Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 3, July 2026, pp: 1047-1059 e-ISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

Dynamic Fine-Tuning Strategy of Deep Learning Models for Lung Disease Classification on Chest X-ray Images

Phuoc-Hai Huynh[1, 2] **and Thi-Diem Truong** [1, 2]

[1] Faculty of Information Technology, An Giang University, An Giang, Vietnam

[2] Vietnam National University HoChiMinh City, HoChiMinh City, Vietnam

Corresponding author: Phuoc-Hai Huynh (e-mail: hphai@agu.edu.vn), **Author(s) Email**: Thi-Diem Truong (email: ttdiem@agu.edu.vn)

Abstract: Lung diseases remain a leading cause of life-threatening illnesses worldwide, particularly in developing countries with limited healthcare resources. In recent years, deep convolutional neural networks (CNNs) have demonstrated strong potential in the automatic interpretation of chest X-ray (CXR) images. However, existing approaches often rely on rigid two-stage fine-tuning or fixed-step progressive unfreezing strategies, which may fail to effectively adapt pretrained representations or destabilize optimization, especially when applied to imbalanced real-world datasets. This study proposes a validation-driven dynamic fine-tuning strategy for transfer learning that adaptively unfreezes network layers based on convergence signals observed on the validation set rather than predefined training epochs. By coupling the timing and depth of adaptation to generalization behavior, the proposed method enables controlled knowledge transfer while mitigating catastrophic forgetting and improving training stability. Experiments were conducted on a large-scale, real-world clinical dataset comprising 15,416 CXR images collected at An Giang Regional General Hospital, Vietnam. The proposed strategy was systematically evaluated across multiple CNN backbones, including Xception, DenseNet121, EfficientNetV2S, InceptionV3, MobileNet, ResNet50, and VGG16. Performance was assessed using overall accuracy and macro-F1 score to address class imbalance. Results demonstrate consistent improvements across all architectures, with a mean accuracy gain of 3.18% compared to conventional fine-tuning (p = 0.02). MobileNet achieved the best performance with 85.1% accuracy and 66.8% macro-F1, while maintaining a compact model size of 73.05 MB. These findings indicate that validation-driven dynamic fine-tuning provides a stable, statistically significant, and deployment-feasible transfer learning mechanism suitable for real-world clinical environments.

Keywords **Dynamic Fine-tuning; Deep Learning; Lung diseases classification; Chest X-ray images**

I. Introduction

Lung diseases are a major public health burden, particularly in developing countries where access to advanced diagnostic facilities is limited [1], [2] . Chest X-ray (CXR) is the most widely used imaging modality for lung disease screening owing to its cost-effectiveness [3], [4]. However, accurate CXR interpretation requires specialized expertise, and diagnostic performance can be affected by inter-reader variability and heavy clinical workload [5], [6]. Deep learning based on convolutional neural networks (CNNs) has emerged as an important tool for medical image analysis [7], [8] . Despite promising results, training CNNs from scratch is often impractical in medical domains because annotated datasets are typically limited in size [9]. Consequently, transfer learning using ImageNet-pretrained models has become a de facto strategy for CXR classification [10], [11].

Most existing fine-tuning pipelines follow a rigid two-stage paradigm: (i) freezing the backbone while training the classification head, followed by (ii) unfreezing all layers for joint training. While effective to some extent, this procedure can be suboptimal in two key aspects. First, it may underutilize the representational capacity of the pretrained backbone: once the classification head saturates, keeping the backbone frozen prevents higher-level feature extractors from adapting to domain-specific radiographic patterns, forcing the model to separate classes using features that remain biased toward ImageNet semantics. Second, unfreezing the entire network simultaneously can destabilize optimization and increase the risk of catastrophic forgetting, because the sudden increase in trainable parameters may cause large gradient updates that overwrite broadly transferable low-level features. For example, subtle CXR cues such as interstitial opacities, pleural lines, or diffuse haze often

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

require adaptation in mid-to-deep layers. If these layers remain frozen, the classifier may rely on weak or spurious signals. Conversely, unfreezing all layers at once may unnecessarily perturb early edge or texture filters.

To address these limitations, we propose a dynamic fine-tuning strategy that adaptively unfreezes backbone layers based on validation convergence signals. This validation-driven approach enables more stable and efficient knowledge transfer from natural-image pretraining to the target CXR domain. The main contributions of this work are as follows:

1. We propose a validation-driven dynamic fine-tuning strategy that performs progressive unfreezing triggered by validation convergence signals, rather than following a predetermined epoch-based schedule as in conventional gradual/progressive unfreezing. This mechanism reduces manual schedule tuning and adapts the depth and timing of backbone adaptation to the observed training dynamics.
2. We systematically evaluate the proposed strategy across multiple CNN backbones for seven-class lung disease classification.
3. We show that the proposed strategy improves classification performance while maintaining computational efficiency, supporting deployment in resource-limited clinical settings.
4. We construct and curate a large-scale real-world clinical dataset comprising 15,416 paired chest X-ray images and clinical notes collected from An Giang Regional General Hospital, Vietnam. This dataset reflects diverse disease patterns and undergoes expert validation and rigorous preprocessing, contributing a valuable benchmark for medical AI research.

II. Related works

Convolutional neural networks (CNNs) have become the dominant paradigm for medical image analysis, particularly in chest X-ray (CXR) interpretation. A comprehensive deep-learning model for chest radiograph interpretation has been shown to improve clinicians' accuracy in routine practice, underscoring radiologist-comparable performance in real-world settings [12]. Subsequent works extended this approach to multi-label classification, covering a wide range of thoracic pathologies [13], [14]. Collectively, these works show that deep learning can extract clinically relevant features from CXR images, while also highlighting persistent challenges such as class imbalance, domain shift, and limited interpretability. Given the limited size of annotated medical datasets, transfer learning from ImageNet-pretrained models has become standard practice. Pre-trained CNN backbones such as VGG, ResNet, DenseNet, and EfficientNet have been widely adopted for CXR analysis [15], [16], [17], [18], [19]. Although transfer learning typically accelerates convergence and improves generalization, the fine-tuning strategy remains a key determinant of downstream performance.

Traditional fine-tuning methods generally follow a two-stage approach: (i) freezing the convolutional backbone while training only the classification head, and (ii) unfreezing all layers for joint optimization [20]. Although simple and effective, this rigid procedure may either underutilize the representational capacity of deeper layers or induce catastrophic forgetting. More sophisticated approaches, including gradual unfreezing and discriminative learning rates [20], [21] have been proposed to mitigate these limitations. However, most of these strategies depend on predefined or heuristic schedules (fixed epoch-based unfreezing plans or manually chosen layer-wise learning-rate ratios), which can be problematic in medical imaging for several reasons.

First, medical images exhibit substantial domain shift from natural images (e.g., grayscale appearance, different texture statistics, and subtle pathology-specific cues), meaning that the optimal depth and timing of adaptation can vary markedly across tasks and datasets. Second, clinical datasets are often smaller, noisier, and more imbalanced, and are subject to institution and acquisition-specific variability (scanner type, exposure, patient positioning, and prevalence patterns). Under such variability, a fixed unfreezing schedule may unfreeze too early (leading to unstable optimization and overfitting) or too late (leading to insufficient domain adaptation), while manually tuned discriminative learning rates may not transfer reliably across backbones or institutions.

