IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 21, 2020, accepted November 29, 2020, date of publication December 2, 2020,
date of current version December 15, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3041867

Image Enhancement for Tuberculosis Detection Using Deep Learning

KHAIRUL MUNADI[1,2,8], (Member, IEEE), KAHLIL MUCHTAR[1,2], (Member, IEEE), NOVI MAULINA[3], AND BISWAJEET PRADHAN[4,5,6,7], (Senior Member, IEEE)

[1]Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia
[2]Telematics Research Center (TRC), Universitas Syiah Kuala, Banda Aceh 23111, Indonesia
[3]Faculty of Medicine, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia
[4]Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney, Sydney, NSW 2007, Australia
[5]Department of Energy and Mineral Resources Engineering, Sejong University, Seoul 05006, South Korea
[6]Center of Excellence for Climate Change Research, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[7]Earth Observation Center, Institute of Climate Change, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia
[8]Tsunami Disaster Mitigation Research Center (TDMRC), Universitas Syiah Kuala, Banda Aceh 23111, Indonesia

Corresponding author: Kahlil Muchtar (kahlil@unsyiah.ac.id)

This work was supported in part by the Institute for Research and Community Services (LPPM), Universitas Syiah Kuala, Indonesia, under Grant 9/UN11.2.1/PT.01.03/PNBP/2020, and in part by the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia through the 2019 World Class Professor (WCP) Programme.

ABSTRACT The latest World Health Organization's (WHO) study on 2018 is showing that about 1.5 million people died and around 10 million people are infected with tuberculosis (TB) each year. Moreover, more than 4,000 people die every day from TB. A number of those deaths could have been stopped if the disease was identified sooner. In the recent literature, important work can be found on automating the diagnosis by applying techniques of deep learning (DL) to the medical images. While DL has yielded promising results in many areas, comprehensive TB diagnostic studies remain limited. DL requires a large number of high-quality training samples to yield better performance. Due to the low contrast of TB chest x-ray (CXR) images, they are often is in poor quality. This work assesses the effect of image enhancement on performance of DL technique to address this problem. The employed image enhancement algorithm was able to highlight the overall or local characteristics of the images, including some interesting features. Specifically, three image enhancement algorithms called Unsharp Masking (UM), High-Frequency Emphasis Filtering (HEF) and Contrast Limited Adaptive Histogram Equalization (CLAHE), were evaluated. The enhanced image samples were then fed to the pre-trained ResNet and EfficientNet models for transfer learning. In a TB image dataset, we achieved 89.92% and 94.8% of classification accuracy and AUC (Area Under Curve) scores, respectively. All the results are obtained using Shenzhen dataset, which are available in the public domain.

INDEX TERMS Image enhancement, convolutional neural network (CNN), ResNet, EfficientNet, binary tuberculosis classification.

I. INTRODUCTION

Millions of people around the world die each year because of tuberculosis (TB). Around 1.5 million deaths due to TB are reported in 2018 alone, according to the WHO's report. It is mainly prevalent in Africa and Southeast Asia. This is a very infectious disease caused by a TB of the bacillus mycobacterium [36]. In recent developments, there are several diagnostic methods that utilize molecular analysis and bacteriological culture. Unfortunately, they are still high-cost, especially for most of developing countries which are affected by the disease. It is also stated that the low-cost and most common diagnostic technique so-called sputum smear microscopy have problems with sensitivity [1]. In recent years, DL has performed well in the area of image recognition and classification, and one of the most popular models are convolutional neural network (CNN) model [37]. DL techniques have been successfully been implemented in many different fields, such as surveillance system [2], face recognition [3], [4], autonomous cars [5], vehicle classification [6], and many others [7]–[9], [50]. In addition, there are already numerous CAD (computer-aided design) systems that use

The associate editor coordinating the review of this manuscript and approving it for publication was Yudong Zhang.

CNNs to diagnose disease [10], [11], but their application to TB detection remains limited [38]. The key problems are the limited size of the public dataset and relying on preset feature models. In recent literature, some automated TB detection systems have utilized the DL technique, such as the works proposed by Liu, *et al.* [14], Lopes and Valiati [12], Lakhani and Sundaram [13], Hwang *et al.* [15], and Rajaraman and Antani [42]. This research focuses on evaluating the effect of the pre-processing step for the performance of the DL technique thoroughly, which has not been reported in the aforementioned literature.

In most of the cases, the training of the new CNN model and the fine tuning of the pre-trained CNN model (transfer learning) are two essential ways to train the CNN model [12], [14]. However, regardless of the training methods employed, the image datasets are typically processed in various ways prior to training the CNN models, such as image cropping and image enhancement. The quality of the image data sets greatly affects the model's performance [33]. In the medical CXR imaging method, the operator's expertise, the patient's own consideration and certain other factors that may cause the imaging effect are not optimal, such as low brightness, low contrast, and bad or blurry information. According to Koo and Cha [16], image pre-processing is important when training CNN models, which can effectively boost the performance of CNN models in classification. Also, the enhancement of images is a very important part of pre-processing. Hence, research into the relationship between image enhancement and the CNN model is important.

The main objective of this work is to evaluate the effect of two different pre-processing approaches (UM [17] and HEF [18]) on the use of pre-trained CNN to detect TB disease. The UM and HEF algorithms have proven to perform well on medical images [19], [20]. This work, hopefully, opens the opportunity to consider other state-of-the-arts enhancement techniques to be performed prior to medical image classification. Therefore, in this research we used a common approach as a baseline and early-work incorporating TB image enhancement and deep learning model. Because of its outstanding accuracy and robustness, we used ResNet [21] and EfficientNet [47] as a DL architecture. The most important contributions are as follows:

- The quantitative analysis of the performance of ResNet and EfficientNet after training of enhanced image datasets through UM and HEF algorithms is presented.
- In addition, several fine-tune parameters, including enhancement parameters, are also carefully discussed. This is the first time, to our knowledge, that all of these observations have been identified in the literature.

The paper is divided as follows: the most important works are described in Section 2. Section 3 discusses the data collection and methods utilized in the image enhancement phase. Section 4 explains the proposed final detection pipeline and its results.

II. RELATED WORKS

A. TB DETECTION USING CONVENTIONAL MACHINE LEARNING ALGORITHM

TB detection is a difficult task that mainly because of various types of manifestations such as large opacities, aggregation, focal lesions, cavities, small opacities, and CXR image nodules. Related papers (especially those using a machine learning algorithm) use color, texture, shape or geometry features to detect the TB from an image [22]–[24]. One of the important work is proposed by Chauhan *et al.* [24]. They implemented a single system for TB detection through CXR images. Their system consists of a series of modules that must obey several steps in order to classify the input image. The dataset begins with a denoising-based pre-processing module, followed by an extraction function. Finally, a model is constructed through the SVM classifier.

B. TB DETECTION USING PRE-TRAINED CNN (TRANSFER LEARNING)

CNN is commonly used for various computer vision applications, especially for classification, detection, and recognition task. Typically, the CNN model consists of several layers, namely pooling, convolutional, and fully-connected (FC) layers. The preceding layer by means of kernels with a predefined, fixed-size receptive field is connected to every layer. The CNN model learns the setup of hyper-parameters from a big data collection to represent the image's global or local characteristics. That model architecture has different layer types and activation functions to display better representational features than human-engineered software.

In recent literature [12], [14], [44], pre-trained CNN (where an ImageNet-trained network is used) shows good performance in the medical domain. For efficiency purposes, the researcher can avoid train more than a million images, which also requires a large amount of memory and computation. This is a method called transfer learning. Transfer learning aims at storing information gained from one domain and applying it to another, still similar domain [39]. It usually takes a lot of time when training from scratch, as the random Gaussian distribution is utilized to initialize all model parameters. Typically, the convergence is achieved with a batch size of 50 images, and after at least 30 epochs [14].

