

Received 20 March 2024, accepted 4 April 2024, date of publication 8 April 2024, date of current version 16 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3386208

## RESEARCH ARTICLE

# A Novel Digitized Microscopic Images of ZN-Stained Sputum Smear and Its Classification Based on IUATLD Grades

SUCI AULIA<sup>1,2</sup>, ANDRIYAN BAYU SUKSMONO<sup>1,3,4</sup>, (Senior Member, IEEE),  
TATI RAJAB MENGKO<sup>1</sup>, AND BACHTI ALISJAHBANA<sup>5</sup>

<sup>1</sup>School of Electrical Engineering and Informatics, Bandung Institute of Technology, Bandung 40132, Indonesia

<sup>2</sup>School of Applied Science, Telkom University, Bandung 40257, Indonesia

<sup>3</sup>Research Collaboration Center for Quantum Technology 2.0, STEI-ITB, Bandung 40132, Indonesia

<sup>4</sup>ITB Research Center on ICT (PPTIK-ITB), Bandung 40132, Indonesia

<sup>5</sup>Medical Faculty, Universitas Padjadjaran, Bandung 40161, Indonesia

Corresponding author: Suci Aulia (suci.aulia.sca@gmail.com)

This work was supported in part by the Indonesia Endowment Fund for Education Agency and School of Electrical and Information Engineering, Bandung Institute of Technology; in part by Riset Unggulan ITB-2023; and in part by the Ministry of Research and Technology and the Ministry of Education and Culture under World Class University (WCU) Program managed by Institut Teknologi Bandung.

**ABSTRACT** Microscopic detection of acid-fast bacilli (AFB) from mycobacterium tuberculosis (MTB) in Ziehl-Neelsen (ZN)-stained sputum samples is a crucial step in the detection of TB (tuberculosis) disease. Pathologists encounter many challenges that may result in incorrect diagnoses, such as the heterogeneous shape and irregular appearance of MTB, low-quality ZN staining, and errors in scanning each of the field of view (FoV) using a conventional microscope. Additionally, multiple manual observations may cause fatigue that leads to human error. Several studies have created microscopic imaging databases of sputum samples, aiding researchers in creating computer-aided diagnosis (CAD) for tuberculosis, which is a promising method that offers timely, reliable, and repeatable assistance. Nevertheless, the implementation of CAD systems for TB diagnosis remains an area of ongoing research and development owing to the lack of microscopic image datasets of sputum samples, which represent whole-slide imaging (WSI) that follows the WHO (World Health Organization) regulations. To address this issue, this study developed a novel digitized microscopic image from sputum smear samples of Indonesian patients in the WSI that conform to the WHO regulation. These images are collected as a Microscopic Imaging Database of Tuberculosis Indonesia (MIDTI). This study also proposed a method based on the YOLOv7 (You Only Look Once seventh version) algorithm to develop a CAD for tuberculosis diagnosis by classifying ZN-stained sputum smear samples into International Union Against Tuberculosis and Lung Disease (IUATLD) grades, which has never been revealed in any previous studies.

**INDEX TERMS** Tuberculosis, Ziehl-Neelsen, acid-fast bacilli, MIDTI, IUATLD, YOLOv7, reparameterization visual geometry group (RepVGG).

## I. INTRODUCTION

Tuberculosis (TB), caused by mycobacterium tuberculosis (MTB), is one of the leading global causes of death, particularly among sensitive groups (such as those with malnutrition, AIDS, and low-income nations). In 2021, the

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

World Health Organization (WHO) regions of the Western Pacific (18%), Africa (23%), and Southeast Asia (45%) had the highest proportions of TB cases, followed by the eastern Mediterranean (8.1%), Americas (2.9%), and Europe (2.2%). Eight of these 30 nations accounted for more than two-thirds of the global total, including 87% of all estimated incident cases worldwide: India (28%), Indonesia (9.2%), China (7.4%), the Philippines (7.0%), Pakistan (5.8%), Nigeria

(4.4%), Bangladesh (3.6%), and the Democratic Republic of Congo (2.9%) [1], [2]. According to the global tuberculosis report by the WHO, Indonesia was the second country after India, which had major contributors to the global TB increase between 2020 and 2021. The government, the medical communities, intellectuals, and other groups must work together to address this issue. The term ‘intellectuals’ refers to researchers who study the development of a system to aid medical staff in diagnosing patients faster and with higher precision, sensitivity, and specificity. Sputum smear microscopy, chest X-rays [3], [4], [5], [6], [7], [8], [9], [10], [11], [12] rapid molecular tests, MRI [13], and culture methods are diagnostic laboratory tests for TB. Sputum smear microscopy is the most common and essential method used in Indonesia because it is the most reliable and cost-effective approach recommended by the WHO for first-line laboratory diagnosis of TB [14]. The primary challenge is that the MTB is a small bacillus (length  $2 - 4\mu\text{m}$ /width  $0.2 - 0.6\mu\text{m}$ ), which needs to be examined adequately in hundreds or thousands of microscopic fields of  $0.5\text{ mm}$  in diameter to be found in a  $2 \times 3\text{ cm}$  of tissue on the slide ( $6 \times 10^{12}$  square microns) [15], [16]. To address this problem, MTB is typically observed under a conventional light microscope (CM) at  $1000\times$  magnification with immersion oil [17]. The process of converting MTB to acid-fast bacilli (AFB) involves several steps, including air drying, heat-fixing, and staining methods using Ziehl-Neelsen (ZN) [14], [18], [19]. In the context of AFB detection, the choice of staining features, such as color, is crucial for creating an effective visual contrast and identifying the AFB in sputum smears. The ZN staining method is specifically designed to target the AFB, which appears red after staining, allowing for clear differentiation from non-AFB and other cellular structures, in contrast to the blue background, as shown in the red boxes in Fig. 1.

After ZN staining, problems can arise due to several factors, such as slide contamination, incorrect staining time, overheating or underheating of the stain, and improper blotting of the slide. For example, over or under heating, the stain causes the AFB color to fade (blur) or become bluish (dark red), which can result in missed detection of AFB as the background. Variations in AFB shape, such as curves and grouping or colonies, are also key features in detection, because they can cause errors in calculating the number of bacteria. These factors can lead to false-negative and false-positive results [20], [21].

Pathologists commonly scan every field of view (FoV) of sputum smear slides with their eyes to identify and compute AFB under CM. Because of the time-consuming and ineffective nature of this process, computer-aided diagnosis (CAD) is a promising method that offers timely, reliable, and repeatable assistance. The CAD system can help pathologists to perform TB diagnostic smear microscopy more accurately and quickly.

Two recent methods for sputum smear microscopy diagnosis have been developed to create CAD for TB: artificial

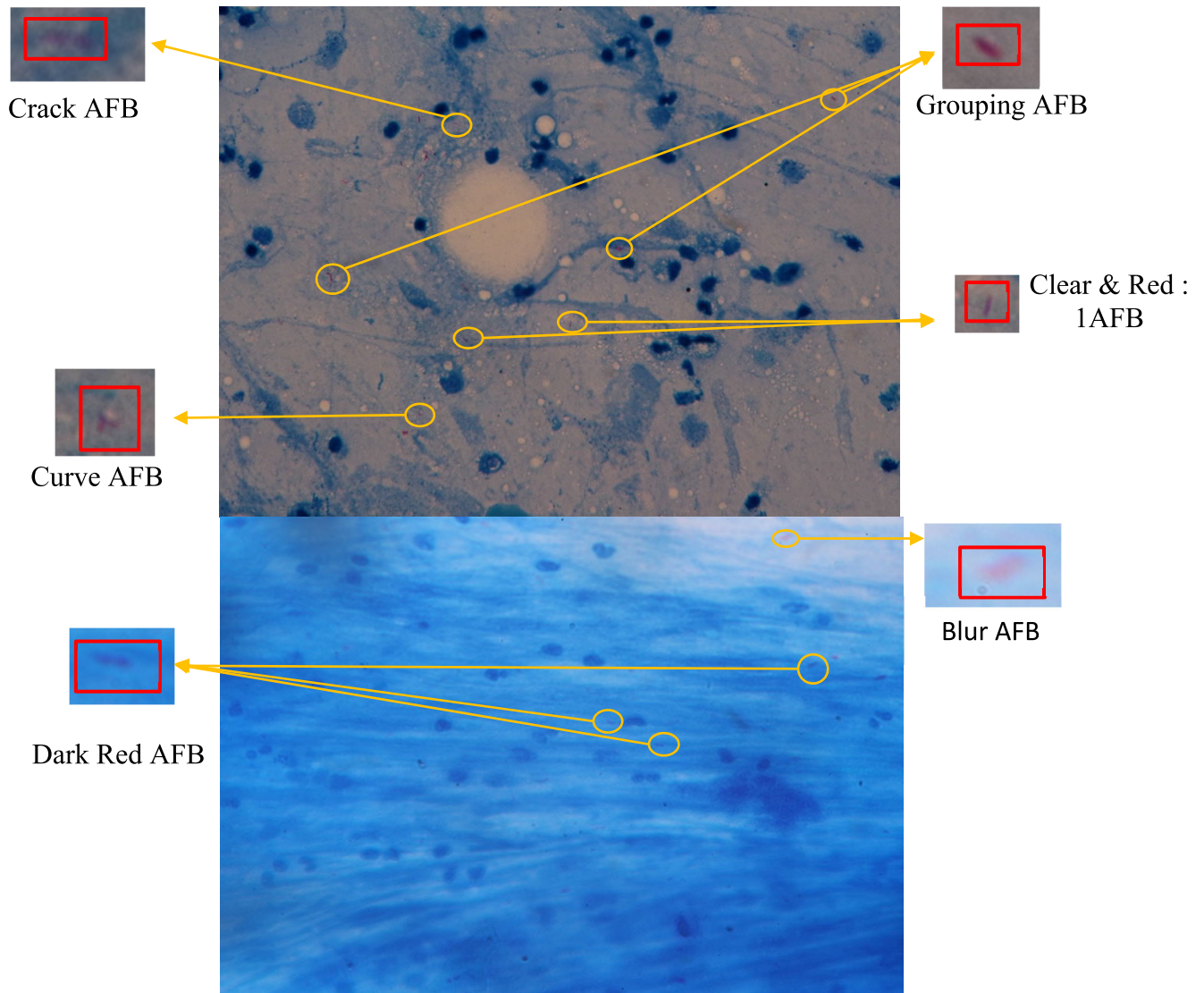
TABLE 1. IUATLD-grade of tuberculosis disease.

TB GRADE/CLASS	NUMBERS OF AFB	$\Sigma$ min OF FoV
NEGATIVE	0 AFB/300FoV	300
SCANTY	1-29 AFB/300FoV	300
1+	>30 AFB/300FoV	300
2+	1-10 AFB/FoV	50
3+	> 10 AFB/FoV	20

intelligence (AI) [15], [22], [23], [24], [25] and mechanical automation [15], [26], [27], [28]. These methods are similar to digital pathology, which uses an image recognition algorithm to detect AFB by digitally scanning the TB smear slides. These two combinations make the examination procedure more sensitive and accurate. Although most CAD studies show improved performance over human evaluation, most are still in the development stage and use “proof-of-concept” systems or laboratory-spiked materials [29]. Most studies that used mechanical automation and artificial intelligence (AI) still needed to place sputum smear slides into International Union Against Tuberculosis and Lung Disease (IUATLD) classes, as shown in Table 2. They classified sputum smear slides into positive (AFB +) and negative (AFB-) case groups. However, skilled professionals still need to count the number of the AFB and analyze the test results based on IUATLD grades. This study aimed to determine how well a CAD system can distinguish sputum smears into five classes based on IUATLD grades: negative, scanty, TB1+, TB2+, and TB3+, to assess the severity and extent of the infection. This classification is primarily used in the context of sputum smear microscopy, which is a common diagnostic test for tuberculosis. Categories were defined based on the presence and quantity of AFB in the sputum smears. AFB count was determined by examining sputum samples under a microscope and counting the number of AFB per FOV, as shown in Table 1 [30], [31].

Our study provides new insights into various areas that have not been covered in previous research, i.e:

- First, we created a new dataset of TB sputum sample images, namely the Microscopic Imaging Database of Tuberculosis Indonesia (MIDTI). MIDTI was categorized based on the IUATLD grade and digitized system, according to the WHO standard. The MIDTI images had dimensions of  $3072 \times 2304$  pixels in (.jpg) format. The MIDTI is an open access data and can be accessed at [https://drive.google.com/drive/folders/1CB1jJbw38yavWcMo3azPeSrV9R9g\\_iW?usp=sharing](https://drive.google.com/drive/folders/1CB1jJbw38yavWcMo3azPeSrV9R9g_iW?usp=sharing). We also provide the AFB annotations on Roboflow, available at <https://universe.roboflow.com/suci-aulia/afb-test-y7>.
- Second, our study proposed CAD for TB based on sputum examination using a method that is based on the YOLOv7. This study classifies sputum smears into IUATLD grades: negative, scanty, TB1+, TB2+, and TB3+, which have never been revealed in previous studies.



