



OPEN ACCESS

EDITED BY

L. J. Muhammad,
Bayero University Kano, Nigeria

REVIEWED BY

Yuanda Zhu,
Independent Researcher, Atlanta, GA,
United States
Seng Hansun,
Multimedia Nusantara University, Indonesia

*CORRESPONDENCE

Mohamudha Parveen Rahamathulla
✉ m.parveen@psau.edu.sa

RECEIVED 27 March 2024

ACCEPTED 29 May 2024

PUBLISHED 27 June 2024

CITATION

Parveen Rahamathulla M, Sam Emmanuel WR,
Bindhu A and Mustaq Ahmed M (2024)
YOLOv8's advancements in tuberculosis
identification from chest images.
Front. Big Data 7:1401981.
doi: 10.3389/fdata.2024.1401981

COPYRIGHT

© 2024 Parveen Rahamathulla, Sam
Emmanuel, Bindhu and Mustaq Ahmed. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

YOLOv8's advancements in tuberculosis identification from chest images

Mohamudha Parveen Rahamathulla^{1*}, W. R. Sam Emmanuel²,
A. Bindhu³ and Mohamed Mustaq Ahmed⁴

¹Department of Basic Medical Science, College of Medicine, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia, ²Department of Computer Science and Research Centre, Nesamony Memorial Christian College, Marthandam, Tamil Nadu, India, ³Department of Computer Science, Infant Jesus College of Arts and Science for Women, Mulagumoodu, Tamil Nadu, India, ⁴Department of Information Technology, The New College, Chennai, Tamil Nadu, India

Tuberculosis (TB) is a chronic and pathogenic disease that leads to life-threatening situations like death. Many people have been affected by TB owing to inaccuracy, late diagnosis, and deficiency of treatment. The early detection of TB is important to protect people from the severity of the disease and its threatening consequences. Traditionally, different manual methods have been used for TB prediction, such as chest X-rays and CT scans. Nevertheless, these approaches are identified as time-consuming and ineffective for achieving optimal results. To resolve this problem, several researchers have focused on TB prediction. Conversely, it results in a lack of accuracy, overfitting of data, and speed. For improving TB prediction, the proposed research employs the Selection Focal Fusion (SFF) block in the You Look Only Once v8 (YOLOv8, Ultralytics software company, Los Angeles, United States) object detection model with attention mechanism through the Kaggle TBX-11k dataset. The YOLOv8 is used for its ability to detect multiple objects in a single pass. However, it struggles with small objects and finds it impossible to perform fine-grained classifications. To evade this problem, the proposed research incorporates the SFF technique to improve detection performance and decrease small object missed detection rates. Correspondingly, the efficacy of the projected mechanism is calculated utilizing various performance metrics such as recall, precision, F1Score, and mean Average Precision (mAP) to estimate the performance of the proposed framework. Furthermore, the comparison of existing models reveals the efficiency of the proposed research. The present research is envisioned to contribute to the medical world and assist radiologists in identifying tuberculosis using the YOLOv8 model to obtain an optimal outcome.

KEYWORDS

tuberculosis, Yolov8 model, Selection Focal Fusion block, attention mechanism and object detection, computer-aided diagnosis

1 Introduction

Globally, tuberculosis (TB) is a severe public health concern caused by *Mycobacterium tuberculosis*. It significantly affects the human lungs (pulmonary TB), (Kotei and Thirunavukarasu, 2024) but also other parts of the body such as the spine, brain, and kidneys (extrapulmonary TB) (Wang et al., 2024). Usually, TB spreads by air; it transmits from one person to another through coughs and sneezes.

It can be detected through different tests, like blood tests, sputum tests, imaging studies, or skin tests (Maipan-uku et al., 2024). It is crucial to identify it at an early stage and control it with proper treatment. Besides, bacteria are considered the most common cause of TB in healthcare settings (Amin et al., 2024). Henceforth, TB is considered to be the leading cause of death among infectious diseases. Chest radiography is playing a vital role in detecting TB at a low cost. Resource-constrained countries are bearing the burden of TB. However, the insufficient availability of expert readers is considered a major concern hampering the application of chest radiography (Hwang et al., 2024). To evade the prediction problem, an effective disease prediction mechanism is needed to avoid the consequences. Recently, most researchers have utilized Artificial Intelligence (AI)-related (Acharya et al., 2024) technology for the disease prediction of TB and non-TB for the capability of better cost-saving techniques and greater scalability. Furthermore, AI technology can systematize prediction problems and monitor the data efficiently. It can identify abnormalities in the data while in access. These AI benefits with Machine Learning (ML) and Deep Learning (DL) methods (Rahman et al., 2020) provide several benefits for classification mechanisms in TB prediction at various stages.

Congruently, several prevailing models have attempted to accomplish better TB and non-TB prediction. For instance, the existing model, a fine-tuning deep neural network, has been used to perform lesion detection for tuberculosis. The utilized task dataset and the results have shown that the convolutional neural network (CNN) model has attained a better mean Area Under Curve (AUC), which represents the better efficiency of the prevailing model (Lu et al., 2022). Similarly, the conventional model has deployed detection mechanisms for Pulmonary Tuberculosis lesions and utilized the TBX-11 dataset. The outcomes of the conventional model have attained 0.5 Intersection Over Union (IoU) of average precision, and the recall rate has attained 77.6% (An et al., 2022). Correspondingly, Computer-Aided Diagnosis (CAD) utilizing computer vision methods and advanced DL models has been used to improve the accuracy using the YOLOv7 object detection architecture. It has used the TBX-11 dataset and employed data augmentation and class weighting techniques to notice the imbalance present in the dataset. The promising results have shown that the conventional model has attained 0.587 mean Average Precision (mAP) (Bista et al., 2023). Likewise, the prevailing model has deployed a TB detection framework and utilized the Montgomery and Shenzhen datasets. The conventional model was evaluated with a k -fold cross-validation technique and attained 0.97 and 0.99 of AUC for the Montgomery and Shenzhen datasets, respectively (Ayaz et al., 2021). Similarly, the prevailing model has developed DL classification and segmentation models (Iqbal et al., 2022) for precise and accurate detection of TB on chest X-ray (CXR) images with the vision of infection utilizing Gradient-weighted Class Activation Mapping (Grad-CAM) heat maps. It has utilized the NIAID TB portal dataset (Ekins and Freundlich, 2011; Acharya et al., 2022) and has applied the Xception model. The results have attained better recall, F1 score, accuracy, and precision (Sharma et al., 2023). Accordingly, classical models have accomplished satisfactory results, but they lack a few limitations, such as mAP, recall, precision, speed, and overfitting of data.

To solve this problem, the proposed research utilizes certain procedures to improve the performance of the YOLOv8 model for TB and non-TB prediction. Initially, the TBX-11 dataset is loaded into the mechanism where the images are utilized to enhance the proposed model's performance. It is processed with image label pre-processing to improve the consistency and efficiency of the image data. Correspondingly, after the image processing, it is forwarded to pre-processing with augmentation. It transfers the utilized data to augmented samples to improve efficiency. Then, processed data is divided into training, validating, and testing data to train the present model and evaluate its performance. Accordingly, the training data are used for the prediction with Selection Focal Fusion (SFF) in the YOLOv8 model with an attention mechanism. After the prediction mechanism, the validating set is used for performance calculation in the present framework. Additionally, the efficiency of the projected model is calculated utilizing performance metrics to examine the proposed prediction performance. The major contribution of the projected model is discussed in the following:

- To utilize YOLOv8 for the prediction of tuberculosis with the Kaggle TBX-11k dataset to enhance the accuracy and computation in the present research.
- To employ the SFF technique in the YOLOv8 model with attention mechanism for prediction of active TB, latent TB, healthy and sick, but non-TB for enhancing the prediction efficiency.
- To calculate the efficiency of the projected system with performance metrics such as precision, recall, mAP, and F1 score.

1.1 Paper organization

The paper's organization is significant for communicating the research findings in a coherent manner. It is organized on the basis of analyzing the existing methods and the approaches applied under the prediction of Active Tuberculosis (ATB), latent TB, and non-TB in Section 2. Section 3 provides the proposed methodology for the current research. Moreover, the results accomplished through the current system are illustrated in Section 4. The outcomes of future research using the current approach are depicted in Section 5.

2 Literature review

This section explains the analysis of various existing models of ML and DL techniques for the prediction of TB and non-attack in the non-TB classification systems. Furthermore, the problem mentioned in the prevailing research is also identified. The conventional model has evaluated its diagnostic performance utilizing ML techniques (Anand et al., 2021) in differentiating ATB from Latent Tuberculosis Infection (LTBI). It has employed the Gradient Boosting Machine (GBM) model for accurate identification of ATB. The results have shown that the GBM model has attained 89.81% accuracy (Luo et al., 2021). Similarly, the prevailing model has detected Pulmonary Tuberculosis (PTB)

and Extrapulmonary Tuberculosis (EPTB) using ML algorithms. It has employed clinical and imaging data from hospitals and incorporated the Decision Tree algorithm. The result has shown that the Decision Tree (DT) algorithm has attained 95.5% classification accuracy (Kaur and Sharma, 2021). In the same way, the existing model has demonstrated a TB detection mechanism (Heyckendorf et al., 2021) utilizing advanced DL models. It has utilized EfficientNetB3 and the CNN model (Dey et al., 2022) for the classification, along with the available CXR dataset. The result has shown that the model has attained 98.7% accuracy (Nafisah and Muhammad, 2024). Correspondingly, the conventional model has diagnosed the ATB from multiplex serological data. It has utilized Machine Intelligence Learning Optimiser (MILO) for the prediction performance. It has incorporated secondary and tertiary datasets and has resulted in 86% accuracy (Rashidi et al., 2021). Contrarily, the existing model has enhanced the weight voting ensemble learning method to aid in diagnostic development for predicting TB infection at an early stage. It has used TB gene expression data for diagnosis. The results have shown that ensemble classifiers, Support Vector machine (SVM), and Naive Bayes (NB) have attained 95, 92, and 87% accuracy, respectively (Osamor and Okezie, 2021).

Congruently, the conventional system has presented a solution for identifying TB by utilizing Bayesian-based CNN (B-CNN). It has been evaluated with two TB benchmark datasets called Shenzhen and Montgomery. The results have revealed that B-CNN (Yusoff et al., 2021) has attained 86.46 and 96.4% accuracy on both datasets, respectively (Abideen et al., 2020). Contrastingly, the classical model has combined clinical indicators and metabolomics along with ML for a precise identification of Smear-Negative Pulmonary Tuberculosis (SNPT) (Xie et al., 2020) and Smear-Positive Pulmonary Tuberculosis (SPPT). The outcome has shown that the model has achieved 83–93% accuracy (Hu et al., 2022). In parallel, the prevailing model has presented an automatic cough classification for TB (Sathitratanacheewin et al., 2020) and utilized ML methods such as KNN, MLP, LR, CNN, and SVM. It has used nested cross-validation (Singh et al., 2021). Among these ML methods, LR has outperformed with 84.54% accuracy (Liu et al., 2017; Pahar et al., 2021). Simultaneously, the conventional method has suggested a DL binary classifier for TB and non-TB diagnosis utilizing chest X-rays. It has employed a 2-step binary DT and has been trained through CNN on the PyTorch frame. The prevailing model has used the Shenzhen dataset, and results have shown 98 and 80% accuracy at both the first and second steps, respectively (Yoo et al., 2020). Similarly, the conventional mechanism has developed a TB and non-TB detection and Drug-Resistant Categorization Diagnosis Decision Support System (TB-DRC-DSS) (Zhu et al., 2023) using the DL ensemble model (Rajaraman and Antani, 2020) with different CNN architectures such as Dense-Net121, mobileNetV2, and EfficientNetB7. It has incorporated the Shenzhen, (Rajaraman et al., 2021) Kaggle, Montgomery, (Le et al., 2022) and Portal datasets. The results have revealed that TB-DRC-DSS has attained 92.8% accuracy (Prasitpuriprecha et al., 2022).

In contrast, the prevailing mechanism has applied automatic AI detection for screening a huge populace and has evaluated its feasibility. It has helped in diagnosing TB utilizing CXR

radiographs. The outcomes of the conventional model have shown that the AI detection model has attained 85% accuracy (Nijati et al., 2021). Similarly, the existing model has presented a deep CNN approach for diagnosing TB utilizing CXR images. It has employed a histogram matching method with CXR images (Verma et al., 2020) for improving detection and accuracy performance for TB detection. The results have shown that the respective mechanism has attained better accuracy and F1 scores (Ignatiu et al., 2023). Similarly, the existing system has presented automatic TB detection using an enhanced DL model with CXR images. For the process, the respective model has used the Shenzhen China (SC) (Yang et al., 2022) and Montgomery County (MC) datasets. The results have shown that the conventional model has attained better accuracy (Simi Margarat et al., 2022). Contrarily, the prevailing model has developed CAD for the classification of TB disease using an X-ray dataset. It has incorporated an Artificial Neural Network (ANN) with SVM for better classification and has attained 94.65 accuracy (Pathak et al., 2022). Concomitantly, the prevailing model has detected TB using LightTBNNet, a lightweight, efficient, and fast deep CNN (Sahlol et al., 2020) from the CXR images. It has been evaluated with two publicly available datasets. The results have shown 90.6% accuracy, a 90.7% F1 score, and a 96.1% ROC curve (Capellán-Martín et al., 2023).

