

Deep Hybrid Framework for Tuberculosis Detection Using Attention-Based EfficientNet and Optimized FE-MobileNet

Dr. Ishani Mishra¹, Renuka Kondabala², Kanamala Suresh³, Krishna Suresh B V N V⁴, Dr. P. Rupa Ezhil Arasi⁵

¹Department of ECE, New Horizon College of Engineering, Bangalore

²Assistant Professor, Department of Information Technology, Nageswara Rao Vignana Jothi Institute of Engineering and Technology, Hyderabad

³Assistant Professor, Sharda School of Computing Science and Engineering, Sharda University, Greater Noida, Uttar Pradesh

⁴Assistant Professor, Department of Computer Science & Engineering, KLEF, KL University, Vaddeswaram, Andhra Pradesh

⁵Professor, Department of AIML, Sri Sairam College of Engineering, Bengaluru

Emails: Ishani.nhce@gmail.com (I.M.), renuka_k@vnrvjiet.in (R.K.), kanamala.suresh@sharda.ac.in (K.S.), krishnasuresh@kluniversity.in (K.S.B.V.N.V.), rupasharansai@gmail.com (P.R.E.A.)

Received: 10th August, 2026; **Revised:** 20th August, 2026; **Accepted:** 22nd August, 2026; **Available Online:** 24th August, 2026

ABSTRACT

One of the most dangerous infections prevailing in today's world is tuberculosis (TB), and therefore, early detection is essential to the successful treatment of TB patients. However, the subjectivity of chest x-rays, low picture quality, and the intricacy of tuberculosis patterns pose problems for the conventional TB diagnosis method. In order to diagnose tuberculosis using chest X-ray images, this study suggests a deep hybrid learning system. In order to make our model robust and generalize well, benchmarking datasets, including Image CLEF and Montgomery County X-ray datasets, are utilized. Chest X-ray images obtained in the dataset are first pre-processed to remove any distortion, increase image quality, and normalize images. During the feature extraction stage, channel and spatial information are utilized to generate discriminatory deep features using an effective Net enhanced with Convolutional Block Attention Module (CBAM). Subsequently, the generated deep features are fed into the designed Feature Enhanced Mobile Net (FE-Mobile Net), where the Feature Enhancement Block (FEB) is introduced to enhance the feature extraction capabilities and increase the classification accuracy. In addition, GWO is employed to optimize the hyperparameters to increase the learning efficiency and performance of the proposed classification model. Finally, the chest X-ray pictures are divided into two classes (Normal and TB) by the improved FE-Mobile Net model. The findings show that the suggested deep learning model is effective at identifying tuberculosis from CXR pictures, providing a sophisticated computer-aided diagnosis model for medical applications.

Keywords:- Tuberculosis, Mobile Net, Efficient Net, Xray, Gray wolf optimizer, Convolutional Block Attention Module, FE-Mobile Net and Chest X-ray

How to cite this article: Mishra I, Kondabala R, Suresh K, B V N V KS, Ezhil Arasi PR. Deep Hybrid Framework for Tuberculosis Detection Using Attention-Based EfficientNet and Optimized FE-MobileNet. Int J Drug Deliv Technol. 2026;16(66s):339-394. DOI: 10.25258/ijddt.16.66s.39

1. INTRODUCTION

One of the global public health problems that plague our planet today is tuberculosis (TB). The cause of the airborne disease is Mycobacterium tuberculosis (MTB), first discovered in 1882. Due to initiatives taken globally to improve TB prevention and treatments, mortality due to TB has been falling around the globe. Nonetheless, despite advancements in medicine and healthcare,

Tuberculosis is still considered one of the leading causes of death from an infectious disease globally. About 9.6 million people suffer from Tuberculosis (TB), with 1.5 million people having died due to TB in the year 2014 alone.[1] It takes about 6 months for successful treatment of TB disease, and for DR-TB treatment, it takes 9 to 18 months [2]. In cases where TB is detected on time and managed properly, there are chances that it can be cured. Chest X-rays (CXRs) are frequently used

to screen for and diagnose pulmonary tuberculosis. In the field of medicine, chest X-rays are usually evaluated by trained health practitioners for the purpose of diagnosing TB [3]. The use of CXRs to detect the disease is often subjective and sometimes ineffective. There is no doubt that there might be subjectivity in diagnosing disease through CXRs. The CXR pictures of TB patients are sometimes confused for other lung diseases of a similar pattern [4].

The challenge associated with establishing the resistance of the suspected TB patient's body to the medications on the first visit is another major hindrance that inhibits the management of MDR-TB [5]. To establish the resistance of a sputum sample to several drugs, drug susceptibility tests must be performed. However, the process requires a properly equipped lab setting and takes four to six weeks to obtain lab results. Sequencing is commonly employed to determine drug susceptibility through the identification of mutation. The techniques for predicting drug susceptibilities using sequencing can be broadly classified as either DA (direct association) or ML (machine learning). DA tools include KvarQ, CASTB, MyKrobe Predictor TB, PhyResSE, TGS-TB, TBProfiler, and SAM-TB [6]. However, these tools are based on known resistance mutations, with several shortcomings such as unidentified mutations, lack of ability to model gene-gene interactions, and poor prediction power as MDR rises. On the other hand, ML techniques utilize both sequencing and DST data to predict antibiotic resistances with the help of logistic regression, random forest, decision tree, and gradient boosting algorithms. Deep learning includes CNNs, MLPs, Wide & Deep models, auto-encoders, graph attention networks, and attention-based neural networks [7].

Even with innovations in imaging and molecular diagnostics, there are several challenges associated with the diagnosis and management of TB infections. Drug susceptibility tests are time-intensive, and chest x-rays interpretation is subjective. Machine learning algorithms that include SVM, Random Forest, Decision Trees, and Neural Network Model (NNM) can be used to diagnose TB and predict drug resistance [8]. Deep learning algorithms, which include Convolutional Neural Networks (CNN), MLP, and attention models, offer more accurate feature extraction and classification performance [9]. The use of pre-planning and optimization models will also help decision-makers to determine the best strategies for treatment. These machine learning models demand large data sets and exhibit poor generalization and transparency properties. Integration of data from imaging and genomics in these models is quite low.

