iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VIII Month of publication: August 2025

DOI: https://doi.org/10.22214/ijraset.2025.73765

# Hybrid Deep Learning Model for Automated Tuberculosis Detection from Chest X-Rays

Hemalatha K[1], Niteesh Kumar[2], Sangmesh[3], Srujan Kumar D[4], Sudarshan Rudrapure[5]
[1]*Assistant Professor, Sapthagiri College of Engineering, Karnataka, India*
[2, 3, 4, 5]*Computer Science and Engineering, Sapthagiri College of Engineering, Karnataka, India*

***Abstract: Tuberculosis (TB) resides a significant health challenge, particularly in regions with limited access to medical resources. It spreads through the air and mainly affects the lungs, and without early detection, it can be deadly. Unfortunately, many areas don't have enough trained healthcare workers or proper screening tools, which delays diagnosis.To help with this issue, we're currently working on building an AI-based system that can automatically detect TB using chest X-ray images. Our approach involves a multi model deep learning system that combines two architectures: DeiT (Data-efficient Vision Transformer) and ResNet-16. DeiT helps the model understand the overall structure of the image through attention mechanisms, while ResNet-16 focuses on capturing detailed features in specific regions, all while keeping the system lightweight and efficient. We're using the TBX11K dataset , featuring three types of chest X-rays: healthy, non-TB but sick, and TB-positive. To make the system more transparent, we're also adding heatmap visualizations using class activation mapping, so we can actually see which parts of the image the model is focusing on while making predictions.***

***Our target is to design a model that's not only accurate—hopefully hitting above 99% in key metrics—but also fast, with a prediction time of under 5 milliseconds. Once complete, we aim to make it suitable for real-time use in hospitals and clinics, especially in areas where medical support is limited.***

***Keywords: Tuberculosis Detection, Chest X-ray, Deep Learning, Vision Transformer (DeiT), ResNet-16, Hybrid Model, Medical Image Analysis, Class Activation Mapping (CAM), TBX11K Dataset, Real-time Diagnosis, Low-resource Healthcare, AI in Medicine.***

## I. INTRODUCTION

Tuberculosis (TB) continues to pose a serious global health risk, accounting for nearly twice the count of deaths as HIV/AIDS in 2022 and ranking as the second most fatal infectious disease after COVID-19. Although medical system advancement have improved treatment methods, the early detection of TB remains a persistent issue, especially in low-resource settings. The lack of skilled healthcare personnel and inadequate diagnostic infrastructure significantly hinders timely diagnosis and intervention. To help bridge this gap, advancements in Artificial intelligence (AI) and Deep Learning (DL) have high scope in medical image analyzing. Convolutional Neural Networks (CNNs) are commonly utilized since they can automatically obtain significant visual features from chest X-rays without manual processing. However, the reliability of these models is often compromised due to variations in image quality from different machines and the limited availability of annotated medical datasets, leading to overfitting and reduced generalization capability. To enhance model accuracy, researchers often employ transfer learning, in which networks initially trained on extensive datasets are later adapted to perform medical tasks. Ensemble learning is another strategy that combines the predictions of multiple models to boost reliability and reduce prediction variance. Even with these techniques, achieving an optimal trade-off between accuracy, speed, and interpretability continues to be a key challenge.

In response to this, our study presents a integrated deep learning approach that integrates a data-efficient Vision Transformer (DeiT) with ResNet-16. The transformer component leverages self-attention mechanisms to learn global features across chest X-ray images, while ResNet-16 extracts fine-grained local patterns through depth-wise convolutions, which help reduce computational overhead. This combined framework is designed to improve diagnostic performance while maintaining a lightweight and fast architecture—ideal for clinical use, particularly in areas with limited access to healthcare technology.

To enhance model transparency, we incorporate class activation maps (CAMs), which generate heatmaps highlighting the regions of the image the model focuses on during prediction. This helps provide visual explanations for medical professionals, supporting informed decision-making.Initial this technique produces are highly promising, with the model achieving over 99% accuracy, along with rapid processing times, making it suitable for real-time TB screening. The efficient design of this system not only improves diagnostic speed and reliability but also offers a scalable solution for TB detection, especially in under-resourced regions. By leveraging modern deep learning innovations, this method aims to significantly advance global TB screening efforts.