Similarly, the conventional model has predicted a bacteriologic confirmation of Mycobacterium TB in children and infants. It has employed ML models for the development of prediction performance. It has used a new dataset, and the result has shown that the prevailing mechanism has attained better values in accuracy metrics (Smith et al., 2023). Contrarily, the suggested model has diagnosed TB from real-world cough recordings and has incorporated the conventional ML models for better prediction. It has utilized a huge dataset of TB and non-TB audio recordings (cough). The outcome has shown satisfactory results (Kafentzis et al., 2023). Contrastingly, the traditional model has assessed an enhancement of image in TB detection utilizing DL methods. It has evaluated image enhancement algorithms, namely, High-Frequency Emphasis Filtering (HEF), Unsharp Masking (UM), and Contrast Limited Adaptive Histogram Equalization (CLAHE). It has used a TB image dataset, and the results have revealed that the prevailing method has attained 89.92% accuracy (Munadi et al., 2020).

2.1 Problem identification

- Several existing studies have been focused on predicting TB and non-TB. However, limited research has focused on the ATB, healthy, LTBI, sick, and non-TB classifications (Kaur and Sharma, 2021).
- Accuracy is an important performance metric utilized to examine the model's performance. Nevertheless, existing models lack accuracy rates and prediction rates (Osamor and Okezie, 2021).
- Due to its primary cause, in TB prediction research, feature selection has been lacking in conventional research (Capellán-Martín et al., 2023).

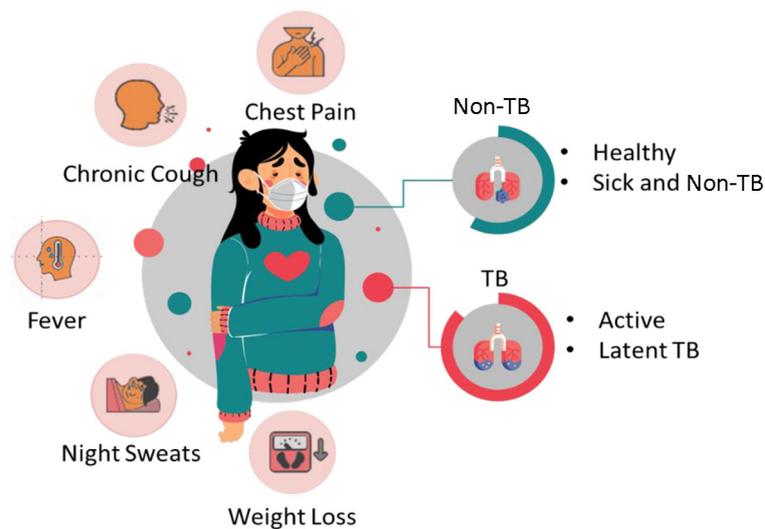


FIGURE 1
Symptoms of TB and non-TB.

3 Materials and methods

In the present world, the severity of TB disease is increasing in most countries. It affects the lives of massive numbers of people. To evade the severity of the disease and the consequences of TB, it is important to recognize the disease. Though the prevailing research is a time-consuming and expensive process, several studies are focused on TB prediction but lack detection rates, overfitting data, and accuracy. To improve the TB prediction, the respective model employed SFF in the YOLOv8 model to classify ATB, LTBI, and non-TB, sick and healthy, through the TBX-11k dataset. Congruently, it is necessary to recognize the major cause of disease for enhanced classification, as TB is the most hazardous lung disease that affects the respiratory system.

Tuberculosis is a contagious bacterial infection that can be caused by bacteria called *Mycobacterium tuberculosis* complex, which is considered one of the oldest diseases. It primarily affects the lungs of humans and is the leading cause of death worldwide. Generally, TB is spread through the air while the infected person sneezes or coughs. The bacteria can be transmitted from one person to another. Some of the common symptoms of TB encompass chest pain, fatigue, chronic cough, fever, night sweats, weight loss, and coughs with blood. Figure 1 depicts the symptoms of TB and non-TB.

Primarily, TB causes severe health problems if it is not treated or if you with a weak immune system. It leads to injury to the lungs and spreads to other parts of the organs like the spine, kidneys, and brain in some cases. The major reason for TB is *Mycobacterium tuberculosis*. It is defined as when a person inhales this bacteria, it can be transmitted to the lungs and multiply, which leads to infection. Moreover, the immune system responds by forming granulomas, which are known as small nodules that contain infection. In several cases, bacteria

remain dormant for years without any symptoms. Conversely, the immune system gets weak, and infection causes illness and becomes active. Though it has severity, TB is a preventable and treatable disease. Early prediction, appropriate treatment, and medication adherence are significant in preventing and managing TB transmission. To overcome the existing research problems, the proposed model employed SFF in the YOLOv8 model to classify ATB, LTBI, and non-TB, sick and healthy, through the TBX-11k dataset. Figure 2 depicts the illustrative representation of the dataset creation.

The proposed mechanism utilizes the TBX-11k dataset to predict TB through the X-ray images. It uses the advantages and incorporates SFF in YOLOv8 to enhance the prediction performance of TB. To calculate the efficiency of the proposed mechanism, performance metrics like precision, F1-score, recall, and accuracy are utilized in projected research. The proposed TB and non-TB prediction mechanisms are depicted in Figure 3.

Figure 3 signifies the present model, which is comprised of several stages, including the selection of datasets, the image label-pre-processing method, training and testing split, and prediction with SFF in the YOLOv8 model. The detailed description of each stage in the proposed system is discussed below:

3.1 Dataset selection

The proposed system uses the TBX-11k dataset from the Kaggle website for the prediction of TB and non-TB in X-ray images, which is publicly available. For assessing the proposed mechanism, the TBX-11k dataset is used in this model, as outlined in Figure 3. Furthermore, the dataset includes the data of both TB and non-TB, and it comprises 11,200 X-ray image samples. There are five divisions in this dataset: Sick but non-TB, Healthy, Active TB,

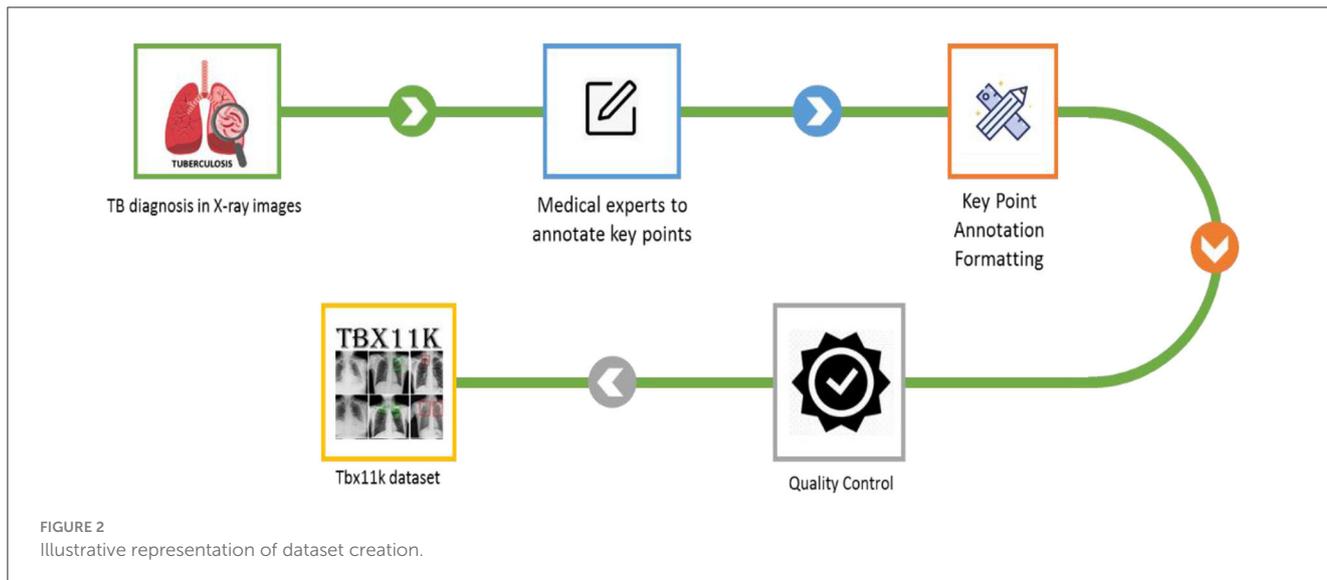


FIGURE 2 Illustrative representation of dataset creation.

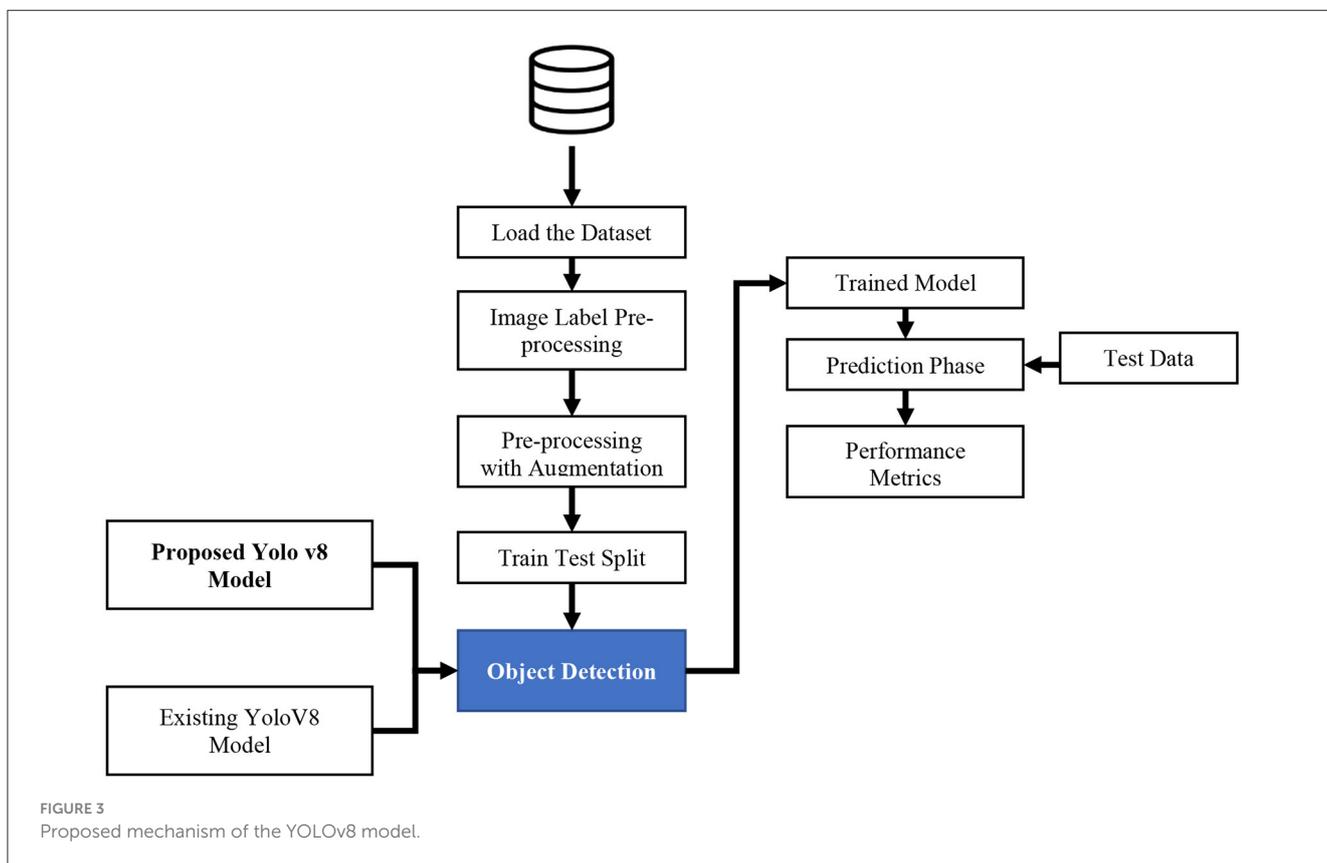


FIGURE 3 Proposed mechanism of the YOLOv8 model.

Latent TB, and uncertain TB. It can be split into a training set, a validation set, and a testing set. The website line of the Kaggle TBX-11k dataset is given below:

<https://www.kaggle.com/datasets/usmanshams/tbx-11/data>.

3.2 Image label pre-processing

The image label pre-processing process refers to the preparation of image data for classification or labeling tasks. Typically, these

steps involved normalizing and transforming the images to confirm that they were suitable for further processing. Image label pre-processing techniques include normalization, resizing, cropping, color-space conversion, and augmentation. Moreover, the pre-processing techniques support improving the consistency and efficiency of the image data and make it more suitable for the labeling and classification processes. In the projected research, the image label pre-processing method is used to convert the features of the dataset to a common scale, enhancing the accuracy and performance of the prediction.

TABLE 1 Hyperparameters training.

Hyperparameters	Value
Epochs	100
Batch size	16
Image size	640
Optimizer	Auto
Momentum	0.937

3.3 Pre-processing with augmentation

The process of pre-processing with augmentation is denoted as a set of techniques that were utilized to modify and enhance the data before it was used for the training process. Augmentation methods are implemented to improve the quantity and diversity of training data that enhance the generalization ability and performance of the model. Moreover, the process involved relating different modifications and transformations to existing data and creating augmented samples. Transformations such as scaling, rotating, cropping, adding noise, flipping, and changing the contrast and brightness of images.