Recent studies [25], [26] have shown that fine tuning of a more complex dataset results in excellent classification and detection performance using a pre-trained ImageNet dataset model. The reason for this training procedure is that CNN receives a general representation of natural images from pre-training. The model adjusts the parameter after the fine-tuning to show the specific features of the individual images, while retaining the ability to display the general image. This training strategy is implicitly implemented, coupled with a sampling of shuffles and cross-validation.

It is noteworthy that most of the previous existing approaches for TB detection have adopted pre-trained CNN steps. As one of the initial work on TB detection through DL,

Hwang *et al.* [27] proposed to collect and train huge private data collection consisting of approximately 10,000 images. Unlike many other areas, the literature of TB detection through DL remains limited. The recent works are proposed by Lakhani and Sundaram [13], Hooda *et al.* [28], Liu *et al.* [14], Lopes and Valiati [12], and Rajaraman and Antani [42] to name a few. Lakhani and Sundaram [13] evaluated the efficacy of CNN for detecting TB. Both untrained and pre-trained on ImageNet were used and augmented with numerous pre-processing techniques. Hooda *et al.* [28] employed 7 convolution layers and 3 FC layers of CNN architecture. They compared three different optimizers, and Adam optimizer achieved the best results. Liu *et al.* [14] proposed a DL method with CNN and transfer learning. They used AlexNet [29] and GoogleNet [30] as pre-trained networks. Different from the aforementioned works, Lopes and Valiati [12] focused on investigation whether pre-trained CNN as feature extractors are able to detect TB disease or not. Recently, Rajaraman and Antani [42] improved the accuracy of TB detection through modality-specific deep learning models. The pre-trained CNNs are employed to train large-scale datasets including the RSNA dataset, Pediatric pneumonia dataset and Indiana dataset.

C. IMAGE ENHANCEMENT EFFECT FOR PRE-TRAINED CNN

Throughout this study, we evaluated the impact of image enhancement on a pre-trained CNN model. This influence has been studied in the literature in various fields. In [31], Tian *et al.* reported that using image enhancement function with Laplace operator can improve the performance of fast R-CNN and R-CNN in the pedestrian detection task. They compared the enhanced image dataset with the Laplace operator and the original image. The enhanced approach could increase the detection rate by 2% and 1% for the two R-CNN models, respectively. The paper's experiments used the transfer learning by fine tuning of the pre-trained R-CNN model. Kuang *et al.* [32] proposed a night-time vehicle detection approach that has three key techniques: night-time image enhancement focused on ROI extraction incorporating light detection for the vehicle. In [33], Chen provided a comprehensive evaluation of image enhancement on RGB and general CXR images. In addition, the accurate detection of retinal blood vessels via deep learning and image enhancement is discussed in a recent paper by Soomro, *et al.* [43]. While several works have been proposed, none of the work thoroughly investigates the effect of image enhancement through pre-trained CNN for TB detection.

III. METHODS

A. DATASET

All the experiments in this research have been tested using the Shenzhen Public Dataset [34]. The Shenzhen data set was collected at Shenzhen Hospital, Guandong Providence, China. The dataset consists of 662 frontal CXR, of which 336 are TB infected and 326 are not disease-infected. All images have a resolution of about 3K×3K pixels. As one of the largest infectious disease hospitals in China, Shenzhen hospital aims at improving both its prevention and treatment. The CXR photos were captured within a period of one month, mostly in September 2012, as part of Shenzhen hospital's daily routine, using a digital diagnosis system from Philips DR. For the Shenzhen dataset, it has a corresponding radiologist reading, which is then considered being ground-truth. In Fig. 1, we visualize the example of CXR images in this dataset.

B. METHODOLOGY

1) TB CXR IMAGE ENHANCEMENT BASED ON UNSHARP MASKING (UM) AND HIGH-FREQUENCY EMPHASIS FILTERING (HEF)

The quality of CXR images is not always good because the patient-related x-ray amount is small considering patient safety [48]. The image enhancement techniques have been used to improve the image quality and achieved desired solution [19]. The widely used methods for improving images include histogram equalization [40], low pass filtering, and high pass filtering [41]. The basic principle of histogram equalization is to enhance the image's visual effect by expanding the image distribution's gray scale spectrum. This allows the process to accomplish the purpose of improving the contrast. As a result, the original image, which has several gray levels, are then synthesized into one gray level, resulting in image information loss. Only image smoothing or sharpening may be accomplished by the high-pass filter or low-pass filter [18]. In image enhancement, it is typically important not only to change the image's dynamic range but also to highlight the image's characteristics. Color Enhancement Technology is an image-based technique designed to increase visual effects through digital image processing and transformation to reflect color quality and enhance contrast [33], [41].

Using two successful image enhancement algorithms; UM and HEF, this research improved the visualization of TB CXR images before DL classification stage. Every method was independently executed and evaluated. It is noteworthy, in the next section, CLAHE will also be evaluated by calculating over-enhancement measures. Fig. 2 shows the illustration of the image enhancement step:

2) UNSHARP MASKING (UM)

UM [17] is an image sharpening technique that produces a mask of the original image using a distorted, or ''unsharp,'' negative image. The unsharp image is then combined with the original positive image, creating a less blurred image than the original image. In other words, UM is a linear filter capable of amplifying an image's high-frequencies.

The first step of the algorithm was to copy the original image and apply Gaussian blur to it. The radius is an important setting when performing Gaussian blur. Radius is related

