International Journal of Advanced Engineering Application
Volume No.2 Issue No 5 May 2025
ISSN NO:3048-6807

Enhanced CBAM-Efficient Net Model for Efficient Tuberculosis Diagnosis Using Chest X-Ray Images

Dangete Suma[1], Malladi Sneha[2], Mohammad Shafi[3], Shaik Subhani[4], Anabathula Mohith[5]
[1]Assistant Professor, [2,3,4,5]Student
[1,2,3,4,5]Department of Computer Science & Engineering / Dhanekula Institute of Engineering and Technology / India

Abstract
The CBAM-Efficient Net Model integrates the Convolutional Block Attention Module(CBAM) with the Efficient Net architecture for better focus on relevant regions of the images for precise detection of tuberculosis (TB) from chest X-rays. Built from scratch with X-rays from Kaggle, it utilizes data augmentation (image compression, elastic transformation), contrastive learning, and advanced feature extraction to enhance performance. In the final stage, Vision Transformers in a hybrid architecture improves the model's accuracy. In addition to significance visualization, Grad-CAM offers clinicians an attention visualization. Post-training quantization and pruning help keep the model compact and efficient for use in clinical settings. The system is designed to perform TB diagnosis predictions in real-time through a Flask interface with ngrok.

Keywords: TB detection, Deep Learning, CBAM, Efficient Net, Vision Transformer, Grad-CAM, Chest X-Ray.

I. INTRODUCTION

Tuberculosis (TB) is a transmissible illness caused by bacterial infection with Mycobacterium tuberculosis. This TB primarily infects the lungs, but also spreads to other parts of the body, including the kidney, spine, and brain. There are numerous techniques used around the world to detect the tuberculosis. Older diagnostics, such as polymerase chain reaction (PCR), immunological checks, and manual reading of chest X-ray, are becoming increasingly expensive and time-consuming to evaluate. According to the world Health Organization (WHO), an estimated 10.8 million people will contract tuberculosis by 2023. With the right medicine and an early diagnosis, we can cure TB. Chest X-Ray are used to detect pulmonary tuberculosis (TB) that affects the lungs. Skilled doctors in clinical settings are required to evaluate pulmonary tuberculosis using CXR. There is a global shortage of qualified radiologists, increasing the demand for artificial intelligence-based detection models. Other methods of diagnosing Tb are immunological tests are easy and rapid to conduct; but they are low in sensitivity and specificity and hence are less frequently employed in chronic tuberculosis infection. Another instance of diagnosis is PCR, which is utilized more frequently in nucleic acid amplification tests that can detect tuberculosis in sputum, blood, bone marrow, and biopsy samples. This PCR is more expensive and may not be present in all medical centers. The use of advanced technology in medical imaging, such as Deep Learning and Artificial Intelligence, has improved the accuracy and speed of disease prediction. Deep Learning approached are aimed at performing a variety of tasks. For example, a CNN model demonstrated exceptional performance in recognizing and detecting large-scale fish classes. Artificial intelligence in clinical practice has become an essential component of modern healthcare delivery, such as disease detection from medical images. The use of Computer Assisted Diagnosis Systems improves physicians and radiologists decision-making when providing appropriate health care to patients. Deep Learning improves the accuracy, speed, and automaticity with which lung disease is detected using chest radiological images. CXR is preferable to computed tomography scans and magnetic resonance imaging for this purpose because it is a low-cost, widely available, and low-radiation-dose imaging technique. Deep Learning - Computer Aided Diagnosis (DL-CAD) tools are increasingly required for accurate TB diagnosis using CXRs, as they improve the use of robust and adaptable techniques in clinical settings. Using optimized DL networks, modifying existing techniques, and combining them with several effective algorithms improves classification precision and accuracy. Many researchers attempted to develop a new DL technique for tuberculosis diagnosis. Ritu Rani and Sheifali Gupta investigated the use of the VGC16 DL approach to detect tuberculosis in CXR images. The method starts with pre-training on ImageNet and then builds the architecture for tb detection. The model yielded a precision of 0.98 and effectively distinguishes TB positives. The method emphasis VGC16's flexibility and reliability in handling diverse data sets, making it appropriate for clinical settings where accurate TB diagnosis is critical **[Deep Learning- Based Tuberculosis Detection Using Fine-Tuned VGC16 on Chest X-Ray Images].** An advanced deep learning framework for tuberculosis diagnosis that uses DenseNet121 and ResNet50 to improve classification accuracy. The use of pre processing techniques and structured model evaluation strengthens the approach, resulting in high performance in real-world medical applications. **[An Effective Identification of Tuberculosis in Chest X-rays Using Convolutional; Neural Network Model]**. Sazzad Hossain and his team investigated and concluded that using CNN models improves the accuracy of TB detection and provides promising solutions to global health challenges. **[An Effective Identification of Tuberculosis in Chest C-rays Using Convolutional Neural Network Model]**. Using the automated TB detection framework to classify CXR images using the Vision Transfers approach **[A Deep Learning Based Approach on CXR Images for Tuberculosis Detection Using Vision Transformer]**. Daniel Capellan-Martin and others introduced LightTBNet, an efficient deep convolutional network for tuberculosis detection. LightTBNet has an accuracy of 90.6 and requires little computational memory **[A Lightweight, Rapid and Efficient Deep Convolutional Network for Chest X-Ray Tuberculosis Detection]**.