Although there have been tremendous advancements made by the current TB diagnosis systems, several limitations exist that hinder their efficiency. For example, because of inadequate feature extraction and a lack of spatial and channel attention, using classical CNNs may restrict the model's capacity to recognize TB patterns from the chest X-ray pictures. The classification models developed using lightweight networks usually have low feature enhancement capabilities, which make it difficult to classify complicated TB patterns. There is a trend of using pre-defined hyperparameters in most studies, leading to poor model performance and computational costs. Moreover, variability of image quality, noise, and contrast across the benchmark datasets influences the robustness of existing models. This inspired the design of an intelligent and robust deep hybrid framework aimed at identifying TB features and improving classification efficiency. The main contribution of the research is listed below;

- To use chest X-ray (CXR) images to automatically identify tuberculosis (TB), an innovative 'Deep Hybrid Framework' is proposed, which integrates Attention-Based EfficientNet and enhanced FE-MobileNet architectures to improve diagnostic accuracy and robustness.
- A robust preprocessing scheme is presented that involves normalization, smoothing, enhancement, and resizing processes in order to improve the quality of CXR images by decreasing noise and making the pattern of tuberculosis more visible.
- An attention-based feature extraction approach is devised by embedding CBAM in the EfficientNet architecture in order for the network to extract important channel and spatial features related to TB pathology from CXRs.
- Another innovation is the introduction of an innovative Feature Enhancement Block in the proposed FE-MobileNet classifier in order to enhance feature representation ability and make discriminative learning better for improved tuberculosis classification.
- The Grey Wolf Optimization algorithm is utilized for optimizing the hyperparameters of the FE-MobileNet model for faster convergence, increased learning rate, and improved classification performance.
- Finally, two benchmark TB datasets, the ImageCLEF Tuberculosis Dataset and the Montgomery County Chest X-ray Dataset, were used to confirm the efficacy of the developed CAD system for categorization into normal and tuberculosis categories.

2. LITERATURE SURVEY

The automated identification of tuberculosis using chest X-ray images has been significantly improved by recent advances in deep learning and medical imaging. Various machine learning and deep learning techniques have been developed to increase detection accuracy and enable early diagnosis. This part of the paper discusses different approaches to tuberculosis detection and their merits and demerits.

A deep learning-based method for early tuberculosis identification utilizing chest X-ray pictures in Internet of Things-based healthcare applications was suggested by G. Simi Margarat et al. [10]. Preprocessing, feature extraction, segmentation using the Adaptive Fuzzy C-Means clustering, and classification using the DBN trained by the AMBO algorithm comprise the architecture (DBN-AMBO). It is tested using the MC and SC dataset, where it attains an accuracy of 99.2. This makes the algorithm superior to all previous methodologies. The objective of the methodology is to help radiologists diagnose TB patients accurately and early. But the architecture has some flaws such as dependency on hand-crafted feature extraction and parameters optimization, and there is a lack of validation on real-life datasets.

Tawsifur Rahman [11] et al. presented TB-CXRNet, a deep learning framework for detecting tuberculosis (TB) and drug-resistant TB using chest X-ray images. For training purposes, 40,000 images were utilized to distinguish between TB, non-TB, and healthy categories, in addition to 3,037 TB images for drug resistance prediction. The proposed architecture involves two stages: first, detecting TB cases, and second, classifying them into drug-sensitive or drug-resistant groups. A CheXNet encoder with Self-MLP and stacking technique were implemented to enhance efficiency. The interpretation of the model is done using Score-CAM. The results showed that accuracy rates were achieved at 93.32%, 87.48%, and 79.59% for multi-class, binary, and three-class classifications, respectively.

Jong Hyuk Lee et al. [12] introduced a deep learning-based automated detection (DLAD) algorithm for identifying active pulmonary tuberculosis from chest radiographs in large-scale screening. The model was tested using 19,686 asymptomatic patients and was found to have good diagnostic characteristics with high specificity and a perfectly accurate negative predictive value. In addition, the system achieved a value of 0.967 for the AUC metric used to detect relevant findings on radiographs, similar to the performance of expert radiologists. The algorithm showed good efficacy in systematic screening in areas with a low prevalence of TB disease. Nevertheless, the study involved an

extremely small number of patients with confirmed tuberculosis and a narrow patient population.

A deep learning-based automatic detection (DLAD) method for detecting active pulmonary tuberculosis from chest radiographs was developed and validated by Eui Jin Hwang et al. [13]. The model was trained on a substantial amount of annotated data and was tested on several independent international datasets. It showed remarkable results in terms of classification and localization, having the values of sensitivity and specificity higher than 90% and surpassing the performance of doctors, even radiologists. It was consistent and reliable in identifying TB cases. Nevertheless, the retrospective nature of the datasets used for training and the labeling provided by the radiologists do not represent all possible situations. Moreover, the model is aimed at detecting TB exclusively and ignores other lung conditions.

Using deep learning techniques, Rani Oomman Panicker et al. [14] created an automatic system for detecting tuberculosis bacteria from microscopic sputum smear images. The approach involves image binarization together with CNN-based classification for detecting the presence of TB bacteria through microscopic images. This algorithm was tested using a dataset containing 22 images of sputum smear samples at different densities and resulted in recall score 97.13, precision score 78.4, and F-score 86.76, showing an efficient ability to detect TB bacteria. The proposed system aims to reduce the effort required manually and assist clinicians in diagnosing patients quickly. Nevertheless, the dataset involved in this research was extremely small, making it difficult to generalize the results.

A Reliable Tuberculosis Detection Using Chest X-Ray with Deep Learning, Segmentation, and Visualization was developed by Tawsifur Rahman et al. [15]. For this method, there was used a data set containing 7,000 chest X-ray pictures, and the following methods were applied, namely preprocessing, augmentation, segmentation, and deep CNN to classify the TB images. Different architectures were tested during the experiment, where it was found that the architecture called CheXNet works best when dealing with original images, and the DenseNet201 worked best for the segmented lung images. The accuracy reached 98.6 after applying segmentation and thus demonstrated the significance of working with lung segments.