Motivated by this gap, we propose a validation-driven dynamic fine-tuning strategy that adaptively unfreezes backbone layers based on convergence signals on the validation set. By coupling unfreezing decisions to observed generalization behavior rather than predetermined epochs, the proposed approach provides a more principled and data-responsive mechanism for transferring ImageNet representations to CXR classification.

III. Method

A. Overall Workflow

The proposed framework adopts a systematic deep learning pipeline for CXR image classification (in Fig.1). Firstly, raw CXR images undergo comprehensive preprocessing and data augmentation to enrich the diversity of the training set. The CNN backbone pre-trained models are employed to capture hierarchical visual representations.

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

The dynamic fine-tuning strategy is appliedto progressively adapt the pre-trained backbone to the target CXR domain. This adaptive process enables the model to refine its feature space toward domain-specific pathological patterns while mitigating overfitting. The resulting feature embeddings are then fed into a lightweight classification head, which outputs the final disease probabilities through a fully connected layer followed by a Softmax activation. The modular design of the system ensures both adaptability and computational efficiency.

Within the CNN backbone, stacked convolutional layers extract local image features such as edges, textures, and high-level structures; pooling layers perform spatial down-sampling to reduce

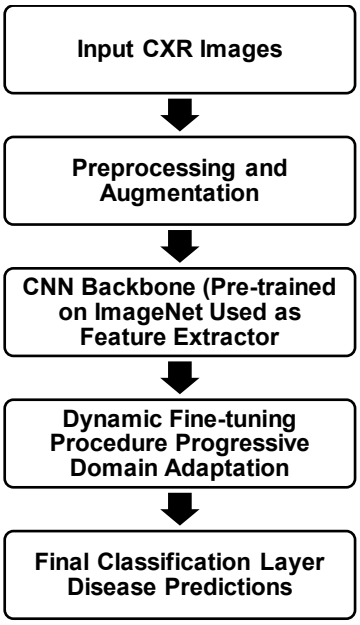

Fig. 1. Proposed framework for disease classification from Chest X-ray images

dimensionality while retaining salient information; and nonlinear activation functions introduce model expressiveness. Finally, the fully connected layers aggregate these learned representations into a compact vector that serves as the input to the prediction.

B. Image Preprocessing

Raw CXR images were first subjected to quality control to remove low-quality or blurry scans. All records were anonymized and assigned unique study identifiers prior to analysis. Each image was standardized to 1024×1024 pixels and intensity-normalized to reduce inter-image exposure variability. During model training, images were resized to the input resolution required by each backbone and converted to the expected input format. For ImageNet-pretrained backbones, we applied the corresponding backbone-specific input preprocessing and replicated the single grayscale channel to three channels when required. No additional denoising or contrast-enhancement operations were applied beyond the augmentation operations reported in Table 2.

C. Dynamic Fine-Tuning Strategy

Fine-tuning pretrained convolutional neural networks (CNNs) is widely used in medical image analysis. However, the conventional two-stage fine-tuning strategy, training a classification head followed by unfreezing all layers simultaneously, often leads to unstable optimization, underfitting in early stages, and catastrophic forgetting in later stages. This issue is particularly severe when training data are limited, and the target domain differs substantially from ImageNet.

To address these limitations, we propose a Dynamic Fine-Tuning Strategy that gradually adapts the pretrained backbone by progressively unfreezing groups of layers according to validation convergence. The key idea is to allow high-level, task-specific features to adapt first, while preserving low-level generic representations.

Let the training and validation datasets be defined as in Eq. (1) as follows [22]. In this formulation, x_i and x_j denote input chest X-ray images, y_i and y_j represent their corresponding class labels, N and N_v denote the number of training and validation samples, respectively, and C is the number of disease categories

$$D_{\text{tr}} = \{(x_i, y_i)\}_{i=1}^{N}, D_{\text{va}} = \{(x_j, y_j)\}_{j=1}^{N_v} \quad (1)$$

The model consists of a pretrained backbone $f(\cdot;\theta)$ and a classification head $g(\cdot;\phi)$. For an input image x, the network prediction is

$$z = g(f(x;\theta);\phi) \quad (2)$$

In Eq. (2) [22] $f(\cdot;\theta)$ denotes the pretrained backbone mapping, $g(\cdot;\phi)$ denotes the classification head, and $z \in \mathbb{R}^C$ represents the logit vector before Softmax normalization. After applying the Softmax function, the predicted class probabilities in Eq.(2) [22]:

$$\hat{p}_c(x) = \frac{\exp(z_c)}{\sum_{r=1}^{C} \exp(z_r)}, c = 1, \ldots, C \quad (3)$$

In Eq. (3) [22], $\hat{p}_c(x)$ denotes the predicted probability that input x belongs to class c, and z_c is the corresponding logit value for class c. The network is trained by minimizing the standard cross-entropy loss on the training set and validation set:

$$\mathcal{L}_{\text{tr}} = \frac{1}{N}\sum_{i=1}^{N}\left[-\log \hat{p}_{y_i}(x_i)\right] \quad (4)$$

Validation performance is evaluated using Eq. (4):

$$\mathcal{L}_{\text{va}} = \frac{1}{N_v}\sum_{j=1}^{N_v}\left[-\log \hat{p}_{y_j}(x_j)\right] \quad (5)$$

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

The validation loss $\mathcal{L}_{va}$ in Eq.(4) [23] is computed analogously to Eq. (3) [23] but evaluated on the validation dataset. Model parameters are updated using gradient descent:

$$\theta \leftarrow \theta - \eta \nabla_\theta \mathcal{L}_{\text{tr}}, \phi \leftarrow \phi - \eta \nabla_\phi \mathcal{L}_{\text{tr}}, \quad (6)$$

In Eq.(6) [24], η denotes the learning rate, and ∇_θ and ∇_ϕ represent the gradients of the training loss with respect to the backbone parameters θ and classification head parameters ϕ, respectively. L denotes the total number of ordered backbone layer groups. To control the adaptation of pretrained features, the backbone parameters θ are divided into L ordered layer groups (Eq.(7) [25]):

$$\theta = \{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(L)}\} \quad (7)$$

where $\theta^{(1)}$ represents shallow layers and $\theta^{(L)}$ represents deep layers close to the classification head. At each unfreezing stage r, where r indexes the unfreezing stage rather than the training epoch. The number of trainable backbone groups is increased as (Eq. (8) [25])

$$m_r = \min(m_{r-1} + k, L) \quad (8)$$

where k is a fixed step size. Accordingly, the set of unfrozen layers is

$$\mathcal{U}_r = \{\theta^{(\ell)} \mid \ell \geq L - m_r + 1\} \quad (9)$$

The set $\mathcal{U}_r$ in Eq. (9) [26] represents the collection of backbone layer groups that are allowed to update during the current training stage. Only parameters in $\mathcal{U}_r$ are updated during training (Eq. (10) [25], [27]):

$$\theta^{(\ell)} \leftarrow \theta^{(\ell)} - \eta \nabla_{\theta^{(\ell)}} \mathcal{L}_{\text{tr}}, \theta^{(\ell)} \in \mathcal{U}_r \quad (10)$$

