IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 3, June 2026, pp. 2567~2580
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i3.pp2567-2580

Diagnosing tuberculosis from X-ray imaging based on contrast limited adaptive histogram equalization

Nguyen Trong Vinh[1], Lam Thanh Hien[1], Ha Manh Toan[2], Do Nang Toan[2]
[1]Faculty of Information Technology, Lac Hong University, Bien Hoa, Vietnam
[2]Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam

Article Info

Article history:

Received May 22, 2025
Revised Feb 11, 2026
Accepted Mar 5, 2026

Keywords:

Chest X-ray
Deep learning
Image enhancement
Neural network
Tuberculosis diagnosis

ABSTRACT

Tuberculosis is a serious threat, and one of the effective data types for diagnosing tuberculosis is chest X-ray data. In this paper, we hypothesize the effect of image enhancement on the effectiveness of deep learning models in the problem of diagnosing pulmonary tuberculosis from chest X-ray images. To clarify the hypothesis, we have designed a data processing process with an image enhancement step using the contrast limited adaptive histogram equalization (CLAHE) technique to enhance the quality of input chest X-ray data, and the experiments were conducted with a standard dataset that was published on the Kaggle system. The evaluation is performed comprehensively with popular convolutional neural network architectures, including DenseNet201, DenseNet121, EfficientNetB0, and MobileNetV2, compared in two scenarios with and without the image enhancement step. Experiments have shown that the image enhancement step effectively improves the classification performance of all models, clearly through important scores such as area under curve (AUC), accuracy, F1-score, precision, and recall. The best result tested is the EfficientNetB0 model with 0.925926 accuracy score, 0.970732 AUC score, 0.904762 precision score, 0.95 recall score, and 0.926829 F1-score. In addition, qualitative analysis using gradient-weighted class activation mapping (Grad-CAM) shows that the resulting models have shown a focus on the lung region, reflecting the interpretability and suitability for radiologist expertise.

Corresponding Author:

Do Nang Toan
Institute of Information Technology, Vietnam Academy of Science and Technology
18 Hoang Quoc Viet, Cau Giay, Hanoi10072, Vietnam
Email: dntoan@ioit.ac.vn

1. INTRODUCTION

One of the most serious threats to human health is tuberculosis, caused by the bacterium mycobacterium tuberculosis. Tuberculosis remains one of the leading causes of death from infectious diseases, especially in developing countries. More specifically, reports on tuberculosis show that there are more than 4,000 deaths and about 30,000 new cases every day, based on data released on World Tuberculosis Day 2022 [1]. Treatment of tuberculosis depends largely on when the disease is detected. More specifically, many different tuberculosis treatment regimens have been proposed [2], such as the six-month or nine-month rifampin, isoniazid, pyrazinamide, and ethambutol (RIPE) regimen or the four-month rifapentine-moxifloxacin regimen. Early detection and timely treatment are important in preventing the spread of tuberculosis and minimizing its public health consequences.

In the medical practice of diagnosing pulmonary tuberculosis, chest X-ray images are a common and effective type of data for diagnosis. Chest X-ray data is low-cost and has the ability to reflect abnormal

characteristics on the lungs, so it is widely used for screening pulmonary tuberculosis. However, the interpretation and analysis of X-ray images require high expertise. This requires highly trained and experienced radiologists, which is always a scarce group in the medical team. Therefore, there is a need to apply computer vision and artificial intelligence technologies to build effective automated pulmonary tuberculosis diagnosis systems. This is also the motivation for applying many traditional methods, such as research [3] using traditional neural network combined with a genetic algorithm, research [4] using an identification tree in tuberculosis diagnosis problems. In 2017, Lakhani and Sundaram [5] used convolutional neural network models such as AlexNet and GoogleNet to classify tuberculosis on chest X-ray, achieving high accuracy. The authors used four de-identified datasets, and achieved the best area under curve (AUC) score with 0.99 also in 2017, Rajpurkar *et al*. [6] presented the CheXNet model. CheXNet is a DenseNet121 model for detecting pneumonia from chest X-ray data learned on the ChestX-ray14, the dataset containing over 100,000 images. The authors noted that the proposed model outperformed human radiologists in detecting pneumonia from chest radiographs with the F1 metric. Next, in 2017, Hooda *et al*. [7] published a study proposing a deep learning-based tuberculosis detection system. In the study, the authors used a neural network consisting of seven convolutional layers and three fully connected layers. Experiments showed that the best case achieved 0.8209 for validation accuracy and 0.9473 for overall accuracy. In 2018, Rajpurkar *et al.* [8] compared the effectiveness of the CheXNet model with human radiologists in diagnosing a variety of pathologies on chest radiographs. Radiologists have an average experience of twelve years, and three senior radiology residents. They came from three different academic institutions. In 2020, Chandra *et al.* [9] applied a hierarchical feature extraction method to detect pulmonary tuberculosis abnormalities in chest X-ray images, through which the authors improved the accuracy of the diagnostic system. In the experiment, the authors achieved an accuracy of approximately 0.956 and an AUC of approximately 0.95 on the Montgomery dataset, and similarly 0.994 and 0.99 on the Shenzhen dataset. In 2019, Pasa *et al*. [10] proposed a simple network to diagnose tuberculosis faster. In the experiment, the authors also discussed that their architecture can preserve the accuracy score versus previous studies. In that year, Meraj *et al.* [11] also presented research using four different convolutional neural networks with Montgomery and Shenzhen datasets to diagnose tuberculosis effectively. In 2023, Toan *et al* [12] published their work designing several training strategies with several models. Their study aimed to evaluate and discuss tuberculosis diagnosis in Vietnamese data, as well as the impact of Vietnamese data with a transfer learning approach.

In 2022, Wong *et al.* [13] introduced a self-attention network specifically designed for detecting pulmonary tuberculosis. The authors also conducted an explainability evaluation. The study used chest X-ray data from multiple countries and the experimental results showed that the model achieved an accuracy of 99.86%. In 2024, Vinh *et al.* [14] focused on the problem of automatic diagnosis of pulmonary tuberculosis in the context of Vietnam. The authors built a custom convolutional neural network combined with a heat-map predictive visualization mechanism. The study used a large X-ray dataset of Vietnamese patients and the experimental results recorded that their model achieved both sensitivity score and specificity score on the test set above 0.9, and the highlighted regions in the heat-map were similar to the symptoms annotated by doctors. In 2024, Patel *et al*. [15] studied the problem of screening for pulmonary tuberculosis from X-ray images by proposing a self-supervised learning network with the ability to explain, to reduce the cost and difficulty in data collection. The results indicated that their system achieved 98.14% accuracy score, thereby strengthening the potential for practical application in community tuberculosis screening. In 2024, Bista *et al.* [16] focused on the problem of detecting pulmonary tuberculosis from X-ray images using the YOLOv7 architecture combined with a convolutional neural network. The study was conducted on the TBX11K dataset, and the results showed that the system achieved map of 0.587. In 2024, Kotei and Thirunavukarasu [17] studied the problem of automatic diagnosis of tuberculosis by proposing a combined model. The study uses the TBX11K dataset and the experimental results show that the model achieves 99.38% accuracy. We also include a comparative table of tuberculosis detection methods that consists of both traditional image processing and CNN-based models. The details are in Table 1.

