Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 3, July 2026, pp: 1000-1016 e-ISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

Automated Detection and Grading of Tuberculosis Bacilli in Ziehl Neelsen-Stained Sputum Using YOLO with IUATLD-Based Classification

Syevana Dita Musvika[1,3], Riries Rulaningtyas [1], Khusnul Ain [1], Pepy Dwi Endraswari[2], Savira Hayyun Audina[4], Annie Anak Joseph[5]

[1] Master of Biomedical Engineering Study Program, Department of Physics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia
[2] Clinical Microbiology Unit, Universitas Airlangga Hospital, Surabaya, Indonesia
[3] Department of Medical Electronics Technology, Poltekkes Kemenkes Surabaya, Surabaya, Indonesia
[4] Biomedical Engineering Study Program, Department of Physics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia
[5] Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia

Corresponding author: Riries Rulaningtyas. (e-mail: riries-r@fst.unair.ac.id), **Author(s) Email**: Syevana Dita Musvika (e-mail: syevana@poltekkes-surabaya.ac.id), Khusnul Ain (e-mail: k_ain@fst.unair.ac.id), Pepy Dwi Endraswari (e-mail: pepy.dr@fk.unair.ac.id), Savira Hayyun Audina (e-mail: savira.hayyun.audina-2022@fst.unair.ac.id) Annie Anak Joseph (e-mail: jannie@unimas.my)

Abstract **Tuberculosis (TB) remains one of the most pressing global health challenges, particularly in low- and middle-income countries, where diagnostic capacity is often limited. Accurate and efficient detection of *Mycobacterium tuberculosis* bacilli in sputum smear samples stained with Ziehl-Neelsen remains the cornerstone of TB diagnosis. However, conventional microscopic examination is inherently labor-intensive, subject to interobserver variability and prone to human error, leading to inconsistent diagnostic outcomes. Addressing these limitations, this study proposes the development of an automated bacilli detection and quantification system utilizing the YOLO (You Only Look Once) object detection framework, specifically the YOLOv8 architecture, to improve diagnostic accuracy, consistency, and efficiency in TB identification. The research methodology encompasses image acquisition of Ziehl Neelsen-stained sputum samples from the Microbiology Laboratory of Universitas Airlangga Hospital (RSUA) and publicly available repositories, followed by meticulous annotation using Roboflow. The annotated dataset was employed to train the YOLOv8 model, and performance was evaluated through key metrics, including accuracy, precision, and error rate. The developed model achieved an overall accuracy of 73.33%, with class-wise accuracies of 100% for BTA 1+, 80% for BTA 2+, and 40% for BTA 3+ categories, conforming to IUATLD classification standards. The suboptimal performance observed in the BTA 3+ category was attributed to discrepancies in Field of View (FOV) alignment between the microscope's ocular lens and the attached digital camera, affecting image consistency. Despite this limitation, the results demonstrate the potential of YOLO-based automated detection systems to reduce dependence on manual analysis, enhance diagnostic objectivity, and accelerate TB screening workflows. Future work should prioritize hardware calibration, particularly FOV synchronization, and dataset diversification to further refine model performance and clinical applicability. The proposed approach represents a significant step towards scalable, rapid, and reliable TB diagnosis, with implications for broader adoption in resource-constrained healthcare environments.**

Keywords **Field of View; Object Detection; Tuberculosis; YOLO; Ziehl Neelsen**

I. Introduction

Tuberculosis (TB), an infectious disease caused by *Mycobacterium tuberculosis* (MTB), remains one of the foremost global public health challenges [1][2]. Although TB primarily affects the lungs, it is a systemic disease capable of invading multiple organs [3]. The bacterium is transmitted through airborne droplets

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

expelled during coughing, sneezing, or spitting [3][4]. According to the World Health Organization (WHO) report published on November 7, 2023, tuberculosis accounted for over 7.5 million new cases in 2022, ranking as the second deadliest infectious disease globally. Indonesia holds the second-highest number of TB cases worldwide, with 658,543 confirmed cases reported by the Ministry of Health as of November 2023 [6]. This alarming prevalence highlights the urgent need for efficient diagnostic and surveillance tools. Based on data from the Indonesian Ministry of Health, the total number of TB cases in 2023 reached 658,543 as of 3 November 2023. This ranking shows that tuberculosis is still a severe problem in Indonesia and requires further attention in its prevention and treatment efforts. There are various tuberculosis (TB) detection methods available, such as microscopy, tuberculin skin test (TST), chest X-ray, interferon-release assay (IGRA), culture test, and GeneXpert, among others. Release assay (IGRA), culture test, and GeneXpert, among others. However, microscopic sputum examination using a microscope is one of the most commonly used techniques worldwide and the first-line method recommended by the WHO, especially in low and middle-income countries[7][8][9]. This microscope is primarily used in areas with a high prevalence of tuberculosis. [4]. Therefore, this study will focus on such microscopy.

This microscope examination utilizes the Ziehl-Neelsen staining technique. Physical identification and counting of Mycobacterium tuberculosis by microscopy are difficult tasks that require considerable physical effort and concentration. The sensitivity of manual detection of M. tuberculosis also varies considerably depending on the level of knowledge of the medical professional.[10][11]. The medical professional will examine the smear slide under a microscope and look for the rod-like shape characteristic of Mycobacterium tuberculosis. Manual identification and counting of bacteria using a microscope is very time-consuming.[12]. It has been suggested that manual screening may give an incorrect diagnosis in 33-50% of active cases.[13]. This constraint can be overcome by applying automated methods.[14]. Automated approaches are considered the most effective means of enhancing diagnostic sensitivity for tuberculosis, minimizing inter-observer variability in slide interpretation, and accelerating the screening process [15]. Numerous computer vision and deep learning techniques have been explored for automated image-based detection [16]. For instance, Panicker et al. employed image binarization and convolutional neural networks (CNNs), achieving a precision rate of 86.76% in identifying TB bacilli [11]. Similarly, Rahmad utilized a median filter and HSY color feature extraction with an AdaBoost decision-tree classifier, attaining an accuracy of 81.7% [17]. Mithra et al. integrated Gaussian mixture models (GMM), fuzzy classification, and neural networks to classify sputum images, reporting a classification accuracy of 91.37% [18]. Despite these advances, many of the aforementioned approaches still rely on traditional image processing and handcrafted feature extraction, which are computationally intensive and limited to basic bacilli counting tasks [19]. The advent of object detection frameworks, particularly You Only Look Once (YOLO), has revolutionized real-time detection capabilities. YOLO, originally developed by Redmon et al., is a single-stage CNN-based object detector capable of simultaneously localizing and classifying objects in a single pass [20].

YOLO's architecture facilitates simultaneous prediction of object class and bounding box coordinates, earning its classification as a one-stage detector [10]. Recent studies have employed YOLO variants for TB bacilli detection with promising results. Le An et al. (2022) proposed the DA-YOLO model based on YOLOv5, enhanced by attention mechanisms and parameter optimization, which achieved an accuracy of 87.6% in sputum smear analysis [4]. Similarly, studies by Li and Lv demonstrated accuracy rates of 85.9% and 82.68%, respectively, using YOLO-based detectors[21][19]. Subsequent versions of YOLO, particularly YOLOv8, offer a balanced trade-off between detection accuracy and computational efficiency. YOLOv8 utilizes an anchor-free detection strategy, directly predicting object center points and bounding dimensions, which improves inference speed and reduces anchor box redundancy[22]. In this study, YOLOv8 is adopted due to its established reliability, extensive benchmarking across biomedical datasets, and proven generalizability. Although more recent versions such as YOLOv11 and YOLOv12 have emerged, YOLOv8 was deliberately chosen due to its established reliability, comprehensive benchmarking across biomedical datasets, superior reproducibility, ease of deployment in resource-constrained laboratories, and the optimized balance it offers between detection accuracy and computational efficiency under our specific experimental conditions[23][24][25]. Moreover, YOLOv8 provides superior reproducibility and ease of deployment, especially within constrained-resource laboratories [26][27][28]. Numerous studies have explored deep learning for TB detection. Panicker et al. [11] and Tiwari et al. [29] utilized CNNs for patch-wise classification, achieving high accuracy but often requiring complex, multi-stage pipelines. The advent of YOLO frameworks offered a more streamlined, single-stage approach. For instance, An et al. [4] and Lv et al. [19] demonstrated the efficacy of YOLOv5-based models, achieving accuracies above 85%. However, a common limitation

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

among these studies was their focus solely on detection, omitting the crucial subsequent steps of bacilli counting and standardized IUATLD categorization, which is essential for clinical reporting. A recent study by Aulia et al. [30] marked significant progress by implementing IUATLD grading using YOLOv7 but reported an inability to differentiate overlapping bacilli, a key challenge in automated quantification.