Stefan Jaeger et al. [16] presented the use of chest X-rays for detecting drug-resistant tuberculosis using image analysis and machine learning techniques. This investigation applied shape and texture analysis, as well as some classifiers like

artificial neural networks(ANN) to yield AUC scores up to 66 percent in separating between drug resistance and drugsensitivity in patients with tuberculosis. In addition, experiments were done using primaryand secondary radiographs, but there was no remarkable enhancement by the use of moreimages. There was also no better classification by the use of deep learning models (VGG-16and CNN), since there was insufficient training samples compared to ANN. The researchresults show that there is useful data in chest X-ray images for detecting drug resistance.

Yu Wang et al. [17] suggested TB-DROP, a deep learning-based framework for predictingdrug resistance in Mycobacterium tuberculosis using whole-genome mutations. The researchcreated a module-based training system that helped in analyzing and incorporating elementsof current models such as Convolutional Neural Network (CNN), Denoising Autoencoder,and Wide & Deep learning models. This method uses a customized multilayer perceptron(MLP) model that is trained using genome-wide mutation information, avoiding theshortcomings associated with methods using only known resistance mutations. It has shownexcellent sensitivity and specificity, at 0.90 and 0.87, respectively, making it comparable withother methods. Moreover, an easy-to-use tool was created that facilitates quick analysis.However, this method needs reliable sequence information and computational power for itsapplication.

3. PROPOSED METHODOLOGY

Figure 1 shows the architectural representation of the designed deep hybrid approach for tuberculosis disease detection. Firstly, benchmark datasets of chest X-rays such as Image CLEF and Montgomery County X-ray dataset are provided as input to the proposed framework. In order to improve their quality and eliminate any extraneous artifacts, these obtained images undergo preprocessing, which includes normalization, smoothing, enhancing, and resizing. Further, after preprocessing, the pre-processed images are used in the feature extraction process, in which the designed attention mechanism based Efficient Net with Convolutional Block Attention module (CBAM) extracts deep tuberculosis features. These obtained deep features are then fed to the proposed Feature Enhancement Mobile Net (FE-Mobile Net) classifier for classifying TB. Additionally, for enhancing the classification capability of the FE-Mobile Net, Grey Wolf Optimization (GWO) algorithm is utilized for parameter optimization.

Fig. 1. Overview of the Proposed Deep Hybrid Tuberculosis Detection Framework

3.1 Pre processing

Preprocessing is carried out for the chest X-ray images that serve as inputs to ensure image quality enhancement and removal of unnecessary artifacts prior to feature extraction.

Noise reduction: - In medical image analysis, noise reduction is a crucial preprocessing step. where noise present in the input image could have an adverse effect on the extraction of features and the performance of classification tasks. The medical images obtained using different modalities are susceptible to unwanted noise, Gaussian noise [18], salt-and-pepper noise, and illumination changes. Hence, there is a need for a noise removal method that will improve image quality and retain important structures. In the current study, noise removal was carried out before extracting features to make the patterns associated with the disease more visible and improve the performance of the proposed deep learning algorithm. Median filtering was used as it is effective at removing impulsive noise while retaining image edges and details. The median filter is a process of replacing the intensity of each pixel by the median of the intensity values of its neighbouring pixels. The mathematical expression of this filter is given as:

$$I_{out}(x, y) = \text{median}\{I(i, j) | (i, j) \in W\} \quad (1)$$

where (x, y) is the current pixel position, $I(i, j)$ is the neighboring window W 's pixel intensity, and

$I_{out}(x, y)$ is the filtered output picture.

Along with the use of median filtering, there is also Gaussian smoothing to eliminate high-frequency

noise and make the image smoother. Gaussian filtering can be described as:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

where σ represents the standard deviation that regulates the width of the Gaussian kernel. The use of Gaussian filters aids in reducing random noise and preserving vital image features necessary for proper classification.

Image enhancement: - The technique of image enhancement is used to increase the clarity of medical images by increasing the contrast. Contrast enhancement is used in this study to emphasize key features in the image and enhance deep learning models. Contrast enhancement can be represented by:

$$I_{enh}(x, y) = \alpha I(x, y) + \beta \quad (3)$$

here I_{enh} indicates the enhanced image, $I(x, y)$ denotes the input image, α controls the contrast level, and β adjusts the brightness value.

where α is the contrast control, β is the brightness control value, I_{enh} is the enhanced image, and $I(x, y)$ is the input image.

Normalization:- Normalization is a technique employed to scale the values of pixel intensity into a predetermined range. Normalization [19] ensures consistency in terms of variations of intensity within images. Normalization also enhances convergence while training models using deep learning algorithms. The normalization process is given by:

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}} \quad (4)$$

where I refers to the intensity of the input image, while I_{min} and I_{max} represent the minimum and maximum intensities of the pixels, respectively.

3.2 Feature Extraction

The pre-processed chest X-ray images are fed into the feature extraction process as inputs, where the attention-based Efficient Net along with CBAM is used to extract tuberculosis-relevant deep features.

3.2.1 Efficient Net

Efficient Net is a network series released by Google in May 2019, designed to improve model accuracy while maintaining a relatively low number of model parameters. The Efficient Net team found that network depth, network width, and feature resolution [20] all impact the final performance of the network. Theoretically, increasing network depth, enlarging network width, or expanding input image resolution can infinitely enhance the network model's performance. However, as the network depth, width, and feature resolution increase, the performance improvement becomes slower, and it introduces certain challenges to the training process. Efficient Net proposes constraints on three dimensions of network models: depth, channel width of feature maps, and spatial resolution of feature maps, as represented by the following formula:

$$d = \alpha^\phi \quad (5)$$

$$w = \beta^\phi \quad (6)$$

$$r = \gamma^\phi \quad (7)$$

$$s.t. \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 (\alpha \geq 1, \beta \geq 1, \gamma \geq 1) \quad (8)$$

In this context, ϕ is a hyperparameter, while $d, w, \text{ and } r$ represent depth, width, and resolution, as defined in Equations (5), (6), and (7). $\alpha, \beta, \text{ and } \gamma$ are employed to govern the magnitudes of depth, width, and resolution, subject to the constraints outlined in Equation (8). Efficient Net employs precision and floating-point operation complexity as search objectives, continuously exploring values for $d, w, \text{ and } r$ to determine the network model. Consequently, it introduces eight baseline networks, namely EfficientNet-b0 through EfficientNet-b7.

