Journal of
Information Systems Engineering
and Business Intelligence

Vol.11, No.2, June 2025
Available online at: http://e-journal.unair.ac.id/index.php/JISEBI

Classification and Counting of Mycobacterium Tuberculosis using YOLOv5

Nia Saurina [1], Nur Chamidah [2]*, Riries Rulaningtyas [3], Aryati Aryati [4]

[1]*Doctoral Study Program of Mathematics and Natural Science, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia*
[1]nia.saurina-2022@fst.unair.ac.id

[1]*Department of Informatic, Faculty of Engineering, Universitas Wijaya Kusuma, Surabaya, Indonesia*
[1]nia.saurina-2022@fst.unair.ac.id

[2]*Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia*
[2]nur-c@fst.unair.ac.id

[2]*Research Group of Statistical Modeling in Life Science, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia*
[2]nur-c@fst.unair.ac.id

[3]*Department of Physics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia*
[3]rulaningtyas@gmail.com

[4]*Department of Clinical Pathology, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia*
[4]aryati@fk.unair.ac.id

Abstract

Background: Indonesia is a nation with the third-highest number of tuberculosis (TB) cases worldwide, after China and India. TB detection has been facilitated using YOLOv5 deep learning framework despite previous studies not having incorporated assessment metrics recommended by International Union Against Tuberculosis and Lung Disease (IUATLD).
Objective: This study aims to present a method for classifying and enumerating *Mycobacterium tuberculosis* by using YOLOv5 architecture with IUATLD evaluation standards. Sputum samples served as the primary medium for identifying the presence of *Mycobacterium tuberculosis*. In addition, the method showed precise delineation of bacterial boundaries to minimize classification inaccuracies and improve edge clarity through YOLOv5.
Methods: Following the acquisition of microscopic images of TB, the data were resized from 1632x1442 to 640x480 pixels. Annotation was performed using YOLOv5 bounding boxes, and the model was subsequently trained as well as tested according to IUATLD guidelines.
Results: During the analysis, YOLOv5-based classification system produced optimal performance. The model achieved 84.74% accuracy, 87.31% precision, and Mean Average Precision (mAP) score of 84.98%. These metrics showed high reliability in identifying *Mycobacterium tuberculosis* in the image dataset.
Conclusion: The classification and quantification of *Mycobacterium tuberculosis* using YOLOv5 framework shows high precision, with mAP score of 84.98%, signifying strong model performance. Additionally, the counting process achieves a MAPE (Mean Absolute Percentage Error) of 0.15%, reflecting excellent prediction accuracy.

Keywords: IUATLD, Tuberculosis, YOLOv5.

Article history: Received 3 Oktober 2024, first decision 31 January 2025, accepted 24 June 2025, available online 22 July 2025

I. INTRODUCTION

High-burden tuberculosis (TB) countries in 2019 were 87% of global TB diagnoses, with eight nations, which is India, Indonesia, China, Philippines, Pakistan, Nigeria, Bangladesh, and South Africa, making up nearly two-thirds of

* Corresponding author

ISSN 2443-2555 (online) 2598-6333 (print)
doi: http://dx.doi.org/10.20473/jisebi.11.2.267-278

these cases. Indonesia reports approximately 250,000 new TB cases each year, with around 100,000 of these leading to death [1]. TB remains the top infectious killer in the country and ranks as the third most common cause of death, after heart disease and acute respiratory infections. Globally, Indonesia is the third most affected country by TB behind China and India. In 2019, the country reported an estimated 842,000 TB cases, including 36,000 co-infected with HIV (Human Immunodeficiency Virus) [2]. Among HIV-negative individuals, TB was responsible for around 107,000 deaths, while 9,400 deaths occurred among those living with HIV [3]. The illness significantly strains the national economy, particularly since 75% of those impacted are in the economically productive age group of 15 to 50 [4]. This leads to job losses and reduced productivity for both patients as well as caregivers. Economic consequences stem from factors such as early mortality, prolonged illness, and related medical as well as non-medical expenditures [5] [6]. According to WHO (World Health Organization) figures from 1999, Indonesia recorded about 583,000 new TB cases annually at that time, with about 140,000 deaths per year [7].

TB is a transmissible disease predominantly caused by *Mycobacterium tuberculosis* [8][9]. Sputum smear microscopy has been recognized as an efficient method for TB detection, facilitating early diagnosis and contributing to disease control efforts [10]. However, accurate TB diagnosis remains complex, as most current methods depend on recognizing the immune reaction of the host rather than directly isolating the pathogen [11][12]. Speedier diagnostic methods are crucial to ensure prompt therapy, which is major to stopping further transmission [13]. Identifying individuals with active TB is important for curbing the spread of the infection and strengthening intervention strategies [14][15].

A wide array of studies have adopted YOLOv5 architecture for object detection and classification tasks. For example, study by [16] improved the recognition of small-scale targets in satellite imagery by incorporating YOLOv5 with Region-based Fully Convolutional Networks (R-FCN). Similarly, results from [17] showed that augmenting YOLOv5 with complementary models significantly increased classification accuracy. In [18], YOLOv5 framework was paired with Local Fully Convolutional Neural Network (LFCN) to better capture distant, minute objects. You Only Look Once (YOLO) model family has also been used in TB diagnostics. According to [19], TB detection system using YOLOv7 was built through a convolutional neural network (CNN) for efficient single-pass object detection. Another study by [20] introduced a unique method by modifying a hybrid network capable of extracting features from pretrained CNN and Vision Transformers (ViT), which were then processed through a dual-layer fully connected architecture based on YOLOv3. Despite these advancements, none of the mentioned methods incorporated evaluation strategies grounded in the guidelines of International Union Against Tuberculosis and Lung Disease (IUATLD). Since 1994, IUATLD, in partnership with WHO, has endorsed unified anti-TB drug treatment protocols aimed at simplifying therapy, improving adherence, and minimizing errors in medication administration [21]. Therefore, applying IUATLD scoring standard in TB sputum diagnostics is crucial to achieving reliable and standardized clinical assessments [22].