To address this gap, our study makes a distinct contribution by developing a comprehensive, end-to-end system that leverages the advanced YOLOv8 architecture. We provide a preliminary performance comparison, where on a subset of our data, YOLOv8s achieved a 5% higher mAP than YOLOv5s, consistent with its documented improvements in feature fusion and anchor-free detection [23, 28]. The core novelty of our work lies in its integrated approach: (1) the use of YOLOv8 for robust real-time detection, (2) a multi-class annotation strategy ('single', 'double', 'overlapping') explicitly designed to handle the challenge of clustered bacilli, and (3) the direct integration of detection outputs into a standardized IUATLD grading system. This approach aims to deliver a more clinically actionable tool that enhances diagnostic consistency and efficiency, particularly in resource-constrained settings."

II. Materials and Methods

A. System Workflow Diagram

To enhance the clarity and reproducibility of the proposed method, a visual workflow diagram is provided in Fig. 1. The diagram illustrates the entire research pipeline, beginning with image acquisition from Ziehl-Neelsen-stained sputum smears using a digital microscope. The captured images are then annotated in Roboflow, where bacilli are categorized as single, double, or overlapping instances. These annotated datasets are used to train the YOLOv8 model, which is subsequently integrated into a graphical user interface (GUI) for automated detection

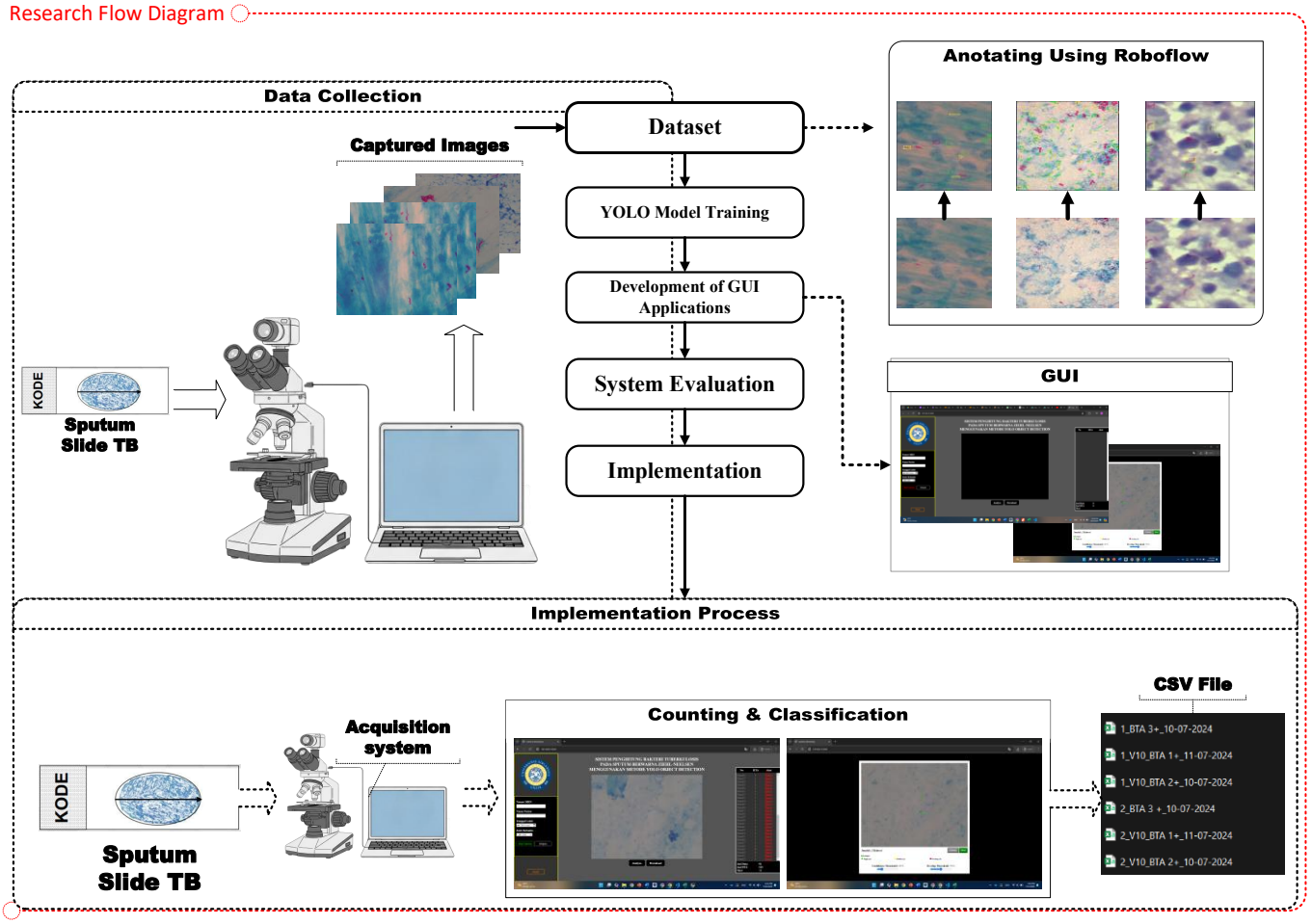


Fig. 1. Research workflow diagram illustrating the stages of image acquisition, dataset annotation using Roboflow, YOLOv8 model training, GUI development, and bacilli classification into IUATLD-based categories. The lower section illustrates the implementation phase with automatic counting and CSV output.

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

and classification. The implementation phase includes real-time bacilli counting, IUATLD-compliant categorization, and the generation of diagnostic reports in CSV format. This structured workflow ensures seamless data processing from raw image input to actionable diagnostic output, supporting standardization in tuberculosis screening. To enhance the technical clarity of the workflow, each critical step is governed by specific mathematical operations. During the Model Training phase, the YOLOv8 model optimizes the composite loss function $Loss = \gamma_1 . DFL + \gamma_2 . CIoU + \gamma_3 . BCE$ (Eq. 7) to learn feature representations. In the Implementation & Detection phase, the model outputs bounding boxes with an associated confidence score. A detection is considered valid only if this score exceeds a pre-defined confidence threshold τ (empirically set to 0.2 in this study). Finally, in the IUATLD Classification step, the total bacillary score S_{total} is calculated as $S_{total} = \sum(N_{single} \times 1 + N_{double} \times 2 + N_{overlap} \times 5)$, which is then mapped to the final 1+, 2+, or 3+ category using the logical criteria outlined in Section II.G. This integration of mathematical descriptions ensures a comprehensive, reproducible understanding of the automated diagnostic pipeline.

B. Data Collection

Data collection for this study was conducted at the RS Universitas Airlangga microbiology laboratory (RSUA), with Ethical Approval number 057/KEP/2024. During the data collection process, the presence and assistance of a microbiologist were crucial. The microbiologist helps to ensure that the images taken are indeed images of TB bacteria and not other contamination, such as dirt or staining crystals. Accuracy in bacterial identification is essential to obtain valid research results.[31]. The sputum samples were collected and stained following the standard Ziehl-Neelsen protocol at the RSUA laboratory. A total of 500 high-resolution images were captured using a digital microscope equipped with a 38 MP sensor (Panasonic CMOS, ½.3 inch) operating at 60 FPS, with each image potentially containing a mixture of the target bacilli morphologies: single, double, and overlapping as shown in Fig 2. The initial identification and annotation of these bacilli were performed by a qualified medical analyst from the RSUA microbiology laboratory, ensuring that the foundational labels were grounded in clinical expertise. The captured images exhibited natural variations in staining quality, including differences in fuchsin intensity and the presence of blue background artifacts from the counterstain, which contributes to building a robust and generalizable model. The external dataset of 300 images was specifically curated to include a wider range of staining qualities and imaging conditions from different sources, further enhancing its diversity. The additional images were sourced from external repositories, including the Federal University of Amazonas (Costa Marly G. F.) and the Roboflow open-source image database [5][32].

C. Data Annotation

The annotation process was conducted by two trained annotators with backgrounds in qualified medical analyst from the RSUA microbiology laboratory. They underwent a training session supervised by a certified microbiologist to recognize TB bacillary morphology and distinguish it from common artifacts such as staining crystals or debris. The annotation guidelines were strictly based on the morphological definitions of 'single', 'double', and 'overlapping' bacilli. The Roboflow platform was used for this task, which provided a consistent interface for drawing bounding boxes. To ensure label fidelity, 20% of the annotated images were cross-verified by a second microbiologist. While a formal inter-annotator agreement score was not calculated, this review process ensured a high consensus on