## II. LITERATURE REVIEW

1) *An Automated Tuberculosis Detection Framework Using Chest X-Ray Imaging Based on a Stochastic Neural Network Approach* (Shabana Urooj et al., 2021).

The authors introduce an artificial neural network (ANN) framework that leverages stochastic learning methods for TB diagnosis via chest X-ray images. The approach integrates random elements—such as stochastic transfer functions and randomized weight initialization—to enhance the model's capacity to process and learn from high-dimensional medical image data. The model exhibited robust performance, attaining an accuracy of 98.45%, sensitivity of 96.12%, and specificity of 98.01% on the Shenzhen and Montgomery chest X-ray datasets. These results suggest that the method is effective at identifying intricate TB characteristics like cavitation, shape irregularities, and tissue density. However , the reliance on randomness introduces variability in results, potentially affecting the model's consistency. Additionally, the study failed to evaluate its system across a broader range of datasets, nor did it benchmark its performance against expert radiologists or other existing AI-based methods.

2) *Towards Automated Tuberculosis Detection Using Deep Learning"* by Sonaal Kant and Muktabh Mayank Srivastava (2018).

The authors explore the application of deep learning techniques for tuberculosis detection based on microscopy images. Rather than relying on chest X-rays, their approach highlights the automatic identification of *Mycobacterium tuberculosis* bacilli using a custom Convolutional Neural Network (CNN) architecture. The model obtaines a recall of 83.78% and a precision of 67.55%, showing potential for reducing manual diagnostic workload and improving consistency in TB screening. However, the system was trained on a relatively small dataset, and its low precision led to a evaluated incidence of false positives. These factors limit the model's clinical applicability and scalability, indicating the need for more fine-tuning and validation on more diverse and larger datasets.

3) *An Enhanced DenseNet-Based Deep Learning Framework for Detecting Tuberculosis in Chest Radiographs, Vo Trong Quang Huy and Chih-Min Lin introduced CBAMWDNet.*

A deep neural architecture that integrates DenseNet with the Convolutional Block Attention Module (CBAM) to enhance feature extraction and improve diagnostic accuracy. The model yielded excellant performance, with an accuracy of 98.80%, sensitivity of 94.28%, and specificity of 95.7%, outperforming many existing architectures. Its enhanced adaptability across multiple datasets indicates its potential for real-world TB detection applications. However, the model's high computational demands make it less practical for deployment in resource-limited environments. Additionally, it depends on large datasets for optimal performance and lacks evaluation in real-time diagnostic settings.

4) *Accurate Tuberculosis Detection in Chest X-Rays Through Deep Learning, Segmentation and Visualization, Tawsifur Rahman and colleagues.*

Explored the consequences of lung segmentation and interpretability techniques on TB detection accuracy. The researchers employed nine pre-trained convolutional neural networks (CNNs) and applied them specifically to segmented lung regions rather than entire chest X-ray images. This approach led to improved diagnostic performance, with DenseNet201 achieving a notable 98.6% accuracy on segmented data. To enhance model transparency, visualization tools like Score-CAM were used, confirming that the networks focused on clinically relevant lung areas during predictions. Although segmentation contributed contributing to increased accuracy and improved interpretability, it introduced added computational complexity and required high-quality, annotated datasets. Therefore, the models may encounter hurdles when applied to low-quality or diverse X-ray inputs.

5) *A Deep Learning Approach to Tuberculosis Detection Using Fine-Tuned VGG16 on Chest X-Ray Images, by Ritu Rani and Sheifali Gupta.*

Proposed a TB detection model based on the VGG16 deep learning architecture. The model remained initially learened on the ImageNet dataset, after which the early convolutional layers were kept static to retain general image features while focusing fine-tuning efforts on deeper layers that capture TB-specific patterns. To enhance classification, the network was extended with a flattening layer succeeded by a binary dense classifier. Further improvements were achieved by selectively unfreezing certain layers, supporting the model in learning more detailed representations related to TB. The final system achieved a exactness of 0.98 and an aggregate accuracy of 98%, demonstrating strong potential for use in automated TB diagnosis. However, the VGG16 architecture is computationally intensive, making it less suitable for mobile or low-power devices.Additionally, the study lacks external validation and explainability features, which could hinder trust and adoption in clinical environments.