During the initial stage ($m_0 = 0$), all backbone layers are frozen and only the classification head ϕ is trained. To stabilize optimization, all batch normalization layers are kept fixed throughout training:

$$\mu^{(t+1)} = \mu^{(t)}, (\sigma^2)^{(t+1)} = (\sigma^2)^{(t)} \quad (11)$$

In Eq. (11), μ and σ^2 denote the running mean and running variance of batch normalization layers, respectively, and t indexes the training iteration. This constraint ensures that batch normalization statistics remain fixed during fine-tuning. If the validation loss does not improve for p consecutive epochs, a convergence plateau is detected, where p denotes the patience parameter controlling the number of consecutive epochs without validation improvement. In this case, additional backbone layers are unfrozen according to Eq. (7) to Eq. (8), and the learning rate is reduced:

$$\eta \leftarrow \eta\,(1 - \text{decay}_x) \quad (12)$$

where $\text{decay}_x = 0.1$ in all experiments. To prevent performance degradation, the model parameters are restored to the best validation checkpoint before continuing training. In Eq. (12) [28], $decay_x \in (0,1)$ denotes the learning-rate decay ratio applied after each unfreezing event to stabilize optimization. The training process stops when all backbone groups are unfrozen or the maximum number of epochs is reached. The final model is selected as

$$(\theta^\star, \phi^\star) = \arg\min_t \mathcal{L}_{\text{va}}^{(t)} \quad (13)$$

In Eq. (13) [27], (θ^*, ϕ^*)denotes the parameter configuration that achieves the minimum validation loss during training. The complete training procedure is summarized in Algorithm 1, which directly follows the update rules and conditions defined in Eq. (10) to Eq. (13).

Algorithm 1. Dynamic Fine-Tuning

(1) Input:

- Training dataset $\mathcal{D}_{\text{tr}}$
- Validation dataset $\mathcal{D}_{\text{va}}$
- Maximum number of epochs n
- Patience parameter p
- Unfreezing step size k
- Learning-rate decay factor decay_x

(2) Output:

- Optimized model parameters (θ*, φ*)

(3) Initialization:

(4) m ← 0 *// current number of unfrozen* backbone groups
(5) U ← ∅ *// set of unfrozen backbone groups*
(6) Initialize learning rate η
(7) $L_{\text{best}} \leftarrow +\infty$ *// best validation loss*
(8) $\text{patience}_{\text{counter}} \leftarrow 0$
(9) $(\theta_{\text{best}}, \varphi_{\text{best}}) \leftarrow (\theta, \varphi)$
(10) Freeze all backbone groups
(11) Freeze all batch normalization layers
(12) DO
(13) FOR epoch t = 1 to n
(14) Train one epoch on $\mathcal{D}_{\text{tr}}$
(15) Update only parameters in U and φ
(16) Compute validation loss L_{va} on $\mathcal{D}_{\text{va}}$
(17) IF $L_{\text{va}} < L_{\text{best}}$ THEN
(18) $L_{\text{best}} \leftarrow L_{\text{va}}$
(19) $(\theta_{\text{best}}, \varphi_{\text{best}}) \leftarrow (\theta, \varphi)$
(20) $\text{patience}_{\text{counter}} \leftarrow 0$
(21) ELSE
(22) $\text{patience}_{\text{counter}} \leftarrow \text{patience}_{\text{counter}} + 1$
(23) END IF
(24) IF $\text{patience}_{\text{counter}} \geq p$ AND m < L THEN
(25) m ← min(m + k, L)
(26) U ← set of deepest m backbone groups
(27) Unfreeze backbone groups in U
(28) Freeze all batch normalization layers
(29) $\eta \leftarrow \eta\,(1 - \text{decay}_x)$
(30) $(\theta, \varphi) \leftarrow (\theta_{\text{best}}, \varphi_{\text{best}})$
(31) $\text{patience}_{\text{counter}} \leftarrow 0$
(32) END IF

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

(33) END FOR
(34) WHILE m < L
(35) Return (θ^*, φ^*) ← (θ_{best}, φ_{best})

Table 1. Distribution of CXR images by disease type and dataset split (train/test/validation).

ID	Disease Type	Total	Train (80%)	Val (10%)	Test (10%)
1	Pneumonia	10398	8318	1040	1040
2	COVID-19	2456	1965	245	246
3	Tuberculosis (TB)	1070	856	107	107
4	Pleural Effusion	718	574	72	72
5	Pneumothorax	404	324	40	40
6	Pulmonary Edema	265	212	26	27
7	Atelectasis	105	85	10	10
	Total	15,416	12,334	1,540	1,542

D. Dataset

The primary dataset is collected at An Giang Regional General Hospital, a major provincial medical center in Southern Vietnam. It comprises 15,416 paired CXR images and corresponding clinical examination notes obtained between January 2018 and December 2023. All images are acquired as part of standard care for patients treated for respiratory illnesses. Eligible cases are those with primary diagnoses in one of seven pulmonary categories: community-acquired pneumonia, COVID-19, tuberculosis, pleural effusion, pneumothorax, cardiogenic pulmonary edema, and atelectasis (see Fig. 2). Diagnoses are confirmed by clinical evaluation and ICD-10 discharge codes. All labels are independently verified by a board-certified radiologist and validated by the hospital's multidisciplinary review board to ensure diagnostic consistency. Label reliability was strengthened through multidisciplinary case review and consultation reports, and the final ground-truth labels were taken from the consensus diagnosis recorded in the EMR. Prior to dataset assembly, a quality-control step excluded non-diagnostic or technically compromised images. Table 1 shows a highly imbalanced class distribution. Pneumonia is the dominant category with 10,398/15,416 images (67.4%), followed by COVID-19 with 2,456 images (15.9%) and tuberculosis with 1,070 images (6.9%). The remaining classes are markedly underrepresented: pleural effusion 718 (4.7%), pneumothorax 404 (2.6%), pulmonary edema 265 (1.7%), and atelectasis 105 (0.7%). Overall, the majority-to-minority ratio (Pneumonia vs. Atelectasis) is approximately 99:1, indicating a severe long-tailed distribution that motivates the use of class-balanced sampling, class-weighted loss, and selective augmentation for minority classes. The dataset underwent preprocessing, including quality control (removing low-quality or blurry images), standardizing image resolution to 1024×1024 pixels, and applying intensity normalization. To comply with privacy regulations, all records are anonymized using unique study identifiers. Table 1 summarizes the class distribution, which exhibits severe imbalance, dominated by pneumonia cases. To avoid data leakage, the dataset is split at the patient level (80% training, 10% validation, 10% test). To minimize potential bias toward majority classes and ensure fair learning across categories, a combination of data-level and algorithm-level strategies is employed during model training.