**FIGURE 1.** Examples of AFB-MTB variations in one FoV (3702 × 2304 pixels). Some problems which come up after ZN-staining; ZN overstaining caused the AFB to have a dark red color, an unfocused acquisition system caused the AFB to blur, an outdated sample caused the AFB to crack and disappear. Besides AFB-color variations, another problem which caused miss leading the screening process was; one of the types of AFB is a curved form, and an AFB group is an accumulation of AFBs nested together or in colonies.

## II. RELATED WORK

Many researchers have described AFB detection techniques for ZN-stained sputum smear images using both conventional and automatic scanning microscopy (Table 2). The creation of an automated AFB detection system generally involves the following steps: (i) capturing and pre-processing images, (ii) segmenting objects, (iii) extracting features, and (iv) classifying them into AFB and non-AFB [30]. Several studies have created microscopic imaging databases from sputum samples to aid researchers in the development of CAD-TB. Nevertheless, only a few of the microscopic image datasets of sputum samples [16], [22], [23], [15], [25], and [28] shown in Table 2 represent whole slide imaging (WSI) or at least 20–300 FoVs/sputum samples following the IUATLD grade that conforms to WHO regulations.

A related study produced microscopic image databases of WSI-representative sputum samples [16]. An Aperio AT Turbo automatic scanner was used to gather and divide 167 slides from the Cedars-Sinai Medical Center, producing 33,672 non-overlapping images. With a pixel size of  $256 \times 256$ , each image was classified as AFB+ and AFB- using CNN (Convolutional Neural Network) that yields 87.13% sensitivity, 87.62% specificity, and 80.18% F1-score. Images with and without AFB were denoted as AFB+ and AFB-, respectively. Another study examined and digitized 441 images from 2016 to 2019 using Aperio AT2, Leica Biosystems scanner, and Hamamatsu Nanozoomer XR was given by Pantanowitz et al. in [22]. In this paper, an original  $100,000 \times 100,000$  pixel WSI size was partitioned into separate patches with a pixel size of  $64 \times 64$  pixels, and

each patch in the AFB detection was classified by the CNN algorithm as AFB+ or AFB -, with a sensitivity of 60% and a specificity of 99.99%.

Lo et al. [23] used an automated whole-slide scanner ScanScope XT (Aperio) to show the WSIs of 613 bacillus-positive and 1202 negative images. Sputum smear slides were obtained from Macky Memorial Hospital. As a result of this study, the Deep CNN model correctly classified the tissue images into AFB+ and AFB-, with a sensitivity of 93.5%, a specificity of 96.3%, and an accuracy of 95.3%. Another study gathered 570 ZN-stained sputum slides from the Department of Pathology of Colentina University Hospital [15]. A Leica Aperio GT450 automatic scanner was used to scan each slide as a whole slide image (WSI). "svs" format. The images were classified as AFB+ or AFB- using an artificial intelligence-based technique with 98.33% accuracy, 100% specificity, and 95.65% sensitivity. Further, Zaizen et al. developed an AI-assisted pathology method using WSI labelled on HALO software to detect AFB in ZN-stained sputum smear slides [25]. Every slide was taken from a pulmonary tuberculosis autopsy case and digitized WSI was acquired using Motic EasyScan with a resolution of 0.25 m/px. However, the outcomes were categorized as AFB + or AFB-.

Another assessment of an AI-based system for AFB detection was used in the same year 2022 by Fu et al. in [27]. Each smear contained 200 images, which differed according to the WHO standard of at least 20-300 FoVs depending on tuberculosis level based on the IUATLD grades [31]. In this study, an AI-based system for AFB detection successfully identified AFB in 5930 TB sputum smears. Each smear was classified into six groups: negative, trace, 1+,2+,3+, and 4+, resulting in an accuracy of 95.2%, a sensitivity of 85.7%, and a specificity of 96.9%. The study in [31] claimed that the microscopic images were digitally and randomly captured; this acquisition did not meet the WHO standard requirement, which is a well-organized scanning flow rule. In addition, the sputum smear classification was not based on IUATLD grades, which were divided into five groups: negative, scanty, TB1+, TB2+, and TB3+. The last update of the study was published in 2023 [28], and an automatic slide scanner from 3DHISTECH was used to produce 28,913 negative images (without AFB) and 18,426 positive images (containing AFB). This study reduced the original image size from  $512 \times 512$  pixels to  $256 \times 256$  pixels to address the analysis effectiveness of the covariance-based classification (ANCOVA) method in training positive and negative images. The aim of this study was to implement AI to classify the WSI of ZN-stained slides as AFB + or AFB -. As shown in Table 2, several studies employing various methods have been conducted to identify AFB in ZN-stained sputum smears. Nevertheless, studies that successfully classified sputum smears into AFB+ and AFB- have primarily used CNN-based approaches. As a variant of the CNN method, object identification techniques, such as Region based CNN (RCNN) and its variants [51] and [46], employ a sequence

of steps to complete the work. This can be slow to run and difficult to optimize because each element must be trained individually by creating the proposed regions and identifying the objects within those regions, which are known as two-stage object detectors [53].

You Only Look Once (YOLO) is one of the most well-known object recognition algorithms because it uses single-stage object detectors that can predict the outcome of an input in only one pass. The YOLO family, which is based on neural network models, is currently the recommended algorithm for fast object detection because it is highly compatible with industrial standards, including precision, lightness, and user-friendly installation conditions [54].

YOLO was first launched in 2015 by Joseph Redmon [55] namely YOLOv1, then developed rapidly into YOLO9000/YOLOv2 in 2016 [56] YOLOv3 in 2018 [57], YOLOv4 in April 2020 [58], Scaled YOLOv4 in November 2020 [59], YOLOv5 in June 2020 (without research paper published, but available online <https://github.com/ultralytics/yolov5>), PP-YOLO in July 2020 [60], YOLOR in May 2021 [61], YOLOX in July 2021 [62], YOLOv6 in June 2022 [63], YOLOv7 in July 2022 [64], and the latest is YOLOv8 in 2023 without research paper published, but available online [65]. Our study employs YOLO-v7 algorithm as a base of method to detect the AFB on MIDTI. The YOLO-v7 outperformed the examined object detectors in terms of accuracy, speed, efficiency as shown in Fig.2.

Fig. 2a shows that in the range of 5-160 FPS, YOLOv7 achieves the highest accuracy among all real-time object detection models. This makes it suitable for real-time applications, where the speed is crucial. Although YOLOv8 exhibits excellent real-time performance, it is assumed that YOLO-v8 concentrates on limited device implementations with high deduction speeds [54].

Fig.2b shows that YOLOv7 is 41% more efficient in terms of parameter usage than YOLOv8 [54]. Thus, YOLOv7 is a better choice for applications where resource constraints are a concern. Additionally, YOLOv8 is still regarded as unofficial because no paper has been published since its release by Ultralytics in 2023 [54].

On the other hand, with a simple architecture comprising a stack of CNN, ReLU (Rectified Linear Unit), and pooling, the Visual Geometry Group (VGG) showed impressive results in image recognition despite being still heavily used for real-world applications in both academia and industry [66]. As a solution, reparameterization VGG (RepVGG) was published in 2021 and achieved over 80% top-1 accuracy on ImageNet compared with RegNetX, ResNet, ResNeXt, and EfficientNet [67], [68]. The RepVGG model is typically used for deep feature extraction and classification [69], and it has recently been proposed to be linearly connected with YOLOv6 [69] and YOLOv7 [70] to achieve a better accuracy and speed trade-off. RepVGG has been added to the YOLOv7 backbone network (RepVGG-Yolov7) to improve the model capacity for feature extraction during training and

**TABLE 2.** List of studies that report identification of the AFB of Mycobacteria tuberculosis in ZN-stained sputum smears using image processing methods, including the development of datasets processes that follows the WSI approach or representing all smear viewpoints, mechanical automation and whether the output is based on IUATLD. (x = no, = yes, NA = not available,  $\phi$  = number of pathologist, output = the result of a study in classifying sputum sample slides, AFB+ = images with AFB, AFB- = images without AFB).

No	YEAR	Authors	Method	$\phi$	WSI	Automatic Scanner	Output
1.	2009	Khutlang et al. [32]	MoG, PCA,KNN	NA	×	×	×
2.	2009	Makkapati et al. [33]	HSV color space	NA	×	×	×
3.	2010	Osman et al. [34]	Color space, K-means, HMLP	NA	×	×	×
4.	2010	Khutlang et al. [35]	Bayes, linear, quadratic, KNN, PNN, SVM	NA	×	×	×
5.	2010	Nayak et al. [36]	HIS color space	NA	×	×	×
6.	2010	Zhai et al. [37]	Color spaces	NA	×	×	×
7.	2011	Riries et al. [38]	JST-BP	NA	×	×	×
8.	2014	Riries et al. [39]	KNN	NA	×	×	×
9.	2015	Ayma et al. [40]	Least Mean Squares and Reduced Rank algorithms.	NA	×	×	×
10.	2016	Riries et al. [41]	Patch, K-mean, Liu's Factor	NA	×	×	×
11.	2016	Shah et al. [42]	Watershed Transform, Creating database called ZNSM-iDB	NA	×	×	×
12.	2017	Riries et al. [43]	SOM, K-Mean clustering	NA	×	×	×
13.	2017	Shah et al. [44]	Watershed Transform, and completed developing ZNSM-iDB[44]	NA	×	×	×
14.	2017	Mithra et al. [45]	Gaussian-Fuzzy-Neural network, ZNSM-iDB[44]	NA	×	×	×
15.	2018	Kuok et al. [46]	Faster RCNN	NA	×	×	×
16.	2018	Sonaal et al. [47]	5 layered FCN, using ZNSM-iDB[44]	NA	×	×	×
17.	2018	Shah et al. [48]	Watershed Transform , using ZNSM-iDB[44]	NA	×	×	×
18.	2018	Rafikha et al. [49]	K-Mean Clustering	NA	×	×	×
19.	2019	Del Carpio et al. [50]	SVM	2	×	MiniVID microscope with CMOS sensor.	×
20.	2019	El-Melegy et al. [51]	Faster RCNN	NA	×	×	×
21.	2019	Mithra et al. [52]	Channel Area Thresholding (CAT), using ZNSM-iDB[44]	NA	×	×	×
22.	2019	Jorge et al. [21]	Bayes Classifier, Kmeans	NA	×	×	×
23.	2020	Yang et al. [16]	CNN, Inception-V3	NA	✓	Aperio AT Turbo scanner: 40x	AFB+ , AFB-
24.	2021	Pantanowitz et al. [22]	DeepCNN	26	✓	Aperio AT2 : 40x Hamamatsu XR:40x ScanScope XT (Aperio)	AFB+ , AFB-
25.	2020	Ming et al. [23]	DCNN	NA	✓	ScanScope XT (Aperio)	AFB+ , AFB-
26.	2021	Ibrahim et al. [24]	AlexNet Models	NA	×	×	AFB+ , AFB-
27.	2022	Zurac et al. [15]	RegNet-X4 modification	7	✓	Leica Aperio GT450.	AFB+ , AFB-
28.	2022	Zaizen et el. [25]	HALO AI software (CNN-based)	6	✓	Motic EasyScan: 0.25 $\mu$ m/px.400x	AFB+ , AFB-
29.	2022	Fu et al. [27]	$\mu$ -Scan software, CNN-based	NA	×	$\mu$ -Scan	negative, trace, 1+,2+,3+, 4+
30.	2022	Huang et al. [26]	$\mu$ -Scan software, CNN-based	NA	NA	$\mu$ -Scan 2.0	AFB+ , AFB-
31.	2023	Nurzynska et el. [28]	CNN	NA	✓	3DHISTECH, Hungary)	AFB+ , AFB-

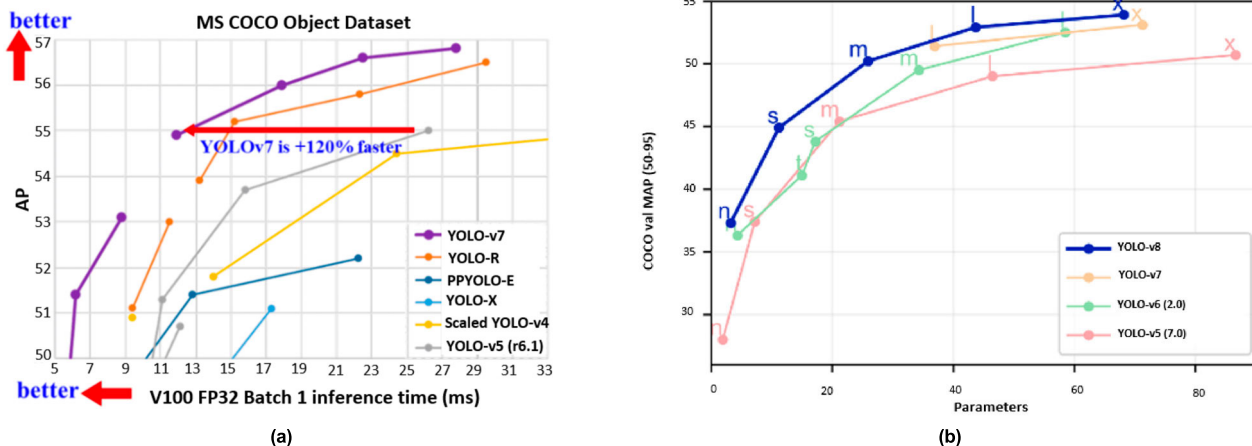
to achieve lossless model compression during inference [70]. According to numerical experiments, the proposed algorithm can identify smoke with a complex background with accuracy of up to 95.1% [70]. Another study combined YOLOv7 with RepVGG to detect beach litter [71]. Based on inference time measurement, Yolov7 and Yolov7+RepVGG resulted in 11.8ms and 8.4ms, respectively. This result shows that RepVGG significantly speeds up the detection and reduces the inference time.