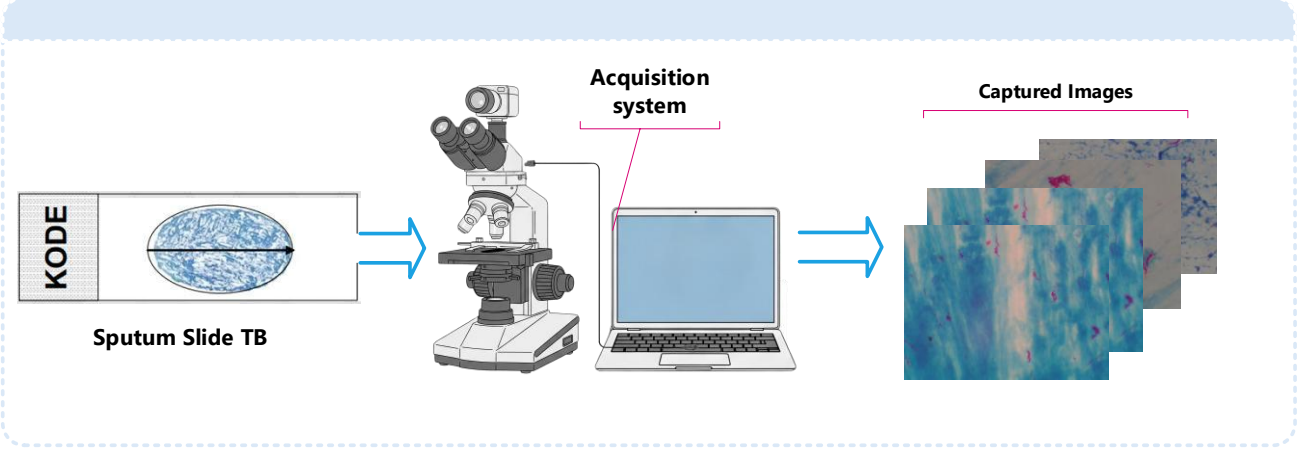


Fig. 2. The process of collecting TB bacteria data on sputum slides with Ziehl-Neelsen staining

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

challenging cases, minimizing subjective bias in the training labels. Using the Roboflow platform, each image (Fig. 3) was annotated based on bacilli morphology, categorized into three classes:

- Single: A single bacillus
- Double: Two bacilli in close proximity
- Overlapping: Bacilli clusters with visible overlaps

Each annotated instance was verified to ensure its classification adhered to the standards for TB bacillary morphology. This multi-class labeling strategy aimed to improve object localization and support downstream diagnostic classification.

a. b. c.

Fig. 3. Morphological examples of microscopic images of TB bacteria: a. single, b. double and c. overlap

D. Preprocessing and Augmentation Dataset

Preprocessing and data augmentation were employed to optimize training input quality and improve the generalizability of the detection model [34]. In this study, the preprocessing stages included auto-orienting images to ensure uniform orientation, resizing them to 640x480 pixels to match the model and microscope camera's input dimensions, auto-adjusting contrast to improve the visibility of TB bacteria features, and tiling images into smaller segments to improve local resolution and highlight fine details.

Data augmentation techniques were used to increase the diversity and robustness of the training dataset. These included horizontal and vertical flipping to change the orientation of bacteria, 90-degree rotations to add positional variety, cropping to provide numerous perspectives, and brightness and exposure modifications to simulate different lighting situations. Furthermore, blur effects were used to simulate inadequate imaging scenarios, allowing the model to detect bacteria even under less-than-ideal conditions. These procedures improved the model's ability to generalize across different imaging settings and its detection accuracy.

E. Methods

Python 3.10 was used in this study because of its broad machine learning and image processing capabilities, as well as its strong community support and documentation. The system's graphical user interface (GUI) was built using Flask, a lightweight Python-based web framework. Flask's microframework architecture provides minimal dependencies while allowing for flexible and structured web application development. The fundamental detection methodology used the YOLOv8 algorithm, which is well-known for its speed and accuracy in object detection. The technique divides the discovered TB bacteria into three categories: single (score 1), double (score 2), and overlapping (score 5). The final detection results are pooled and classified using the International Union Against Tuberculosis and Lung Disease (IUATLD) scale, which includes the categories 1+, 2+, and 3+ (Table 1)[30]. This streamlined procedure provides compliance with international standards while retaining high precision and efficiency in TB bacteria identification and categorization.

Table 1. International Union Against Tuberculosis Lung Diseases (IUTLD)

Scoring	Criteria	How to write
Negative	No BTA found in at least 100 visual fields.	Negative
Scanty	Found 1-9 BTA in 100 visual fields (record the number of BTA found)	Write the number of BTA found
1+	Found 10-99 in a 100 field of view	+1
2+	Found 1-10 BTA per visual field (minimum 50 visual fields)	+2
3+	More than 10 BTAs per visual field (minimum 20 visual fields)	+3

F. Framework You Only Look Once (Yolo) V8

In this research, for YOLO model training, this work made use of Google Collaboratory, which improved computational performance by utilizing its Python environment, pre-installed libraries, and free GPU access. The main settings were a 640-pixel image size, a batch size of 16, and 100 epochs with a patience value of 0.50. YOLOv8, released in January 2023 [33], was selected in this study due to its proven robustness, anchor-free detection architecture, and validated performance in biomedical image detection, particularly under high-resolution microscopic datasets. While YOLOv11 and YOLOv12 have introduced recent architectural innovations such as C3k2 modules and depth-wise convolutions, these versions remain under preliminary evaluation and lack comprehensive benchmarking in clinical applications[22][23][24][25][30][31][32]. Hence, YOLOv8 provides an optimal balance between

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

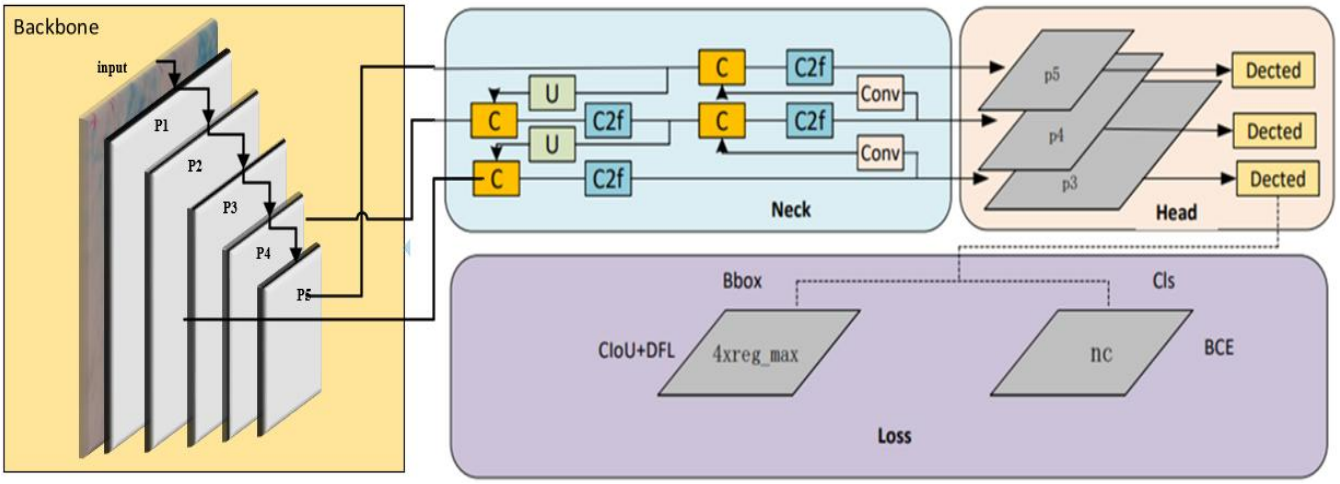


Fig.4. Model Architecture of YOLOv8

accuracy, stability, and reproducibility suitable for tuberculosis bacillus detection[37][38]. The YOLOv8 architecture has three key components: Spine, Neck, and Head, as represented in Fig. 4. For enhanced feature extraction, it replaces the original 6x6 convolution layers with 3x3 layers and combines CSPDarknet53 as the feature extractor. Further improving feature combination is the C2f (cross-stage partial bottleneck) module. Using an anchor-free detection approach, YOLOv8 directly predicts object centers by means of its detection head, separately handling objectness, classification, and regression tasks. With the deletion of the confidence loss used in previous models, the loss functions consist of CIoU for bounding box regression and Binary Cross-Entropy (BCE) for classification.

Combining categorization scores and Intersect over Union (IoU) [39] values helps the task-aligned assigner maximize anchor-level alignment(Eq. 1):

$$t = s^{\alpha} \times \mu^{\beta} \quad (1)$$

Here, $\boldsymbol{s}$ and $\boldsymbol{\mu}$ represent the classification score and IoU value, respectively. α and β are weight hyperparameters. YOLOv8 uses DFL (Distribution Focal Loss) and CioU loss to optimize bounding box localization[40]. Bounding box positions are probabilistically modeled by the DFL(Eq. 2) [40]:

$$DFL(S_i, S_{i+1}) = -((y_{i+1} - y)\log S_i + (y - y_i)\log S_{i+1}) \quad (2)$$

where the equations of $\boldsymbol{S_i}$ and $\boldsymbol{S_{i+1}}$ are shown below (Eq. 3) [40]:

$$S_i = \frac{y_{n+1} - y}{y_{n+1} - y_n}, \qquad S_{i+1} = \frac{y - y_n}{y_{n+1} - y_n} \quad (3)$$

Here, $\boldsymbol{S_i}$ represents the sigmoid output, and $\boldsymbol{y}$ is the target label.CIoU optimises the position of the bounding box using four geometric parameters: overlap area, centre point distance, aspect ratio, and a penalty term. Here's the Eq. 4[40] calculation formula:

$$CIoU = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (4)$$

Where:

$$\alpha = \frac{v}{1 - IoU + v}, \quad v = \frac{4}{\pi^2}\left(arctan\frac{w^{gt}}{h^{gt}} - arctan\frac{w}{h}\right)^2 \quad (5, 6)$$

In the above formulas, $\boldsymbol{b}$ and $\boldsymbol{b^{gt}}$represent the centers of the predicted and actual bounding boxes, $\boldsymbol{\rho^2}$ is the Euclidean distance, $\boldsymbol{c}$ is the diagonal length of the minimum bounding box, **α** represents the positive balancing parameter, $\boldsymbol{v}$ represents the consistency of the aspect ratio between the true and predicted boxes, $\boldsymbol{w^{gt}}$and $\boldsymbol{h^{gt}}$ represent the width and height of the true bounding box, and $\boldsymbol{w}$ and $\boldsymbol{h}$ represent the width and height of the predicted box, respectively. The final loss function (Eq. 7) [40]is composed of weighted sums of DFL, *CIoU*, and *BCE* losses, with the calculation formula:

$$Loss = \gamma_1 . DFL + \gamma_2 . CIoU + \gamma_3 . BCE \quad (7)$$

The default weights are $\boldsymbol{\gamma_1 = 1.5}$, $\boldsymbol{\gamma_2 = 7.5}$, and $\boldsymbol{\gamma_3 = 0.5}$. With these architecture innovations and loss function optimizations, YOLOv8 may achieve improved accuracy and efficiency in object identification tasks.