In general, detecting pulmonary tuberculosis lesions on X-ray images is challenging. On the one hand, this comes from the image features that are often not really clear, sometimes with a mild level of damage, and can be mixed with anatomical structures such as ribs, heart, or blood vessels. On the other hand, the quality of X-ray images in practice can be affected by different shooting conditions due to differences in equipment as well as the way the technician controls it. This leads to certain variations related to image quality, such as the level of contrast between damaged areas and healthy lung tissue. All of the above problems lead to certain difficulties for deep learning models when recognizing abnormalities. Accordingly, a feasible solution to this problem is to improve image quality based on local contrast enhancement, thereby clarifying the details in the image. With that goal, the research team focused on applying the contrast limited adaptive histogram equalization (CLAHE) technique, which helps increase the discrimination in low-contrast X-ray regions while trying not to cause noise amplification in high-contrast regions. This is expected to help

improve the classification performance in the problem of diagnosing pulmonary tuberculosis on chest X-rays. In this research context, CLAHE-based preprocessing helps improve the performance of convolutional neural network models when analyzing chest X-ray images. CLAHE helps to clarify the structural details in the lung region that may be obscured by noise or uneven brightness. As a result, pathological features such as tuberculosis lesions, nodules, or lung tissue abnormalities become more prominent and are easily recognized by the convolutional neural network model. Compared with preprocessing based on global histogram equalization, CLAHE balances and retains more important local features. At the same time, unlike edge-enhancing preprocessing that emphasizes the boundary, CLAHE provides more detailed information on the entire lung texture to help convolutional neural network models.

Our study is set in the context of developing an effective deep learning solution for diagnosing pulmonary tuberculosis from chest X-ray data. In the study, we will evaluate the role of image enhancement, and specifically CLAHE, in boosting the diagnostic quality of deep learning models. To conduct a comprehensive evaluation, we will compare the performance of four popular deep learning architectures, including DenseNet201, DenseNet121, EfficientNetB0, and MobileNetV2, on the task of classifying pulmonary tuberculosis from chest X-rays, in two scenarios: with and without image enhancement. The main contributions of our research include: i) proposing a simple and efficient processing procedure that can be easily integrated into tuberculosis diagnosis support systems using chest X-ray images; ii) clarifying the hypothesis of image quality enhancement through evaluating the classification performance on four different popular deep learning models, namely DenseNet201, DenseNet121, EfficientNetB0, and MobileNetV2; and iii) discussing the qualitative evaluation through gradient-weighted class activation mapping (Grad-CAM) of the concentration regions of deep learning models when applying the proposed method.

Table 1. Some tuberculosis diagnosis studies

Authors	Year	Model
Elveren and Yumuşak [3]	2011	Traditional neural network and genetic algorithm
Dongardive *et al.* [4]	2011	Identification tree
Lakhani and Sundaram [5]	2017	AlexNet and GoogleNet
Rajpurkar *et al*. [6]	2017	CheXNet
Hooda *et al*. [7]	2017	A specific CNN
Rajpurkar *et al*. [8]	2018	CheXNeXt
Chandra *et al.* [9]	2020	Hierarchical feature extraction scheme
Pasa *et al.* [10]	2019	A simple CNN
Meraj *et al*. [11]	2019	GoogleNet, ResNet50, VGG-16, and VGG-19.
Toan *et al*. [12]	2023	ResNet50, ResNet34, AlexNet, and DenseNet121
Wong *et al*. [13]	2022	TBNet
Vinh *et al*. [14]	2024	A custom CNN
Patel *et al*. [15]	2024	A self-supervised learning network
Bista *et al*. [16]	2024	YOLOv7
Kotei *et al*. [17]	2024	Data efficient image transformer and ResNet16

2. METHOD

2.1. Dataset

In this research, we used the Pulmonary chest X-ray abnormalities dataset [18] publicly available on the Kaggle platform. The dataset was built based on data from Montgomery County and Shenzhen Hospital. This dataset was built as one of the efforts to serve the research on classifying tuberculosis from chest X-ray data. The image dataset consists of 800 images, of which 662 images are from the Shenzhen set and 138 images are from the Montgomery set. Besides, the label distribution in the Shenzhen set includes 326 normal images and 336 tuberculosis images. The label distribution in the Montgomery set includes 80 normal images and 58 tuberculosis images. The label distribution is illustrated in Figure 1.

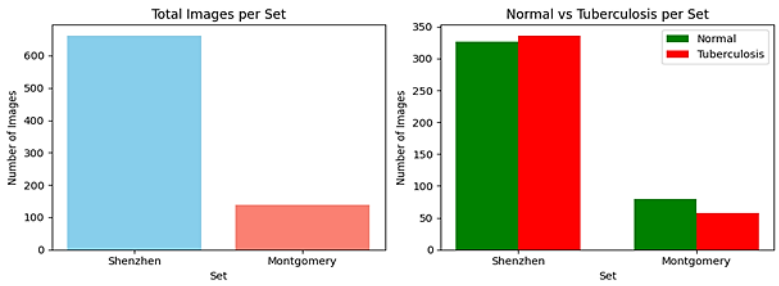

Figure 1. The label distribution in datasets

2.2. Models

Next, EfficientNet [19] was proposed as a family of convolutional neural networks with notable improvements. DenseNet is a well-known deep neural network introduced by Huang *et al*. [20] in 2017. In the work, the authors presented a design in which each layer is directly linked to previous layers. Thus, the input of each layer will include not only the output of its direct previous layer, but also the outputs from other previous layers, as described in Figure 2. This is a design that aims to reuse features and handle gradient vanishing. In this study, we use two versions of DenseNet, including DenseNet121, and DenseNet201.