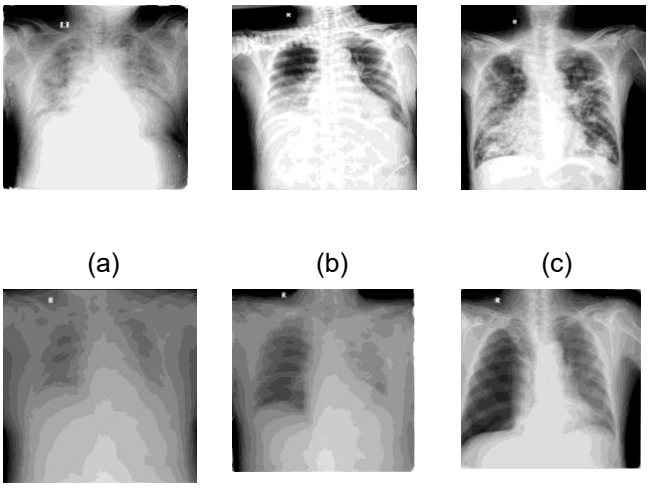

Fig. 2. Representative examples of Chest X-ray images from the An Giang clinical dataset, illustrating cases across multiple disease categories (Pneumonia (a), COVID-19 (b), Tuberculosis (c), Pleural Effusion (d), Pneumothorax (e) and Pulmonary Edema (f)). All images are anonymized and shown in their original frontal projection.

Weighted Sampling: A class-balanced sampling approach is implemented during mini-batch construction to guarantee that each batch contains a representative proportion of minority-class samples [29]. This mechanism helps maintain gradient diversity and prevents the network from being disproportionately influenced by the abundant pneumonia cases [30].

Class-Weighted Loss Function: The categorical cross-entropy loss is reweighted using Eq. (14) [31] class-specific coefficients inversely proportional to class frequencies [32], [33]:

$$w_i = \frac{N}{C \times n_{in}} \quad (14)$$

where N denotes the total number of samples, C the number of classes, and n_i the sample count of class i. This cost-sensitive weighting approximates a balanced

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

training objective by increasing the penalty for minority-class errors. Recent studies have further investigated weighted cross-entropy variants and weighting frameworks for long-tailed recognition, including validation-informed weighting mechanisms and balanced loss designs tailored to imbalanced medical imaging tasks [34], [35]. In our setting, Eq. (14) is chosen as a simple and stable baseline that introduces no additional loss-specific hyperparameters, and it integrates cleanly with our class-balanced sampling strategy and dynamic fine-tuning procedure.

Data Augmentation: To further enhance minority-class diversity, selective augmentation techniques were applied exclusively to underrepresented categories. These transformations aimed to increase sample variability and reduce overfitting tendencies among low-frequency classes [36]. The detailed augmentation operations and their parameter settings are summarized in Table 2.

Table 2. Data Augmentation Techniques and Parameter Settings

Techniques	Parameter Settings
Random Rotation	±10°
Brightness Adjustment	[0.8, 1.2]
Contrast Normalization	[0.9, 1.1]

IV. Results

We evaluate the classification performance of the proposed model on the X-ray image dataset. This study compares the proposed dynamic fine-tuning strategy with traditional fine-tuning approaches. The proposed method is evaluated on multiple pre-trained CNN backbones: Xception [37], DenseNet121 [21], EfficientNetV2S [38], InceptionV3 [39], MobileNet [40], ResNet50 [41], and VGG16 [42].

E. Experiments Setup

To compare and demonstrate the effectiveness of proposed models, we constructed seven deep learning models, including InceptionV3, MobileNet, ResNet50, VGG16, DenseNet121, Xception, and EfficientNetV2S. We use Python and several libraries to implement and evaluate our models. We built and trained the deep learning models using TensorFlow [43] and Keras [44]. All experiments were conducted on a Linux (Ubuntu 22) system equipped with an Intel Core i7-4790 CPU (3.6 GHz, 4 cores), 32 GB of RAM, and an NVIDIA GTX 1080Ti GPU (11 GB GDDR5X, 3584 CUDA cores). We measured classification accuracy and F1 scores on the test set, and recorded model load time (seconds) and on-disk model size (MB) to evaluate computational efficiency.

F. Training Configuration

To evaluate the robustness of the proposed dynamic fine-tuning strategy, we tuned four key hyperparameters: the patience parameter (p), the unfreezing step size (k), the learning-rate decay factor ($decay_x$), and the maximum training-epoch budget (n). For the patience parameter, we evaluated p = {3, 5, 7, 10}. Small patience values triggered premature unfreezing and unstable convergence, whereas larger values slowed adaptation without measurable performance gains. The best trade-off between convergence stability and classification performance was obtained with p = 5. For the unfreezing step size, we evaluated k = {1, 3, 5, 7}. Small increments resulted in slow adaptation, while larger increments occasionally caused overfitting in deeper architectures. An intermediate value of k = 3 provided the most consistent performance across all evaluated backbones.

Learning-rate decay factors of 10%, 30%, 50%, and 70% were also investigated. A 10% decay consistently achieved the best balance between optimization stability and final classification accuracy. Therefore, $decay_x$ = 0.1 was adopted in all subsequent experiments. To determine an appropriate training budget, we evaluated n = {50, 75, 100, 125} epochs. Extending training beyond 100 epochs produced only marginal performance improvements while substantially increasing computational cost. Consequently, n = 100 was selected as the optimal epoch budget. The final configuration used throughout the study was p = 5, k = 3, $decay_x$ = 0.10, and n = 100. This configuration consistently demonstrated stable convergence and strong performance across all evaluated CNN backbones.

Table 3. Performance of models trained with traditional fine-tuning (train n layers unfrozen).

Fine-Tuned Models	ACC (%)	F1 (%)	NB Layers (n)	Load Time (s)	Size (MB)
Xception	83	59.1	70	20.3	457.89
DenseNet121	82.1	56.5	174	45.8	182.88
EfficientNetV2S	75.5	28.2	260	65.66	449.95
InceptionV3	81.3	47.3	159	40.65	410.23
MobileNet	81.7	54.2	29	11.92	71.68
ResNet50	76	25.6	5	26.14	461.7
VGG16	80.2	34.4	10	5.35	258.18

G. Classification results

To further evaluate the effectiveness of the proposed dynamic fine-tuning strategy, we compared its performance against the traditional fine-tuning approach, in which only a fixed number of top layers (the layers closest to the classification head) of each backbone are unfrozen during training. For each backbone, the value of n reported in Table 3 was

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

Table 4. Performance of models trained with Dynamic fine-tuning

Dynamic Fine-Tuned Models	ACC (%)	F1 (%)	Load Time (s)	Size (MB)
Xception	83.2	64.8	20.18	383.07
DenseNet121	83.4	61.1	42.48	183.5
EfficientNetV2S	81.8	55.5	78.1	436.61
InceptionV3	83.1	62.8	39.08	419.68
MobileNet	**85.1**	**66.8**	**12.5**	**73.05**
ResNet50	82.9	63.5	21.59	467.92
VGG16	82.6	59	7.61	329.32

determined through a validation-based sensitivity analysis. Specifically, we evaluated a small set of candidate unfreezing depths by progressively increasing the number of trainable layers from the top of the backbone while keeping the remaining layers frozen. For each candidate configuration, models were trained under identical optimization settings, and the final $n$was selected based on the best validation performance while avoiding early overfitting.

