

Automated Tuberculosis Detection from Chest X-Ray Images Using Traditional Digital Image Processing Techniques

Anders Emmanuel Tan

*Department of Computer Science and Electronics
Gadjah Mada University
Yogyakarta, Indonesia
andersemmuaneltan@mail.ugm.ac.id*

Indratanaya Budiman

*Department of Computer Science and Electronics
Gadjah Mada University
Yogyakarta, Indonesia
indratanaya.b@mail.ugm.ac.id*

Evan Razzan Adytaputra

*Department of Computer Science and Electronics
Gadjah Mada University
Yogyakarta, Indonesia
evanrazzanadytaputra2006@mail.ugm.ac.id*

Daffa Maulana Siddiq

*Department of Computer Science and Electronics
Gadjah Mada University
Yogyakarta, Indonesia
main@dims.slmall.me*

Abstract—Tuberculosis (TBC) has become and remains a significant global challenge for thousands of years, causing deaths worldwide, particularly in countries not yet developed. This disease spreads through air, meaning that it is an airborne disease. It spreads through active TBC carrier by coughs or sneeze that releases water droplets containing the bacteria. Its infectious nature and the fatal potential that it brought if not treated immediately means that fast and accurate early detection becomes the key to decrease mortality rate related to TBC [1], [2]. Chest X-ray (CXR) imaging is a widely used method to screen for pulmonary TBC, but it requires manual inference from experts (i.e. doctors) to observe and determine TBC Positivity. And as proven by research, this faces challenges of subjectiveness, variable observer, and time consumption [5], [6]. In addition, TBC radiological patterns often include other lung diseases, and with it potential misclassification [7]. This paper presents an automated TBC detection system utilising traditional digital image processing (DIP) techniques following the AESFERM (Acquisition, Enhancement, Segmentation, Feature Extraction, Representation, Matching) framework. Our approach to this uses Gaussian smoothing, histogram equalisation, and Laplacian sharpening for image enhancement, Otsu thresholding and morphological operations for segmenting the lung, and feature extraction utilising Local Binary Patterns (LBP), Gray-Level-Co-occurrence Matrix (GLCM), and Histogram of Oriented Gradients (HOG). The system returned promising results in distinguishing between normal and chest X-rays affected by TBC, giving an alternative that is less computationally expensive compared to deep learning methods, which gives an advantage for edge devices and resource constraint.

Index Terms—Tuberculosis Detection, Digital Image Processing, Chest X-Ray Analysis, Computer-Aided Diagnosis, Medical Image Processing

I. INTRODUCTION

Tuberculosis (TBC) is a serious infectious disease caused by *Mycobacterium tuberculosis*. It has become one of the most significant global health problems for thousands of years [1],

[2]. TBC remains a leading cause of deaths worldwide, with fast and accurate early detection and treatment crucial for controlling and reduce disease spread and reducing mortality rates [1].

One of the most used method in detection or screening pulmonary TBC is Chest X-ray (CXR) [3], [4]. CXR check allows lung conditions visualisation, helping in identifying earlier signs of TBC infection. In clinical practice, the interpretation of CXR image is usually done by medical personnel or experienced radiologist, which brings its own sets of limitations. These include, but not limited to subjective interpretations, a lot of time, and possibilities for errors or inter-observer opinion variability [5], [6].

There is more to this challenge. Unfortunately, CXR images of TBC patients often show radiological patterns similar to other lung diseases such as pneumonia or lung cancer. And relying on manual interpretation may frequently lead to misclassification for the reasons earlier. [7]. This situation can have serious consequences as misdiagnosis not only leads to inappropriate therapy but may also worsen health conditions and increase disease transmission risk (do remember that TBC is an airborne disease). The problem is further amplified by the shortage of radiologists, particularly in low-resource developing countries that are most affected by the high prevalence of TBC.