To increase the performance of sensitivity and specificity measurements, we proposed a model to detect AFB in sputum images by adding RepVGG to the YOLOv7 backbone network.

### III. MATERIAL AND METHODS

#### A. MATERIAL DATA

In this study, the MIDTI consisting of 15,046 images, was acquired and digitalized from 67 sputum smear samples.



**FIGURE 2.** The performance of YOLO-v7 with earlier YOLO variations, (b) YOLO-v8 comparison with earlier YOLO variation based on its model (n = nano, s = small, m = medium, l = large, x = extra) [57].

All 67 ZN-stained sputum smear samples were collected from the RC3ID (Research Center for Care and Control of Infectious) Diseases, Medical Faculty, Padjadjaran University (UNPAD), Bandung, Indonesia. Sample collection, preparation, sputum smear microscopy scanning and diagnosis, and handling of the sputum smear slide for digitization were performed according to the RC3ID guidelines, with ethical approval number 4/UN6.KEP/EC/2023. Some of the images captured using the digitizing system are shown in Fig.3.

The image acquisition system used was an Olympus CX-31 microscope equipped with a 700D DSLR camera provided by the Biomedical Laboratory at the School of Electrical Engineering and Informatics, Bandung Institute of Technology (BME-ITB).

The acquisition of microscopic images or digitized sputum smears followed the WHO and the Gold Standard of RC3ID-UNPAD to guarantee that smears are read sequentially following the WSI approach or represent all smear viewpoints. We move to the next FoV and scan the slides with visible stains from left to right, if the current FoV is empty. Once we had the best focus on AFB, we captured the FoV. Sputum smear samples with a grade of negative, scanty, and TB1+ must have at least 300 FoVs in their image collection. The minimum number of images for sputum smear samples with 2+ and TB3+ was 100 and 20, respectively.

All sample data and digitalization of the MIDTI data were collected from October 2020 to August 2022. MIDTI comprises 15,046 images from 67 samples of ZN-stained sputum smear slides. The image size had a resolution of 3,072 × 2,304 pixels in (.jpg) file format and RGB color. A total of 15,046 images were divided into two groups. In the first group, 5,687 images from 27 smears were digitalized and analyzed for the training process. In the second group, 9,359 images were obtained from 40 blind samples that were used for testing. The distribution of the number of samples for each class was almost equal, with a proportional distribution

**TABLE 3.** Distribution of 67 samples and 15.046 images on MIDTI for each class (grade) in IUATLD.

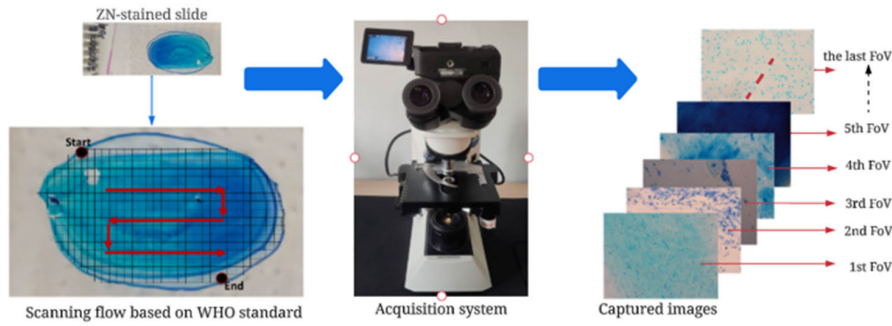
Class	Number of samples			Number of images		
	Train	Test	Total	Train	Test	Total
Negative	5	8	13	1,518	2,403	3,921
Scanty	4	8	12	1,200	2,419	3,619
TB1+	5	8	13	1,395	2,409	3,804
TB2+	6	8	14	870	975	1,845
TB3+	7	8	15	704	1,153	1,857
Total	27	40	67	5,687	9,359	15,046

representative of all ZN staining categories (Fig.4), as shown in Table 3.

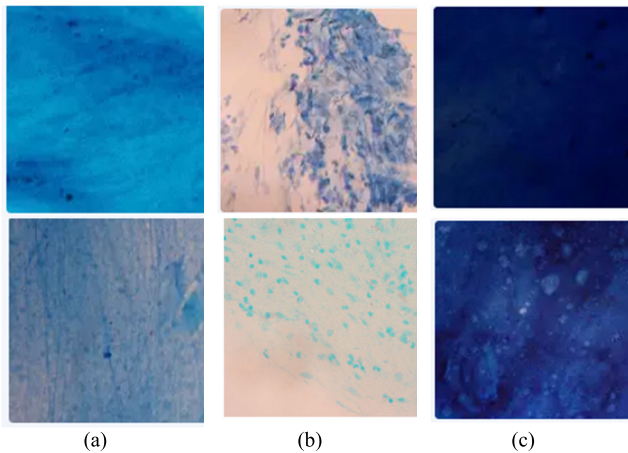
**B. DATA ANNOTATION**

Three pathologists (NR, MI, and IZ) with various level of experience in RC3ID diagnosed 67 sputum samples and annotated or labelled each sample based on IUATLD grades: negative, scanty, TB1+, TB2+, and TB3+. The pathologists provided 27 samples with label identity of the diagnostic results for training, and the 40 remaining samples were used for testing as blind datasets or without label identity.

Throughout the training process using the MIDTI training datasets, AFB objects were annotated using bounding boxes (BB) for feature extraction, using Roboflow as the digital annotation tool (www.roboflow.com). Roboflow Annotate offers a user-friendly interface for annotating images with tools like polygons and bounding boxes to identify objects. The platform supports both model-assisted labeling and human annotation. Labeling bacteria, i.e. the AFB, on RoboFlow requires great expertise in tuberculosis pathology. In this study, three pathologists with years of experience in the tuberculosis laboratory at RC3ID UNPAD annotated each bacterium on 371 images in Roboflow to establish a baseline. Images were randomly selected from the MIDTI training data, which were proportionally representa-



**FIGURE 3.** Acquisition system for MIDTI. ZN-stained slides were prepared by the tuberculosis laboratory at RC3ID UNPAD with red-path digitalized scanning using an Olympus CX-31 mounted with DSLR 700D provided by BME ITB.



**FIGURE 4.** Background variations in sputum smear samples: (a) good staining, (b) less staining, and (c) oversteining.

tive based on variations in the ZN staining intensity, AFB shape, and color degradation.

The quality of ZN staining can cause AFB color variations, as shown in Fig.1, the captured image results in Fig.3, and the details in Fig.4, which affect the image acquisition results of the MIDTI. For accurate results, the quality of ZN staining is important. Staining can be graded as good, less, or overstained according to how the stained smears appear under a microscope by the pathologist, as shown in Figures 4a, 4b, and 4c, respectively.

**C. DATA VALIDATION**

As explained in the Data Annotation section, pathologists at RC3ID provided a blind dataset to test the proposed model. After the system in this study classified the entire blind dataset into IUATLD grades, pathologists compared the results with those of their previous manual screening as the ground truth using CM ZEISS with 1000x magnification and oil immersion.

**D. CLASSIFICATION METHODS**

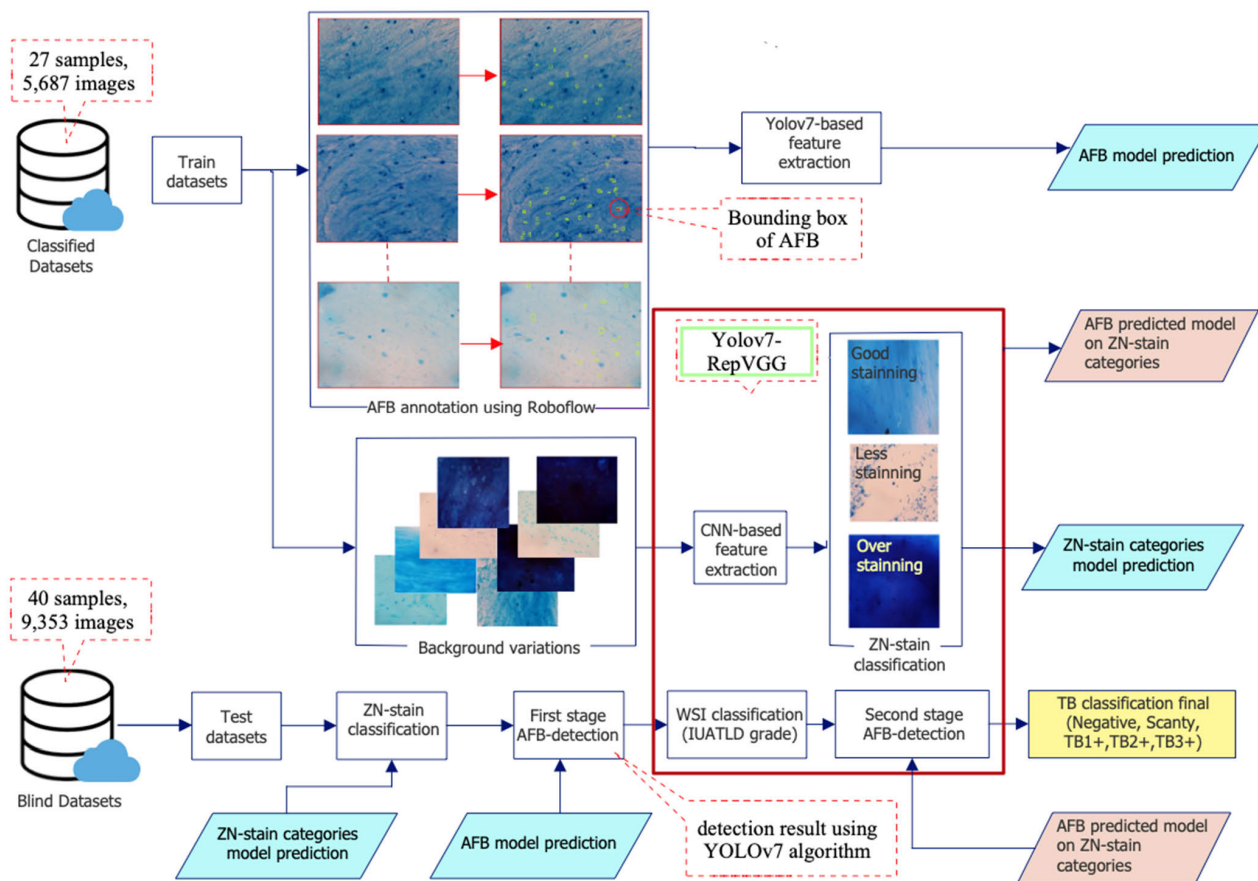
We proposed a study to classify ZN-stained sputum smear slides in MIDTI into five classes based on IUATLD grades

(Table 1). Our active learning includes two stages: the first stage is AFB detection using YOLOv7 and the second stage is an optimization in AFB detection using YOLOv7-RepVGG, as shown in Fig.5.

The following stages are the process of transforming each sputum sample from MIDTI to IUATLD by using the proposed model:

1. At first, a feature extraction of AFB was conducted by annotating AFB using bounding boxes (BB) format in Roboflow by the TB pathologists. All annotation results or all BB AFB results from 371 images in Roboflow can be downloaded in YOLOv7 format (YOLOv7PyTorch.txt) as an AFB annotation package.
2. The packaging was then trained on Yolov7. During the training, the model should learn to detect and localize AFB in sputum microscopic images. Optimizing the hyperparameters produced excellent AFB prediction modelling, which was subsequently used for testing. This stage result is called the AFB model-predicted result, as shown in Fig.5.
3. To optimize the hyperparameters on Yolov7, we adjust the tuning parameters, including input size, learning rate, number of batch sizes, epochs, activation function, and momentum during training [72].
4. The best AFB-model-predicted value was then used for the data-test MIDTI. The model identifies and counts the AFB in each image or the FoV. The total AFB of WSI/sample sputum was used to determine the grade of TB disease based on the IUATLD, as shown in Table 1.
5. Validation of the performance of Yolov7 by pathologists. Measure the performance using appropriate metrics and refine the model as needed to ensure accurate detection and counting of AFB.