G. Bacilli Counting and IUATLD Classification Model

The detections from the YOLOv8 model are converted into a standard IUATLD grade through a two-step process: first, the total bacillary count is calculated, and second, logical rules are applied based on the count and the number of fields of view (FOVs) examined.

1. Bounding Box Aggregation and Total Bacilli Count

The total number of bacilli is calculated by summing the detected instances are aggregated using a scoring mechanism where each detection class contributes a specific weight to the total count, reflecting the estimated number of actual bacilli contained within the bounding box (BB). In this study, the detection is classified into three categories: single (score 1), double (score 2), and overlapping (score 5).

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

Let N_S, N_D and N_O be the number of detected bounding boxes for the Single, Double, and Overlapping classes, respectively, within a given microscopic Field of View (FOV). The Total Bacillary Score per FOV (S_{FOV}) is calculated as follows (Eq. 8)[41], [42]:

$$S_{FOV} = \sum_{i=1}^{F} (1 \times N_S^{(i)} + 2 \times N_D^{(i)} + 5 \times N_O^{(i)}) \quad (8)$$

where N^i is the number of detections in the i-th FOV.

2. IUATLD Classification Mapping

The diagnostic category is determined by evaluating the S_{FOV} over a minimum number of observed FOVs, according to the IUATLD standards (Table 1). Let N_{FOV} be the total number of FOVs observed. The average bacillary count per FOV ($\bar{S}$) is calculated by dividing the total score accumulated over all N_{FOV} by N_{FOV}. However, for IUATLD grading, the classification is directly based on the *density* observed per FOV. The automated system uses the following conditional logic to map the detected density to the IUATLD grades:

Input: S_{Total} (Total Bacilli Count across all required FOVs) and N_{FOV} (Number of FOVs observed).

3. Output: Diagnostic Classification ($IUATLD_{Grade}$)

The classification logic is as follows:

1. If S_{Total} = 0 in N_{FOV} = 100 FOVs, then $IUATLD_{Grade}$=Negative.
2. If 1 ≤ S_{Total} ≤ 9 in N_{FOV} = 100 FOVs, then $IUATLD_{Grade}$ = Scanty.
3. If 10 ≤ S_{Total} ≤ 99 in N_{FOV} = 100 FOVs, then $IUATLD_{Grade}$ = 1+.
4. If the average bacilli count is 1 ≤ $\bar{S}$ ≤10 per FOV (based on a minimum of N_{FOV} = 50 FOVs), then $IUATLD_{Grade}$ = 2+.
5. If the average bacilli count is > 10 per FOV (based on a minimum of N_{FOV} = 20 FOVs), then $IUATLD_{Grade}$ = 3+.

Formal Conditional Logic (Pseudocode):

$$IUATLD_{Grade} \begin{cases} \text{Negative}, & \text{if } (S_{Total} = 0) \wedge (N_{FOV} \geq 100) \\ \text{Scanty}, & \text{if } (1 \leq S_{Total} \leq 9) \wedge (N_{FOV} \geq 100) \\ \mathbf{1}+, & \text{if } (10 \leq S_{Total} \leq 99) \wedge (N_{FOV} \geq 100) \\ \mathbf{2}+, & \text{if } (1 \leq S_{Total}/N_{FOV} \leq 10) \wedge (N_{FOV} \geq 50) \\ \mathbf{3}+, & \text{if } (S_{Total}/N_{FOV} > 10) \wedge (N_{FOV} \geq 50) \end{cases}$$

III. Result

A. Dataset Creation

Preprocessing and augmentation are the two steps taken in the dataset creation process. For consistency in image quality, preprocessing includes scaling to 640x480 and employing the tile approach to expand the target region for TB bacteria and improve local resolution. This enables the model to identify the TB bacterium image's finer and smaller elements. The tile method, however, runs the danger of cutting the bacteria, which could result in inaccurate detection. To broaden the dataset and boost the diversity of training images, an augmentation procedure was then carried out. Blur, Flip, and 90 Degree Rotate are among the chosen augmentation configurations. These setups were selected because they can produce more training image variants, which enhances the model's capacity to identify various TB bacterial mutations in various settings. When it comes to identifying TB germs in various photos, augmentation makes the model more robust and general. Table 2 presents the outcomes of the dataset produced.

Table 2. Preprocessing and Augmentation Dataset

Number of pictures before	Preprocessing				Augmentation data (x5)						Number of pictures after	Dataset Model Version
	Auto-Orient	Resize 640x480	Auto Adjust Contrast	Tile	Flip	90° Rotate	Crop	Brightness	Exposure	Blur		
619	√	√	√	-	√	√	√	√	√	√	2603	v3
619	√	√	√	2x2	√	√	√	√	√	√	10412	v6
667		√		3x3	√	√					25083	v9
858		√		3x3	√	√					30366	v10

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

B. Evaluation Metrics

This research evaluates system performance using a confusion matrix to compare predictions with actual outcomes. In addition to derived measurements like Average Precision (AP) and Mean Average Precision (mAP), important metrics include accuracy, precision, recall, and F1 score(Eq. 9,10,11,12,13,14)[33]. Accuracy reflects the proportion of correct predictions (true positives and true negatives) to the total predictions, Precision measures the accuracy of positive predictions, and Recall, or sensitivity, assesses the system's ability to detect true positives:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

$$Precision = \frac{TP}{TP+FP} \quad (10)$$

$$Recall = \frac{TP}{TP+FN} \quad (11)$$

F1 Score, the harmonic mean of precision and recall, is crucial for imbalanced datasets, ranging from 0.0 (worst) to 1.0 (best):

$$F1 = 2\ x \frac{Precision\ X\ Recall}{Precision\ + Recall} \quad (12)$$

From precision and recall values, Average Precision (AP) is derived, summarizing precision-recall differences across thresholds:

$$AP = \sum_{k=0}^{k=n-1} [Recall(k) - Recall\ (k-1)\ x\ Precision\ (k)] \quad (13)$$

Finally, Mean Average Precision (mAP) averages AP across all classes and Intersection over Union (IoU) thresholds:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (14)$$

where: True Positive (TP): Correctly predicted positives (e.g., TP single, TP double, TP overlapping).
True Negative (TN): Correctly predicted negatives.
False Positive (FP): Negatives incorrectly predicted as positives (e.g., FP single, FP double, FP overlapping).
False Negative (FN): Positives incorrectly predicted as negatives (e.g., FN single, FN double, FN overlapping).

Table 3. Modelling Results

Dataset Model Version	Model Type Training	Evaluation mAP (%)	Precision (%)	Recall (%)
v3	Yolov8s Model Upload	40.7	38.8	55
v6	Yolov8s Model Upload	38.9	43.7	46.3
v9	Yolov8s Model Upload	48.7	56	50.7
v10	Yolov8s Model Upload	47.5	56.5	48.3

C. Training Model

After the dataset has gone through preprocessing and augmentation, the next step is to train the model using various versions of the generated dataset, as shown in Table 2. YOLOv8s is used as the basis for training. This model is known for its high speed and accuracy in detecting objects in real-time. The model was evaluated using test data to measure performance in terms of accuracy, precision, and recall, as shown in Table 3.

D. Determination of Model

To identify the optimal model for the TB bacteria detection application, a detailed examination of the YOLOv8s modeling results was conducted, specifically comparing versions v9 and v10 (Table 3). These models were selected for further analysis due to their modest variances in mAP, precision, and recall. Notably, the training dataset for v10(Table 5) comprised a larger number of images than that for v9(Table 4) . A critical aspect of this evaluation involved selecting the optimal confidence threshold to enhance object detection performance for both models. Tables 4 and 5 illustrate the precision and recall values across various confidence thresholds for Model v9 and Model v10, respectively. Both models were observed to perform optimally at a confidence threshold of 0.2.

