

An Analysis of Image Enhancement Effects on Convolutional Neural Network-based Pulmonary Tuberculosis Detection

Devito Susilo¹, and Wahyono^{2*}.

¹Undergraduate Program in Computer Science, Universitas Gadjah Mada, Yogyakarta, Indonesia

²Department of Computer Science and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia

Abstract. Pulmonary Tuberculosis (TB) is a primary global infectious disease. Diagnosing TB patients involves medical examination and chest X-ray (CXR) imaging. This CXR image creates an opportunity to utilize machine learning to help physicians and radiologists diagnose TB suspects. Due to the inconsistency of image quality, image enhancement is one of the preprocessing steps to overcome the poor quality of the image. This study examines the effects of several image enhancement techniques, i.e., Histogram Equalization (HE), Contrast Limited Adaptive Histogram Equalization (CLAHE), and Fast Fourier Transform (FFT). These enhanced images are input for a Convolutional Neural Network (CNN). InceptionV3 is a transfer learning architecture with ImageNet as the pre-trained model. The image dataset consists of 3,500 normal and 3,500 tuberculosis CXR images. The best performance, in terms of accuracy and processing time, is achieved by the CLAHE enhancement technique, increasing accuracy by 4.57% compared to the original images as input and a processing time of 5.6 ms faster per testing image. A deeper analysis shows despite FFT achieving high performance, the processing time increases by 14.4 ms compared to the original image processing time. This study concluded that each image enhancement needs to consider the characteristics of the images.

1 Introduction

Pulmonary Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis*, affecting human lung tissue. Infection occurs by inhaling infectious particles from close contact with infected individuals [1]. According to the World Health Organization (2022), pulmonary tuberculosis remains a major global infectious disease, ranking second only to COVID-19 during the COVID-19 pandemic. In 2021, an estimated 10.6 million people worldwide were infected, with a 3.6% increase in cases per 100,000 population compared to 2020. The number of tuberculosis-related deaths also rose from 2019 to 2021, with an estimated 1.6 million deaths recorded by the end of 2021 [2].

Diagnosing suspected TB patients typically involves a medical history examination and a chest X-ray (CXR) image. The results of the CXR image assist physicians in diagnosing various respiratory diseases, including pulmonary tuberculosis. When analyzing chest X-ray images, doctors focus on segmenting images showing infection indications. Infected lungs typically exhibit lesions and grayish-white shadows [3]. However, there are often errors in chest X-ray screening, which has led to utilizing machine learning techniques for image processing, segmentation, analysis, and accurate tuberculosis detection [4]. Machine learning with image analysis has been widely employed to assist physicians and radiologists in enhancing diagnostic

decision-making. Deep learning techniques, such as Convolutional Neural Networks (CNNs), have also been utilized because they can leverage large datasets and extract high-accuracy data features. CNNs have demonstrated good performance and can be employed to aid in tuberculosis diagnosis [5].

This problem leads to inconsistent image data quality that can be influenced by various factors [6]. In neural networks, achieving optimal results necessitates good image data. This realization has led to effectively implementing image enhancement techniques to improve image quality before processing. Image enhancement improves image features, allowing computers to study dynamic feature coverage effectively. It also reduces noise and blurring, hindering model performance [7]. This study aims to explore the impact of image enhancement on tuberculosis detection using a CNN-based model and to analyze the effects of the enhancement on the accuracy and processing time of this CNN-based tuberculosis detection.

2 Related Works

Related research has been conducted on TB detection using chest X-ray image data. In [8], the author studied TB detection using deep machine learning, segmentation, and visualization. The results of this research compared nine CNN architectures. In [9] and [10], the authors of these papers also conducted studies

*Corresponding author : wahyo@ugm.ac.id

using CNN with the latest CNN architectures for TB screening with chest X-ray images. A study was also conducted on using transfer learning in tuberculosis prediction [11].