However, with technological advancement, there is hope. In recent years, Computer-Aided Diagnosis (CAD) systems have emerged as promising solutions for assisting medical personnel and radiologist to aid TBC detection based on CXR images. Research by [8] demonstrates that AI-based systems can achieve TBC detection accuracy of 85%, significantly higher than radiologists without AI assistance (62%). This brings good news, as the research points out a possibly

better ways to detect TBC while simultaneously reducing costs for training new radiologists and subject to less errors. Additionally, several studies have explored various image processing techniques for TBC detection, including hybrid feature approaches [10] and chest X-ray enhancement methods [11].

Proving to be substantially better and safer than manual interpretation, deep learning approaches have shown impressive results. And with it, substantial computational resources and large annotated datasets. And it takes a lot of time, money, and work to improve deep learning models and set everything up for the hospital.

What sets this paper apart is that it explores an alternative approach using traditional digital image processing techniques that offer interpretability, computational efficiency, and effectiveness even with limited data. Our work builds upon existing research in feature representation learning [12] and feature detection methods [14], but focuses specifically on a comprehensive pipeline using well-established image processing algorithms. And of course, the method does not need as much computational power required by deep learning models.

The main contributions of this work include:

- A complete TBC detection pipeline using traditional digital image processing techniques
- Implementation of the AESFERM framework specifically tailored for TBC detection in chest X-rays
- Comprehensive feature extraction combining texture, statistical, and shape information
- Evaluation of the approach on a publicly available TBC chest X-ray dataset

II. PROPOSED METHODOLOGY

The primary constraint that applies in the making of this paper is to stick to traditional digital image processing methods and should not utilise machine learning. The limitation forces the usage of traditional machine learning, and the TBC detection system follows the comprehensive AESFERM framework (Acquisition, Enhancement, Segmentation, Feature Extraction, Representation, Matching). This structured approach ensures systematic processing of chest X-ray images from initial acquisition through final classification. And hopefully, resulting in accurate result that can be used to help medical personnel to better identify early stages of TBC and provide treatment quickly.

A. Image Acquisition

Image Acquisition. The stage that starts them all. This stage is the initialisation and involves acquiring and standardising chest X-ray images for consistent processing. For this, we utilise the publicly available Tuberculosis Chest X-ray Database from Kaggle [9]. In addition to that, we also use grayscale conversion which simplifies processing while preserving the essential morphological information needed for TBC detection, as color information is redundant in X-ray analysis, due to its grayscale nature.

Before processing the images, each of them undergoes grayscale conversion to simplify processing while preserving essential morphological information. This process of standardisation ensures that all input images have consistent dimensions and color space, thus forming a reliable foundation for the next processing stages.

B. Image Enhancement

Having acquired the image, the next step is to perform image enhancement. The Image enhancement stage aims to improve visual quality and highlight features that are relevant to TBC detection. Our enhancement pipeline uses three sequential operations:

1) *Gaussian Smoothing:* We apply Gaussian smoothing with a 5×5 kernel to reduce random noise inherent in X-ray acquisition:

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

where $\sigma = 0$ for uniform smoothing. This operation makes sure that important lung structures are preserved all while minimizing acquisition artifacts that are present in the picture.

We choose this method, as gaussian smoothing effectively reduces high-frequency noise without significantly blurring important edges and structures. This is really great because X-ray images often contain quantum noise from the acquisition process (the X-ray shot) that can interfere with accurate segmentation and feature extraction. The 5×5 kernel size provides optimal balance between noise reduction and detail preservation for 512×512 images.

2) *Laplacian Sharpening:* Lung CXR often includes parts that are not used for the machine learning model. We want to focus on the lung part, since it is where the TBC virus lives. Edge detection is used to identify the edges (perimeter) of the lung and feed it for the next stage. Edge enhancement is achieved through Laplacian sharpening:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (2)$$

The sharpening operation emphasizes boundaries and fine details characteristic of TBC manifestations, such as cavity walls and lesion contours. This will aid in the segmentation process later on.

Laplacian sharpening specifically enhances edges and fine details that are critical for identifying TBC-infected lung signs like cavitations, infiltrates, and pleural effusions. These features often have subtle intensity transitions that need enhancement for reliable detection. The method works by amplifying high-frequency components where these important diagnostic features reside.