Once an AFB identification method using the YOLOv7 algorithm was employed, the complex AFB color variations, as shown in Fig.1, produce many incorrect candidates. Therefore, to optimize the performance of sputum smear slide classification into IUATLD grades, a ZN staining classification system was proposed in this study, as illustrated in Fig.5. The input image underwent a filtering procedure based on the three ZN-staining categories during the testing



**FIGURE 5.** Active learning of the proposed model process for TB classification based on IUATLD grading. The top-left figure shows the training procedure using 27 sputum samples, which were labeled by the pathologists based on the IUATLD grades. The training procedure is divided into two parts: the feature extraction of AFB using Roboflow for AFB annotation and the feature extraction of background variations (good, less, or overstained). The middle part with the red box shows an enhanced AFB identification technique that combines AFB detection outcomes using YOLOv7 with background variation classification using Rep-VGG. The bottom part of the figure displays testing on MIDTI test data, where all 40 blind samples are categorized into IUATLD levels (Negative, Scanty, TB1+, TB2+, and TB3+).

phase. Following the ZN staining classification step, the system counted the number of AFBs in each FOV and accumulated them in the WSI to classify them according to IUATLD grades. A large amount of red degradation was caused by artifacts in the background, which were misclassified as AFB. The CNN-based algorithm RepVGG was then added to YOLOv7 for second-stage AFB detection (YOLOv7-RepVGG), which filtered the AFB and non-AFB from the first-stage AFB detection results. As shown by the yellow block in Fig.5, the AFB detection technique in the second step was the outcome of classifying sputum samples according to the IUATLD grades.

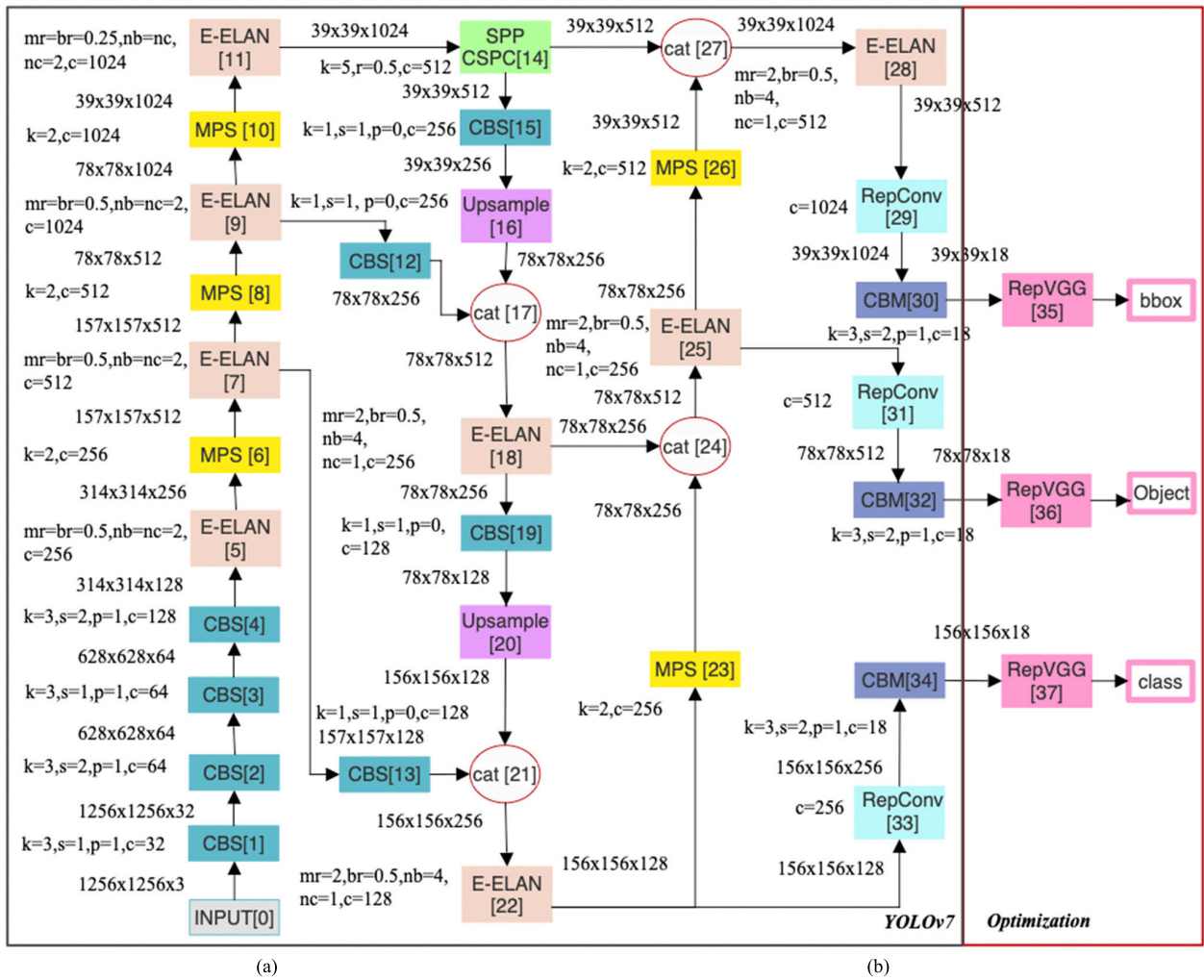
**E. PROPOSED MODEL**

The YOLO-v7, as a continued development over the earlier YOLO series, offers an excellent combination of accuracy and processing time. The YOLO-V7 network consists of three primary modules: backbone, neck, and head. In the YOLO-v7 backbone network, the optimization that reforms the previous YOLO is an Extended Efficient Layer Aggregation Networks (E-ELAN). E-ELAN is a computational block that can

proceed directly from different groups to acquire unique characteristics without breaking the original gradient route [73]. To preserve an ideal structural model, the YOLOv7 neck network is a new method for scaling concatenation-based models to satisfy various application requirements. This scaling changes some model features to produce models of various sizes, such as depth scaling (number of stages), width scaling (number of channels), and resolution scaling (input image size) [64]. The YOLOv7 architecture uses reparameterized convolution (Repconv) without an identity connection [74]. The YOLOv7 head network has three head blocks, consisting of a lead head and an auxiliary head. The head network in the previous YOLO only produced YOLO loss, based on the calculation of the prediction result. Meanwhile, in YOLOv7, the lead head is responsible for generating the final output, whereas the auxiliary head is responsible for supporting middle-layer training.

This study proposes a model to identify AFB by adding RepVGG to the head block of the YOLOv7 architecture (YOLOv7-RepVGG), as shown in Fig.6, which has the following architecture:





**FIGURE 6.** Block diagram of the proposed model: (a) Backbone-neck of YOLOv7 architecture; (b) insertion of RepVGG on YOLOv7-Head block. The backbone is the main body of the YOLOv7 network (steps 1-11), which produces the feature extractions. The feature extraction data from the backbone is blended and combined in steps 12 through 27 of the neck process, preparing it for the head prediction stage (steps 28-33). The red box illustrates the optimization process for identifying AFB by adding RepVGG to the head-block part of the YOLOv7 (steps 35-37). CBS is a combination of three components: Convolutional, Batch Normalization and SiLU, E-ELAN (Extended Efficient Layer Aggregation Networks) is a group of convolution, MP (Max Pooling) is an operation used to reduce the spatial dimensions of an image, SPPCSPC (Spatial Pyramid Pooling & Convolutional Spatial Pyramid Pooling) is an advance of pooling technique which more robust to object deformations, CBM is a combination of Convolutional, Batch Normalization and Sigmoid, RepConv (Reparameterization Convolution) is a reshaping of convolution layer to enhancing the feature maps, RepVGG (Reparameterization Visual Geometry Group) aims to improve the performance of feature maps by reshaping the VGG convolutional network.

- The input image was resized to a fixed size of  $1,256 \times 1,256$  pixels, which can lead to a loss of vital information such as blurring or loss of pixel information. To address this issue, YOLOv7 uses nine aspect ratio anchor boxes during training, which allows for the detection of a wider range of object shapes, colors, and sizes than previous YOLO versions. This technique helps make accurate predictions for objects of various shapes, colors, and sizes.
- After the image was resized, it was run through several convolutional layers to extract the features and save spatial information, such as Convolutional, Batch Normalization and SiLU (Sigmoid Linear Units) activation (CBS); Spatial Pyramid Pooling & Convolutional Spa-

- tial Pyramid Pooling (SPPCSPC); and Convolutional, Batch Normalization and Sigmoid (CBM).
- The suggested extended ELAN (E-ELAN) maintains the gradient transmission path of the original architecture while utilizing group convolution to improve the cardinality of the additional proposed features. It also reorganizes and merges the cardinality of the features of other groups. The architecture of the transition layer remains unchanged, whereas E-ELAN only modifies the architecture of the computing block. The E-ELAN technique increases the number of channels and cardinality of computing blocks using group convolution. In the computational layer, E-ELAN applied the same group of parameters and channel multipliers to each

computing block. The feature map produced from each computational block was concatenated after shuffling into groups ( $g$ ) in accordance with the specified group parameter ( $g$ ). At this point, the number of channels in each feature map group was equivalent to the number of channels in the original architecture. To complete the merging cardinality, E-ELAN adds ( $g$ ) groups of feature maps [64].

- YOLOv7 uses max pooling (MP) to compress the feature size.
- YOLOv7 uses reparameterized convolution (RepConv). RepConv is a convolutional layer that integrates the identity connections,  $3 \times 3$  convolution, and  $1 \times 1$  convolution.
- YOLOv7 uses a coarse auxiliary, and is suitable for lead loss. In the YOLOv7 output, as part of the information given for predicting the result of object detection, each grid cell predicts the bounding boxes ( $BB$ ) as well as the dimensions and confidence scores. The confidence score ( $conf.score$ ) determines the presence of an inside bounding box and the likelihood of prediction results with the ground truth. The presence of an object can have a probability  $p(object)$  range of 0 to 1. The likelihood was calculated based on the mean average precision of the likelihood, which in turn was calculated based on the mean average precision of the Intersection over Union (IoU) criteria ( $IoU_{pred}^{truth}$ ). The confidence score formula is calculated using (Equation 1-2) [54]:

$$conf.score = p(object) \times IoU_{pred}^{truth} \quad (1)$$

$$IoU_{pred}^{truth} = \frac{truth \cap pred}{truth \cup pred} \quad (2)$$

This algorithm can obtain multiple detections of the same object. As an example, it may identify three BBs for the same object, each with IoU probabilities of 0.8, 0.9, and 0.3. The loss function formula for YOLOv7 is given by (Equation 3) [55]:

$$\begin{aligned} \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{I}_{ij}^{obj} \left[ (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \\ + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{I}_{ij}^{obj} \left[ \left( \sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left( \sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right] \\ + \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{I}_{ij}^{obj} \left( C_i - \hat{C}_i \right)^2 \\ + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{I}_{ij}^{obj} \left( C_i - \hat{C}_i \right)^2 \\ + \sum_{i=0}^{S^2} \mathbb{I}_i^{obj} \sum_{c \in classes} (p_i(c) - \hat{p}_i(c))^2 \end{aligned} \quad (3)$$

where  $\lambda_{coord}$  is a constant value of bounding box coordinate predictions and  $\lambda_{noobj}$  is a constant value of confidence

## NMS Algorithm

Procedure NMS (B,c)