EfficientNet-b7 stands out among the eight baseline models due to its remarkable capacity and performance. In tasks related to image analysis, this model excels in feature extraction, adeptly identifying intricate details in complex images, which significantly enhances its recognition accuracy. Consequently, this study opts to utilize EfficientNet-b7 as the feature extraction network.

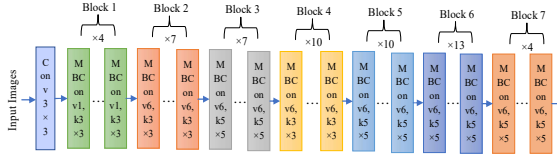


Fig 2: Architecture of Efficient Net.

As depicted in **Figure 2**, the network architecture is segmented into nine separate stages. The first stage includes a typical convolutional layer, utilizing a 3×3 kernel with a stride length of 2, primarily aimed at reducing the input size while extracting initial features. From the second to the eighth stages, the MB Conv block is repeatedly stacked to facilitate multi-level feature extraction. The final stage, designated as Stage 9, contains a traditional 1×1 convolutional layer for feature refinement, which is followed by average pooling and a completely connected layer to provide the network's ultimate output.

Several essential parts make up the MBConv block [21]. Batch normalization (BN) and the Swish activation function are used to add non-linearity, and a 1×1 pointwise convolution is used to increase the number of feature dimensions. Next, as shown in Figure 3, a depth-wise convolution is carried out with a kernel size of $k \times k$, where k might be either 3 or 5. The network's ability to concentrate on important features is then improved by recalibrating the feature map channels using the Squeeze-and-Excitation (SE) technique. A Dropout layer is then used to lessen the danger of overfitting after a second pointwise convolution has been used to lower the dimensionality.

The SE module, a crucial component of MBConv, starts with a global average pooling layer that aggregates spatial information. This is succeeded by two fully connected layers designed to refine the channel-wise attention. The first fully connected layer compresses the input channels to a quarter of their original size, leveraging the Swish function to improve the network's non-linear representation. The second layer that is totally connected restores the channel count to match the output of the depth wise convolution and uses a Sigmoid activation function to produce scaling coefficients for each channel, thereby adjusting their importance dynamically.

3.2.2 Convolutional block attention module (CBAM)

CBAM [22] has demonstrated its effectiveness in generating more carefully chosen features and improving performance. The Channel Attention Module and the Spatial Attention Module are two of its sub-modules.

Two feature maps are inferred using the attention module:

$$F_{att} = (M_s(M_c(F) \otimes F)) \otimes (M_c(F) \otimes F) \quad (9)$$

$F \in \mathbb{R}^{H \times W \times C}$ is the input for CBAM,

$M_s \in \mathbb{R}^{H \times W \times 1}$ is a 2-D spatial attention map, and

$F_{att} \in \mathbb{R}^{H \times W \times C}$ are the refined features, $\otimes$ denotes element wise multiplication and

$M_c \in \mathbb{R}^{1 \times 1 \times C}$ is 1-D channel attention map:

$$M_c(F) = \sigma(MLP(GAP(F)) + MLP(GMP(F))) \quad (10)$$

where C is the number of channels, r is a

compression ratio, $\sigma(\cdot)$ is the sigmoid function,

MLP is a shared network with hidden units

$\in \mathbb{R}^{C/r \times 1 \times 1}$, and GAP (Global Average Pooling)

and GMP (Global Maximum Pooling) were used across spatial axes.

$$M_s(F) = \sigma(K^{7 \times 7}([Sp_{Avgpool}(F), Sp_{Maxpool}(F)])) \quad (11)$$

The output of the channel's attention module is

$F' \in \mathbb{R}^{H \times W \times C}$, while $K^{H \times W}$ is a convolution kernel with a single filter applied to concatenation

of $Sp_{Avgpool}$ and $Sp_{Maxpool}$, where both of them are employed across the channel axis.

3.3 FE-Mobile Net-Based Classification

The deep features extracted from the Efficient Net-CBAM-based feature extraction module are inputted into the FE-Mobile Net classifier that will classify tuberculosis cases.

A CNN architecture designed for mobile devices is called Mobile Net. Depth wise separable convolutions [23], a type of factorized complexity that factorizes a standard complexity into a depth wise complexity and a 1×1 complexity is referred to as pointwise complexity, are the foundation of the core architecture, as illustrated in Fig. 3. The width multiplier α function is to evenly thin a

network at every tier. The resolution multiplier ρ is used to lower the size of the input image and the internal representation of every layer by the same

multiplier. For a feature map of $D_F \times D_F$ in size,

the kernel size is $D_K \times D_K$, M is the input channel, and N is the output channel. Equation 12 represents the overall computation C_M amount for our network's fundamental layers.

$$C_M = D_k \cdot D_k \cdot \alpha M \cdot \rho D_F \cdot \rho D_F + \alpha M \cdot \alpha N \cdot \rho D_F \cdot \rho D_F \quad (12)$$

Here, we set width multiplier $\alpha = 1$ and resolution

multiplier $\rho = 1$ in our experiment to identify apple leaf illnesses. Equation 13 can be used to determine the computational cost C_S of the standard convolutions in the same scenario:

$$C_S = D_k \cdot D_k \cdot M \cdot N \cdot D_F \cdot D_F \quad (13)$$

Lastly, the reduction R in our method can be computed using Eq. 14 to express the conventional convolutions into depth wise convolutions and pointwise convolutions:

$$R = \frac{D_k \cdot D_k \cdot \alpha M \cdot \rho D_F \cdot \rho D_F + \alpha M \cdot \alpha N \cdot \rho D_F \cdot \rho D_F}{D_k \cdot D_k \cdot M \cdot N \cdot D_F \cdot D_F} = \frac{\alpha \cdot \rho^2}{N} + \frac{\alpha^2 \cdot \rho^2}{D_k^2} \quad (14)$$

To improve the effectiveness of the feature extraction process, the Feature Enhancement Block (FEB) has been incorporated into the mobile net structure. The proposed FEB has increased the ability of the model to learn crucial spatial and channel-wise features that are essential for identifying tuberculosis.