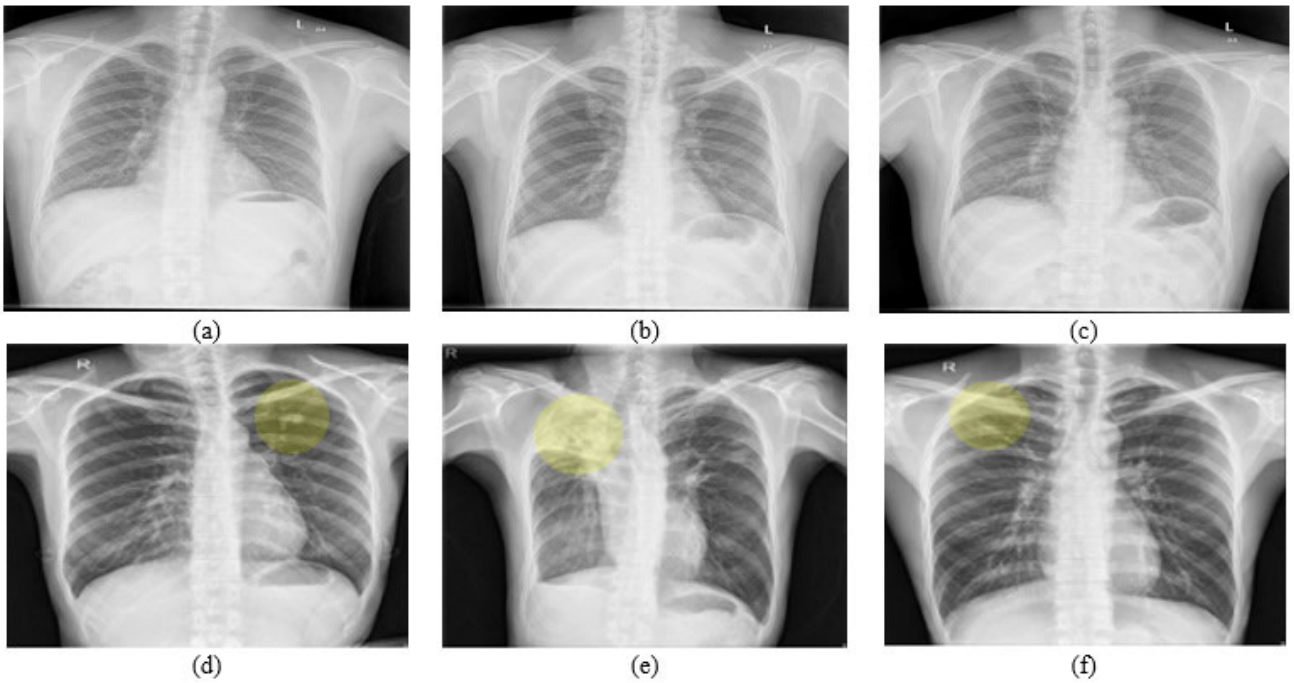


FIGURE 1. Example of CXRs in the Shenzhen dataset. (a) The top left CXR is from a 28 year-old male, (b) center image from a 39-year-old male and (c) the top rights from a 38-year-old male (no TB). The bottom three images are cases of TB: (d) the left CXR is from a 28 year-old female (left secondary TB), (e) center image is from a 50 year-old male (bilateral secondary TB with right upper atelectasis; 2.right pleural adhesions; 3.left compensatory emphysema) and (f) the right CXR is from a 32-year-old male (secondary TB in the right upper field).

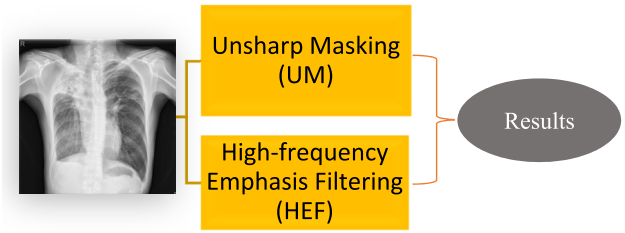


FIGURE 2. Image processing of gray images.

to blur intensity as it defines the size of the edges.

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

where i and j are in horizontal and vertical axis, respectively, distance from the origin. Furthermore, σ is the standard deviation of the Gaussian distribution.

Next, we subtracted the blurred image from the original image, and we will only get the blurred edges. This is what we called the "unsharp mask." Finally, after applying the following equation, the enhanced image is collected:

$$I_{um\_enhanced} = I_{ori} + \text{amount} * (I_u) \quad (2)$$

where $I_{um\_enhanced}$, I_{ori}, and I_u are final result after applying UM algorithm, the original image and unsharp image, respectively. To be clear, the radius and the amount (in Eq.(2)) are described below:

- Amount can be thought of as how much contrast is added to the edges (how much dark or light it will be).
- Radius influences the size of the edges to be improved or how large the edge rims are, so a smaller radius improves the accuracy of smaller scale.

Different guidelines exist for starting values of these parameters, and the significance may vary from one implementation to another. In our experiment on TB CXR images, the best two values for radius and amount, respectively, are 5 & 2. Fig.3 provides an illustration between the original image and the enhanced image by UM:

3) HIGH-FREQUENCY EMPHASIS FILTERING (HEF)

HEF [18] is a technique that uses a Gaussian high-pass filter to emphasize and accentuate the edges. The edges tend to be expressed in the high-frequency spectrum because they have more significant changes in intensity. This technique produces a low contrast image, and the use of histogram equalization is necessary to increase sharpness and contrast.

The filter step of the algorithm is to apply the Gaussian high-pass filter (intensity depends on the setting of the radius) (see Eq. (3)). As shown in Eq. (4), the image has to go through the transformation of Fourier and the filter function is calculated on it. We will have the image filtered after the inverse transformation (see Eq. (5)). Fig. 4 illustrates the step-by-step process of the HEF technique. In addition, the comparison of each technique is shown in Fig. 5.

$$G\_Filter(i, j) = 1 - e^{-D^2(i,j)/2D_0^2} \quad (3)$$

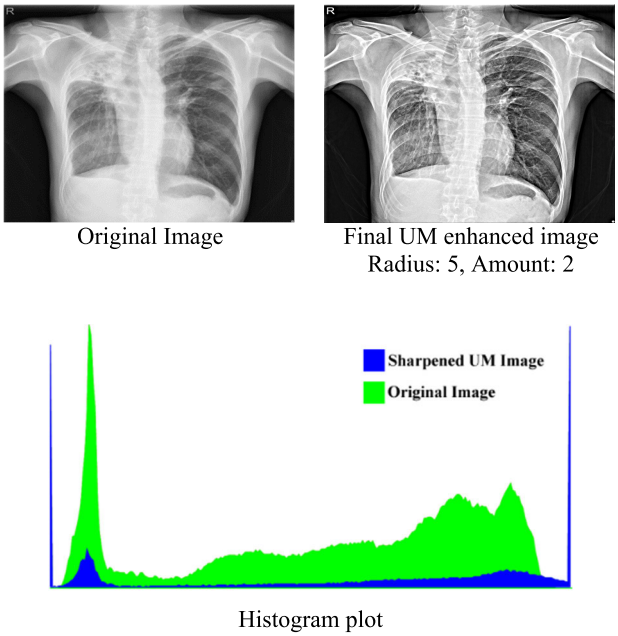


FIGURE 3. Illustration of the image comparison between original image and enhanced image through UM.

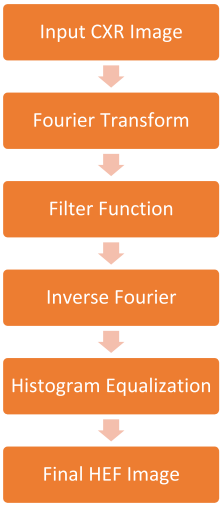


FIGURE 4. Using Fourier-domain filtering to apply HEF.

where D_0 is the cut off distance. The Fourier transform of f(x,y), denoted by F(i,j), for x and i = 0,1,2…M − 1 and y and j = 0,1,2…N − 1, is given by the equation:

$$F(i,j) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)\, e^{-j2\pi(\frac{ix}{M}+\frac{jy}{N})} \tag{4}$$

The inverse Fourier transform is given by:

$$f(x,y) = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} F(i,j)\, e^{j2\pi(\frac{ix}{M}+\frac{jy}{N})} \tag{5}$$

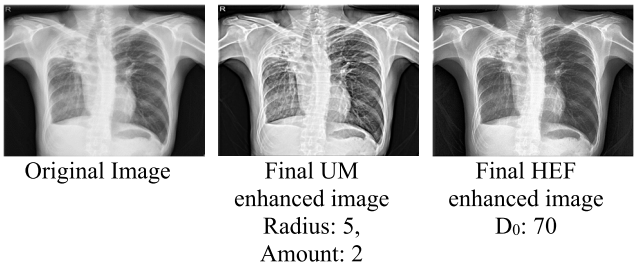


FIGURE 5. The image comparison between original TB CXR, enhanced UM, and enhanced HEF images.