Regarding previous research on image enhancement, [12] demonstrated that enhancement techniques such as contrast-limited Adaptive Histogram Equalization (CLAHE) and Histogram Equalization (HE) significantly impact the training data in deep learning models, resulting in increased accuracy. In [13], the author conducted research on the effect of image enhancement methods on deep learning models that aimed to evaluate the impact of two methods, namely Unsharp Masking (UM) and High-Frequency Emphasis Filtering (HEF). In [14], the author employed various spatial and frequency techniques for image enhancement in COVID-19 diagnosis models. In [15], the author also conducted research on the utilization of Gray Level Co-occurrence Matrix (GLCM), Discrete Wavelet Transform (DWT), and Local Binary Pattern (LBP) algorithms.

3 Experiments

3.1 Dataset

In this experiment, we use the dataset in [8] that consists of 3,500 normal CXR images and 3,500 tuberculosis CXR images gathered from several CXR image datasets. The original size of the images is 512×512 pixels. In order to accommodate the input requirement of the model, these images are resized into 300×300 pixels. After resizing, these images are split into 75% training, 15% validation, and 10% test data shown in Table 1.

Table 1. Dataset Split

	Class	
	Normal	Tuberculosis
Training (75%)	2,625	2,625
Validation (15%)	525	525
Test (10%)	350	350

3.2 Image Enhancement

In this study, there are three enhancement techniques used, i.e., Histogram Equalization (HE), Contrast Limited Adaptive Histogram Equalization (CLAHE), and Fast Fourier Transform (FFT). HE and CLAHE are used as contrast enhancement, while FFT is expected to perform a sharpness enhancement. These enhancements are used in order to see whether the contrast or sharpness of an image will be suitable for the CNN model. The diagram proposed is shown in Figure 1.

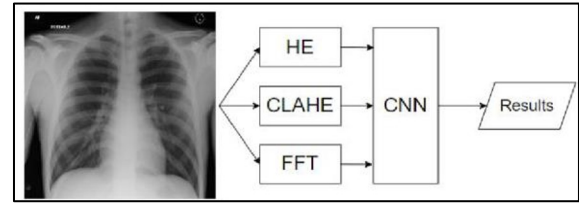


Fig. 1. Proposed Tuberculosis Detection with Image Enhancements

a) *Histogram Equalization (HE)*: Histogram Equalization (HE) is a technique used to enhance the contrast quality of an image by redistributing the intensity values [16]. The working principle of this technique involves transforming the pixel intensity distribution in the image to achieve a more uniform distribution.

b) *Contrast Limited Adaptive Histogram Equalization (CLAHE)*: Contrast Limited Adaptive Histogram Equalization (CLAHE) is an enhancement of the Adaptive Histogram Equalization technique to address excessive contrast enhancement [17]. This method introduces a clip limit to the histogram, representing the maximum height threshold of the histogram [18].

c) *Fast Fourier Transform (FFT)*: Fast Fourier Transform (FFT) is an algorithm used to compute the discrete Fourier transform (DFT) efficiently [19]. It works by transforming the function from the spatial domain to the frequency domain, and the reverse transformation (inverse) is used to implement changes. In image processing applications, FFT can be applied to sharpen images through high-pass filtering, as shown in Figure 2. The high-pass filters used include Ideal High-Pass Filter (IHPF), Butterworth High-Pass Filter (BHPF), and Gaussian High-Pass Filter (GHPF).

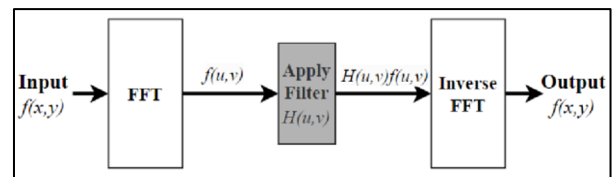


Fig. 2. Filtering in Fast Fourier Transform

3.3 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is a specialization of neural networks designed to process data with grid-like topologies, such as 2D-pixel grids found in image data [20]. This type of neural network is based on mathematical convolution operations. InceptionV3 architecture developed by Google [21] is utilized for this study. The usage of this specific CNN due to its transfer learning capability with ImageNet as a pre-trained model is shown in Figure 3. The input images for this architecture are 300×300 pixels and normalized.