Fig. 1 shows the structure of YOLOv5 model [16], which is composed of three essential parts, namely Backbone, PANet, and Output modules. Backbone functions as a feature extractor from the input image, using BottleNeckCSP blocks to maintain critical information while improving computational performance. Spatial Pyramid Pooling (SPP) mechanism is incorporated to capture features at various scales, increasing detection reliability. PANet (Path Aggregation Network) component refines the features extracted by Backbone through several stages for classification.

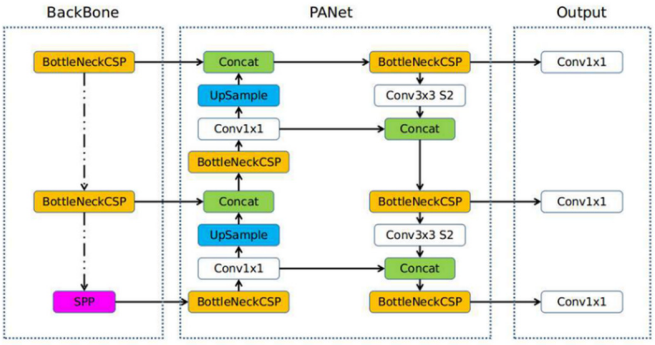

Fig. 1. Methodology of Classifying and Counting *Mycobacterium Tuberculosis*

This study proposes a classification and quantification system for *Mycobacterium tuberculosis* using YOLOv5 architecture to solve the pending issue according to the evaluation guidelines set by IUATLD. The analysis centers on sputum specimens as the principal source for bacterial detection, since this method accurately identifies the predicted

bacterial regions. Through the capabilities of YOLOv5, the framework is designed to reduce misclassification and improve the sharpness as well as reliability of object boundary recognition in the classification workflow.

The analysis contributes by addressing two major questions, namely how can YOLO architecture be applied to classify TB from sputum samples, and how IUATLD standard is used for counting *Mycobacterium tuberculosis*, respectively.

The layout of this manuscript is arranged as follows, where Section II delivers an overview of existing studies, and Section III explains the proposed methods in depth. In addition, Section IV describes the results and offers an interpretation of the experimental data, and Section V concludes the study.

II. METHODS

The architecture of the proposed YOLO model was shown in Fig. 2. Initially, microscopic images of TB were gathered and resized from the original resolution of 1632×1442 to 640×480 pixels. These images were then annotated with bounding boxes following YOLO framework. The dataset passed through a training and detection phase aimed at classifying *Mycobacterium tuberculosis*. In the final stage, the number of detected bacteria was quantified based on the output of the model and evaluated using IUATLD assessment guidelines.

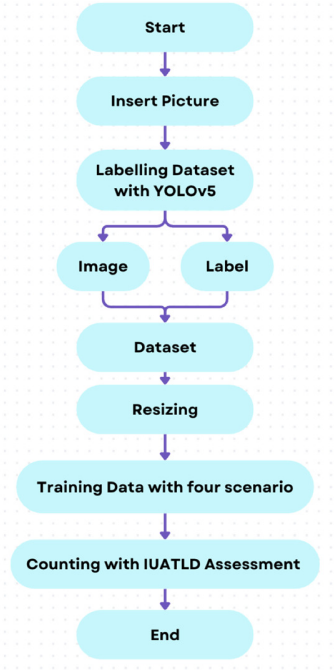

Fig. 2. Methodology of Classify and Counting *Mycobacterium tuberculosis*

A. Data Collection

This study used a dataset comprising images of *Mycobacterium tuberculosis* provided by the Clinical Pathology Department of Dr. Soetomo Hospital in partnership with Airlangga University. A total of 1,265 samples were collected to support the training phase of YOLO-based detection model. These images were obtained through Ziehl-Neelsen staining, a standard method for visualizing acid-fast bacilli in TB diagnosis. The dataset was curated specifically for model training and included a balanced split for both training as well as validation purposes. During the process, ethical approval for data use was granted under exemption reference number 53/EC/KEPK/FKUA/2023.

B. Insert Picture

During the process, images used in this study were inserted into the model. The images contained the location of *Mycobacterium tuberculosis* coordinates in the form of the x and y axes. This step was performed to gather all the visual data processed and analyzed in subsequent steps. Following the process, ensuring that the images were correctly inserted was crucial for the accuracy and reliability of the study.

C. Resize

After importing the images, adjustment was conducted to produce a standardized resolution. This resizing process ensured uniformity across all images, which was crucial for reliable analysis and model performance. By standardizing the dimensions of the images, the data became compatible with YOLOv5 framework, facilitating effective processing. This step was critical for preserving data consistency and enabling the accurate detection as well as quantification of *Mycobacterium tuberculosis* by the model.

D. Labelling Data

The annotation process, namely labelling, included assigning descriptive metadata to identify biological features, such as structural patterns shown in the dataset. The images were manually labeled to capture the precise coordinates of bounding boxes representing ground truth, which later served as a benchmark against the predictions generated by the model. Following the process, model training was performed over 100 epochs using a batch size of 16, with ongoing monitoring of both the loss function and performance metrics. Fig. 3 shows that YOLO model was applied to detect *Mycobacterium tuberculosis* in individual frames of images by generating bounding boxes localizing the bacteria. These boxes were defined by coordinate points (X, Y), serving as classification labels. The quantity of annotations per input varied, depending on how single or multiple overlapping bacterial regions were present. During the analysis, the labeling phase was completed in five working days with guidance from a lab technician at the Clinical Pathology Department of Dr. Soetomo Hospital in Surabaya.