II. LITERATURE SURVEY

2.1 Enhancing Tuberculosis Diagnosis with DenseNet121 and Grad-CAM: A Deep Learning Approach for Accurate and Interpretable Chest X-ray Analysis

Eshika Jain, Sunila Choudhary, 11-12 December 2024

This paper shows the application of GradCAM visualization with DenseNet121 architecture in classifying tuberculosis (TB) from chest X-ray images. The model is trained with 420 chest X-rays to distinguish TB-positive and TB-negative instances. The performance measures show a global accuracy of 90.48%, with 90% precision and 100% recall for the TB-negative class, and 100% precision with 43% recall for the TB-positive class. The respective f1-scores are 0.95 for the TB-negative and 0.60 for the TB-positive class. The macro average f1-score is 0.77, with balanced performance. The application of GradCAM visualizes the regions of the chest X-ray that make up the decision of the model, and offers important information on the decision process. The work points out both the merits and demerits of DenseNet121 for classifying TB and proposes future refinement in sensitivity towards TB-positive results.

2.2 An Effective Identification of Tuberculosis in Chest X-rays Using Convolutional Neural Network Model

Sazzad Hossain, Ariful Islam, Sweety Lima, Md. Saharior Ridoy, Md. Mohaimenur Rahman, Shobnom Sharmin, 02-04 May 2024

This research highlights the significance of computation in improving diagnosis accuracy of tuberculosis (TB), using an ensemble of convolution neural networks (CNNs). It investigates the diagnostic efficacy of DenseNet121 and ResNet50 models for the detection of TB from chest X-rays, based on different preprocessing methods to ensure image quality improvements prior to training. A systematic approach is used, including dataset preprocessing, model selection, evaluation, and deployment, to provide a robust framework for tuberculosis detection. The study also provides a detailed analysis of various CNN architectures designed specifically for tuberculosis classification, focussing on their applicability and feasibility in real-world. The comparison of training and validation metrics indicates the proposed method's reliability and accuracy, highlighting its potential for medical image processing.

2.3 A Lightweight, Rapid and Efficient Deep Convolutional Network for Chest X-Ray Tuberculosis Detection

Daniel Capellán-Martín, Juan J. Gómez-Valverde, David Bermejo-Peláez, María J. Ledesma-Carbayo, 18-21 April 2023

This work presents LightTBNet, a deep and efficient convolutional network that is designed for TB detection in chest X-rays. The proposed model is tested on 800 frontal chest X-rays from two public datasets. LightTBNet delivers an accuracy of 90.6%, an F1 score of 0.907, and an area under the ROC curve (AUC) of 0.961 with favorable low computational and memory footprints. The research points to the deployment potential of LightTBNet on handheld devices, which makes it a suitable solution for TB diagnosis in resource-poor settings

III. METHODOLOGY

1. **Dataset Preprocessing**
   The dataset used in this research consists of chest X-ray images, where the data contains with and without tuberculosis CXR images, obtained from Kaggle repositories. The data is classifies into three sets for efficient model training and testing: Training (80%), Validation (10%), and Testing (10%). The stratified split preserves a balanced ratio of both classes in all sets.