3.4 Data splitting

In the YOLOv8 model, data splitting is the process of dividing the dataset into discrete subsets for processes such as training, validation, and testing. It is a common activity in ML for evaluating the model's performance to prevent overfitting and unseen data. The training set is utilized to train the YOLOv8 model, and the validation set is used to tune the hyperparameters of the model and observe its performance. The testing set is to calculate the final performance of the trained model on unseen data. Table 1 depicts the hyperparameters and their values for training.

3.5 Object detection

The YOLOv8 model is defined as an object detection algorithm. It aims to detect and classify the objects accurately. It is referred to as an improvement model on previous versions of the YOLO model that offers better accuracy and performance. Moreover, the model evaluation is based on mAP, which is considered one of the most utilized evaluation metrics for object detection. The mAP took the average precision (AP) on classes and computed them at the pre-mentioned IoU threshold. The proposed model assesses the generalization and robustness capabilities of models by mAP scores that are calculated among the diverse test scenarios, highlighting the significance of the YOLOv8 model in TB and non-TB detection.

3.5.1 Conventional YOLOv8 model

YOLOv8 is a cutting-edge and advanced model that provides higher detection speed and accuracy. The YOLOv8 is known as a real-time object detection mechanism that utilizes a single Neural

Network (NN) to predict class probabilities and bounding boxes directly from entire images in a single evaluation. Significantly, it is based on Darknet architecture and uses layers of convolutional layers along with skip connections to enhance speed and accuracy. The YOLOv8 model incorporates different improvements from previous versions, like advanced data augmentation techniques and feature pyramid networks, to attain better performance in the task of object detection. Figure 4 illustrates the architecture of the existing YOLOv8 model.

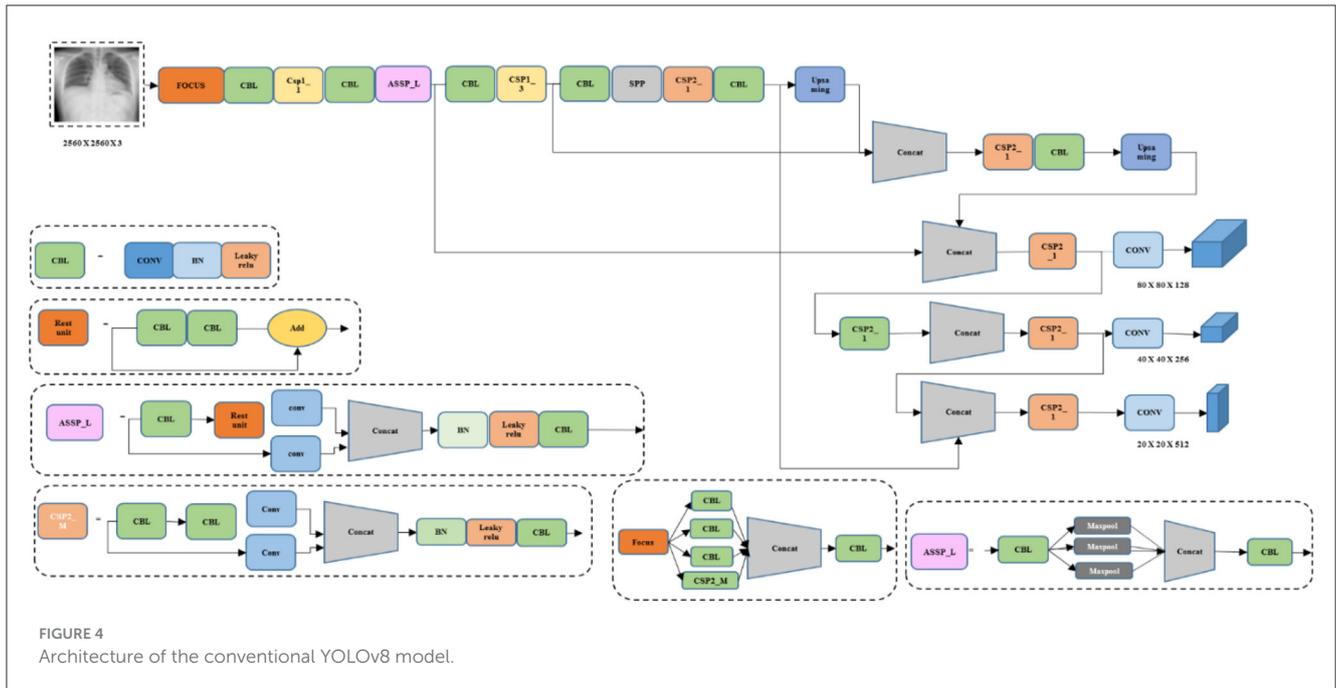
Figure 4 deliberates the architecture of conventional YOLOv8, a DL model. It is capable of processing the inputs in different forms, such as real-time camera images, videos, and images. Moreover, it comprises convolutional layers (conv), concatenation layers (concat), and three max pooling layers. Though it performed with satisfactory results, it lacked accuracy and speed. Thus, the proposed model incorporates several techniques to improve the performance of the proposed model.

3.5.2 Proposed YOLOv8 model

The YOLOv8 model is known as a popular object detection algorithm that is utilized in system vision tasks encompassing TB detection. Primarily, it is designed for object detection; thus, the proposed research is adopted for predicting TB. This model is known for its fast inference speed, which allows for the processing of the image data rapidly and is particularly useful in the diagnosis process. The YOLOv8 model has the ability to detect multiple objects in a single pass. It is beneficial in cases where TB abnormalities or lesions are presented in various areas of the lungs or other affected organs. In addition, the model provides relatively high accuracy in object-detecting tasks, which has been trained on huge datasets, and uses Deep Neural Networks (DNN) to identify TB. Though the YOLOv8 model performs better, it struggles with small objects and finds an inability to perform fine-grained classifications. To evade this problem, the proposed research incorporates the SFF technique to improve the detection performance and decrease the small object missed detection rates. Figure 5 illustrates the architecture of the YOLOv8 model.

Figure 5 signifies the architecture of the proposed YOLOv8 model, which incorporates improvements for high detection accuracy while maintaining high efficiency and speed. Several key modifications include the Cross-Stage Partial Bottleneck (C2f module) and the detection head, including independent branches, activation functions, and loss functions. The C2f module is incorporated to merge high-level features with contextual data effectively. It is achieved by concatenating the bottle net block output, which contains two 3×3 convolutions along with residual associations. The YOLOv8 model adopted the detection head for eliminating the requirement for pre-defined anchor boxes and predict the object centers directly. Figure 6 shows the schematic diagram of the C2f module.

Figure 6 illustrates the schematic diagram of the C2f module and the bottleneck of the proposed model. The C2f module is utilized to improve quality after feature fusion. It is known as an improvement of the original C3 module, which is considered a benefit of the YOLOv8 model with high gradient information. It decreases one convolutional layer and makes use of the



bottleneck module to extend the gradient branch to acquire high gradient flow data when ensuring lightweight. The main idea of the given structure is to improve time of multi-scale fusion and the probability of attaining a higher detection accuracy rate. Since the model is lightweight, B3-N3 and B4-N4 are included, and one unit is used. The respective process can be explained below.

$$Net_5^{out} = C2f(Concat(Conv(Net_5^{in}), B_5^{out}), n)$$

$$Net_{in}^{out} = C2f(Concat(Conv(Net_i^{in}), B_{in}^{out}, P_{in}^{out}), n)$$

where Conv and C2f are represented as conforming module operations. Net, B, and P corresponded to backbone feature maps, as PAN and FPN, respectively; n denotes the number of C2f uses. On the other hand, the attention mechanism called Decoupled Fully Connected (DFC) avoids the prevailing attention algorithms for feature capture at long distances and computational complexity. Its execution is deliberated in Equations (1) and (2).

$$Y' = X * Feat_{1*1} \tag{1}$$

$$Y = Concat([Y', Y' * Feat_{dep}]) \tag{2}$$

where $Feat_{dep}$ is represented as a depth-wise convolution. The depth-wise convolution acts upon feature maps to attain a linear transformation process. In addition, the DFC attention mechanism is used directly as a deeply separable structure, with the simplest structure being to acquire an attention map along with information. The

particular measuring process is expressed in the following Equations (3) and (4):

$$\alpha'_{hw} = \sum_{h'=1}^H Feat_{h,h',w}^H \odot X_{h',w} \quad h = 1, 2, \dots, H, w = 1, 2, \dots, W \tag{3}$$

$$\alpha_{hw} = \sum_{w'=1}^W Feat_{w,h,w'}^W \odot \alpha'_{hw'} \quad h = 1, 2, \dots, H, w = 1, 2, \dots, W \tag{4}$$

where Feat refers to the depth-wise separable convolution process, which is classified into vertical and horizontal directions; α' is an attention map in the vertical direction and α is an attention map on α' in the horizontal direction. Therefore, feature vectors are transformed linearly to estimate three matrices, namely V, K, and Q. The calculation process is expressed in Equations (5)–(7).

$$Q = X^{rev} Wid^Q \tag{5}$$

$$K = X^{rev} Wid^{rev} \tag{6}$$

$$V = X' Wid^V \tag{7}$$

Then attention to relation from place to place is obtained through structuring a directed graph to localize its related region. The particular implementation is as follows: V and Q for each place are processed by region to obtain the levels Q^r and K^r . After that, Q^r and K^r dot products are estimated to obtain the adjacency matrix

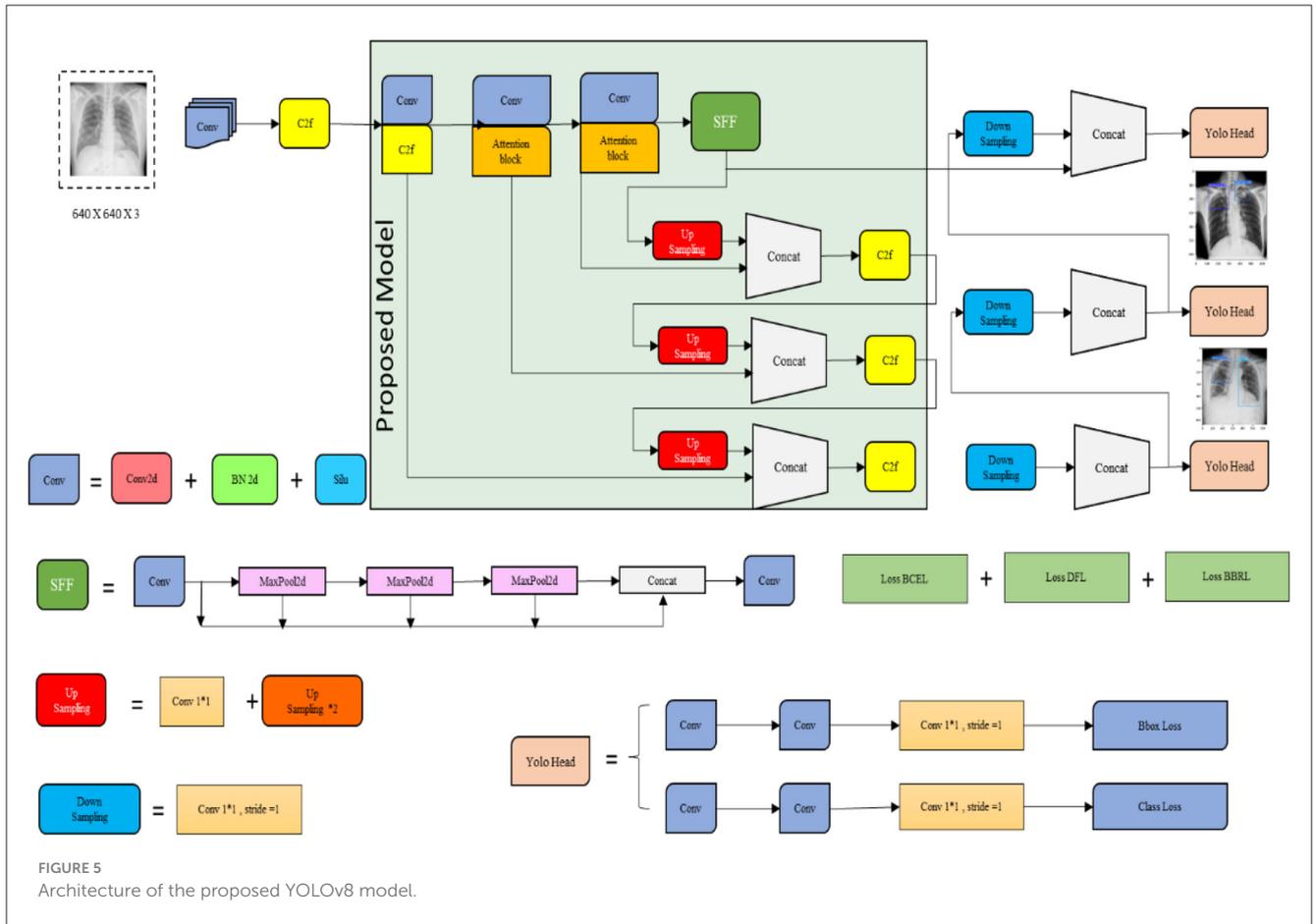


FIGURE 5 Architecture of the proposed YOLOv8 model.