6) *Automated Pneumoconiosis Detection Using Multilevel Deep Features Extracted from Chest X- Ray Radiographs, by Devnath and colleagues.*

Proposed a hybrid diagnostic model that combines a **Convolutional Neural Network (CNN)** for features extracted via SVM for classification purposes . The CNN was designed to capture **multi-scale spatial features** from chest X-ray images, which allowed the model to learn detailed patterns relevant to pneumoconiosis detection. These layered features were then used by the SVM to classify the presence or absence of the disease. This approach outperformed various classic machine learning approaches and showed strong diagnostic accuracy. However, the model appeared to be specialized for pneumoconiosis and may not generalize well to other lung conditions such as tuberculosis. Additionally, the system lacked interpretability features—an important aspect for clinical adoption—and the CNN-SVM architecture may face challenges in scalability for large-scale deployment.

7) *A Novel Method for Tuberculosis Detection in Chest Rays Using Deep Neural Features Optimized with Artificial Ecosystem-Based Algorithms*,( Sahlol and colleagues.)

Introduced integrated model that merges the lightweight MobileNet architecture with Artificial Ecosystem-Based Optimization (AEO) for enhanced identification of TB in chest X-ray images. The approach focuses on enhancing feature selection by applying AEO to the deep features extracted by MobileNet, thereby boosting classification accuracy without heavily increasing computational demands. This fusion of deep learning with bio-inspired optimization proved effective in refining relevant feature sets and maintaining efficiency. However, the use of metaheuristic algorithms like AEO adds complexity to the system and may result in inconsistent performance across varying datasets. Additionally, necessity for meticulous parameter adjustment and expert oversight poses challenges for scalability and routine clinical deployment.

8) *Deep Learning-Based Image Enhancement for Tuberculosis Detection.*, Munadi et al.

Explored how image preprocessing techniques can improve the effectiveness of TB detection models. The authors applied several enhancement methods—such as Unsharp Masking (UM), High-Frequency Emphasis Filtering (HEF), and Contrast Limited Adaptive Histogram Equalization (CLAHE)—to chest radiographs to make lung abnormalities more visually distinct. These upgraded images were then utilized to train a deep learning model, which achieved enhanced classification accuracy relative to models trained on unprocessed images. The study demonstrated that thoughtful preprocessing can significantly impact the performance of deep learning systems in medical image analysis. However, the added enhancement steps increase computational requirements, potentially limiting the model's suitability for real-time applications. Moreover, certain preprocessing techniques may unintentionally distort critical image features, their performance may change according to the quality and type of imaging equipment used.

9) *A Deep Learning-Based Approach for Tuberculosis Identification Using Chest Radiography* , Hooda and colleagues.

Proposed a deep convolutional neural network (CNN) comprising 27 layers and leveraging connections to identify tuberculosis automatically through chest Radiographs. The model utilized a semi-supervised learning strategy that combined lesion-level localization with whole-image classification, allowing it to capture both detailed and high-level patterns associated with TB. By incorporating residual connections, the architecture addressed vanishing gradient issues and enabled deeper network training. While the model demonstrated strong learning capabilities, its computational complexity and risk of overfitting—especially on smaller or imbalanced datasets—pose challenges. Moreover, the lack of model interpretability and the necessity for high computaional resources limit its practical use in real-world clinical settings, particularly in low-resource environments.

10) *A Novel Stacked Model Generalization for Enhanced Tuberculosis Detection in Chest Radiographs*, Rajaraman and colleagues.

Presented a fusion-based leaning strategy that brings together the predictive power of multiple pre-trained CNN architectures, incorporating VGGNet , ResNet and GoogleNet. The standalone model outputs were aggregated applying a Support Vector Machine (SVM) to make the final classification, resulting in enhanced accuracy and robustness. This stacked ensemble method improved generalization across diverse test sets by leveraging the unique strengths of each architecture. However, the approach is computationally intensive, requiring significant processing power and memory, which makes it less suitable for real-time or resource-limited deployment. Additionally, the ensemble's complexity reduces interpretability, making it difficult for medical professionals to trace the reasoning behind specific predictions. The process of training, managing, and updating such multi-model systems also involves a higher operational burden.

