throughout the epochs, reaching a peak of 94.41% at epoch 14.

Journal of Theoretical and Applied Information Technology
30th June 2025. Vol.103. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

This trend indicates that the model continued to learn and improve based on the validation set. Training loss decreased significantly during the early epochs, reaching a low of 0.028 at epoch 10, suggesting the model's strong fit to the training data. However, validation loss fluctuated between 0.2465 and 0.3179 across epochs, signaling partial overfitting.

By epoch 6, the model achieved 100% training accuracy, confirming that it memorized the training data. However, the lower validation accuracy indicates that overfitting limited its ability to generalize. While a constant learning rate of 1e-4 was maintained during training, reducing the learning rate over time could enhance validation accuracy and mitigate overfitting.

In general, EfficientNetB0 demonstrated a high level of reliability, achieving exceptional performance on the validation set. Nevertheless, its generalization ability could be improved through stronger regularization techniques, data augmentation, and hyperparameter tuning. Addressing the disparity between training and validation performance is critical for enhancing the model's robustness and ensuring its applicability in real-world scenarios.

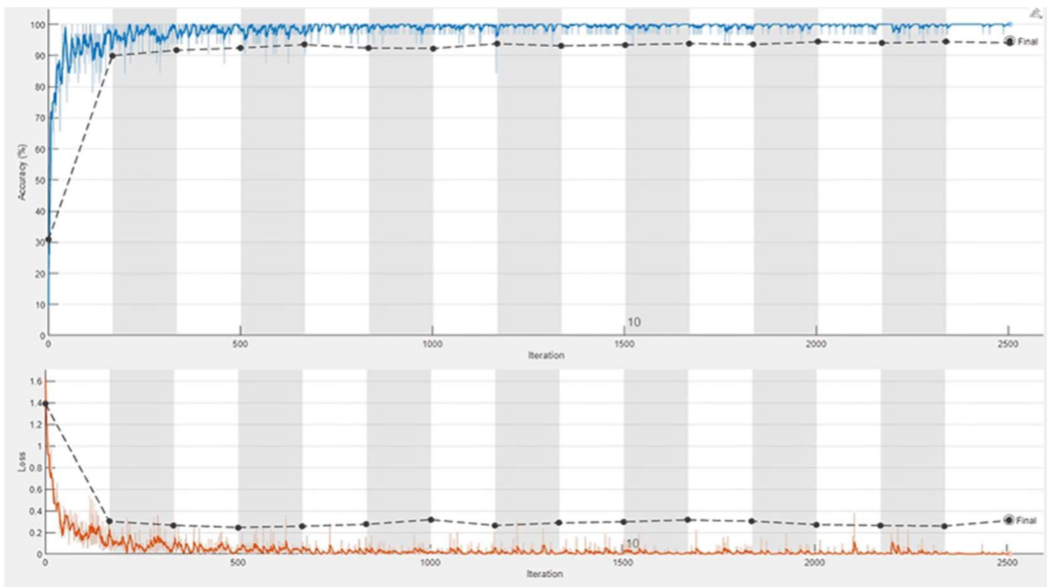
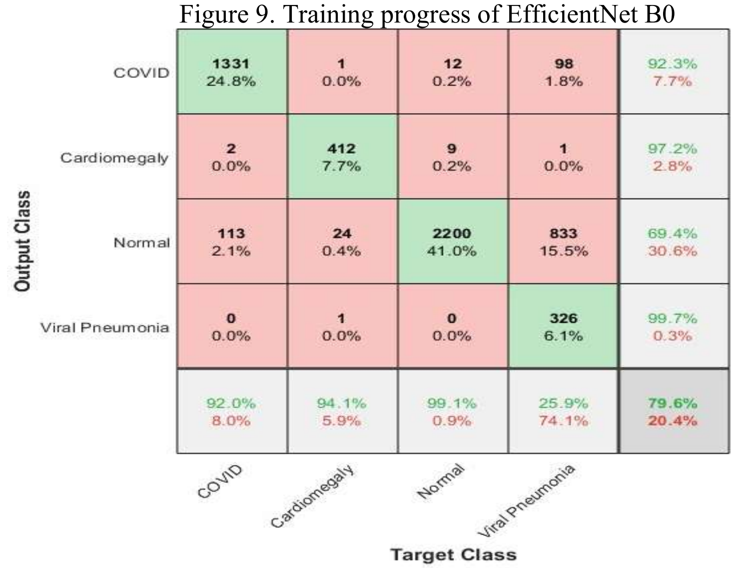