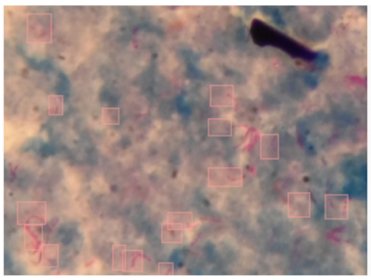

Fig. 3. Dataset Labelling with Bounding Box

E. Proposed Classification Method

The dataset annotated in the previous step served as the foundation for training the model to detect patterns, producing optimized weights for object identification in visual data. This process was conducted using YOLO algorithm. Model development and execution were conducted in PyCharm incorporated development environment. During the analysis, training included images specifically labeled to show *Mycobacterium tuberculosis* lesions. In the process, model parameters were iteratively adjusted to minimize classification errors, which signified the disparity between actual labels and model predictions. The accuracy of the annotated bounding boxes was reviewed and verified by two subject matter experts, namely the Head of the Clinical Pathology Laboratory and a professor from the Clinical Pathology Department.

YOLOv5 used a combination of loss functions during the training phase, including those for classification, bounding box precision (localization), and objectness confidence, which estimated the probability concerning the presence of an object [23]. Despite YOLO family occasionally producing higher error rates than some other detection models, its ability to recognize object-like features in background regions contributed to strong generalization across varying object categories [24], [25], [26]. This version of YOLOv5 included several architectural upgrades over predecessors [27], [28], and was organized into three primary components, namely backbone, neck, and head. Backbone, which was built on the Darknet-53 framework handled feature extraction. In addition, neck improved these features by combining spatial and contextual details from various resolutions [16]. Head contained three distinct branches that generated multi-scale predictions. To refine detection results, the model also applied methods to eliminate overlapping bounding boxes [29].

The model known as YOLOv5, a refined successor in YOLO family included several critical developments that improved its performance in detection and classification tasks. Optimized for real-time operations, it offered fast inference speeds with minimal impact on accuracy [16], [30]. The advanced feature extraction of this model was powered by components such as Spatial Pyramid Pooling-Fast (SPPF) layer, which helped the tool capture multi-scale characteristics effectively [24]. Additionally, the model was also adaptable and lightweight, allowing users to select versions modified to specific accuracy as well as hardware requirements, on high-performance systems or constrained devices [31], [32]. Another feature was its flexible structure, which supported easy modifications for custom detection

needs [33]. YOLOv5 had been successfully applied across various sectors, such as detecting objects in dim lighting, identifying small targets in aerial views, and tracking motion in sports scenarios [24].

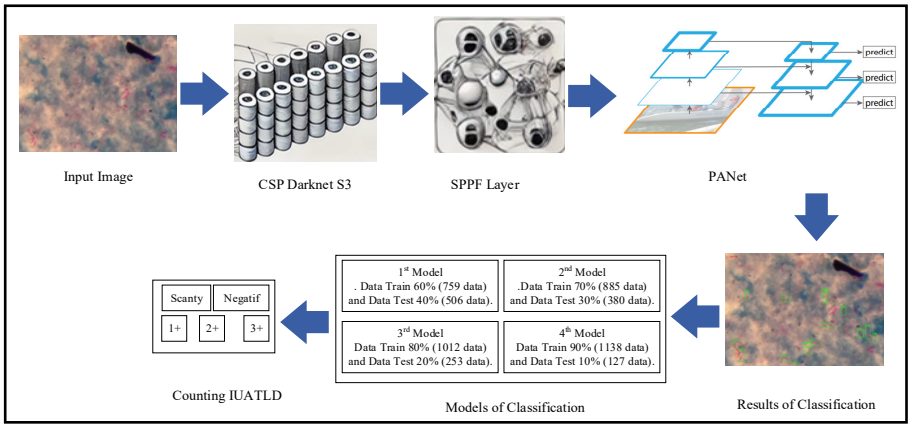

Fig. 4. Tuberculosis Classification Model

Fig. 4 shows the architecture of TB Classification system during the study. Input Image was the initial phase, where images were introduced into YOLOv5 framework for object detection and classification. Following the process, backbone Network (CSPDarknet53) as the foundational feature extractor processed input images to generate informative feature maps. The use of Cross Stage Partial (CSP) Networks in this process improved computational efficiency and minimized redundant operations. SPPF (Spatial Pyramid Pooling Fast) Layer was used to capture contextual information across multiple spatial scales, aiding in the recognition of objects of varying sizes in the images. In addition, Neck Component (PANet – Path Aggregation Network) combined feature maps from different resolution levels. By aggregating multi-scale data, the tool ensured that the final feature representations used for detection were both comprehensive and relevant. The model generated predictions in the form of bounding boxes, which showed object locations along with class probabilities identifying the object types to provide information crucial for the subsequent quantification of *Mycobacterium tuberculosis*. In line with this discussion, classification Models could evaluate five different IUATLD grouping models. IUATLD-Based Counting included categorizing patients by the number of TB bacilli present. During the study, the proposed model incorporated YOLOv5 with IUATLD-based quantification standards to perform both classification and enumeration of *Mycobacterium tuberculosis*.

F. Classification using YOLO

During the training phase, the model was developed using a dataset containing labeled images that marked the presence of *Mycobacterium tuberculosis*. Fig. 5 shows that the training output included samples processed with YOLOv5 model. Each detection was assigned a confidence value between 0 and 1, showing how certain the model was about the presence of the bacterium in the predicted bounding box, as higher values represented greater confidence. These scores were crucial for eliminating low-certainty predictions and ensuring that only detections meeting a predefined threshold were considered for evaluation. During the process, YOLOv5 was used to assign a probability score to each detection, estimating the possibility that a region contained *Mycobacterium tuberculosis*. Predictions with higher confidence were considered more trustworthy for both classification and counting purposes. To ensure accurate detection and quantification, specific evaluation criteria were applied, considering bounding box precision, confidence levels, and object localization accuracy. Moreover, several major performance metrics were used to evaluate the effectiveness of the model. These metrics included Accuracy, reflecting the rate of correct detections, Precision, measuring the proportion of true positives out of all positive results, and Recall, which showed the sensitivity of the model in identifying real cases of *Mycobacterium tuberculosis*. Others included mAP (Mean Average Precision), which reviewed performance across varying confidence levels, and MAPE (Mean Absolute Percentage Error), assessing count precision using IUATLD guidelines. These metrics were calculated using verified statistical methods to ensure reliability. Additionally, senior pathologists manually validated the results, comparing the outcome with IUATLD guidelines to strengthen the credibility of the results.