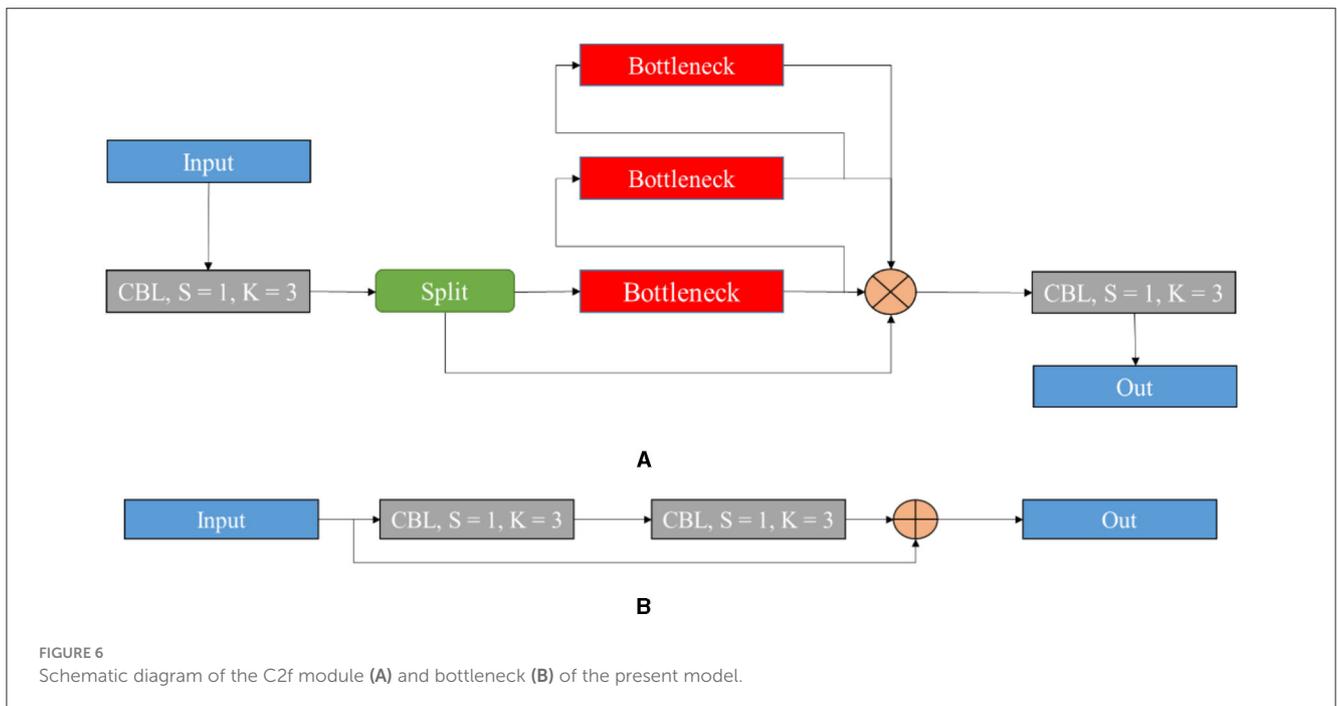


FIGURE 6 Schematic diagram of the C2f module (A) and bottleneck (B) of the present model.

and Adj^{rev} is used to calculate the inter-place correlation, and the equation is

$$Adj^{rev} = Q^{rev} (K^{rev})^T$$

Further, Adj^{rev} is clipped and the minimum relevant token in Adj^{rev} is filtered out at a coarse-grained level. The top k relevant regions in Adj^{rev} have been recollected to acquire a matrix of routing index, I^{rev} which is deliberated in Equation (8).

$$I^{rev} = \text{topkIndex}(A^{rev}) \tag{8}$$

Subsequently, at a fine-grained level, token-to-token attention is utilized. This attention is focused on k routing regions that are indexed and collecting all the V and K tensors in k regions to obtain V^{foc} and K^{foc} that are calculated through Equations (9) and (10).

$$K^{foc} = \text{focal}(K, I^{rev}) \tag{9}$$

$$V^{foc} = \text{focal}(V, I^{rev}) \tag{10}$$

The collected V^{foc} and K^{foc} are processed along with the attention. A Local Context Enhancement (LCE) term is employed to obtain the output tensor. The formula is deliberated in Equation (11).

$$\text{Out} = \text{Attention}(Q, K^{foc}, V^{foc}) + \text{LCE}(V) \tag{11}$$

For the size of the input feature $h \times w \times c$, the required SFFs are utilizing the size of regular convolution $k \times k$ in Equation (12). C represents the number of channels of input data.

$$\text{SFFs}_{\text{Conv}} = h \times w \times k^2 \times c^2 \tag{12}$$

A deep convolutional kernel makes the sliding operations over the input channel space to estimate the output channel features. The SFFs for deep convolution are estimated as the output channel features that are calculated through Equation (13).

$$\text{SFFs}_{\text{DWConv}} = h \times w \times k^2 \times c \tag{13}$$

Correspondingly, deep convolution requires other computational costs and point-by-point convolution to compensate for decreasing accuracy after the convolution function. It introduces Partial Convolution (PConv), which utilizes regular convolution to process the operation on continuous features in the input channel. c_p is the channel number in input features. The formula for measuring SFFs of pConv is indicated in Equation (14).

$$\text{SFF}_{\text{sPConv}} = h \times w \times k^2 \times c_p^2 \tag{14}$$

c_p , $\frac{1}{4}$ of a number of feature channels (input) c , the SFFs of pConv are $\frac{1}{16}$ of the conventional convolution. This decreases the number of memory accesses when minimizing the parameters. Firstly, pConv has been utilized to swap the two 1×1 convolutional layers in the SFF block, which enhances the receptive field when making the original module more efficient and faster. Secondly, residual concatenation has been added to the last two convolutional layers in the block to develop the output information features and decrease the effective feature loss. It can optimize the present model's performance. Due to the anchor-free idea usage, the YOLOv8 loss function is greatly different from the previous YOLOv5 series. The optimization was comprised of two parts, namely regression and classification. The regression part utilizes the Boundary Box Regression Loss (BBRL) and Distribution Focal Loss (DFL), and the classification loss utilizes Binary Cross Entropy Loss (BCEL). The loss function is expressed in the following Equation (15):

$$\text{func}_{\text{loss}} = \lambda_1 \text{func}_{\text{BCEL}} + \lambda_2 \text{func}_{\text{DFL}} + \lambda_3 \text{func}_{\text{BBRL}} + \lambda_3 \text{func}_{\text{Out}} \tag{15}$$

The loss of prediction set is necessarily the cross-entropy loss, and the calculation can be deliberated in Equation (16).

$$\text{func}_{\text{BCEL}} = \text{weg}[\text{class}](-x[\text{cls}] + \log(\sum_j \exp(x[j]))) \tag{16}$$

where cls is a number of categories, $\text{weg}[\text{class}]$ refers to each class weight, and x refers to the probability value after the sigmoid activation. DFL is an optimization of the focal loss function that generalizes discrete outcomes of classification into unremitting outcomes through integration. Hence, the calculation is expressed in the following Equation (17).

$$\text{func}_{\text{DFL}}(S_i, S_{i+1}) = -((y_{i+1} - y) \log(S_i) + (y - y_i) \log(S_{i+1})) \tag{17}$$

where y_i and y_{i+1} represent values from the left and right sides that are near the consecutive labels y . Simultaneously, while the prediction box contains a high degree of coincidence with the target box, the loss function made the model obtain better generalization capability with less training intervention through weakening the geometric factors penalty. A 2-layer attention mechanism and active non-monotonic FM apparatus are utilized. The expression is provided in Equations (18) and (19).

$$\text{func}_{\text{BBRL}} = (1 - \frac{\text{Wid}_i \text{Heg}_i}{S_u}) \exp(\frac{(x_p - x_{gt})^2 + (y_p - y_{gt})^2}{(\text{Wid}_g^2 + \text{Heg}_g^2)}) \gamma \tag{18}$$

$$\gamma = \frac{\beta}{\delta \alpha^{\beta - \delta}} \tag{19}$$

where β refers to the abnormality degree of the predicted box. The small degree of abnormality denotes that the anchor box

quality is higher. Henceforth, β is used to build the non-monotonic focal number assigned small gradient gains for predicting large outliers' boxes. α and δ are referred to as hyperparameters; x_p and y_p denote coordinate values of the prediction box, x_{gt} and Y_{gt} represent ground truth coordinate values. The corresponding wid and heg values denote the width and height of the two boxes, respectively. The process can be expressed as

$$S_u = wid_{heg} + wid_{gt}heg_{gt} - Wid_iHeg_i.$$

3.5.3 SFF

The Selection Focal Fusion (SFF) block is utilized in the proposed research. It is a technique utilized in the YOLOv8 model that is popular in object detection algorithms. It improves the accuracy of detection by selectively fusing features from three layers of the network. In the YOLOv8 model, a network architecture consists of multiple detection layers at various scales. Each of the detection layers is accountable for detecting objects of various sizes. SFF is applied to three detection layers and employs an attention mechanism for featuring the image data to improve their performance. Furthermore, focal aggregation classifies the images at both horizontal and vertical levels and provides the attention mechanism that is presented for the optimization of the backbone network that enhances the attention of the model to critical information. The proposed SFF substantially decreases the missed detection rates of small objects and improves the detection performance of the present model.

3.5.4 Prediction phase

The prediction phase is a commonly used technique by researchers to determine the efficiency of the proposed system. In the prediction phase, the algorithm is processed using the test data, which will reveal the performance of the present model. Finally, the system's efficiency is calculated using certain performance matrices, such as mAP, F1-score, precision, and recall, to evaluate the efficiency of the projected model.

4 Results

4.1 EDA

Exploratory Data Analysis (EDA) supports examining the process of data comprehension in detail and aids in learning different characteristics of data. It helps in representing the data statistically. Various visual representations are deliberated for the identification of patterns and styles, such as heat maps, histograms, box plots, bar charts, and scatter plots. Furthermore, EDA is significantly utilized for outliers, anomalies, error detection, and other suspicious patterns or assumptions in the data. Therefore, it can be utilized for data comprehension before creating assumptions. [Figure 7](#) shows the sample data used in the proposed research, which consists of lungs.

[Figures 7, 8](#) deliberate the sets of training and validation of image data from the dataset. From [Figure 8](#), the validation set of

images shows the level of TB in the lungs; [Figure 9](#) signifies the information about manually labeling objects in the dataset. Here, the red box indicates active TB, salmon pink indicates healthy, the yellow box indicates sick, and the orange box indicates latent TB.

[Figure 9](#) shows the analysis of the dataset, which is surmised. The dataset is comprised of a huge number of small objects that can exist in an uneven and dense distribution. The first subfigure illustrates the number of objects of each type in the dataset. The second subfigure shows the distribution of center point coordinates of objects in bounding boxes in the dataset. The third subfigure represents a scatter plot of the corresponding height and width of the object in the bounding box. Moreover, [Figure 10](#) presents the correlogram label.

[Figure 10](#) deliberates the correlogram label. A correlogram is known as a graphical representation that shows correlation coefficients among variables in the dataset. Moreover, the label in the correlogram denotes features or variables that are being examined. The YOLOv8 model endured training on the TBX-11k dataset for its challenging X-ray samples and limited data. [Figure 9](#) shows the xywh of the dataset, where x and y are positions and h and w refer to height and width. The estimation has encompassed training and validation sets, which reveal the classification and box prediction loss on the validation set. [Figure 11](#) represents the learning curves of box classification and prediction loss for training and validating the dataset.

[Figure 11](#) depicts the learning curves of box classification and prediction loss for training and validating the dataset. To address the overfitting, the annotated data aimed to augment the size of the samples for improved generalization. During the training, it is important to notice that optimal weights are saved and the checkpoint weights are retained for deployment. Despite this, [Figure 11](#) shows the better performance of the proposed model, attaining a prediction loss below 0.06 and a classification loss below 0.1 on the validation set, displaying its proficiency in localizing and identifying the lesions in lung images.

4.2 Performance metrics

Performance metrics are primarily used for observing the efficiency of the projected research by utilizing various metrics like recall rate, precision, mAP, and F1-score value.