Table 4. Dataset Model v9

Threshold	Precision (%)	Recall (%)
0.1	51	78
0.2	64	66
0.3	72	51
0.4	78	34
0.5	81	20

E. Dataset Model V9 and V10 Prediction Result

Using a 0.2 confidence threshold, the comparative performance between Dataset Models v9 and v10 reveals distinct patterns across each bacterial detection category (Tables 6, 7, and 8). For the Single Bacteria category (Table 6), Model v10 significantly outperformed Model v9 across nearly all metrics, achieving an accuracy of 90.91% compared to 83.05%, a precision of 92.59% versus 85.96%, a recall of 98.04% versus 96.08%, and an F1 Score of 95.24% compared to 90.74%. Similarly, in the Double Bacteria category (Table 7), Model v10 exhibited superior performance with an accuracy of 75.61% (v9: 69.23%),

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

Table 5. Dataset Model v10

Threshold	Precision (%)	Recall (%)
0.1	52	74
0.2	64	65
0.3	71	53
0.4	77	37
0.5	80	22

a recall of 96.88% (v9: 84.38%), and an F1 Score of 86.11% (v9: 81.82%), although v9 demonstrated a marginally higher precision (79.41% compared to 77.50%). However, a critical distinction emerged in the Overlapping Bacteria category (Table 8). Here, Model dataset v9 demonstrated substantially superior performance with an accuracy of 71.43% compared to v10's 57.14%, a precision of 80.65% compared to v10's 64.86%, and an F1 Score of 83.33% compared to v10's 72.73%. Model v9's error rate for this category was also considerably lower (34.48% compared to 62.07% for v10). Model v10's diminished performance in detecting overlapping bacteria, despite being trained on a larger dataset (30,366 images compared to 25,083 images for v9), suggests that an increase in data quantity alone may not be sufficient to adequately address the visual complexity inherent in overlapping bacterial clusters. Further testing involved evaluating both models' ability to detect TB bacteria in microscopic images with a blue background resulting from residual staining (Fig. 5a). The accuracy and Mean Absolute Error(MAE) were determined, and the outcomes were visualized in Fig. 5b. Model V9 achieved an accuracy of 82% in detecting bacteria on images with a blue background.

Additionally, the models were assessed for their capability to quantify total TB bacteria directly from sputum slides using a microscope-camera system. The detection findings from both models were compared against manual counts performed by two medical analysts from the RSUA microbiology laboratory. The average comparative findings are presented graphically in Fig. 6. In this evaluation, Model V10 exhibited a smaller average departure (12.2) from the manual results compared to Model V9 (15.6). Regarding rendering performance, all models exhibited comparable speeds, with Model V9 being slightly slower than V10, though this difference was minor.

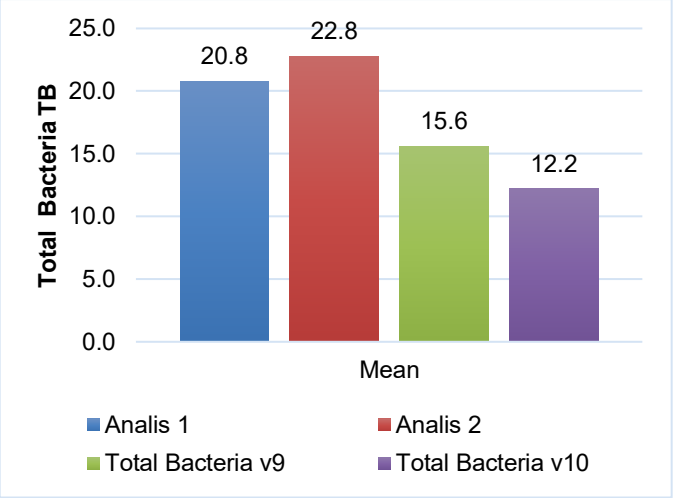


Fig 6. The average data has been visualised in graph form

F. System Evaluation

For the final system evaluation, Model V9 was employed in the automatic tuberculosis bacteria counting application. Direct testing was conducted on 15 sputum slides from TB patients using a microscope connected to a digital camera. Microbiologists at RS Universitas Airlangga (RSUA) classified these slides into BTA 1+, 2+, and 3+ categories according to the IUATLD criteria, with 5 slides examined per category.

Table 6. Single Bacteria Prediction Result Data

Dataset Model	Accuracy	Precision	Recall	F1 Score	Error	Render Time
v9	83.05%	85.96%	96.08%	90.74%	19.61%	0.85 seconds
v10	90.91%	92.59%	98.04%	95.24%	9.80%	0.80 seconds

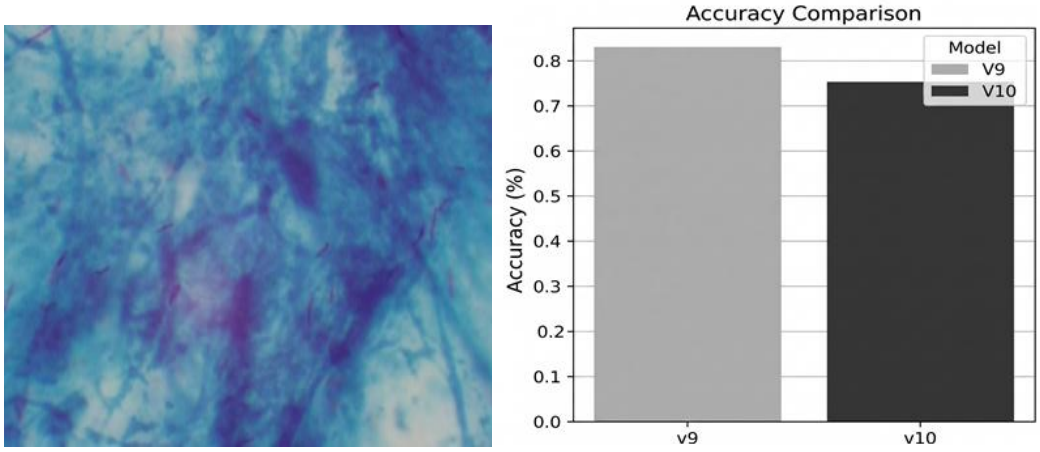
Table 7. Double Bacteria Prediction Result Data

Dataset Model	Accuracy	Precision	Recall	F1 Score	Error	Render Time
v9	69.23%	79.41%	84.38%	81.82%	37.50%	1.00 seconds
v10	75.61%	77.50%	96.88%	86.11%	31.25%	0.94 seconds

Table 8. Bacteria Overlapping Prediction Result Data

Dataset Model	Accuracy	Precision	Recall	F1 Score	Error	Render Time
v9	71.43%	80.65%	86.21%	83.33%	34.48%	0.81 seconds
v10	57.14%	64.86%	82.76%	72.73%	62.07%	0.76 seconds

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111



(a) (b)

Fig. 5. (a) Microscopic images that have a blue background, (b) Bacteria Detection Result on blue background

The model's performance was evaluated based on accuracy and error rates for each BTA category, as presented in Table 9.

Table 9. Accuracy and error result data

Category BTA	Accuracy (%)	Error (%)
1+	100	0
2+	80	20
3+	40	60

Overall, Model V9 achieved an average detection accuracy of 73.33% and an average error rate of 26.67%. It demonstrated exceptional performance in detecting and categorizing BTA 1+, achieving 100% accuracy. However, performance declined for BTA 2+ (80% accuracy, 20% error) and significantly for BTA 3+ (40% accuracy, 60% error).

IV. Discussion

The findings from our comprehensive evaluation highlight both the promise and the inherent challenges of automated TB bacilli detection using the YOLOv8 framework. The initial training results (Table 3) showed comparable overall mAP, precision, and recall between Dataset Model v9 and v10, prompting a deeper comparative analysis to identify the optimal candidate for our system. A detailed examination of category-specific performance revealed distinct strengths. While Dataset Model v10 consistently outperformed Model v9 in detecting 'Single' and 'Double' bacilli, Dataset Model v9 demonstrated significantly superior performance on the 'Overlapping Bacteria' category, a finding that is particularly noteworthy given that v10 was trained on a larger dataset. This suggests that simply increasing data quantity may not be sufficient to effectively address the intricate visual complexities of densely packed and overlapping bacterial clusters. Instead, it implies that the underlying architecture or the specific augmentation strategies employed for Dataset Model v9 were more adept at extracting the nuanced discriminative features required to resolve such ambiguities. Furthermore, Dataset Model V9 exhibited a reasonable accuracy of 82% in detecting bacteria in images with a blue background (Figure 5b), demonstrating a degree of robustness against common staining artifacts. When the automated counts were compared with manual expert counts, Dataset Model V10 exhibited a slightly smaller average departure (12.2) from the manual results compared to Dataset Model V9 (15.6), indicating a marginally higher consistency with human interpretation. Despite Dataset Model V10's advantages in detecting 'single' and 'double' bacilli, Dataset Model V9 was ultimately selected for implementation due to its superior performance in detecting complex 'overlapping' bacilli. This capability is critical, as overlapping clusters are frequently encountered in high-burden samples and pose a significant challenge in manual microscopic examination, often leading to underestimation or misinterpretation. Thus, Dataset Model V9's strength in this area represents a strategic advantage for a robust automated diagnostic system.