In Table 3 and Table 4, Load Time (s) denotes the time required to load the trained model into memory for inference, while Size (MB) refers to the on-disk model

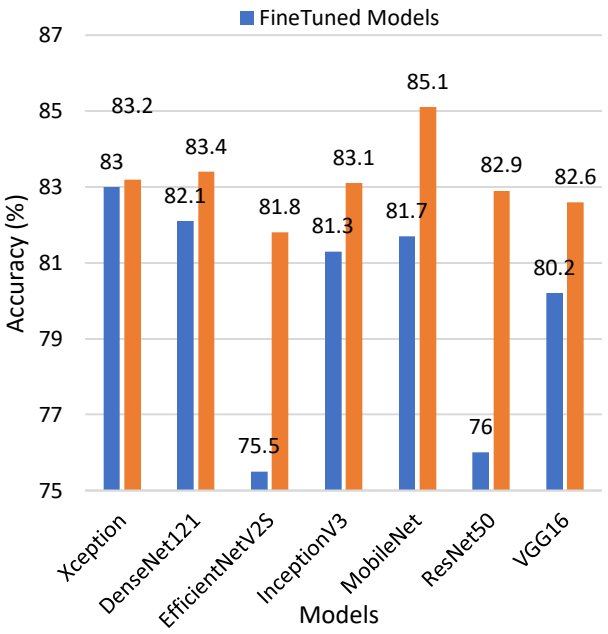

Fig. 3. Comparison of classification accuracy (%) between traditional fine-tuning models and proposed dynamic fine-tuning models across different CNN backbones.

size (the serialized model file after training). Model size is reported as a practical proxy for deployment efficiency, reflecting storage requirements and memory footprint during model loading in resource-constrained clinical environments.

Table 4 presents the performance of convolutional neural network backbones when trained using the proposed Dynamic Fine-Tuning strategy. Results show that all evaluated models benefited from the dynamic strategy. MobileNet achieved the highest accuracy of 85.1%, with low computational cost (load time 12.5 s, memory usage increase 73.05 MB), making it particularly suitable for deployment in real-world, resource-constrained clinical environments. DenseNet121 (83.4%), Xception (83.2%), and InceptionV3 (83.1%) also achieved strong performance, while maintaining moderate computational demands. DenseNet121 is the most memory-efficient among the high-performing models, requiring only 183.5 MB. Although EfficientNetV2S achieved slightly lower accuracy (81.8%), it still showed a notable improvement compared to its traditional fine-tuning baseline, confirming the robustness of the proposed method across architectures. ResNet50 (82.9%) and VGG16 (82.6%) also achieved substantial gains under the dynamic approach, despite their lower baseline performance.

In addition to overall accuracy, the F1-score is reported to provide a more informative evaluation under the severely imbalanced class distribution of the dataset. As shown in Table 3 and Table 4, models trained with the proposed Dynamic Fine-Tuning strategy consistently achieve higher F1-scores compared to traditional fine-tuning across all evaluated backbones, indicating improved balance between precision and recall. MobileNet achieves the highest F1-score (66.8%) under the dynamic strategy, consistent with its top-performing accuracy. This result indicates that the lightweight architecture, when combined with controlled layer unfreezing, is able to maintain strong discriminative power across classes without overfitting to majority labels. Similarly, Xception, InceptionV3, and DenseNet121 demonstrate consistent F1-score improvements, reflecting more stable precision–recall trade-offs across various models.

In terms of deployment efficiency, the models exhibit substantial differences in loading time and storage size. MobileNet demonstrates the fastest loading time (12.5 s) and the smallest model size (73.05 MB), highlighting its advantage in speed and suitability for resource-constrained environments. Conversely, larger architectures such as EfficientNetV2S (436.61 MB, 78.1 s) and ResNet50 (467.92 MB, 21.59 s) require considerably more computational resources and memory, which may hinder real-time applications. Nevertheless, these increases in complexity do not yield a significant improvement in accuracy, indicating that MobileNet

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

achieves an optimal trade-off between performance and computational efficiency for chest X-ray image classification tasks.

Fig. 3. and Fig. 4. illustrate the comparative performance between the traditional fine-tuning strategy (blue bars) and the proposed Dynamic Fine-Tuning strategy (orange bars) across seven CNN backbones. The results clearly show that the proposed method consistently improves classification accuracy in all architectures. The most significant gains are observed in ResNet50 (from 76.0% to 82.9%), EfficientNetV2S (from 75.5% to 81.8%), and MobileNet (from 81.7% to 85.1%). Smaller yet meaningful improvements are also observed in DenseNet121, Xception, InceptionV3, and VGG16, demonstrating the robustness of the method across models of varying depth and complexity. Overall, the figure highlights that Dynamic Fine-Tuning provides a systematic improvement in accuracy over conventional fine-tuning, with MobileNet achieving the highest performance among all evaluated backbones.

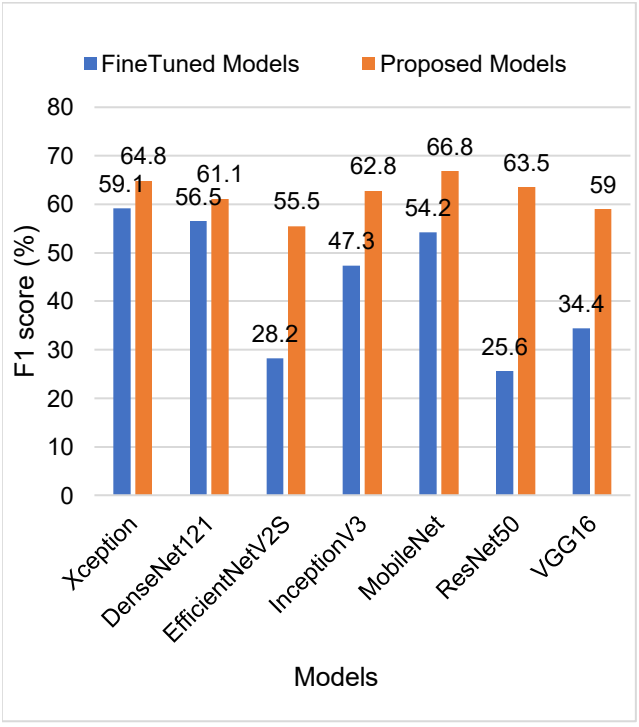

Fig. 4. Comparison of classification F1 scores (%) between traditional fine-tuning models and proposed dynamic fine-tuning models across different CNN backbones.

Fig. 5. shows the normalized confusion matrix for the MobileNet model trained with dynamic fine-tuning. High diagonal values indicate effective discrimination between classes with distinctive radiographic patterns, while misclassifications are primarily confined to diseases with overlapping visual features, such as pleural and pulmonary abnormalities. However, it also indicates that misclassifications for minority classes such as Pulmonary Edema, Tuberculosis, and Pleural Effusion are predominantly assigned to the Pneumonia category. These errors can be explained by overlapping radiographic patterns, including bilateral or perihilar opacities and diffuse haziness shared by Pulmonary Edema and Pneumonia, as well as patchy consolidations in early Tuberculosis that resemble infectious infiltrates. Mild Pleural Effusion may also appear as subtle basal opacities or costophrenic angle blunting, which can be difficult to distinguish from lower-lobe pneumonia on single-view CXRs. Although dynamic fine-tuning improves overall feature adaptation, these subtle inter-class distinctions remain challenging under severe class imbalance and limited samples for rare diseases.

