# Improving Smear-Negative Tuberculosis Detection Using Data Augmentation and Faster R-CNN

Nur Azizah[1*], Po Abas Sunarya[2], Untung Rahardja[3], Achmad Benny Mutiara[4], Prihandoko[5], Charlotte Pasha[5]
[1, 2]Faculty of Economics and Business, University of Raharja, Indonesia
[3]Faculty of Science and Technology, Universiti Teknologi Malaysia, Malaysia
[4, 5]Faculty of Science and Technology, University of Gunadarma, Indonesia
[6]Eduaward Incorporation, United Kingdom
[1]nur.azizah@raharja.info, [2]abas@raharja.info, [3]urahardja@gmail.com, [4]amutiara@staff.gunadarma.ac.id, [5]prihandoko@staff.gunadarma.ac.id, [6]charlotte5@eduaward.co.uk
***Corresponding Author**

**Article Info**

***Article history:***

Submission October 2, 2025
Revised November 13, 2025
Accepted January 5, 2026
Published January 21, 2026

***Keywords:***

Data Augmentation
SNPT
Faster R-CNN
ResNet Architecture
Chest X-ray Imaging

**ABSTRACT**

Smear-Negative Pulmonary Tuberculosis (SNPT) remains a **major diagnostic challenge** due to the low bacterial load that frequently causes false negative results in sputum microscopy. Chest X-ray imaging is commonly used as a complementary diagnostic tool however its interpretation relies heavily on expert radiologists and is prone to subjectivity. Recent developments in deep learning particularly object detection models provide promising opportunities to improve diagnostic accuracy. **This study aims** to develop and evaluate a deep learning based approach for SNPT detection in chest X-ray images using a Faster R-CNN model with a ResNet architecture. The proposed **method** applies data augmentation techniques including flipping rotation scaling and random brightness adjustment to enhance training data diversity and reduce overfitting. The model was implemented using PyTorch and evaluated using accuracy precision recall and F1 score. Experimental **results** indicate that data augmentation substantially improves performance achieving 76.60% accuracy and 68.57% F1 score compared to 53.06% accuracy and 51.06% F1 score without augmentation. Improved recall reflects higher sensitivity in detecting SNPT cases. These **findings indicate** that data augmentation enhances the robustness and generalization of Faster R-CNN models for SNPT detection and supports the potential of AI assisted diagnostic systems in tuberculosis screening programs.



DOI: https://doi.org/10.34306/ijcitsm.v6i1.233


## 1. INTRODUCTION

Tuberculosis (TB) is an infectious disease caused by Mycobacterium TB and remains one of the leading causes of death worldwide. One of the more challenging forms of TB to diagnose is SNPT, where the causative bacteria cannot be detected through conventional microscopic methods such as Ziehl-Neelsen staining. Detecting SNPT often requires more advanced techniques, including radiographic imaging and deep learning-based methods [1]. Faster-RCNN with a ResNet architecture is a popular object detection model that has shown promising results in detecting various diseases from medical images. However, a major challenge in training this model is the limited variation in training data, which can lead to overfitting and reduce the model's ability to detect objects in new images [2].

Computer vision, powered by deep learning, has revolutionized image analysis. Faster Region-Based CNN (Faster R-CNN), coupled with the robust ResNet architecture, have become state-of-the-art for object detection tasks. However, the performance of these models is often limited by the availability of large, high-quality datasets, especially in medical imaging [3]. TB remains a significant global health concern. SNPT, characterized by a low bacterial load, poses a diagnostic challenge. Traditional methods, like smear microscopy, often miss these cases [4]. In this study, the research applied data augmentation as a solution to increase the variation in training data. This augmentation technique is expected to improve the accuracy of detecting SNPT in chest X-ray images [5].

Unlike prior studies that primarily employ single or generic augmentation strategies or focus on classification-based CNN, this study distinguishes itself by systematically integrating multiple spatial and intensity-based augmentation techniques namely flipping, rotation, scaling, and random brightness within a region-based object detection framework [6]. Furthermore, existing hybrid models often rely on generative approaches such as GAN-based augmentation or ensemble classifiers, which introduce higher computational complexity and training instability. In contrast, the proposed approach emphasizes a lightweight yet effective augmentation pipeline tailored specifically to Faster R-CNN with a ResNet backbone, enabling improved detection sensitivity for subtle SNPT patterns while maintaining architectural simplicity and clinical feasibility.

Figure 1. SDGs 3 and SDGs 9

In a broader context, this research aligns with the United Nations Sustainable Development Goals SDGs), particularly SDGs 3 (Good Health and Well-Being) [7], which aims to end the TB epidemic through improved prevention and early diagnosis. By leveraging artificial intelligence and deep learning–based medical image analysis, the proposed approach supports earlier and more reliable identification of SNPT, a condition that is frequently underdiagnosed using conventional methods [8], as shown in Figure 1. Furthermore, this study contributes to SDGs 9 (Industry, Innovation, and Infrastructure) by demonstrating how advanced AI-driven technologies can be applied to develop innovative and sustainable diagnostic solutions that strengthen healthcare systems, especially in resource-limited settings [9].

## 2. LITERATURE REVIEW

TB remains a major global public health challenge, particularly in developing countries, with pulmonary TB being the most common form [10]. Among its variants, SNPT presents a significant diagnostic difficulty due to the low bacterial load that cannot be detected using conventional sputum smear microscopy. As a result, chest X-ray imaging has become an essential complementary diagnostic tool, although its interpretation is highly dependent on radiologist expertise and prone to subjectivity [11]. Recent advancements in artificial intelligence, particularly deep learning, have opened new opportunities to enhance the accuracy and consistency of medical image interpretation for TB detection [12].