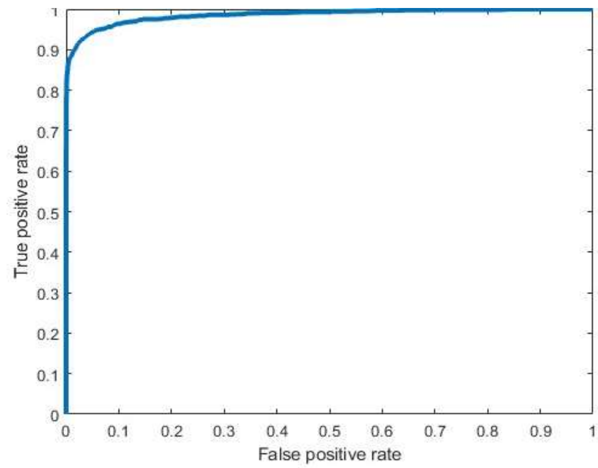

Figure 9. Training progress of EfficientNet B0




(a)

Journal of Theoretical and Applied Information Technology
30th June 2025. Vol.103. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195




(b)

*Figure 10. A. Confusion Matrix And B. ROC Curve Of Efficientnet B0*

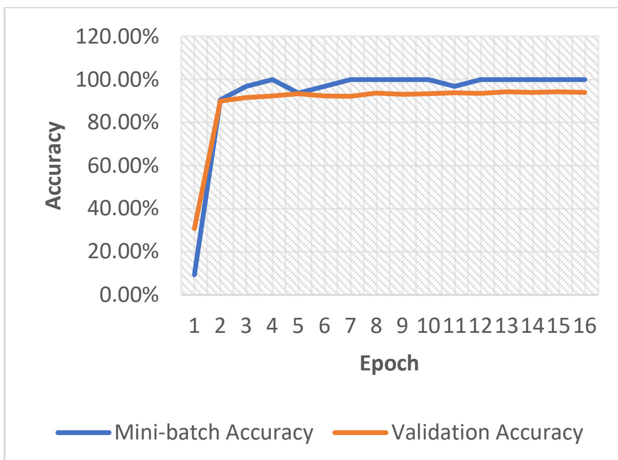



Figure 11. Accuracy Vs Epoch

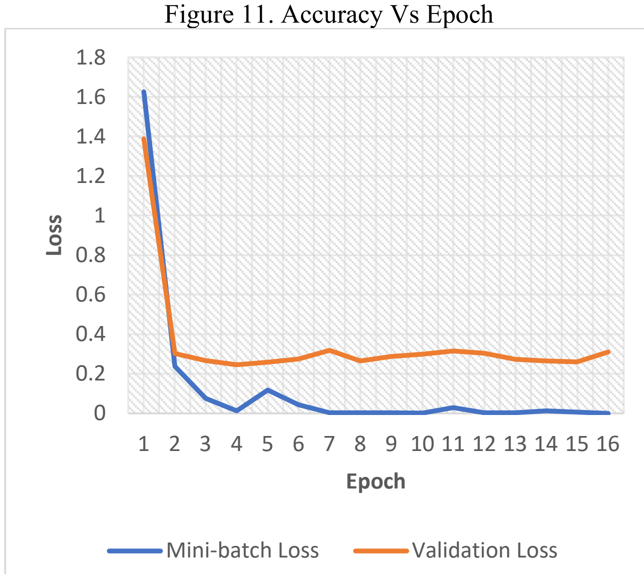



*Figure 12. Loss Vs Epoch*

Journal of Theoretical and Applied Information Technology
30th June 2025. Vol.103. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