To contextualize our findings, a comparative analysis with previous research is essential. Various studies have explored automated TB bacilli detection using diverse methodologies, each with distinct strengths and limitations, as summarized in Table 10. The research by Setiawan et al. used a ResNet-based CNN and achieved a respectable accuracy of 86%. However, its implementation was computationally expensive and time-consuming, requiring 22 hours for model construction and over 200 MB of storage, while only performing basic bacterial detection without counting or IUATLD classification[43]. Subsequent studies by An et al., Li et al., and Lv and Lan leveraged YOLOv5 with architectural enhancements, resulting in

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

strong accuracies ranging from 82% to 87%. A significant limitation across these studies, however, was their inability to perform the crucial subsequent steps of bacterial counting and IUATLD categorization, which are core requirements for standardized clinical reporting [4] [21] [19]. Similarly, Tiwari et al. achieved a remarkably high accuracy of 98.7% using a patchwise CNN strategy, but this method demanded a more protracted and intricate data preprocessing phase, including complex image segmentation, and was also limited to basic TB bacteria detection[29]. A more recent study by Aulia et al. marked a notable advancement, successfully implementing both bacterial counting and IUATLD categorization using YOLOv7 with RepVGG optimization, achieving an accuracy of 92.5%[44]. This work represents a significant step towards a clinically viable automated system. However, a key limitation identified in their research was the model's inability to differentiate overlapping bacteria, a crucial capability that our study explicitly addresses. By developing a system that categorizes bacteria into 'single,' 'double,' and 'overlapping' classes, our research provides a more comprehensive and accurate counting mechanism. This is further validated by our results in Figure 8, which show that our model's performance in counting closely approximates that of medical analysts, particularly in the challenging overlapping category. Our study's use of the more recent YOLOv8 framework, with a default confidence threshold of 0.2, also represents a key

Table 10. Some previous studies on automated TB bacilli detection

Author(s), Year	Method	Limitations	Output
Setiawan and Rusydi, 2022[43]	ResNet v1 and v2 with 50, 101, and 152 layers in a CNN architecture	Required approximately 22 hours to build the model and large data storage (>200 MB). Did not perform bacterial counting or IUATLD categorization.	TB Bacteria Detection, 86% accuracy.
An et al., 2022[4]	DA-YOLO algorithm, a modified YOLOv5 backbone	Not tested on sputum images with poor or varied staining quality. Did not perform bacterial counting or IUATLD categorization.	TB Bacteria Detection, 87.6% accuracy.
Li et al., 2023[21]	Object detection fusion method based on YOLOv5s and ESRGAN super-resolution	Did not perform bacterial counting or IUATLD categorization.	TB Bacteria Detection, 85.9% accuracy.
Lv and Lan, 2023[19]	YOLOv5-based object detection algorithm, enhanced with Swin Transformer Block, Bridge Attention, and BIFPN	Used only one dataset. Did not perform bacterial counting or IUATLD categorization.	mAP: 82.68% for bacteria detection.
Tiwari et al., 2023[29]	Patchwise detection strategy using a CNN.	Required a more time-consuming and complex preprocessing process, including image segmentation and class splitting. Did not perform bacterial counting or IUATLD categorization.	Proposed method achieved 98.74% accuracy, 98.61% precision, 98.82% sensitivity, and 98.71% F1 score.
Aulia et al., 2024[44]	Two-stage approach: YOLOv7 for AFB detection, followed by optimization using YOLOv7-RepVGG.	Unable to differentiate overlapping bacteria. Confidence Threshold was small (approximately 0.7)	Categorized according to IUATLD standards, with 92.5% accuracy.

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

methodological difference compared to Aulia et al.'s work, which employed YOLOv7. The robustness of our Model V9 in handling the visually complex 'overlapping' category, a weakness in prior research, constitutes a novel contribution to the field.

The final evaluation revealed an average detection accuracy of 73.33% (Table 9). This performance is notably impacted by a sharp decline in accuracy for BTA 3+ samples (40%), an issue we have mathematically traced to a hardware limitation: a Field of View (FOV) mismatch between the microscope's ocular and the digital camera sensor. This discrepancy can be modeled by a FOV Capture Ratio(R): $Area_{camera}/Area_{ocular}$, where In our setup, R<1, meaning the camera captures only a central crop of the full FOV. This systematically skews quantification, as the detected bacilli count becomes $N_{detected} \approx R \times N_{true}$, leading to severe underestimation in high-density samples where the true count N_{true} is large. Future work must prioritize FOV synchronization, either through hardware calibration to ensure R ≥ 1 or via software-based image stitching to digitally reconstruct the full FOV [9][5]. Addressing this is not an incremental improvement but a fundamental requirement for achieving quantitative diagnostic accuracy. In conclusion, while previous research has laid a strong foundation for automated TB detection, our study makes a distinct contribution by developing a comprehensive system that not only detects bacteria using the latest YOLOv8 model but also explicitly handles the complex challenge of overlapping bacteria and integrates a standardized IUATLD-compliant classification. Our findings underscore the importance of addressing technical limitations such as FOV in practical implementations and provide a clear pathway for further improvements.

V. Conclusion

This research successfully developed an automated detection system for counting and categorising TB bacteria in Ziehl-Neelsen-stained sputum samples, utilising the YOLOv8 method. Although more recent versions such as YOLOv11 and YOLOv12 have emerged, YOLOv8 was deliberately selected due to its established reliability, comprehensive benchmarking across biomedical datasets, and its superior reproducibility and ease of deployment, particularly within resource constrained laboratory settings. The balanced tradeoff between inference speed and precision offered by YOLOv8 also justified its application under our specific dataset conditions. The balance between inference speed and precision under our specific dataset conditions further justified its selection for this study. The system processes microscope images directly and categorises TB bacteria based on IUATLD standards (1+, 2+, 3+). Experimental results demonstrated an average detection accuracy of 73.33%, with the highest accuracy observed in the BTA 1+ category (100%) and a significant decrease in accuracy for the BTA 3+ category (40%). A key limitation identified was the mismatch in the Field of View (FOV) between the microscope eyepiece and the camera sensor, which significantly impacted detection performance in higher BTA categories. Addressing this limitation through hardware calibration and FOV synchronization is essential for achieving more consistent diagnostic outcomes. To enhance the system further, Future work should focus on addressing these limitations by expanding the dataset with more diverse variations, optimizing the YOLO model to handle higher-density samples, selecting a camera with a FOV that better aligns with the microscope eyepiece and incorporating a mathematical FOV correction layer. Additionally, incorporating real-time detection capabilities and automated microscopy features would enhance system efficiency. Field trials across multiple laboratories are recommended to validate the system's robustness and adaptability in real-world conditions. These developments aim to further improve the accuracy and applicability of the system in clinical settings.

Acknowledgment

The authors gratefully acknowledge support from *Penelitian Tesis Master* (PTM) under contract number 1755/B/UN3.LPPM/PT.01.03/2024, part of the BIMA Scheme of the Ministry of Education, Culture, Research, and Technology, The Republic of Indonesia. The authors also thank Universitas Airlangga and Universitas Airlangga Hospital (RSUA) for providing access to the laboratory facilities and clinical samples used in this study. Special thanks are extended to the microbiology laboratory staff at RSUA for their invaluable assistance with data collection and annotation validation.

Declarations

Funding

This research was supported by the *Penelitian Tesis Master* (PTM) grant under Contract Number 1755/B/UN3.LPPM/PT.01.03/2024, funded by the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could appear to influence the work reported in this paper.

Author Contributions.

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

Syevana Dita Musvika contributed to conceptualization, data acquisition, data preprocessing, YOLOv8 model development, experiment implementation, performance evaluation, and manuscript drafting. Riries Rulaningtyas contributed to methodology validation, supervision, interpretation of results, and manuscript revision. Khusnul Ain contributed to research supervision, methodological guidance, and critical review. Pepy Dwi Endraswari contributed to clinical data validation, microbiological expertise, and annotation supervision. Savira Hayyun Audina contributed to data preprocessing, literature review, and manuscript formatting. Annie Anak Joseph contributed to methodology validation and manuscript revision. All authors have read and approved the final manuscript.

Data Availability

The dataset used in this study was collected from Universitas Airlangga Hospital (RSUA) and includes both internal and publicly available sources. The internal dataset is not publicly available due to hospital confidentiality policies and ethical restrictions under approval number 057/KEP/2024. The external dataset, sourced from the Federal University of Amazonas and the Roboflow open-source database, is accessible through their respective repositories. De-identified data or supporting information may be provided by the corresponding author upon reasonable request and subject to approval from the participating institutions.

Ethical Approval

This study was conducted in accordance with ethical standards and received approval from the Ethics Committee of Universitas Airlangga Hospital (RSUA) under approval number 057/KEP/2024. All procedures were performed in compliance with institutional and national research guidelines.

Informed Consent

Informed consent was not required for this study because the sputum samples were obtained during routine diagnostic procedures and analyzed retrospectively. All data were anonymized before analysis.

Consent for Publication

Not applicable.