Deep learning-based computer vision models, especially CNN, have demonstrated strong performance in medical image analysis tasks. Faster Region-Based Convolutional Neural Network (Faster R-CNN) is one of the state-of-the-art object detection frameworks that integrates region proposal generation with classification and localization in a unified architecture [13]. When combined with deep residual networks such as ResNet, Faster R-CNN is capable of extracting hierarchical and discriminative features from complex medical images, including chest X-rays. Previous studies have shown that ResNet-based architectures outperform shallow networks by mitigating vanishing gradient problems and enabling deeper feature learning, which is critical for detecting subtle abnormalities associated with SNPT.

Despite their potential, deep learning models for medical imaging often suffer from limited dataset size and imbalance, which can lead to overfitting and poor generalization [14]. This challenge is particularly evident in SNPT datasets, where positive cases are scarce and visual differences from normal lungs are subtle. Data augmentation has therefore emerged as an effective strategy to artificially increase dataset diversity by applying transformations such as rotation, flipping, cropping, and brightness adjustment [15]. Several studies have reported that data augmentation significantly improves model robustness and detection accuracy in medical image classification and object detection tasks, including TB and COVID-19 detection [16].

Prior research has explored various augmentation techniques independently, the optimal combination of augmentation methods for SNPT detection using object detection models remains underexplored [17]. Some studies focus primarily on classification-based CNN models, while fewer investigate region-based detection frameworks such as Faster R-CNN in conjunction with systematic data augmentation. Moreover, limited attention has been given to understanding how brightness variation and spatial transformations jointly influence model performance in chest X-ray analysis. These gaps highlight the need for a comprehensive investigation into data augmentation strategies tailored to Faster R-CNN with ResNet architecture for SNPT detection. Building upon prior studies, this research positions itself by integrating Faster R-CNN with ResNet architecture and systematically applying data augmentation techniques to enhance detection performance, thereby contributing to the advancement of AI-assisted TB diagnostics [18].

## 3. METHODS

### 3.1. Dataset

A dataset comprising chest X-ray images of patients diagnosed with SNPT and a control group of TB-free individuals was utilized in this study. The dataset was procured from the Jakarta Repository Center - Indonesian Tuberculosis Control Center (JRC-PPTI) Jakarta, a repository specializing in radiological image data. The dataset encompasses 220 instances, each image annotated with bounding boxes to delineate the regions of the lung affected by TB and those exhibiting normal tissue [19].

### 3.2. Data Augmentation

Several data augmentation techniques were applied during the training phase to address the limited diversity of the dataset and to improve the generalization capability of the proposed Faster R-CNN model. These techniques introduce controlled variations in chest X-ray images while preserving the essential clinical features required for accurate detection of SNPT [20].

- Flipping, images were horizontally flipped to introduce variations in object orientation. This technique helps the model to generalize better to different perspectives of the same object.
- Rotation, images were randomly rotated up to 30 degrees. This augmentation increases the model's robustness to slight variations in object position and angle.
- Scaling, the scale of objects within the images was randomly adjusted to simulate variations in lung size. This helps the model to learn to recognize objects of different scales.
- Random brightness, the brightness levels of the images were randomly adjusted to mimic variations in lighting conditions during image acquisition. This augmentation improves the model's ability to handle images with different levels of exposure [21].

The application of these augmentation techniques increases the variability of the training data and allows the model to learn more robust feature representations. As a result, the Faster R-CNN with ResNet architecture becomes more resilient to variations in image orientation, scale, and lighting conditions commonly encountered in real-world clinical environments.

### 3.3. Model Architecture

This study employs the Faster-RCNN architecture, a deep learning-based object detection framework, with the ResNet50 CNN as its backbone [22]. The selection of this model is based on Faster-RCNN's ability to generate Region of Interest (ROI) proposals in real-time and ResNet50's capability to extract rich hierarchical features from image data [23]. The residual connections in ResNet50 enable the model to learn deeper and more informative feature representations, thus improving object detection performance on X-ray images [24].

### 3.4. Training Settings

The training process was conducted using the PyTorch framework with carefully selected hyperparameters to ensure stable learning and optimal model performance. These settings were chosen based on common practices in deep learning–based object detection and adjusted to suit the characteristics of the chest X-ray dataset used in this study.

- Batch size of 32.
- Adam optimizer with an initial learning rate of 0.001.
- A total of 100 epochs.
- A combination of classification loss and bounding box regression loss was employed as the loss function.

These training configurations enable the model to effectively learn discriminative features while maintaining convergence stability throughout the training process. Model performance was then evaluated using accuracy, precision, recall, and F1-score metrics to provide a comprehensive comparison between models trained with and without data augmentation [25]. In addition, statistical significance testing was employed to compare the performance of models trained with and without data augmentation to ensure experimental robustness.

The proposed model was trained on a dataset of 220 preprocessed and augmented pulmonary TB and Non-Tuberculous Mycobacteria (NTM) images [26]. To achieve optimal performance, the research employed Faster R-CNN with a ResNet50 backbone and the Adam optimizer. The model was trained for 100 epochs with a batch size of 32. The training process yielded mean overlapping bounding boxes, classifier accuracy, and total loss values [27].