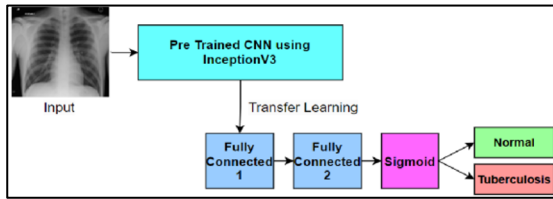


Fig. 3. Proposed CNN for Tuberculosis Detection

4 Results

This study uses three types of images as CNN's input image, i.e., original, HE, CLAHE, and FFT-enhanced images. These images are fed into the CNN architecture with transfer learning from ImageNet as default for the InceptionV3 model. The performance of these images is measured using the model's accuracy and loss. The accuracy is measured using the formula as follows:

$$Accuracy = \frac{TP+FN}{All\ Data} \quad (1)$$

where TP is True Positive, and FN is False Negative.

Regarding the hyperparameter, this study conducted a manual hyperparameter tuning. The tuning process regards the suitable hyperparameter with balanced accuracy, loss, and processing time per image using original images in milliseconds (ms). The parameter used for the InceptionV3 model is shown in Table 2.

The focus of this study is to explore the effects of each enhancement in CNN-based pulmonary tuberculosis detection. In this tuned hyperparameter, the original image has a 94,98% training accuracy and 94,14% testing accuracy. The processing time for each image in training is 69.3 ms, and testing is 74.2 ms.

4.1 Contrast Enhancement

In the contrast enhancement scenario, HE and CLAHE are performed to the images and fed into the CNN as input. In Figure 4, it is observed that CLAHE gives more clarity than HE. In CLAHE, the images are enhanced with clip limit 4.0 and tile grid size 16×16 as default.

Table 2. Tuned Hyperparameter

Hyperparameter	Value
Model	2 Fully Connected Layer (64 Units with 0.2 Dropout Rate)
Learning Rate	0,0001
Optimizers	Adam
Batch Size	32
Epochs	15

HE achieved high performance with 97.12% training accuracy and 96.29% testing accuracy. An increasing performance was also shown by CLAHE, with 98.66% training accuracy and 98.71% testing accuracy. This indicates that CLAHE-enhanced images are better for

the CNN architecture to learn the feature of each image in comparison to HE-enhanced, as shown in Figure 5.