Finally, the contrast of the image was adjusted with a simple histogram equalization (Hist_Eq):

$$I_{hef\_sharpened} = (I_{ori} + (G\_Filter)) * (Hist\_Eq) \tag{6}$$

4) PRE-TRAINED ResNet AND EfficientNet MODELS

The input of CNN is assumed to be in the form of an image. This allows us to encode certain properties in the network. In our work, since the input is a grayscale image, the array of pixels will be height × width × depth (e.g., 224 × 224 × 1). For generating useful features from the images data, specific parameters were modified in the Conv layers throughout the training. Concurrently, the collection of learned parameters in the FC layers classify the extracted features in the target classes (normal image or TB image). Conv layers acquire visual features from raw input images in a hierarchical manner. The low-level features, such as shapes or edges, are obtained from lower layers, while high-level visual features, such as object part are obtained from higher-layers. The images are labeled as vectors. In our case, the label is binary, either the value '1' for the TB image or '0' for the other image.

The experiment selected the ResNet-18, ResNet-50, and EfficientNet-B4 models in transfer learning. ResNet [21] is a neural network that uses skip connections to leap over several layers (as shown in Fig. 6). The purpose to leap over layers is to prevent the issue of vanishing gradients by reusing activations from the previous layer before the neighboring layer knows its weights. The efficiency of the extracted features on the network depends on the depth of the architecture. The ResNet is made up of various layers in which each layer is usually inserted numerous times. Note that for the design of the ResNet-50, each 2-layer block in the 34-layer net is replaced by a 3-layer bottleneck block [21].

FIGURE 6. Illustration of a ResNet. A layer k-1 is skipped over activation from *k*-2.

By implementing a compound scaling approach in all network dimensions, i.e. distance, depth, and resolution, due

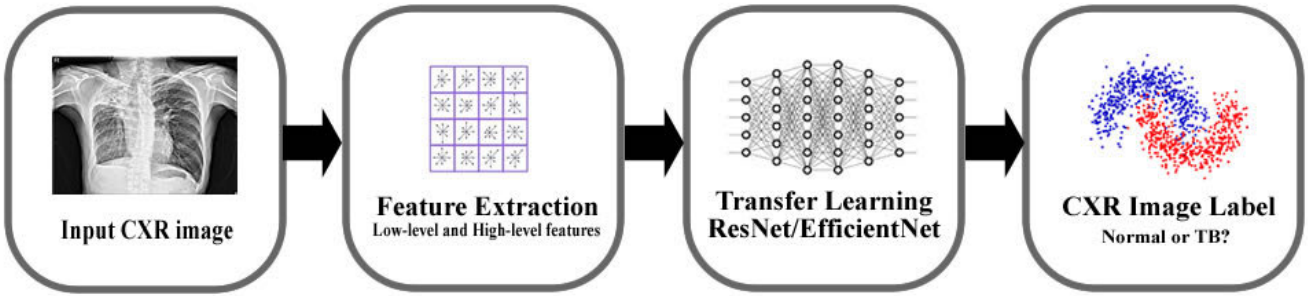


FIGURE 7. The deep learning-based architecture for TB detection through ResNet and EfficientNet-B4.

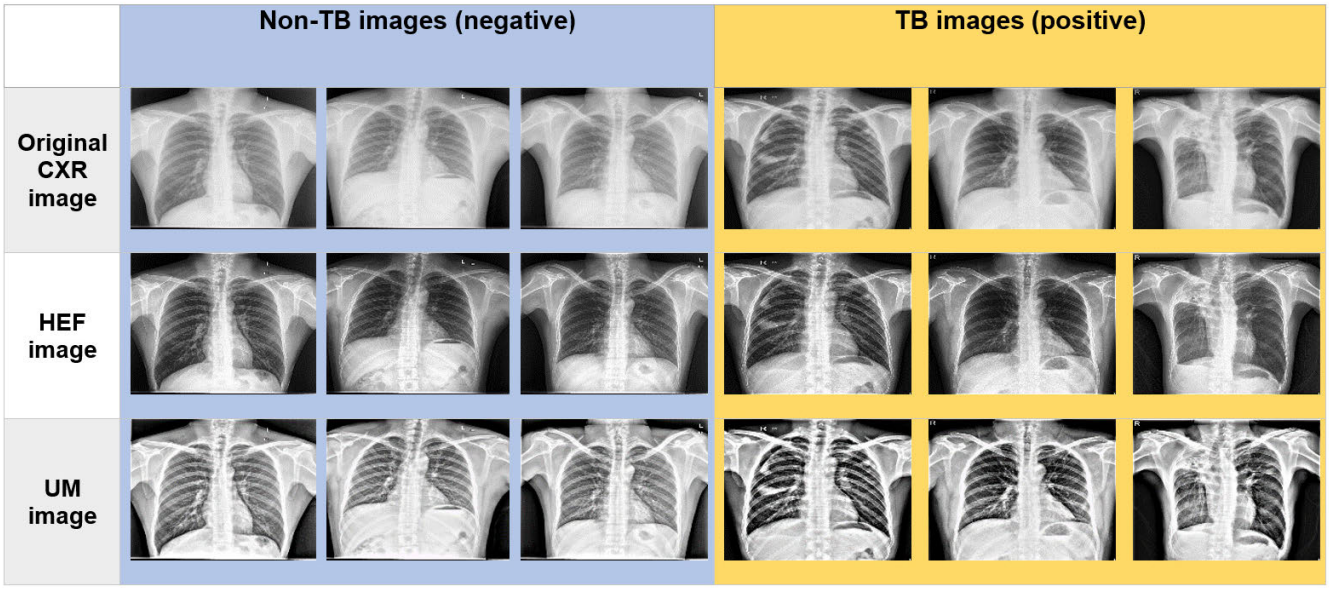


FIGURE 8. Some example of testing images during experiment.

to its superiority in prediction efficiency, EfficientNet [47] has attracted attention. EfficientNet is a collection of models, namely EfficientNet-B0 to B7 that were derived from the base network (so-called EfficientNet-B0). The advantages of EfficientNet are expressed in two ways, it provides higher precision, and also increases the performance of the model by reducing the parameters and FLOPS (Floating Point Operations per second). Fig. 7 explains the deep learning-based architecture used in our work.

First, the initial resizing of the input images was performed prior to enhancing the images. This will output a uniform size of 640 × 480 pixels of enhanced input images. Next, we used transfer learning from a pre-trained network, and the image was resized to 224 × 224 pixels (as commonly used by pre-trained models on the ImageNet database). In other words, a crop of 224 × 224 of enhanced image was used. Specifically, a crop of 224 × 224 image or horizontal flip was sampled randomly, and subtract the mean of that pixel across all images. This is called per-pixel mean subtraction. Fig. 8 visualizes some example of testing images during experiment. In table 1, we show the parameter used in the training stage.

TABLE 1. Parameter used in the training stage.

Parameter	ResNet	EfficientNet-B4
Batch Size	6	2
Num of workers	10	10
Max Epoch	10	10
Learning rate	0.01	0.01
Learning rate decay	0.5	0.5
Learning rate stepsize	3	3
Weight decay	$1x10^{-5}$	$1x10^{-5}$
Crop scale	3.5	3.5
Image Size	224	224

5) ENHANCED IMAGE QUALITY ASSESMENT AND PARAMETER SETTING

The primary intention of the IQA is to automatically determine the quality of an image in accordance with a good assessment of human quality. It involves two main purposes,

firstly, it is used to determine the quality of the optical image and to change the setting automatically to gain the best quality. Second, in order to completely optimize the algorithm and achieve an optimized design in preprocessing to offer the best parameter setting, the algorithm can be embedded in image processing. This allows the selection of a relatively good parameter for a specific dataset. To be specific, while generalization of enhancement parameters (D0, radius, amount, and etc.) is not straightforward, we involve over-enhancement measure, LOE, which quantifies the unnaturalness. As a result, the good parameter values are then set to produce the enhanced image (prior to training via deep learning network). It is noteworthy, the idea of incorporating this measurement was recently introduced by Bai and Reibman [49]. Therefore, in our work, we incorporate a well-known image enhancement related IQAs so-called Lightness Order Error (LOE) [46]. The differences in the order of lightness between the original and its enhanced image show the level of naturalness. When the lightness order is high, the LOE score is high, so a high potential for the enhanced image suffers from unnatural enhancement of contrast, and vice versa.