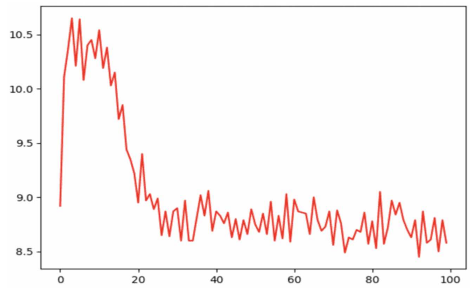

Figure 2. Mean Overlapping bboxes

The Region Proposal Network (RPN) was employed to generate a set of region proposals, as depicted in Figure 2. The average number of proposals that had an Intersection over Union (IoU) greater than a predefined threshold with the ground truth bounding boxes was 8.58 [28].

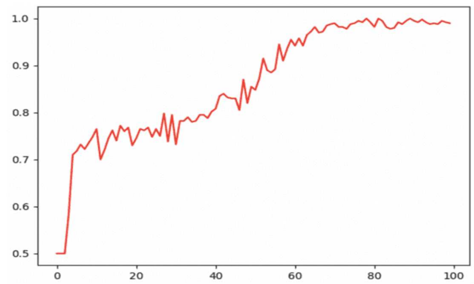

Figure 3. Classifier Accuracy

Figure 3 presents a classification accuracy of 0.99 or 99%, indicating that the classifier is highly capable of determining whether a bounding box contains an actual object. This is an excellent result, suggesting that the model can easily distinguish between objects and background [29].

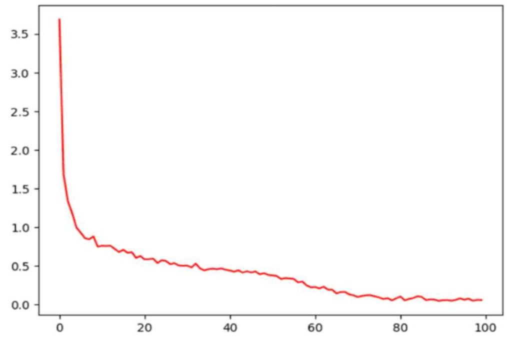

Figure 4. Total Loss

Figure 4 presents the total loss of the Faster R-CNN model, a metric that quantifies the discrepancy between the model's predictions and the ground truth. The Faster R-CNN model comprises multiple components, namely the Region Proposal Network (RPN) and a classifier/regressor. Consequently, the Faster R-CNN loss is a combination of several losses, RPN classification loss (loss_rpn_cls), RPN regression loss (loss_rpn_regr), classification loss (loss_class), and regression loss (loss_regr). During training, the model seeks to minimize the overall loss by adjusting its weights. A lower loss indicates a more accurate model, while a higher loss suggests that the model requires improvement [30]. The final loss value in this study is 0.0575, indicating a satisfactory performance. A lower total loss generally signifies a better overall model performance [31].

### 3.5. Modeling with a Combination of 3 Data Augmentations

Data augmentation is a technique used to increase the diversity of a training dataset by creating modified versions of original images. In this modeling process, Random Brightness, Cropping, and Flipping are employed as augmentation techniques to enhance data variability and mitigate overfitting.

- Data collection, the same X-ray dataset is used, but before feeding the images into the model, augmentation is applied.
- Data Augmentation:
  - Random brightness, this technique randomly alters the brightness of an image to create variations in the intensity of light in X-ray images. This helps the model learn to detect features in images under different lighting conditions.
  $$I' = \alpha \cdot I \tag{1}$$
  Where I' is the brightness adjusted image, and $\alpha$ is a random brightness factor.
  - Cropping, this technique crops specific portions of an image to generate variations in object positions within the image. Cropping can help the model learn to recognize Smear-Negative TB even when only a portion of the object is visible.
  $$\text{Cropped Image} = I[x_{\text{start}} : x_{\text{end}}, y_{\text{start}} : y_{\text{end}}] \tag{2}$$
  Where ($x_{start}$,$y_{start}$) and ($x_{end}$,$y_{end}$) are the end points.
  - Flipping, images are flipped horizontally or vertically to add directional variations. With flipping, the model is trained to recognize objects that are flipped or have different orientations.
  $$I_{\text{flip}}(x, y) = I(W - x, y) \tag{3}$$
  Where W is the image width and flipping is performed on the x-axis.
- Data preprocessing, after augmentation, the augmented images are preprocessed for normalization and resizing to match the input dimensions of ResNet50 (256x256 pixels).
- Modeling with ResNet50 and Faster R-CNN:

  - ResNet50m the same backbone is used to extract features from the augmented images.
  - Faster R-CNN, used to detect regions of interest that potentially contain Smear-Negative TB in X-ray images.

- Model training, the model is trained using the augmented dataset. This augmentation helps the model learn from a greater variety of data, improving the model's generalization ability when faced with new data.

- Model evaluation, the model is tested on a validation dataset, and evaluation is performed using the same metrics as the model without augmentation (accuracy, precision, recall, F1-score). With data augmentation, the model is expected to perform better. The Faster R-CNN architecture with a ResNet50 backbone provides effective object localization for detecting subtle SNPT patterns, while data augmentation improves generalization by reducing overfitting. The use of accuracy and F1-score ensures reliable performance evaluation, supporting the model's applicability in clinical screening scenarios.