B is the list of initial detection boxes,

$B_{nms} \leftarrow \emptyset$  # empty set

$discard \leftarrow False$  # discarded

**for**  $b_i \in B$  **do** # Repeat over every single box.

**for**  $b_j \in B$  **do** # begin a new loop for comparison with  $b_i$

**if** same ( $b_j, b_i$ )  $> \lambda_{nms}$  **then** #  $IoU_{b_j} > IoU_{b_i}$

**if** score ( $c, b_j$ )  $>$  score ( $c, b_i$ ) **then** #  $IoU_{b_j} > IoU_{b_i}$

$discard \leftarrow True$  # discarded  $b_i$

**if not**  $discard$  **then**

$B_{nms} \leftarrow B_{nms} \cup b_i$  # select  $b_i$  as final

**return**  $B_{nms}$  # Repeat all steps for the remaining boxes, then return to the final result.

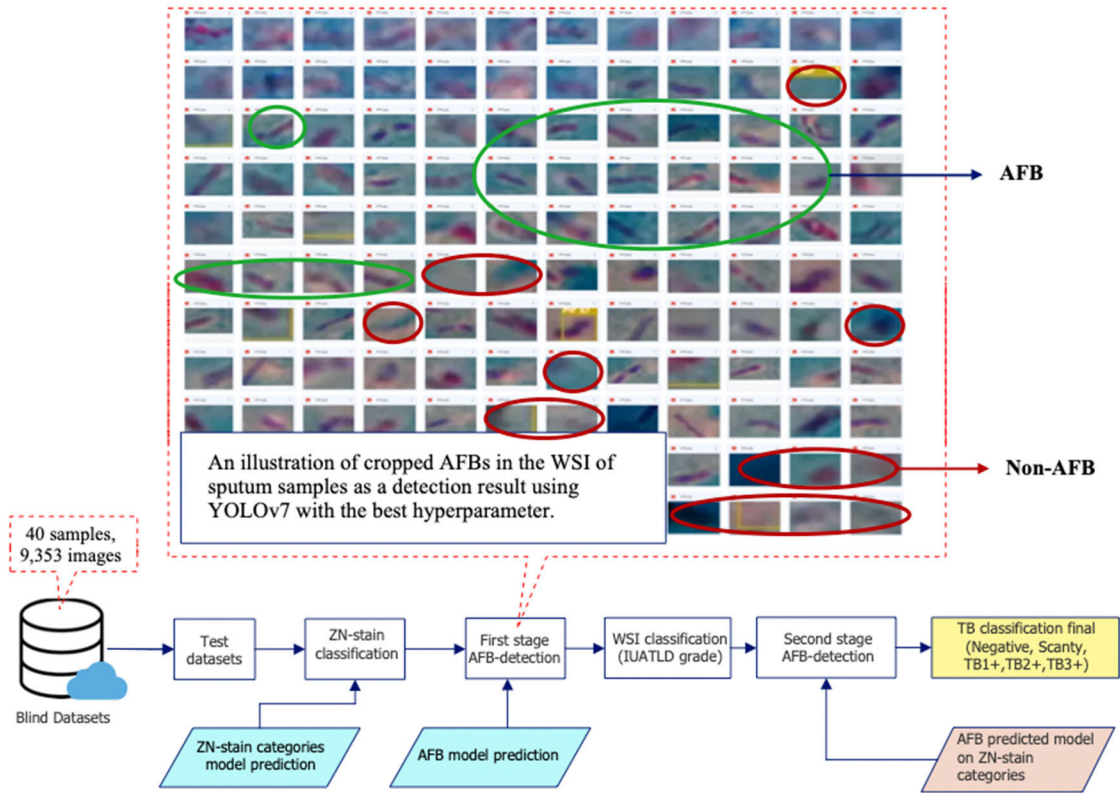
predictions for boxes that do not contain objects.  $S^2 = S_x S_y$  is the size of the grid in the input image,  $B$  is the number of boxes.  $\mathbb{I}_i^{obj}$  denotes if object ( $obj$ ) appears in grid  $cell-i$ ,  $\mathbb{I}_{ij}^{obj}$  indicates the  $cell-j^{th}$   $BB$  prediction in the  $cell-i$  controls for prediction.  $(x_i, y_i)$  is the box centre in the grid cell, and  $(\hat{x}_i, \hat{y}_i)$  its normalization,  $(w, h)$  is the predicted width and height according to the entire image and  $(\hat{w}_i, \hat{h}_i)$  is its normalization.  $C_i$  is class probabilities, and  $\hat{C}_i$  is its normalization.  $p_i(c)$  and  $\hat{p}_i(c)$  are the predicted class probabilities ( $c$ ) in the grid  $cell-i$  and its normalization, respectively.

- YOLOv7 uses non-maximum suppression (NMS) to delete duplicates, which selects the box with the highest IoU probability as the prediction object. The NMS pseudocode is as follows [74].
- RepVGG uses  $W^{(1)} \in \mathbb{R}^{C_1 \times C_2 \times 3 \times 3}$  for the kernel of a  $1 \times 1$  branch and  $W^{(3)} \in \mathbb{R}^{C_2 \times C_1 \times 3 \times 3}$  for the kernel of a  $3 \times 3$  convolution (conv) layer, where  $C_1$  and  $C_2$  are the input and output channels, respectively. RepVGG use  $\mu^3, \sigma^3, \gamma^3, \beta^3$  represents the accumulated mean, standard deviation, learning scaling factor, and bias of the batch normalization (BN) layer following  $3 \times 3$  conv. Meanwhile, RepVGG use  $\mu^1, \sigma^1, \gamma^1, \beta^1$  for the BN following  $1 \times 1$  conv, and use  $\mu^0, \sigma^0, \gamma^0, \beta^0$  for the identity branch. Let  $*$  be the convolution operator,  $M^1 \in \mathbb{R}^{N \times C_1 \times H_1 \times W_1}$ ,  $M^2 \in \mathbb{R}^{N \times C_2 \times H_2 \times W_2}$  be the input and output, respectively. If  $C_1 = C_2, H_1 = H_2, W_1 = W_2$ , then  $M^2$  is defined by [48].

$$\begin{aligned} M^2 = bn \left( M^1 * W^3, \mu^3, \sigma^3, \gamma^3, \beta^3 \right) \\ + bn \left( M^1 * W^1, \mu^1, \sigma^1, \gamma^1, \beta^1 \right) \\ + bn \left( M^1, \mu^0, \sigma^0, \gamma^0, \beta^0 \right) \end{aligned} \quad (4)$$

Nevertheless, this study did not use the identity branch; therefore, there are only two terms in the eq.4. Formally,  $\forall 1 < i < C_2$ , in which  $bn$  is the inference-time BN function defined by (Equation 5)

$$\begin{aligned} bn = (M, \mu, \sigma, \gamma, \beta)_{:,i,:} \\ = (M_{:,i,:} - \mu_i) \frac{\gamma_i}{\sigma_i} + \beta_i \end{aligned} \quad (5)$$



**FIGURE 7.** Several cropped non-AFB objects misclassified as AFB (red markers) and true classified AFB (green markers) on the sputum WSI in the upper part. The bottom of the figure illustrates the AFB detection process in each field of view and the classification of sputum samples on MIDTI based on IUATLD grades.

**F. PERFORMANCE MEASUREMENT**

The following equations (6)–(9) were used to test how well the suggested method classified the ZN-stained sputum samples into five classes of TB based on the IUATLD grade.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

$$\text{Recall(Sensitivity)} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (8)$$

$$\text{F1-score} = \frac{2 \times \text{precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (10)$$

where:

- True Positive (TP): Positive classes were classified as positive, scanty as scanty, TB1+ as TB1+, TB2+ as TB2+, and TB3+ as TB3+.
- True Negative (TN): The negative class was classified as negative.
- False Negative (FN): A negative class was classified as positive (scanty, TB1+, TB2+, or TB3+).
- False Positive (FP): A positive class (Scanty, TB1+, TB2+, or TB3+) was classified as a negative class.

**IV. RESULTS**

In the first-stage (stage-1) of AFB detection, YOLOv7 was used with the best hyperparameters to classify sputum smears based on the IUATLD grades for all MIDTI test data. The best hyperparameters used in this study were as follows: batch size = 4, epochs = 110, 0.01 learning rate, 2/grid for the number of anchors, IoU threshold = 0.5, and confidence score = 0.17. The test results on all datasets (all MIDTI test data) achieved an accuracy level of 77.5% and had low values in the negative 50%, scanty 62.5%, and TB1+ 75%, as shown in Table 4. A specificity value of 44.44% in the first experiment showed that out of the 40 sputum samples examined using YOLOv7, the system successfully classified 4/8 of negative class samples. This problem occurs because many objects detected were misclassified as AFB, as illustrated in Fig.7 shows the total number of objects detected as AFB by YOLOv7 from a sputum sample. Based on the validation results by the pathologist, green round marks represent true AFB and red round marks represent non-AFB. This misclassification of AFB significantly affects the ultimate outcomes of the IUATLD scale-based sputum sample classification. Meanwhile, the TB2+ and TB3+ classifications achieved the highest accuracy of 100%. This happened because there were a lot of AFB/FoV samples in the TB2+ and TB3+ classes, which was still close to

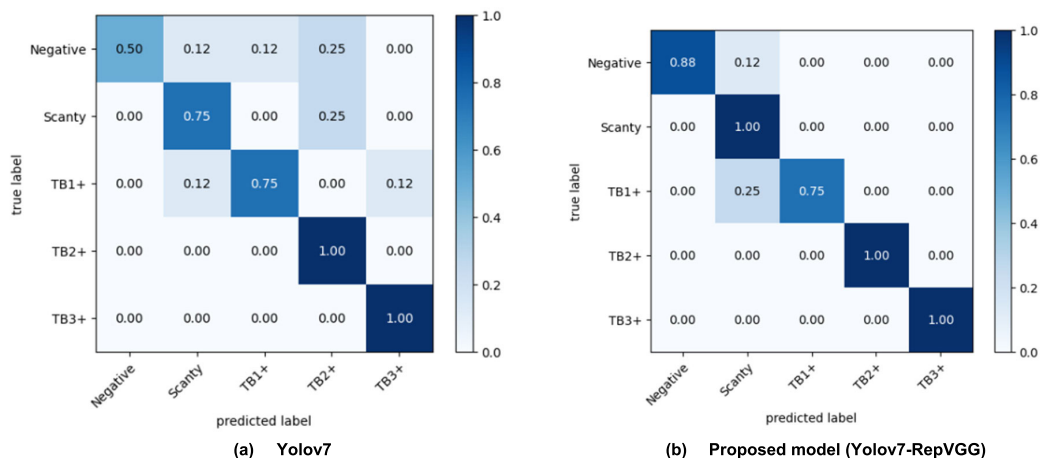


FIGURE 8. Confusion matrix classification results of 40 sputum samples for each class based on the IUATLD scale.

TABLE 4. ZN-stained sputum sample classification based on the IUATLD grade ( $\sum s =$  number of samples).

IUATLD-grades	(a) Yolov7 on All test dataset		(b) Yolov7 on ZN-Filtered dataset	
	$\sum s$	Acc.	$\sum s$	Acc.
Negative	8	50%	5	80%
Scanty	8	62.50%	7	71%
TB 1+	8	75%	6	100%
TB 2+	8	100%	8	100%
TB 3+	8	100%	8	100%
Total	40		34	
Accuracy		77.5%		91.18%
Sensitivity		87.10%		96.43%
Specificity		44.44%		57.14%

the minimum accumulated amount of AFB/sputum samples based on the IUATLD scale.

Based on the first-stage AFB detection, we concluded that the classification process for sputum samples cannot be done if the ZN-stained sample is excessively dark or overstained. Table 3 (column b) shows the results of the test, which was performed using only 34 sputum samples. Sputum samples were selected based on ZN categories with good staining quality in the MIDTI test data only. When the YOLOv7 was used to detect AFB, the accuracy increased from 77.5% to 91.18%, the sensitivity increased from 87.10% to 96.43%, and the specificity increased from 44.44% to 57.14%. This proved that the difference in background color due to the quality of ZN staining could lead to misclassification of AFB.

As an optimization step (AFB detection stage-2) to increase the performance of sensitivity and specificity measurements, the main idea was to reduce the number of misclassified AFBs from the AFB detection stage-1 result.

The proposed model for AFB detection in stage-2 is illustrated in the red block in Fig.5 and 6. The optimization stages are as follows:

1. The features of AFB characteristics for each ZN staining category (good staining, less staining, and

TABLE 5. Measurement of sputum sample (MIDTI) classification based on the IUATLD scale.

Parameter	Yolov7	Proposed Model
No. Sample	40	40
True Positive	22	30
False Positive	10	2
True Negative	4	7
False Negative	4	1
Precision	68.75%	93.75%
Recall	84.62%	96.77%
F1-score	75.86%	95.24%
Specificity	28.57%	77.78%
Accuracy	65%	92.50%

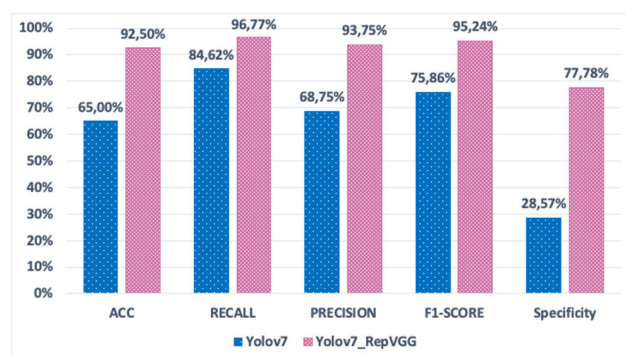


FIGURE 9. Comparison of sputum classification samples using YOLOV7 and YOLOV7-RepVGG.