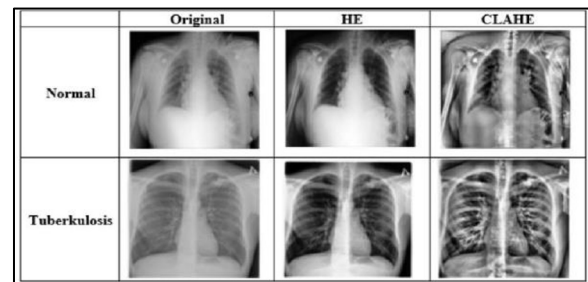


Fig. 4. Contrast-Enhanced CXR Images

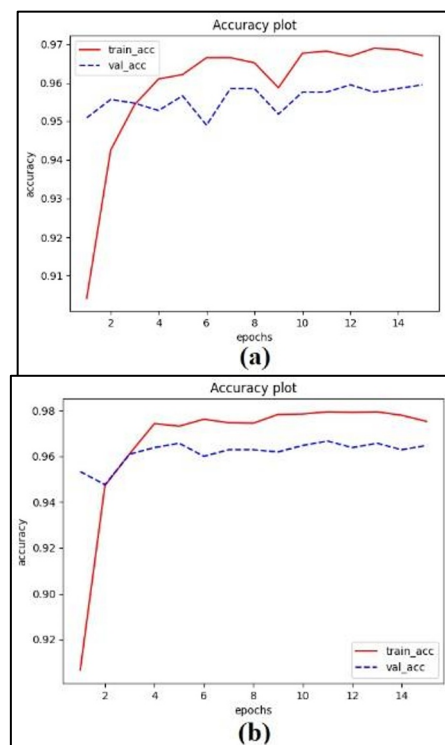


Fig. 5. (a) HE Accuracy Plot **(b)** CLAHE Accuracy Plot

Furthermore, the processing time for HE and CLAHE shows a consistent performance. The processing time for HE is 69.3 ms on training per image and 80 ms on testing per image. Meanwhile, CLAHE achieves faster processing time with 67.8 ms on training per image and 68.6 ms. It can be observed that CLAHE performs better with higher accuracy and faster processing time.

4.2 Sharpness Enhancement

For the sharpness enhancement, this study performs the Fast Fourier Transform (FFT) with three high-pass filters, i.e., Ideal High-Pass Filter (IHPF), Butterworth High-Pass Filter (BHPF), and Gaussian High-Pass Filter (GHPF). These filters enhance the images in the frequency domain, as can be seen in Figure 6.

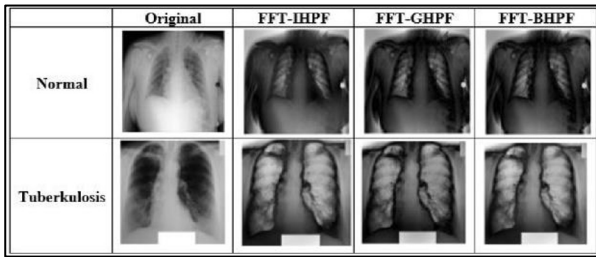


Fig. 6. FFT Enhanced CXR Images

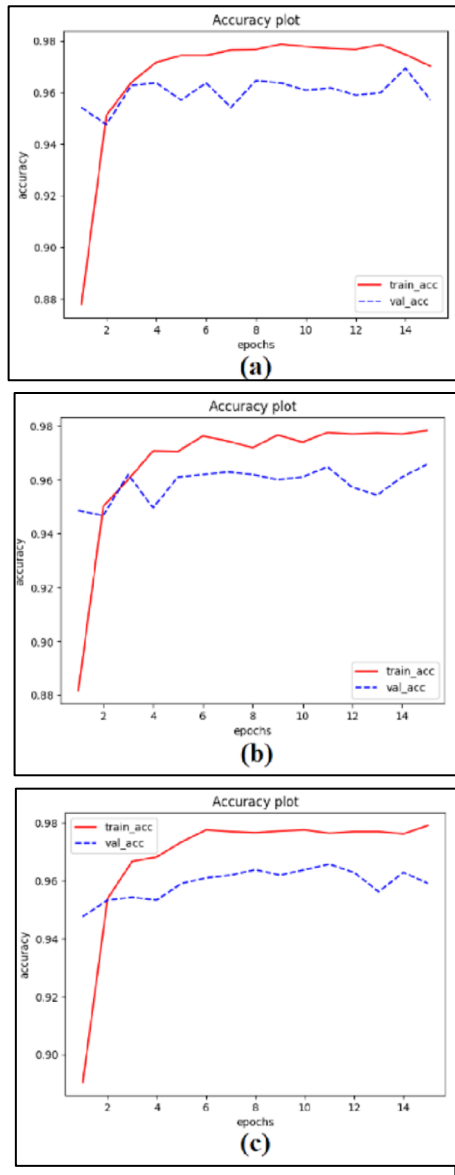


Fig. 7. (a) FFT-IHPF (b) FFT-GHPF (c) FFT-BHPF

Based on the observation of several samples of these enhanced CXR images in Figure 6, the enhancement result is generally similar across three high-pass filters performed. Despite this observation, the performance of CNN with the input of these enhanced images is different, as shown in Fig 5. FFT-IHPF achieves 97.24% training accuracy and 98.29% testing accuracy. FFT-GHPF achieves 98.43% training accuracy and 98.14% testing accuracy. Meanwhile, FFT-BHPF achieves 98.71% training accuracy and 97.86% testing accuracy. Based on this result, it shows that each high pass filter

on FFT fluctuates the performance of CNN of each enhanced image. The accuracy plot shown in Figure 7 also indicates how distinguished the CNN is in learning the features of images.