4.2.1 Recall

The term recall is deliberated as the reclusive of the production metric that estimates the total of accurate positive categories made out of all the optimistic categories. It is calculated with the following equation:

$$\text{Recall} = \frac{\text{True_Pos}}{\text{False_Neg} + \text{True_Pos}}$$

4.2.2 Precision

The term precision is denoted as the covariance unit of technique, which results from the appropriately recognized

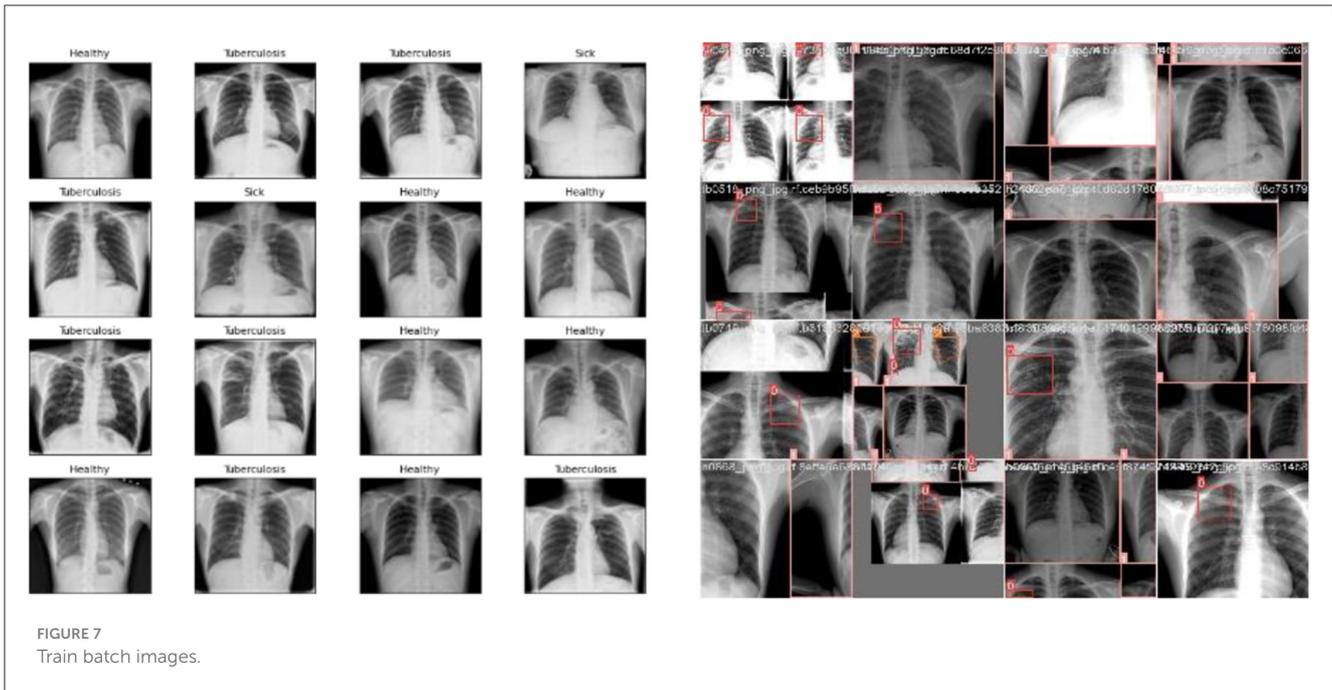


FIGURE 7
Train batch images.

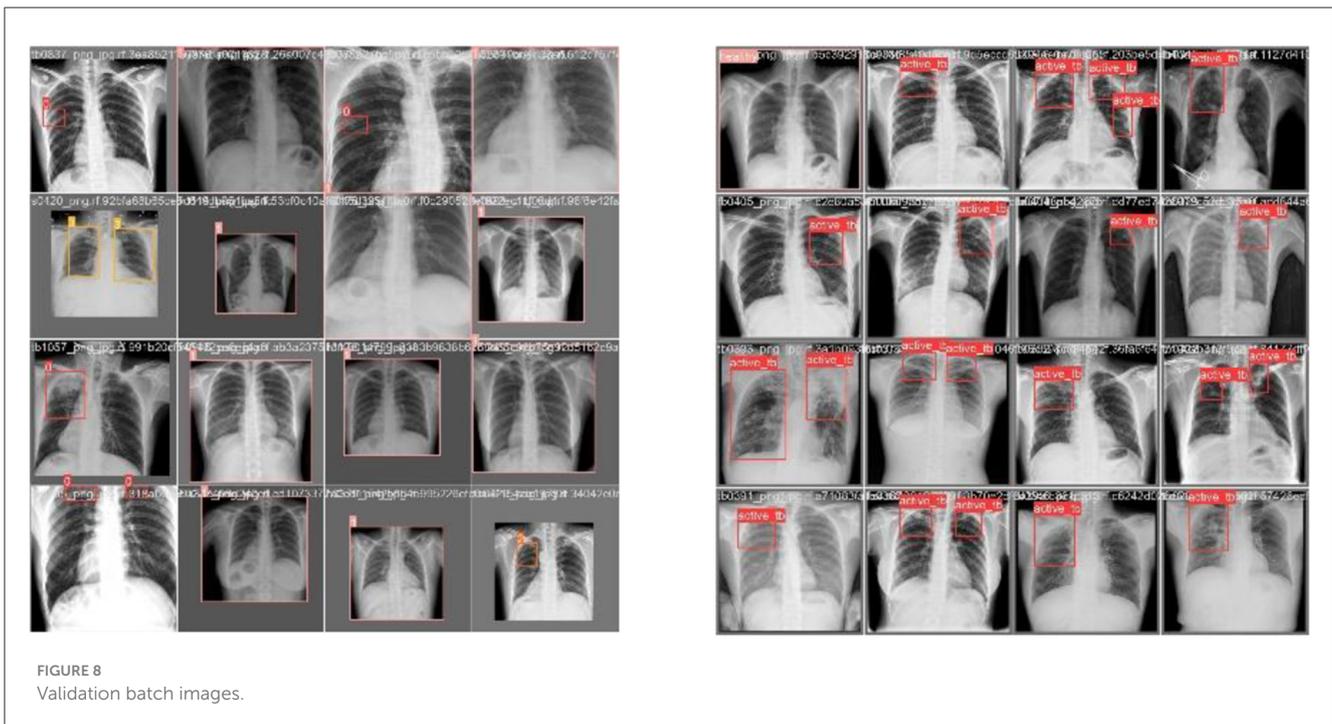


FIGURE 8
Validation batch images.

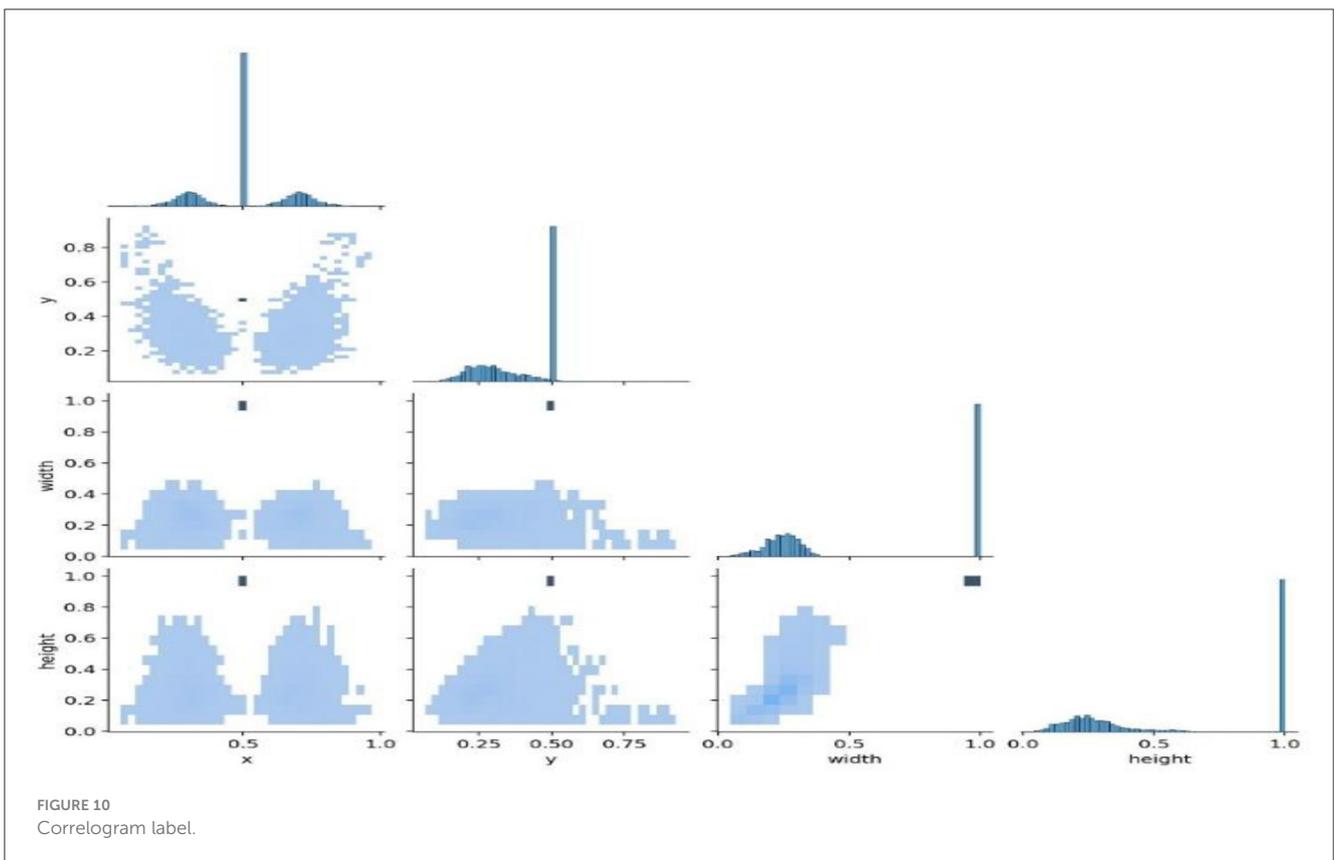
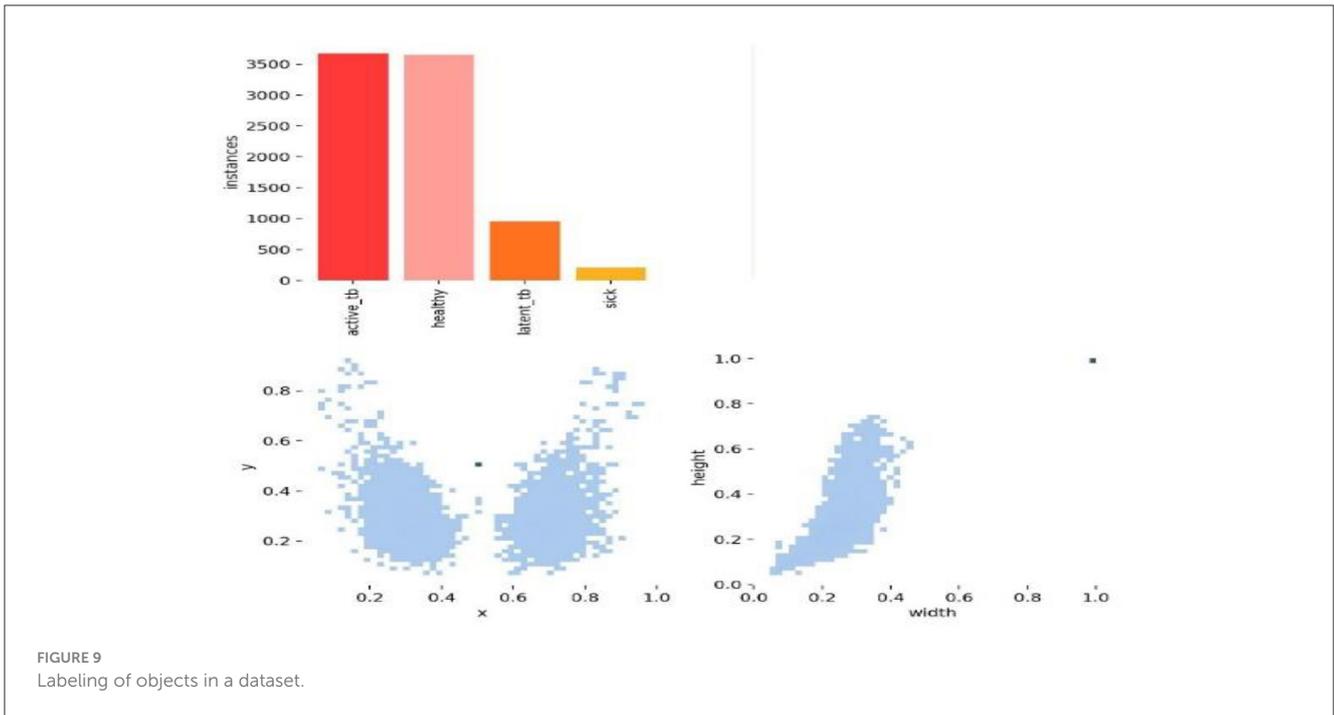
cases (*Truepos*) to the total group of cases that are accurately categorized (*Truepos + Falsepos*). It includes the repeatability and reproducibility of the capitals. Equation (20) depicts the formula for precision.

$$\text{Precision} = \frac{\text{True_Pos}}{\text{False_Pos} + \text{True_Pos}} \quad (20)$$

4.2.3 F1-score

The F1-score is represented as a measure of the appropriate mean rate of recall and precision value. The mathematical equation for the F1-score is depicted in Equation (21).

$$\text{F1 - score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (21)$$



4.2.4 mAP

Mean Average Precision (mAP) is computed utilizing Equation (2). It is a broadly accepted performance metric

for models of object detection. The mAP is evaluated by considering the mean of AP for each of the classes. AP for each class k is determined by measuring the area

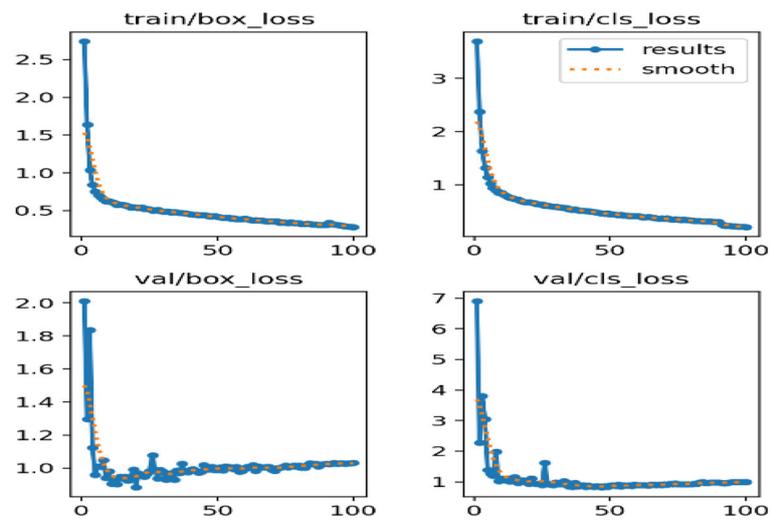


FIGURE 11 Curves of box classification and prediction loss for validating and training sets.