G. Counting Mycobacterium tuberculosis

In this study, data splitting was conducted using the holdout validation method, commonly known as data partitioning [34]. The method automatically separated the dataset into three groups, namely training, validation, and

testing subsets. Independent test data played a crucial role in classification workflows, allowing accurate evaluation of model performance after training. This method helped determine how well the model could adapt to previously unobserved samples [35]. The dataset comprised 1,265 images, divided into three subsets and saved in designated directories. During training, the model was trained to learn object detection patterns from these labeled inputs. To assess performance across diverse configurations, four partitioning schemes were explored. Scenario 1 assigned 60% (759 images) to training, and 20% each to validation (253) as well as testing (253). Additionally, Scenario 2 used 70% for training (1,096 images), with the remaining 15% allocated to validation (60) as well as testing (62). In Scenario 3, 80% of the data (1,096 images) was used for training, 10% for validation (62), and 15% for testing (60). Lastly, Scenario 4 applied 90% which was equivalent to 1,138 images for training, with 0.5% each for validation (127) and testing (127).

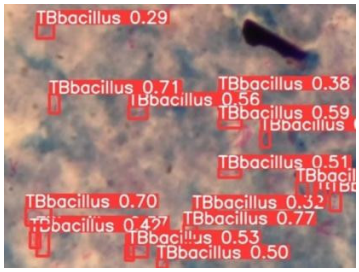

Fig. 5 Result of training the Image of *Mycobacterium tuberculosis* with YOLOv5

The quantification of *Mycobacterium tuberculosis* was derived from the classification output, implemented across four distinct TB counting scenarios. The grading of bacilli present in sputum samples followed the standards established by IUATLD. According to these guidelines, the categories were arranged in the following order. Negative was no acid-fast bacilli (AFB) shown across 100 microscopic fields, 1+ represented detection of 10 to 99 AFBs in 100 fields, 2+ was between 1 to 9 AFBs per field across at least 50 fields, and 3+ represented more than 10 AFBs observed in each field across a minimum of 20 fields. Participant eligibility in this study was limited to newly diagnosed smear-positive pulmonary TB (Category I) patients aged 17 to 60 who were initiating standard first-line anti-TB therapy. Following the discussion, individuals diagnosed with extrapulmonary TB were excluded from this evaluation [36]

H. Evaluation

During the evaluation stage, the system measured how well the trained model performed when applied to previously unobserved data excluded from the training process. This testing was conducted only after the model had passed through full training. Moreover, accuracy was assessed using mAP metric—a widely recognized benchmark for evaluating object detection models, including Regional (R)-CNN, Faster R-CNN, SSD (Single Shot Multibox Detector), and similar architectures [37].

This study used three primary metrics to evaluate model performance, namely Precision, Recall, and mAP. Both mAP and F1-Score were used to assess the effectiveness of the model. Moreover, mAP was widely adopted in object detection tasks as it quantified the ability of the model to correctly locate and classify objects in images. It was derived by calculating Average Precision (AP) for each category and averaging these results [24]. The computation of mAP followed the formulation described in Equation (1).

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

where:
N : Number of *AP*
: Average precision
Equations (2), (3), and (4) showed *A*P.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$F1 - Score = \frac{2xPrecisionxRecall}{Precision+Recall} \quad (5)$$

Where:
TP : True Positive
TN : True Negative
FP : False Positive
FN : False Negative

During the analysis, mAP evaluation standard was shown in Table 1. A score of more than 70% signified that the model achieved high accuracy. Scores ranging between 50% and 70% showed the model performed at an acceptable level. On the other hand, mAP value less than 50% reflected poor model performance.

TABLE 1
MAP EVALUATION STANDARD [36]

mAP (%)	Value
mAP > 70%	Extremely Accurate: The model performed with high accuracy in object detection and classification.
50% ≤ mAP ≤ 70%	Decent Accuracy: The model showed good performance generally, but it was not yet fully optimized.
mAP < 50%	Unreliable Accuracy: The model showed significant difficulties in correctly detecting and classifying objects.

Several statistical metrics were used to evaluate model accuracy, including Mean Absolute Deviation (MAD), Mean Squared Error (MSE), and MAPE. Among these, MAPE was widely recognized as a preferred metric for assessing forecast performance. The model evaluation standard included values less than 10% showed highly accurate predictions, and values between 10% and 20% signified good accuracy. A range of 20% to 50% reflected moderate accuracy, and values exceeding 50% were considered poor forecasts, as shown in Table 2 [38].

TABLE 2
MAPE EVALUATION STANDARD [38]

MAPE (%)	Value
MAPE < 10%	Highly Accurate Forecast
10 - 20%	Good Forecast
20 - 50%	Reasonable Forecast
MAPE > 50%	Inaccurate Forecast

III. RESULTS

Data partitioning included separating the dataset into several subsets to facilitate both training and evaluation of the model. These different configurations aimed to identify which split achieved the highest mAP, serving as an indicator of the optimal performance setup of the model.

TABLE 3
CONFUSION MATRIX OF CLASSIFICATION *MYCOBACTERIUM TUBERCULOSIS* WITH YOLOV5

Data Train	True Positive (TP)	False Positive (FP)	False Negative (FN)	True Negative (TN)
Data Train 60%	398	92	98	171
Data Train 70%	413	60	75	337
Data Train 80%	510	75	80	347
Data Train 90%	547	98	95	398

Table 3 shows the confusion matrices for *Mycobacterium tuberculosis* classification task across four experimental scenarios. Using Equations (1) to (4), the analysis computed the evaluation metrics, namely accuracy, precision, recall, F1 score, and mAP. In 1st scenario, the model produced 398 true positives (TP), 92 false positives (FP), 98 false negatives (FN), and 171 true negatives (TN), respectively. The 2nd recorded 413 TP, 60 FP, 75 FN, and 337 TN. The 3rd scenario had 510 TP, 60 FP, 75 FN, and 347 TN. Lastly, 4th scenario achieved 547 TP, 98 FP, 80 FN, and 171 TN.