3) *Histogram Equalization:* In addition to the two methods above, we also utilise histogram equalisation, which alters contrast. Global contrast improvement is performed by re-distributing intensity values to span the full dynamic range (equalisation). This is enhancement particularly beneficial to visualise the subtle TBC signs like small infiltrates and early cavitation.

You see, Chest X-rays often suffer from poor contrast due to variations in exposure parameters and patient anatomy. Histogram equalization maximizes the contrast throughout the image, making subtle TBC findings more apparent. This is particularly important for early-stage TBC where pathological changes may be minimal and poorly contrasted against normal lung tissue.

C. Image Segmentation

After acquiring the image and enhance it, we need to segment the lung away from the rest of the pictures, because we only need the lung part to analyse TBC positiveness. For that reason, accurate lung segmentation is crucial for precisely isolating the region of interest. Our segmentation approach combines multiple techniques:

1) *Otsu Thresholding*: This is an advanced thresholding method. Otsu's method is great for automatic optimal threshold calculation, and effective at separating lung tissue from background by minimizing intra-class intensity variance. The resulting binary mask is inverted to position lungs as the foreground, thus the region of interest. This method is amazing. Otsu's method automatically determines the optimal threshold without manual intervention, making it suitable for automated systems. It works particularly well with chest X-rays because there is typically good separation between the dark lung fields and brighter surrounding tissues. The bimodal intensity distribution problem commonly found in chest radiographs is handled effectively.

2) *Morphological Processing*: Morphological operations like dilation and erosion functions to better process the image. The dilation effectively "fills the hole" that are present in the image, while erosion "erodes" the image. Combining it in different order would create opening and closing methods, which are effective to remove noise. Morphological operations refine the initial segmentation:

- **Opening** (13×13 elliptical kernel): Removes small noise artifacts and separates connected components
- **Closing** (19×19 elliptical kernel): Fills holes in lung parenchyma and smooths boundaries

The elliptical kernel shape approximates the natural curvature of lung boundaries better than rectangular kernels, as lung have a shape that somewhat resembles elips. Rectangular kernels are not the best to use here, as using it will cause a lot of excess area that are not of interest. We use opening to remove small noise particles and separate lungs from other thoracic structures, while closing fills small gaps within the lung fields that may result from blood vessels or pathology. The specific kernel sizes (13×13 and 19×19) were the optimal performance values for 512×512 pixels chest X-rays images.

3) *Connected Component Analysis*: We identify lung regions through connected component analysis with spatial constraints:

$$\text{Lung candidates} = \{C_i | w \times 0.2 < x_{\text{centroid}} < w \times 0.8\} \quad (3)$$

where only components near the central chest region are considered, and the two largest valid components are selected as left and right lungs.

The spatial constraint (centroid between 20% and 80% of image width) ensures that only anatomically plausible lung regions are selected, effectively eliminating artifacts in the region of interest. Selecting the two largest components corresponds to the left and right lungs, which in CXR images are typically the largest dark regions in a chest X-ray. This approach is more robust to variations in patient positioning and image quality.

D. Feature Extraction and Representation

After getting through acquisition, enhancement, and segmentation, the image will be extracted to obtain comprehensive features from the segmented lung regions using three more methods that were inspired by successfull medical image analysis implementations:

1) *Local Binary Patterns (LBP)*: LBP is a method that captures local texture patterns using uniform patterns with $P = 8$ and $R = 1$:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) \cdot 2^p \quad (4)$$

The resulting 59-bins histogram from the process provides rotation-invariant texture descriptors.

The reason why we use LBP is that it effectively captures micro-textural patterns that are characteristic of TBC manifestations such as granulomas, cavitations, and fibrotic changes. The uniform pattern variant ($P=8$, $R=1$) provides a good compromise between discriminative power and feature dimensionality. And the greath thing about LBP is that it is rotation invariance. And that is particularly valuable since TBC patterns can appear at various orientations.