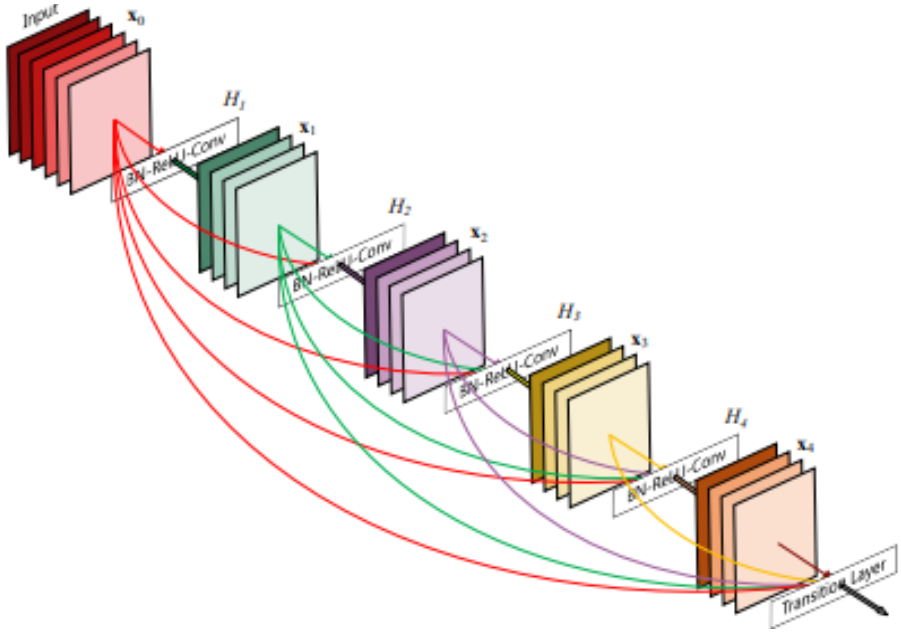

Figure 2. A 5-layer dense block [20]

In this study, we use the EfficientNetB0 version, which is a basic version and does not require too many resources. Another case we used is MobileNet [21]. This is a lightweight CNN model family designed with mobile devices in mind. The design uses the depth-wise separable convolutions mechanism as described in Figure 3, which allows for reduced computations and parameter counts while maintaining a certain level of performance. The version we tested in this study is MobileNetV2.

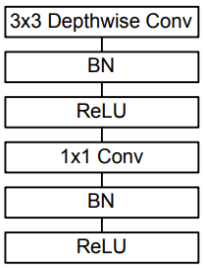

Figure 3. Depth-wise separable convolutions [21]

2.3. The data processing procedure

Our data preprocessing pipeline is described in Figure 4. First, the image is read as a grayscale image. The image data is then normalized to the pixel values in the range [0, 1] by dividing the pixel values by 255. Next, the image is resized according to a predetermined parameter. Finally, the image is reformatted to be compatible with the four-dimensional tensor input [1, 1, height, width] of the deep learning library. Usually, in the preprocessing process, some issues can be considered such as lung segmentation or noise

reduction. In this study, data normalization has supported noise reduction, which helps to improve image quality and support the model to learn more effectively.

After loading the image, the next step is to perform contrast enhancement. The technique used in this step is CLAHE [22]. This enhances the local contrast in images, thereby making the details in chest X-ray data, for example in the lung regions, clearer. This transformation is performed directly on the graphics processing unit (GPU) with the help of functions in the Kornia library [23].

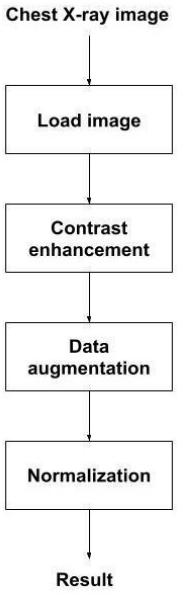

Figure 4. The data preprocessing pipeline

Next, the image will be passed through a data augmentation transformation step to enhance the diversity of the input image for the training phase of the deep learning model. Specifically, the image will be applied with random transformations including horizontal flipping, random left-right rotation, random slight translation in both horizontal and vertical directions, and random brightness change. These transformations are performed non-uniformly for each image in each batch, and the entire pipeline is executed on the GPU to increase processing speed. Finally, pixel value is normalized using mean and standard deviation calculated from ImageNet. This is a common normalization step when feeding images into neural networks. The content is also shown in pseudocode in Algorithm 1. Using these transformations is expected to help the model learn more general features and improve its ability to resist overfitting. In the experiment, in order to demonstrate the hypothesis of the role of image enhancement, the chest X-ray data processing algorithm was designed to enable or disable contrast enhancement with CLAHE. Thus, in the experiments, we will perform two options of data processing: with or without using the contrast enhancement step.

Algorithm 1. The data processing procedure

```
Input: image file path file_path
Output: image tensor img_tensor
Process:
1: gray_image = load_gray_image(file_path)
2: img_tensor = convert_tensor(gray_image)
3: normalize_0_1(img_tensor)
4: resize(img_tensor)
5: enhance_contrast(img_tensor)
6: augment_transform(img_tensor)
7: resize(img_tensor)
8: normalize_mean_stddev(img_tensor)
```

2.4. Experiment configuration

In this study, CLAHE was set with clipLimit as 2 and tileGridSize as 8×8. This ensures that the cells are small enough to highlight the lung texture details and abnormal areas, but still maintain the continuity of the lung tissue without tearing it too much. The clipLimit value of 2 helps to avoid over-enhancing the

contrast, which can cause an imbalance between light and dark areas and make the image look artificial and unnatural. Therefore, setting these parameters contributes to supporting the convolutional neural network model to learn and recognize more accurately.

In this research, preprocessed chest X-ray data is fed into models to perform pulmonary tuberculosis diagnosis classification. As shown in Figure 5, the process includes two options A and B. Option A uses the original image without contrast enhancement and option B uses the image that has been contrast enhanced using the CLAHE method. It allows us to assess the effectiveness and validate our research hypothesis on the effect of contrast enhancement using the CLAHE method on the pulmonary tuberculosis recognition performance of models.

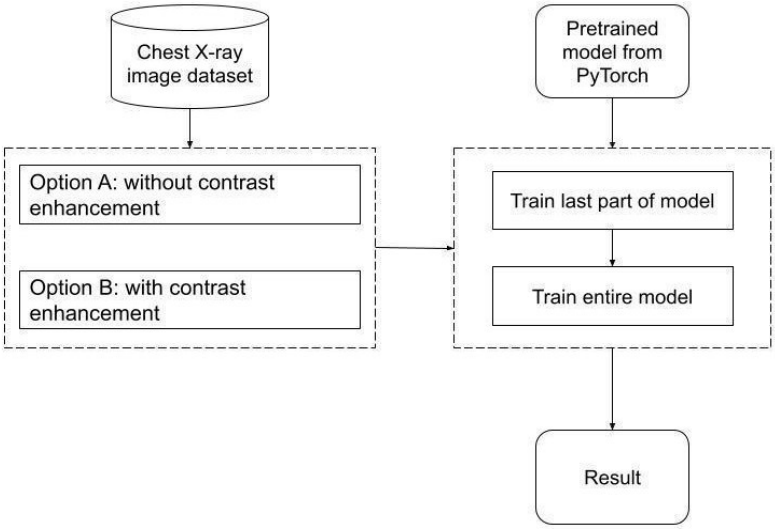

Figure 5. The training workflow

In the experiment, the selected convolutional neural network models include DenseNet121, DenseNet201, EfficientNetB0, and MobileNetV2. These are all popular architectures in image recognition problems and are suitable for medical applications. The experimental results are expected to show that integrating contrast enhancement using the CLAHE method can improve the pulmonary tuberculosis recognition performance of deep learning models by highlighting the texture regions in the lungs. In addition, the validation on many different convolutional neural network architectures helps confirm the scalability and stability of our proposed method in the paper, creating a basis for practical application in pulmonary tuberculosis diagnosis support systems from chest X-ray image data.