Overall, MobileNet demonstrates satisfactory

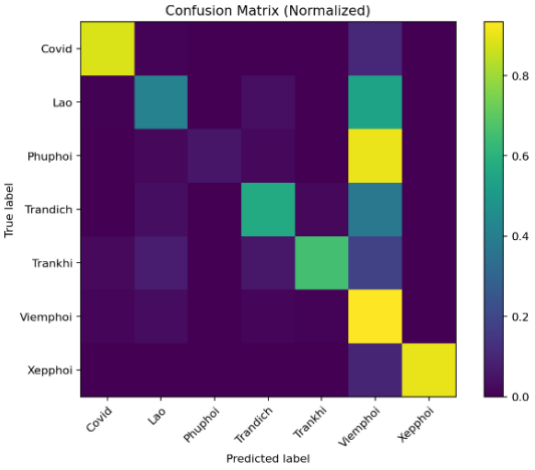

Fig. 5. Confusion matrix (Normalized) of the Dynamic Fine-Tuning for MobileNet model.

performance for common thoracic diseases while maintaining a lightweight structure and fast inference speed, making it suitable for automated diagnostic support systems in medical imaging applications. In summary, the experimental results demonstrate that the Dynamic Fine-Tuning strategy consistently outperforms traditional fixed-layer fine-tuning, yielding higher accuracy while maintaining computational efficiency across diverse CNN backbones.

We conduct a paired sample t-test across seven CNN backbones (Xception, DenseNet121, EfficientNetV2S, InceptionV3, MobileNet, ResNet50, and VGG16). The test yields a mean accuracy improvement of 3.18% with a standard deviation of 1.23%. The resulting t-statistic is t(6) = 3.01, corresponding to a p-value of 0.02, indicating a statistically significant difference at the 5% significance level. Although the sample size (n = 7 backbones) is relatively small, the paired design controls for

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

architectural variability, strengthening the validity of the comparison. This suggests that the proposed adaptive unfreezing mechanism consistently enhances model generalization across different architectures. The comparison results are summarized in Table 3, Table 4 and the statistical test validates the robustness and reliability of the observed improvements.

V. Discussion

The experimental results compiled in Table 5 provide a comprehensive performance evaluation of various state-of-the-art CNN architectures utilizing our proposed validation-driven dynamic fine-tuning strategy. It compares our dynamic fine-tuning strategy with key studies referenced in the Related Works section. Most prior works rely on public benchmarks such as ChestX-ray14 or ChestX-ray8, which involve multi-label annotations and relatively milder class imbalance. These studies often report higher AUC or accuracy values, but only under more balanced distributions, with curated preprocessing, or with ensemble-based architectures. In contrast, our study evaluates performance on a real-world clinical dataset from An Giang Regional General Hospital comprising 15,416 CXR images across seven single-label disease categories with a severely long-tailed distribution. Under such conditions, classification is inherently more challenging, particularly for minority classes.

Recent studies further contextualize our findings. An attention-based DenseNet201 model achieved per-class AUCs of 89–96% across 14 multi-label thoracic diseases [45]. However, their framework relies on public data and increased computational complexity. The authors in [17] achieved 96.4% accuracy using fine-tuned VGG16, though their evaluation likely involved fewer classes and a more balanced dataset. Unlike these architecture- or ensemble-focused approaches, our contribution lies in an optimization-level strategy that is architecture agnostic. The proposed validation-driven dynamic fine-tuning consistently improves performance across heterogeneous backbone architectures without modifying model structure or introducing additional parameters. The observed performance gains can be interpreted from an optimization perspective. In deeper networks such as ResNet50 and EfficientNetV2S, rigid two-stage fine-tuning may either insufficiently adapt mid-to-deep representations or destabilize training when all layers are unfrozen simultaneously [20]. The validation-triggered unfreezing mechanism allows the model to expand its trainable depth only after convergence plateaus, thereby reducing abrupt gradient shifts and improving feature specialization [21]. This controlled adaptation appears particularly beneficial for long-tailed datasets, where minority-class discrimination requires gradual refinement of higher-level representations rather than aggressive global

Table 5. Comparison with related Chest X-ray lung disease classification studies

Study	Dataset/Classes	Method/Fine-tuning	Results (Acc./AUC/F1)	Key Differences vs. Present Work
[12]	ChestXray14 multi-label	CheXNet/DenseNet121 2-stage FT	Pneumonia AUC: 0.888	Multi-label public data; fixed conventional fine-tuning
[13]	ChestXray/multi-label	BB-GCN graph CNN	High multi-label mAP	Graph-based; heavy compute; no Fine-Tuning ablation.
[14]	ChestXray/multi-label	Dual-weighted CNN	Improved multi-label F1	Loss-focused; fixed backbone.
[18]	Custom CXR/4–5 diseases	DenseNet+MobileNet ensemble	ACC ~94%	Ensemble heavy; fewer classes.
[20]	Small medical	Traditional 2-stage Fine-Tuning	Improved small data	Rigid schedule (our baseline).
[21]	ChestXray8/8 diseases	Gradual LR Fine-Tuning	ACC ~90–95%	Fixed epoch schedule.
Our Proposed	An Giang Hosp. 7 classes	**Dynamic Fine-Tuning MobileNet**	**ACC: 85.1%, 66.8% F1**	Real-world severe imbalance; lightweight (73 MB).

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

updates [34], [35]. MobileNet achieved the highest overall accuracy (85.1%) while maintaining minimal deployment overhead (73 MB, 12.5 s loading time), demonstrating an advantageous balance between predictive performance and computational efficiency. Importantly, the performance improvement is not merely attributable to architectural design, as consistent gains are observed across all seven backbones. This indicates that the dynamic strategy enhances the robustness of transfer learning rather than favoring a specific network family. However, computational efficiency remains architecture-dependent, with heavier models such as EfficientNetV2S and ResNet50 incurring higher memory and loading costs despite improved accuracy.

Despite promising results, several limitations should be acknowledged. First, severe class imbalance remains a major challenge, as evidenced by minority-class misclassifications into the dominant pneumonia category. While weighted sampling and loss reweighting mitigate bias to some extent, representation imbalance persists in long-tailed settings. Second, the dataset originates from a single institution, potentially introducing acquisition- or demographic-specific bias that may limit external generalization. Third, the study does not include uncertainty quantification or explainability analyses, which are critical for clinical adoption and risk-aware decision support.

Future work will explore advanced imbalance-aware strategies, including synthetic oversampling (e.g., SMOTE [46], ADASYN [47]) and focal or misclassification-guided losses, integrated with the proposed dynamic fine-tuning framework. Multi-center validation and domain adaptation techniques will also be essential for evaluating generalizability across cross-institutional shifts. More broadly, the validation-driven unfreezing principle introduced in this work may be applicable to other medical imaging tasks characterized by limited labeled data, domain shift, and long-tailed distributions, providing a scalable and architecture-independent optimization paradigm for clinical AI systems.

VI. Conclusion

This study proposes a dynamic fine-tuning strategy for CNN-based lung disease classification from chest X-ray images and demonstrates that validation-guided layer unfreezing improves both accuracy and computational efficiency. By adaptively controlling the depth and timing of backbone adaptation based on validation convergence, the method reduces reliance on manually designed fine-tuning schedules and yields more stable training behavior across architectures. The proposed strategy consistently outperforms traditional fine-tuning, with MobileNet achieving an accuracy of 85.1% while maintaining low resource requirements. Its compatibility with lightweight architectures and controlled model size supports efficient deployment on resource-constrained systems. Future work will address class imbalance through learning-level strategies, conduct multi-center validation, and explore explainable and uncertainty-aware models. Overall, the proposed framework provides a scalable and practical approach for real-world chest X-ray analysis.