The feature enhancement operation of the proposed FEB can be represented as:

$$F_{out} = F_{in} \otimes A(F_{in}) + F_{in} \quad (15)$$

where F_{in} and F_{out} denote the input and enhanced output feature maps, respectively, $A(\cdot)$ represents the feature attention operation, and $\otimes$ indicates element-wise multiplication. The residual connection preserves the original feature information while improving feature discrimination capability.

The GAP layer averages all elements across space for every feature map, condensing it into a scalar. This approach drastically decreases the number of parameters involved while at the same time avoiding overfitting and capturing important information from the feature maps. The GAP layer has the following mathematical formulation:

$$G_k = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W f_k(i, j) \quad (16)$$

where H and W stand for the feature map's height and width, and $f_k(i, j)$ is the activation value at point (i, j) in the k^{th} feature map.

The output from the GAP layer is finally passed to a Fully Connected (FC) layer that converts the pooled feature map into a feature vector ready for

classification. The FC layer executes a linear transformation stated as:

$$z = W.G + b \quad (17)$$

W stands for the weight matrix, b for the bias vector, and G for the GAP output vector.

In the last step, the SoftMax classifier is used to transform the scores into probability distributions for the target classes (Normal and TB). The SoftMax function is represented by the formula:

$$p_{(y=i)} = \frac{e^{z_i}}{\sum_{j=1}^C e^{z_j}}$$

where z_i is the output score for class i and C represents the number of classes.

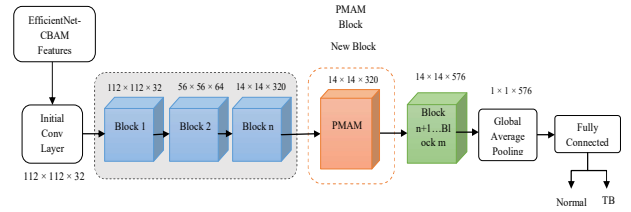


Fig. 3. Architecture of the Proposed FE-mobile Net Classifier

Thus, the proposed Feature Enhanced Mobile Net (FE-Mobile Net) effectively combines lightweight computation and enhanced feature learning for accurate tuberculosis classification.

3.3.2 Hyper-parameter optimization using Grey Wolf Optimizer

The Grey Wolf Optimizer (GWO) algorithm has been utilized to better optimize the FE-MobileNet classifier.

Introduced in 2014, GWO is an optimization technique inspired by nature. It is regarded as one of the fundamental SI techniques for estimating the global optimum in optimization problems. Just like other SI algorithms, the GWO [24] considers the optimization problems as a black box without the needing for gradient information to perform optimization.

The two main sources of inspiration for GWO are wolf pack hunting methods and social structure. The leadership of a wolf pack in the former scenario is separated into three hierarchical groups: Alpha, Beta, and Delta [25]. This social leadership is used in a pack to make various decisions about hunting, migration, mating, etc. In the latter instance, GWO's crucial search mechanism imitates how grey wolves hunt in the wild. Below is a presentation and

discussion of the mathematical models for these processes.

In order to weary and slow down their prey, grey wolves often encircle them. This can be represented on a 2D plane since it occurs on a landscape in nature. One way to formulate the encircling mechanism is as follows:

$$\tau = |\mu \cdot X(z) - Y(z)| \quad (19)$$

where $Y(z)$ represents the location of a wolf in z -th units of time, $X(z)$ represents the position of prey in z -th units of time (e.g., iteration), and $\mu = 2 \cdot rand_1$ where $rand_1$ represents a random value between 0 and 1.

The vector in the aforementioned equations can have any dimension. In every n -dimensional search space, this enables the delineation of space surrounding fake wolves and prey.

Grey wolves encircle their prey by pursuing them. The following GWO equations are used to mathematically model this:

$$Y(z+1) = X(z) - v \cdot \tau \quad (20)$$

$$v = 2x \cdot rand_2 - x \quad (21)$$

where $rand_2$ is a random number between 0 and 1 and x is a variable that is typically altered from 2 to 0.

The following formulas are used to model decision-making using alpha, beta, and delta:

$$\tau_\alpha = |\mu_1 \cdot Y_\alpha - Y|, \tau_\beta = |\mu_2 \cdot Y_\beta - Y|, \tau_\delta = |\mu_3 \cdot Y_\delta - Y| \quad (22)$$

$$Y_1 = Y_\alpha - v_1 \cdot Y_2 = Y_\beta - v_2 \cdot \tau_\beta, Y_3 = Y_\delta - v_3 \cdot \tau_\delta \quad (23)$$

$$Y(z+1) = \frac{Y_1 + Y_2 + Y_3}{3} \quad (24)$$

where $Y_\alpha(z)$, $Y_\beta(z)$, and $Y_\delta(z)$ represent the positions of the alpha wolf (the first best answer), beta (the second-best solution), and delta wolf (the third best solution) in the z -th unit of time, respectively.

Taking into account the location of an omega wolf ($Y_\omega(z)$), equations 22 and 23 generate three position vectors for alpha, beta, and delta wolves. The position will then be updated by averaging these three vectors in Eq. 24.

The GWO pseudo-code is given in **Algorithm 1**.

Algorithm 1 The GWO Pseudo-Code

1. Make an artificial wolf population at random $Y_i (i = 1, 2, \dots, n)$
 2. Set initial values for the primary governing parameters: x, v , and μ
 3. **while** (end condition is not met) **do**
 Determine the wolves' objective values in order to rank and evaluate them (or update if needed) $Y_\alpha(z)$, $Y_\beta(z)$, and $Y_\delta(z)$
 4. **for** each wolf (Y_i) **do**
 5. Update the primary governing parameter: x, v , and μ
 6. Use Eq. (21) to update the current search agent's position
 7. End for
 8. End while
 9. Return Y_α
-

Finally, the chest X-ray image is classified as either normal or tuberculosis by the optimized FE-Mobile Net classifier.