During the training of the classification model, we perform two steps for transfer learning. We use a pre-trained model with ImageNet provided in the PyTorch library [24]. We replace the last classification layer to suit our problem. In the first two epochs, only the last part of the model is trained, while the remaining layers retain their initial weights. From the third epoch, the entire model is trained, allowing all model parameters to be updated every epoch.

In terms of optimization, architectures are trained using the cross-entropy loss function combined with the Adam algorithm [25]. The learning rate is adjusted using the one cycle mechanism [26]. This setting is expected to help the learning rate change reasonably during the training process, avoid falling into local extremes, and help the model converge faster. In addition, training is also supported by the gradient accumulation technique, which helps the model train more efficiently under limited GPU memory conditions.

In this study, to evaluate the training results of the deep learning model for classifying tuberculosis from chest X-ray images, we used the model performance evaluation indexes, including AUC, accuracy, F1-score, precision, and recall. In addition, to better understand how the model makes predictions, the Grad-CAM technique [27] helps generate heatmaps from the X-ray images. Grad-CAM will calculate the gradient of the last classification layer on the spatial feature aspect to determine the probability map to assess the model's concentration on the image. In this study, Grad-CAM heatmaps can help us check whether the model really pays attention to the lung regions or lesions that are characteristic of tuberculosis, thereby increasing reliability and aiming to support clinical validation.

3. RESULTS AND DISCUSSION

This section illustrates the experimental results of four models, including DenseNet121, DenseNet201, MobileNetV2, and EfficientNetB0, trained with two different training scenarios, with or without contrast enhancement. We call the option without contrast enhancement option A and the option with contrast enhancement option B. Thus, option B is an improvement over option A.

3.1. Choosing parameters of contrast enhancement

To support the robustness of parameter choice, we conduct some preliminary experiments to evaluate some parameter choices. Specifically, with reference to the study [28], some parameter choices are conducted including (clipLimit =2, tileGridSize = 4×4), (clipLimit =2, tileGridSize =8×8), (clipLimit = 4, tileGridSize =16×16), and (clipLimit =10, tileGridSize =32×32). For each parameter choice, we will conduct experiments with four models, including DenseNet121, DenseNet201, MobileNetV2, and EfficientNetB0. For each parameter selection case, we average the scores for accuracy, AUC, F1-score, precision, and recall across all four models. Details are shown in Table 2.

Table 2. The average performance of parameter choices

(clipLimit, tileGridSize)	Accuracy	AUC	Precision	Recall	F1-score
(2, 4×4)	0.882716	0.96189	0.896469	0.875	0.881505
(2, 8×8)	0.91358	0.976067	0.925217	0.9	0.910923
(4, 16×16)	0.876543	0.934756	0.880295	0.86875	0.874169
(10, 32×32)	0.876543	0.947561	0.871951	0.88125	0.876005

From Table 2, we can see that the average scores of the case of choosing parameters clipLimit as 2 and tileGridSize as 8×8 have a certain superiority. Specifically, the average AUC value of the case of choosing parameters clipLimit as 2 and tileGridSize as 8×8 reaches the highest position with a value of 0.976067, the average AUC values of the other cases are all slightly lower with 0.96189 of the case clipLimit =2 and tileGridSize =4×4, 0.934756 of the case clipLimit = 4 and tileGridSize =16×16, 0.947561 of the case clipLimit =10 and tileGridSize =32×32. Considering another metric case, the average precision value of the case with parameter selection clipLimit =2 and tileGridSize =8×8 reached the highest position with the value of 0.925217, the average precision values of the other selection cases were all slightly lower with 0.896469 of the case with clipLimit =2 and tileGridSize =4×4, 0.880295 of the case with clipLimit =4 and tileGridSize =16×16, 0.871951 of the case with clipLimit =10 and tileGridSize =32×32. Thus, based on the preliminary testing to evaluate some parameter choices, the choice with clipLimit =2 and tileGridSize =8×8 achieved the highest test score. This is also consistent with the parameter selection considerations presented above.

3.2. Quantitative comparison between training with and without contrast enhancement

Figure 6 clearly demonstrates that the training strategy using option B brings consistent and substantial improvements through all key performance metrics compared to the training strategy using option B. Specially, the average accuracy increases from 0.885802 to 0.913580, while the average AUC improves from 0.950762 to 0.976067, pointing out a comprehensively enhanced capability to classify between normal and tuberculosis data. Furthermore, the average precision increases from 0.877250 to 0.925217, and the average recall improves from 0.89375 to 0.9. This describes that the training strategy using option B not only reduces false positives but also minimizes false negatives. This is important because missing a real case of tuberculosis can cause serious public health consequences. The rise in F1-score from 0.885332 to 0.910923 helps us confirm that models trained under the training strategy using option B can achieve a more balanced performance, not requiring a trade-off between sensitivity and specificity. It also reflects our hypothesis for our proposed study. We continue to see the average performance metrics regarding true positive, false negative, true negative, and false positive between models trained using options A and B in Figure 7.

In Figure 7, the average performance metrics from models trained using option B show slightly better performance than option A across all scores, including true positive, false negative, true negative, and false positive. Therefore, it clearly demonstrates that option B brings more accuracy in distinguishing between tuberculosis and normal samples. Specifically, the average score of true positive rate increased slightly from 44.1% to 44.4%, and the average score of false negative rate decreased from 5.2% to 4.9%. These scores indicate that the models in option B were better at detecting real tuberculosis cases. On the other hand, the average score of false positive rate decreased from 6.2% to 3.7%, while the average score of true negative rate increased from 44.4% to 46.9%. These scores indicate that the models in option B were better at distinguishing between tuberculosis and normal images. Although the improvements are slight, using option B is generally preferred. This is because it shows comprehensive improvement, reducing both false alarms and missed detections.