## III. METHODOLOGY

### *A. Dataset Preparation*

In this study, the TBX11K dataset is used, which contains chest X-ray images categorized into three groups: healthy individuals, patients with non-TB illnesses, and confirmed TB cases. To clean the data and remove visual noise such as unwanted dark spots or machine annotations, preprocessing techniques like image rescaling and cropping are applied. To make the model more robust and to reduce overfitting, techniques of data augmentation are used. These include flipping the images, zooming in or out, rotating, shifting horizontally or vertically, and applying slight distortions. These transformations help simulate real-world variations. The overall dataset is then divided into three subsets—80% for training the model, 10% for validating it, and the remaining 10% for testing its performance on unseen data. This ensures balanced learning and fair evaluation.

### *B. Model Architecture*

The proposed model combines two powerful architectures: the Data-efficient Image Transformer (DeiT) and a smaller ResNet-16 network. The DeiT module helps the model understand broad patterns and structures in the chest X-rays by using a self-attention mechanism, which captures relationships between distant parts of the image. ResNet-16 complements this by focusing on finer local features, especially helpful in identifying subtle signs of tuberculosis in lung regions. Its residual connections allow deeper learning without losing important information. This combination ensures the model captures both overall structure and detailed features, making it highly efficient. It also reduces the number of parameters compared to heavier models, improving its speed and adaptability. These components collectively create a compact and powerful hybrid network.

### *C. Training Strategy*

The model is developed using the Adam optimizer, which is well-suited for deep learning because of its capability to adjust learning rates dynamically. The starting learning rate is set to 0.001 to allow steady and stable progress. To prevent overfitting—especially in deep networks—a dropout layer with a 0.2 rate is used, stochastically reactivating neurons while training. The training lasts for 15 epochs, which balances learning with training time. Fine-tuning is performed using hyperparameter optimization techniques, adjusting settings like batch size and learning rate to maximize accuracy. Early stopping is also used to halt training if no improvement is observed over time, saving computational resources. Combined, these components form a compact but highly effective hybrid network..

### *D. Feature Extraction and Visualization*

Once the model appears to be trained, feature extraction is performed using a global average pooling(GAP) layer, which condenses each feature map into a single number. This helps minimize model size while maintaining critical information needed for classification. To make the model's decision-making process more transparent, heatmaps are generated using Class Activation Mapping (CAM). These visualizations highlight which regions of the X-ray the model focuses on when making predictions. This interpretability is vital in medical applications, as it allows healthcare professionals to understand and trust the model's reasoning. Such visual tools can also serve as a second opinion in clinical diagnoses. Overall, this approach balances performance with explainability, which is often lacking in deep learning models.

### *E. Performance Indicators*

The proposed model is tested using several key metrics to get a complete picture of its effectiveness. Accuracy measures how many total predictions were correct, while sensitivity (or recall) focuses on the model's ability to detect TB-positive cases correctly. Specificity measures how well the model avoids false positives by correctly identifying healthy individuals, and precision tells us how many predicted TB cases were actually correct. In addition to these, the count of parameters is tracked to assess model complexity, and the duration required to examine each image is measured to determine efficiency. Collectively, these matrix validate that the model is accurate and well-suited for real-time deployment. It shows promising results across all these evaluation areas.

### *F. Comparison and Validation*

To show the efficacy of the hybrid approach, its performance is compared against several well-known CNN models like VGG16, ResNet, and DenseNet, all trained based on the same dataset. The developed model consistently outperforms them in accuracy, recall, and precision, demonstrating its superiority.

Graphical results and statistical tests support these claims, showing that the improvements are significant and not by chance. Moreover, the hybrid model shows better generalization, meaning it performs well even on data it hasn't seen before. While many deep learning models struggle with consistency, this one maintains stable performance across different test cases. The results suggest that the proposed architecture is not only innovative but also practical for clinical use.