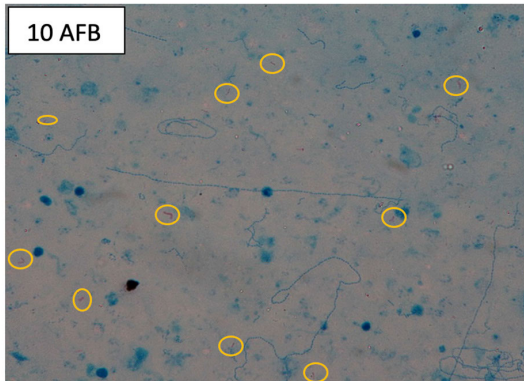
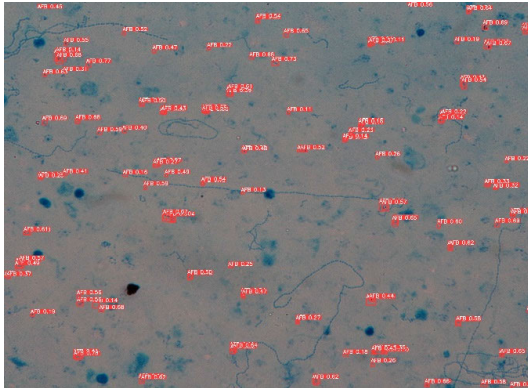
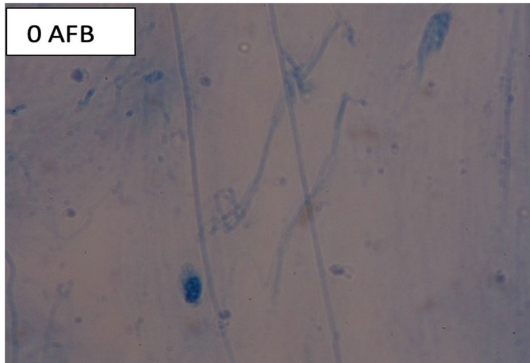
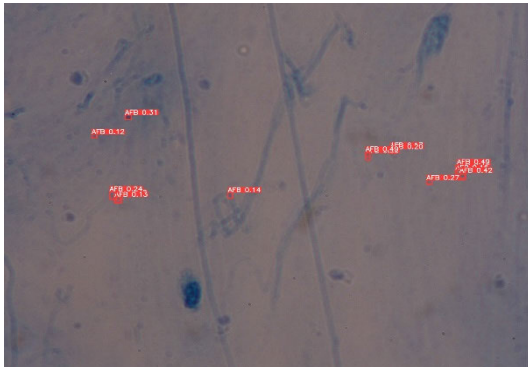
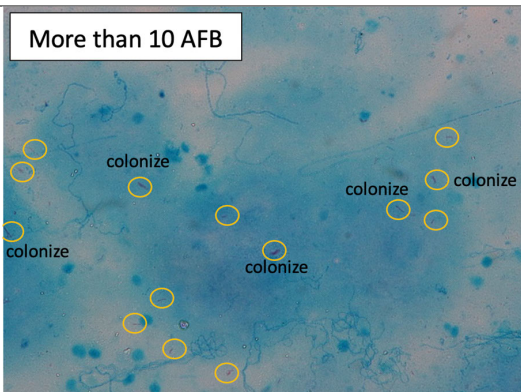
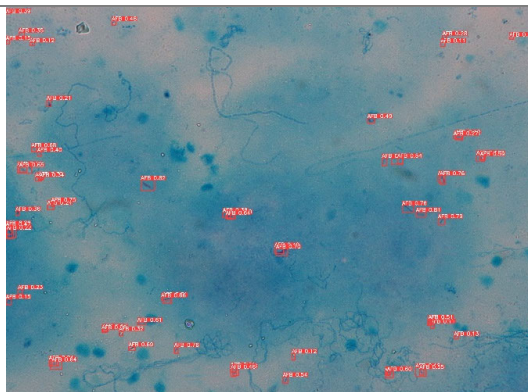
overstaining) were extracted using the RepVGG algorithm.

2. Adding RepVGG modelling results in the first point above the head-block part of the YOLOv7 algorithm (YOLOv7-RepVGG).

Finally, the proposed model was used for the AFB detection stage-2. The test was conducted to classify all 40 sputum samples on the MIDTI (test data) into IUATLD grade. Fig.8 shows the comparison of the performance results of Yolov7 and our proposed model for the five-class IUATLD.

Table 5 and Fig.9 display the output values of the performance results obtained from testing using the Yolov7

**TABLE 6.** Results of proposed model of AFB detection in MIDTI using YOLOv7-RepVGG.

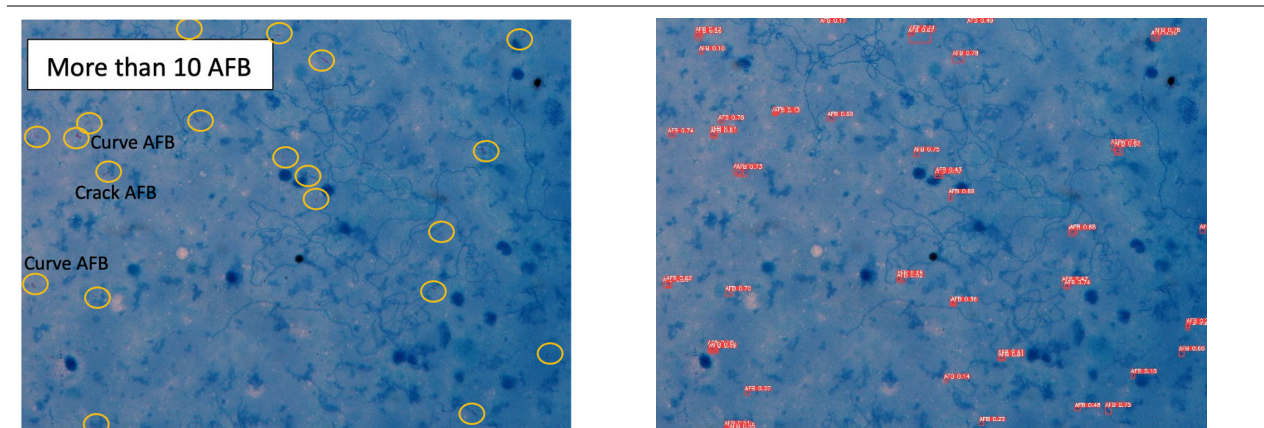
<b>ZN staining category: Less staining</b>	
Annotation AFB by pathologists	AFB detection results using YOLOv7-RepVGG
	
	
<b>ZN staining category: Good staining</b>	
Annotation AFB by pathologist	AFB detection results using YOLOv7-RepVGG
	

and proposed models. The visualization of the proposed model resulted (AFB detection per FoV), as shown in Table 6.

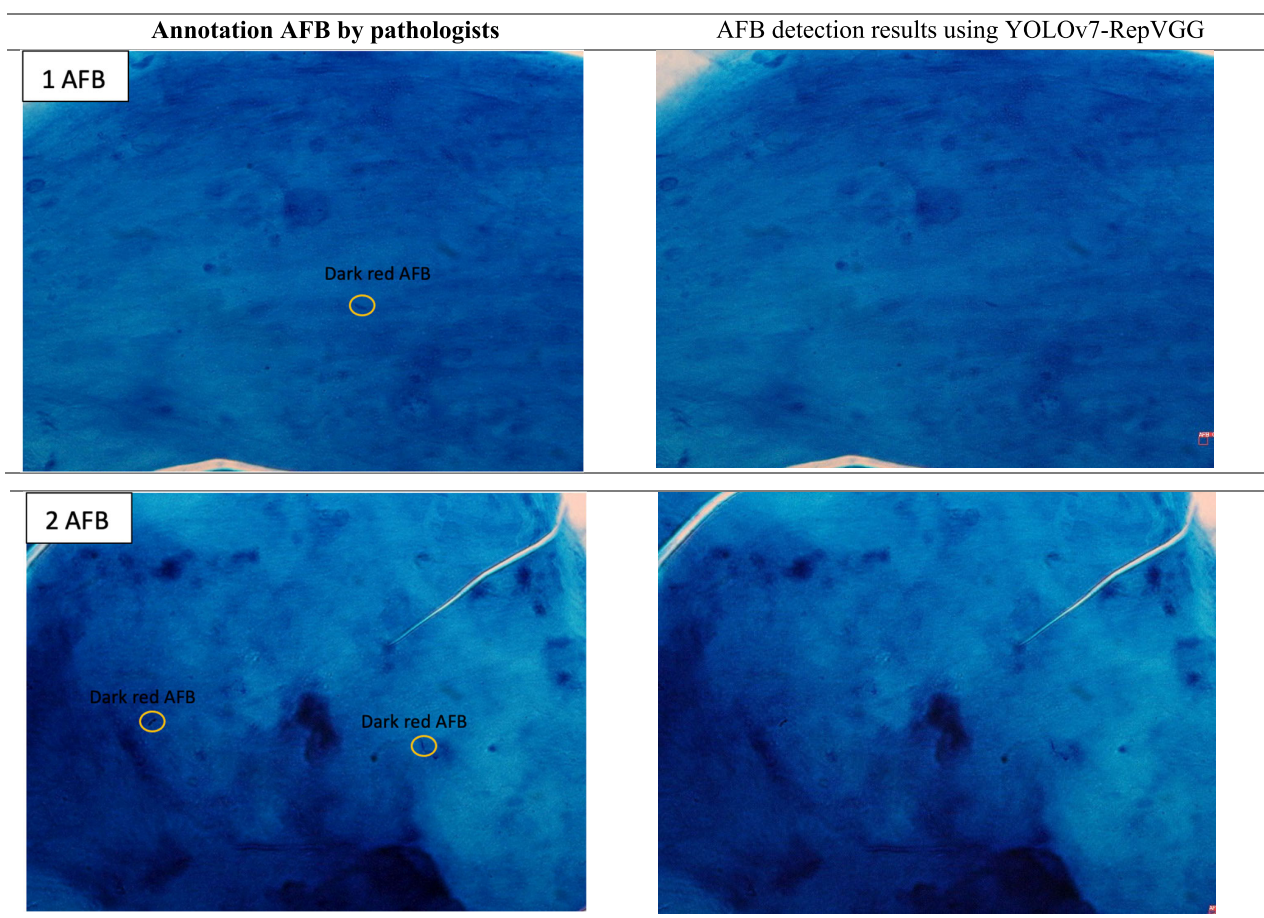
As shown in Table 4, the proposed model outperformed YOLOv7 in every performance metric, including the accuracy, specificity, recall or sensitivity, F1-score, precision,

and recall. As previously mentioned, the test results for class TB2+ and TB3+ sputum sample calcification had a high level of sensitivity using both YOLOv7 and YOLO-Rep VGG, this was conducted because the number of AFB detected still close to the minimum accumulated

**TABLE 6.** (Continued.) Results of proposed model of AFB detection in MIDTI using YOLOv7-RepVGG.



ZN staining category: Overstaining



amount of AFB/sputum samples based on the IUATLD scale (Table 1).

The classification of sputum samples as TB2+ based on IUATLD standardization (Table 1) is that the average number detected is approximately 1–9 AFB/FoV for the first 50 FoVs, and TB3+ is if the average number detected is more than 10 AFB/FoV for the first 20 FoVs.

The large number of bacteria in the TB2+ and TB3+ classes often causes AFB colonization. However, the number of colonized AFBs was not accurately calculated experimentally; instead, they were considered as a single AFB object in the proposed model, which presents a challenge for future research.

## V. CONCLUSION

This study successfully built a database of microscopic images from sputum smear samples of Indonesian patients called MIDTI. MIDTI can facilitate the development of computer-aided diagnosis (CAD) of tuberculosis. The MIDTI consists of 15,067 microscopic images acquired from 67 sputum samples obtained from the WSI. The dataset image size was  $3,072 \times 2,304$  pixels in jpeg format.

A total of 67 sputum samples were obtained from RC3ID UNPAD, and microscopic image acquisition was performed at 20–300 viewing points for each sample according to the WHO standard. The acquisition system used facilities from the ITB Biomedical Laboratory, including an Olympus CX-31 microscope connected to a modified DSLR 700D camera. This study also proposed a model using a method based on the YOLOv7 to detect AFB as part of CAD for sorting sputum samples in MIDTI into IUATLD grades. The results showed that the proposed model could help pathologists diagnose TB in sputum samples as negative, scanty, TB1+, TB2+, and TB3+. However, this study had limitations in terms of accurately counting AFB, especially if they were present in colonies. This could lead to challenges in precisely determining the number of AFB objects in future research.

## APPENDIX

Table 6 shows the visualization of the proposed model that detects AFB per FoV. The left column shows the annotation AFB by three pathologists (NR, MI, and IZ) with varying experience levels in the TB laboratory RC3ID-UNPAD. The column on the right shows the AFB detection results obtained using the proposed model.

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

## SUPPLEMENTARY MATERIALS

The Microscopic Imaging Database of Tuberculosis Indonesia (MIDTI) is available database, which can be accessed at [https://drive.google.com/drive/folders/1CB1jJbw38yavWcMo3azPeSrVk9R9g\\_iW?usp=sharing](https://drive.google.com/drive/folders/1CB1jJbw38yavWcMo3azPeSrVk9R9g_iW?usp=sharing).