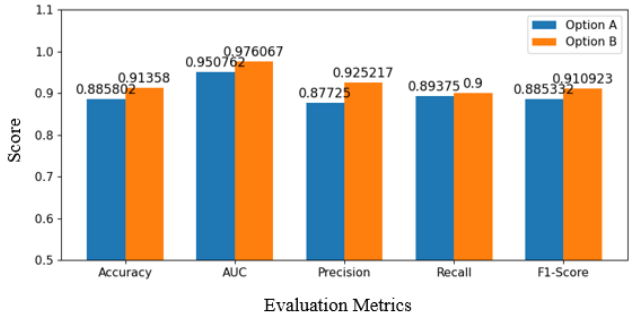

Figure 6. Average performance metrics between models trained using options A and B

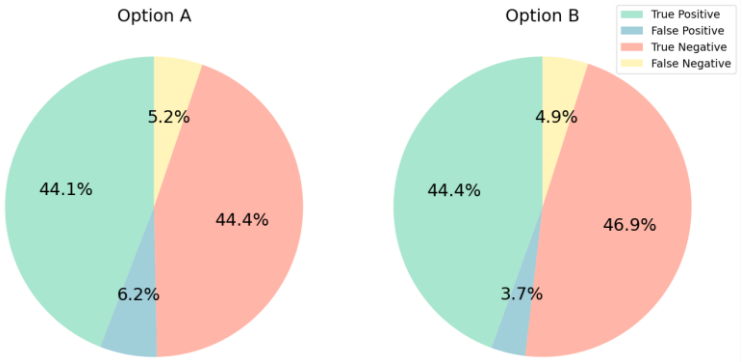

Figure 7. Average scores of true positive, false negative, true negative, and false positive between models trained using options A and B

3.3. Analysis Grad-CAM heatmaps of the training strategy using option B

We consider the Grad-CAM heatmaps of some correctly recognized cases by deep learning models trained with the strategy using option B in Figure 8. In which, the first line belongs to the EfficientNetB0 model, the second line belongs to the DenseNet201 model, the third line belongs to the DenseNet121 model, and the fourth line is the results of the MobileNetV2 model. The analysis of the obtained Grad-CAM shows that deep learning models diagnosing these correctly recognized cases often focus on the lung area, that is, the anatomical area directly related to pulmonary tuberculosis. This reflects that the deep learning models have learned image features that are related to the semantics of pulmonary tuberculosis. This contributes to strengthening the explainability and clinical plausibility of the model and supporting experts in evaluating the model results.

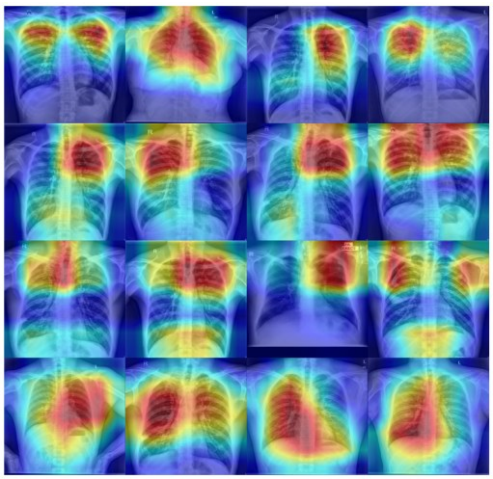

Figure 8. Grad-CAM visualization from models trained using option B

3.4. Quantitative comparison among models

Next, we would see the quantitative diagnostic performance metrics. All scores were assessed on the Pulmonary chest X-ray abnormalities dataset, a labeled tuberculosis dataset built based on data from the Shenzhen dataset and the Montgomery dataset. Table 3 shows the average performance of the models in the experiment, with the average accuracy score, the average AUC score, the average F1-score, the average precision score, and the average recall score.

Table 3 shows that EfficientNetB0 has the best overall performance. EfficientNetB0 reaches some highest scores. The average accuracy reaches 0.913580 with 0.901235 in the case without CLAHE-based contrast preprocessing and 0.925926 in the case with CLAHE-based contrast preprocessing, so using CLAHE outperforms 0.024691. The average F1-score reaches 0.913415 with 0.9 in the case without CLAHE-based contrast preprocessing and 0.926829 in the case with CLAHE-based contrast preprocessing, so using CLAHE outperforms 0.026829. The average recall reaches 0.925 with 0.9 in the case without CLAHE-based contrast preprocessing and 0.95 in the case with CLAHE-based contrast preprocessing, so using CLAHE outperforms 0.05. They indicate a higher ability to correctly diagnose tuberculosis cases than the other architectures in this study. MobileNetV2 stands out with the highest average AUC of 0.973171 and the highest average precision of 0.92577, indicating that this model is particularly strong in reducing false positives. The AUC increases from 0.964024 to 0.982317, and the precision increases from 0.880952 to 0.970588 with CLAHE-based contrast preprocessing. Meanwhile, DenseNet201 has the lowest performance in this study, indicating a lower classification ability than the other models, and this may be from the restricted dataset. Overall, EfficientNetB0 is the most prominent model among our experiments.

Table 3. The average performance of the models

Model	Accuracy	AUC	Precision	Recall	F1-score
EfficientNetB0	0.91358	0.967378	0.902381	0.925	0.913415
DenseNet201	0.882716	0.96189	0.886538	0.875	0.880696
DenseNet121	0.901235	0.95122	0.890244	0.9125	0.901235
MobileNetV2	0.901235	0.973171	0.92577	0.875	0.897165

Essentially, the models all achieved fairly good average scores. It can be partly explained by how specific architectural features relate to tuberculosis image characteristics. EfficientNet, with its compound scaling strategy, can help exploit subtle features such as lesion texture in the lung region. Besides, DenseNet utilizes feature reuse mechanisms through dense connections, which may be good for identifying the lesions' characteristics and contrast distribution.

Although all the tested convolutional neural network models performed well in detecting pulmonary tuberculosis on chest X-ray data, EfficientNetB0 outperformed some other models in the research experiments. This difference can be partly explained by the EfficientNet architecture, in which the simultaneous balance between depth, width, and resolution helps the model learn better. On the other hand, deeper models such as DenseNet201 may not be optimal because of the large number of parameters, and here the test dataset is not large in size. In addition, the dataset used was generated from two data sources, Montgomery County and Shenzhen Hospital. In theory, this also makes data preprocessing useful in improving the different image quality between the datasets.

We take a closer look at the EfficientNetB0 model with the training strategy using option B as shown in Figure 9, with the confusion matrix on the left side (Figure 9(a)) and the ROC curve on the right side (Figure 9(b)). The confusion matrix points out the detailed values with the true positive of 38, the true negatives of 37, the false positives of 4, and the false negatives of 2. From there, many metrics such as accuracy, F1-score, precision, and recall can be calculated. Basically, the model gives high and balanced performance. On the one hand, the overall classification accuracy is quite good with a value of 0.925926. On the other hand, the F1-score of 0.926829 shows the balance in the diagnostic ability of the model. More specifically, the recall score of 0.95 shows a low rate of missing cases that are actually tuberculosis, while the precision score of 0.904762 shows the accuracy in the range of cases classified as tuberculosis.