In this work, the LOE score allows us to prove the best approach among tested methods (HEF, UM, and CLAHE). To be clear, the LOE score is utilized in order to find the best parameter setting in each method. Fig. 9 shows sample of an image tested with LOE. Notice that, Fig. 9(b) is a well enhanced contrast image has LOE = 341.09, while Fig. 9(d) is suffering from unnatural contrast enhancement has LOE = 1852.1.

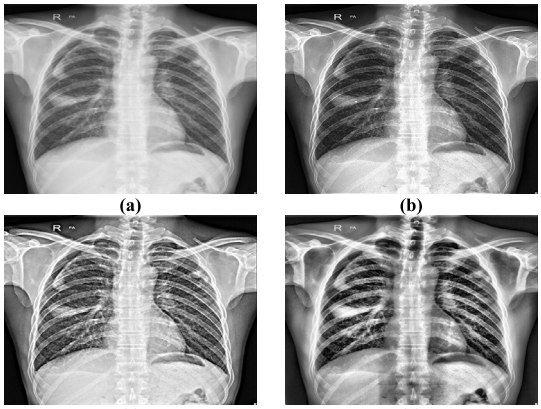


FIGURE 9. An example of enhanced image and its LOE Score. (a) Original Image, (b) HEF (LOE = 341.09), (c) UM (LOE = 837.89), and (d) CLAHE (LOE = 1852.1).

IV. RESULTS AND DISCUSSION

We evaluated the impact of image enhancement by calculating the LOE (Lightness Order Error) [46] score among enhancement methods (the HEF, UM, and CLAHE methods). The analysis shows that HEF can produce the best score in terms of naturalness enhancement. Notice that the enhanced image via CLAHE suffers from unnatural enhancement and led to very poor accuracy during the transfer learning stage (see Fig. 10 and 11).

TABLE 2. Best parameter for HEF, UM and CLAHE.

Method	Parameter	LOE Score
HEF	D0: 70	**341.09**
UM	Radius: 5, Amount: 2	837.89
CLAHE	Window size: 100 Clip Limit: 150	1852.1

The validation accuracy is calculated at the end of the Epoch. It aims to improve generalization, robustness, and prevent overfitting. The data is split into 60% training and 40% validation sets. Training was conducted using categorical cross-entropy as the error function and with the ResNet and EfficientNet batches of 6 and 2 samples, respectively. With the SGD (stochastic gradient descent) optimizer, which only calculates on a small subset or random collection of data examples, we trained for 10 epochs. Besides, the weight decay was used to generalize the model better. To be specific, weight decay will penalize large weights and effectively limit the freedom of our model. In our setting, the weight decay value is set to 0.00001. Since the dataset is small (e.g. in the case of the Shenzhen dataset), the validation analysis is very useful in estimating the model's accuracy. It is because of performance between validation sets that could be significantly changed.

Using Shenzhen CXR images as a dataset, experiments with the EfficientNet-B4, ResNet-50, and ResNet-18 model were performed using HEF, UM, and CLAHE image enhancement algorithms. As discussed by Lopes and Valiati [12], the ResNet model can outperform other models on Shenzhen public datasets. Because of the data variance, certain orders of magnitude are lower, and very large networks are not needed to improve the accuracy [12]. However, in our experiment, the pre-trained EfficientNet-B4 model can obtain more stable performances in enhanced images via HEF and UM methods (see Fig. 10).

When we enhance training images, the image contrast is increased, the disparity between the TB areas was observed and the remaining regions were more noticeable. In addition, the observed edge region and outline were more distinct. As a result, the trained model can extract better features, and contour details of the obtained region. We compared our proposed pipeline with the works from Lopes and Valiati [12], Jaeger *et al.* [35], Hwang *et al.* [27], Rajaraman and Antani [42], and original data without enhancement (baseline). In terms of accuracy (table 3), the proposed pre-trained EfficientNet-B4 with UM outperforms previous works from Lopes and Valiati [12], Jaeger *et al.* [35], and Hwang *et al.* [27] with the result of 89.92. In terms of AUC (table 3), our proposed method attained the competitive result, with an AUC of 94.8. While our work cannot outperform the results proposed by [42], our results still can be very competitive. It is noteworthy that the previous works by Lopes and Valiati [12]

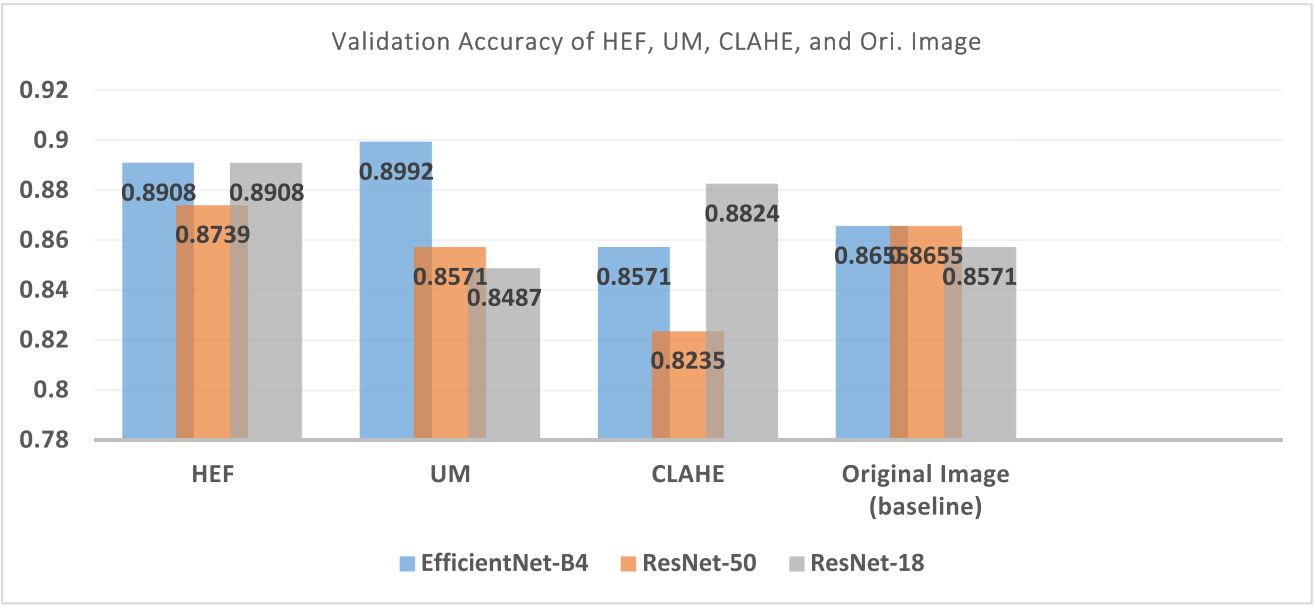


FIGURE 10. Validation accuracy of each image enhancement method through transfer learning.