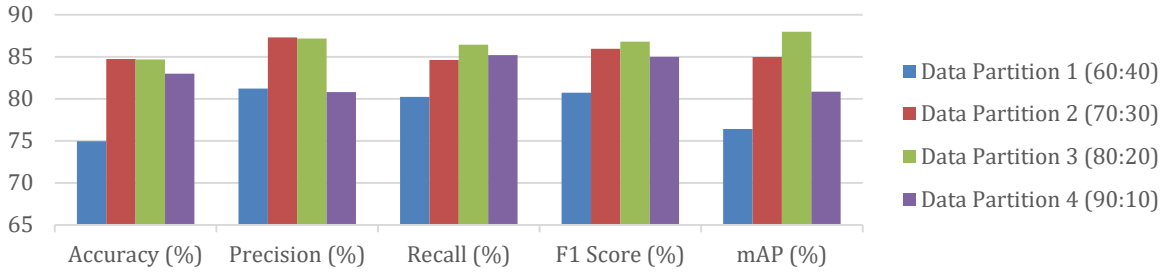

Fig. 6. Results from Dataset Split Evaluation

Fig. 6 shows the results of four different testing scenarios, each associated with distinct performance metrics. The 1st through 4th scenario achieved accuracy scores of 74.96%, 84.74%, 84.68%, and 83%. Precision was 81.22%, 87.31%, 87.17%, and 84.80%. The recall reached 80.24%, 84.63%, 86.44%, and 85.20%. Corresponding F1-scores were 80.73%, 85.95%, 86.80%, and 85%, while mAP scores obtained were 76.41%, 84.98%, 87.98%, as well as 80.85%.

Fig. 7 shows the training graph from the classification of *Mycobacterium tuberculosis* using YOLOv5 model. Train/Box_Loss curve signified a steady downward trend, decreasing from approximately 2.0 to 1.5, showing improved accuracy in drawing bounding boxes. Similarly, Train/Cls_Loss curve dropped consistently from around 1.6 to 0.8, showing improved model performance in object classification. The consistent reduction in both loss curves showed stable learning without major fluctuations. This showed a strong sign that the model was training effectively without significant signs of overfitting or underfitting. Furthermore, the validation curves—val/box_loss and val/cls_loss—also followed a similar downward trajectory, reflecting good generalization to unseen data. This grouping between training and validation losses confirmed that the model was learning from the training data and adapting well to external datasets.

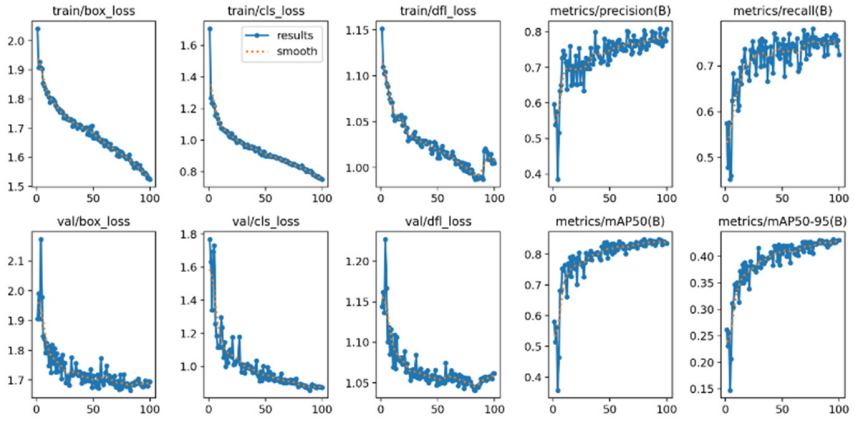

Fig. 7 Training Graph of *Mycobacterium tuberculosis* Image Classification Using YOLOv5

IV. Discussion

Most existing studies focused on detecting and diagnosing TB to determine its presence or absence, without following the evaluation process with IUATLD assessment standards. In observation, *Mycobacterium tuberculosis* appeared as elongated, red-colored organisms against a blue background. However, accurate interpretation of sputum microscopy results should follow IUATLD grading system [24]. IUATLD offered a globally accepted framework with standardized criteria for diagnosing and categorizing TB. Incorporating these guidelines improved the validity and comparability of results across studies. Additionally, IUATLD method showed comprehensive benchmarks for identifying and classifying TB bacilli. Embedding these standards into YOLOv5 detection model could significantly refine the classification and enumeration process, leading to improved precision when identifying *Mycobacterium tuberculosis* in diagnostic imagery.

The results of the acid-fast bacilli examination in principle observed TB microorganisms, namely *Mycobacterium tuberculosis*, as the cause of TB under a microscope after staining. This allowed the examination data to be obtained

in the form of the number of sputum TB adjusted to IUATLD scale represented in negative and positive (1+, 2+, 3+), as shown in Table 4 [39].

TABLE 4
SCALE IUATLD [38]

Score	Criteria
Negative	No acid-fast bacilli were detected across a minimum of 100 observed microscopic fields.
Scanty	1-9 acid-fast bacilli were found in 100 visual fields
1+	10-99 acid-fast bacilli were found in 100 visual fields
2+	1-10 acid-fast bacilli were found in 50 visual fields
3+	More than 10 acid-fast bacilli were found in 20 visual fields.

IUATLD incorporation or calculation into the prediction results from YOLO had transformed through these structured steps: First was raw YOLOv5 Classification, where the model detected *Mycobacterium tuberculosis* using bounding boxes and confidence scores. Second was AFB Count Extraction, where the output from YOLOv5 was processed to count the number of Acid-Fast Bacilli (AFB) in microscopic images. Third step was IUATLD Score Mapping to detect AFB was translated into IUATLD grades using the following criteria in Table 4.