From the ROC curve on the right side shown in Figure 9, it is seen that the model achieved 0.970732 AUC score. This means the model achieved a good ability to discriminate between chest X-rays with and without tuberculosis. The high ROC curve, along with the AUC value quite close to 1, shows a good performance over the entire range of classification thresholds. This can be beneficial when, in real-world applications, it may be necessary to adjust the thresholds. From the evaluations in the scope of this study, the EfficientNetB0 model with the training strategy using option B is a good choice in further research or entering the implementation of practical medical applications.

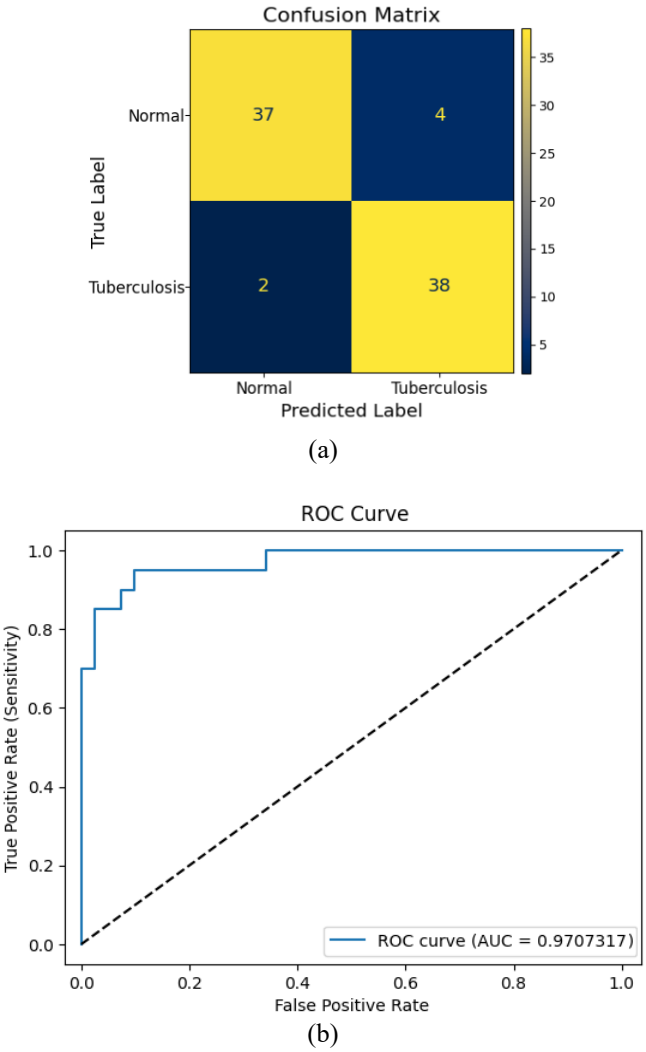

Figure 9. EfficientNetB0 model with the training strategy using option B for (a) confusion matrix and (b) receiver operating characteristic (ROC) curve

This result is similar to another study that also used CLAHE for preprocessing [29]. The authors used EfficientNetV2 to detect pulmonary tuberculosis in a simulated emergency department environment and achieved an AUC of 0.878. When assessed on external data, the system maintained stable performance with AUC of 0.838 on the Montgomery set and 0.806 on the Shenzhen set. Compared to our study, both demonstrate the potential of applying image preprocessing to the performance of deep learning models in pulmonary tuberculosis detection, and here it is CLAHE-based preprocessing.

Besides these results, we also augment with additional relevant recent studies that reported improvements with CLAHE. Besides these results, we also augment with additional relevant recent studies that reported improvements with CLAHE. In study [28], the authors presented their work on eye detection in thermal images that used CLAHE to improve performance. According to the paper, when experimenting with models, the authors reached an average of precision as 0.990250 in the case of raw thermal images of human faces. Besides, the authors reached an average of precision as 0.998000 in the case of CLAHE-enhanced thermal images. Therefore, applying image enhancement with CLAHE improved the average precision score of the models by 0.007750. In study [30], the authors applied image quality enhancement with CLAHE to the problem of breast cancer diagnosis from mammogram images. The authors used the model based on VGG-16 and experimented on a total of 1,459 mammograms. When testing in the case of raw images with artefact removal, the authors achieved 88.45% of the test accuracy and 88.62% of the F1-score. When applying CLAHE once, the authors achieved 94.63% of the test accuracy and 94.74% of the F1-score, and when applying CLAHE twice, the authors achieved 95.21% of the test accuracy and 93.28% of the F1-score. Therefore, applying image enhancement with CLAHE improved the test accuracy by a maximum of 6.76%

with CLAHE twice, and improved the F1-score by a maximum of 4.66% with CLAHE once. In study [31], the authors showed their work on diabetic retinopathy detection. By applying CLAHE, the authors reached the highest accuracy of 99.69%, which improved by 2% from the case of not applying CLAHE.

To discuss the model's interpretability and practical utility, we conducted an experiment with some images containing expert radiologist annotations. These images come from the Vietnam National Lung Hospital and were taken with the assistance of radiologists. Some results are described in Figure 10.

Figure 10 shows some visualization results using Grad-CAM with the EfficientNetB0 model with the training strategy using option B on images from the Vietnam National Lung Hospital. Specifically, the left column is the results of contrast enhancement, the middle column is the image with pulmonary tuberculosis lesions marked by radiologists, and the right column is the visualization results using Grad-CAM. In the marked areas, the blue area is the infiltrated area (consolidation), the pink area is the large nodule area (size >2 mm), and the orange area is the calcified pleural thickening area. We can see that the visualization results using Grad-CAM have a certain similarity compared to the areas marked by the doctor. This indicates that the trained model results have shown the ability to highlight the location of lesions on the test image. There are also examples where the Grad-CAM highlighted range is not really accurate, perhaps with small lesions or even highlighted areas falling into areas without lesions. This shows that there are cases where the model does not show the focus on the lesion when making an inference. An important point to mention here is the difference in the spatial domain of the data, and this has also been discussed in the study [12]. The model was trained with data from international sources and tested with local data of Vietnamese people from the Vietnam National Lung Hospital. Therefore, the difference in the spatial domain of this data affected the effectiveness of the model on the test data in Vietnam. Thus, the resulting model has shown some results reflecting the location of the lesions compared to the areas marked by radiologists. In future studies, to have a more specific evaluation as well as to orient towards a product that can be applied in hospitals in Vietnam, a focus of our research will be to conduct more research with Vietnamese data, as well as generalization of the model on data domains.