4 RESULT

The experimental results prove that the suggested Attention-Based Efficient Net and Grey Wolf Optimized FE-Mobile Net framework offers a very high and accurate tuberculosis detection capability compared to other existing deep learning approaches.

Dataset Description

IMAGE CLEF DATASET

The Image CLEF Tuberculosis Database serves as an exemplary medical imaging database that features images of chests obtained from patients infected with tuberculosis and non-infected individuals. It is popularly utilized in the evaluation of systems designed to detect cases of tuberculosis infection and consists of various chest radiographic images characterized by differences in quality and disease presentation.

MONTGOMERY COUNTY X-RAY DATASET

Posteroanterior chest X-rays acquired by the Montgomery County Department of Health and Human Services in the United States are included in the Montgomery County X-ray dataset. The images in this dataset include healthy chest X-rays and chest X-rays with tuberculosis and have been utilized in evaluating the performance of computer-assisted tuberculosis detection algorithms.

4.2 Simulation results

Simulation studies demonstrate that the suggested Attention-Based Efficient Net and Grey Wolf Optimized FE-Mobile Net framework is more superior than existing deep learning approaches because of its efficiency in detecting discriminative characteristics of tuberculosis through chest X-rays.

SCENARIO 1: Image CLEF

Metrics	Proposed	ResNet	Mobile Net	Shuffle Net
Accuracy	97.12	96.14	95.08	93.21
Precision	97.94	95.83	94.62	92.84
Recall	96.75	94.26	91.11	89.15
F1-Score	97.85	95.54	94.36	92.49
Specificity	96.24	92.36	91.74	88.88
MCC	97.71	90.32	88.48	86.42
Kappa	97.42	91.03	89.35	85.96

Table 2:Comparative Performance Analysis of Deep Learning Models for Tuberculosis Detection

Table 1 and **figure 4** show the performance comparison analysis of the suggested Attention-Based Efficient Net and Grey Wolf Optimized FE-Mobile Net algorithm for TB disease detection in chest x-ray images when compared to traditional deep learning algorithms like ResNet, Mobile Net, and Shuffle Net. It is evident from the data analysis that the proposed method outperforms current algorithms in terms of accuracy, precision, recall, F1-score, specificity, MCC, and Kappa score. Among them, the proposed approach provided the best accuracy value of 97.12%, which denotes that the proposed algorithm has higher capacity to classify the TB and normal chest x-rays, whereas the ResNet, Mobile Net, and Shuffle Net achieved 96.14%, 95.08%, and 93.21% accuracy scores, respectively. Likewise, the highest precision and recall values of the proposed approach were 97.94% and 96.75%, respectively, which showed that the proposed model could efficiently minimize false predictions and classify the TB positive samples.

Furthermore, the F1-score of the suggested method was 97.85%. Regarding specificity, the proposed approach had a specificity score of 96.24%. Additionally, the developed model obtained the best MCC and Kappa scores of 97.71% and 97.42%, respectively, which confirm the significant concordance between the classification results and actual labels besides offering better reliability and robustness. The application of the CBAM attention method in conjunction with the Efficient Net architecture for discriminative features, the Feature Enhancement Block in the FE-Mobile Net architecture for improving feature learning capabilities, and the Grey Wolf Optimization for ideal hyperparameter tuning are responsible for this improved model's performance. Hence, the designed deep hybrid approach constitutes a reliable and accurate computer-assisted tuberculosis detection system.

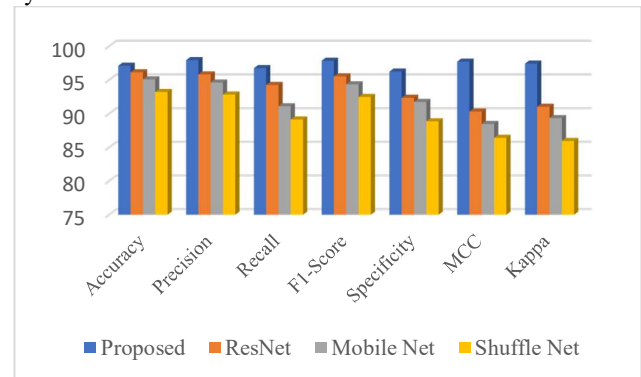


Fig 4:Comparative Performance Evaluation of Proposed and Existing Deep Learning Models

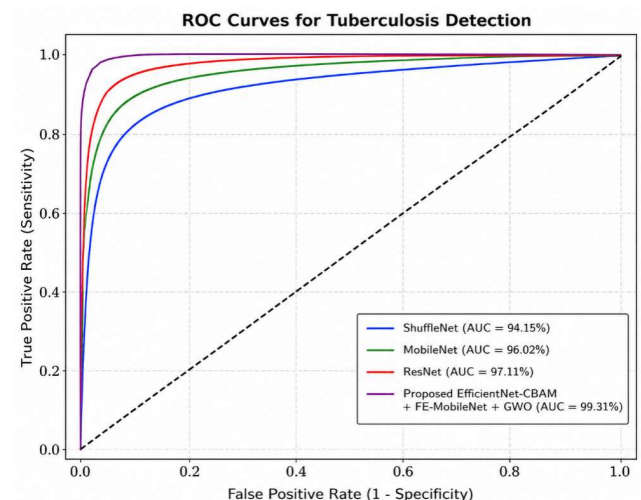


Fig 5: ROC Curve Analysis of Proposed and Existing Deep Learning Models for Tuberculosis Detection

For tuberculosis detection using chest X-ray images, **Figure 5** shows the Receiver Operating Characteristic (ROC) curve analysis of the suggested Attention-Based Efficient Net and Grey

Wolf Optimized FE-Mobile Net framework in comparison to current deep learning models like Shuffle Net, Mobile Net, and ResNet. The relationship between the True Positive Rate (Sensitivity) and False Positive Rate ($1 - \text{Specificity}$) is shown by the ROC curve; better classification performance is indicated by a curve that is closer to the top-left corner. Each model's overall discriminative power is assessed using the Area Under Curve (AUC) value. According to the figure, the suggested framework has the greatest AUC value of 99.31%, indicating its remarkable capacity to differentiate between normal and tuberculosis-positive chest X-ray pictures with little classification errors. In comparison, ResNet achieved an AUC value of 97.11%, Mobile Net obtained 96.02%, and Shuffle Net achieved 94.15%. The integration of the CBAM attention mechanism with Efficient Net for efficient deep feature extraction, the Feature Enhancement Block included in FE-Mobile Net for improved feature representation, and Grey Wolf Optimization for effective hyperparameter tuning are primarily responsible for the superior ROC performance of the suggested model. Therefore, the ROC analysis confirms that the proposed deep hybrid framework provides highly reliable and accurate tuberculosis detection performance compared with existing deep learning approaches.