2) *Gray-Level Co-occurrence Matrix (GLCM)*: GLCM extracts statistical texture features across multiple scales and orientations:

- **Distances**: [1, 2, 3] pixels for multi-scale analysis
- **Angles**: $[0, \pi/4, \pi/2, 3\pi/4]$ for orientation invariance
- **Features**: Contrast, correlation, energy, homogeneity, ASM, variance, dissimilarity, entropy

GLCM provides statistical measures of texture that complement LBP's local patterns. The multiple distances capture texture information at different scales. And that is very important because TBC manifestations vary in size from small nodules to large consolidations. The four angles ensure rotation invariance, while the eight Haralick features provide comprehensive characterization of tissue properties affected by TBC, such as heterogeneity (entropy), uniformity (energy), and local variations (contrast).

3) *Histogram of Oriented Gradients (HOG)*: Another feature extraction method that we use is HOG. The HOG method captures shape and edge information with the following parameters:

- **Orientations**: 9 bins

- **Pixels per cell:** 8×8
- **Cells per block:** 2×2 with L2-Hys normalization

HOG effectively captures the structural information and edge patterns that are crucial for identifying TBC-related anatomical changes, further increasing accuracy after LBP and GLCM. The 8×8 pixel cells provide sufficient spatial resolution to detect local shape variations, all while the 2×2 cell blocks with L2-Hys normalization provide illumination invariance. The 9 orientation bins offer good angular resolution for capturing directional features of TBC manifestations such as linear fibrotic streaks or irregular cavity boundaries.

E. Machine Learning

after extracting the necessary features and labels, the first 80% of the extracted features and labels were used for training while the rest of the data was used for testing the accuracy of the model. The model selected by the researcher is Logistic Regression which is a supervised machine learning algorithm that finds an optimal sigmoid function that will be used to determine the probability of a data point falling into a certain binary class, given the value of certain attributes. This algorithm best fits the case since the model would only need to predict whether the selected x-ray image for the input shows a healthy lung or a lung with TBC based on certain attributes such as the ASM, variance, and dissimilarity value of the received input.

III. EXPERIMENTAL SETUP AND PRELIMINARY RESULTS

A. Dataset and Evaluation Metrics

The Kaggle dataset used for the training section was the public Kaggle dataset containing 4,200 chest X-ray images (3,500 normal cases, 700 TBC cases) in PNG format with 512×512 pixel resolution provided by Tawsifur Rahman [9]. This Kaggle dataset provides a substantial number of clinically validated cases with balanced distribution between normal and TBC classes. The fixed 512×512 resolution ensures consistency in the processing and eliminates variations that occurs due to different image dimensions. For the testing set, we borrowed 100 healthy lung and 100 TBC lung images from the Kaggle page provided by Muhammad Rehan [15]. Below are the metrics we used to evaluate the performance of our model:

1) *Accuracy:* The accuracy tells us how many results the model predicted correctly from the 20% test data provided, it is calculated by dividing the sum of the total true positives and true negatives with the total amount of data in the test split.

2) *confusion matrix:* The confusion matrix is a 2×2 table that shows how many predictions fall into the true positives, false positives, true negatives, and false negatives category. This gives valuable information about the behavior of the model, where it most often makes mistakes, and in which areas did the model not learn properly.

3) *Precision:* Precision gives information about: out of all the samples that the model predicted as positive, how many were actually positive. it is calculated using the formula:

$$\frac{\text{True positives}}{\text{True positives} + \text{False positives}} \quad (5)$$

4) *F1-Score:* The F1-score is a metric that provides a single measure of a model's performance by calculating the harmonic mean of its precision and recall. It offers a balanced view of the trade-off between false positives and false negatives. It is calculated using the formula:

$$2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (6)$$

5) *Sensitivity and specificity:* Sensitivity calculates out of all the positive cases how many did the model actually detect, a low sensitivity means that the model will not be able to diagnose or identify TBC patients correctly. Sensitivity is calculated using the formula:

$$\frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (7)$$

Meanwhile specificity determines how well the model is able to identify healthy patients, low specificity means that the model would often misdiagnose healthy lungs as TBC patients. Specificity is calculated by:

$$\frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}} \quad (8)$$

6) *ROC curve and AUC:* The Receiver Operating Characteristic (ROC) curve measures how well the model is able to separate the 2 classes (TBC and normal), the x-axis of the ROC represents the ratio of false positives and the amount of TBC lungs in the dataset, the y-axis of the ROC represents the ratio of true positives and the amount of healthy lungs in the dataset. The area under curve (AUC) gives the probability that a randomly chosen positive example gets a higher score than a randomly chosen negative example, an AUC score of 0.5 means that the model has just been randomly guessing instead of making calculated decisions based on data, a high AUC score (0.8) would mean that the model has been learning well.