## IV. DISCUSSION

Integrating neural network model for tuberculosis identification has brought a fresh perspective to medical image analysis. By combining architectures like ResNet and DeiT, the system takes advantage of both detailed local features and broader image context, improving its ability to identify TB accurately from chest X-rays. This hybrid approach has shown stronger performance compared to using a single model, especially when dealing with complex or subtle lung patterns. The use of heatmaps to visualize decision areas also helps clinicians understand what the model is focusing on, making it more trustworthy for real-world use. However, there are still some hurdles. The system requires significant computing power, and building reliable models depends heavily on having access to large, well-labeled datasets. Additionally, different hospitals use different X-ray machines, which might affect model performance. To make this technology more practical, future reasearsh should focus lightweight, adaptable models that work well even in low-resource areas. Improving explainability and testing across varied patient groups will also be important for building confidence in clinical settings.

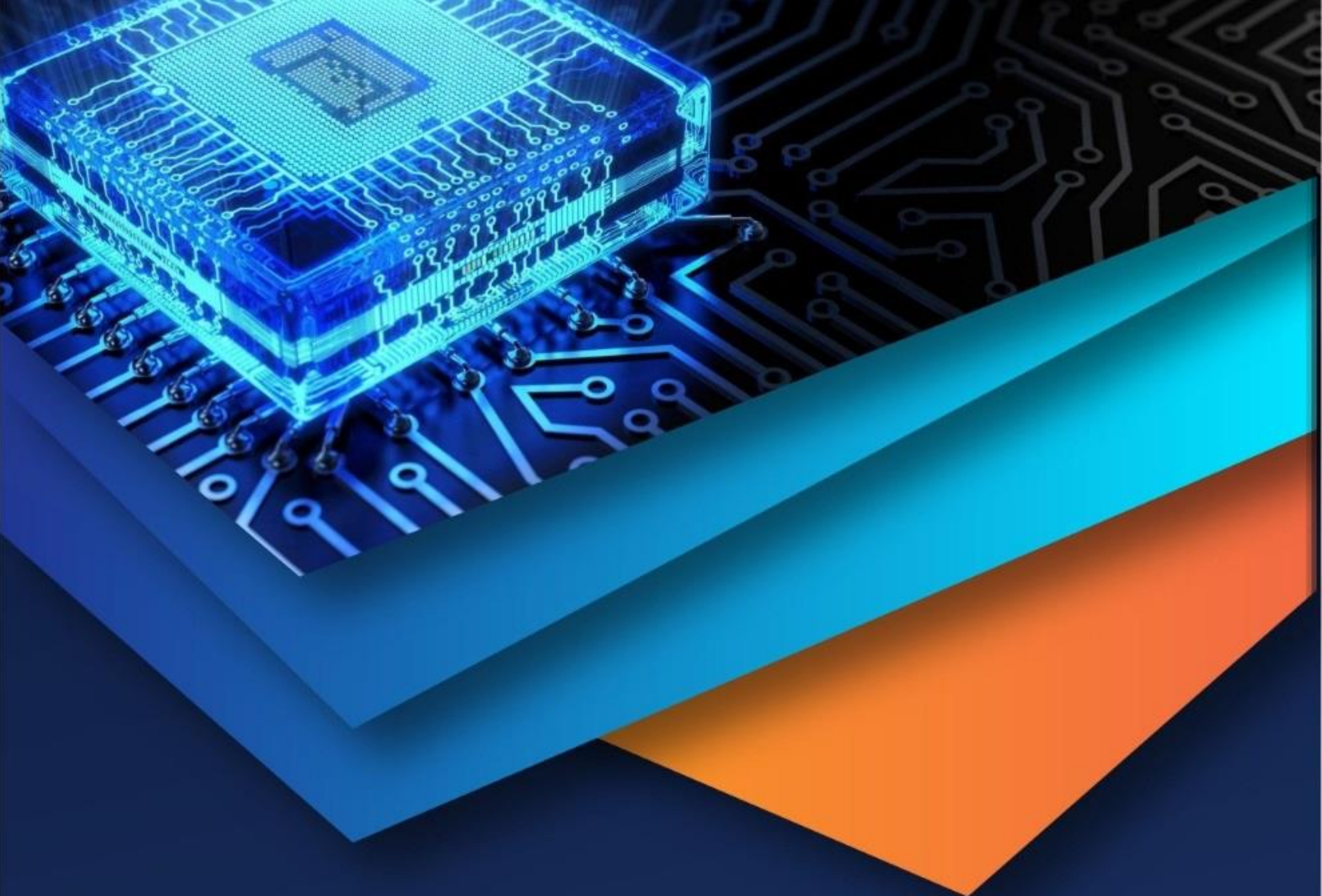
.