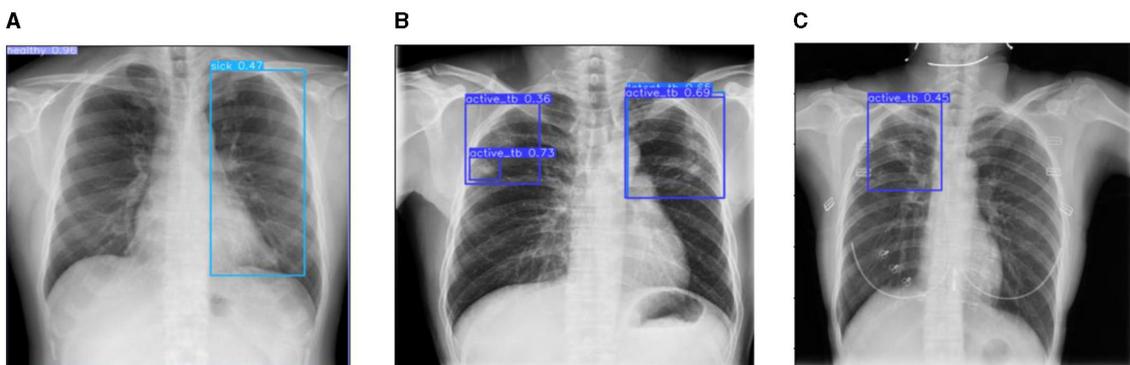


FIGURE 12 Prediction results of existing YOLOv8 model. (A) X-ray with both healthy and sick regions. (B, C) X-ray with active TB and latent TB.

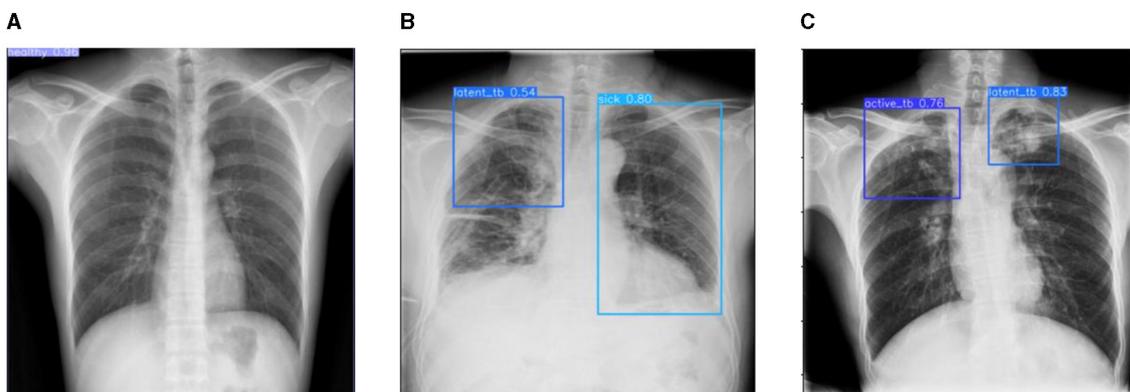


FIGURE 13 Prediction results of the proposed YOLOv8 model. (A) X-ray with healthy lung. (B) X-ray with sick region. (C) X-ray with latent TB.

TABLE 2 Validating results of proposed system.

Class	Images	Instances	Box (P)	R	mAP50	mAP50-95
All	357	449	0.711	0.622	0.657	0.475
Active_TB	357	207	0.773	0.56	0.676	0.32
Healthy	357	182	1	0.995	0.995	0.994
Latent_TB	357	41	0.176	0.146	0.107	0.0567
Sick	357	19	0.895	0.789	0.849	0.529

under the precision-recall curve. mAP offers a single score that considers precision, recall, and IoU, avoiding bias in performance.

$$mAP = \frac{1}{n} \sum_{k=1}^{k=n} AP_k$$

4.3 Experimental results

This section discusses the results accomplished by the proposed mechanism for predicting Active TB, Sick but non-TB, Healthy, and Latent TB with the Kaggle TBX-11k dataset. Furthermore, the outcomes obtained in the internal comparison of the present research YOLOv8 model and existing data. Figures 12, 13 signify various illustrations of lungs and the prediction results of the proposed research.

Figure 12 represents the existing sample image data from the TBX-11k dataset. After the implementation, the prediction results are shown in Figure 13. It predicted active TB, latent TB, sick, and healthy among the X-ray images. Here, the lavender box indicates healthy, the light blue box indicates sick, the violet box indicates active TB, and the dark blue box indicates Latent TB in Figure 12. The results for both healthy and sick were 0.47%, which is less than the proposed model. However, Figure 13A shows a healthy lung, and (b) shows a sick lung with a 0.80% better prediction. The latent_TB in Figure 12 predicts 0.65% whereas, whereas latent_TB in Figure 13C results in 0.83% of. Likewise, the active_TB results in 0.36%, 0.73%, 0.69%, and 0.45% in Figure 12, whereas the proposed model in Figure 13 results in 0.76%, which shows a higher prediction than the existing model. While validating, the results are attained at 0.2 ms of speed, 2.0 ms of inference of pre-process, 0.0 ms of loss, and 1.1 ms of post-process per image. Furthermore, Table 2 depicts the validating results of the proposed system.

Table 2 illustrates the proposed system for validating results. It highlights the overall prediction results that comprised 449 instances, 0.711 of precision, 0.622 of recall, 0.657 of mAP50, and 0.475 of mAP50-95. Additionally, it shows results for predictions such as active TB, Latent TB, healthy, and sick.

4.4 Comparative analysis

The section illustrates the comparative analysis of the proposed mechanism with the existing approaches depending on the

TABLE 3 Comparative analysis of proposed research (Bista et al., 2023).

Model		All classes (mAP@0.5)
YOLOv7	Base model with class imbalance	0.249
	Base model with image weights	0.211
	Base model along with minority class image augmentation	0.280
	Putting it all together and evolving the hyperparameter	0.587
YOLOv8	Base model	0.627
	Selective Focal Fusion (SFF) Block in YOLOv8	0.657

TABLE 4 Comparative outcomes of proposed system.

All classes at mAP	Proposed model	Existing model
F1 confidence	0.66	0.63
Precision confidence	1	0.96
Recall confidence	0.76	0.71
Precision-recall confidence	0.65	0.62

performance metrics. Table 3 deliberates the comparative analysis of YOLOv8 with the YOLOv7 model at mAP values.

From Table 3, the base model of YOLOv8 has attained 0.627 mAP values, whereas the proposed model of SFF in YOLOv8 has attained 0.657 mAP values. It is proven that the present model attains 0.37 higher than the base model, which represents its efficiency. Additionally, Table 4 discusses the comparative outcomes of the proposed system.

In Table 4, the comparative outcomes of the existing and proposed models are presented. It shows the proposed research has attained better outcomes than the prevailing model in all classes at mAP values. It attained 0.03 of F1 confidence, 0.04 of precision confidence, 0.05 of recall confidence, and 0.03 of precision-recall confidence more than the prevailing model. Additionally, Figure 14 deliberates the performance metrics comparison.

Figure 14 illustrates that the proposed model attained better values than the prevailing model in all classes at mAP values. It attained 0.66 of F1 confidence, 1 of precision confidence, 0.76 of Recall confidence, and 0.65 of precision-recall confidence, which is more than the prevailing model values.

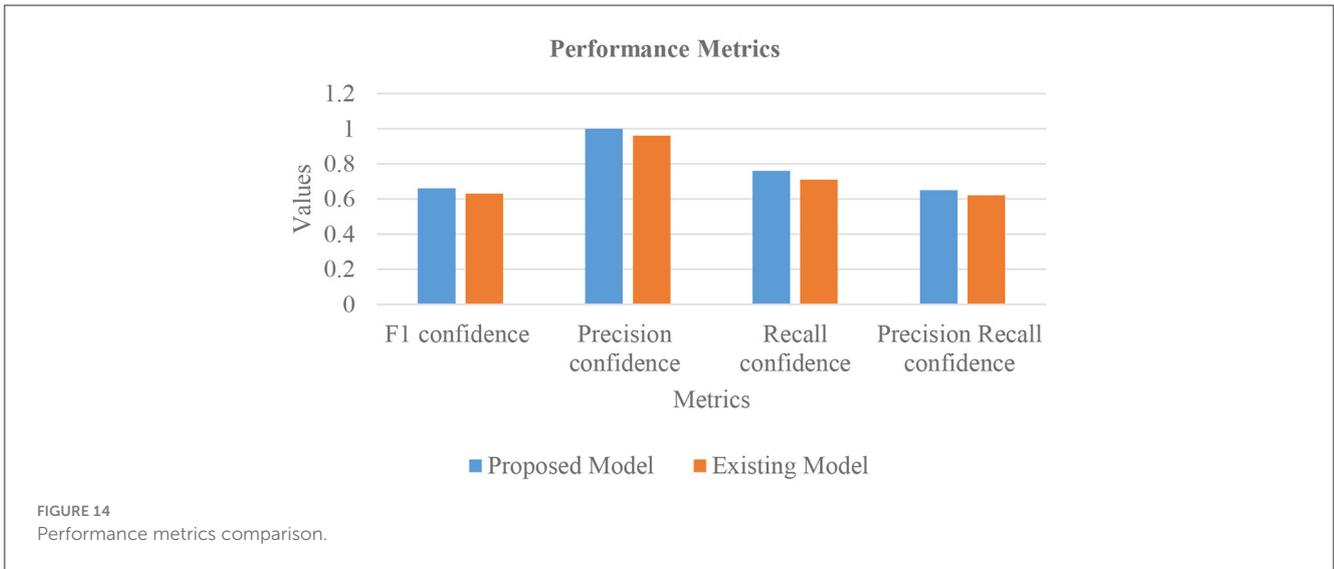


FIGURE 14 Performance metrics comparison.

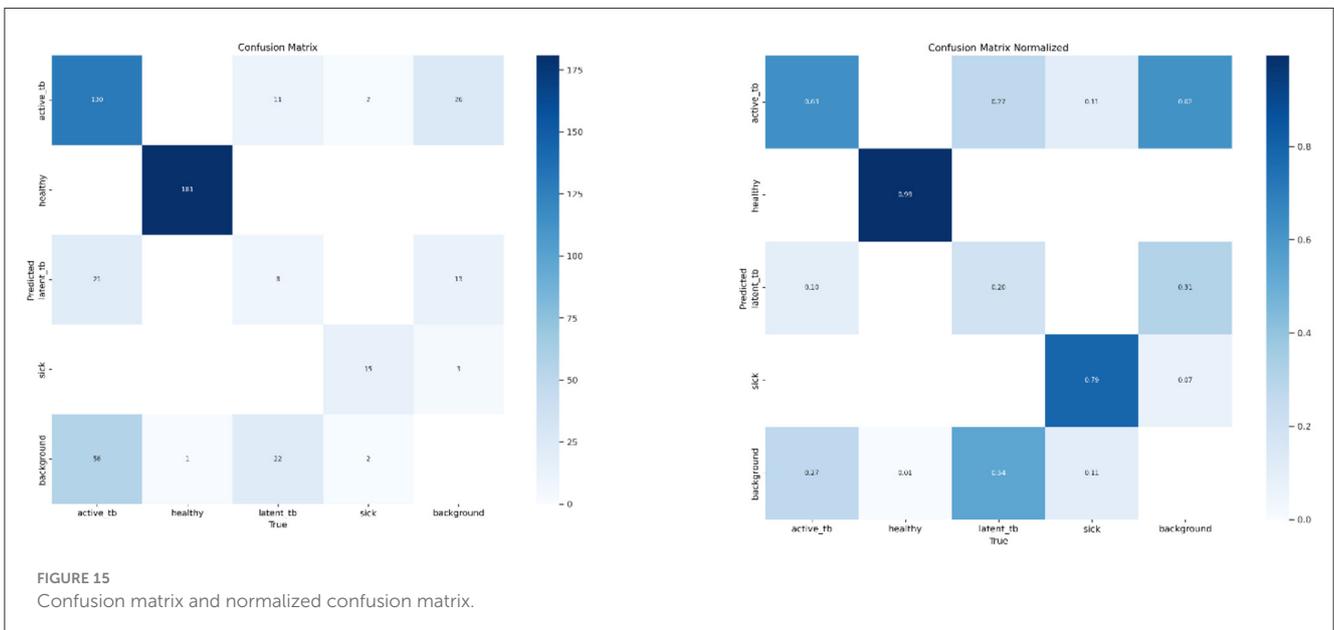


FIGURE 15 Confusion matrix and normalized confusion matrix.

