International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(RESEARCH ARTICLE)

Check for updates

Two-stage deep learning pipeline for tuberculosis detection in chest radiographs

H. S. Esbergenov [1,*] and P. B. Nurimov [2]

[1] *Department of Artificial Intelligence, Tashkent University of Information Technologies named after Muhammad ibn Musa al-Khwarizmi, Tashkent, Uzbekistan.*
[2] *Department of Artificial Intelligence Technologies, Nukus State Technology University, Nukus, Uzbekistan.*

International Journal of Science and Research Archive, 2026, 19(02), 1592-1602

Publication history: Received on 20 April 2026; revised on 27 May 2026; accepted on 29 May 2026

Article DOI: https://doi.org/10.30574/ijsra.2026.19.2.1222

Abstract

Tuberculosis (TB) is a critical public health challenge, particularly in environmentally devastated regions such as Karakalpakstan, Uzbekistan. We propose a two-stage deep learning pipeline for automated TB detection and lesion localization using the TBX11K benchmark dataset. In Stage 1, three CNN backbones — ResNet-18, ResNet-50, and EfficientNet-B0 — classify chest X-ray images into healthy, sick-non-TB, and TB-positive categories. In Stage 2, YOLO11 localizes TB lesion regions via bounding box prediction on TB-positive images. All three backbones exceed 99% accuracy, while YOLO11 achieves a mAP@0.5 of 0.7227 at 1.3 ms inference speed. The lightweight ResNet-18 remains competitive with deeper models, supporting deployment in resource-limited clinical settings.

Keywords: Tuberculosis Detection; Chest X-Ray; Convolutional Neural Networks; YOLO11; Object Detection; TBX11K; Transfer Learning; Computer-Aided Diagnosis

1. Introduction

Tuberculosis (TB) remains one of the world's deadliest infectious diseases, with 8.2 million new cases reported in 2023 — the highest since WHO began global monitoring — again surpassing COVID-19 as the leading infectious disease killer [1].

The burden of TB is especially severe in Karakalpakstan, an autonomous republic in western Uzbekistan. The desiccation of the Aral Sea — described by the United Nations as the largest man-made environmental disaster in the world [14] — has caused catastrophic public health consequences for its 1.5 million residents. Toxic dust storms from the exposed seabed carry salt and pesticide residues across the region, severely degrading respiratory health [15]. TB incidence in Karakalpakstan has been reported at approximately 220 per 100,000 people — more than three times the WHO epidemic threshold of 50–70 per 100,000 [16]. Médecins Sans Frontières, operating in the region since 1998, has documented high rates of multidrug-resistant TB directly linked to the Aral Sea disaster [17], and a 2024 study confirmed that Karakalpakstan consistently records the highest TB rates in Uzbekistan [18].

Chest X-ray (CXR) radiography is the most widely used TB screening tool, yet manual interpretation requires experienced radiologists and is prone to error — particularly in resource-constrained environments. Deep learning offers a scalable solution. ResNet [2] and EfficientNet [3] have become standard backbones for medical image classification through transfer learning, while YOLO11 [4] — released in 2024 — provides state-of-the-art real-time object detection. Liu et al. introduced TBX11K [5], the largest annotated TB benchmark, enabling end-to-end pipeline evaluation.

* Corresponding author: H.S.Esbergenov

We propose a two-stage pipeline that classifies CXR images with CNN backbones and localizes TB lesions with YOLO11. The contributions of this work are: (1) a systematic comparison of ResNet-18, ResNet-50, and EfficientNet-B0 for TB classification; (2) the first application of YOLO11 to the TBX11K detection benchmark; and (3) a complete, clinically interpretable pipeline motivated by the urgent TB screening needs of Karakalpakstan.

The paper is organized as follows: Section II reviews related work; Section III describes the dataset and methodology; Section IV presents results; Section V discusses findings; Section VI concludes.

2. Related work

2.1. Tuberculosis Detection Using Deep Learning

The application of deep learning to TB detection from CXR images has gained significant momentum in recent years. Hansun et al. conducted a systematic literature review on machine and deep learning for TB detection on chest X-rays, confirming the increasing adoption of CNN-based approaches and identifying the Shenzhen and Montgomery datasets as the most widely used benchmarks in the field [6]. However, these earlier datasets are limited in size and annotation quality, restricting the development of more sophisticated detection models.

Liu et al. introduced the TBX11K dataset to address the lack of large-scale annotated TB data, proposing reformed object detectors for simultaneous image classification and TB area detection, and establishing the first standardized benchmark for computer-aided TB diagnosis [5]. Building on this work, Bista et al. applied the YOLOv7 architecture to TBX11K, employing class weights and data augmentation techniques to address dataset imbalance, achieving a mean average precision of 0.587 [7]. More recently, Rahamathulla et al. investigated YOLOv8 for TB identification from chest images, utilizing its ability to detect multiple objects in a single pass, while noting limitations in small object detection and fine-grained classification [8]. Sharma et al. developed deep learning segmentation and classification models for accurate TB detection in CXR images, incorporating gradient-weighted class activation mapping for visualization of infected regions [10].