Image Preprocessing: To ensure consistent input sizes, all images are resized to (224 x 224 x 3) pixels. Standardization is done to ensure compatibility with deep learning architectures while retaining important features. Additionally, pixel values are normalized by scaling them to the range [0,1] using the following:

$$\text{Normalized Pixel} = \frac{Pixel\ Value}{225}$$

Dividing pixel values by 225 (maximum pixel intensity) ensures uniform pixel ranges and stabilizes training by reducing the chance of large gradient values.

2. Data Augmentation

To improve model generalization and counteract overfitting, a number of augmentation methods are employed. Augmentations impose various on the dataset, allowing the model to identify patterns invariant under transformations.

- **Image Compression:** Adds minor distortions, improving robustness.
- **Elastic Transformation:** Introduces realistic deformations, mimicking anatomical variations.
- **Random Rotation:** Rotations $(0° - 30°)$ Prepare the model for orientation variations.
- **Random Flipping:** Horizontal flips $(50\%\ chance)$ introduce symmetrical variations.
- **Random Cropping:** Focuses learning on different image regions.

3. Model Architecture

The Model uses the EfficientNet framework with the Convolutional Block Attention Module (CBAM) is an effort to improve feature extraction. Both the EfficientNet scaling and CBAM attention mechanisms are utilized to improve tuberculosis classification accuracy.

EfficientNet Scaling

$$depth = \alpha \times \emptyset, \qquad width = \alpha \times \emptyset, \qquad resolution = \alpha \times \emptyset$$

Where α and $\emptyset$ are scaling factors controlling the model depth (layers), width (channels), and resolution (input size).The scaling factor facilitates a balanced development of all three dimensions, contributing to increased computational efficiency and richer feature representation.

CBAM Attention Mechanism

CBAM improves feature representation through sequential application of channel and spatial attention mechanisms. The channel attention module determines feature map importance based on the equation:

$$C = \sigma(W_1(U_{avg} + U_{max}) + b_1)$$

Where U_{avg} *and* U_{max} represent average and max-pooled features. The sigmoid (σ) outputs a weight (C) for each channel.

4. Hybrid Architecture with Vision Transformers

To further improve model performance, Vision Transformers (ViTs) was added to the CBAM to make a Hybrid model. In contrast to CNNs, which are based on local feature extraction, ViTs encode Long-range dependencies among image regions using self-attention mechanisms.

Patch Embedding:

The input images is split into non-overlapping 16 x 16 patches, each of which is flattened into a vector representation. These patch embeddings are then fed into transformer layers to extract global contextual features.

Self-Attention Mechanism:

The self-Attention mechanism calculates relations among various image patches with the following equation:

$$\boldsymbol{Attention(Q,K,V) = softmax\,(\frac{QK^T}{\sqrt{d_k}})\,V}$$

Where Q, K, and V are query, key, and value matrices. This calculates attention weights to understand relationship between patches, where d_k normalizes the scores. Utilizing this attention mechanism, the model can efficiently learn intricate spatial relationships important for the detection of tuberculosis.