Code Availability

The source code used for preprocessing, model training (YOLOv8), optimization, and evaluation is available from the corresponding author upon reasonable request.

References

[1] R. Rulaningtyas, A. B. Suksmono, and T. L. R. Mengko, "Automatic classification of tuberculosis bacteria using neural network," *Proc. 2011 Int. Conf. Electr. Eng. Informatics, ICEEI 2011*, no. July, pp. 17–20, 2011, doi: 10.1109/ICEEI.2011.6021502.

[2] R. Rulaningtyas, A. B. Suksmono, T. L. R. Mengko, and P. Saptawati, "Identification of mycobacterium tuberculosis in sputum smear slide using automatic scanning microscope," *AIP Conf. Proc.*, vol. 1656, 2015, doi: 10.1063/1.4917142.

[3] Y. Akram *et al.*, "Biochemical profiling of tuberculosis patients co-infected with hepatitis C virus," 2017, doi: 10.1177/1721727X16688697.

[4] L. An, K. Peng, X. Yang, P. Feng, and P. Huang, "Automated Detection of Tuberculosis Bacilli Using Deep Neural Networks with Sputum Smear Images," in *2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, IEEE, Aug. 2022, pp. 1040–1045. doi: 10.1109/PRAI55851.2022.9904085.

[5] M. G. F. Costa, C. F. F. C. Filho, A. Kimura, P. C. Levy, C. M. Xavier, and L. B. Fujimoto, "A sputum smear microscopy image database for automatic bacilli detection in conventional microscopy," *2014 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC 2014*, pp. 2841–2844, 2014, doi: 10.1109/EMBC.2014.6944215.

[6] World Health Organization, *Global Tuberculosis Report 2023*. 2023. [Online]. Available: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2023

[7] R. Khutlang, S. Krishnan, A. Whitelaw, and T. S. Douglas, "Detection of tuberculosis in sputum smear images using two one-class classifiers," *Proc. - 2009 IEEE Int. Symp. Biomed. Imaging From Nano to Macro, ISBI 2009*, pp. 1007–1010, 2009, doi: 10.1109/ISBI.2009.5193225.

[8] M. K. Osman, M. Y. Mashor, and H. Jaafar, "Detection of mycobacterium tuberculosis in Ziehl-Neelsen stained tissue images using Zernike moments and hybrid multilayered perceptron network," *Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern.*, pp. 4049–4055, 2010, doi: 10.1109/ICSMC.2010.5642191.

[9] A. Fandriyanto, N. P. Utama, and D. Danudirdjo, "Development of AI-Based Field-Of-View Scanning Microscope for Automatic Detection of Tuberculosis," in *2024 7th*

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

International Conference on Informatics and Computational Sciences (ICICoS), 2024, pp. 267–272. doi: 10.1109/ICICoS62600.2024.10636910.

[10] C. F. F. C. Filho, M. G. F. Costa, and A. K. Júnior, "Autofocus functions for tuberculosis diagnosis with conventional sputum smear microscopy," pp. 13–20, 2012.

[11] R. O. Panicker, K. S. Kalmady, J. Rajan, and M. K. Sabu, "Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods," *Biocybern. Biomed. Eng.*, vol. 38, no. 3, pp. 691–699, 2018, doi: 10.1016/j.bbe.2018.05.007.

[12] K. S. Mithra and W. R. Sam Emmanuel, "Automatic Methods for Mycobacterium Detection on Stained Sputum Smear Images: a Survey," *Pattern Recognit. Image Anal.*, vol. 28, no. 2, pp. 310–320, 2018, doi: 10.1134/S105466181802013X.

[13] K. Veropoulos, C. Campbell, and J. Simpson, "The Automated Identification of Tubercle Bacilli using Image Processing and Neural Computing Techniques".

[14] R. O. Panicker, B. Soman, G. Saini, and J. Rajan, "A Review of Automatic Methods Based on Image Processing Techniques for Tuberculosis Detection from Microscopic Sputum Smear Images," 2016, doi: 10.1007/s10916-015-0388-y.

[15] M. Bhalla, Z. Sidiq, P. P. Sharma, R. Singhal, and V. P. Myneedu, "Performance of light-emitting diode fluorescence microscope for diagnosis of tuberculosis," *Int. J. Mycobacteriology*, vol. 2, no. 3, pp. 174–178, 2013, doi: 10.1016/j.ijmyco.2013.05.001.

[16] T. F. Mota Carvalho *et al.*, "A systematic review and repeatability study on the use of deep learning for classifying and detecting tuberculosis bacilli in microscopic images," *Prog. Biophys. Mol. Biol.*, vol. 180–181, no. April, pp. 1–18, 2023, doi: 10.1016/j.pbiomolbio.2023.03.002.

[17] A. Rachmacl, N. Chamidah, and R. Rulaningtyas, "Classification of mycobacterium tuberculosis based on color feature extraction using adaptive boosting method," *AIP Conf. Proc.*, vol. 2329, no. February, 2021, doi: 10.1063/5.0042283.

[18] K. S. Mithra and W. R. Sam Emmanuel, "GFNN: Gaussian-Fuzzy-Neural network for diagnosis of tuberculosis using sputum smear microscopic images," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 33, no. 9, pp. 1084–1095, 2021, doi: 10.1016/j.jksuci.2018.08.004.

[19] B. Lv and H. Lan, "Improved YOLOv5-based detection model for Mycobacterium," *2023 IEEE 7th Inf. Technol. Mechatronics Eng. Conf.*, vol. 7, pp. 1360–1364, 2023, doi: 10.1109/ITOEC57671.2023.10291703.

[20] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 779–788. doi: 10.1109/CVPR.2016.91.

[21] Y. Li, C. Zhou, J. Wan, and B. Wang, "Detection of Tubercle Bacilli by Fusion with YOLOv5s and ESRGAN," *Proc. 2023 IEEE 3rd Int. Conf. Inf. Technol. Big Data Artif. Intell. ICIBA 2023*, vol. 3, no. Iciba, pp. 659–663, 2023, doi: 10.1109/ICIBA56860.2023.10165501.

[22] C. Liang, Z. Zhang, X. Zhou, B. Li, S. Zhu, and W. Hu, "Rethinking the Competition Between Detection and ReID in Multiobject Tracking," *IEEE Trans. Image Process.*, vol. 31, pp. 3182–3196, 2022, doi: 10.1109/TIP.2022.3165376.

[23] M. L. Ali and Z. Zhang, "The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection," *Computers*, vol. 13, p. 336, 2024, doi: 10.3390/computers13120336.

[24] A. Awad and S. A. Aly, "Early Diagnosis of Acute Lymphoblastic Leukemia Using YOLOv8 and YOLOv11 Deep Learning Models," 2024, [Online]. Available: https://arxiv.org/pdf/2410.10701

[25] Y. Luo, Y. Du, Z. Wang, J. Mo, W. Yu, and S. Dou, "DScanNet: Packaging Defect Detection Algorithm Based on Selective State Space Models," *Algorithms*, vol. 18, no. 6, p. 370, 2025, doi: 10.3390/a18060370.

[26] W. Haolin, "CSDN: A Context-Gated Self-Adaptive Detection Network for Real-Time Object Detection," 2025, [Online]. Available: http://arxiv.org/abs/2506.17679

[27] H. Chen, G. Zhou, and H. Jiang, "Student Behavior Detection in the Classroom Based on Improved YOLOv8," 2023.

[28] R. Sapkota, Z. Meng, M. Churuvija, X. Du, Z. Ma, and M. Karkee, "Comprehensive performance evaluation of yolov12, yolo11, yolov10, yolov9 and yolov8 on detecting and counting fruitlet in complex orchard environments," *arXiv Prepr. arXiv2407.12040*, 2024.

[29] M. Tiwari, M. Patankar, V. Chaurasia, M. Shandilya, A. Kumar, and A. Potnis, "Detection of Tuberculosis Bacilli Using Deep Learning," *1st IEEE Int. Conf. Innov. High Speed Commun.*

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

Signal Process. IHCSP 2023, pp. 492–496, 2023, doi: 10.1109/IHCSP56702.2023.10127220.

[30] S. Aulia, A. B. Suksmono, T. R. Mengko, and B. Alisjahbana, "A Novel Digitized Microscopic Images of ZN-Stained Sputum Smear and Its Classification Based on IUATLD Grades," *IEEE Access*, vol. 12, pp. 51364–51380, 2024, doi: 10.1109/ACCESS.2024.3386208.

[31] H. A. Devon *et al.*, "A psychometric toolbox for testing validity and reliability," *J. Nurs. Scholarsh.*, vol. 39, no. 2, pp. 155–164, 2007, doi: 10.1111/j.1547-5069.2007.00161.x.

[32] Detection TB, "Tuberculosis 2 Dataset," Roboflow Universe.

[33] G. Jun *et al.*, "A Novel real-time arrhythmia detection model using YOLOv8," pp. 1–17.

[34] D. Grantor, "Application of Computer Vision for Underwater Litter Detection APPLICATION OF COMPUTER VISION FOR UNDERWATER LITTER DETECTION MASTER ' S THESIS," 2025.