B. Preliminary Results

The LBP algorithm will be used to capture textures such as the spots or lesions of the lungs by comparing each pixel to its neighbors and encoding micro-patterns, the GLCM algorithm measures how often pairs of pixel intensities occur at certain distances and angles, the HOG feature representation algorithm will be used to detect and extract edges and also gradient strength patterns to outline the structural shape and contour of the lungs. To find the combination that gives the most optimal output, the researcher has tuned various combinations of the feature representation algorithms.

The results from table 1 show that combining multiple feature extraction methods significantly improves the model's ability to distinguish between normal and tuberculosis chest

Method	Accuracy	Precision	F1-score	Sensitivity	Specificity
HOG + LBP	93.33%	0.8189	0.7879	0.7591	0.9673
GLCM + LBP	95.59%	0.8906	0.8604	0.8321	0.9801
GLCM + HOG	95.23%	0.8760	0.8496	0.8248	0.9772
HOG	93.33%	0.8240	0.7863	0.7518	0.9687
LBP	83.69%	0	0	0	1
GLCM	95.23%	0.8760	0.8496	0.8248	0.9772
GLCM + HOG + LBP	96.9%	0.9237	0.9030	0.8832	0.9858

Table I

COMPARISON OF DIFFERENT COMBINATIONS OF FEATURE REPRESENTATION

X-ray images. Individually, HOG and GLCM achieve solid performance (93.33% and 95.23%), while LBP alone gives poor accuracy as the results show that it seems to be incapable of detecting normal lungs. Pairing GLCM with HOG and HOG with LBP does not seem to improve GLCM or HOG alone at all, meanwhile pairing GLCM with LBP slightly improves the performance of GLCM. However, when all methods are combined together, the weaknesses of each algorithm are covered and this results in the most optimal performance with an accuracy of 96.9%, highest true positives (121) and lowest false negatives (16).

C. Performance Analysis

By using the pipeline that gave us the highest accuracy (LBP + GLCM + HOG), the model managed to achieve the following results:

Metrics	results
Accuracy	96.9%
True positives	694
False positives	9
False negatives	18
True negatives	119
Precision	0.9237
F1-score	0.9030
Specificity	0.9858
Sensitivity	0.8832
AUC Score	0.989
Training accuracy	99.9%
Validation accuracy	96.3%
Correct predictions on test set	393/400
Confidence average	0.97
False negatives (test set)	5
False positives (test set)	2

Table II

COMPARISON OF DIFFERENT COMBINATIONS OF FEATURE REPRESENTATION

These results indicate that the model successfully learned the relevant features and is capable of strongly separating the two classes, as demonstrated by its low false-positive and false-negative rates, as well as its high AUC, specificity, and sensitivity values. However, there seems to be a significant difference between the training accuracy and validation accuracy (3.6%), this indicates that the model might have been fed a little too much data or was given a dataset that was too detailed or too representative of specific patterns which causes it to rely on the patterns on the given train set a little too much, therefore causing the model to experience a slight overfitting. To prove the quality of the model in practice, we have used the model on the Kaggle dataset provided by TJIPTJ [16] where

we took 200 cases of healthy lung images and 200 cases of TBC lung images. The model correctly classified 393 out of 400 samples, achieving an average confidence score of 0.97, which suggests that it effectively captured the distinguishing characteristics of healthy versus TBC lungs. The model also showed a reasonable ability to approximate the evaluation of its own performance, as its average confidence level is similar to the overall accuracy. However the approximation was not entirely perfect, table II shows higher specificity than sensitivity (implying more false positives), whereas the external test results showed more false negatives than false positives. Overall, these findings suggest that while the model demonstrates very strong feature learning performance, future improvements should focus on reducing its sensitivity and specificity imbalance and dealing with the overfitting observed during validation to ensure even more consistent and generalizable predictions across varied datasets.