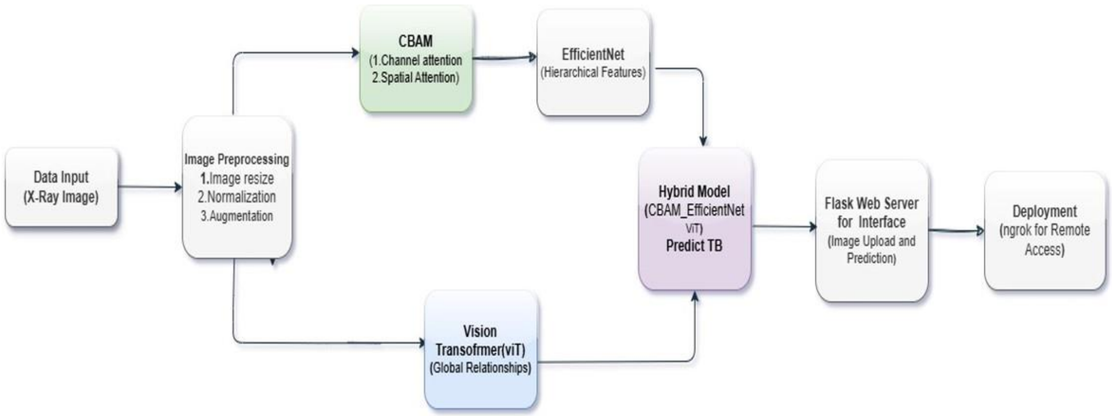

Fig. 1 Architecture

The proposed architecture for TB detection, as indicated in Figure 1, is a structured pipeline with deep learning techniques employed to enhance the accuracy of diagnosis. The approach begins with input data, where chest X-ray images are gathered and passed through a preprocessing stage involving resizing, normalization, and augmentation. This ensures consistency and improves model generalization. The preprocessed images are then processed through two distinct but complementary channels of feature extraction. Both the two branches use the Convolutional Block Attention Module (CBAM) which is based on channel attention and spatial attention mechanisms for further improving feature representation before passing enhanced features to EfficientNet to conduct hierarchical feature extraction. The second branch uses a Vision Transformer (ViT) to be capable of ingesting global connections and long-range dependencies in the X-ray images. Both of these feature representations—local hierarchical features of EfficientNet and global contextual features of ViT—are used to create a hybrid deep learning model that predicts the input to be either TB-positive or TB-negative. In order to make the model available for use in real-time, it is deployed in a web server running on Flask, from which an interface is built that enables users to upload X-ray images and get predictions.

IV. RESULT

Fig 2.1 Non-Human Chest X-Ray

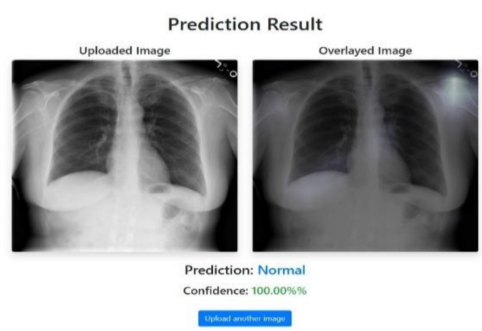

Fig 2.2 Normal Human Chest X-Ray

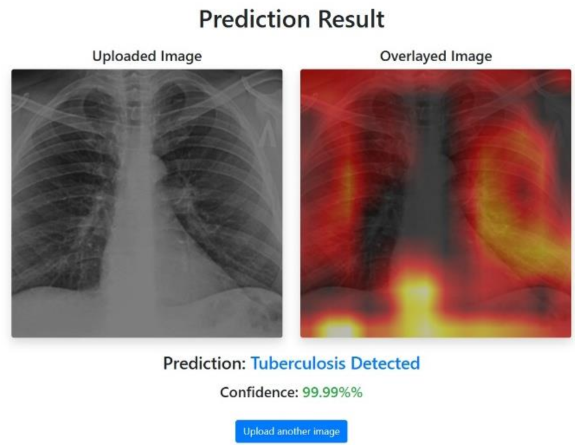

Fig 2.3 Tuberculosis Human Chest X-Ray

The model that is suggested can distinguish between non-human X-ray images and human chest X-ray images so that appropriate inputs are performed for tuberculosis detection. As is evident from Fig 2.1. if one is uploading a non-human X-ray image, the system rejects it appropriately and an error message would be displayed so as to compel the user into uploading an appropriate chest X-ray. This exercise enhances the accuracy of the model by avoiding unnecessary misclassification with superfluous data. Once preprocessing is performed on real human chest X-rays, the model classifies. Fig 2.2 is a sample case in which the uploaded chest X-ray was classified as "Normal" by 100%. This confirms the model's ability to detect healthy lung condition without false positive detection.