*Table 5 Detailed Metrics Of Efficientnet B0*

| Class | Precision (%) | Recall (%) | F1-Score (%) | Accuracy (%) |
|---|---|---|---|---|
| **COVID** | 92.05 | 92.30 | 92.17 | 95.79 |
| **Cardiomegaly** | 94.06 | 97.40 | 95.70 | 99.31 |
| **Normal** | 99.05 | 69.40 | 81.62 | 81.52 |
| **Viral Pneumonia** | 25.93 | 99.69 | 41.16 | 82.62 |

The performance metrics derived from the confusion matrix provide a comprehensive evaluation of the EfficientNet B0 model's classification ability across different classes shown in table 5. The model demonstrates high precision and recall for the COVID and Cardiomegaly classes, with F1-scores of 92.17% and 95.70%, respectively, indicating its robustness in identifying these conditions with minimal false positives and negatives. Similarly, the Normal class achieved an exceptional precision of 99.05%; however, its relatively lower recall (69.40%) and F1-score (81.62%) suggest potential misclassification of "Normal" cases into other classes. The Viral Pneumonia class, despite achieving a remarkable recall of 99.69%, suffers from low precision (25.93%) and F1-score (41.16%), highlighting significant challenges with false positives. The model's accuracy varies across classes, ranging from 81.52% (Normal) to 99.31% (Cardiomegaly), indicating a need for further refinement, particularly in imbalanced or overlapping classes. These results emphasize the model’s strengths and limitations, forming a foundation for optimization and comparative analysis with state-of-the-art techniques

c. Ensemble method

The integration of Places365 GoogLeNet and EfficientNetB0 using the AdaBoost ensemble method yielded highly promising results in lung disease classification. Among the many methods, we choose the AdaBoost due to its less computation power and successful in improving the classification performance. In AdaBoost, unlike Bagging, whereby models are trained on distinct subsets of the training data, and their predictions averaged to reduce variance, the algorithm assigns greater weight to previously misclassified inputs, resulting in a more effective boosted approach when the base learners, in this case GoogLeNet and EfficientNetB0, are already capable of classifying specifically the features of interest. In contrast to Stacking, which trains meta-learner on the results of several models and thus significantly increases computational complexity, AdaBoost uses weighted aggregation which is a good trade between performance gain and feasibility for real world usage. The ensemble model achieved an accuracy rate of 92.47%, demonstrating its efficacy in distinguishing between the various classes within the dataset. Moreover, the validation accuracy peaked at 96.48%, highlighting the model's robustness and strong performance on unseen data. This represents a significant improvement compared to the individual performance of each constituent model. The AdaBoost ensemble effectively mitigated the overfitting challenges observed in the standalone models, resulting in a more balanced overall performance. By leveraging the complementary strengths of Places365 GoogLeNet’s advanced multi-scale feature extraction and EfficientNetB0’s computational efficiency, the ensemble method enhanced classification accuracy and generalizability. These findings underscore the potential of ensemble techniques, such as AdaBoost, in improving the performance of deep learning models for medical image classification. The demonstrated improvement in classification accuracy highlights the advantages of combining diverse architectures to address the challenges of overfitting and class imbalance. This approach provides a compelling direction for future research, with implications for broader applications in medical imaging and other complex classification tasks.

d. Comparative Analysis

The ensemble model outperforms both Places 365 GoogLeNet and EfficientNet B0 individually, as evident from the overall accuracy and per-class performance metrics. The ensemble approach achieved a higher combined accuracy of 93.16%, effectively balancing the strengths of both models. For COVID, the ensemble model maintained an accuracy of 93.16%, while Cardiomegaly achieved an improved accuracy of 99.31%. The Normal class reached 89.13%, and Viral Pneumonia showed an accuracy of 94.27%. These results reflect significant improvements over the individual models, demonstrating the effectiveness of the ensemble technique in leveraging complementary strengths.