4.5 Performance analysis

The performance of the proposed algorithm is examined using evaluation metrics like Recall, Precision, F1-score, and accuracy. Likewise, a Confusion Matrix (CM) is utilized for identifying the performance of the proposed research. It encapsulates and envisages the performance of the classification algorithm. Hence, the CM signifies how many predictions are right and wrong as per the class. Figure 8 deliberates the CM of the respective system.

Figure 15 shows the CM of the proposed research. It represents the actual and correct prediction of the model. Figures 16, 17 show the F1 confidence curve and Precision confidence curve, respectively. Correspondingly, the confidence curves are associated with performance metrics that illustrate the performance of the proposed mechanism at various confidence levels.

The F1-score is a measurement of the accuracy of the proposed model in classification tasks, taking into account both recall and

precision. It combines these two metrics into a single value to obtain a balanced evaluation. Figure 17 deliberates the precision confidence curve; which focuses on positive prediction proportions made by the proposed research. The *p*-confidence curve shows the relationship between the precision value and the threshold value. Similarly, Figure 18 refers to the precision-recall curve, and Figure 19 exemplifies the recall-confidence curve.

Figure 18 deliberates the trade-off between recall and precision. Precision refers to the proportions of correctly classified positive instances out of all instances that are classified as positive. Recall refers to the correctly classified positive instances proportions out of the entire actual positive instances. The PR curve shows how the precision and recall points change as the classification threshold value varies. Similarly, in Figure 19, the recall-confidence curve denotes the association between the confidence threshold and the recall of the proposed research. They both help in assessing the ability of the proposed research to identify the positive instances

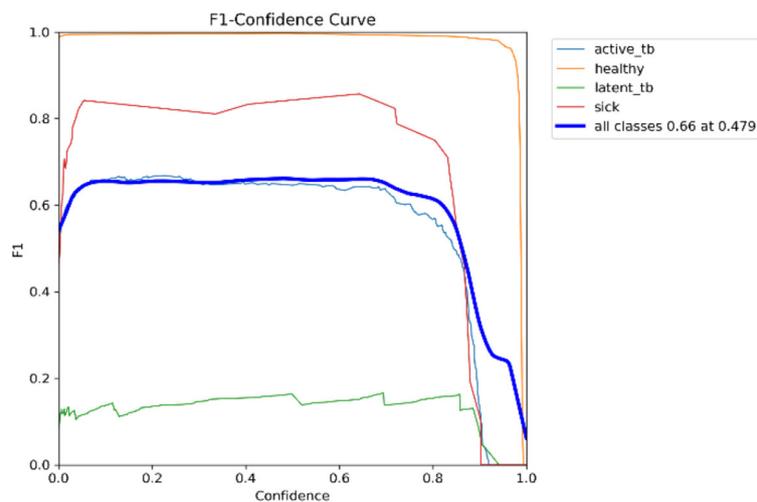


FIGURE 16
F1-curve.

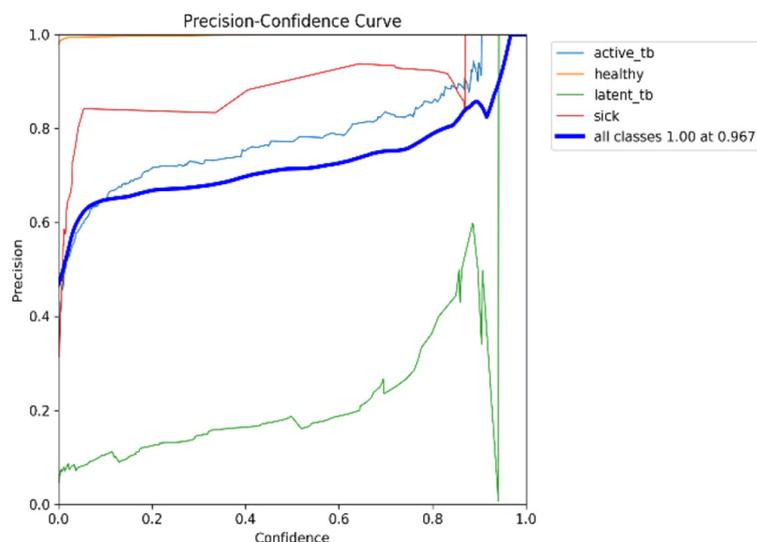


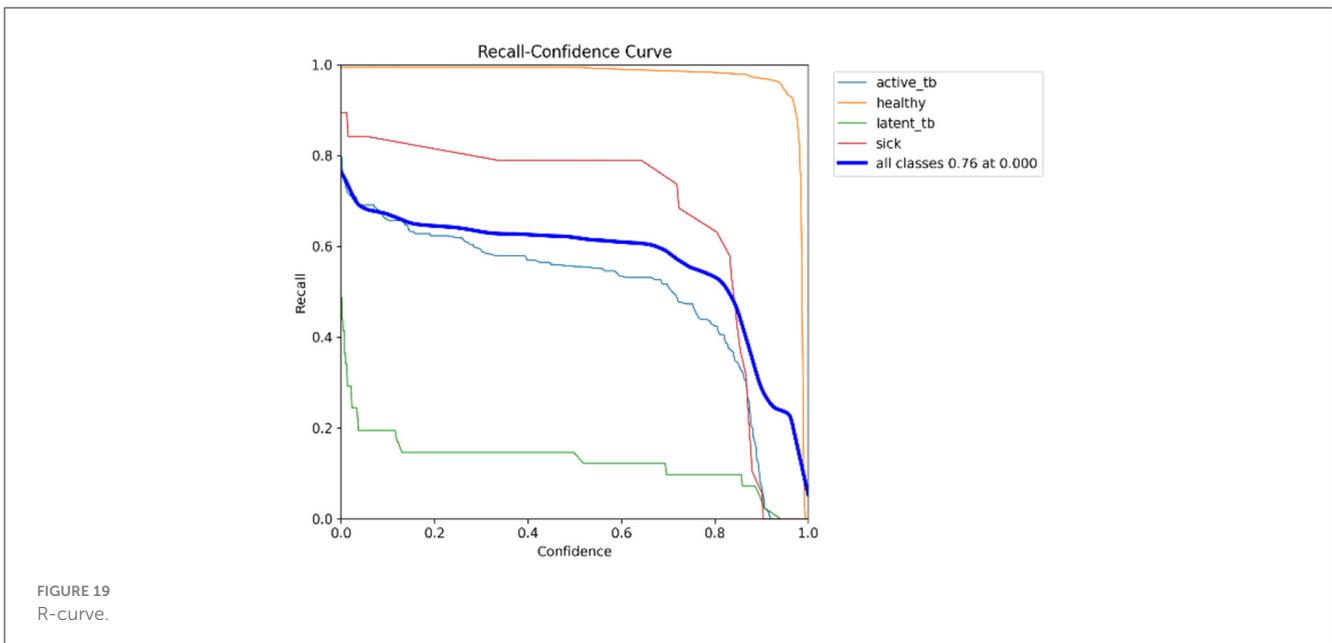
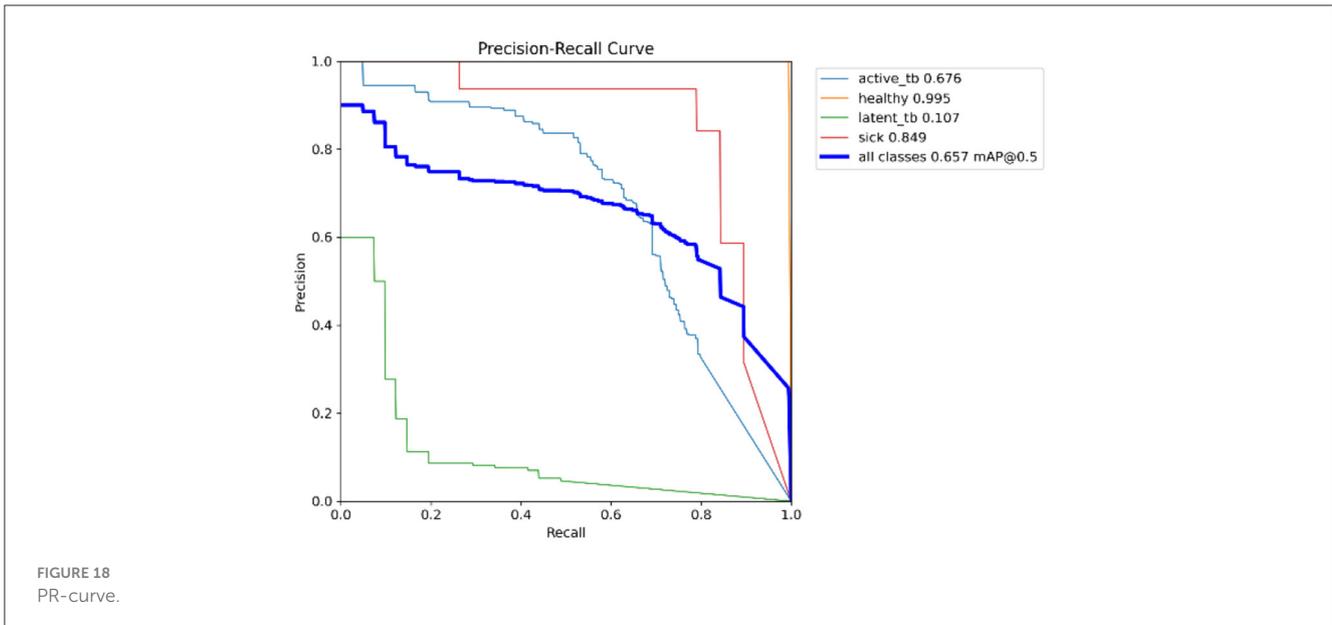
FIGURE 17
P-curve.

correctly and provide a precise understanding of its confidence threshold performance.

5 Discussion

The contributions of the respective research and the utilized dataset are indicated in this section. Various types of research are focused on the prediction of TB through its CXR images (Ekins and Freundlich, 2011; Acharya et al., 2022), cough, and different tests like blood tests, sputum tests, imaging studies, skin tests (Maipan-uku et al., 2024), etc. In order to decrease the severity of the disease, it is important to diagnose the disease at an early stage by predicting TB. According to that, the primary cause of TB is Mycobacteria,

which affects the human respiratory system. All the functions of the lungs will be affected, like chronic coughs, sneezes, sweating, chest pain, and fever. Correspondingly, predicting tuberculosis is necessary to consider a precise treatment to reduce the disease's consequences. For this purpose, the proposed mechanism focused on predicting Active TB, Latent TB, Healthy and Sick but non-TB through the Kaggle TBX-11k dataset. This dataset is comprised of chest radiograph samples that are related to TB and non-TB. Moreover, mAP is the prime metric for object detection. The proposed prediction model, with the advantage of efficient results from a smaller dataset, attains an mAP of 0.657, which is higher than other prevailing models. Besides, the dataset used is insufficient for prevailing research. For this reason, the respective model is examined with the traditional models, which reveal



the efficiency of the proposed research. In the comparison, the proposed method achieved higher mAP values than conventional models and can assist radiologists and experts in providing effective TB prediction. Besides, it is envisioned to improve the life quality of TB patients.

6 Conclusion

Tuberculosis is a pathogenic and deadly disease caused by bacterial Mycobacteria. It mostly affects the human respiratory system and can even lead to death. Many people have been affected by TB owing to inaccuracy, late diagnosis, and deficiency of treatment. Hence, early and accurate diagnosis is a significant solution to preventing and checking tuberculosis. Applying technologies to support the medical business plays a primary

role in improving accuracy and speed in prediction methods. A CXR is a major diagnostic tool utilized to screen for disease. However, the visual inspection accuracy through human experts is time-consuming and limited. To resolve this problem, the proposed research on SFF in the YOLOv8 model aimed to overcome the limitations and improve the prediction accuracy for TB detection, thereby reducing transmission risks and facilitating early intervention. Correspondingly, the utilization of the YOLOv8 architecture enabled TB pattern prediction and accurate detection in CXR images. The TBX-11k dataset is used to demonstrate its effectiveness in distinguishing sick but non-TB, active TB, Latent TB, and healthy cases. The data imbalance is mitigated through data augmentation and class weights methods that are comprised of image augmentation and focal loss, which result in robust generalization and improved performance of the proposed framework. After integrating the SFF into the YOLOv8 model,

the respective research findings indicate that the projected system attained a promising mAP value of 0.657 on the validation set, making it an assistive tool for radiologists. Constantly, the outcome of the comparative analysis signifies that the respective model has outperformed the existing models. Though the proposed model performed efficiently, it did not encompass the large data due to the limited resources available at the time, and limited data are considered limitations. The usage of a wide range of data and associating it to the analysis and understanding the TB characteristics and patterns can be reached in future research.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

MP: Conceptualization, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. WS: Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. AB: Conceptualization, Formal analysis, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. MM: Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing.