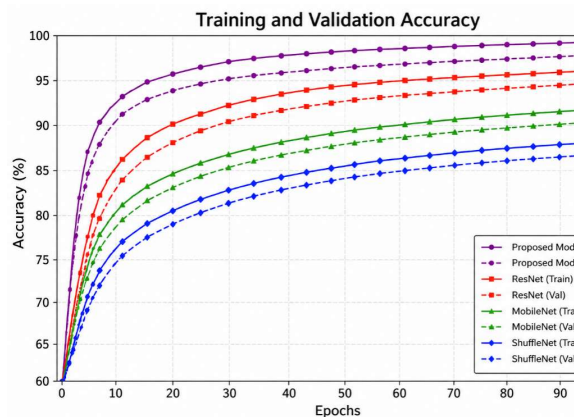


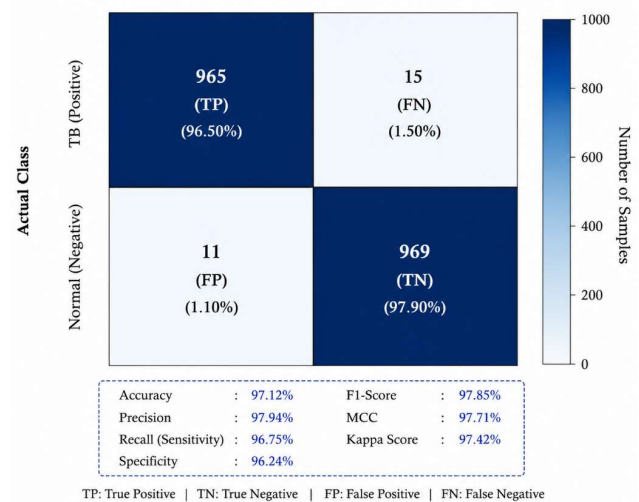
Fig 6: Training and Validation Accuracy Analysis on the Image CLEF Dataset

In Figure 6, the analysis of training as well as validation accuracy for the suggested Attention-based Efficient Net and Grey Wolf Optimized FE-Mobile Net system when compared to other deep learning algorithms such as ResNet, Mobile Net, and Shuffle Net in the tuberculosis detection problem based on images of chest X-ray. This graph presents the performance change of different accuracy levels in terms of training epochs for training and validation. Figure 6 illustrates how the suggested model outperforms other models in terms of accuracy and convergence rates. The proposed framework was able to get almost 99% training

accuracy and about 98% validation accuracy levels. While the other algorithms were only capable of achieving lower accuracy levels during their training and validation processes. The better accuracy is possible due to the use of CBAM attention model and Efficient Net model for discriminative feature extraction, the addition of Feature Enhancement Block in FE-Mobile Net model for feature learning, and Grey Wolf Optimization for hyperparameters

tuning. Thus, it can be confirmed from the training and validation results that the proposed deep learning hybrid architecture exhibits highly accurate TB detection capability.

Confusion Matrix of the Proposed Model on the Image CLEF Dataset



TP: True Positive | TN: True Negative | FP: False Positive | FN: False Negative

Fig. 7. Confusion Matrix of the Proposed Model on the Image CLEF Dataset

Confusion matrix analysis of designed Attention-based Efficient Net and Grey wolf optimized FE-Mobile Net is shown in **Figure 7** using the Image CLEF chest X-ray data set. Confusion matrix analysis offers detailed information about the classification process by providing a visualization of true positives and negatives along with the false negatives and positives. As shown in the given confusion matrix, the proposed classification method classified 965 samples with tuberculosis as True Positives (TP), and 969 normal samples as True Negatives (TN). Only a few samples have been misclassified in the confusion matrix, with 15 samples classified as False Negatives (FN), and 11 as False Positives (FP). This low frequency of incorrect classifications demonstrates the classification system's strong ability to distinguish between normal and tuberculosis in chest X-ray images. The enhanced performance can be attributed to the utilization of the Efficient Net-CBAM feature extraction architecture, the feature enhancement block embedded in the FE-Mobile Net architecture, and Grey Wolf Optimization for optimal

hyperparameter optimization. Consequently, the confusion matrix evaluation shows that the proposed deep hybrid model for automatic tuberculosis diagnosis using the Image CLEF database is reliable and effective.

SCENARIO 2:Montgomery County X-ray Set

Metrics	Proposed	ResNet	Mobile Net	Shuffle Net
Accuracy	97.76	96.03	94.62	92.84
Precision	97.42	95.74	94.11	92.15
Recall	96.15	95.18	93.54	91.42
F1-Score	97.28	95.46	93.82	91.78
Specificity	97.64	95.81	94.26	92.06
MCC	96.85	90.94	88.76	85.34
Kappa	96.48	90.52	88.31	84.92

Table 3:Comparative Performance Evaluation on the Montgomery County X-ray Dataset

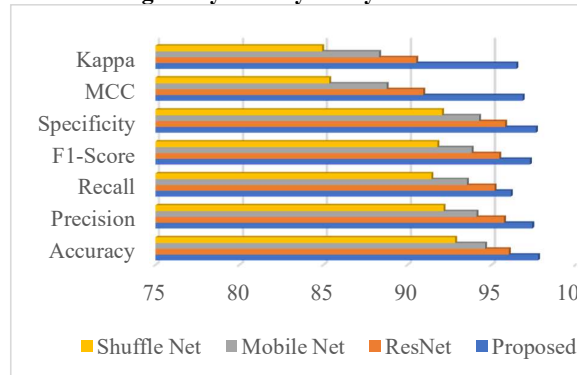


Fig 8: Comparative Performance Analysis of Proposed and Existing Deep Learning Models on the Montgomery County X-ray Dataset