During the process, IUATLD classification was applied to a dataset consisting of 1,265 images, which were collected from 13 different patients. Each set of approximately 100 images represented sputum sample results from an individual patient. Consequently, the dataset enabled IUATLD-based classification for a total of 13 patients. IUATLD assessment shown in Table 4 was according to the classification output derived from the second scenario, which produced the highest accuracy among all scenarios.

The following reviewed IUATLD classification outcomes for each patient based on the number of visual fields falling into specific scoring categories. The 1st Patient was 1+ (98), 2+ (70), 3+ (28), IUATLD Score was 3+. The 2nd Patient was 1+ (98), 2+ (76), 3+ (22), IUATLD Score was 2+. In addition, the 3rd Patient had 1+ (100), 2+ (74), 3+ (26), with IUATLD Score of 3+. The 4th Patient was 1+ (98), 2+ (78), 3+ (20), having IUATLD Score of 2+. The 5th Patient was 1+ (96), 2+ (66), 3+ (30), IUATLD Score was 3+. The 6th Patient had 1+ (96), 2+ (66), 3+ (30), with IUATLD Score of 3+. The 7th Patient was 1+ (99), 2+ (69), 3+ (30), having IUATLD Score of 3+. The 8th Patient had 1+ (98), 2+ (68), 3+ (30), IUATLD Score was 3+. The 9th Patient was 1+ (99), 2+ (82), 3+ (17), IUATLD Score was 2+. The 10th Patient had 1+ (98), 2+ (73), 3+ (25), with IUATLD Score of 3+. The 11th Patient was 1+ (98), 2+ (75), 3+ (23), having IUATLD Score of 3+. The 12th Patient was 1+ (98), 2+ (70), 3+ (28), IUATLD Score was 3+. The 13th Patient had 1+ (37), 2+ (27), 3+ (10), with IUATLD Score of 1+. These results showed IUATLD-based classification effectiveness when incorporated with YOLOv5 model, enabling accurate bacillary quantification per patient.

TABLE 5.
IUATLD OUTCOMES FROM THE SECOND DATA SPLIT EVALUATION

NO	Number of Visual Fields with Condition 1+	Number of Visual Fields with Condition 2+	Number of Visual Fields with Condition 3+	SCORE IUATLD
1	98	70	28	3+
2	98	76	22	3+
3	100	74	26	3+
4	98	78	20	3+
5	97	73	24	3+
6	96	66	30	3+
7	99	69	30	3+
8	98	68	30	3+
9	99	82	17	2+
10	98	73	25	3+
11	98	75	23	3+
12	98	70	28	3+
13	37	27	10	1+

Among the 13 patients analyzed, 1 individual was categorized as 1+, 7 were classified as 2+, and 5 fell into the 3+ category according to IUATLD classification results, as shown in Table 5. Referring to [39], MAPE score of 84.98% signified exceptionally accurate forecasting performance. In this study, MAPE value was derived by comparing *Mycobacterium tuberculosis* counts identified by the proposed detection model with those recorded manually by a laboratory assistant from the Clinical Pathology Department, Dr. Soetomo Hospital, Universitas Airlangga. The discrepancy between the predicted and actual values, commonly referred to as the residual, was used in calculating

forecasting accuracy. A frequently used method for this was MAPE, which expressed the average error as a percentage. During the analysis, MAPE was calculated by taking the absolute error percentages across all observations and averaging the outcome over the total number of measured periods [24]. The mathematical formula used to compute MAPE was provided in Equation (5).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{Ai-Fi}{Ai} \right| x100\% \quad (5)$$

Information:
n : Sample Size
Ai : Actual Data Value
Fi : Forecasting Data Value

Table 6 shows that the 1st and 2nd scenario had the same mAP value of 76.41% and 0.46% MAPE. The 3rd scenario had mAP value of 87.98% and MAPE of 0.38% while the 4th achieved 80.85% mAP as well as 0.46% MAPE.

TABLE 6.
RESULTS OF MAPE

No	YOLO	Scenario	mAP (%)	MAPE (%)
1	YOLOv5	1st scenario	76.41%	0.46%
2		2nd scenario	84.98%	0.15%
3		3rd scenario	87.98%	0.38%
4		4th scenario	80.85%	0.46%

This study used two evaluation metrics, namely mAP for identifying the most accurate classification performance, and MAPE for determining the most reliable IUATLD calculation outcome. Table 1 shows that the second scenario achieved the highest mAP score at 84.98%, signifying it provided the most effective classification results. However, this study also represented the importance of IUATLD-based evaluation. According to Table 2, MAPE value less than 10% signified highly accurate predictions. The 3rd scenario produced the lowest MAPE value of 0.15%, showing it achieved the most precise IUATLD forecasting outcome. Therefore, as the 3rd scenario was observed in classification, the 2nd scenario showed superior accuracy in TB grading.

V. CONCLUSIONS

In conclusion, this study used YOLO framework to identify *Mycobacterium tuberculosis*. The dataset was initially labeled under a single classification class, and after labeling, a preprocessing step was conducted including the arrangement of images to guide the detection of *Mycobacterium tuberculosis* in each frame. The annotated dataset was then used to train the model, generating weight values essential for object detection. Among the four partitioning strategies evaluated, the second scenario produced the best performance, achieving 84.74% accuracy, 87.31% precision, and mAP of 84.98%. In addition to this outcome, the elevated mAP reflected the strong detection capability of the model. When assessment was conducted using MAPE, the system showed a low error rate of 0.15%, signifying high prediction accuracy. Based on IUATLD grading applied to 13 patient cases, results showed 1 case classified as 1+, 7 as 2+, and 5 categorized under the 3+ grade.