References

- Abideen, Z. U., Ghafoor, M., Munir, K., Saqib, M., Ullah, A., Zia, T., et al. (2020). Uncertainty assisted robust tuberculosis identification with Bayesian convolutional neural networks. *IEEE Access* 8, 22812–22825. doi: 10.1109/ACCESS.2020.2970023
- Acharya, V., Choi, D., Yener, B., and Beamer, G. (2024). Prediction of tuberculosis from lung tissue images of diversity outbred mice using jump knowledge based cell graph neural network. *IEEE Access* 12, 17164–17194. doi: 10.1109/ACCESS.2024.3359989
- Acharya, V., Dhiman, G., Prakasha, K., Bahadur, P., Choraria, A., Prabhu, S., et al. (2022). AI-assisted tuberculosis detection and classification from chest X-rays using a deep learning normalization-free network model. *Comput. Intell. Neurosci.* 2022:2399428. doi: 10.1155/2022/2399428
- Amin, S. U., Taj, S., Hussain, A., and Seo, S. (2024). An automated chest X-ray analysis for COVID-19, tuberculosis, and pneumonia employing ensemble learning approach. *Biomed. Signal Process. Control* 87:105408. doi: 10.1016/j.bspc.2023.105408
- An, L., Peng, K., and Yang, X. (2022). “An improved deep network of pulmonary tuberculosis lesions detection based on YOLO,” in *2022 IEEE 24th Int Conf on High Performance Computing & Communications; 8th Int Conf on Data Science & Systems; 20th Int Conf on Smart City; 8th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)* (Hainan: IEEE), 2368–2372.
- Anand, A., Anandan, K. R., Jayaraman, B., and Thai, M. T. N. T. (2021). “Simple neural network based TB classification,” in *CLEF (Working Notes)* (Bucharest), 1145–1150.
- Ayaz, M., Shaukat, F., and Raja, G. (2021). Ensemble learning based automatic detection of tuberculosis in chest X-ray images using hybrid feature descriptors. *Phys. Eng. Sci. Med.* 44, 183–194. doi: 10.1007/s13246-020-00966-0
- Bista, R., Timilsina, A., Manandhar, A., Paudel, A., Bajracharya, A., Wagle, S., et al. (2023). Advancing tuberculosis detection in chest X-rays: a YOLOv7-based approach. *Information* 14:655. doi: 10.3390/info14120655
- Capellán-Martín, D., Gómez-Valverde, J. J., Bermejo-Peláez, D., and Ledesma-Carbayo, M. (2023). A lightweight, rapid and efficient deep convolutional network for chest X-ray tuberculosis detection. *arXiv: 2309.02140*. doi: 10.1109/ISBI53787.2023.10230500
- Deey, S., Roychoudhury, R., Malakar, S., and Sarkar, R. (2022). An optimized fuzzy ensemble of convolutional neural networks for detecting tuberculosis from Chest X-ray images. *Appl. Soft Comput.* 114:108094. doi: 10.1016/j.asoc.2021.108094
- Ekins, S., and Freundlich, J. S. (2011). Validating new tuberculosis computational models with public whole cell screening aerobic activity datasets. *Pharm. Res.* 28, 1859–1869. doi: 10.1007/s11095-011-0413-x
- Heyckendorf, J., Marwitz, S., Reimann, M., Avsar, K., DiNardo, A. R., Günther, G., et al. (2021). Prediction of anti-tuberculosis treatment duration based on a 22-gene transcriptomic model. *Eur. Respir. J.* 58, 1–13. doi: 10.1183/13993003.03492-2020
- Hu, X., Wang, J., Ju, Y., Zhang, X., Qimanguli, W., Li, C., et al. (2022). Combining metabolome and clinical indicators with machine learning provides some promising diagnostic markers to precisely detect smear-positive/negative pulmonary tuberculosis. *BMC Infect. Dis.* 22:707. doi: 10.1186/s12879-022-07694-8
- Hwang, E. J., Jeong, W. G., David, P.-M., Arentz, M., Ruhwald, M., Yoon, S. H., et al. (2024). AI for detection of tuberculosis: Implications for global health. *Radiol. Artif. Intell.* 6:e230327. doi: 10.1148/ryai.230327
- Ignatius, J. L. P., Selvakumar, S., Paul, K. G. J. L., Kailash, A. B., Keertivaas, S., Prajan, S. A. R., et al. (2023). Histogram matched chest X-rays based tuberculosis detection using CNN. *Comp. Syst. Sci. Eng.* 44, 81–97. doi: 10.32604/csse.2023.025195
- Iqbal, A., Usman, M., and Ahmed, Z. (2022). An efficient deep learning-based framework for tuberculosis detection using chest X-ray images. *Tuberculosis* 136:102234. doi: 10.1016/j.tube.2022.102234
- Kafentzis, G. P., Tetsing, S., Brew, J., Jover, L., Galvosas, M., Chaccour, C., et al. (2023). Predicting tuberculosis from real-world cough audio recordings and metadata. *arXiv.* (2023). doi: 10.48550/arXiv.2307.04842
- Kaur, R., and Sharma, A. (2021). An accurate integrated system to detect pulmonary and extra pulmonary tuberculosis using machine learning algorithms. *Intel. Artif.* 24, 104–122. doi: 10.4114/intartif.vol24iss68pp104-122
- Kotei, E., and Thirunavukarasu, R. (2024). A comprehensive review on advancement in deep learning techniques for automatic detection of tuberculosis from chest X-ray images. *Arch. Comp. Methods Eng.* 31, 455–474. doi: 10.1007/s11831-023-09987-w

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

This project is sponsored by Prince Sattam bin Abdulaziz University (PSAU) as part of funding for its SDG Roadmap Research Funding Programme project number PSAU/2023/SDG/84.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Le, T.-M., Nguyen-Tat, B.-T., and Ngo, V. M. (2022). Automated evaluation of tuberculosis using deep neural networks. *Eur. Union Digit. Library* 9:e4. doi: 10.4108/eetinis.v8i30.478
- Liu, C., Cao, Y., Alcantara, M., Liu, B., Brunette, M., Peinado, J., et al. (2017). *TX-CNN: detecting tuberculosis in chest X-ray images using convolutional neural network*. Beijing: IEEE
- Lu, X., Yan, A., Chang, E. Y., Hsu, C. N., McAuley, J. J., Du, J., et al. (2022). *Semi-supervised Multi-Label Classification with 3D CBAM Resnet for Tuberculosis Cavern Report*. Elsevier Limited.
- Luo, Y., Xue, Y., Song, H., Tang, G., Liu, W., Bai, H., et al. (2021). Machine learning based on routine laboratory indicators promoting the discrimination between active tuberculosis and latent tuberculosis infection. *J. Infect.* 84, 648–657. doi: 10.1016/j.jinf.2021.12.046
- Maipan-uku, J. Y., and Cavus, N., and Sekeroglu, S. (2024). Short-term tuberculosis incidence rate prediction for europe using machine learning algorithms. *J. Optimiz. Ind. Eng.* 16, 213–219. doi: 10.22094/joie.2023.1988443.2079
- Munadi, K., Muchtar, K., Maulina, N., and Pradhan, B. (2020). Image enhancement for tuberculosis detection using deep learning. *IEEE Access* 8, 217897–217907. doi: 10.1109/ACCESS.2020.3041867
- Nafisah, S. I., and Muhammad, G. (2024). Tuberculosis detection in chest radiograph using convolutional neural network architecture and explainable artificial intelligence. *Neural Comput. Appl.* 36, 111–131. doi: 10.1007/s00521-022-07258-6
- Nijjati, M., Zhang, Z., Abulizi, A., Miao, H., Tuluhong, A., Quan, S., et al. (2021). Deep learning assistance for tuberculosis diagnosis with chest radiography in low-resource settings. *J. Xray. Sci. Technol.* 29, 785–796. doi: 10.3233/XST-210894
- Osamor, V. C., and Okezie, A. F. (2021). Enhancing the weighted voting ensemble algorithm for tuberculosis predictive diagnosis. *Sci. Rep.* 11:14806. doi: 10.1038/s41598-021-94347-6
- Pahar, M., Klopper, M., Reeve, B., Warren, R., Theron, G., Niesler, T., et al. (2021). Automatic cough classification for tuberculosis screening in a real-world environment. *Physiol. Meas.* 42:105014. doi: 10.1088/1361-6579/ac2fb8
- Pathak, K. C., Kundaram, S. S., Sarvaiya, J. N., and Darji, A. (2022). Diagnosis and analysis of tuberculosis disease using simple neural network and deep learning approach for chest X-ray images. *Track. Prev. Dis. Artif. Intell.* 206, 77–102. doi: 10.1007/978-3-030-76732-7_4
- Prasitpuriprecha, C., Jantama, S. S., Preeprem, T., Pitakaso, R., Srichok, T., Khonjun, S., et al. (2022). Drug-resistant tuberculosis treatment recommendation, and multi-class tuberculosis detection and classification using ensemble deep learning-based system. *Pharmaceuticals* 16:13. doi: 10.3390/ph16010013
- Rahman, T., Khandakar, A., Abdul Kadir, M., Rejaul Islam, K., Islam, K. F., Mazhar, R., et al. (2020). Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. *IEEE Access* 8, 191586–191601. doi: 10.1109/ACCESS.2020.3031384
- Rajaraman, S., and Antani, S. K. (2020). Modality-specific deep learning model ensembles toward improving TB detection in chest radiographs. *IEEE Access* 8, 27318–27326. doi: 10.1109/ACCESS.2020.2971257
- Rajaraman, S. G. Z., Folio, L. A., and Antani, S. (2021). Chest x-ray bone suppression for improving classification of tuberculosis-consistent findings. *Diagnostics* 11:840. doi: 10.3390/diagnostics11050840
- Rashidi, H. H., Dang, L. T., Albahra, S., Ravindran, R., and Khan, I. H. (2021). Automated machine learning for endemic active tuberculosis prediction from multiplex serological data. *Sci. Rep.* 11:17900. doi: 10.1038/s41598-021-97453-7
- Sahlol, A. T., Abd Elaziz, M., Tariq Jamal, A., Damaševičius, R., and Farouk Hassan, O. (2020). A novel method for detection of tuberculosis in chest radiographs using artificial ecosystem-based optimisation of deep neural network features. *Symmetry* 12:1146. doi: 10.3390/sym12071146
- Sathitratanchewin, S., Sunanta, P., and Pongpirul, K. (2020). Deep learning for automated classification of tuberculosis-related chest X-Ray: dataset distribution shift limits diagnostic performance generalizability. *Heliyon* 6:e04614. doi: 10.1016/j.heliyon.2020.e04614
- Sharma, V., Gupta, S. K., and Shukla, K. K. (2023). Deep learning models for tuberculosis detection and infected region visualization in chest X-ray images. *Intell. Med.* 1–10. doi: 10.1016/j.imed.2023.06.001
- Simi Margarat, G., Hemalatha, G., Mishra, A., Shaheen, H., Maheswari, K., Tamijesvelan, S., et al. (2022). Early diagnosis of tuberculosis using deep learning approach for IOT based healthcare applications. *Comput. Intell. Neurosci.* 2022:3357508. doi: 10.1155/2022/3357508
- Singh, V., Gourisaria, M. K., Harshvardhan, G., and Singh, V. (2021). *Mycobacterium tuberculosis detection using CNN ranking approach*, in *Advanced Computational Paradigms and Hybrid Intelligent Computing: Proceedings of ICACCP 2021* (Singapore: Springer), 583–596.
- Smith, J. P., Milligan, K., McCarthy, K. D., Mchembere, W., Okeyo, E., Musau, S. K., et al. (2023). Machine learning to predict bacteriologic confirmation of *Mycobacterium tuberculosis* in infants and very young *Children* 2:e0000249. doi: 10.1371/journal.pdig.0000249
- Verma, D., Bose, C., Tufchi, N., Pant, K., Tripathi, V., Thapliyal, A., et al. (2020). An efficient framework for identification of Tuberculosis and Pneumonia in chest X-ray images using Neural Network. *Proc. Comput. Sci.* 171, 217–224. doi: 10.1016/j.procs.2020.04.023
- Wang, Y., Jiang, Z., Liang, P., Liu, Z., Cai, H., and Sun, Q. (2024). TB-DROP: deep learning-based drug resistance prediction of *Mycobacterium tuberculosis* utilizing whole genome mutations. *BMC Genom.* 25:167. doi: 10.1186/s12864-024-10066-y
- Xie, Y., Wu, Z., Han, X., Wang, H., Wu, Y., Cui, L., et al. (2020). Computer-aided system for the detection of multicategory pulmonary tuberculosis in radiographs. *J. Healthc. Eng.* 2020:9205082. doi: 10.1155/2020/9205082
- Yang, F., Lu, P. X., Deng, M., Wang, Y. X. J., Rajaraman, S., Xue, Z., et al. (2022). Annotations of lung abnormalities in the Shenzhen chest X-ray dataset for computer-aided screening of pulmonary diseases. *Data* 7:95. doi: 10.3390/data7070095
- Yoo, S., Geng, H., Chiu, T. L., Yu, S. K., Cho, D. C., Heo, J., et al. (2020). Study on the TB and non-TB diagnosis using two-step deep learning-based binary classifier. *J. Instrum.* 15:10011. doi: 10.1088/1748-0221/15/10/P10011
- Yusoff, M., Saaidi, M. S. I., Afendi, A. S. M., and Hassan, A. M. (2021). Tuberculosis x-ray images classification based dynamic update particle swarm optimization with CNN. *J. Hunan Univ. Nat. Sci.* (2021) 48, 1–10.
- Zhu, Z., Tao, J., Wang, S., Zhang, X., and Zhang, Y. (2023). TBDLNet: a network for classifying multidrug-resistant and drug-sensitive tuberculosis. *arXiv*. e12815. doi: 10.1002/eng2.12815