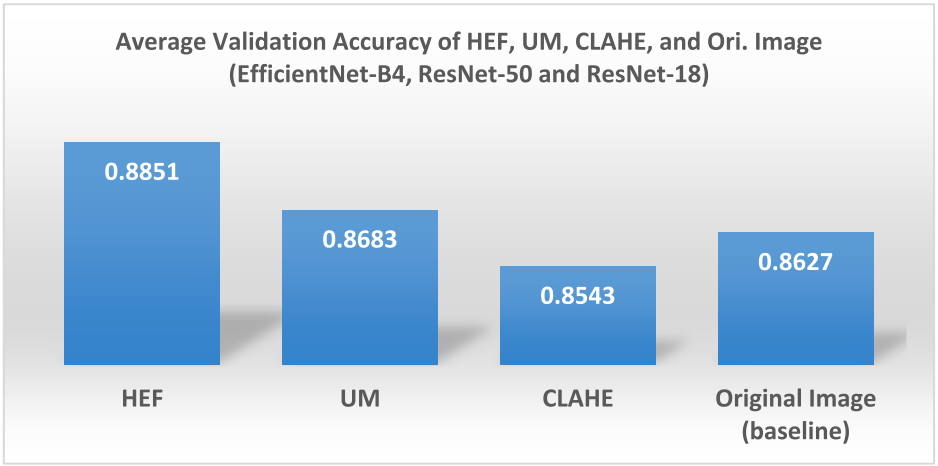


FIGURE 11. Average validation accuracy through transfer learning (pre-trained of EfficientNet-B4, ResNet-50 and ResNet-18 models).

TABLE 3. Accuracy of previous works, proposed EfficientNet-B4 with UM and HEF image.

	Lopes and Valiati [12]	Hwang, et al. [27]	Jaeger, et al. [35]	Rajaraman and Antani [42]	Proposed EfficientNet-B4 with UM	Proposed EfficientNet-B4 with HEF
Parameter	-	-	-	-	Rad: 5, Amo: 2	D_0:70
Shenzhen Datasets	84.7%	83.7%	84%	94.1%	89.92%	89.08%

and Rajaraman and Antani [42] utilize ensemble approaches. Ensemble approaches use multiple DL algorithms to achieve better predictive efficiency than any DL algorithm could have achieved by itself. Computational costs are increased because of the inclusion of the multi-stage deep learning ensemble technique. The use of ensemble approaches is however out of the scope of this work and focuses on evaluating the impact of image enhancement on each model. As visualized in the comparative bar charts (Fig. 10-13), our proposed approach achieved results that are relatively competitive in many cases.

Finally, we analyzed the time of computation. For training and testing, we used a NVIDIA® GeForce GTX 1050 Ti with memory 4GB, it roughly takes 14 minutes nonstop to train the

TABLE 4. Comparison – AUC.

	Lopes and Valiati [12]	Hwang, et al. [27]	Jaeger, et al. [35]	Rajaraman and Antani [42]	Our proposed EfficientNet-B4 with UM	Our proposed EfficientNet-B4 with HEF
Shenzhen Datasets	92.6%	92.6%	90%	99%	94.8%	93.81%

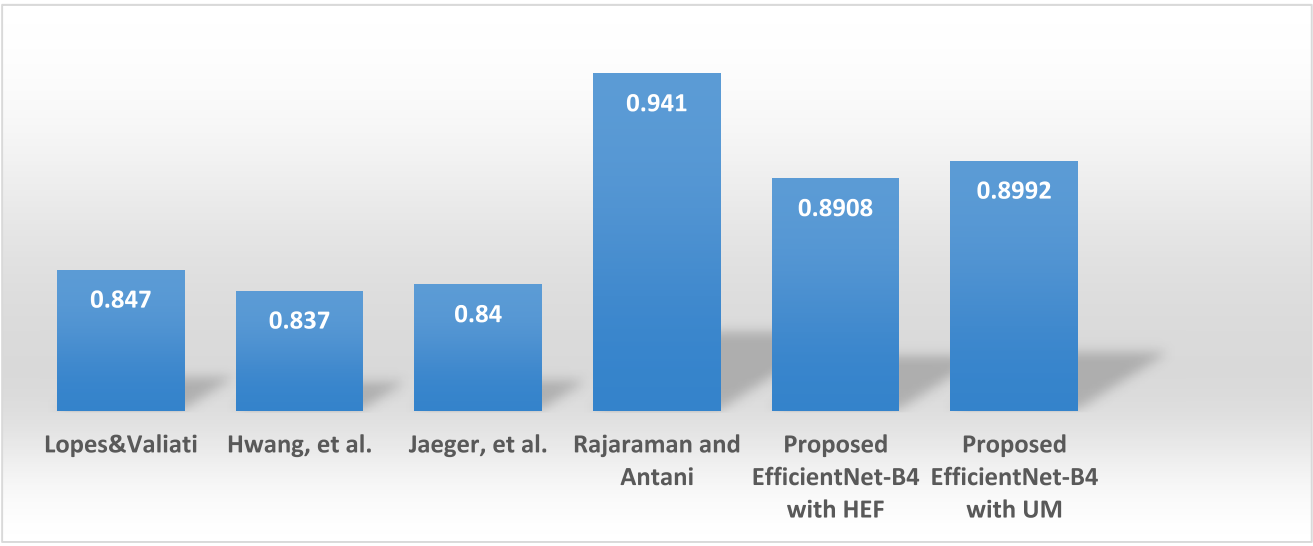


FIGURE 12. Comparison of accuracy.

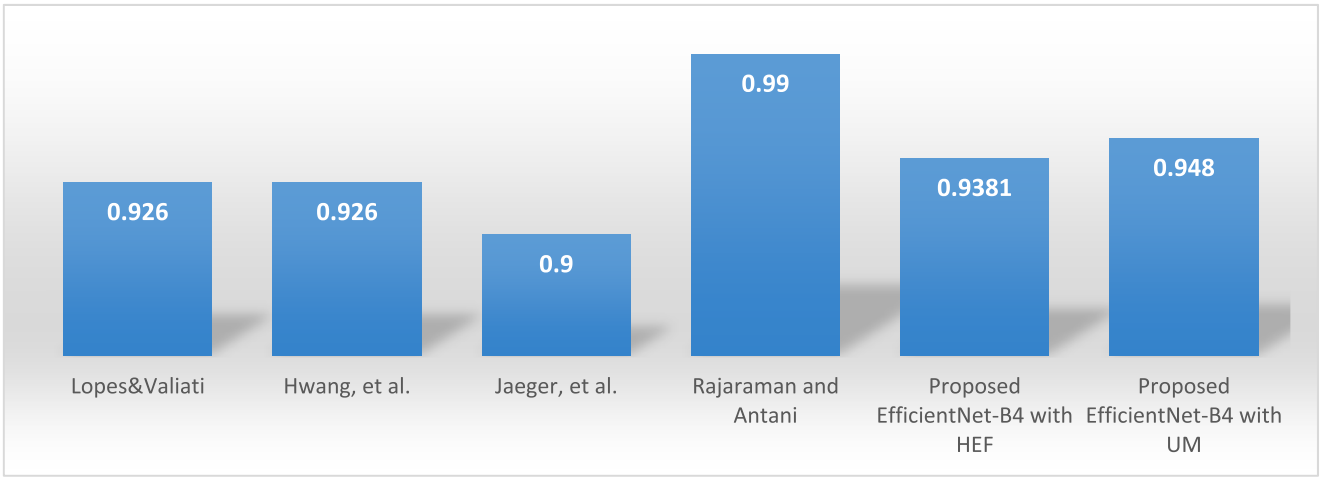


FIGURE 13. Comparison of AUC.

model. The PyTorch 1.2 was employed as a framework. The inference of one CXR image takes less than one minute by using GPU. This research has its limitations. By employing the CNN network, directly feed the original image resolution will require higher GPU and memory quality. In addition, the process would increase the training time. Therefore, the images were down-sampled to 224 × 224 pixels before being fed into the pre-trained networks. Indeed, accuracy may improve with the use of higher resolution images, especially for subtle findings [13].