## V. CONCLUSION

Detecting tuberculosis through deep learning has shown great promise in improving diagnosis, especially in areas where medical resources are limited. The integration of advanced models like CNNs and transformers has led to better exactness and potential to visualize decision areas using heatmaps. These improvements not only make the system more effective but also help build trust by making the results easier to interpret. Techniques like data augmentation and feature optimization have further improved the model's ability to handle diverse medical images.Still, there are some challenges that need attention. The requirement for large labeled datasets, high processing power, and inconsistent image quality across sources can affect model performance. Despite these issues, the findings indicate that AI-powered tools can become valuable aids for doctors in diagnosing TB. Future research should aim to make these systems lighter, more transparent, and easier to apply in real-world clinics. With continuous development and support from healthcare professionals and researchers, deep learning can play a major role in reducing the burden of tuberculosis globally.

## REFERENCES

[1] E. Kotei and R. Thirunavukarasu, "Tuberculosis Detection From Chest X-Ray Image Modalities Based on Transformer and Convolutional Neural Network," IEEE Access, vol. 12, pp. 97417–97427, Jul. 2024, doi: 10.1109/ACCESS.2024.3428446.
[2] World Health Organization, "Tuberculosis (TB)," WHO, 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/tuberculosis
[3] C. S. Preetham and N. C. Krishnan, "Deep learning approaches for tuberculosis detection using chest X-ray images," Computers in Biology and Medicine, vol. 137, p. 104783, 2021.
[4] A. K. Jaiswal, P. Tiwari, A. Kumar, S. Gupta, and V. Khanna, "Classification of tuberculosis bacteria using deep convolutional neural network," Journal of Medical Systems, vol. 43, no. 7, pp. 1–7, 2019.
[5] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, "Training data-efficient image transformers & distillation through attention," in Proc. Int. Conf. Machine Learning (ICML), pp. 10347–10357, 2021.
[6] 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), pp. 770–778, 2016.
[7] H. T. Goh, B. Ong, and L. Wang, "TBX11K: A large chest X-ray dataset for tuberculosis screening," arXiv preprint arXiv:2102.07661, 2021.
[8] L. Devnath, S. Luo, P. Summons, and D. Wang, "Automated detection of pneumoconiosis with multilevel deep features learned from chest X-ray radiographs," Computers in Biology and Medicine, vol. 129, Art. no. 104125, Feb. 2021.
[9] A. T. Sahlol, M. A. Elaziz, A. T. Jamal, R. Damaševičius, and O. F. Hassan, "A novel method for detection of tuberculosis in chest radiographs using artificial ecosystem-based optimisation of deep neural network features," Symmetry, vol. 12, no. 7, p. 1146, Jul. 2020.
[10] K. Munadi, K. Muchtar, N. Maulina, and B. Pradhan, "Image enhancement for tuberculosis detection using deep learning," IEEE Access, vol. 8, pp. 217897–217907, 2020R. Hooda, S. Sofat, S. Kaur, A. Mittal, and F. Meriaudeau, "Deep-learning: A potential method for tuberculosis detection using chest radiography," in Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA), Sep. 2017, pp. 497–502
[11] S. Rajaraman et al., "A novel stacked generalization of models for improved TB detection in chest radiographs," in Proc. IEEE EMBC, Jul. 2018, pp. 718–721
[12] M. Ayaz, F. Shaukat, and G. Raja, "Ensemble learning based automatic detection of tuberculosis in chest X-ray images using hybrid feature descriptors," Physiological Engineering and Science in Medicine, vol. 44, no. 1, pp. 183–194, Mar. 202
[13] Vision Transformer (ViT) - Attention is All You Need for Vision https://www.youtube.com/watch?v=1fLE3J66U68
[14] Tuberculosis Detection using Deep Learning in Python | CNN Project https://www.youtube.com/watch?v=zCz2i2LE2b8
[15] Convolutional Neural Networks (CNNs) Explained https://www.youtube.com/watch?v=YRhxdVk_sIs



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY