

TBNet: learning from scratch and limited training data, a CNN based tuberculosis bacilli detection

Ali Suryaperdana Agoes¹, Winarno²

¹Informatics Engineering, STMIK AMIK Bandung, Bandung, Indonesia

²Department of Physics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia

Article Info

Article history:

Received Nov 18, 2022

Revised May 25, 2023

Accepted Aug 30, 2023

Keywords:

Bacilli

Convolutional neural network

Deep learning

Object detection

Tuberculosis

ABSTRACT

Tuberculosis (TB) is an infectious disease caused by the micro-bacteria. Several studies that have been conducted previously aimed to reduce the burden of observing tuberculosis bacilli using the digital image processing method. In this study, we proposed a newly developed convolutional neural network (CNN) based deep learning model to detect tuberculosis bacilli in sputum smear images. Recent advances in deep learning apply large scale image dataset to achieve convergent weight model. However, medical image dataset commonly available in relatively small quantity. In contrary with common deep learning approach, our model is capable to learn from our small dataset which consist of highly diverse hue and contrast of sputum smear images. Furthermore, its performance is proven to be reliable to detect sputum smear image content, which are TB bacillus and debris.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Winarno

Department of Physics, Faculty of Science and Technology, Universitas Airlangga

Jl. Dr. Ir. H. Soekarno, Mulyorejo, Kec. Mulyorejo, Surabaya, Jawa Timur 60115, Indonesia

Email: winarno@fst.unair.ac.id

1. INTRODUCTION

Tuberculosis (TB) is caused by *bacillus mycobacterium tuberculosis* [1]. Although it is a curable disease, the death rate of TB is up to 13.57% in 2018 [1]. Moreover, one-third of the estimated incident cases (3 million) remain unknown to the health system due to an underreporting of detected cases and underdiagnosis [1]. Most of the underreporting and underdiagnosis cases happened in India (25%), Nigeria (12%), Indonesia (10%), and Philippines (8%) [1]. Therefore, improvements in the accessibility of TB diagnosis and treatment are urgently required in these countries.

There are several established methods to diagnose TB i.e., microscopic analysis, polymerase chain reaction (PCR), and electronic nose system [2]. Among other methods, sputum smear microscopy is the most used method, especially in developing countries, because it is simple, low cost, and easy to maintenance [3], [4]. Prepared and stained sputum specimen is analyzed manually under a microscope. TB bacilli is a gram-positive bacillus, on gram-stained smears specimen of a patient with TB, bacilli can be detected [5]. TB bacilli counted manually depending on technician observation. Generally, laboratory technician needs about 40 min – 3 h to examine 100 fields of view on each prepared specimen [5], [6]. The number of detected TB bacilli used to determine the diagnosis [2]. A prompt and precise diagnosis method of TB is essential to control the individual treatment and prevent further contagion [7]. The manual microscopic method is tiring and time-consuming, gives varies sensitivity and high false-negative rate detection [8], [9].

The raw image processing of conventional algorithms without learning is limited. Deep learning has substantially better performance than conventional methods due to its automatic fast learning features [10], [11]. Deep learning perform better accuracy in image classification, semantic segmentation, object detection, and

simultaneous localization and mapping (SLAM) [12]. Convolutional neural network (CNN) is the most popular deep learning based method due to its remarkable improvement in prediction performance using big data and plentiful computing resources [11]. CNN has pushed the boundaries of what was possible i.e., automatic detection of TB bacilli by CNNs successfully identified both single and touching bacilli [13]. Most of the recently advanced object detection lies on fine-tuned of pre-trained CNNs on ImageNet [14]. This process can generate the final model quickly and requires way fewer instance-level annotated training data than the classification task [15]. Yet, several limitations are undeniable, i.e., limited design space on network structures, learning/optimization bias, and domain mismatch [15]. Therefore, a deeply supervised object detector (DSOD) framework is developed to overcome these problems. DSOD framework can learn object detectors from scratch [16]. As the author knowledge, there are no research about TB bacilli detection based on the DSOD framework reported. These facts show a great research opportunity. Therefore, in this work, we proposed an enhanced architecture of the DSOD framework for TB bacillus detection based on the deep learning model.

Process for collecting TB sputum smear for deep learning dataset is not easy. Domain specific data frequently limits the preparation of dataset. The usage right for medical resource often being exclusively owned by single institution. While TB patient number distribution are varied from area to area [1]. Making it impossible to assemble one huge dataset. On the contrary, deep learning methods often needs images in the order of thousand to train a model. Another point to mark is the difficulty for preparing image annotation for the ground truth. Process involved in ground truth preparation for image classification, object detection, and segmentation are increased in complexity. These jobs are often done by human. Therefore, the more images that are worked on, the more hindrances will be encountered.

In this paper, we proposed a newly developed architecture with deep supervision for TB bacillus detection trained on a limited resource. First, we prepare sputum smear dataset. It consists of normal and overstained microscopical images. Then, we present a deep learning model with deep supervision designed specifically for TB detection called TBnet. We also provided a performance comparison between the TBnet with SSD, MobileNet SSD, Peelenet and DSOD model. The qualitative and quantitative studies showed that the new model has better performance in detecting TB bacilli.

2. RELATED WORKS

Much research regarding TB bacilli detection have been reported. Rulaningtyas *et al.* [17] developed an automatic classification of TB bacilli using a neural network. First, feature extraction was done to find the morphology (shape) of TB bacilli. The features arranged into a vector and submitted to the neural network. Then, bacilli classified by the backpropagation method. Although this research showed good TB bacilli classification results, this method used handcrafted feature vectors to discriminate bacilli pixels from non-bacilli pixels based on the morphology shape. Therefore, the performance heavily depends on bacilli features [13]. Khutlang *et al.* [4] proposed an automatic detection of TB Bacilli based on two one-class classifiers. The first stage classification was done used a one-class pixel classifier. The object output filtered based on the object area. Features (Fourier, moment, eccentricity) were extracted from the remaining objects. Then, the second one-class object classification was done in different feature sets. The mixture of Gaussians performed the best result in first stage classification, but the accuracy of object outline detection is low, resulting in a low percentage of correctly classified pixels (75.74%). Ghosh *et al.* [18] proposed an automatic TB diagnosis by hybrid (crisp and fuzzy data representation) approach. Sputum image was pre-processed before segmented using a gradient-based region growing technique to find the accurate contour of TB bacilli. Then, the features (shape, color, and granularity) of TB bacilli extracted to generate individual fuzzy classification. Finally, the individual classification was combined to strengthen the diagnosis. The result showed quite high sensitivity (93.9%) and specificity (88.2%). Unfortunately, overlapped bacilli were failed to identify.

Recently, deep learning based research taking active part in the field of medical image analysis including image classification, object detection and image segmentation. Computer aided system for TB diagnosis is also influenced by this advance. Quinn *et al.* [19] presented a work for detecting TB bacilli in the microscope field of view that is captured with a mobile phone. They divide image into many small patches fed into CNN based network. The model classified each of boxes. This process is known for taking too much computation resources because the system needs to compute all bounding box's representation inside the image. They used prior knowledges training before the TB bacilli training for classification jobs done [13] used 2 folds segmentation to detect TB bacilli. In the first step, foreground and background were separated using Otsu's method to extract pixel with bacilli's color tendency. Next step, a CNN based segmentation method classified pixels inside the patches containing the objects. Using this approach, an end-to-end model training could not be achieved. Although their model setup did not need a prior training. Their result of TB bacilli segmentation accuracies relies on color feature classification during the first step. In more recent paper, Trilaksana *et al.* [20]

proposed the use of faster R-CNN to tackle the patch division steps before the classification process. Their approach still depends on the prior knowledge transfer learning to produce the result.

Automatic CNN based TB bacilli detection relies on the database. Costa *et al.* [21] introduced a two-parts database of sputum smear images i.e., i) autofocus database and ii) segmentation, classification database. Segmentation and classification were done by annotated the objects with geometric shapes: a circle is for true bacillus, a rectangle for agglomerated bacillus, and a polygon for doubtful bacillus. However, the ‘agglomerated bacillus’ and ‘doubtful bacillus’ segmentation and classification are still not satisfying. Shah *et al.* [22] introduced the Ziehl-Neelsen sputum smear microscopy image database which consists of seven categories datasets. The database can be used for the development of an efficient algorithm for autofocusing, autostitching, and automatic bacilli segmentation and grading. Even so, the number of images is limited for automatic TB detection. Moreover, the available database annotated images on a clean smear which differs from common sputum smear slides in Indonesia. The nature of the image background is one of the important factors that affect TB detection performance. Therefore, Trilaksana *et al.* [20] introduced a sputum smear images database with divers smearing background: clear and definitive bacilli to the highly cluttered and stained bluish background. There are three kinds of annotated images i.e., TB bacteria, non TB bacteria, and stainresidues. However, further development is needed to add more data. Commonly, the sputum smear images from laboratory technicians are overstained due to the dye’s quality and the process of specimen preparation. In this work, we introduce our database which has been adjusted to this medical fact.

3. TBNET ARCHITECTURE

In this section, we introduce our CNN based model to detect tuberculosis bacilli. Our model comprises of two parts. First part is the feature extraction block, then it’s followed by prediction block. Our first network part is slight adjustment from lightweight color depth semantic segmentation (LICODS) [23]. The main difference is our model network consist only one branch to incorporate single image coming from input side. The first layer in comb block involving two convolutional layer and one max pooling layer. Here, we use 64 layers of depth for each convolutional filter compared 14 in the original network. Denser filters contribute more information to relay in further progress in the network. Further, we handle the propagated information contradictorily with [23]. Right after the pooling layer in the transition block, we divide the main branch into 3 separate sub-branches in contrast with original LICODS [23]. First sub-branch conveys the information right after the first expansion block. Second sub-branch squeezes the features maps into 64 depth layers. Third sub-branch is the expansion block part. In the end of first expansion block the second and third sub-branches are concatenated to collect multi feature map results. Expansion block two followed expansion block one with one a smaller number of depth wise convolutional layer. This fashion is repeated until three times. The first sub-branch then merged with three expansion block result. Table 1 shows our feature extraction block detail. feature extraction block consists of comb block, transition block, shortcut, and four expansion blocks.

Table 1. TBNet feature extraction block

Layer	Layer	Output size (input 3×240×306)
Comb block	2×2 conv, stride 2, pad 0	64×120×153
	3×3 conv, stride 2, pad 1	64×120×153
	2×2 max pool, stride 2, pad 0	3×120×153
	Concatenate	131×120×153
Transition block	1×1 conv, stride 1, pad 0	64×120×153
	1×1 conv, stride 1, pad 0	128×120×153
	2×2 max pool, stride 2, pad 0	128×60×77
Shortcut	1×1 conv, stride 1, pad 0	64×60×77
Expansion block 1	1×1 conv, stride 1, pad 0	32×60×77
	4× (conv depth wise)	32×60×77
Expansion block 2	1×1 conv, stride 1, pad 0	32×60×77
	3× (conv depth wise)	32×60×77
Expansion block 3	1×1 conv, stride 1, pad 0	32×60×77
	2× (conv depth wise)	32×60×77
Expansion block 4	1×1 conv, stride 1, pad 0	32×60×77
	1× (conv depth wise)	32×60×77
Conv	1×1 conv, stride 1, pad 0	128×60×77

Next block in our network is the prediction block. Here, we employ similar approach for DSOD Net [15] to serve multiple scale corresponding with different size feature maps and combined with dense

structure prediction. We carefully follow the dense prediction structure presented in [15]. Our network utilizes larger resolution feature map prediction structure compared the original dense structure. The largest feature maps resolution is 60×77 . We increased the feature maps resolution to handle small object in digital sputum images. The prediction structure positioned after the prediction block except for the largest feature maps which connected to convolution layer after expansion block concatenation. In each prediction layer comprise of multi-scale information forwarded from different stages of layers.

4. DATASETS

TB bacilli expectorated in human's sputum is visible in individually or clustered. The mycobacterium tuberculosis within stained sputum smear under a microscope seen as red-purplish colored blob with the shape of rectangular over the blue background, shown in the Figures 1(a) and (b) shows sputum smear images consisting of mycobacterial cell's wall comprises of a substance composed of mycolic acid. These is a β -hydroxy carboxylic acids with chain lengths up to 90 carbon atoms. The property of acid fastness is related to the carbon chain length of mycolic acid found in any species [24]. The mycolic acid raise a barrier to dye entering, this problem is usually overcome by adding a lipophilic agent to a concentrated aqueous solution and partly by heating [25]. We use TB positive stained sputum smear as its slides. 20 different images are acquired from spatially shifted view of fields. We prepare a microscope to observe sputum smear on objective glass. Then, we turn on the computer connected to the microscope. We set computer's programs to open an observation program that is used to capture sputum microscopic images. After we turn on the computer, we open the standard cell application that is used to capture the sputum image, we set the cellsens application to match the magnification of the microscope. The magnification that we use on the microscope is 1000 magnifications. After everything is ready, we place the sputum smear on the preparation table on a microscope, after that we just need to observe and shift the preparation to move to the next layer. We record the digital form of microscope's field of view under the supervision of two physician.

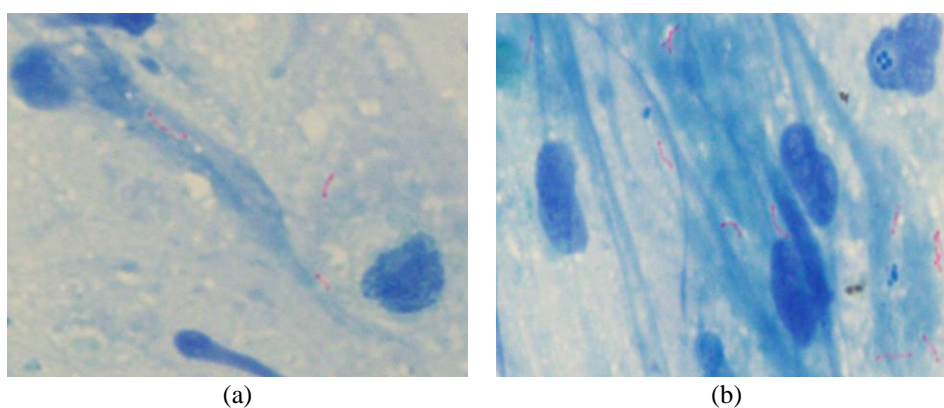


Figure 1. Sputum smear images after the dye has fixed and examined under the microscope: (a) sputum smear image with light blue background and (b) sputum smear images with darker blue background

Our tuberculosis database (TBDB) contains of 350 images. Our data is a portray of a routine smear examination by expertise at medical facility. The process for data preparation is contributing for the dye's smear evenness quality. One example is during the patient's sputum spreading on the middle 1/3 slide. The pressure applied into the applicator are susceptible to the intensity change. Another example is the quality of dye. Therefore, images in the TBDB are composed of normal-stained and over-stained sputum smears. We prepare the ground truth for TB bacilli detection by human input. We use two labels to annotate images, which are: TB bacilli and debris. The annotation mark for overlapped bacillus is carried out one bacillus after another, and there is no distinguishing mark for any of them. We mark area containing one of three labels inside the image using labelImg program. Figures 2(a) and (b) shows the annotated sputum smear images. Data validation is established during annotation process by an expert. Afterward, the marking result is checked by the researcher which has the expertise on TB sputum smear examination. The two-step validation is to ensure the ground truth correctness. We obtain the total of 3,102 bacillus with shape and color intensity variation. Its intensity ranged from faint to strong red-purplish color. The Bacillus shapes are also varied, from single to in group bacillus.

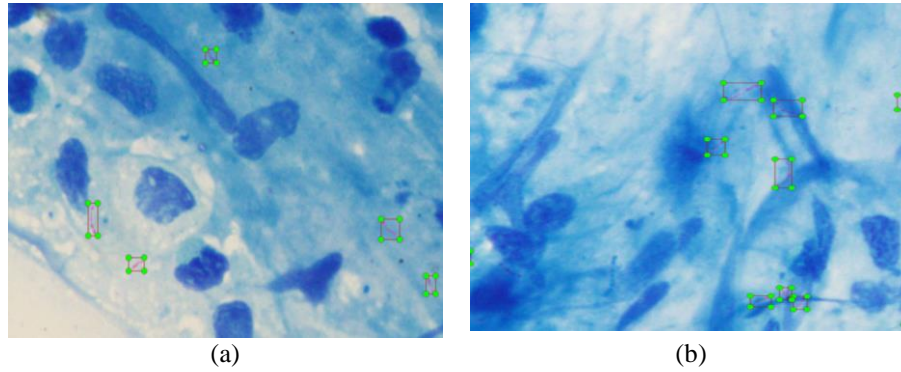


Figure 2. TB bacillus after annotation process on sputum smear images: (a) TB bacillus appear with its original dye and clear shape in front of faint background and (b) TB bacillus appear with slightly red-purplish

5. RESULTS

In this section, we present our finding and result for our proposed TBNNet for detection TB bacilli in sputum smear images. We analyzed our network using metric stated in [25]. In (1) is an expression for the Precision value. Precision value is an explanation toward model capability to recover only the precise object completed with its location across all detection results. True positive (TP) define object predicted correctly with the intersection over union (IoU) value above some points. False positive (FP) define object predicted in false manner. In (2) is an expression to find the recall value. Recall value is the model capability to recover all the precise object in observation. True positive (TP) define object predicted correctly with the IoU value above some threshold points. In (3) is used for calculating IoU. Where B_p is a prediction bounding box and B_{gt} is a ground truth bounding box. Subsequently, precision \times recall curve is composed of each predicted class. Lastly, the average precision (AP) number is calculated with (5). Table 2 shows quantitative detection result over several models for comparison.

$$Precision = \frac{TP}{TP+FP} = \frac{TP}{\text{all detections}} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} = \frac{TP}{\text{all ground truths}} \quad (2)$$

$$IOU = \frac{\text{area}(B_p \cap B_{gt})}{\text{area}(B_p \cup B_{gt})} \quad (3)$$

$$\sum_{r=0}^1 (r_{n+1} - r_n) p_{interp}(r_{n+1}) \quad (4)$$

$$p_{interp}(r_{n+1}) = p(\tilde{r}) \quad (5)$$

Table 2. Sputum smear object detection result, shown result is in (%)

Model	mAP	TB bacilli	Debris
TBNNet	76.42	75.87	76.97
DSOD	73.49	73.70	73.28
PeleeNet	60.34	52.21	68.46
MobileNetSSD	57.48	53.65	61.31
SSD	67.22	46.68	56.95

We evaluated and compared our TBNNet quantitative result toward several leading object detection model available, namely SSD, MobileNetSSD, and PeleeNet. We ran our test images in a single pass for each model training result. Our model performs comparably well throughout our tests. The mAP number give 3% significant different toward our based model DSOD, and more than 10% toward traditional approach model for training such as SSD and PeleeNet. These number continues at each class average precision result. Our approach to increase the feature maps resolution twice as big as original DSOD in final block perform outstanding. Figures 3(a) to (j) (in appendix) shows detail performance for each object detection across our test models.

Figures 4(a) to (d) provides qualitative comprehension comparison toward our TBNNet performance. Here, we use a test image taken from outside our datasets. Similar result with our previous test is shown, TBNNet perform well across our 1000 test images. Figure 4(a) shows TBNNet managed separate two debris

inside the images quite well. This performance continues for TB Bacilli detection. Two TB Bacillus in close perimeter are detected individually.

The other models in our comparison failed to detect TB Bacilli individually, except in Figure 4(d) SSD detection model. This behavior is predicted, for the reason that our model and SSD share the common prediction block basis. Except that TBNets backbone comprises more efficient feature maps filter parameter, also a dense prediction structure inherited from DSOD being adopted. MobileNetSSD perform poorly to detect objects in across our dataset test images, therefore we exclude the detection results in the Figure 4. MobileNetSSD failed to produce meaningful boundaries around our designated object.

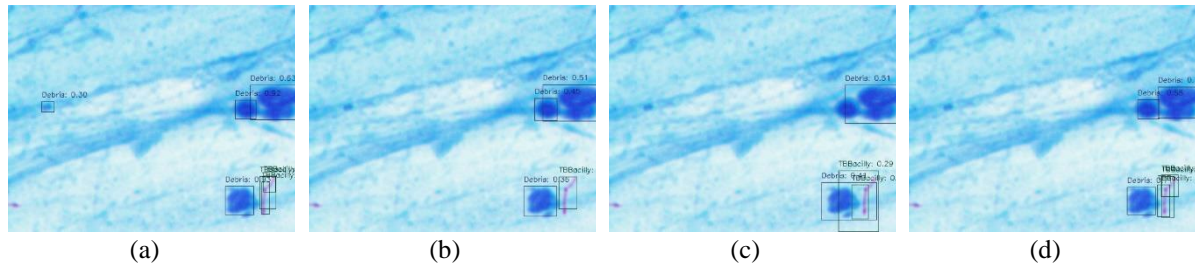


Figure 4. Images from TB bacilli and debris detection results: (a) TBNets, (b) DSOD, (c) PeleeNet, and (d) SSD, respectively

6. CONCLUSION

In this paper we deliver our strategy to develop a CNN based model to detect object, especially TB bacillus and debris in a sputum smear image. Our strategies to develop detection model comprises a CNN backbone that accommodate a fast-training method without any prior knowledge toward dataset being present. Also, a comprehend understanding to construct a model architecture with limited resources come to our attention. Our result show that our model fit in our initial intention to minimally train a model. Furthermore, limited number of training image data also did not negatively impact our model performance. This training fashion support computer vision approach especially convolutional based neural network in a way efficient detection method for the limited image dataset and training session.

ACKNOWLEDGEMENTS

Authors would like to thank Faculty of Science and Technology, Universitas Airlangga for financial support through Penelitian Unggulan Fakultas (PUF) scheme.

APPENDIX

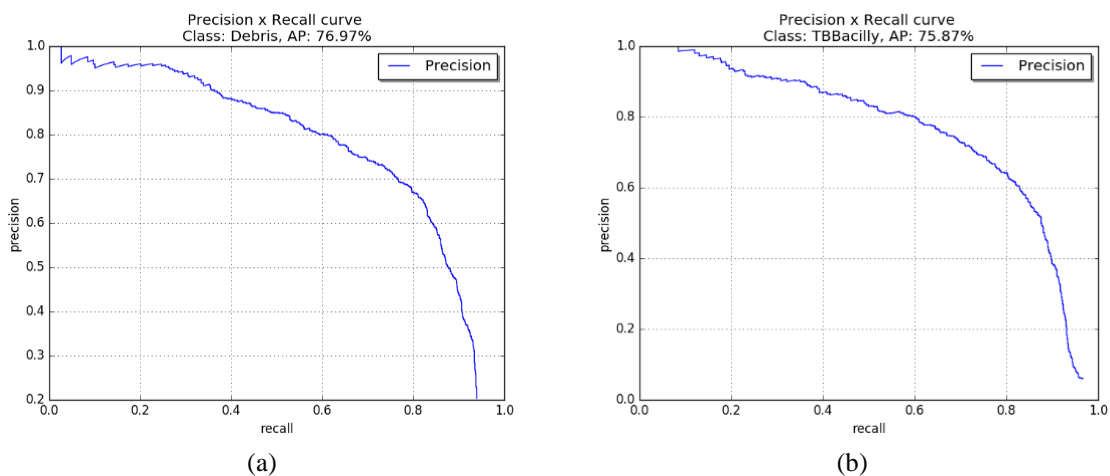


Figure 3. Left side figure shows precision x recall curve performance for Debris detection. Right side figure shows precision x recall curve performance for TB Bacilli detection; (a) and (b) TBNets

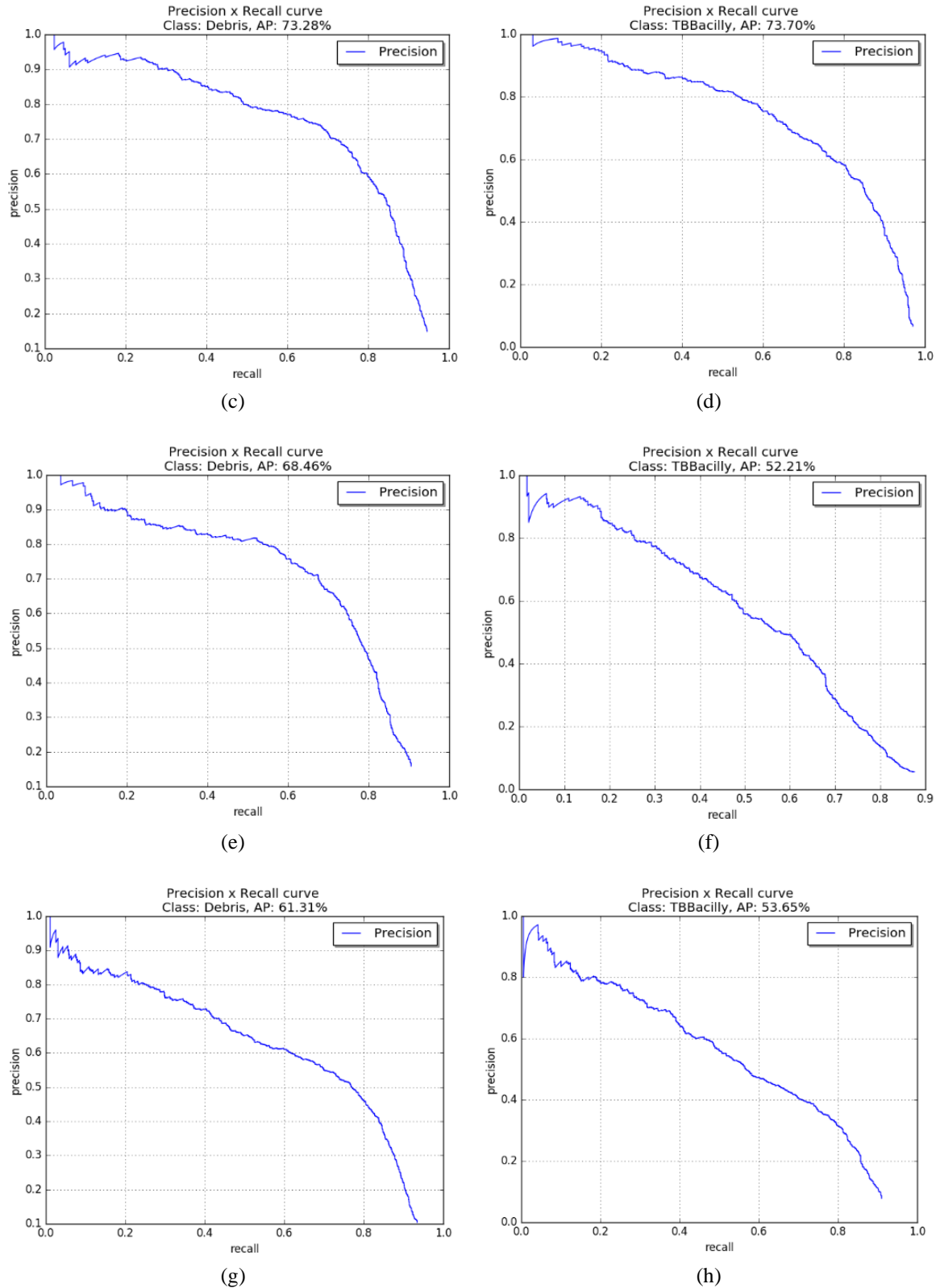


Figure 3. Left side figure shows precision x recall curve performance for Debris detection. Right side figure shows precision x recall curve performance for TB Bacilli detection; (c) and (d) DSOD, (e) and (f) PeleeNet, (g) and (h) MobileNetSSD (continue)

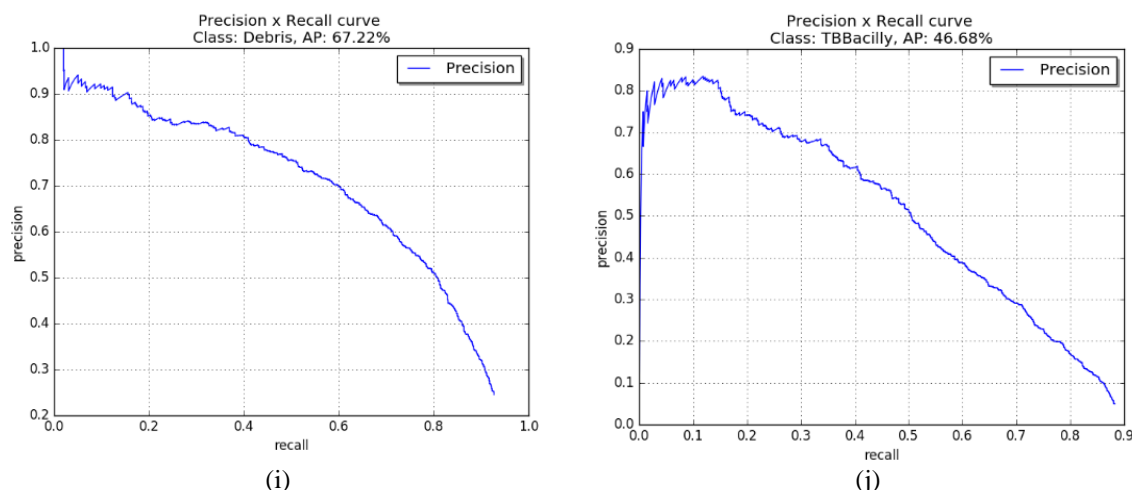


Figure 3. Left side figure shows precision x recall curve performance for Debris detection. Right side figure shows precision x recall curve performance for TB Bacilli detection; (i) and (j) SSD (continue)




REFERENCES

- [1] WHO, "WHO | Global tuberculosis report 2019," 2020, Available online: <https://www.who.int/publications/i/item/9789241565714>, 2020. Date accessed: 10 August 2022.
- [2] L. N. Sahenda, M. H. Pumomo, I. K. E. Purnama, and I. D. G. H. Wisana, "Comparison of Tuberculosis Bacteria Classification from Digital Image of Sputum Smears," in *2018 International Conference on Computer Engineering, Network and Intelligent Multimedia, CENIM 2018 - Proceeding*, Nov. 2018, pp. 20–24. doi: 10.1109/CENIM.2018.8711386.
- [3] R. O. Panicker, B. Soman, G. Saini, and J. Rajan, "A Review of Automatic Methods Based on Image Processing Techniques for Tuberculosis Detection from Microscopic Sputum Smear Images," *Journal of Medical Systems*, vol. 40, no. 1, pp. 1–13, Jan. 2016, doi: 10.1007/s10916-015-0388-y.
- [4] R. Khutlang, S. Krishnan, A. Whitelaw, and T. S. Douglas, "Detection of tuberculosis in sputum smear images using two one-class classifiers," in *Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009*, Jun. 2009, vol. 40, no. 1, pp. 1007–1010. doi: 10.1109/ISBI.2009.5193225.
- [5] A. Whitelaw, "Microbiological testing for Mycobacterium tuberculosis," in *Tuberculosis*, Elsevier, 2009, pp. 169–178. doi: 10.1016/B978-1-4160-3988-4.00018-4.
- [6] M. Sotaquirá, L. Rueda, and R. Narvaez, "Detection and quantification of bacilli and clusters present in sputum smear samples: A novel algorithm for pulmonary tuberculosis diagnosis," in *Proceedings - 2009 International Conference on Digital Image Processing, ICDIP 2009*, Mar. 2009, pp. 117–121. doi: 10.1109/ICDIP.2009.59.
- [7] C. T. Sreeramareddy, K. V. Panduru, J. Menten, and J. Van den Ende, "Time delays in diagnosis of pulmonary tuberculosis: A systematic review of literature," *BMC Infectious Diseases*, vol. 9, no. 1, p. 91, Dec. 2009, doi: 10.1186/1471-2334-9-91.
- [8] A. Kimura, M. G. F. Costa, C. F. F. Costa Filho, L. B. M. Fujimoto, and J. Salem, "Evaluation of autofocus functions of conventional sputum smear microscopy for tuberculosis," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*, Aug. 2010, vol. 9, no. 1, pp. 3041–3044. doi: 10.1109/IEMBS.2010.5626143.
- [9] R. F. Grossman, P. R. Hsueh, S. H. Gillespie, and F. Blasi, "Community-acquired pneumonia and tuberculosis: DIFFERENTIAL diagnosis and the use of fluoroquinolones," *International Journal of Infectious Diseases*, vol. 18, no. 1, pp. 14–21, Jan. 2014, doi: 10.1016/j.ijid.2013.09.013.
- [10] J. M. R. S. Tavares and P. R. Fernandes, "Correction: New developments on computational methods and imaging in biomechanics and biomedical engineering (Lecture Notes in Computational Vision and Biomechanics, (2019) 33, 10.1007/978-3-030-23073-9)," in *Lecture Notes in Computational Vision and Biomechanics*, vol. 33, 2019, p. C1. doi: 10.1007/978-3-030-23073-9_11.
- [11] S. Rabhi, J. Jakubowicz, and M. H. Metzger, "Deep Learning versus Conventional Machine Learning for Detection of Healthcare-Associated Infections in French Clinical Narratives," *Methods of Information in Medicine*, vol. 58, no. 1, pp. 31–41, Jun. 2019, doi: 10.1055/s-0039-1677692.
- [12] N. O'Mahony *et al.*, "Deep Learning vs. Traditional Computer Vision," in *Advances in Intelligent Systems and Computing*, vol. 943, 2020, pp. 128–144. doi: 10.1007/978-3-030-17795-9_10.
- [13] 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 and Biomedical Engineering*, vol. 38, no. 3, pp. 691–699, 2018, doi: 10.1016/j.bbe.2018.05.007.
- [14] H. C. Shin *et al.*, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, May 2016, doi: 10.1109/TMI.2016.2528162.
- [15] Z. Shen, Z. Liu, J. Li, Y. G. Jiang, Y. Chen, and X. Xue, "DSOD: Learning Deeply Supervised Object Detectors from Scratch," in *Proceedings of the IEEE International Conference on Computer Vision*, Oct. 2017, vol. 2017-October, no. 5, pp. 1937–1945. doi: 10.1109/ICCV.2017.212.
- [16] Z. Shen, Z. Liu, J. Li, Y. G. Jiang, Y. Chen, and X. Xue, "Object Detection from Scratch with Deep Supervision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 398–412, Feb. 2020, doi: 10.1109/TPAMI.2019.2922181.
- [17] R. Rulaningtyas, A. B. Suksmono, and T. L. R. Mengko, "Automatic classification of tuberculosis bacteria using neural network," in *Proceedings of the 2011 International Conference on Electrical Engineering and Informatics, ICEEI 2011*, Jul. 2011, pp. 1–4. doi: 10.1109/ICEEI.2011.6021502.




- [18] P. Ghosh, D. Bhattacharjee, and M. Nasipuri, "A hybrid approach to diagnosis of tuberculosis from sputum," in *International Conference on Electrical, Electronics, and Optimization Techniques, ICEEOT 2016*, Mar. 2016, pp. 771–776. doi: 10.1109/ICEEOT.2016.7754790.
- [19] J. A. Quinn, R. Nakasi, P. K. B. Mugagga, P. Byanyima, W. Lubega, and A. Andama, "Deep Convolutional Neural Networks for Microscopy-Based Point of Care Diagnostics," 2016. doi: 10.48550/arXiv.1608.02989.
- [20] H. Trilaksana, G. N. Dwimudyari, A. S. Agoes, and D. B. Widhyatmoko, "Sputum smear images database: A resource for deep learning study based to detect Bacilli for TB diagnose," in *AIP Conference Proceedings*, 2020, vol. 2314, p. 40013. doi: 10.1063/5.0036388.
- [21] 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 *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, Aug. 2014, pp. 2841–2844. doi: 10.1109/EMBC.2014.6944215.
- [22] M. I. Shah *et al.*, "Ziehl–Neelsen sputum smear microscopy image database: a resource to facilitate automated bacilli detection for tuberculosis diagnosis," *Journal of Medical Imaging*, vol. 4, no. 2, p. 27503, Jun. 2017, doi: 10.1117/1.jmi.4.2.027503.
- [23] A. S. Agoes, Z. Hu, and N. Matsunaga, "Licods: A cnn based, lightweight rgb-d semantic segmentation for outdoor scenes," *International Journal of Innovative Computing, Information and Control*, vol. 15, no. 5, pp. 1935–1946, 2019, doi: 10.24507/ijicic.15.05.1935.
- [24] M. Jadhav, P. P. Janrao, and K. V. Kale, "Acid Fast Stain Sputum Smear Images Data Set for Diagnosis of Tuberculosis," 2013.
- [25] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (VOC) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, Jun. 2010, doi: 10.1007/s11263-009-0275-4.

BIOGRAPHIES OF AUTHORS



Ali Suryaperdana Agoes    received his M.T. degree from Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia in 2014. He gets his Doctoral degree from Kumamoto University, Japan in 2018. Since 2018 until 2020, he was with HMS-Global, Japan. At present time, he is a lecturer in STMIK AMIK Bandung. He is also the head of Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) at STMIK AMIK Bandung. His research interest includes deep learning model optimization for computer vision implementation, 3D surfaces reconstruction and Metaverse's visual layer. His email address is ali@stmik-amikbandung.ac.id.



Winarno    is a lecturer in the Physics Department, Faculty of Science and Technology, Universitas Airlangga. He has graduated in Physics from Universitas Airlangga and received his Master of Engineering from Institut Teknologi Sepuluh Nopember Surabaya. His research interests are computer vision, machine learning, electronics, and interfacing. He can be contacted at email: winarno@fst.unair.ac.id.