V. CONCLUSION

Incorporating DL technique with Unsharp Masking (UM) and High-Frequency Emphasis Filtering (HEF) image enhancement, this paper uses EfficientNet-B4, ResNet-50 and ResNet-18 in order to train the TB images and improve the detection accuracy. The experiments showed that the accuracy of the proposed idea is very competitive. Moreover, in terms of the AUC and accuracy, we also thoroughly compared the results with previous works, the proposed idea achieved better results. The use of an image enhancement

system to pre-process the TB images will thus allow the tested pre-trained network to learn better model. Future works will evaluate more image enhancement techniques in order to show a more significant effect of enhancement on DL models. Moreover, a comprehensive subjective judgment and preference from the medical expert will also be analyzed.

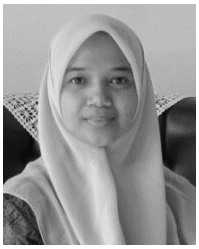
REFERENCES

[1] C. C. Leung, "Reexamining the role of radiography in tuberculosis case finding," *Int. J. Tuberculosis Lung Disease*, vol. 15, no. 10, p. 1279, Oct. 2011.
[2] R. Nawaratne, D. Alahakoon, D. D. Silva, and X. Yu, "Spatiotemporal anomaly detection using deep learning for real-time video surveillance," *IEEE Trans. Ind. Informat.*, vol. 16, no. 1, pp. 393–402 Jan. 2020.
[3] C. Huang, Y. Li, C. L. Chen, and X. Tang, "Deep imbalanced learning for face recognition and attribute prediction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 11, pp. 2781–2794, Nov. 2019.
[4] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker, "Feature transfer learning for face recognition with under-represented data," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 5704–5713.
[5] D. Li, D. Zhao, Q. Zhang, and Y. Chen, "Reinforcement learning and deep learning based lateral control for autonomous driving [Application Notes]," *IEEE Comput. Intell. Mag.*, vol. 14, no. 2, pp. 83–98, May 2019.
[6] N. Nasaruddin, K. Muchtar, and A. Afdhal, "A lightweight moving vehicle classification system through attention-based method and deep learning," *IEEE Access*, vol. 7, pp. 157564–157573, 2019.
[7] M. S. Hossain, M. Al-Hammadi, and G. Muhammad, "Automatic fruit classification using deep learning for industrial applications," *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 1027–1034, Feb. 2019.
[8] S. B. Block, R. D. D. da Silva, L. Dorini, and R. Minetto, "Inspection of imprint defects in stamped metal surfaces using deep learning and tracking," *IEEE Trans. Ind. Electron.*, early access, Apr. 9, 2020, doi: 10.1109/TIE.2020.2984453.
[9] L. Li and X. Feng, "Face anti-spoofing via deep local binary pattern," in *Proc. Deep Learn. Object Detection Recognit.*, 2019, pp. 91–111.
[10] N. R. Varghese and N. R. Gopan, "Performance analysis of automated detection of diabetic retinopathy using machine learning and deep learning techniques," in *Proc. Int. Conf. Innov. Data Commun. Technol. Appl.*, 2020, pp. 156–164.
[11] A. K. Jaiswal, P. Tiwari, S. Kumar, D. Gupta, A. Khanna, and J. J. P. C. Rodrigues, "Identifying pneumonia in chest X-Rays: A deep learning approach," *Measurement*, vol. 145, pp. 511–518, Oct. 2019.
[12] U. K. Lopes and J. F. Valiati, "Pre-trained convolutional neural networks as feature extractors for tuberculosis detection," *Comput. Biol. Med.*, vol. 89, pp. 135–143, Oct. 2017.
[13] 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.
[14] C. Liu, Y. Cao, M. Alcantara, B. Liu, M. Brunette, J. Peinado, and W. Curioso, "TX-CNN: Detecting tuberculosis in chest X-Ray images using convolutional neural network," in *Proc. ICIP*, 2017, pp. 2314–2318.
[15] E. J. Hwang, S. Park, K.-N. Jin, J. I. Kim, S. Y. Choi, J. H. Lee, J. M. Goo, J. Aum, J.-J. Yim, and C. M. Park, "Development and validation of a deep learning-based automatic detection algorithm for active pulmonary tuberculosis on chest radiographs," *Clin. Infectious Diseases*, vol. 69, no. 5, pp. 739–747, Sep. 2019.
[16] K.-M. Koo and E.-Y. Cha, "Image recognition performance enhancements using image normalization," *Hum.-centric Comput. Inf. Sci.*, vol. 7, no. 1, pp. 1–11, Nov. 2017.
[17] A. Polesel, G. Ramponi, and V. J. Mathews, "Image enhancement via adaptive unsharp masking," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 505–510, Mar. 2000.
[18] A. Bundy and L. Wallen, "High-emphasis filtering," in *Catalogue of Artificial Intelligence Tools*. Berlin, Germany: Springer, 1984, p. 47.
[19] A. Wubuli, J. Zhen-Hong, Q. Xi-Zhong, Y. Jie, and N. Kasabov, "Medical image enhancement based on shearlet transform and unsharp masking," *J. Med. Imag. Health Informat.*, vol. 4, no. 5, pp. 814–818, Oct. 2014.
[20] K. Panetta, Y. Zhou, S. Agaian, and H. Jia, "Nonlinear unsharp masking for mammogram enhancement," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 6, pp. 918–928, Nov. 2011.
[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[22] N. Singh and S. Hamde, "Tuberculosis detection using shape and texture features of chest X-Rays," in *Proc. Innov. Electron. Commun. Eng.*, Feb. 2019, pp. 43–50.
[23] 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 Syst. Appl.*, vol. 158, Nov. 2020, Art. no. 113514.
[24] A. Chauhan, D. Chauhan, and C. Rout, "Role of gist and PHOG features in computer-aided diagnosis of tuberculosis without segmentation," *PLoS ONE*, vol. 9, no. 11, Nov. 2014, Art. no. e112980.
[25] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016.
[26] Y. Cao, C. Liu, B. Liu, M. J. Brunette, N. Zhang, T. Sun, P. Zhang, J. Peinado, E. S. Garavito, L. L. Garcia, and W. H. Curioso, "Improving tuberculosis diagnostics using deep learning and mobile health technologies among resource-poor and marginalized communities," in *Proc. IEEE 1st Int. Conf. Connected Health, Appl., Syst. Eng. Technol. (CHASE)*, Jun. 2016.
[27] S. Hwang, H.-E. Kim, and H.-J. Kim, "A novel approach for tuberculosis screening based on deep convolutional neural networks," *Proc. SPIE Med. Imag.*, vol. 9785, Mar. 2016, Art. no. 97852W.
[28] R. Hooda, S. Sofat, S. Kaur, A. Mittal, and F. Meriaudeau, "Deep-learning: A potential method for tuberculosis detection using chest radiography," in *Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA)*, Sep. 2017, pp. 497–502.
[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 84–90.
[30] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
[31] Q. Tian, G. Xie, Y. Wang, and Y. Zhang, "Pedestrian detection based on laplace operator image enhancement algorithm and faster R-CNN," in *Proc. 11th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI)*, Oct. 2018, pp. 1–5.
[32] H. Kuang, L. Chen, F. Gu, J. Chen, L. Chan, and H. Yan, "Combining Region-of-Interest extraction and image enhancement for nighttime vehicle detection," *IEEE Intell. Syst.*, vol. 31, no. 3, pp. 57–65, May 2016.
[33] X. Chen, "Image enhancement effect on the performance of convolutional neural networks," Dept. Comput., Blekinge Inst. Technol., Karlskrona, Sweden, Tech. Rep., 2019.
[34] S. Jaeger, S. Candemir, S. Antani, Y.-X. J. Wáng, P.-X. Lu, and G. Thoma, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," *Quant. Imag. Med. surgery*, vol. 4, no. 6, p. 475, 2014.
[35] S. Jaeger, A. Karargyris, S. Candemir, L. Folio, J. Siegelman, F. Callaghan, Z. Xue, K. Palaniappan, R. K. Singh, S. Antani, G. Thoma, Y.-X. Wang, P.-X. Lu, and C. J. McDonald, "Automatic tuberculosis screening using chest radiographs," *IEEE Trans. Med. Imag.*, vol. 33, no. 2, pp. 233–245, Feb. 2014.
[36] E. Cambau and M. Drancourt, "Steps towards the discovery of mycobacterium tuberculosis by robert koch, 1882," *Clin. Microbiol. Infection*, vol. 20, no. 3, pp. 196–201, Mar. 2014.
[37] W. Zhiqiang and L. Jun, "A review of object detection based on convolutional neural network," in *Proc. 36th Chin. Control Conf. (CCC)*, Jul. 2017, pp. 11104–11109.
[38] S. Kulkarni and S. Jha, "Artificial intelligence, radiology, and tuberculosis: A review," *Academic Radiol.*, vol. 27, no. 1, pp. 71–75, Jan. 2020.
[39] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.
[40] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *J. VLSI signal Process. Syst. signal, image video Technol.*, vol. 38, no. 1, pp. 35–44, 2004.
[41] S. S. Agaian, K. Panetta, and A. M. Grigoryan, "Transform-based image enhancement algorithms with performance measure," *IEEE Trans. Image Process.*, vol. 10, no. 3, pp. 367–382, Mar. 2001.