2.2. CNN Backbones for Medical Image Classification

Transfer learning with pretrained CNN backbones has become the standard approach for medical image classification tasks. Kim et al. conducted a comprehensive review of transfer learning with CNNs for medical image classification, demonstrating that pretrained ImageNet models consistently outperform models trained from scratch, particularly in low-data regimes, and identifying ResNet and EfficientNet as among the most effective backbone architectures [9]. He et al. introduced the residual learning framework, demonstrating that networks with skip connections are easier to optimize and can achieve significantly higher accuracy with increased depth, establishing ResNet as a foundational architecture in medical imaging [2]. Tan and Le proposed EfficientNet, showing that systematically balancing network depth, width, and input resolution through compound scaling achieves superior accuracy with substantially fewer parameters compared to conventional CNN scaling approaches [3].

Kim et al. demonstrated that multi-class lung disease classification using EfficientNet-based architectures with transfer learning achieves competitive diagnostic performance on CXR datasets [11]. Sethanan et al. proposed an ensemble deep learning approach for computer-aided diagnosis of drug-resistant TB, achieving state-of-the-art accuracy improvements over standard CNN models including EfficientNet variants [12].

2.3. YOLO-Based Object Detection in Medical Imaging

The YOLO family of detectors has increasingly been applied to medical image analysis due to its real-time inference capability and strong detection performance. Several studies have demonstrated the effectiveness of YOLO-based approaches for lesion localization in chest radiographs, including detection of pneumonia, COVID-19, and TB regions. YOLO11, released by Ultralytics in 2024, introduces an improved backbone and neck architecture with enhanced feature extraction capabilities, offering improved precision and efficiency over its predecessors [4]. Hirao et al. proposed a two-stage AI-CAD pipeline for TB screening using lung segmentation followed by CNN-based classification and localization, demonstrating the clinical utility of multi-stage detection frameworks [13].

2.4. Gap in Existing Literature

Despite significant progress, several gaps remain. First, most existing studies evaluate classification and detection as isolated tasks, without integrating them into a cohesive clinical pipeline. Second, systematic comparisons of lightweight versus deeper backbone architectures on the TBX11K benchmark are limited. Third, to the best of our knowledge,

YOLO11 has not yet been applied to the TBX11K dataset. This work addresses all three gaps by proposing a two-stage pipeline that unifies backbone-based classification with YOLO11 detection, and provides a rigorous comparative evaluation of three CNN architectures.

3. Dataset and methodology

3.1. Dataset

This study utilizes the TBX11K dataset, introduced by Liu et al. [5], which is currently the largest publicly available benchmark for computer-aided TB diagnosis. The dataset contains 11,200 chest X-ray images with corresponding bounding box annotations for TB areas, representing a significant expansion over previously available TB datasets. All images are standardized to a resolution of 512x512 pixels. The dataset is organized into three subsets: 6,600 images for training, 1,800 images for validation, and 3,302 images for testing.

Images are categorized into four groups: healthy, sick-but-non-TB, active TB, and latent TB. In this work, the active TB and latent TB categories are merged into a single TB-positive class, resulting in a three-class classification problem: healthy, sick-non-TB, and TB. Images labeled as uncertain TB are excluded from training to maintain label quality. The class distribution in the training set comprises 3,000 healthy, 3,000 sick-non-TB, and 600 TB images, reflecting a 5:1 class imbalance between negative and positive TB cases.

For the detection stage, bounding box annotations are available for 599 training images and 200 validation images, all belonging to the TB-positive class. Two TB subtypes are annotated: Active Tuberculosis and Obsolete Pulmonary Tuberculosis. In this work, both subtypes are merged into a single detection class — TB — consistent with the binary clinical screening objective.

3.2. Two-Stage Pipeline Overview

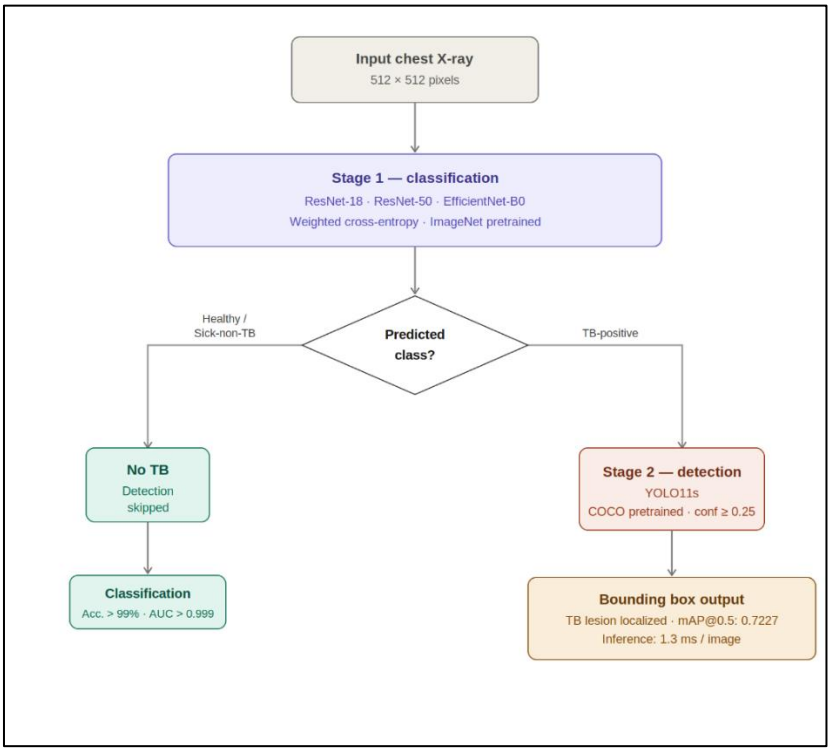

Figure 1 Overview of the proposed two-stage TB detection pipeline. Stage 1 classifies each CXR into healthy, sick-non-TB, or TB-positive using a CNN backbone. TB-positive images are forwarded to Stage 2 where YOLO11 localizes lesion regions via bounding box prediction

The proposed pipeline consists of two sequential stages as illustrated in Fig. 1. In Stage 1, a CNN backbone classifies each input CXR image into one of three classes. Images classified as TB-positive are forwarded to Stage 2, where a

YOLO11 detector localizes the lesion regions by predicting bounding boxes. Images classified as healthy or sick-non-TB are not passed to the detector, reducing unnecessary computation and false positive detections.