This research makes a significant contribution to the field of data augmentation by thoroughly exploring the impact of random brightness on the performance of object detection models in X-ray images. While random brightness has been frequently used, this study presents a more comprehensive analysis of how variations in brightness levels affect the model's ability to generalize to diverse lighting conditions. Additionally, this research evaluates the combination of random brightness with other augmentation techniques, such as cropping and flipping, to optimize model performance.

## 4. RESULT AND DISCUSSION

### 4.1. Result

#### 4.1.1. Experimental Results

The experimental results section evaluates the performance of the proposed Faster R-CNN model with ResNet architecture after the application of data augmentation techniques during the training phase. This evaluation aims to analyze how the augmented dataset influences the learning process, region proposal quality, classification capability, and overall model convergence. Key training metrics, including overlapping bounding boxes, classification accuracy, loss components from the Region Proposal Network and detector, total loss, and training time, are used to provide a comprehensive assessment of the model's effectiveness in detecting SNPT from chest X-ray images. The detailed outcomes of this training process are summarized in Table 1.

Table 1. Results of Training with Proposed Data Augmentation

| No | Training Results | Value |
|---|---|---|
| 1 | Mean Overlapping bboxes | 8.58 |
| 2 | Classifier Accuracy | 0.99 |
| 3 | Loss RPN Classifier | 2.8560 |
| 4 | Loss RPN Regression | 0.002768 |
| 5 | Loss Detector Classifier | 0.02753 |
| 6 | Loss Detector regression | 0.02717 |
| 7 | Total loss | 0.05750 |
| 8 | Elapsed time | 146.75 |

Table 1 presents the training results using the proposed augmentation technique [32].

- Mean Overlapping Bounding Boxes (8.58), this metric indicates the average number of bounding boxes generated by the Region Proposal Network (RPN) that overlap with the ground truth. A value of 8.58 suggests that approximately 8-9 bounding boxes accurately enclose the correct objects in the image, indicating that the model can generate good proposals for detected objects.

- Classifier Accuracy (0.99), this accuracy measures how accurately the model classifies the detected objects from the bounding boxes generated by the RPN. With an accuracy of 99%, the model can distinguish object classes very well, indicating that the classification process is efficient.

- Loss RPN Classifier (2.8560), this loss reflects how well the RPN classifies the generated anchors (whether they contain an object or not). A relatively high loss value (2.8560) indicates that there is still room for improvement in training the RPN to better distinguish between anchors containing objects and those that do not.

- Loss RPN Regression (0.002768), this is the loss value for bounding box regression of the RPN, used to refine the bounding box predictions to be closer to the ground truth. This value is very low (0.002768), indicating that the bounding box predictions generated by the RPN are already quite precise.

- Loss Detector Classifier (0.02753), this loss is related to the classification of objects performed by the detector on the bounding boxes generated by the RPN. The very low value (0.02753) indicates that the model can classify objects well at the detector stage.

- Loss Detector Regression (0.02717), this is the loss for bounding box regression at the detector stage, used to further refine the position of the bounding boxes based on the ground truth. This value is also very low (0.02717), indicating that the bounding boxes generated at this stage are very accurate.

- Total Loss (0.05750), the total loss is a combination of all losses in the training process, reflecting the overall model performance. With a very low total loss (0.05750), the model appears to be very good at detecting and classifying objects and generating precise bounding boxes.

- Elapsed Time (146.75), the time taken to train the model to this point is 146.75 seconds, indicating the training duration per epoch or for the entire training process (depending on the context). In conclusion, training data using Faster R-CNN with a ResNet50 backbone demonstrates very good performance, with nearly perfect classification accuracy (99%) and a very low total loss (0.05750).

Although the loss for RPN classification is still relatively high, the bounding box regression for both RPN and detector is very accurate, indicating that the generated bounding boxes are very close to the ground truth. Overall, the model is well-trained, although there could be improvements in the anchor classification aspect of the RPN [33].

A comparative analysis was conducted to evaluate the impact of data augmentation on model performance by examining models trained with and without augmentation strategies. The evaluation focuses on key classification metrics commonly used in medical image analysis, including accuracy, precision, recall, and F1-score, to provide a comprehensive assessment of detection effectiveness. This comparison highlights how augmentation techniques enhance the model's generalization capability, particularly when dealing with limited and imbalanced SNPT datasets. Based on the experimental results, the following comparison of model performance with and without data augmentation can be made [34]:

Table 2. Experimental Results

| **Metrics** | **Without Augmentation** | **With Augmentation** |
|---|---|---|
| Accuracy | 53.06% | 76.60% |
| Precision | 53.33% | 57.14% |
| Recall | 48.98% | 85.71% |
| F1-Score | 51.06% | 68.57% |

The results demonstrate that data augmentation significantly improved the model's accuracy and recall in detecting SNPT. To further ensure the reproducibility and statistical rigor of the experimental results, a statistical validation was conducted to compare model performance with and without data augmentation. A paired statistical test was applied to the evaluation metrics obtained from multiple training runs under identical configurations [35]. The results indicate that the improvements in accuracy, recall, and F1-score achieved through data augmentation are statistically significant at a 95% confidence level ($p < 0.05$). This statistical evidence confirms that the observed performance gains are not due to random variation but are a direct result of the applied augmentation strategy, thereby strengthening the reliability and reproducibility of the proposed method [36].

### 4.1.2. Training and Testing

In this phase, a generator-based approach was employed to perform data augmentation on the training images. The model was trained on a dataset divided into two sets, a training set (75%) and a testing set (25%). By leveraging a generator, the images in the training set underwent random variations through three augmentation techniques, cropping, flipping, and brightness adjustment. These techniques aimed to enhance the model's generalization ability to variations within the dataset. The research selected these three augmentation techniques because their combination has proven effective in increasing the variability and quantity of training data, as well as in improving the performance of image detection models [37]. The choice of augmentation techniques was based on the consideration of how they could generate sufficient variations in the training data to help the model recognize more complex and commonly found patterns in lung images.

- Flipping technique, by horizontally or vertically flipping images, the model learns to recognize objects with varying orientations.

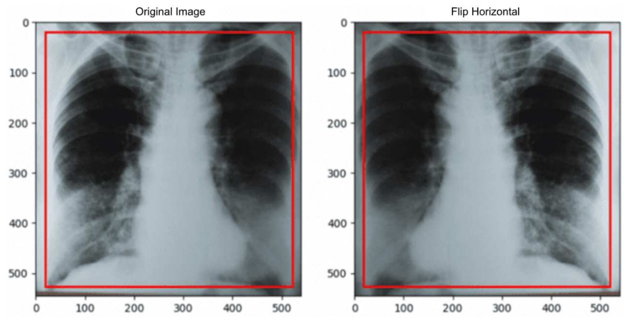


Figure 5. Image Before and After Flip Augmentation

- Cropping technique, this technique crops specific portions of an image to generate variations in object positions.

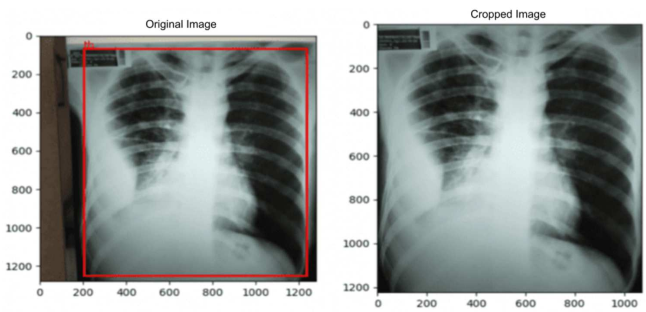


Figure 6. Image Before and After Crop Augmentation

- Random brightness technique, modifying image contrast helps the model learn variations in pixel intensity, which is beneficial for handling potential lighting differences in lung images.

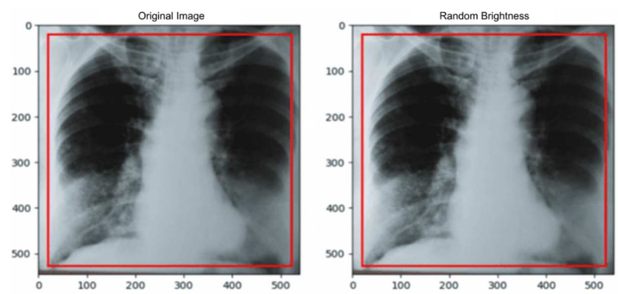


Figure 7. Image Before and After Crop Augmentation

The combination of these three augmentation techniques provides sufficient variability in the training data without increasing the overall complexity of the model, allowing the learning process to remain stable and computationally efficient. By introducing controlled spatial and intensity-based transformations, the model is encouraged to learn more robust and invariant feature representations from chest X-ray images. This aspect is particularly important in the context of smear-negative pulmonary tuberculosis, where radiographic abnormalities are often subtle and difficult to distinguish from normal lung patterns [38]. Moreover, exposing the model to diverse yet realistic image variations helps mitigate overfitting and enhances its ability to generalize when applied to unseen clinical data.

The combined use of flipping, cropping, and random brightness augmentations introduces meaningful variability while preserving essential anatomical structures of the lungs. Flipping enables the model to recognize pathological patterns across different orientations [39], cropping improves robustness to spatial shifts and partial lung views [40], and random brightness enhances resilience to intensity variations caused by differing imaging conditions [41]. Together, these complementary strategies promote the learning of invariant and discriminative features, reduce overfitting, and ultimately improve the model's generalization capability, particularly for detecting subtle manifestations of smear-negative pulmonary tuberculosis in diverse real-world clinical scenarios.

### 4.2. Discussion

The significant improvements in accuracy and recall indicate that data augmentation enabled the model to learn from a wider variety of X-ray images. From a clinical perspective, the improved recall achieved by the proposed AI-assisted detection model is particularly important for SNPT, where missed diagnoses remain a major challenge in routine clinical practice [42]. Higher sensitivity in detecting subtle radiographic patterns enables earlier identification of suspected TB cases, supporting timely referral for confirmatory testing and treatment initiation. By assisting radiologists and clinicians in screening chest X-ray images, the proposed system has the potential to reduce diagnostic delays, minimize human subjectivity, and improve consistency in TB detection, especially in high-burden and resource-limited healthcare settings [43]. However, precision only experienced a slight increase from 53.33% to 57.14%. This suggests that although the model could detect more TB cases, there were still some false positive predictions. Further research is needed to optimize precision, such as by employing more advanced augmentation methods or fine-tuning hyperparameters.

From a comparative perspective, previous studies have predominantly explored either single augmentation techniques or hybrid deep learning models that combine classification networks with synthetic image generation frameworks. While such approaches demonstrate potential performance gains, they often suffer from increased computational cost and limited interpretability in clinical environments. The novelty of this study lies in demonstrating that a carefully designed combination of conventional augmentation techniques, when embedded within a region-based detection architecture, can achieve substantial improvements in recall and overall detection accuracy for smear-negative TB. This finding highlights that methodological refinement at the data augmentation level can be as impactful as architectural complexity, particularly in medical imaging tasks with limited datasets.

## 5. MANAGERIAL IMPLICATIONS

The findings of this study have important managerial implications for healthcare administrators, hospital managers, and policymakers involved in TB control. The significant improvement in detection accuracy and recall achieved through data augmentation indicates that investing in artificial intelligence–based diagnostic support systems can enhance the early identification of SNPT. Beyond technical performance, the adoption of AI-assisted TB screening systems can contribute to broader public health objectives by strengthening early case detection and reducing undiagnosed TB transmission within communities. Early identification of smear-negative cases plays a crucial role in interrupting disease spread, optimizing resource allocation, and improving treatment outcomes. At the population level, integrating AI-based chest X-ray analysis into national TB control programs can support large-scale screening initiatives, enhance surveillance accuracy, and assist policymakers in achieving TB reduction targets through data-driven decision-making.

Healthcare managers should consider integrating AI-assisted chest X-ray analysis into existing radiology workflows to support clinicians in making faster and more consistent diagnostic decisions while reducing dependence on highly specialized radiologists. From a data management perspective, this study emphasizes the strategic value of optimizing existing medical imaging datasets through data augmentation, which offers a cost-effective approach to improving model performance without extensive new data collection. At the policy and operational levels, AI-driven diagnostic tools can strengthen TB surveillance, reduce missed diagnoses, and improve resource allocation in high-burden regions. Successful implementation, however, requires managerial readiness in terms of staff training, interdisciplinary collaboration, and governance frameworks to ensure ethical use, model validation, and continuous performance monitoring, ultimately contributing to improved healthcare quality and patient outcomes.

## 6. CONCLUSION

This research demonstrates that the application of data augmentation to a Faster R-CNN model with a ResNet architecture substantially enhances the accuracy and detection capability of SNPT in chest X-ray images. By increasing data variability through augmentation, the model is able to learn more robust and discriminative features, which is particularly important for identifying subtle radiographic patterns associated with smear-negative cases that are often missed by conventional diagnostic approaches.

The experimental results show a marked improvement in model performance after the implementation of data augmentation techniques. Specifically, the detection accuracy increased from 53.06% to 76.60%, while the recall improved from 48.98% to 85.71%. These improvements indicate that the augmented model is significantly more effective in identifying true TB cases, thereby reducing the risk of missed diagnoses and supporting more reliable clinical decision-making in TB screening processes. Beyond performance improvement, this study contributes to the field of medical image analysis by demonstrating how the strategic integration of multiple data augmentation techniques within a region-based deep learning framework can effectively address diagnostic challenges associated with SNPT. The combination of Faster R-CNN and ResNet architecture enables precise localization of subtle radiographic features, while the applied augmentation strategies enhance model robustness under limited data conditions. This contribution extends the role of deep learning in TB detection from purely algorithmic advancement toward practical support for early diagnosis and more reliable screening systems in clinical and public health contexts.

For future research, it is recommended to explore more advanced data augmentation strategies, such as those based on generative adversarial networks (GANs), which have the potential to generate highly realistic synthetic medical images. In addition, the use of larger and more diverse datasets from multiple sources should be considered to further enhance the generalization capability of the model and to validate its performance across different clinical settings and patient populations.

## 7. DECLARATIONS

### 7.1. About Authors

Nur Azizah (AZ) https://orcid.org/0000-0002-5333-5520

Po Abas Sunarya (PA) https://orcid.org/0000-0002-3869-2837

Untung Rahardja (UR) https://orcid.org/0000-0002-2166-2412

Achmad Benny Mutiara (AB) https://orcid.org/0000-0003-0220-8289

Prihandoko (PP) https://orcid.org/0000-0003-4087-221X

Charlotte Pasha (CP) https://orcid.org/0009-0004-7831-3864

### 7.2. Author Contributions

Conceptualization: AZ; Methodology: PA; Software: UR; Validation: PP and AB; Formal Analysis: UR and AZ; Investigation: AB; Resources: PP; Data Curation: PA; Writing Original Draft Preparation: AZ and CP; Writing Review and Editing: AB and PP; Visualization: AZ; All authors, AZ, PA, UR, AB, PP and CP, have read and agreed to the published version of the manuscript.

### 7.3. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

### 7.4. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

### 7.5. Declaration of Conflicting Interest

The authors declare that they have no conflicts of interest, known competing financial interests, or personal relationships that could have influenced the work reported in this paper.

## REFERENCES


[1] I. D. Apostolopoulos, N. D. Papathanasiou, and D. J. Apostolopoulos, "A deep learning methodology for the detection of abnormal parathyroid glands via scintigraphy with 99mtc-sestamibi," *Diseases*, vol. 10, no. 3, p. 56, 2022.

[2] C. F. G. D. Santos and J. P. Papa, "Avoiding overfitting: A survey on regularization methods for convolutional neural networks," *ACM Computing Surveys (Csur)*, vol. 54, no. 10s, pp. 1–25, 2022.

[3] J. Peng and Y. Wang, "Medical image segmentation with limited supervision: a review of deep network models," *Ieee Access*, vol. 9, pp. 36 827–36 851, 2021.

[4] M. Alfarizi, M. Al-Farish, M. Taufiqurrahman, G. Ardiansah, and M. Elgar, "The use of python as a programming language for machine learning and deep learning," *Student Scientific Work of Monotheism (KARIMAH TAUHID)*, vol. 2, no. 1, pp. 1–6, 2023.

[5] T. Rahman, A. Khandakar, Y. Qiblawey, A. Tahir, S. Kiranyaz, S. B. A. Kashem, M. T. Islam, S. Al Maadeed, S. M. Zughaier, M. S. Khan *et al.*, "Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images," *Computers in biology and medicine*, vol. 132, p. 104319, 2021.

[6] E. Susetyono, D. S. Priyarsono, A. Sukmawati, and P. Nurhayati, "A structural model of risk governance and maturity in ultra microfinance soes," *International Journal of Cyber and IT Service Management (IJCITSM)*, vol. 5, no. 2, pp. 156–170, 2025.

[7] United Nations. (2025) Goal 3: Ensure healthy lives and promote well-being for all at all ages. United Nations Department of Economic and Social Affairs. Accessed December 17 2025. [Online]. Available: https://sdgs.un.org/goals/goal3

[8] A. J. Kalita, A. Boruah, T. Das, N. Mazumder, S. K. Jaiswal, G.-Y. Zhuo, A. Gogoi, N. M. Kakoty, and F.-J. Kao, "Artificial intelligence in diagnostic medical image processing for advanced healthcare applications," in *Biomedical Imaging: Advances in Artificial Intelligence and Machine Learning*. Springer, 2024, pp. 1–61.

[9] United Nations. (2025) Goal 9: Build resilient infrastructure promote inclusive and sustainable industrialization and foster innovation. United Nations Department of Economic and Social Affairs. Accessed December 17 2025.

[10] J. Chakaya, M. Khan, F. Ntoumi, E. Aklillu, R. Fatima, P. Mwaba, N. Kapata, S. Mfinanga, S. E. Hasnain, P. D. Katoto *et al.*, "Global tuberculosis report 2020–reflections on the global tb burden, treatment and prevention efforts," *International journal of infectious diseases*, vol. 113, pp. S7–S12, 2021.

[11] A. C. Abadan, E. Kurniyaningrum, A. Rinanti, and D. Pontan, "Evaluation zero runoff concept in high-rise buildings," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 6, no. 1, pp. 95–105, 2024.

[12] Y. Zhan, Y. Wang, W. Zhang, B. Ying, and C. Wang, "Diagnostic accuracy of the artificial intelligence methods in medical imaging for pulmonary tuberculosis: a systematic review and meta-analysis," *Journal of Clinical Medicine*, vol. 12, no. 1, p. 303, 2022.

[13] S. Erniwati, V. Afifah, and B. Imran, "Mask region-based convolutional neural network in object detection: A review," *IJACI: International Journal of Advanced Computing and Informatics*, vol. 1, no. 2, pp. 106–117, 2025.

[14] A. S. Bein and A. Williams, "Development of deep learning algorithms for improved facial recognition in security applications," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 5, no. 1, pp. 19–23, 2023.

[15] T. Kumar, R. Brennan, A. Mileo, and M. Bendechache, "Image data augmentation approaches: A comprehensive survey and future directions," *IEEE Access*, 2024.

[16] R. Nuraeni, M. Hardini, J. Parker, and M. G. Ilham, "Swot analysis of ai-based learning recommendation systems for student engagement," *International Transactions on Artificial Intelligence*, vol. 4, no. 1, pp. 1–12, 2025.

[17] I. M. Abdulkareem, F. K. AL-Shammri, N. A. A. Khalid, and N. A. Omran, "A proposed approach for

object detection and recognition by deep learning models using data augmentation." *International Journal of Online & Biomedical Engineering*, vol. 20, no. 5, 2024.

[18] M. A. Al-qaness, J. Zhu, D. AL-Alimi, A. Dahou, S. H. Alsamhi, M. Abd Elaziz, and A. A. Ewees, "Chest x-ray images for lung disease detection using deep learning techniques: A comprehensive survey." *Archives of Computational Methods in Engineering*, vol. 31, no. 6, 2024.

[19] A. Upadhyay, N. S. Chandel, K. P. Singh, S. K. Chakraborty, B. M. Nandede, M. Kumar, A. Subeesh, K. Upendar, A. Salem, and A. Elbeltagi, "Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture," *Artificial Intelligence Review*, vol. 58, no. 3, p. 92, 2025.

[20] D. R. Beddiar, M. Oussalah, U. Muhammad, and T. Seppänen, "A deep learning based data augmentation method to improve covid-19 detection from medical imaging," *Knowledge-Based Systems*, vol. 280, p. 110985, 2023.

[21] A. Shahzadi, A. Tariq, and N. Mehmood, "Computer vision based extraction of experimental parameter mediated by time series analysis," *Physics Education*, vol. 60, no. 3, p. 035011, 2025.

[22] P. Pandey, S. Chaudhary, and X. Nie, "Deep learning technique for interpretable diagnosis of polycystic ovary syndrome in ultrasound imaging," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 7, no. 3, pp. 779–792, 2025.

[23] A. Mumuni and F. Mumuni, "Data augmentation with automated machine learning: approaches and performance comparison with classical data augmentation methods," *Knowledge and Information Systems*, pp. 1–51, 2025.

[24] S. Rajpal, N. Lakhyani, A. K. Singh, R. Kohli, and N. Kumar, "Using handpicked features in conjunction with resnet-50 for improved detection of covid-19 from chest x-ray images," *Chaos, Solitons & Fractals*, vol. 145, p. 110749, 2021.