Acknowledgment

This research is funded by Vietnam National University HoChiMinh City (VNU-HCM) under grant number C2024-16-02.

Data Availability

The clinical chest X-ray dataset analyzed in this study was obtained from An Giang Regional General Hospital under institutional data-access approval and is not publicly available due to patient privacy and data protection regulations. De-identified data may be made available by the corresponding author upon reasonable request, subject to permission from the hospital administration and applicable ethical and legal requirements.

Author Contribution

Phuoc-Hai Huynh contributed to conceptualization, methodology, algorithm formulation, software development, data collection, model implementation, and drafting the manuscript.
Thi-Diem Truong contributed to supervision, project administration, program execution, and analysis of experimental results. Both authors participated in reviewing, editing, and approving the final manuscript.

Declarations

Ethical Approval

This study used fully anonymized retrospective chest X-ray data provided by An Giang Regional General Hospital. Institutional data-access approval was obtained from the hospital administration for the research use of the dataset. Because the study involved only secondary analysis of de-identified data and no identifiable patient information was accessed, formal ethics committee review and informed consent requirements were waived in accordance with applicable institutional and national regulations.

Consent for Publication Participants.

Because all data were fully anonymized prior to use and no individual-level information is disclosed, informed consent for publication was not required.

Competing Interests

The authors declare no competing interests.

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

References

[1] World Health Organization, 'Fourth WHO consultation on the translation of tuberculosis research into global policy guidelines: meeting report', 2024.

[2] P. Jackson, S. Z. Muyanja, and T. Siddharthan, 'Health Equity and Respiratory Diseases in Low- and Middle-Income Countries', *Clin. Chest Med.*, vol. 44, no. 3, pp. 623–634, Sep. 2023, doi: https://doi.org/10.1016/j.ccm.2023.03.015.

[3] C. Ortiz-Toro, A. García-Pedrero, M. Lillo-Saavedra, and C. Gonzalo-Martín, 'Automatic detection of pneumonia in chest X-ray images using textural features', *Computers in Biology and Medicine*, vol. 145, p. 105466, 2022, doi: 10.1016/j.compbiomed.2022.105466.

[4] 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', 2024, *Springer*. doi: https://doi.org/10.1007/s42979-024-02751-2

[5] G. F. Abbott, 'Challenging Chest Radiograph Interpretation', in *Diseases of the Chest, Heart and Vascular System 2025-2028: Diagnostic Imaging*, Springer, 2025, pp. 169–187. doi: 10.1007/978-3-031-83872-9_14.

[6] I. Pershin, T. Mustafaev, D. Ibragimova, and B. Ibragimov, 'Changes in Radiologists' Gaze Patterns Against Lung X-rays with Different Abnormalities: a Randomized Experiment', *Journal of Digital Imaging*, vol. 36, no. 3, pp. 767–775, Jan. 2023, doi: 10.1007/s10278-022-00760-2.

[7] E. Kotei and R. Thirunavukarasu, 'A Comprehensive Review on Advancement in Deep Learning Techniques for Automatic Detection of Tuberculosis from Chest X-ray Images', Jan. 28, 2024. doi: 10.1007/s11831-023-09987-w.

[8] A. W. Salehi *et al.*, 'A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope', Mar. 29, 2023, *MDPI*. doi: 10.3390/su15075930.

[9] S. Fayaz, S. Z. Ahmad Shah, N. M. ud din, N. Gul, and A. Assad, 'Advancements in Data Augmentation and Transfer Learning: A Comprehensive Survey to Address Data Scarcity Challenges', *Recent Advances in Computer Science and Communications*, vol. 17, no. 8, pp. 14–35, Nov. 2024, doi: 10.2174/0126662558286875231215054324.

[10] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, 'Transfer learning for medical image classification: a literature review', 2022. doi: 10.1186/s12880-022-00793-7.

[11] G. Ayana, K. Dese, A. M. Abagaro, K. C. Jeong, S.-D. Yoon, and S. Choe, 'Multistage transfer learning for medical images', *Artificial Intelligence Review*, vol. 57, no. 9, p. 232, 2024. https://doi.org/10.1007/s10462-024-10855-7

[12] J. C. Y. Seah *et al.*, 'Effect of a comprehensive deep-learning model on the accuracy of chest x-ray interpretation by radiologists: a retrospective, multireader multicase study', *The Lancet Digital Health*, vol. 3, no. 8, pp. e496–e506, 2021, doi: 10.1016/S2589-7500(21)00106-0.

[13] G. Wang, P. Wang, J. Cong, K. Liu, and B. Wei, 'BB-GCN: A Bi-modal Bridged Graph Convolutional Network for Multi-label Chest X-Ray Recognition', 2023, doi: https://doi.org/10.48550/arXiv.2302.11082

[14] Y. Jin, H. Lu, W. Zhu, and W. Huo, 'Deep learning based classification of multi-label chest X-ray images via dual-weighted metric loss', *Computers in Biology and Medicine*, vol. 157, p. 106683, May 2023, doi: 10.1016/j.compbiomed.2023.106683.

[15] M. Blain *et al.*, 'Determination of disease severity in COVID-19 patients using deep learning in chest X-ray images', vol. 27, no. 1, p. 20. doi: 10.5152/dir.2020.20205

[16] S. Sajed, A. Sanati, J. E. Garcia, H. Rostami, A. Keshavarz, and A. Teixeira, 'The effectiveness of deep learning vs. traditional methods for lung disease diagnosis using chest X-ray images: A systematic review', *Applied Soft Computing*, vol. 147, p. 110817, Nov. 2023, doi: 10.1016/j.asoc.2023.110817.

[17] S. Sharma and K. Guleria, 'A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks', in *Procedia Computer Science*, 2022, pp. 357–366. doi: 10.1016/j.procs.2023.01.018.

[18] A. Tekerek and I. A. M. Al-Rawe, 'A Novel Approach for Prediction of Lung Disease Using Chest X-ray Images Based on DenseNet and MobileNet', *Wireless Personal Communications*, pp. 1–15, May 2023, doi: 10.1007/s11277-023-10489-y.

[19] H. Nakrani, E. Q. Shahra, S. Basurra, R. Mohammad, E. Vakaj, and W. A. Jabbar, 'Advanced Diagnosis of Cardiac and Respiratory Diseases from Chest X-Ray Imagery Using Deep Learning Ensembles', *Journal of Sensor and Actuator Networks*, vol. 14, no. 2, p. 44, 2025, doi: 10.3390/jsan14020044.

[20] N. Spolaôr *et al.*, 'Fine-tuning pre-trained neural networks for medical image classification in small clinical datasets', *Multimedia Tools and Applications*, vol. 83, no. 9, pp. 27305–27329, 2024, doi: 10.1007/s11042-023-16529-w.

[21] C. Bu *et al.*, 'Layer-Wise Learning Rate Optimization for Task-Dependent Fine-Tuning of

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

Pre-Trained Models: An Evolutionary Approach', vol. 4, no. 4, pp. 1–23. https://doi.org/10.1145/3689827

[22] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, 'Transfer learning for medical image classification: a literature review', 2022. doi: 10.1186/s12880-022-00793-7.