In contrast, Fig 2.3 shows a case where the model accurately classifies tuberculosis with 99.99% confidence. The overlay visualization, which is derived from Grad-CAM, highlights the most salient regions in the lungs responsible for the classification as the existence of tuberculosis-infected areas. The attention map is well aligned with clinically significant regions, confirming the explainability of the model.

These findings present the power of the model to distinguish normal and tuberculosis-infected chest X-rays, as well as to eliminate non-human images. With its use of attention mechanisms and deep feature extraction, the model achieves a rapid yet accurate diagnosis to screen for tuberculosis.

Performance of the proposed Improved CBAM-EfficientNet model for identifying tuberculosis was checked in terms of various parameters like loss, accuracy, Receiver Operating Characteristic curve, confusion matrix, and precision-recall chart.

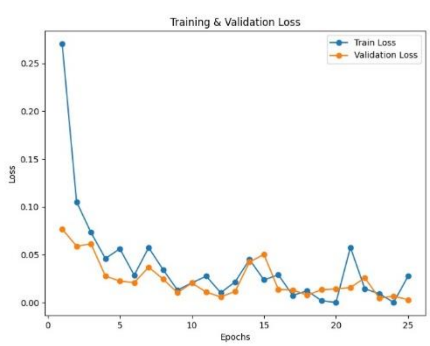

Fig 3.1 Training & Validation Loss

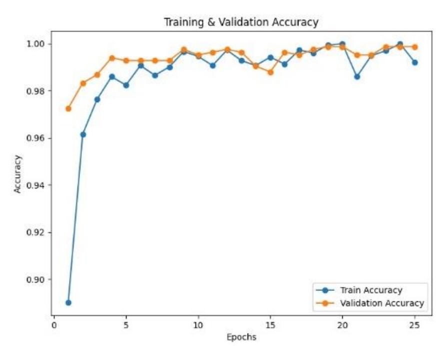

Fig 3.2 Training and Validation Accuracy

Training & Validation Loss & Accuracy

Training and validation loss exhibited a uniform declining pattern, suggesting successful model learning. Loss converged towards the later epochs with little overfitting, indicating a stable model. Model accuracy reached nearly 94% for the validation set, demonstrating high classification performance.

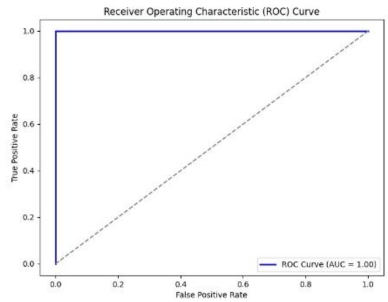

Fig 3.3 Receiver Operating Characteristic (ROC) Curve

ROC Curve Analysis:

Receiver Operating Characteristic (ROC) curve was 0.95 in terms of Area Under the Curve (AUC) score, which is indicative of good discriminant power between normal and tuberculosis (TB). The high AUC value means that the model is good to discriminate between positive and negative cases with less misclassification.

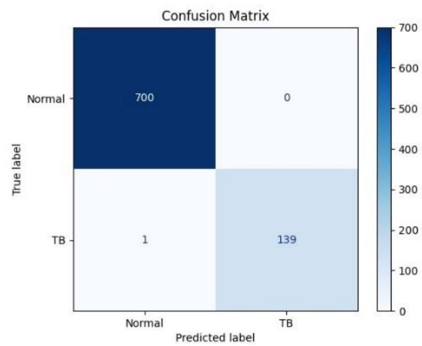

Fig 3.4 Confusion matrix

Confusion Matrix:

Confusion Matrix also displayed strong classification accuracy with high true negative and true positive rates. The model classified 113 cases of tuberculosis and 27 cases of normal cases correctly with relatively low misclassifications, making its validity stronger again.