3.3. Stage 1: Classification

3.3.1. Backbone Architectures:

Three CNN backbones are evaluated for the classification stage, all initialized with ImageNet pretrained weights and fine-tuned on the TBX11K training set.

ResNet-18 [2] is a lightweight residual network with 18 layers and approximately 11 million parameters. Its residual connections enable stable gradient flow during training, making it an efficient baseline for image classification tasks.

ResNet-50 [2] is a deeper variant of the ResNet family with 50 layers and approximately 25 million parameters. It employs bottleneck residual blocks that improve representational capacity while maintaining computational tractability.

EfficientNet-B0 [3] is the baseline model of the EfficientNet family, containing approximately 5 million parameters. It is constructed using neural architecture search and compound scaling, achieving strong accuracy with significantly fewer parameters than ResNet variants.

For all three backbones, the final fully connected layer is replaced with a linear layer mapping to three output classes. A Global Average Pooling layer precedes the classification head.

3.3.2. Training Configuration:

All models are trained under a unified protocol to ensure fair comparison. The training configuration is summarized in Table I.

Table 1 Classification Training Hyperparameters

Hyperparameter	**Value**
Input resolution	512 x 512
Batch size	64
Optimizer	AdamW
Learning rate	1e-4
LR scheduler	CosineAnnealingLR
Weight decay	1e-4
Epochs	30
Loss function	Weighted CrossEntropyLoss

3.3.3. Class Imbalance Handling

To address the 5:1 class imbalance between negative and TB-positive samples, two complementary strategies are applied. First, class weights are computed inversely proportional to class frequency and incorporated into the cross-entropy loss function. Second, a WeightedRandomSampler is applied during training to oversample minority class instances, ensuring balanced mini-batches.

3.3.4. Data Augmentation

The following augmentations are applied during training: random horizontal flip (p=0.5), random rotation (+/-10 degrees), and random color jitter (brightness and contrast +/-0.2). Validation images are not augmented. All images are normalized using ImageNet mean and standard deviation values.

3.4. Stage 2: Detection

3.4.1. Dataset Preparation

A balanced detection dataset is constructed from the TBX11K annotations. All 599 TB training images with bounding box annotations are included. To provide the model with negative context and reduce false positives, an equal number of healthy and sick-non-TB images are randomly sampled — 599 from each class — yielding a total of 1,797 training images in a 1:1:1 ratio. For validation, all 200 annotated TB images are included alongside 200 healthy and 200 sick-non-TB images, totaling 600 validation images. Bounding box annotations are originally provided in COCO format [x, y, width, height] and converted to YOLO format [cx, cy, width, height] with all coordinates normalized to the range [0, 1].

3.4.2. YOLO11 Architecture

YOLO11s [4] is selected as the detection model. YOLO11 employs an improved backbone and neck architecture featuring C3k2 bottleneck blocks and a C2PSA attention module, which enhances feature aggregation across multiple scales. The model contains approximately 9.4 million parameters and operates at 21.5 GFLOPs, making it well suited for deployment on both research and clinical hardware.

3.4.3. Detection Training Configuration:

The YOLO11s model is initialized with pretrained COCO weights and fine-tuned on the constructed TB detection dataset. The training configuration is summarized in Table II.

Table 2 YOLO11 Detection Training Hyperparameters

Hyperparameter	**Value**
Input resolution	512 x 512
Batch size	32
Optimizer	AdamW (auto)
Learning rate	2e-3
Epochs	50
Early stopping patience	10
Confidence threshold	0.25
IoU threshold	0.7
Pretrained weights	COCO

3.4.4. Evaluation Metrics:

For Stage 1, classification performance is evaluated using accuracy, macro-averaged F1-score, macro-averaged AUC-ROC, and per-class precision, recall, and F1-score. For Stage 2, detection performance is evaluated using mean average precision at IoU threshold 0.5 (mAP@0.5), mAP averaged over IoU thresholds from 0.5 to 0.95 (mAP@0.5:0.95), precision, recall, and inference time per image. All evaluations are performed on the official TBX11K validation set.

4. Experimental results

4.1. Classification Results

Table III presents the classification performance of the three evaluated backbone architectures on the TBX11K validation set. All models are trained under identical conditions as described in Section III.