Places 365 GoogLeNet displayed strong precision for COVID and Normal (91.57% and 91.13%) but had slightly lower recall (85.54% and 83.95%), leading to some misclassified true cases. EfficientNet B0 excelled in recall for Viral Pneumonia (99.69%) and Cardiomegaly (97.40%) but struggled with precision for Viral Pneumonia (25.93%). The ensemble model harmonized these disparities, achieving more balanced precision and recall across classes. For instance, the ensemble achieved an F1-score of 93.16%, indicating reliable classification across all classes, outperforming both individual models. Cardiomegaly's F1-score improved from 85.38% (Places 365 GoogLeNet) and 95.70% (EfficientNet B0) to a more stable value in the ensemble, benefiting from reduced false positives and negatives.

e. Individual Model Strengths and Combined Performance

Places 365 GoogLeNet demonstrated high precision for most classes, particularly COVID (91.57%) and Normal (91.13%), making it effective in reducing false positives. It also provided balanced performance across classes, albeit with slightly lower recall values. On the other hand, EfficientNet B0 showcased an exceptional recall for Viral Pneumonia (99.69%) and Cardiomegaly (97.40%), ensuring nearly all true cases were correctly identified. Additionally, it achieved high precision for Normal cases (99.05%), though this was offset by lower recall.

By integrating GoogleNet's precision and EfficientNet B0's recall, the ensemble model achieved better balance and overall performance. The ensemble approach addressed misclassification issues present in individual models, as evident in improved F1-scores and accuracy metrics. It proved robust to class imbalances and overlapping features between conditions, leveraging complementary strengths of both models to deliver enhanced results.

f. Implications, Limitations, and Future Research Directions

The ensemble approach highlights the potential of combining diverse architectures to enhance classification performance in medical imaging tasks. Improved diagnostic accuracy for conditions like COVID and Cardiomegaly can assist in early and reliable detection, aiding healthcare professionals in timely decision-making. The balanced performance across classes suggests that the ensemble model can generalize better across varying data distributions. Despite these improvements, the ensemble model's computational complexity is higher due to the need for running multiple models, which may limit its deployment in resource-constrained environments. Additionally, certain classes (e.g., Viral Pneumonia with lower precision in EfficientNet B0) still face challenges with misclassification. The reliance on labeled datasets may also limit the model's performance when deployed on unseen or noisy data. Future research directions should focus on optimizing lightweight ensemble techniques or knowledge distillation methods to reduce computation overhead while retaining performance. Advanced data augmentation techniques or synthetic data generation could be employed to address class imbalances and improve model robustness. Incorporating explainable AI techniques could help better understand the model's decision-making process, particularly for critical misclassifications. Fine-tuning the ensemble model on domain-specific datasets from other medical imaging tasks could also evaluate its generalizability. Finally, validating the model on larger, diverse, and real-world datasets will be crucial to assess its robustness and scalability.

The ensemble model successfully leverages the strengths of Places 365 GoogLeNet and EfficientNet B0, demonstrating improved accuracy, precision, and recall across multiple classes. For instance, the ensemble model achieved a combined accuracy of 93.16%, with notable improvements in precision and recall across all target conditions. While limitations such as computational complexity and misclassification challenges remain, the results underscore the potential of ensemble methods in advancing medical imaging classification. Future work should focus on optimizing the ensemble's efficiency and generalizability to further enhance its clinical utility.