[23] M. Byra, 'Breast mass classification with transfer learning based on scaling of deep representations', *Biomed. Signal Process. Control*, vol. 69, p. 102828, 2021. https://doi.org/10.1016/j.bspc.2021.102828

[24] M. Abadi *et al.*, 'TensorFlow: A system for large-scale machine learning', May 2016, [Online]. https://doi.org/10.48550/arXiv.1605.08695

[25] N. Spolaôr *et al.*, 'Fine-tuning pre-trained neural networks for medical image classification in small clinical datasets', *Multimed. Tools Appl.*, vol. 83, no. 9, pp. 27305–27329, 2024, doi: 10.1007/s11042-023-16529-w.

[26] C. Bu *et al.*, 'Layer-Wise Learning Rate Optimization for Task-Dependent Fine-Tuning of Pre-Trained Models: An Evolutionary Approach', vol. 4, no. 4, pp. 1–23. https://doi.org/10.1145/3689827

[27] L. Prechelt, 'Early Stopping - But When?', in *Proceedings of Machine Learning Research*, 1998, pp. 55–69. doi: 10.1007/3-540-49430-8_3.

[28] L. N. Smith, 'Cyclical Learning Rates for Training Neural Networks', in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, Mar. 2017, pp. 464–472. doi: 10.1109/WACV.2017.58.

[29] K. Alshamrani and H. A. Alshamrani, 'An Efficient Dual-Sampling Approach for Chest CT Diagnosis', pp. 239–253. Doi: https://doi.org/10.2147/jmdh.s472170

[30] E. Chamseddine, N. Mansouri, M. Soui, and M. Abed, 'Handling class imbalance in COVID-19 chest X-ray images classification: Using SMOTE and weighted loss', vol. 129, p. 109588. doi: https://doi.org/10.1016/j.asoc.2022.109588

[31] M. Byra, 'Breast mass classification with transfer learning based on scaling of deep representations', *Biomedical Signal Processing and Control*, vol. 69, p. 102828, 2021. doi: https://doi.org/10.1016/j.bspc.2021.102828

[32] T. Geroski, O. Pavić, L. Dašić, D. Milovanović, M. Petrović, and N. Filipović, 'SoftLungX: leveraging transfer learning with convolutional neural networks for accurate respiratory disease classification in chest X-ray images', *Journal of Big Data*, vol. 11, no. 1, p. 146, 2024. doi: https://doi.org/10.1186/s40537-024-01018-0

[33] S. Roy, M. Tyagi, V. Bansal, and V. Jain, 'Svd-clahe boosting and balanced loss function for covid-19 detection from an imbalanced chest x-ray dataset', *Computers in Biology and Medicine*, vol. 150, p. 106092, 2022. doi: https://doi.org/10.1016/j.compbiomed.2022.106092

[34] Y.-X. Wu, K. Du, X.-J. Wang, and F. Min, 'Misclassification-guided loss under the weighted cross-entropy loss framework', *Knowledge and Information Systems*, vol. 66, no. 8, pp. 4685–4720, 2024.

[35] D. R. Beddiar, M. Oussalah, U. Muhammad, and T. Seppänen, 'A deep learning based data augmentation method to improve COVID-19 detection from medical imaging', vol. 280, p. 110985. doi: https://doi.org/10.1016/j.knosys.2023.110985

[36] D. A. Dablain, C. Bellinger, B. Krawczyk, and N. V Chawla, 'Efficient augmentation for imbalanced deep learning', in *2023 IEEE 39th International Conference on Data Engineering (ICDE)*, 2023, pp. 1433–1446. doi: https://doi.ieeecomputersociety.org/10.1109/ICDE55515.2023.00114

[37] F. Chollet, 'Xception: Deep learning with depthwise separable convolutions', in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258. doi: https://doi.org/10.1109/cvpr.2017.195

[38] M. Tan and Q. V. Le, 'EfficientNetV2: Smaller Models and Faster Training', in *Proceedings of Machine Learning Research*, PMLR, 2021, pp. 10096–10106. doi: https://doi.org/10.48550/arXiv.2104.00298

[39] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, 'Rethinking the Inception Architecture for Computer Vision', in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826. doi: 10.1109/CVPR.2016.308. https://doi.org/10.48550/arXiv.1512.00567

[40] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, 'MobileNetV2: Inverted Residuals and Linear Bottlenecks', in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 4510–4520. doi: 10.1109/CVPR.2018.00474.

[41] K. He, X. Zhang, S. Ren, and J. Sun, 'Deep residual learning for image recognition', in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[42] K. Simonyan and A. Zisserman, 'Very deep convolutional networks for large-scale image recognition', in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015. doi: https://doi.org/10.48550/arXiv.1409.1556

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352

[43] M. Abadi *et al.*, 'TensorFlow: A system for large-scale machine learning', May 2016. https://doi.org/10.48550/arXiv.1605.08695
[44] M. Watson *et al.*, 'KerasCV and KerasNLP: Multi-framework Models', *Journal of Machine Learning Research*, vol. 25, no. 375, pp. 1–10, 2024. doi: 10.1007/s42979-024-02763-0
[45] U. Chutia, A. S. Tewari, J. P. Singh, and V. K. Raj, 'Classification of Lung Diseases Using an Attention-Based Modified DenseNet Model', *Journal of Imaging Informatics in Medicine*, vol. 37, no. 4, pp. 1625–1641, Mar. 2024, doi: 10.1007/s10278-024-01005-0.
[46] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, 'SMOTE: Synthetic Minority Over-sampling Technique', Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, Jun. 2002, doi: 10.1613/jair.953.
[47] Haibo He, Yang Bai, E. A. Garcia, and Shutao Li, 'ADASYN: Adaptive synthetic sampling approach for imbalanced learning', in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, IEEE, Jun. 2008, pp. 1322–1328. doi: 10.1109/IJCNN.2008.4633969.

Author Biography

Phuoc-Hai Huynh received his B.S. in Information Technology from An Giang University (2007), and his M.S. and Ph.D. in Information Systems from Can Tho University (2014, 2021). He is currently the Head of the Software Engineering Department at the Faculty of Information Technology. His research focuses on data mining and bioinformatics, specifically addressing challenges in complex datasets (high-dimensional, small-sample, and imbalanced). He is also dedicated to applying machine learning techniques in medical and healthcare domains to improve diagnostics and personalized healthcare solutions.

Thi-Diem Truong received her Engineer's degree in Informatics from Can Tho University, Vietnam, in 2003. In 2010, she obtained her Master's degree in Information Systems from Can Tho University, Vietnam. She is currently a lecturer at the Faculty of Information Technology, An Giang University. Her research interests include data mining and medical informatics, with particular emphasis on complex data analytics, including high-dimensional, small-sample, and imbalanced datasets. Her work focuses on the development of robust and interpretable machine learning frameworks for healthcare applications. She is especially interested in leveraging artificial intelligence techniques to enhance diagnostic accuracy, improve predictive performance, and advance personalized healthcare systems.

Manuscript received 19 November 2025; Revised 10 January 2026; Accepted 25 February 2026; Available online 2 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1352