Table 3 Classification Performance on TBX11K Validation Set

Model	Accuracy	Macro-F1	AUC (macro)	TB Precision	TB Recall	TB F1
ResNet-18	0.9939	0.9910	0.9991	0.9949	0.9700	0.9823
ResNet-50	0.9956	0.9935	0.9998	0.9899	0.9850	0.9875
EfficientNet-B0	0.9944	0.9921	0.9995	0.9850	0.9850	0.9850

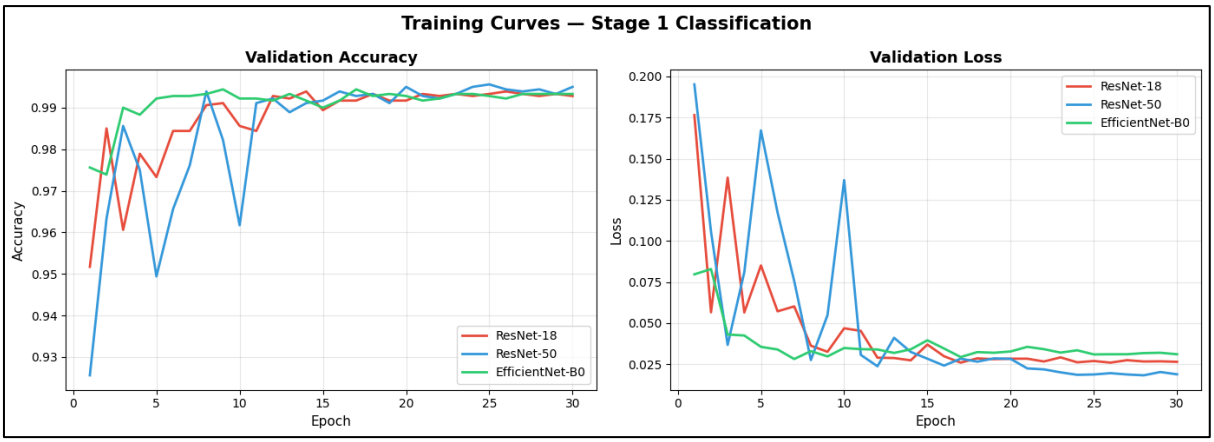

Figure 2 Validation accuracy and loss curves for ResNet-18, ResNet-50, and EfficientNet-B0 over 30 training epochs

All three models achieve accuracy above 99%, with macro-averaged AUC scores exceeding 0.999, indicating strong discriminative capability across all three classes. ResNet-50 attains the highest overall accuracy of 99.56% and the highest macro-F1 score of 0.9935. EfficientNet-B0 achieves competitive performance with 99.44% accuracy despite having the fewest parameters among the three models. ResNet-18, the lightest backbone evaluated, achieves 99.39% accuracy — only 0.17% below ResNet-50 — demonstrating that lightweight architectures are sufficient for this classification task.

The TB class, being the most clinically critical, is analyzed separately. ResNet-50 and EfficientNet-B0 both achieve a TB recall of 0.9850, meaning that 3 out of 200 TB validation images are misclassified. ResNet-18 achieves a TB recall of 0.9700, corresponding to 6 misclassified TB images. In a clinical screening context, recall is the most important metric as it reflects the model's ability to correctly identify true TB cases and minimize missed diagnoses.

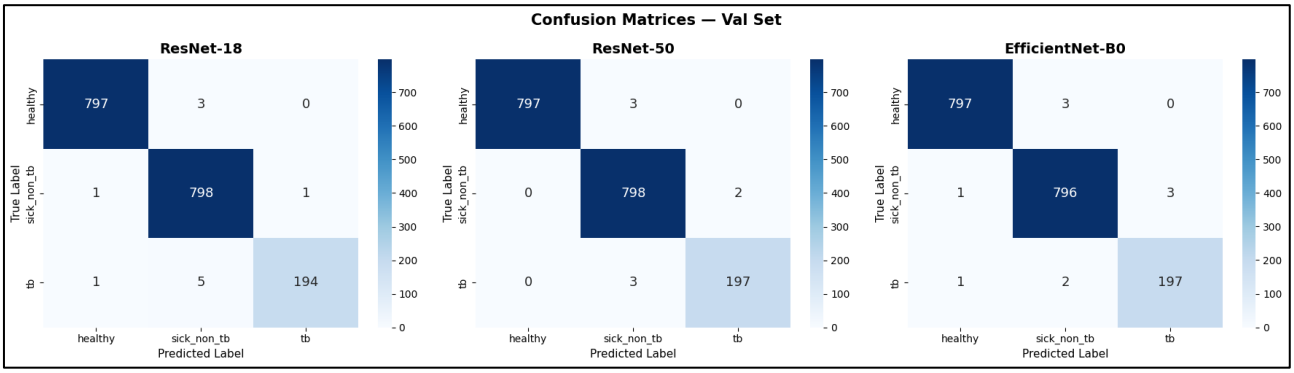

Figure 3 Confusion matrices for ResNet-18, ResNet-50, and EfficientNet-B0 on the TBX11K validation set

The confusion matrices in Fig. 3 confirm that misclassifications are rare and occur primarily between the TB and sick-non-TB classes, which is expected given the visual similarity between pulmonary abnormalities of different etiologies. No TB image is misclassified as healthy by any model, which is a clinically important finding.

4.2. Detection Results

Table IV presents the detection performance of YOLO11s on the TBX11K validation set.