Author Contributions: *Nia Saurina*: Conceptualization, Data Curation, Writing - Original Draft, Writing - Review & Editing, Supervision. *Nur Chamidah*: Conceptualization, Investigation, Methodology. Validation *Riries Rulaningtyas*: Investigation, Data Curation, Supervision. *Aryati Aryati*: Investigation, Data Curation.

All authors have read and agreed to the published version of the manuscript.

Funding: This study received financial support from Beasiswa Pendidikan Indonesia (BPI) under PUSLAPDIKTI, as specified in the individual decision letter No. 00170/BPPT/BPI.06/9/2023 for the year 2023.

Conflicts of Interest: The first author is a doctoral student at Airlangga University and a BPI scholarship recipient.

Data Availability: Data not available due to ethical restrictions.

Informed Consent: There were no human subjects.

Institutional Review Board Statement: This research has been approved to be carried out at RSUD Dr. Soetomo Surabaya with the issuance of a letter of ethical suitability issued by the Regional Hospital Association of Clinical Pathology and Laboratory Medicine Specialists (PATKLIN). Dr Soetomo Surabaya - Faculty of Medicine, Universitas Airlangga. The letter was issued on February 21 2023, with number 53/EC/KEPK/FKUA/2023. Researchers ensure that the security and confidentiality of the data obtained in this research is guaranteed, both when storing the data and when publishing the research results.

Animal Subjects: There were no animal subjects.

ORCID:
Nia Saurina: https://orcid.org/0000-0001-5715-9037
Nur Chamidah: https://orcid.org/0000-0003-1592-4671
Riries Rulaningtyas: https://orcid.org/0000-0001-7058-1566
Aryati Aryati: https://orcid.org/0000-0001-9055-0177

REFERENCES

[1] C. B. Kartasasmita, "Epidemiologi Tuberkulosis," *Sari Pediatri*, vol. 11, no. 2, p. 124, Nov. 2016, doi: 10.14238/sp11.2.2009.124-9.
[2] Abigail Wright and Matteo Zignol, "Anti-Tuberculosis Drug Resistance in the World: Fourth Global Report : The Who/Iuatld Global Project on Anti-Tuberculosis Drug Resistance Surveillance," 2009.
[3] Indonesian Ministry of Health, "Profil Kesehatan Indonesia 2022 ," 2023.
[4] WHO, "Global Tuberculosis Report 2024," 2024.
[5] Kementerian Kesehatan RI, "GERMAS Gerakan Masyarakat Hidup Sehat," 2017.
[6] Erlina Burhan, *Pedoman Nasional Pelayanan Kedokteran Tata Laksana Tuberkulosis*. Kemenkes RI, 2020.
[7] D. Collins, F. Hafidz, and D. Mustikawati, "The economic burden of tuberculosis in Indonesia," *The International Journal of Tuberculosis and Lung Disease*, vol. 21, no. 9, pp. 1041–1048, Sep. 2017, doi: 10.5588/ijtld.16.0898.
[8] N. Chamidah, Y. S. Yonani, E. Ana, and B. Lestari, "Identification the number of Mycobacterium tuberculosis based on sputum image using local linear estimator," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 5, pp. 2109–2116, Oct. 2020, doi: 10.11591/eei.v9i5.2021.
[9] K. K. Chopra and S. Singh, "Tuberculosis: Newer diagnostic tests: Applications and limitations," *Indian Journal of Tuberculosis*, vol. 67, no. 4, pp. S86–S90, Dec. 2020, doi: 10.1016/j.ijtb.2020.09.025.
[10] T. Venkatappa *et al.*, "Association of Mycobacterium tuberculosis infection test results with risk factors for tuberculosis transmission," *J Clin Tuberc Other Mycobact Dis*, vol. 33, p. 100386, Dec. 2023, doi: 10.1016/j.jctube.2023.100386.
[11] T. N. Susilawati and R. Larasati, "A recent update of the diagnostic methods for tuberculosis and their applicability in Indonesia: a narrative review," *Medical Journal of Indonesia*, vol. 28, no. 3, pp. 284–91, Oct. 2019, doi: 10.13181/mji.v28i3.2589.
[12] Nia Saurina, Nur Chamidah, Riries Rulaningtyas, and Aryati Aryati, "Improved GrabCut Algorithm for Classify Mycobacterium Tuberculosis," Universitas Respati Yogyakarta, 2025.
[13] K. M. Tamez-Torres *et al.*, "Mycobacterium tuberculosis complex bacteremia among HIV and non-HIV patients in a Mexican tertiary care center," *The Brazilian Journal of Infectious Diseases*, vol. 22, no. 5, pp. 387–391, Sep. 2018, doi: 10.1016/j.bjid.2018.10.274.
[14] 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.
[15] N. Saurina, N. Chamidah, R. Rulaningtyas, and A. Aryati, "Multimodal deep learning from sputum image segmentation to classify Mycobacterium tuberculosis using IUATLD assessment," *Bulletin of Electrical Engineering and Informatics*, vol. 14, no. 2, pp. 1579–1590, Apr. 2025, doi: 10.11591/eei.v14i2.9250.
[16] W. Wu *et al.*, "Application of local fully Convolutional Neural Network combined with YOLO v5 algorithm in small target detection of remote sensing image," *PLoS One*, vol. 16, no. 10, p. e0259283, Oct. 2021, doi: 10.1371/journal.pone.0259283.
[17] Y. Li, M. Ni, and Y. Lu, "Insulator defect detection for power grid based on light correction enhancement and YOLOv5 model," *Energy Reports*, vol. 8, pp. 807–814, Nov. 2022, doi: 10.1016/j.egyr.2022.08.027.
[18] J. Wu, L. Su, Z. Lin, Y. Chen, J. Ji, and T. Li, "Object Detection of Flexible Objects with Arbitrary Orientation Based on Rotation-Adaptive YOLOv5," *Sensors*, vol. 23, no. 10, p. 4925, May 2023, doi: 10.3390/s23104925.
[19] R. Bista *et al.*, "Advancing Tuberculosis Detection in Chest X-rays: A YOLOv7-Based Approach," *Information*, vol. 14, no. 12, p. 655, Dec. 2023, doi: 10.3390/info14120655.
[20] A. Khalfaoui, A. Badri, and I. EL Mourabit, "Comparative study of YOLOv3 and YOLOv5's performances for real-time person detection," in *2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, IEEE, Mar. 2022, pp. 1–5. doi: 10.1109/IRASET52964.2022.9737924.
[21] Cinantya Paramita, Catur Supriyanto, Amalia, and Khalivio Rahmyanto Putra, "Comparative Analysis of YOLOv5 and YOLOv8 Cigarette Detection in Social Media Content," *Scientific Journal of Informatics*, vol. 11, no. 2, pp. 341–352, May 2024, doi: 10.15294/sji.v11i2.2808.
[22] L.-D. Quach, K. N. Quoc, A. N. Quynh, and H. T. Ngoc, "Evaluating the Effectiveness of YOLO Models in Different Sized Object Detection and Feature-Based Classification of Small Objects," *Journal of Advances in Information Technology*, vol. 14, no. 5, pp. 907–917, 2023, doi: 10.12720/jait.14.5.907-917.
[23] D. Reis, J. Kupec, J. Hong, and A. Daoudi, "Real-Time Flying Object Detection with YOLOv8," May 2024.
[24] A. K. Rozi, "Evaluation of Implementation of Occupational Health Safety Management System (OHSMS) in PT. X, Magetan, East Java," in *Proceedings of the 2nd International Conference Postgraduate School*, SCITEPRESS - Science and Technology Publications, 2018, pp. 550–555. doi: 10.5220/0007546805500555.