The AFB annotations on Roboflow are available at <https://universe.roboflow.com/suci-aulia/afb-test-y7>

## REFERENCES

- [1] *Global Tuberculosis Report 2022*, WHO, Geneva, Switzerland, 2022.
- [2] J. Chakaya, M. Khan, F. Ntoumi, E. Aklillu, R. Fatima, P. Mwaba, N. Kapata, S. Mfinanga, S. E. Hasnain, P. D. M. C. Katoto, A. N. H. Bulabula, N. A. Sam-Agudu, J. B. Nachega, S. Tiberi, T. D. McHugh, I. Abubakar, and A. Zumla, "Global tuberculosis report 2020—Reflections on the global TB burden, treatment and prevention efforts," *Int. J. Infectious Diseases*, vol. 113, pp. 7–12, Dec. 2021, doi: 10.1016/j.ijid.2021.02.107.
- [3] S. Urooj, S. Suchitra, L. Krishnasamy, N. Sharma, and N. Pathak, "Stochastic learning-based artificial neural network model for an automatic tuberculosis detection system using chest X-ray images," *IEEE Access*, vol. 10, pp. 103632–103643, 2022, doi: 10.1109/ACCESS.2022.3208882.
- [4] T. Xu and Z. Yuan, "Convolution neural network with coordinate attention for the automatic detection of pulmonary tuberculosis images on chest X-rays," *IEEE Access*, vol. 10, pp. 86710–86717, 2022, doi: 10.1109/ACCESS.2022.3199419.
- [5] R. Mehrrotraa, M. A. Ansari, R. Agrawal, P. Tripathi, M. B. Bin Heyat, M. Al-Sarem, A. Y. M. Muaad, W. A. E. Nagmeldin, A. Abdelmaboud, and F. Saeed, "Ensembling of efficient deep convolutional networks and machine learning algorithms for resource effective detection of tuberculosis using thoracic (Chest) radiography," *IEEE Access*, vol. 10, pp. 85442–85458, 2022, doi: 10.1109/ACCESS.2022.3194152.
- [6] N. Ravin, S. Saha, A. Schweitzer, A. Elahi, F. Dako, D. Mollura, and D. Chapman, "Mitigating domain shift in AI-based TB screening with unsupervised domain adaptation," *IEEE Access*, vol. 10, pp. 45997–46013, 2022, doi: 10.1109/ACCESS.2022.3168680.
- [7] W. Chokchaithanakul, P. Punyabukkana, and E. Chuangsuwanich, "Adaptive image preprocessing and augmentation for tuberculosis screening on out-of-domain chest X-ray dataset," *IEEE Access*, vol. 10, pp. 132144–132152, 2022, doi: 10.1109/ACCESS.2022.3229591.
- [8] F. Kulwa, C. Li, X. Zhao, B. Cai, N. Xu, S. Qi, S. Chen, and Y. Teng, "A state-of-the-art survey for microorganism image segmentation methods and future potential," *IEEE Access*, vol. 7, pp. 100243–100269, 2019, doi: 10.1109/ACCESS.2019.2930111.
- [9] Z. Ul Abideen, M. Ghafoor, K. Munir, M. Saqib, A. Ullah, T. Zia, S. A. Tariq, G. Ahmed, and A. Zahra, "Uncertainty assisted robust tuberculosis identification with Bayesian convolutional neural networks," *IEEE Access*, vol. 8, pp. 22812–22825, 2020, doi: 10.1109/ACCESS.2020.2970023.
- [10] K. Munadi, K. Mughtar, N. Maulina, and B. Pradhan, "Image enhancement for tuberculosis detection using deep learning," *IEEE Access*, vol. 8, pp. 217897–217907, 2020, doi: 10.1109/ACCESS.2020.3041867.
- [11] A. F. M. Saif, T. Imtiaz, C. Shahnaz, W.-P. Zhu, and M. O. Ahmad, "Exploiting cascaded ensemble of features for the detection of tuberculosis using chest radiographs," *IEEE Access*, vol. 9, pp. 112388–112399, 2021, doi: 10.1109/ACCESS.2021.3102077.
- [12] T. Rahman, A. Khandakar, M. A. Kadir, K. R. Islam, K. F. Islam, R. Mazhar, T. Hamid, M. T. Islam, S. Kashem, Z. B. Mahbub, M. A. Ayari, and M. E. H. Chowdhury, "Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization," *IEEE Access*, vol. 8, pp. 191586–191601, 2020, doi: 10.1109/ACCESS.2020.3031384.
- [13] Y. Cao, J. Mao, H. Yu, Q. Zhang, H. Wang, Q. Zhang, L. Guo, and F. Gao, "A novel hybrid active contour model for intracranial tuberculosis MRI segmentation applications," *IEEE Access*, vol. 8, pp. 149569–149585, 2020, doi: 10.1109/ACCESS.2020.3016746.
- [14] D. Azadi, T. Motallebirad, K. Ghaffari, and H. Shojaei, "Mycobacteriosis and tuberculosis: Laboratory diagnosis," *Open Microbiol. J.*, vol. 12, no. 1, pp. 41–58, Mar. 2018, doi: 10.2174/1874285801812010041.
- [15] S. Zurac, C. Mogodici, T. Poncu, M. Trăscău, C. Popp, L. Nichita, M. Cioplea, B. Ceachi, L. Sticlaru, A. Cioroianu, M. Busca, O. Stefan, I. Tudor, A. Voicu, D. Stanescu, P. Mustatea, C. Dumitru, and A. Bastian, "A new artificial intelligence-based method for identifying mycobacterium tuberculosis in Ziehl–Neelsen stain on tissue," *Diagnostics*, vol. 12, no. 6, p. 1484, Jun. 2022, doi: 10.3390/diagnostics12061484.
- [16] M. Yang, K. Nurzynska, A. E. Walts, and A. Gertych, "A CNN-based active learning framework to identify mycobacteria in digitized Ziehl–Neelsen stained human tissues," *Computerized Med. Imag. Graph.*, vol. 84, Sep. 2020, Art. no. 101752, doi: 10.1016/j.compmedimag.2020.101752.
- [17] Y. N. Law, H. Jian, N. W. S. Lo, M. Ip, M. M. Y. Chan, K. M. Kam, and X. Wu, "Low cost automated whole smear microscopy screening system for detection of acid fast bacilli," *PLoS ONE*, vol. 13, no. 1, Jan. 2018, Art. no. e0190988, doi: 10.1371/journal.pone.0190988.
- [18] G. Rasool, A. M. Khan, R. Mohy-Ud-Din, and M. Riaz, "Detection of mycobacterium tuberculosis in AFB smear-negative sputum specimens through MTB culture and GeneXpert MTB/RIF assay," *Int. J. Immunopathol. Pharmacol.*, vol. 33, Jan. 2019, Art. no. 205873841982717, doi: 10.1177/2058738419827174.
- [19] M. Singh, G. V. Pujar, S. A. Kumar, M. Bhagyalalitha, H. S. Akshatha, B. Abuhajja, A. R. Alsoud, L. Abualigah, N. M. Beeraka, and A. H. Gandomi, "Evolution of machine learning in tuberculosis diagnosis: A review of deep learning-based medical applications," *Electronics*, vol. 11, no. 17, p. 2634, Aug. 2022, doi: 10.3390/electronics11172634.
- [20] K. Ratsavong, F. Quet, F. Nzabintwali, J. Diendéré, J. Seberty, M. Strobel, and Y. Buisson, "Usefulness and limits of Ziehl–Neelsen staining to detect paragonimiasis in highly endemic tuberculosis areas," *Parasite Epidemiology Control*, vol. 2, no. 1, pp. 1–7, Feb. 2017, doi: 10.1016/j.parepi.2016.12.001.

- [21] J. L. Díaz-Huerta, A. D. C. Téllez-Anguiano, M. Fraga-Aguilar, J. A. Gutiérrez-Gnecchi, and S. Arellano-Calderón, "Image processing for AFB segmentation in bacilloscopies of pulmonary tuberculosis diagnosis," *PLoS ONE*, vol. 14, no. 7, Jul. 2019, Art. no. e0218861, doi: [10.1371/journal.pone.0218861](https://doi.org/10.1371/journal.pone.0218861).
- [22] L. Pantanowitz, U. Wu, L. Seigh, E. LoPresti, F.-C. Yeh, P. Salgia, P. Michelow, S. Hazelhurst, W.-Y. Chen, D. Hartman, and C.-Y. Yeh, "Artificial intelligence-based screening for mycobacteria in whole-slide images of tissue samples," *Amer. J. Clin. Pathol.*, vol. 156, no. 1, pp. 117–128, Jun. 2021, doi: [10.1093/ajcp/aaqaa215](https://doi.org/10.1093/ajcp/aaqaa215).
- [23] C.-M. Lo, Y.-H. Wu, Y.-C. Li, and C.-C. Lee, "Computer-aided Bacillus detection in whole-slide pathological images using a deep convolutional neural network," *Appl. Sci.*, vol. 10, no. 12, p. 4059, Jun. 2020, doi: [10.3390/app10124059](https://doi.org/10.3390/app10124059).
- [24] A. U. Ibrahim, E. Guler, M. Guvenir, K. Suer, S. Serte, and M. Ozsoz, "Automated detection of mycobacterium tuberculosis using transfer learning," *J. Infection Developing Countries*, vol. 15, no. 5, pp. 678–686, May 2021, doi: [10.3855/jidc.13532](https://doi.org/10.3855/jidc.13532).
- [25] Y. Zaizen, Y. Kanahori, S. Ishijima, Y. Kitamura, H.-S. Yoon, M. Ozasa, H. Mukae, A. Bychkov, T. Hoshino, and J. Fukuoka, "Deep-learning-aided detection of mycobacteria in pathology specimens increases the sensitivity in early diagnosis of pulmonary tuberculosis compared with bacteriology tests," *Diagnostics*, vol. 12, no. 3, p. 709, Mar. 2022, doi: [10.3390/diagnostics12030709](https://doi.org/10.3390/diagnostics12030709).
- [26] H.-C. Huang, K.-L. Kuo, M.-H. Lo, H.-Y. Chou, and Y. E. Lin, "Novel TB smear microscopy automation system in detecting acid-fast bacilli for tuberculosis—A multi-center double blind study," *Tuberculosis*, vol. 135, Jul. 2022, Art. no. 102212, doi: [10.1016/j.tube.2022.102212](https://doi.org/10.1016/j.tube.2022.102212).
- [27] H.-T. Fu, H.-Z. Tu, H.-S. Lee, Y. E. Lin, and C.-W. Lin, "Evaluation of an AI-based TB AFB smear screening system for laboratory diagnosis on routine practice," *Sensors*, vol. 22, no. 21, p. 8497, Nov. 2022, doi: [10.3390/s22218497](https://doi.org/10.3390/s22218497).
- [28] K. Nurzynska, D. Li, A. E. Walts, and A. Gertych, "Multilayer outperforms single-layer slide scanning in AI-based classification of whole slide images with low-burden acid-fast mycobacteria (AFB)," *Comput. Methods Programs Biomed.*, vol. 234, Jun. 2023, Art. no. 107518, doi: [10.1016/j.cmpb.2023.107518](https://doi.org/10.1016/j.cmpb.2023.107518).
- [29] J. J. Lewis, V. N. Chihota, M. van der Meulen, P. B. Fourie, K. L. Fielding, A. D. Grant, S. E. Dorman, and G. J. Churchyard, "'Proof-of-concept' evaluation of an automated sputum smear microscopy system for tuberculosis diagnosis," *PLoS ONE*, vol. 7, no. 11, Nov. 2012, Art. no. e50173, doi: [10.1371/journal.pone.0050173](https://doi.org/10.1371/journal.pone.0050173).
- [30] 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," *Biocybernetics Biomed. Eng.*, vol. 38, no. 3, pp. 691–699, 2018, doi: [10.1016/j.bbe.2018.05.007](https://doi.org/10.1016/j.bbe.2018.05.007).
- [31] S. Aulia, T. Mengko, A. B. Suksmono, and B. Alisjahbana, "Artefact segmentation in microscopy images for automatic acid-fast bacilli quantification," *ICIC Exp. Lett.*, vol. 17, no. 3, pp. 367–374, 2023, doi: [10.24507/icicel.17.03.367](https://doi.org/10.24507/icicel.17.03.367).
- [32] R. Khutlang, S. Krishnan, A. Whitelaw, and T. S. Douglas, "Detection of tuberculosis in sputum smear images using two one-class classifiers," in *Proc. IEEE Int. Symp. Biomed. Imag., From Nano Macro*, Jun. 2009, pp. 1007–1010, doi: [10.1109/ISBI.2009.5193225](https://doi.org/10.1109/ISBI.2009.5193225).
- [33] V. Makkapati, R. Agrawal, and R. Acharya, "Segmentation and classification of tuberculosis bacilli from ZN-stained sputum smear images," in *Proc. IEEE Int. Conf. Autom. Sci. Eng.*, Aug. 2009, pp. 217–220, doi: [10.1109/COASE.2009.5234173](https://doi.org/10.1109/COASE.2009.5234173).
- [34] 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," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, Oct. 2010, pp. 4049–4055, doi: [10.1109/ICSMC.2010.5642191](https://doi.org/10.1109/ICSMC.2010.5642191).
- [35] R. Khutlang, S. Krishnan, R. Dendere, A. Whitelaw, K. Veropoulos, G. Learmonth, and T. S. Douglas, "Classification of mycobacterium tuberculosis in images of ZN-stained sputum smears," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 4, pp. 949–957, Jul. 2010, doi: [10.1109/TTTB.2009.2028339](https://doi.org/10.1109/TTTB.2009.2028339).
- [36] R. Nayak, V. P. Shenoy, and R. R. Galigeke, "A new algorithm for automatic assessment of the degree of TB-infection using images of ZN-stained sputum smear," in *Proc. Int. Conf. Syst. Med. Biol.*, Dec. 2010, pp. 294–299, doi: [10.1109/ICSMB.2010.5735390](https://doi.org/10.1109/ICSMB.2010.5735390).
- [37] Y. Zhai, Y. Liu, D. Zhou, and S. Liu, "Automatic identification of mycobacterium tuberculosis from ZN-stained sputum smear: Algorithm and system design," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, Dec. 2010, pp. 41–46, doi: [10.1109/ROBIO.2010.5723300](https://doi.org/10.1109/ROBIO.2010.5723300).
- [38] R. Rulaningtyas, A. B. Suksmono, and T. L. R. Mengko, "Automatic classification of tuberculosis bacteria using neural network," in *Proc. Int. Conf. Electr. Eng. Informat.*, Jul. 2011, pp. 1–4, doi: [10.1109/ICEEL.2011.6021502](https://doi.org/10.1109/ICEEL.2011.6021502).
- [39] A. Edafe, O. O. Mumini, and S. Oluwarotimi, "A genetic-neuro-fuzzy inferential technique for diagnosis of tuberculosis," in *Proc. Workshop Pervasive Wireless Healthcare*, New York, NY, USA, Jun. 2015, pp. 39–44, doi: [10.1145/2757290.2757299](https://doi.org/10.1145/2757290.2757299).
- [40] 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," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 2841–2844, doi: [10.1109/EMBC.2014.6944215](https://doi.org/10.1109/EMBC.2014.6944215).
- [41] R. Rulaningtyas, A. B. Suksmono, T. L. R. Mengko, and P. Saptawati, "Nearest patch matching for color image segmentation supporting neural network classification in pulmonary tuberculosis identification," in *Proc. AIP Conf.*, 2016, Art. no. 120002, doi: [10.1063/1.4943354](https://doi.org/10.1063/1.4943354).
- [42] M. I. Shah, S. Mishra, M. Sarkar, and S. K. Sudarshan, "Automatic detection and classification of tuberculosis bacilli from camera-enabled smartphone microscopic images," in *Proc. 4th Int. Conf. Parallel, Distrib. Grid Comput. (PDGC)*, Dec. 2016, pp. 287–290, doi: [10.1109/PDGC.2016.7913161](https://doi.org/10.1109/PDGC.2016.7913161).
- [43] R. Rulaningtyas, A. B. Suksmono, T. L. R. Mengko, and P. Saptawati, "Colour segmentation of multi variants tuberculosis sputum images using self organizing map," *J. Phys., Conf.*, vol. 853, May 2017, Art. no. 012012, doi: [10.1088/1742-6596/853/1/012012](https://doi.org/10.1088/1742-6596/853/1/012012).
- [44] M. I. Shah, S. Mishra, V. K. Yadav, A. Chauhan, M. Sarkar, S. K. Sharma, and C. Rout, "Ziehl–Neelsen sputum smear microscopy image database: A resource to facilitate automated bacilli detection for tuberculosis diagnosis," *J. Med. Imag.*, vol. 4, no. 2, Jun. 2017, Art. no. 027503, doi: [10.1117/1.jmi.4.2.027503](https://doi.org/10.1117/1.jmi.4.2.027503).
- [45] 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, Nov. 2021, doi: [10.1016/j.jksuci.2018.08.004](https://doi.org/10.1016/j.jksuci.2018.08.004).
- [46] C. Kuok, M. Horng, Y. Liao, N. Chow, and Y. Sun, "An effective and accurate identification system of mycobacterium tuberculosis using convolution neural networks," *Microsc. Res. Technique*, vol. 82, no. 6, pp. 709–719, Jun. 2019, doi: [10.1002/jemt.23217](https://doi.org/10.1002/jemt.23217).
- [47] S. Kant and M. M. Srivastava, "Towards automated tuberculosis detection using deep learning," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Nov. 2018, pp. 1250–1253, doi: [10.1109/SSCI.2018.8628800](https://doi.org/10.1109/SSCI.2018.8628800).
- [48] M. I. Shah, S. Mishra, M. Sarkar, and C. Rout, "Automatic detection and classification of tuberculosis bacilli from ZN-stained sputum smear images using watershed segmentation," in *Proc. Int. Conf. Signal Process.*, 2018, pp. 1–4, doi: [10.1049/cp.2016.1459](https://doi.org/10.1049/cp.2016.1459).
- [49] R. Raof, M. Y. Mashor, and S. S. M. Noor, "Segmentation of TB bacilli in Ziehl–Neelsen sputum slide images using k-means clustering technique," *CSRID*, vol. 9, no. 2, p. 63, Mar. 2018, doi: [10.22303/csr.9.2.2017.63-72](https://doi.org/10.22303/csr.9.2.2017.63-72).
- [50] C. Del Carpio, E. Dianderas, M. Zimic, P. Sheen, J. Coronel, R. Lavarello, and G. Kemper, "An algorithm for detection of tuberculosis bacilli in Ziehl–Neelsen sputum smear images," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 9, no. 4, p. 2968, Aug. 2019, doi: [10.11591/ijece.v9i4.pp2968-2981](https://doi.org/10.11591/ijece.v9i4.pp2968-2981).
- [51] M. El-Melegy, D. Mohamed, T. ElMelegy, and M. Abdelrahman, "Identification of tuberculosis bacilli in ZN-stained sputum smear images: A deep learning approach," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 1131–1137, doi: [10.1109/CVPRW.2019.00147](https://doi.org/10.1109/CVPRW.2019.00147).
- [52] K. S. Mithra and W. R. S. Emmanuel, "Automated identification of mycobacterium Bacillus from sputum images for tuberculosis diagnosis," *Signal, Image Video Process.*, vol. 13, no. 8, pp. 1585–1592, Nov. 2019, doi: [10.1007/s11760-019-01509-1](https://doi.org/10.1007/s11760-019-01509-1).
- [53] M. M. Taye, "Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions," *Computation*, vol. 11, no. 3, p. 52, Mar. 2023, doi: [10.3390/computation11030052](https://doi.org/10.3390/computation11030052).
- [54] M. Hussain, "YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection," *Machines*, vol. 11, no. 7, p. 677, Jun. 2023, doi: [10.3390/machines11070677](https://doi.org/10.3390/machines11070677).
- [55] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 779–788, doi: [10.1109/CVPR.2016.91](https://doi.org/10.1109/CVPR.2016.91).