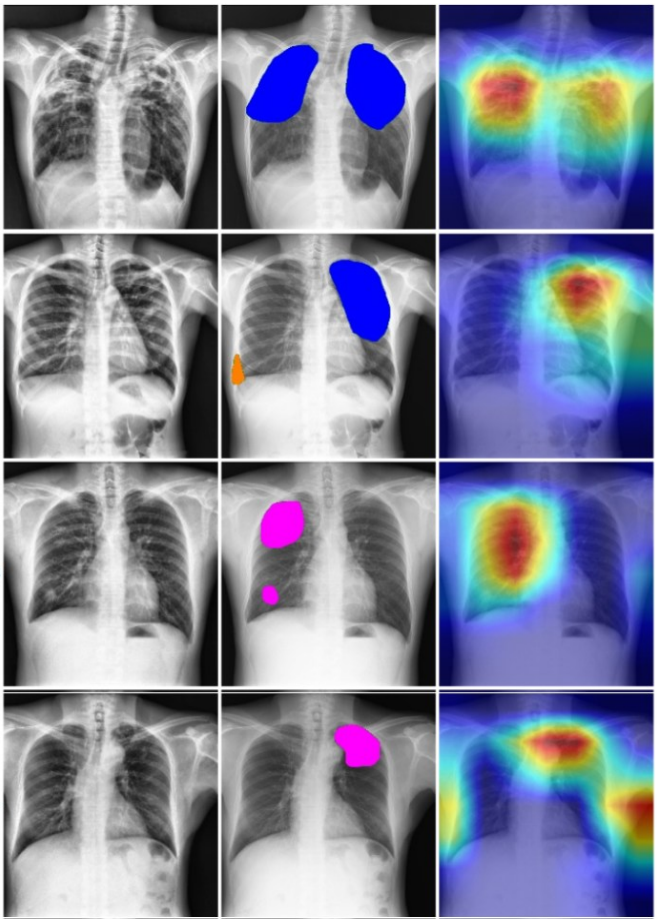

Figure 10. Grad-CAM visualization from the EfficientNetB0 model with the training strategy using option B with images from the Vietnam National Lung Hospital

4. CONCLUSION

The paper presents a study aimed at building an automatic diagnosis solution for pulmonary tuberculosis from X-ray data sources. In it, we hypothesize the role of image enhancement and build a data processing process based on that hypothesis with the CLAHE technique. Experimental results show that applying our process helps to improve classification performance stably and significantly on many deep learning models. Important indicators such as AUC, accuracy, F1-score, precision, and recall all have clear improvements. In addition, Grad-CAM analysis shows that the model works well with clinically meaningful features, reflecting the model's explanatory ability in the problem. Based on this study, integrating the convolutional neural network model with CLAHE-based preprocessing into the diagnostic support system can help radiologists detect pulmonary tuberculosis lesions faster and more accurately. CLAHE improves image quality, thereby supporting the AI model. As a result, the diagnostic support system can provide early warnings to doctors and reduce the danger of missing cases, which is especially useful in environments with high workloads or shortages of specialized human resources. For deployment, the system can be optimized and deployed on mobile devices or edge devices to serve tuberculosis screening in the community, especially in rural areas. Setting up CLAHE-based preprocessing helps improve the quality of images taken from mobile X-ray machines. With the ability to process quickly, this solution can provide timely results, helping lower-level doctors or community health workers have more basis to decide on timely referral or initial treatment. In the near future, CLAHE-based preprocessing will still be a simple, effective, and easy-to-implement method for improving the quality of chest X-ray data in research and clinical applications. However, CLAHE cannot solve all other problems such as noise caused by the scanner or the quality of images from different types of devices. In addition, sometimes the images will become fake and unnatural if used excessively, affecting the reliability of the AI system. Therefore, in the future, CLAHE-based preprocessing can be improved or combined with more advanced techniques such as deep learning preprocessing. In future studies, we plan to continue to expand the testing on many different datasets with other lung diseases to more comprehensively evaluate the generalizability of the model and the hypothesis about the role of image enhancement. These studies are expected to aim at building a comprehensive diagnostic support system and bringing research results to socio-economic life.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Nguyen Trong Vinh		✓	✓			✓	✓	✓	✓		✓			✓
Lam Thanh Hien	✓				✓	✓				✓			✓	
Ha Manh Toan	✓		✓				✓	✓	✓					✓
Do Nang Toan	✓	✓		✓	✓					✓		✓	✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are openly available in Kaggle at https://www.kaggle.com/datasets/kmader/pulmonary-chest-xray-abnormalitie, reference number [18].

REFERENCES

[1] World Health Organization, "World tuberculosis day 2022: invest to end TB. save lives," *WHO.int*. Accessed: Aug. 27, 2025. [Online]. Available: https://www.who.int/campaigns/world-tb-day/2022

[2] Centers for Disease Control and Prevention, "Treating tuberculosis," *CDC.gov*. Accessed: Aug. 27, 2025. [Online]. Available: https://www.cdc.gov/tb/treatment/index.html

[3] E. Elveren and N. Yumuşak, "Tuberculosis disease diagnosis using artificial neural network trained with genetic algorithm," *Journal of Medical Systems*, vol. 35, no. 3, pp. 329–332, Jun. 2011, doi: 10.1007/s10916-009-9369-3.

[4] J. Dongardive, A. Xavier, K. Jain, and S. Abraham, "Classification and rule-based approach to diagnose pulmonary tuberculosis," in *Advances in Computing and Communications*, Springer, Berlin, Heidelberg, 2011, pp. 328–339, doi: 10.1007/978-3-642-22709-7_34.

[5] P. Lakhani and B. Sundaram, "Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks," *Radiology*, vol. 284, no. 2, pp. 574–582, Aug. 2017, doi: 10.1148/radiol.2017162326.

[6] P. Rajpurkar *et al.*, "CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning," 2017, *arXiv: 1711.05225*.

[7] R. Hooda, S. Sofat, S. Kaur, A. Mittal, and F. Meriaudeau, "Deep-learning: a potential method for tuberculosis detection using chest radiography," in *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Sep. 2017, pp. 497–502, doi: 10.1109/ICSIPA.2017.8120663.

[8] P. Rajpurkar *et al.*, "Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists," *PLOS Medicine*, vol. 15, no. 11, Nov. 2018, doi: 10.1371/journal.pmed.1002686.

[9] T. B. Chandra, K. Verma, B. K. Singh, D. Jain, and S. S. Netam, "Automatic detection of tuberculosis related abnormalities in chest X-ray images using hierarchical feature extraction scheme," *Expert Systems with Applications*, vol. 158, Nov. 2020, doi: 10.1016/j.eswa.2020.113514.