[35] R.-W. Bello, P. A. Owolawi, E. A. van Wyk, and C. Tu, "Object Detection Algorithms for Digital Imaging Applications: A Review," *Image Sensors - Digit. Imaging Syst. Appl. [Working Title]*, no. Cv, 2025, doi: 10.5772/intechopen.1010205.

[36] M. Lei *et al.*, "SoftHGNN: Soft Hypergraph Neural Networks for General Visual Recognition," pp. 1–13, 2025, [Online]. Available: http://arxiv.org/abs/2505.15325

[37] Y. Li *et al.*, "Lite-YOLOv8: a more lightweight algorithm for Tubercle Bacilli detection," *Med. Biol. Eng. Comput.*, vol. 63, no. 1, pp. 195–211, 2025, doi: 10.1007/s11517-024-03187-9.

[38] S. Gupta, R. Jindal, C. L. Biji, and J. Dheeba, "Enhanced Tuberculosis Detection Using Deep Neural Network on Microscopic Images," in *Fifth Congress on Intelligent Systems*, S. Kumar, E. A. Mary Anita, J. H. Kim, and A. Nagar, Eds., Singapore: Springer Nature Singapore, 2025, pp. 173–188.

[39] R. Y. Ju and W. Cai, "Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm," *Sci. Rep.*, vol. 13, no. 1, pp. 1–15, 2023, doi: 10.1038/s41598-023-47460-7.

[40] X. Li *et al.*, "Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection," in *Advances in Neural Information Processing Systems*, 2020, pp. 1–14.

[41] R. Blakemore *et al.*, "A multisite assessment of the quantitative capabilities of the Xpert MTB/RIF assay.," *Am. J. Respir. Crit. Care Med.*, vol. 184, no. 9, pp. 1076–1084, Nov. 2011, doi: 10.1164/rccm.201103-0536OC.

[42] A. B. Witarto *et al.*, "AI-Based Analysis of Ziehl-Neelsen-Stained Sputum Smears for Mycobacterium tuberculosis as a Screening Method for Active Tuberculosis.," *Life (Basel, Switzerland)*, vol. 14, no. 11, Nov. 2024, doi: 10.3390/life14111418.

[43] A. W. Setiawan and M. I. Rusydi, "Detection of Mycobacterium Tuberculosis Using Residual Neural Network," *2022 Int. Semin. Intell. Technol. Its Appl. Adv. Innov. Electr. Syst. Humanit. ISITIA 2022 - Proceeding*, pp. 7–11, 2022, doi: 10.1109/ISITIA56226.2022.9855300.

[44] S. Aulia, A. B. Suksmono, T. R. Mengko, and B. Alisjahbana, "A Novel Digitized Microscopic Images of ZN-Stained Sputum Smear and Its Classification Based on IUATLD Grades," *IEEE Access*, vol. 12, pp. 51364–51380, 2024, doi: 10.1109/ACCESS.2024.3386208.

Author Biography

Syevana Dita Musvika is a dedicated faculty member at the Department of Electromedical Engineering, Health Polytechnic of the Ministry of Health Surabaya (Poltekkes Kemenkes Surabaya), where he contributes to both teaching and curriculum development. He earned his Diploma IV in Electromedical Engineering in 2012, demonstrating early proficiency in medical device technology. He further advanced his expertise by completing his Master's degree in Biomedical Engineering at Universitas Airlangga in 2024, with a research focus on deep learning applications for medical image analysis. His primary research interests encompass biomedical instrumentation, biosignal processing, medical image processing, and applied electromedical technologies. Beyond his academic duties, he is actively engaged in scholarly publishing and serves as a board member of the Indonesian Association of Electromedical Engineering Professionals (IKATEMI), where he helps advance the profession nationally. Through his teaching, research, and professional service, he is committed to improving healthcare technology in Indonesia.

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

Riries Rulaningtyas is an accomplished academic and researcher in the field of biomedical engineering, currently serving as an associate professor in the Biomedical Engineering Study Program, Department of Physics, Faculty of Science and Technology, Universitas Airlangga, Indonesia. She earned her bachelor's and master's degrees from the prestigious Institut Teknologi Sepuluh Nopember (ITS), Surabaya, where she built a strong foundation in engineering principles. She then pursued her Ph.D. in Biomedical Engineering at the School of Electrical Engineering and Informatics, Institut Teknologi Bandung (ITB), one of Indonesia's leading programs in the field. Her doctoral research further refined her expertise in advanced signal and image analysis techniques. Her primary research interests encompass medical signal processing, medical image processing, and artificial intelligence, with a particular focus on developing computational solutions for diagnostic challenges in tuberculosis and other diseases. Throughout her career, she has actively contributed to numerous research projects, supervised graduate students, and published extensively in international journals. She is also deeply committed to advancing biomedical engineering education in Indonesia, mentoring the next generation of researchers and practitioners. Her work has significantly contributed to the application of AI in medical diagnostics, particularly in resource-limited healthcare settings.

Khusnul Ain is a distinguished Professor in the Biomedical Engineering Study Program, Department of Physics, Faculty of Science and Technology, Universitas Airlangga, Indonesia. He has built an impressive and diverse academic foundation, beginning with his undergraduate studies in nuclear engineering, for which he received the S.T. degree from Gadjah Mada University, Indonesia, in 1995. He then expanded his expertise into the physical sciences, earning an M.Si. degree from the same university in 2002. Driven to deepen his engineering knowledge, he pursued a Doctoral degree in Engineering Physics from the prestigious Bandung Institute of Technology, Indonesia, which he completed in 2014. This unique combination of nuclear engineering, physical sciences, and engineering physics gives him a distinctive interdisciplinary perspective. His primary research interests include biomeasurement, electrical impedance, computational modeling, and bioinstrumentation, with a particular emphasis on developing non-invasive diagnostic tools. Throughout his career, Professor Ain has published extensively in peer-reviewed journals, supervised numerous graduate theses, and led various nationally funded research projects. He is widely respected for his contributions to advancing biomedical instrumentation in Indonesia and continues to be actively involved in both academic and professional communities, fostering innovation at the intersection of physics, engineering, and medicine.

Dr. Pepy Dwi Endraswari is a clinical microbiologist at Universitas Airlangga Hospital and a lecturer at the Department of Medical Microbiology, Universitas Airlangga, Indonesia. She is also a member of the Indonesian Association of Clinical Microbiologists (Perhimpunan Ahli Mikrobiologi Klinik Indonesia, PAMKI). Dr. Endraswari earned her medical degree from Universitas Brawijaya and completed both her Master's degree in Basic Medical Sciences and her clinical microbiology residency at Universitas Airlangga. Her research interests include fungal infections, antimicrobial resistance, and the development of antifungal agents from natural products. She is particularly interested in biofilm formation by Candida species and has conducted studies on anti-biofilm strategies. In addition, she is involved in tuberculosis research, focusing on the relationship between TB and opportunistic infections.

Savira Hayyun Audina is an ambitious undergraduate student pursuing a degree in Biomedical Engineering at Universitas Airlangga, Indonesia. She has developed a strong passion for integrating cutting-edge technology with healthcare services to create effective and accessible solutions that ultimately improve human well-being. Her academic training has equipped her with a solid foundation in both engineering principles and medical sciences, enabling her to tackle complex, interdisciplinary problems that lie at the intersection of these two critical fields. Beyond her coursework, she is actively involved in research projects that explore the application of artificial intelligence and medical imaging for diagnostic purposes. She is also deeply committed to collaboration, seeking opportunities to work alongside experienced professionals and forward-thinking organizations at the forefront of healthcare technology. By contributing to innovative projects that push the boundaries of what is possible in medical diagnostics and treatment, Savira aims to play a meaningful role in shaping the future of healthcare. Her

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111

dedication, curiosity, and drive position her as a promising emerging researcher in the biomedical engineering community.

Annie Anak Joseph is a respected senior lecturer in the Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Malaysia Sarawak (UNIMAS), where she has been contributing to academic and research excellence since 2006. She began her academic journey by earning her bachelor's degree in Electronic and Electrical Engineering from College University Technology Tun Hussein Onn, Malaysia, in 2005. Demonstrating remarkable academic dedication, she immediately pursued graduate studies and received her Master of Science (M.S.) degree from Universiti Sains Malaysia in 2006. She then embarked on her doctoral studies abroad, earning her Doctor of Engineering (Dr.Eng.) degree in Electrical and Electronic Engineering from the prestigious Kobe University, Japan. Following the completion of her Ph.D., she was promoted to senior lecturer in 2014, recognizing her growing contributions to research and teaching. Her research interests lie at the forefront of computational intelligence, including online learning, concept drift, feature extraction, and machine learning. She is particularly focused on developing adaptive algorithms that can handle dynamic, real-world data streams. Throughout her career, Dr. Annie has supervised numerous graduate students, secured research funding, and published extensively in international journals and conferences. She is also actively involved in collaborative research networks across Southeast Asia and Japan, fostering cross-border innovation in machine learning applications.

Manuscript received 2 August 2025; Revised 10 February 2026; Accepted 24 May 2026; Available online 1 July 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1111