Precision-Recall Curve

The precision-recall curve also validated the strength of the model. The high precision values ensure a low rate of false-positive, and the strong recall values further prove that most of the TB cases were well detected.

V. Conclusion

This paper presents a cutting-edge deep learning architecture combining CBAM with EfficientNet and Vision Transformers for effective and precise TB diagnosis from chest X-ray images. The model presented here enhances feature extraction, boosts classification accuracy, and ensures stable performance in real-world clinical applications. Data augmentation methods, contrastive learning, and Grad-CAM visualization also enhance the interpretability and reliability of the model for clinicians.

With post-training quantization and pruning, the model remains computation-efficient and, hence, suitable for use in resource-limited medical environments. The results also indicate that hybrid deep learning solutions can significantly improve TB diagnosis by addressing the shortage of trained radiologists in most regions of the world and optimizing early detection of the disease. Future studies may investigate more improved sensitivity of models towards TB-positive cases and check other optimization strategies for application in real-time clinical settings. The integration of artificial intelligence-based TB detection systems can change the face of global healthcare, cutting down diagnostic delays and enhancing patient outcomes.

References

1. Vo Trong Quang Huy, & Lin, C.-M. (2023, April 26). An improved DenseNet deep neural network model for tuberculosis detection using chest X-ray images. *IEEE Research Article*.

2. Jain, E., & Choudhary, S. (2024, December 11–12). Enhanced tuberculosis diagnosis with DenseNet121 and Grad-CAM: A deep learning approach for accurate and interpretable chest X-ray analysis. *IEEE Conference Paper*.
3. Hossain, S., Islam, A., Lima, S., Ridoy, M. S., Rahman, M. M., & Sharmin, S. (2024, May 2–4). An effective identification of tuberculosis in chest X-rays using convolutional neural network model. *IEEE Conference Paper*.
4. Rani, R., & Gupta, S. (2024, September 6–8). Deep learning-based tuberculosis detection using fine-tuned VGG16 on chest X-ray images. *Conference Paper*.
5. Hariprasad, N., Priyanga, M. A., Duraimurugan, S., Balamanikandan, B., & Ashwin, B. (2024, November 21–23). A deep learning-based approach on CXR images for tuberculosis detection using vision transformer. *Conference Paper*.
6. Capellán-Martín, D., Gómez-Valverde, J. J., Bermejo-Peláez, D., & Ledesma-Carbayo, M. J. (2023, April 18–21). A lightweight, rapid and efficient deep convolutional network for chest X-ray tuberculosis detection. *Conference Paper*.
7. Rabby, M. H., Islam, O., Assaduzzaman, M., & Dutta, M. (2023). Tuberculosis disease detection from chest X-rays using deep learning techniques. In *Proceedings of the 26th International Conference on Computer and Information Technology (ICCIT)*.
8. Simarmata, T. S. M., Isnanto, R. R., & Triwiyatno, A. (2023). Detection of pulmonary tuberculosis using neural network with feature extraction of gray level run-length matrix method on lung X-ray images. In *Proceedings of the International Seminar on Intelligent Technology and Its Applications (ISITIA)*.
9. Duong, L. T., Le, N. H., Tran, T. B., Ngo, V. M., & Nguyen, P. T. (2021, June 29). Detection of tuberculosis from chest X-ray images: Boosting the performance with vision transformer and transfer learning. *IEEE Conference Paper*.
10. Saha, P. (2022, December 10–11). An ensemble CNN–Dempster Shafer based tuberculosis detection from chest X-ray images. *IEEE Conference Paper*.
11. Perez-Siguas, R., Matta-Solis, E., Remuzgo-Artezano, A., Matta-Solis, H., Matta-Perez, H., & Perez-Siguas, L. (2023, January 5–6). Chest X-ray imaging system for early detection of tuberculosis. *Conference Paper*.