Table 4 YOLO11s Detection Performance on TBX11K Validation Set

Model	mAP@0.5	mAP@0.5:0.95	Precision	Recall	Inference (ms)	Params (M)
YOLO11s	0.7227	0.3473	0.7592	0.6529	1.3	9.4

YOLO11s achieves a mAP@0.5 of 0.7227, indicating that approximately 72% of TB lesion regions are correctly localized at an IoU threshold of 0.5. The mAP@0.5:0.95 score of 0.3473 reflects the stricter COCO-style evaluation averaged across multiple IoU thresholds. The model achieves a precision of 0.7592 and a recall of 0.6529. The inference speed of 1.3 milliseconds per image demonstrates the real-time capability of the model.

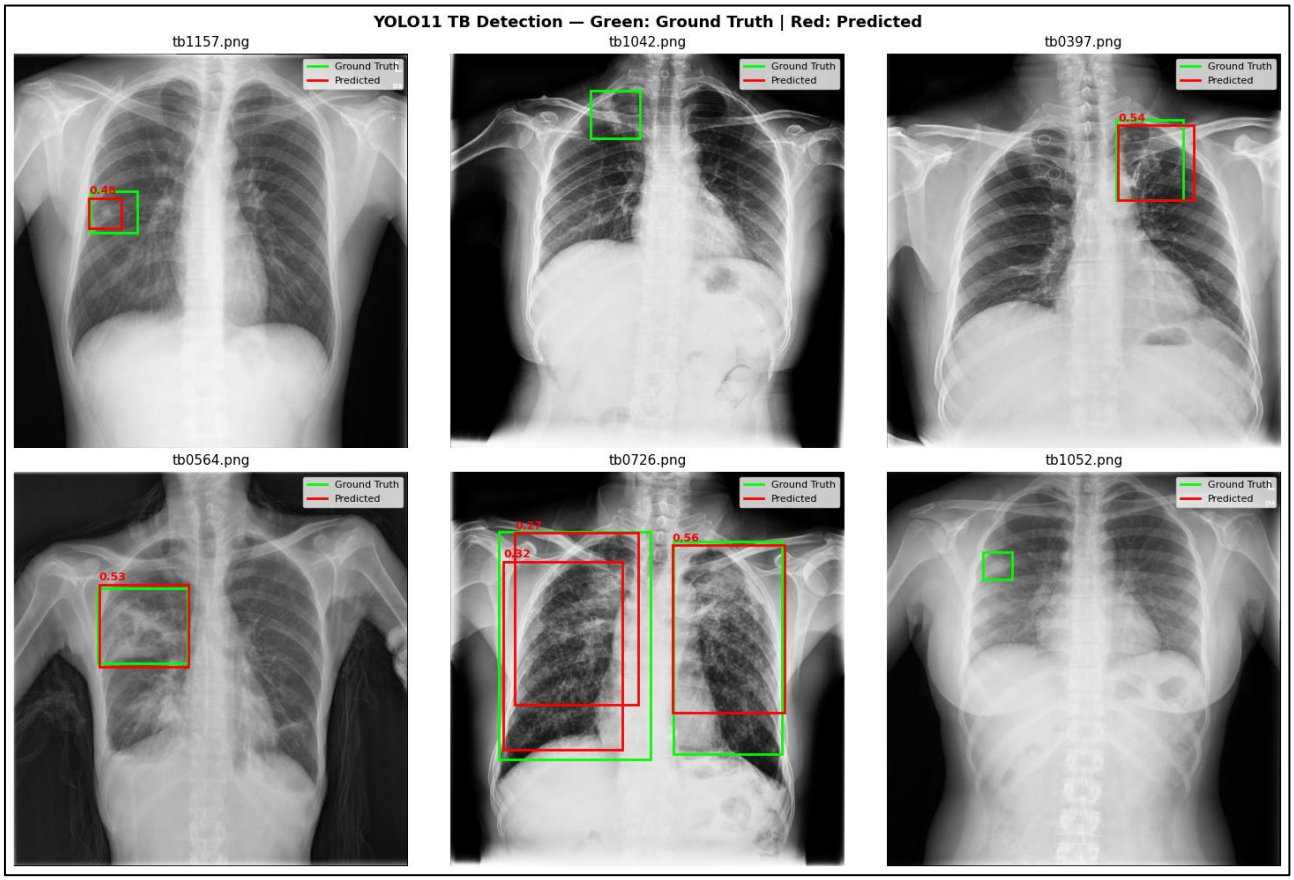

Figure 4 Sample YOLO11s predictions on TB validation images. Green boxes indicate ground truth annotations. Red boxes indicate model predictions with associated confidence scores

Fig. 4 presents qualitative detection results on six representative TB validation images. The results illustrate four distinct detection behaviors: successful localization with strong spatial overlap in tb0397 and tb0564, missed detections in tb1042 and tb1052, positional offset in tb1157, and over-detection in tb0726 where three predicted boxes span the region covered by two ground truth annotations.

4.3. Qualitative Demo Evaluation

To further assess the generalizability of the proposed pipeline, a qualitative evaluation is conducted on two unseen chest X-ray images sourced outside the TBX11K dataset. These images are processed through the full two-stage pipeline without any preprocessing or fine-tuning. Since ground truth annotations are not available for these images, only predicted bounding boxes are shown.