In Table 3, it is evident that the combined approach of Places 365 GoogLeNet and EfficientNetB0 in the model achieves an accuracy of 97.48%, surpassing the range of 88.23% to 94.41% observed in other models. This indicates that using this ensemble approach significantly improves performance. The proposed model also exhibits a recall rate of 95.24%, meaning it excels at identifying positive COVID-19 cases compared

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to the recall rates of 86.22% to 93.46% observed in other models. Additionally, the proposed model demonstrates precision and F1 score values of 94.67% and 96.24%, respectively, showcasing its strong positive predictive value as well as a balanced combination of precision and recall. The specificity of the proposed model stands at 95.46%, which is comparable to other models, affirming its capability to accurately identify negative cases. Overall, the ensembled model outperforms all other deep learning models across various evaluation metrics, particularly in terms of accuracy, recall, precision, and F1 score. This clearly demonstrates how integrating Places 365 GoogLeNet and EfficientNetB0 enhances COVID-19 classification beyond the benchmarks set by other models.

*Table 3: Comparison Of The Proposed Model And DL-Based Models For COVID-19 Classification:*

| Model | Precision | Recall | F1-score | Specificity | Accuracy |
|---|---|---|---|---|---|
| Inception-ResNet [39] | 89 | 86.22 | 95.53 | 89.72 | 88.23 |
| DenseNet-201 [40] | 91 | 91.48 | 95.46 | 95.63 | 92.23 |
| Loey et al. [41] | 93 | 92.65 | 93.11 | 92.81 | 94.41 |
| Sakib et al. [42] | 92 | 93.41 | 92.24 | 93.46 | 92.24 |
| Shibly et al. [43] | 94 | 92.44 | 93.42 | 92.29 | 93.78 |
| Proposed Ensemble (Places 365 GoogLeNet + EfficientnetB0) | 94.67 | 95.24 | 96.24 | 95.46 | 97.48 |

## 6. CONCLUSION

In light of the COVID-19 pandemic, this research focused on developing a robust diagnostic tool for respiratory diseases, specifically targeting COVID-19, cardiomegaly, and viral pneumonia. By combining Places 365 GoogLeNet and EfficientNetB0 using the AdaBoost ensemble method, we achieved significant progress in classifying chest X-ray images. The ensemble model demonstrated superior performance, reaching a validation accuracy of 96.48% and an overall accuracy of 97.48%, outstripping the individual models. Notably, the ensemble method effectively mitigated overfitting and improved generalization by balancing the strengths of both models. This approach was particularly beneficial for conditions such as COVID-19 and cardiomegaly, achieving notable precision and recall improvements. These findings highlight the potential of ensemble techniques to enhance classification accuracy and generalization in medical image analysis, offering a promising tool for healthcare professionals. Recent advancements in deep learning have significantly enhanced the classification of lung diseases using chest X-ray images. Deepak and Bhat (2025) introduced a multi-stage deep learning approach that effectively classifies various lung conditions, demonstrating the potential of such methodologies in improving diagnostic accuracy. Our study aligns with these findings, as our ensemble framework combining GoogLeNet and EfficientNetB0 has shown substantial improvements in detecting lung diseases in COVID-19 patients. This convergence underscores the transformative impact of AI-driven approaches in medical diagnostics, particularly in addressing the complexities introduced by the COVID-19 pandemic. Moving forward, future research should focus on validating this framework with more diverse datasets, conducting real-world trials, and exploring lightweight optimization techniques to ensure its broader applicability. This work has the potential to significantly impact clinical practice by

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

providing accurate, reliable, and timely diagnostics, improving patient outcomes in critical healthcare settings.

## 7. REFERENCES


[1] Siddiqi, Raheel, & Sameena Javaid(2024). "Deep learning for pneumonia detection in chest x-ray images: A comprehensive survey." Journal of imaging 10.8: 176.

[2] Shriwas, Pradeep Kumar, B. Sundaravadivazhagan, & Yogita Manish Patil(2025). "Role of AI in Diagnosis of Viral Infection." In Role of Artificial Intelligence, Telehealth, and Telemedicine in Medical Virology, Singapore: Springer Nature, Singapore ,pp. 31-51.

[3] Godbin, A. Beena, & S. Graceline Jasmine(2024). "Leveraging Radiomics and Genetic Algorithms to Improve Lung Infection Diagnosis in X-Ray Images using Machine Learning." IEEE Access.