- [56] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. 30th IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 6517–6525, doi: [10.1109/CVPR.2017.690](https://doi.org/10.1109/CVPR.2017.690).
- [57] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.
- [58] A. Bochkovskiy, C.-Y. Wang, and H.-Y. Mark Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934*.
- [59] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Scaled-YOLOv4: Scaling cross stage partial network," Nov. 2020.
- [60] X. Long, K. Deng, G. Wang, Y. Zhang, Q. Dang, Y. Gao, H. Shen, J. Ren, S. Han, E. Ding, and S. Wen, "PP-YOLO: An effective and efficient implementation of object detector," 2020, *arXiv:2007.12099*.
- [61] C. Y. Wang, I. H. Yeh, and H. Y. M. Liao, "You only learn one representation: Unified network for multiple tasks," *J. Inf. Sci. Eng.*, vol. 39, no. 3, pp. 691–709, 2023, doi: [10.6688/JISE.202305\\_39\(3\).0015](https://doi.org/10.6688/JISE.202305_39(3).0015).
- [62] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: Exceeding YOLO series in 2021," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2021, pp. 1–7.
- [63] C. Li, L. Li, H. Jiang, K. Weng, Y. Geng, L. Li, Z. Ke, Q. Li, M. Cheng, W. Nie, Y. Li, B. Zhang, Y. Liang, L. Zhou, X. Xu, X. Chu, X. Wei, and X. Wei, "YOLOv6: A single-stage object detection framework for industrial applications," 2022, *arXiv:2209.02976*.
- [64] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2022, pp. 7464–7475.
- [65] G. Jocher, A. Chaurasia, and J. Qiu, "Ultralytics YOLO (version 8.0.0) [computer software]," Ultralytics, Los Angeles, CA, USA, Tech. Rep., 2023. [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [66] X. Ding, X. Zhang, N. Ma, J. Han, G. Ding, and J. Sun, "RepVGG: Making VGG-style ConvNets great again," 2021, *arXiv:2101.03697*.
- [67] X. Chu, L. Li, and B. Zhang, "Make RepVGG greater again: A quantization-aware approach," 2022, *arXiv:2212.01593*.
- [68] K. Weng, X. Chu, X. Xu, J. Huang, and X. Wei, "EfficientRep: An efficient repvgg-style ConvNets with hardware-aware neural network design," 2023, *arXiv:2302.00386*.
- [69] L. Mayats-Alpay, "Artificial intelligence for automatic detection and classification disease on the X-ray images," 2022, *arXiv:2211.08244*.
- [70] X. Chen, Y. Xue, Q. Hou, Y. Fu, and Y. Zhu, "RepVGG-YOLOv7: A modified YOLOv7 for fire smoke detection," *Fire*, vol. 6, no. 10, p. 383, Oct. 2023, doi: [10.3390/fire6100383](https://doi.org/10.3390/fire6100383).
- [71] H. Zhu, D. Zhu, X. Qin, and F. Guo, "Efficient and accurate beach litter detection method based on QSB-YOLO," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 7, pp. 311–322, 2023. [Online]. Available: [www.ijacsa.thesai.org](http://www.ijacsa.thesai.org)
- [72] I. S. Isa, M. S. A. Rosli, U. K. Yusof, M. I. F. Marzuki, and S. N. Sulaiman, "Optimizing the hyperparameter tuning of YOLOv5 for underwater detection," *IEEE Access*, vol. 10, pp. 52818–52831, 2022, doi: [10.1109/ACCESS.2022.3174583](https://doi.org/10.1109/ACCESS.2022.3174583).
- [73] S. Zhou, K. Cai, Y. Feng, X. Tang, H. Pang, J. He, and X. Shi, "An accurate detection model of *Takifugu rubripes* using an improved YOLO-V7 network," *J. Mar. Sci. Eng.*, vol. 11, no. 5, p. 1051, May 2023, doi: [10.3390/jmse11051051](https://doi.org/10.3390/jmse11051051).
- [74] J. Terven and D. Cordova-Esparza, "A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS," 2023, *arXiv:2304.00501*.



**SUCI AULIA** received the bachelor's and master's degrees in informatics engineering from the Telkom Institute of Technology (currently Telkom University), Bandung, Indonesia, in 2010 and 2012, respectively. She is currently pursuing the Ph.D. degree with the School of Electrical Engineering and Informatics, Bandung Institute of Technology, Bandung. Since 2010, she has been with Telkom University, as a Lecturer with the School of Applied Science. Her research interests include signal and image processing, machine learning, artificial intelligence, and biomedical engineering.



**ANDRIYAN BAYU SUKSMONO** (Senior Member, IEEE) received the B.Sc. degree in physics and the M.T. degree in electrical engineering from Bandung Institute of Technology (ITB), Bandung, Indonesia, in 1990 and 1996, respectively, and the Ph.D. degree in engineering from The University of Tokyo, Japan, in 2022. From 1996 to 2005, he was with ITB as an Instructor; from 2005 to 2009, he was an Associate Professor; and since 2009, he has been a Professor. He joined the School of Electrical Engineering and Informatics, ITB. His primary research interests include signal processing, imaging, and quantum computing. He is currently a Professional Member of ACM. He was a recipient of several international research funds, grants, scholarships, and fellowships from RCAST-Tokyo University, Monbukagakusho, JSPS, AIGRP, Asahi Glass, and the Hitachi Foundation. In 2004, he received the Best Paper Award in Theoretical Development from APNNA, for his work on spatiotemporal ultra-wideband neuro-beamforming. He was also a recipient of the Outstanding Faculty Award of ITB, in June 2007; the Outstanding Faculty Award of the Republic of Indonesia, in August 2007; the ITB Innovation Award, in 2017; the ITB Research Award, in 2020; the 114 Indonesia Innovation Award, in 2022; and the 115 Indonesia Innovation Award, in 2023. His work on quantum algorithms for finding the Hadamard matrix was included in Nature's Scientific Report 100 in Physics 2020 (eighth place) and 2023 (second place), Editor's Choice (2023) and is listed in the Quantum Algorithm Zoo.



**TATI RAJAB MENGKO** received the combined Ir. and B.S. degrees in electrical engineering from Bandung Institute of Technology, Bandung, Indonesia, in 1977, and the Dr.-Eng. degree from ENSERG-INPG, Grenoble, France, in 1985. In 1978, she joined the School of Electrical Engineering and Informatics, Bandung Institute of Technology, and a Professor of image processing with the School of Electrical Engineering and Informatics, in 2006. She was the Head of the Biomedical Engineering Research Division, in 2022. In 2015, she was granted the Innovation Award from ITB because of her contribution in developing a noninvasive vascular analyzer device. Her research interests include image processing and instrumentation in biomedical engineering. She has chaired numerous conferences, including the International Conference on Instrumentation, Communication, Information Technology, and Biomedical Engineering (ICICI-BME).



**BACHTI ALISJAHBANA** received the Medical degree from the Faculty of Medicine, Universitas Indonesia, Jakarta, in 1989, and the Ph.D. degree from the University Medical Center Nijmegen, in 2007. He has conducted his fieldwork in the central highlands of Papua, from 1990 to 1993. He started to actively contribute to a TB research project with a Dutch collaborator. Simultaneously, he started collaborating with Naval Medical Research Unit II to study Dengue. These two studies continue to provide opportunities for the follow-up studies of many subsequent Ph.D. students. Subsequently, he continued to lead many young investigators in projects related to TB, HIV, Dengue, and other general infectious diseases with the Medical Faculty of UNPAD. Practical research in TB has provided him with the opportunity to be the Chairperson of the TB Operational Research Group, which promotes the implementation of research to improve the TB program.

...