REFERENCES

- [1] World Health Organization, "Global Tuberculosis Report 2019," WHO, Geneva, Switzerland, 2019.
- [2] S. K. Sharma and A. Mohan, "Tuberculosis: From an incurable scourge to a curable disease—journey over a millennium," Indian J. Med. Res., vol. 137, no. 3, p. 455, 2013.
- [3] C. Silverman, "An appraisal of the contribution of mass radiography in the discovery of pulmonary tuberculosis," Amer. Rev. Tuberculosis, vol. 60, no. 4, pp. 466–482, 1949.
- [4] A. H. van't Hoog, H. K. Meme, K. F. Laserson, J. A. Agaya, B. G. Muchiri, W. A. Githui, L. O. Odeny, B. J. Marston, and M. W. Borgdorff, "Screening strategies for tuberculosis prevalence surveys: The value of chest radiography and symptoms," PLoS ONE, vol. 7, no. 7, Jul. 2012, Art. no. e38691.
- [5] A. P. Brady, "Error and discrepancy in radiology: Inevitable or avoidable?" Insights Imag., vol. 8, no. 1, pp. 171–182, Feb. 2017.
- [6] A. J. Degnan, E. H. Ghobadi, P. Hardy, E. Krupinski, E. P. Scali, L. Stratchko, A. Ulano, E. Walker, A. P. Wasnik, and W. F. Auffermann, "Perceptual and interpretive error in diagnostic radiology—Causes and potential solutions," Academic Radiol., vol. 26, no. 6, pp. 833–845, Jun. 2019.
- [7] M. van Cleeff, L. Kivihya-Ndugga, H. Meme, J. Odhiambo, and P. Klatser, "The role and performance of chest X-ray for the diagnosis of tuberculosis: A cost-effectiveness analysis in Nairobi, Kenya," BMC Infectious Diseases, vol. 5, no. 1, p. 111, Dec. 2005.
- [8] Y. Liu, J. Wu, X. Wu, J. Wang, Y. Lu, H. Zhang, and Y. Xu, "Deep learning assistance for tuberculosis diagnosis with chest radiography in low-resource settings," European Respiratory Journal, vol. 58, no. 6, p. 2100633, 2021.
- [9] Qatar University, University of Dhaka, and University of Malaya, "Tuberculosis (TBC) Chest X-ray Database," Kaggle, 2018. [Online]. Available: <https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tbc-chest-xray-dataset>
- [10] I. A. Ahmed, E. M. Senan, H. S. A. Shatnawi, Z. M. Alkhraisha, and M. M. A. Al-Azzam, "Multi-Techniques for Analyzing X-ray Images for Early Detection and Differentiation of Pneumonia and Tuberculosis Based on Hybrid Features," Diagnostics, vol. 13, no. 4, art. 814, 2023.
- [11] M. R. Tarambale and N. S. Lingayat, "Chest X-ray Enhancement for the Proper Extraction of Suspicious Lung Nodule," Int. J. Innov. Eng. Res. Technol., vol. 3, no. 11, pp. 1-9, 2021.
- [12] S. Ghimire, S. Kashyap, J. T. Wu, A. Karargyris, and M. Moradi, "Learning Invariant Feature Representation to Improve Generalization across Chest X-ray Datasets," arXiv, 2020.
- [13] O. Gozes and H. Greenspan, "Deep Feature Learning from a Hospital-Scale Chest X-ray Dataset with Application to TBC Detection on a Small-Scale Dataset," arXiv, 2019.
- [14] Deepanshu T., "Introduction to Feature Detection and Matching," Medium, 2019.
- [15] Muhammad Rehan. "Chest X-Ray Dataset". Available: <https://www.kaggle.com/datasets/muhammadrehan00/chest-xray-dataset>

- [16] TJIPTJ. "Chest X-Ray (Pneumonia,Covid-19,Tuberculosis)". Available: <https://www.kaggle.com/datasets/jtiptj/chest-xray-pneumoniacovid19tuberculosis>