[25] J. Terven, D.-M. Córdova-Esparza, and J.-A. Romero-González, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Mach Learn Knowl Extr*, vol. 5, no. 4, pp. 1680–1716, Nov. 2023, doi: 10.3390/make5040083.
[26] Z. Z. Qin *et al.*, "Can artificial intelligence (AI) be used to accurately detect tuberculosis (TB) from chest X-rays? An evaluation of five AI products for TB screening and triaging in a high TB burden setting," May 2021.
[27] Tetsuya Asakawa, Riku Tsuneda, Kazuki Shimizu, Takuyuki Komoda, and Masaki Aono, "Caverns Detection and Caverns Report in Tuberculosis: lesion detection based on image using YOLO-V3 and median based multi-label multi-class classification using SRGAN,"
[28] A. N.S., S. D., and R. K. S., "Naive Bayesian fusion based deep learning networks for multisegmented classification of fishes in aquaculture industries," *Ecol Inform*, vol. 61, p. 101248, Mar. 2021, doi: 10.1016/j.ecoinf.2021.101248.
[29] J. Wang *et al.*, "Research on Improved YOLOv5 for Low-Light Environment Object Detection," *Electronics (Basel)*, vol. 12, no. 14, p. 3089, Jul. 2023, doi: 10.3390/electronics12143089.
[30] A. Khalfaoui, A. Badri, and I. El Mourabit, "An improved YOLOv5 for real-time human detection in infrared images," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 32, no. 2, p. 1078, Nov. 2023, doi: 10.11591/ijeecs.v32.i2.pp1078-1085.
[31] Z. Mustafa and H. Nsour, "Using Computer Vision Techniques to Automatically Detect Abnormalities in Chest X-rays," *Diagnostics*, vol. 13, no. 18, p. 2979, Sep. 2023, doi: 10.3390/diagnostics13182979.
[32] G. Fekadu, Y. Wang, and J. H. S. You, "Standard diagnostics with and without urine-based lipoarabinomannan testing for tuberculosis disease in HIV-infected patients in a high-burden setting–A cost-effectiveness analysis," *PLoS One*, vol. 18, no. 7, p. e0288605, Jul. 2023, doi: 10.1371/journal.pone.0288605.
[33] L. H. Thai, T. S. Hai, and N. T. Thuy, "Image Classification using Support Vector Machine and Artificial Neural Network," *International Journal of Information Technology and Computer Science*, vol. 4, no. 5, pp. 32–38, May 2012, doi: 10.5815/ijitcs.2012.05.05.
[34] R. Bai, F. Shen, M. Wang, J. Lu, and Z. Zhang, "Improving Detection Capabilities of YOLOv8-n for Small Objects in Remote Sensing Imagery: Towards Better Precision with Simplified Model Complexity," Jun. 22, 2023. doi: 10.21203/rs.3.rs-3085871/v1.
[35] D. N. B. Tran, T. T. T. Nguyen, H. V. Le, H. X. Nguyen, V. P. Tran, and D. Q. Ta, "Testing neural networks assessment based on data-driven using well log data in Cuu Long basin, offshore Vietnam," *Science & Technology Development Journal - Engineering and Technology*, 2021, doi: 10.32508/stdjet.v4iSI3.584.
[36] L. P. K. Dewi, "PEMERIKSAAN BASIL TAHAN ASAM UNTUK MEMBANTU MENEGAKKAN DIAGNOSIS PENYAKIT TUBERKULOSIS," *International Journal of Applied Chemistry Research*, vol. 1, no. 1, p. 16, Sep. 2020, doi: 10.23887/ijacr.v1i1.28716.
[37] 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.
[38] J. Li *et al.*, "Forecasting the Tuberculosis Incidence Using a Novel Ensemble Empirical Mode Decomposition-Based Data-Driven Hybrid Model in Tibet, China," *Infect Drug Resist*, vol. Volume 14, pp. 1941–1955, May 2021, doi: 10.2147/IDR.S299704.
[39] E. Kee, J. J. Chong, Z. J. Choong, and M. Lau, "A Comparative Analysis of Cross-Validation Techniques for a Smart and Lean Pick-and-Place Solution with Deep Learning," *Electronics (Basel)*, vol. 12, no. 11, p. 2371, May 2023, doi: 10.3390/electronics12112371.