[4] Salmi, Mabrouka, Dalia Atif, Diego Oliva, Ajith Abraham, & Sebastian Ventura(2023). "Handling imbalanced medical datasets: review of a decade of research." Artificial Intelligence Review 57, no. 10: 273.

[5] Shelke A, Inamdar M, Shah V, Tiwari A, Hussain A, Chafekar T, Mehendale N (2021) Chest X-ray classification using deep learning for automated COVID-19 screening. SN Comput Sci 2(4):1–9

[6] Foody, g. M., mcculloch, m. B., & yates, w. B. (1995). The effect of training set size and composition on artificial neural network classification. International Journal of Remote Sensing, 16(9), 1707–1723. https://doi.org/10.1080/01431169508954507

[7] Rahman T, Chowdhury ME, Khandakar A, Islam KR, Islam KF, Mahbub ZB et al (2020) Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. Appl Sci 10(9):3233

[8] Kermany DS, Goldbaum M, Cai W, Valentim CC, Liang H, Baxter SL et al (2018) Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 172(5):1122–1131.

[9] Qjidaa M, Mechbal Y, Ben-Fares A, Amakdouf H, Maarouf M, Alami B, Qjidaa H (2020) Early detection of COVID-19 by deep learning transfer model for populations in isolated rural areas. In 2020 International Conference on Intelligent Systems and Computer Vision (ISCV). IEEE, pp 1–5

[10] Rajaraman S, Candemir S, Kim I, Thoma G, Antani S (2018) Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. Appl Sci 8(10):1715

[11] Szegedy, C., Liu, N. W., Jia, N. Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1–9. https://doi.org/10.1109/cvpr.2015.7298594.

[12] Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., & Torralba, A. (2017). Places: a 10 million image database for scene recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(6), 1452–1464. https://doi.org/10.1109/tpami.2017.2723009 .

[13] Tan, Mingxing and Quoc V. Le.(2019). "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." ArXiv abs/1905.11946 : n. pag.

[14] Shelke, Ankita, et al.2021. "Chest X-ray classification using deep learning for automated COVID-19 screening." SN computer science 2.4: 300.

[15] Bar Y, Diamant I, Wolf L, Lieberman S, Konen E, Greenspan H (2018) Chest pathology identification using deep feature selection with non-medical training. Comput Methods Biomech Biomed Engin: Imaging Visualization 6(3):259–263

[16] Qiblawey, Y., Tahir, A., Chowdhury, M. E., Khandakar, A., Kiranyaz, S., Rahman, T., ... & Ayari, M. A. (2021). Detection and severity classification of COVID-19 in CT images using deep learning. Diagnostics, 11(5), 893.

[17] Barhoom, Alaa MA (2019) Pneumonia diagnosis using deep learning, PhD dissertation., Al-Azhar University-Gaza

[18] Karar ME, Hemdan EED, Shouman MA (2021) Cascaded deep learning classifier for computer-aided diagnosis of COVID-19 and pneumonia diseases in X-ray scans. Complex Intell Syst 7(1):235–247

[19] Pandit MK, Banday SA, Naaz R, Chishti MA.(2021) Automatic detection of COVID-19 from chest radiographs using deep learning. Radiography ;27(2):483–9.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[20] Sekeroglu B, Ozsahin I.(2020). Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks. SLAS TECHNOLOGY: Translating Life Sciences Innovation.25(6):553-565. doi:10.1177/2472630320958376

[21] Khan, Asif Iqbal, Junaid Latief Shah, and Mohammad Mudasir Bhat.(2020). "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images." Computer methods and programs in biomedicine 196: 105581.

[22] Yamac¸ M, Ahishali M, Degerli A, Kiranyaz S, Chowdhury ME, Gabbouj M.(2021). Convolutional sparse support estimator-based covid-19 recognition from x-ray images. IEEE Trans Neural Networks Learn Syst 2021;32(5):1810–20.