Additionally, it is found that the processing time is varied for each high-pass filter. The processing time for FFT-IHPF is 85.3 ms on training per image and 88.6 ms on testing per image. FFT-GHPF uses 87.3 ms on training per image and 91.4 ms on testing per image. Lastly, FFT-BHPF uses 87.7 ms on training per image and 87.3 ms on testing per image. This increasing time is due to noise or artifacts that FFT enhancement results in, potentially making it harder for CNN to extract the feature. This result shows that FFT-IHPF enhanced images perform better in the CNN with higher accuracy and faster processing time.

5 Conclusions

This paper implements three image enhancement methods: HE, CLAHE, and FFT. These enhanced images are used as input on InceptionV3 architecture with transfer learning using ImageNet. The best performance, in terms of testing accuracy and processing time, is achieved using CLAHE enhancement. This technique increases the testing accuracy by 4.57% in comparison to the original images as input. Similar to accuracy, the processing time is also 5.6 ms faster per testing image. The comparison of the performance of each enhanced image is shown in Table 3.

In regards to the accuracy of FFT techniques, despite it showing a high accuracy result, the processing time is increased. The highest performance, in terms of testing accuracy, is shown by FFT-IHPF, with an increase of 4.15% in comparison to the original images as input.

Table 3. Overall Performance

Techniques	Training Accuracy	Training Time per Image (ms)	Testing Accuracy	Testing Time per Image (ms)
Original	94.98%	69.3	94.14%	74.2
HE	97.12%	69.3	96.29%	80
CLAHE	98.66%	67.8	98.71%	68.6
FFT-IHPF	97.24%	85.3	98.29%	88.6
FFT-GHPF	98.43%	87.3	98.14%	91.4
FFT-BHPF	98.71%	87.7	97.86%	87.3

Despite this significant increase, the processing time per testing image also increased by 14.4 ms compared to the original testing image processing time. A deeper examination through three FFT techniques shows that despite different high-pass filters being performed, the results of CNN training can differ. This indicates that each filter generates enhanced images with other characteristics.

This study concludes that performing image enhancement can lead to several effects. On this CNN-based pulmonary detection, contrast enhancement is superior in accuracy and processing time, with CLAHE being more suitable. Compared with sharpness enhancement, although FFT achieves a high precision, the processing time is increased. Further study is advisable to study other CNN architecture and add classes with a suitable enhancement.

The Department of Computer Science and Electronics, the Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Indonesia, funded the dissemination and publication of this research.