[42] S. Rajaraman and S. K. Antani, "Modality-specific deep learning model ensembles toward improving TB detection in chest radiographs," *IEEE Access*, vol. 8, pp. 27318–27326, 2020.
[43] T. A. Soomro, A. J. Afifi, A. A. Shah, S. Soomro, G. A. Baloch, L. Zheng, M. Yin, and J. Gao, "Impact of Image Enhancement Technique on CNN Model for Retinal Blood Vessels Segmentation," *IEEE Access*, vol. 7, pp. 158183–158197, 2019.
[44] I. D. Apostolopoulos and T. A. Mpesiana, "Covid 19: Automatic detection from X ray images utilizing transfer," *Phys. Eng. Sci. Med.*, vol. 43, pp. 635–640, 2020.
[45] T. Ahmed Soomro, T. Mahmood Khan, M. A. U. Khan, J. Gao, M. Paul, and L. Zheng, "Impact of ICA-based image enhancement technique on retinal blood vessels segmentation," *IEEE Access*, vol. 6, pp. 3524–3538, 2018.
[46] S. Wang, J. Zheng, H.-M. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3538–3548, Sep. 2013.
[47] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," 2019, *arXiv:1905.11946*. [Online]. Available: http://arxiv.org/abs/1905.11946
[48] M. Zeng, Y. Li, Q. Meng, T. Yang, and J. Liu, "Improving histogram-based image contrast enhancement using gray-level information histogram with application to X-ray images," *Optik*, vol. 123, no. 6, pp. 511–520, Mar. 2012.
[49] C. Bai and A. R. Reibman, "Controllable image illumination enhancement with an over-enhancement measure," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 385–389.
[50] S. Wang, H. Cheng, L. Ying, T. Xiao, Z. Ke, H. Zheng, and D. Liang, "DeepcomplexMRI: Exploiting deep residual network for fast parallel MR imaging with complex convolution," *Magn. Reson. Imag.*, vol. 68, pp. 136–147, May 2020.

KHAIRUL MUNADI (Member, IEEE) received the B.Eng. degree in electrical engineering from the Sepuluh Nopember Institute of Technology, Surabaya, Indonesia, in 1996, and the M.Eng. and Ph.D. degrees in electrical engineering from Tokyo Metropolitan University (TMU), Japan, in 2004 and 2007, respectively. From 1996 to 1999, he was a System Engineer with Alcatel Indonesia. Since 1999, he has been a Lecturer with the Electrical and Computer Engineering Department, Universitas Syiah Kuala (Unsyiah), Banda Aceh, Indonesia, where he has been a Professor, since August 2019. From March 2007 to March 2008, he was a Visiting Researcher in information and communication systems engineering with the Faculty of System Design, TMU. He was a Visiting Scholar with the Department of Computer Engineering, Suleyman Demirel University (SDU), Isparta, Turkey, in 2016. His research interests include multimedia signal processing, knowledge-based management, and disaster management. He is a member of APSIPA.

KAHLIL MUCHTAR (Member, IEEE) received the B.S. degree in information technology from the School for Engineering of PLN's Foundation (STT-PLN), Jakarta, Indonesia, in 2007, the M.S. degree in computer science and information engineering from Asia University, Taichung, Taiwan, in 2012, and the Ph.D. degree in electrical engineering from National Sun Yat-sen University (NSYSU), Kaohsiung, Taiwan. He was appointed as the Chairperson of the Telematics Research Center (TRC), Universitas Syiah Kuala, Banda Aceh, Indonesia. He was an AI Research Scientist with Nodeflux, Jakarta. He is currently a Lecturer and a Researcher with the Department of Electrical and Computer Engineering, Syiah Kuala University. His research interests include computer vision and image processing. He served as a Publication Chair for the IEEE ICELTICs 2017 and 2018. He received the 2014 IEEE GCCE Outstanding Poster Award and IICM Taiwan 2017 The Best of Ph.D. Dissertation Award.



NOVI MAULINA received the M.D. degree from Syiah Kuala University, Aceh, in 2011, and the M.MSc. degree in clinical medicine with Kaohsiung Medical University, Taiwan, in 2015. From 2011 to 2012, she worked as a Medical Staff of an international NGO (The Salvation Army) for Tsunami victims in West Aceh. She joined the Microbiology Department, Syiah Kuala University, as a Lecturer and a Researcher, in 2017. Her thesis topic was about ocular surface inflammation and ultrasound biomicroscopy (UBM) measurement of eyelids from thyroid orbitopathy patients.

BISWAJEET PRADHAN (Senior Member, IEEE) received the Habilitation degree in remote sensing from the Dresden University of Technology, Germany, in 2011. Since 2015, he has been as the Ambassador Scientist with the Alexander Humboldt Foundation, Germany. He is currently the Director of the Centre for Advanced Modeling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT. He is also a Distinguished Professor with the University of Technology Sydney. He is an internationally established Scientist in the fields of geospatial information systems (GIS), remote sensing, image processing, complex modeling/geo-computing, machine learning and soft-computing applications, natural hazards, and environmental modeling. Out of his more than 550 articles, more than 475 have been published in science citation index (SCI/SCIE) technical journals. He has authored eight books and 13 book chapters. He was a recipient of the Alexander von Humboldt Fellowship from Germany and the Alexander von Humboldt Research Fellowship from Germany. He received 55 awards in recognition of his excellence in teaching, service, and research, since 2006, and the World Class Professor by the Ministry of Research, Technology and Higher Education, Indonesia, from 2018 to 2020. He is also an associate editor and an editorial member of more than eight ISI journals. From 2016 to 2020, he was listed as the World's Most Highly Cited Researcher by Clarivate Analytics Report as one of the world's most influential minds.