[23] Nahiduzzaman Md, Goni MOF, Anower MS, Islam MR, Ahsan M, Haider J, et al.(2021) ‘‘A novel method for multivariant pneumonia classification based on hybrid CNN-PCA based feature extraction using extreme learning machine with CXR images”, IEEE. Access;9:147512–26.

[24] Chandra TB, Verma K, Singh BK, Jain D, Netam SS.(2021) Coronavirus disease (COVID-19) detection in chest x-ray images using majority voting-based classifier ensemble. Expert Syst Appl;165 113909

[25] Sahlol, Ahmed T., et al.(2020) "A novel method for detection of tuberculosis in chest radiographs using artificial ecosystem-based optimization of deep neural network features." Symmetry 12.7: 1146.

[26] Nahiduzzaman M, Islam MR, Hassan R. ChestX-Ray6: Prediction of multiple diseases including COVID-19 from chest X-ray images using convolutional neural network.

[27] Serte S., Serener A.(2020) Early pleural effusion detection from respiratory diseases including COVID-19 via deep learning. IEEE. 2020;2020:1–4.

[28] Chandra, Tej Bahadur, et al.(2020) "Automatic detection of tuberculosis-related abnormalities in Chest X-ray images using hierarchical feature extraction scheme." Expert Systems with Applications 158: 113514.

[29] Rahman, Tawsifur, et al.(2020) "Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization." IEEE Access 8 : 191586-191601.

[30] Rajpurkar, Pranav, et al.(2017) "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning." arXiv preprint arXiv:1711.05225.

[31] Huang, Gao, et al.(2017) "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition.

[32] Van der Maaten, Laurens, and Geoffrey Hinton.(2008) "Visualizing data using t-SNE." Journal of Machine Learning Research 9.11.

[33] Rahman, Muhammad, et al. (2021)"Deep pre-trained networks as a feature extractor with XGBoost to detect tuberculosis from chest X-ray." Computers & Electrical Engineering 93: 107252.

[34] Chen, Tianqi, et al.(2015) "Xgboost: extreme gradient boosting." R package version 0.4-2 1.4: 1-4.

[35] Tan, Mingxing, and Quoc Le.(2019) "Efficientnet: Rethinking model scaling for convolutional neural networks." International conference on machine learning. PMLR.

[36] Kanayama, A., Tsuchihashi, Y., Otomi, Y. et al.(2022) Association of severe COVID-19 outcomes with radiological scoring and cardiomegaly: findings from the COVID-19 inpatients database, Japan. Jpn J Radiol 40, 1138–1147. https://doi.org/10.1007/s11604-022-01300-2

[37] M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I. Reaz, M. T. Islam,(2020) “Can AI help in screening Viral and COVID-19 pneumonia?” IEEE Access, Vol. 8, pp. 132665

[38] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM.(2017) ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR 2017.

[39] Chen, Yunfeng, et al.(2022) "Classification of lungs infected COVID-19 images based on inception-ResNet." Computer methods and programs in biomedicine 225 : 107053.

[40] Akl, Ahmed A., et al.(2023) "A hybrid CNN and ensemble model for COVID-19 lung infection detection on chest CT scans." Plos one 18.3: e0282608.

[41] Loey M, Smarandache FM, Khalifa NE. M Khalifa N E,(2020) ‘‘Within the lack of chest

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning”. Symmetry Apr. 2020;12(4):651.

[42] Sakib, Sadman, Tahrat Tazrin, Mostafa M. Fouda, Zubair Md Fadlullah, and Mohsen Guizani.(2020) "DL-CRC: deep learning-based chest radiograph classification for COVID-19 detection: a novel approach." Ieee Access 8: 171575-171589.

[43] Shibly KH, Dey SK, Islam MTU, et al. (2020)COVID faster R-CNN: A novel framework to Diagnose Novel Coronavirus Disease (COVID-19) in X-Ray images. Inf Med Unlocked Aug.;20 100405.

[44] Deepak, G. D., & Bhat, S. K. (2025). A multi-stage deep learning approach for comprehensive lung disease classification from x-ray images. *Expert Systems with Applications*, 127220.