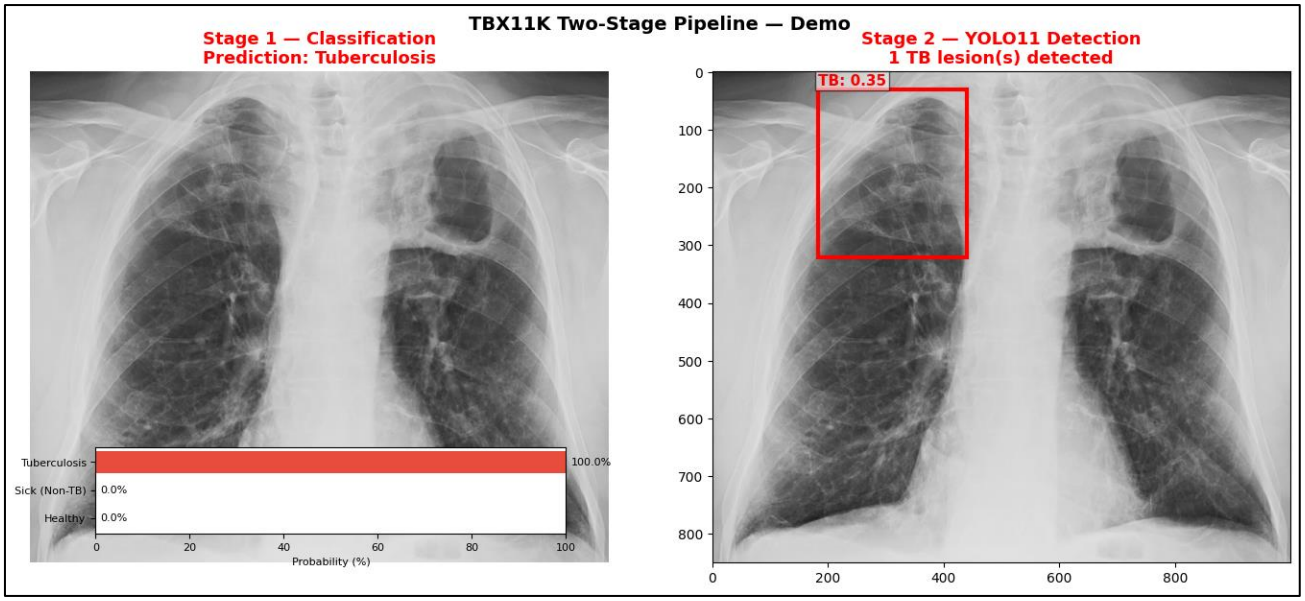

Figure 5 Pipeline output on an unseen TB-positive chest X-ray [19]. Left: Stage 1 classification result with probability distribution. Right: Stage 2 YOLO11 lesion detection with predicted bounding box

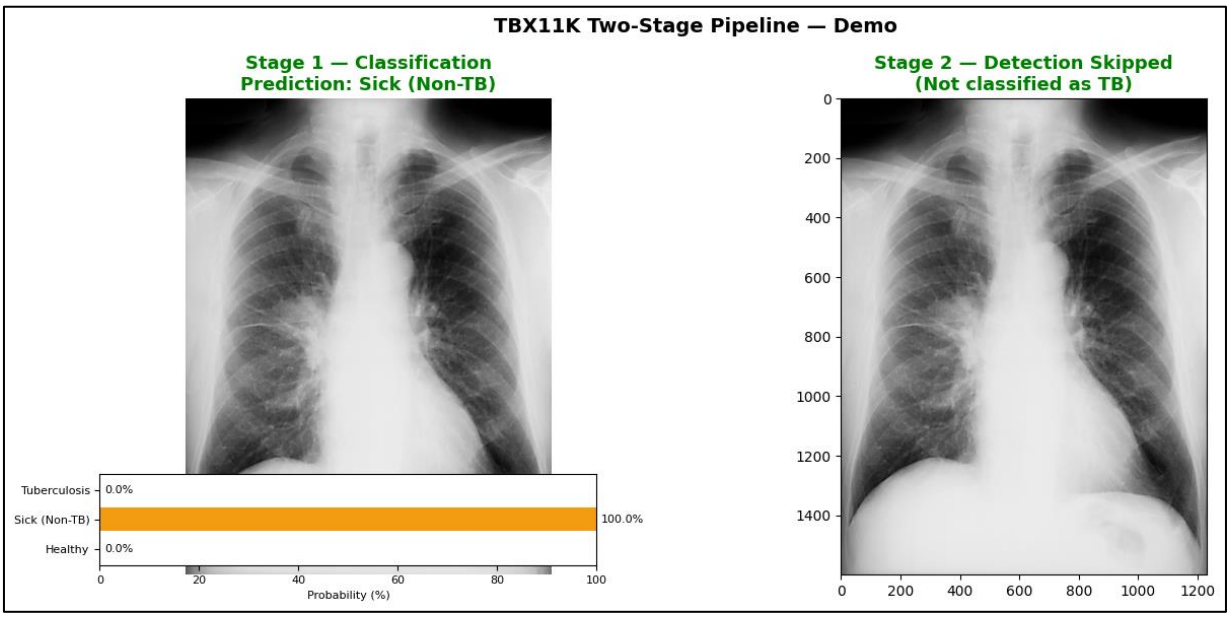

Figure 6 Pipeline output on an unseen sick-non-TB (lung cancer) chest X-ray [20]. Left: Stage 1 classification result. Right: Stage 2 detection skipped as image was not classified as TB-positive

The first image, a TB-positive case sourced from Wikimedia Commons [19], is correctly classified by Stage 1 and the YOLO11 detector successfully localizes the lesion region in Stage 2. The second image, a lung cancer case sourced from MyLungCancerTeam [20], is correctly classified as sick-non-TB in Stage 1 and the detection stage is appropriately skipped, demonstrating the filtering capability of the two-stage design.