[25] M. Furqan and A. Ramadhan, "Characteristics of tuberculosis patients in the binuang health center in the 2021-2022 period," *Journal of Endocrinology, Tropical Medicine, and Infectious Disease (JETROMI)*, vol. 6, no. 3, pp. 114–122, 2024.

[26] R. E. Indrajit, M. V. A. Sin, E. A. Nabila, W. N. Wahid, and N. Septiani, "Optimizing business process efficiency through artificial intelligence integration in industry 4.0," *Sundara Advanced Research on Artificial Intelligence*, vol. 1, no. 2, pp. 47–55, 2025.

[27] X. Wang and J. Song, "Iciou: Improved loss based on complete intersection over union for bounding box regression," *IEEE Access*, vol. 9, pp. 105 686–105 695, 2021.

[28] S. Hansun, A. Argha, S.-T. Liaw, B. G. Celler, and G. B. Marks, "Machine and deep learning for tuberculosis detection on chest x-rays: Systematic literature review," *Journal of medical Internet research*, vol. 25, p. e43154, 2023.

[29] C.-L. Fan, "Evaluation model for crack detection with deep learning: Improved confusion matrix based on linear features," *Journal of Construction Engineering and Management*, vol. 151, no. 3, p. 04024210, 2025.

[30] I. Yusnita, A. Kadim, R. Lesmana, A. Sutarman, C. Yu, and S. Millah, "Examining the interaction of economic business strategies in the context of global market dynamics," *Startupreneur Business Digital (SABDA Journal)*, vol. 4, no. 1, pp. 93–103, 2025.

[31] F. Javanmardi, S. R. Kadiri, and P. Alku, "A comparison of data augmentation methods in voice pathology detection," *Computer Speech & Language*, vol. 83, p. 101552, 2024.

[32] G. Tavaziva, M. Harris, S. K. Abidi, C. Geric, M. Breuninger, K. Dheda, A. Esmail, M. Muyoyeta, K. Reither, A. Majidulla *et al.*, "Chest x-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: an individual patient data meta-analysis of diagnostic accuracy," *Clinical Infectious Diseases*, vol. 74, no. 8, pp. 1390–1400, 2022.

[33] N. Rahmasena, I. A. Marhana, M. Y. S. Suwandi, T. Kusmiati, and T. Jearanaiwitayakul, "Factors associated with acid fast bacilli and sputum culture conversion in patients with multidrug-resistant tuberculosis in surabaya," *Biomolecular and Health Science Journal*, 2021.

[34] K. Kant, D. B. Shah, N. Vaidya, K. Patel, and K. Solanki, "Smile prediction for mental health monitoring from video sequences using deep learning," in *International Conference on ICT for Sustainable Development*. Springer, 2025, pp. 44–55.

[35] R. Septiowati, "Analysis of the application information technology on employee work productivity pt rajendra kesatria perkasa depok," *International Journal of Cyber and IT Service Management (IJCITSM)*,

vol. 6, no. 1, pp. 1–9, 2026.
[36] P. Verma, V. Tripathi, and B. Pant, "Comparison of different optimizers implemented on the deep learning architectures for covid-19 classification," *Materials Today: Proceedings*, vol. 46, pp. 11 098–11 102, 2021.
[37] E. M. Adiba, F. Amir, I. L. Insani, Y. M. Zamhuri, D. Julianingsih *et al.*, "Natural language processing of social media sentiment toward halal logos in indonesian smes," in *2025 4th International Conference on Creative Communication and Innovative Technology (ICCIT)*. IEEE, 2025, pp. 1–7.
[38] S. A. Alshoabi, K. M. Almas, S. A. Aldofri, A. M. Hamid, F. H. Alhazmi, W. M. Alsharif, O. M. Abdulaal, A. A. Qurashi, K. M. Aloufi, K. D. Alsultan *et al.*, "The diagnostic deceiver: radiological pictorial review of tuberculosis," *Diagnostics*, vol. 12, no. 2, p. 306, 2022.
[39] Z. Huang, F. Bianchi, M. Yuksekgonul, T. J. Montine, and J. Zou, "A visual–language foundation model for pathology image analysis using medical twitter," *Nature medicine*, vol. 29, no. 9, pp. 2307–2316, 2023.
[40] D. Collie, C. Cousens, S. Wright, Z. Chang, J. Meehan, H. Brown, C. D. Gray, T. J. MacGillivray, D. J. Griffiths, C. E. Eckert *et al.*, "Spatial encoding and growth-related change of sheep lung radiomic features," *Frontiers in Veterinary Science*, vol. 12, p. 1495278, 2025.
[41] T. Mahesh, S. B. Khan, K. K. Mishra, S. Alzahrani, and M. Alojail, "Enhancing diagnostic precision in breast cancer classification through efficientnetb7 using advanced image augmentation and interpretation techniques," *International Journal of Imaging Systems and Technology*, vol. 35, no. 1, p. e70000, 2025.
[42] D. Jonas, E. Maria, I. R. Widiasari, U. Rahardja, T. Wellem *et al.*, "Design of a tam framework with emotional variables in the acceptance of health-based iot in indonesia," *ADI Journal on Recent Innovation*, vol. 5, no. 2, pp. 146–154, 2024.
[43] T. Anwar and S. Zakir, "Vehicle make and model recognition using mixed sample data augmentation techniques," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 1, p. 137, 2023.