References

1. W. Cruz-Knight and L. Blake-Gumbs, "Tuberculosis: An Overview," *Primary Care: Clinics in Office Practice*, vol. 40, no. 3, pp. 743–756, 2013, doi: <https://doi.org/10.1016/j.pop.2013.06.003>.
2. WHO, *Global Tuberculosis Report 2022*. Geneva: World Health Organization, 2022. [Online]. Available: <https://www.who.int/publications/i/item/9789240061729>
3. J. P. Zellweger, G. Sotgiu, M. Corradi, and P. Durando, "The diagnosis of latent tuberculosis infection (LTBI): currently available tests, future developments, and perspectives to eliminate tuberculosis (TB): The diagnosis of latent tuberculosis infection (LTBI)," *La Medicina del Lavoro | Work, Environment and Health*, vol. 111, no. 3, pp. 170–183, 2020.
4. Q. H. Nguyen et al., "Deep Learning Models for Tuberculosis Detection from Chest X-ray Images," in *2019 26th International Conference on Telecommunications (ICT)*, 2019, pp. 381–385. doi: 10.1109/ICT.2019.8798798.
5. P. Chhikara, P. Singh, P. Gupta, and T. Bhatia, "Deep Convolutional Neural Network with Transfer Learning for Detecting Pneumonia on Chest X-Rays," in *Advances in Bioinformatics, Multimedia, and Electronics Circuits and Signals*, 2020, pp. 155–168.
6. A. P. Brady, "Error and discrepancy in radiology: inevitable or avoidable?" *Insights into Imaging*, vol. 8, 2017, doi: 10.1007/s13244-016-0534-1.
7. X. Chen, "Image enhancement effect on the performance of convolutional neural networks," *Dissertation*, p. 40, 2019. Available: <https://urn.kb.se/resolve?urn=urn:nbn:se:bth-18523>
8. T. Rahman et al., "Reliable Tuberculosis Detection Using Chest X-Ray With Deep Learning, Segmentation and Visualization," *IEEE Access*, vol. 8, pp. 191586–191601, 2020, doi: 10.1109/ACCESS.2020.3031384.
9. D. Das, K. C. Santosh, and U. Pal, "Inception-based Deep Learning Architecture for Tuberculosis Screening using Chest X-rays," in *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021, pp. 3612–3619. doi: 10.1109/ICPR48806.2021.9412748.
10. O. Faruk et al., "Novel and Robust Approach to Detect Tuberculosis Using Transfer Learning," *J Healthc Eng*, 2021, doi: 10.1155/2021/1002799.
11. M. H. Ishtiaq et al., "Deep learning to predict Pulmonary Tuberculosis from Chest Poster Anterior Radiographs of Lungs," in *2022 14th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*, 2022, pp. 1–5. doi: 10.1109/MACS56771.2022.10022769.
12. Wahyono, "Analysis of the Effect of Image Enhancement on X-Ray Image-Based Detection of COVID-19," *Jurnal Teknologi Informasi*, vol. 22, no. 1, pp. 186–194, 2023, [in Indonesian] doi: 10.33633/tc.v22i1.7195.
13. K. Munadi, K. Muchtar, N. Maulina, and B. Pradhan, "Image Enhancement for Tuberculosis Detection Using Deep Learning," *IEEE Access*, vol. 8, pp. 217897–217907, 2020, doi: 10.1109/ACCESS.2020.3041867.
14. A. K. Khuzani, M. Heidari, and S. A. Shariati, "COVID-Classifer: an automated machine learning model to assist in the diagnosis of COVID-19 infection in chest X-ray images," in *Sci Rep*, 2021, vol. 11, no. 9887. doi: 10.1038/s41598-021-88807-2.
15. S. M. Fati, E. M. Senan, and N. ElHakim, "Deep and Hybrid Learning Technique for Early Detection of Tuberculosis Based on X-ray Images Using Feature Fusion," *Applied Sciences*, vol. 12, no. 14, 2022, doi: 10.3390/app12147092.
16. R. C. Gonzalez and R. E. Woods, *Digital Image Processing (4th Edition)*. Prentice-Hall, Inc., 2014.
17. K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization," in *Graphics Gems IV*, USA: Academic Press Professional, Inc., 1994, pp. 474–485.
18. A. M. Reza, "Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement," *The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology*, vol. 38, 2004, doi: 10.1023/B:VLSI.0000028532.53893.82.
19. R. N. Bracewell, *The Fourier Transform and Its Applications 3rd Edition*. New York: McGraw-Hill, 1999.
20. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. The MIT Press, 2016. [Online]. Available: <https://mitpress.mit.edu/9780262035613/deep-learning/>
21. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision." *arXiv*, 2015. doi: 10.48550/ARXIV.1512.00567