4.4. Comparative Discussion

Among the classification backbones, ResNet-18 presents the most favorable trade-off between performance and computational cost. Despite having more than twice fewer parameters than ResNet-50, its accuracy difference is marginal (0.17%), and its TB recall difference is only 1.5 percentage points. This finding suggests that increasing model complexity beyond ResNet-18 yields diminishing returns for this particular task, which has practical implications for deployment in resource-limited clinical settings.

5. Discussion

5.1. Classification Performance Analysis

The classification results demonstrate that all three evaluated backbone architectures achieve exceptional performance on the TBX11K dataset, with accuracy values exceeding 99% and macro-averaged AUC scores above 0.999. These results suggest that the visual patterns distinguishing healthy, sick-non-TB, and TB-positive chest radiographs are sufficiently discriminative for pretrained CNN models to learn effectively, even with a relatively small number of TB-positive training samples (600 images).

A key finding of this study is that ResNet-18, despite being the lightest architecture evaluated, achieves performance closely comparable to ResNet-50 and EfficientNet-B0. The marginal accuracy gap of 0.17% between ResNet-18 and ResNet-50 does not justify the additional computational cost of the deeper model in a clinical screening context. This observation is consistent with findings in the broader medical imaging literature, where lightweight models have been shown to perform competitively on tasks involving structurally constrained image domains such as chest radiography [9]. The relatively low visual complexity of CXR images may reduce the benefit of deeper feature hierarchies, making ResNet-18 a practical and efficient choice for Stage 1 of the pipeline.

The TB class recall is the most clinically significant metric in this study, as a missed TB diagnosis carries greater risk than a false positive in a screening setting. ResNet-50 and EfficientNet-B0 achieve a TB recall of 0.9850, while ResNet-18 achieves 0.9700. Although the difference is small in absolute terms, it represents 3 additional missed TB cases per 200 validation images for ResNet-18. Future work may explore whether post-hoc threshold calibration or ensemble methods can further reduce missed TB detections across all three models.

5.2. Detection Performance Analysis

The YOLO11s detector achieves a mAP@0.5 of 0.7227 on the TBX11K validation set. This result represents a meaningful improvement over prior YOLO-based work on the same dataset. Bista et al. [7] reported a mAP of 0.587 using YOLOv7 on TBX11K, while Rahamathulla et al. [8] applied YOLOv8 without reporting comparable mAP scores on TBX11K directly. The improvement from 0.587 to 0.7227 suggests that the architectural advancements in YOLO11 — including the C2PSA attention module and improved feature pyramid neck — contribute meaningfully to TB lesion localization.

The recall of 0.6529 indicates that approximately 35% of annotated TB lesions are not detected at the default confidence threshold of 0.25. Several factors may explain this limitation. First, TB lesions in chest radiographs are often small, irregularly shaped, and low in contrast relative to surrounding tissue, making precise localization inherently difficult. Second, the limited number of annotated training images (599 boxes) constrains the model's ability to generalize. Third, the model was trained for only 50 epochs and showed consistent improvement through the final epoch, suggesting that additional training could yield further gains.

The inference speed of 1.3 milliseconds per image on an A100 GPU demonstrates that the detection stage is well within real-time operational requirements. Even on less powerful hardware, YOLO11s is expected to maintain clinically acceptable inference speeds, supporting deployment in resource-limited settings.

5.3. Two-Stage Pipeline Evaluation

The proposed two-stage pipeline offers several advantages over single-stage end-to-end detection approaches. By first filtering images through a high-accuracy classifier, the detection stage receives only TB-positive images, significantly reducing the number of false positive bounding box predictions that would otherwise occur on healthy or non-TB images. The pipeline also provides two levels of clinical interpretability: Stage 1 produces a class probability distribution, and Stage 2 produces spatial bounding boxes highlighting specific lung regions of concern.

5.4. Limitations

Several limitations of this study should be acknowledged. First, the TBX11K test set ground truth labels are not publicly released, and all evaluations are therefore conducted on the official validation set. Second, the relatively small number of annotated TB detection images constrains the generalizability of the YOLO11 detector. Third, all experiments are conducted on a single GPU (NVIDIA A100) in a cloud environment, and inference time benchmarks may differ on clinical deployment hardware. Fourth, this study does not evaluate the pipeline on external datasets, and cross-institutional generalization remains to be validated. Fifth, the demo evaluation on out-of-distribution images revealed that the

classifier occasionally misclassified healthy images, suggesting potential sensitivity to image acquisition differences between the TBX11K distribution and real-world CXR images.

6. Conclusion

This paper presented a two-stage deep learning pipeline for automated TB detection and lesion localization in chest radiographs using the TBX11K dataset. ResNet-18, ResNet-50, and EfficientNet-B0 were compared for three-class classification, while YOLO11s localized TB lesions via bounding box prediction.

All three backbones exceeded 99% accuracy with AUC above 0.999. ResNet-18 achieved competitive results with the fewest parameters, supporting lightweight deployment. YOLO11s achieved mAP@0.5 of 0.7227 — a significant improvement over prior YOLO-based work on TBX11K — at 1.3 ms inference speed.