[10] F. Pasa, V. Golkov, F. Pfeiffer, D. Cremers, and D. Pfeiffer, "Efficient deep network architectures for fast chest X-ray tuberculosis screening and visualization," *Scientific Reports*, vol. 9, no. 1, Apr. 2019, doi: 10.1038/s41598-019-42557-4.

[11] S. S. Meraj, R. Yaakob, A. Azman, S. N. M. Rum, A. S. A. Nazri, and N. F. Zakaria, "Detection of pulmonary tuberculosis manifestation in chest X-rays using different convolutional neural network (CNN) Models," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 1, pp. 2270–2275, Oct. 2019, doi: 10.35940/ijeat.A2632.109119.

[12] H. M. Toan, L. T. Hien, N. D. Vinh, and D. N. Toan, "Detecting tuberculosis from Vietnamese X-ray imaging using transfer learning approach," *Computers, Materials & Continua*, vol. 74, no. 3, pp. 5001–5016, 2023, doi: 10.32604/cmc.2023.033429.

[13] A. Wong, J. R. H. Lee, H. R.-Khah, A. Sabri, A. Alaref, and H. Liu, "TBNet: a tailored, self-attention deep convolutional neural network design for detection of tuberculosis cases from chest X-Ray images," *Frontiers in Artificial Intelligence*, vol. 5, Apr. 2022, doi: 10.3389/frai.2022.827299.

[14] N. T. Vinh, L. T. Hien, H. M. Toan, N. D. Vinh, and D. N. Toan, "Tuberculosis diagnosis and visualization with a large Vietnamese X-ray image dataset," *Intelligent Automation & Soft Computing*, vol. 39, no. 2, pp. 281–299, 2024, doi: 10.32604/iasc.2024.045297.

[15] N. Patel, A. Wong, and A. Ebadi, "Empowering tuberculosis screening with explainable self-supervised deep neural networks," in *2024 International Conference on Machine Learning and Applications (ICMLA)*, Dec. 2024, pp. 794–797, doi: 10.1109/ICMLA61862.2024.00113.

[16] R. Bista *et al.*, "Advancing tuberculosis detection in chest X-rays: a YOLOv7-based approach," *Information*, vol. 14, no. 12, Dec. 2023, doi: 10.3390/info14120655.

[17] E. Kotei and R. Thirunavukarasu, "Tuberculosis detection from chest X-ray image modalities based on transformer and convolutional neural network," *IEEE Access*, vol. 12, pp. 97417–97427, 2024, doi: 10.1109/ACCESS.2024.3428446.

[18] K. S. Mader, "Pulmonary chest X-ray abnormalities," *Kaggle.com*. Accessed: Aug. 27, 2025. [Online]. Available: https://www.kaggle.com/datasets/kmader/pulmonary-chest-xray-abnormalities

[19] B. Koonce, *Convolutional neural networks with swift for TensorFlow*. Berkeley, CA: Apress, 2021, doi: 10.1007/978-1-4842-6168-2.

[20] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.

[21] A. G. Howard *et al.*, "MobileNets: efficient convolutional neural networks for mobile vision applications," Apr. 2017, *arXiv:* 1704.04861.

[22] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems*, Elsevier, 1994, pp. 474–485, doi: 10.1016/B978-0-12-336156-1.50061-6.

[23] E. Riba, D. Mishkin, D. Ponsa, E. Rublee, and G. Bradski, "Kornia: an open source differentiable computer vision library for PyTorch," in *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Mar. 2020, pp. 3663–3672, doi: 10.1109/WACV45572.2020.9093363.

[24] PyTorch Foundation, "PyTorch: an open source machine learning framework," *Pytorch.org*. Accessed: Aug. 27, 2025. [Online]. Available: https://pytorch.org/

[25] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *3rd International Conference for Learning Representations*, 2015. doi: 10.48550/arXiv.1412.6980.

[26] L. N. Smith, "A disciplined approach to neural network hyper-parameters: part 1 - learning rate, batch size, momentum, and weight decay," Apr. 2018, *arXiv: 1803.09820.*

[27] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: visual explanations from deep networks via gradient-based localization," in *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct. 2017, pp. 618–626, doi: 10.1109/ICCV.2017.74.

[28] P. J. and S. A, "Synergistic fusion: an integrated pipeline of CLAHE, YOLO models, and advanced super-resolution for enhanced thermal eye detection," *PLOS One*, vol. 20, no. 7, Jul. 2025, doi: 10.1371/journal.pone.0328227.

[29] C.-H. Wang *et al.*, "Deep learning–based diagnosis of pulmonary tuberculosis on chest X-ray in the emergency department: a retrospective study," *Journal of Imaging Informatics in Medicine*, vol. 37, no. 2, pp. 589–600, Jan. 2024, doi: 10.1007/s10278-023-00952-4.

[30] S. Montaha *et al.*, "BreastNet18: a high accuracy fine-tuned VGG16 model evaluated using ablation study for diagnosing breast cancer from enhanced mammography images," *Biology*, vol. 10, no. 12, Dec. 2021, doi: 10.3390/biology10121347.

[31] S. Phimphisan and N. Sriwiboon, "A customized CNN architecture with CLAHE for multi-stage diabetic retinopathy classification," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 18258–18263, Dec. 2024, doi: 10.48084/etasr.8932.

BIOGRAPHIES OF AUTHORS

Nguyen Trong Vinh joined a Master's degree in Information Technology at Lac Hong University, Vietnam, and received a degree in 2011. In 2022, he joined the Ph.D. program at Lac Hong University. He is currently working at Lac Hong University as the Head of the Training Department. His research interests consist of computer vision, machine learning, and deep learning. He can be contacted at email: trongvinh@lhu.edu.vn.

Lam Thanh Hien joined a M.Sc. in applied informatics at the INNOTECH Institute, France, and earned a degree in 2004. In 2017, he received a Ph.D. degree at the Vietnam Academy of Science and Technology. Now, he is working at Lac Hong University in the headmaster's role. His studies interests relate to computer vision, machine learning, and deep learning. He can be contacted at email: lthien@lhu.edu.vn.

Ha Manh Toan studied applied mathematics and informatics at the College of Science, Vietnam National University, Hanoi, and earned a degree in 2009. In 2015, he received an M.Sc. degree from the University of Engineering and Technology, Vietnam National University, Hanoi. Now, he is working as a researcher at the Vietnamese Academy of Science and Technology. His study interests relate to computer vision, machine learning, and deep learning. He can be contacted at email: hmtoan@ioit.ac.vn.

Do Nang Toan learned applied mathematics and informatics at Hanoi University and received a degree in 1990. In 2001, he received a Ph.D. degree from the Vietnam Academy of Science and Technology. Now, he is working as an associate professor at the Vietnamese Academy of Science and Technology. His study interests relate to computer vision, machine learning, and virtual reality. He can be contacted at email: dntoan@ioit.ac.vn.