This work is directly motivated by the TB crisis in Karakalpakstan, Uzbekistan, where decades of Aral Sea desertification have created one of the most severe environmental health emergencies in Central Asia, driving TB rates far beyond internationally recognized epidemic levels. The exposed seabed continues to release toxic dust across the region, placing an extraordinary burden on an already strained healthcare system with limited access to trained radiologists. The proposed pipeline — requiring only a standard chest X-ray and operating in real time — is specifically designed for such resource-limited clinical environments. Future work will focus on collecting a region-specific Karakalpakstan CXR dataset, improving detection recall, and pursuing clinical validation in collaboration with the Ministry of Health of the Republic of Karakalpakstan and international partners.

Compliance with ethical standards

Acknowledgments

The authors thank the creators of the TBX11K dataset for making it publicly available.

Disclosure of Conflict of Interest

No conflict of interest to be disclosed.

Statement of Ethical Approval

This study is based solely on a publicly available dataset. No ethical approval was required.

References

[1] World Health Organization, "Global Tuberculosis Report 2024," WHO Press, Geneva, Switzerland, 2024. [Online]. Available: https://www.who.int/publications/i/item/9789240083851

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, Jun. 2016, pp. 770–778.

[3] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proc. 36th Int. Conf. Machine Learning (ICML), Long Beach, CA, 2019, pp. 6105–6114.

[4] G. Jocher and J. Qiu, "Ultralytics YOLO11," version 11.0.0, Ultralytics, 2024. [Online]. Available: https://github.com/ultralytics/ultralytics

[5] Y. Liu, Y.-H. Wu, Y. Ban, H. Wang, and M.-M. Cheng, "Rethinking Computer-Aided Tuberculosis Diagnosis," in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), 2020, pp. 2646–2655.

[6] 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," J. Med. Internet Res., vol. 25, p. e43154, 2023. doi: 10.2196/43154

[7] R. Bista, A. Timilsina, A. Manandhar, A. Paudel, A. Bajracharya, and S. Wagle, "Advancing Tuberculosis Detection in Chest X-rays: A YOLOv7-Based Approach," Information, vol. 14, no. 12, p. 655, 2023. doi: 10.3390/info14120655

[8] M. P. Rahamathulla, W. R. Sam Emmanuel, A. Bindhu, and M. Mustaq Ahmed, "YOLOv8's advancements in tuberculosis identification from chest images," Front. Big Data, vol. 7, p. 1401981, Jun. 2024. doi: 10.3389/fdata.2024.1401981

[9] H. E. Kim et al., "Transfer learning for medical image classification: a literature review," BMC Med. Imaging, vol. 22, no. 1, p. 69, Apr. 2022. doi: 10.1186/s12880-022-00793-7

[10] V. Sharma, Nillmani, S. K. Gupta, and K. K. Shukla, "Deep learning models for tuberculosis detection and infected region visualization in chest X-ray images," Intell. Med., vol. 4, no. 2, pp. 104–113, May 2024. doi: 10.1016/j.imed.2023.06.001

[11] S. Kim, B. Rim, S. Choi, A. Lee, S. Min, and M. Hong, "Deep Learning in Multi-Class Lung Diseases' Classification on Chest X-ray Images," Diagnostics, vol. 12, no. 4, p. 915, Apr. 2022. doi: 10.3390/diagnostics12040915

[12] K. Sethanan et al., "Computer-aided diagnosis using embedded ensemble deep learning for multiclass drug-resistant tuberculosis classification," Front. Med., vol. 10, p. 1122222, Jun. 2023. doi: 10.3389/fmed.2023.1122222

[13] S. Hirao et al., "Applicability of artificial intelligence-based computer-aided detection (AI-CAD) for pulmonary tuberculosis to community-based active case finding," BMC Infect. Dis., vol. 24, no. 1, Jan. 2024. doi: 10.1186/s12879-023-08933-w

[14] United Nations, "Human Security in the Aral Sea Region," UN Human Security Unit, 2023. [Online]. Available: https://www.un.org/humansecurity

[15] O. Ataniyazova, "Health and Ecological Consequences of the Aral Sea Crisis in Karakalpakstan," ReliefWeb / IRIN Report, UN Office for the Coordination of Humanitarian Affairs, 2003.

[16] Medecins Sans Frontieres (MSF), "The Aral Sea Disappears While Tuberculosis Climbs," MSF Field Report, Nukus, Karakalpakstan, Uzbekistan, 2003. [Online]. Available: https://www.msf.org/aral-sea-disappears-while-tuberculosis-climbs

[17] MSF, "MSF in Karakalpakstan: Improving Care for People Through High Impact Research," MSF Science Portal, 2022. [Online]. Available: https://scienceportal.msf.org

[18] N. Davlyatova et al., "Role of ethnicity and residency in active tuberculosis in Karakalpakstan: study protocol of a matched case-control study," BMJ Open Respir. Res., vol. 11, no. 1, e002554, Oct. 2024. doi: 10.1136/bmjresp-2024-002554

[19] Wikimedia Commons, "Chest radiograph of miliary tuberculosis,"[Online]. Available: https://upload.wikimedia.org/wikipedia/commons/7/70/Chest_radiograph_of_miliary_tuberculosis_2.jpg [Accessed: 2024].

[20] MyLungCancerTeam, "Lung Cancer X-Ray Photos: Examples of Different Types of Results," [Online]. Available: https://www.mylungcancerteam.com/resources/lung-cancer-x-ray-photos-examples-of-different-types-of-results [Accessed: 2024].