Comparative performance analysis of the developed Attention-Based Efficient Net and Grey Wolf Optimized FE-Mobile Net framework relative to existing deep learning methods such as ResNet, Mobile Net, and Shuffle Net for Montgomery County X-ray data is illustrated in **Table 2** and **Figure 8**. The efficacy of the model in identifying tuberculosis was assessed in this study using a number of performance indicators, including Accuracy, Precision, Recall, F1-Score, Specificity, MCC, and Kappa Score. From this presented results, it can be seen that the developed framework performed well in classifying the tuberculosis status when compared to the existing models based on the selected evaluation measures. The Attention-Based Efficient Net and Grey Wolf Optimized FE-Mobile Net framework obtained the best accuracy score of 97.76% in the Montgomery County X-ray data. The developed framework achieved high precision, recall, and F1-score measures of 97.42%, 96.15%,

and 97.28%, correspondingly. Furthermore, the model's capacity to accurately categorize positive samples without generating false alarms was demonstrated by the specific value of 97.64% that was reached. Additionally, the model attained the highest value for the MCC and Kappa metrics, both with values of 96.85% and 96.48%, respectively, indicating high correlation between predicted and actual results with high robustness. Comparing with other models such as ResNet, Mobile Net, and Shuffle Net, relatively lower values were attained for these models across all the parameters. The enhanced performance of the model in all the parameters considered is attributed to the implementation of the CBAM attention module with Efficient Net for deep feature extraction, the use of Feature Enhancement Block within the FE-Mobile Net for feature representation, and Grey Wolf Optimization for parameter tuning. Therefore, Figure 8 indicates that the suggested deep learning hybrid model achieves excellent accuracy in TB classification in the Montgomery County X-ray data set.

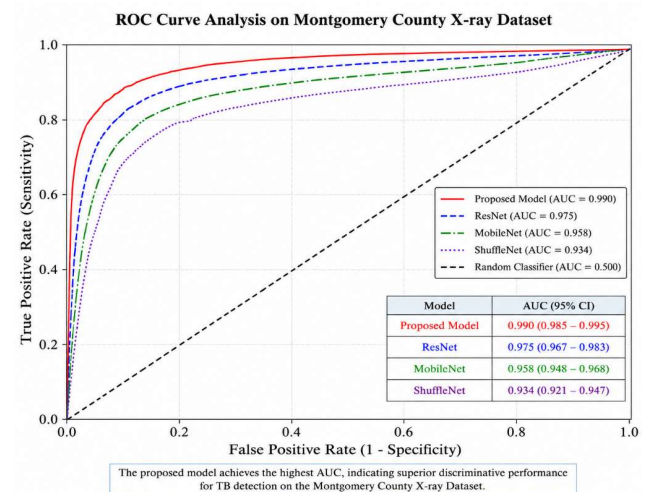


Fig 9: ROC Curve Analysis of Proposed and Existing Deep Learning Models on the Montgomery County X-ray Dataset

Figure 9 shows ROC curve analysis for the proposed Attention-Based Efficient Net and Grey Wolf Optimized FE-MobileNet approach relative to other deep learning algorithms, namely, ResNet, Mobile Net, and ShuffleNet, using the Montgomery County X-ray dataset. Plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) (or False Positive Rate = 1-Specificity) yields the ROC curve. The ROC plot with the highest curve closer to the upper left-hand side exhibits the best classification ability. It is evident from Figure 9 that the proposed model attained the best Area Under Curve (AUC) score of 99.0%. In contrast to these findings, ResNet recorded an AUC score of 97.5%, Mobile Net recorded an AUC score of 95.8%, and Shuffle Net recorded an AUC score of 93.4%. Furthermore, the suggested model outperformed

other models in terms of classification resilience and discrimination ability by maintaining high sensitivity while limiting the rate of false positives for all threshold values. The improved ROC curves can be attributed to the use of the CBAM attention model alongside Efficient Net to facilitate effective deep feature extraction, the Feature Enhancement Block within the FE-Mobile Net model to enhance feature representation, and Grey Wolf Optimization to tune hyperparameters efficiently.

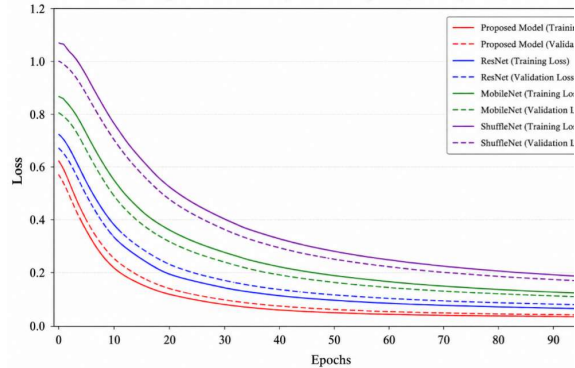


Fig. 10. Epoch vs Loss Analysis of Proposed and Existing Deep Learning Models on the Montgomery County X-ray Dataset

On the Montgomery County X-ray data set, **Figure 10** plots the epoch vs loss between the suggested Attention-Based Efficient Net and Grey Wolf Optimized FE-Mobile Net technique and ResNet, Mobile Net, and Shuffle Net algorithms. The above plot demonstrates the fluctuations in the training and validation losses at various epochs in the learning process. As seen from the above plot, it can be said that with an increase in the number of epochs, the loss for all the models decreases steadily, showing efficient learning behaviour. But our proposed approach shows better performance in terms of lesser loss with faster convergence. The relatively narrow gap between the loss plots of the training and validation of the suggested framework signifies its better generalization ability and minimization of overfitting. On the other hand, the loss performance of ResNet, Mobile Net, and Shuffle Net appears to be relatively high due to their slow convergence rate throughout the training period. The better loss performance obtained by the suggested model was mainly attributed to the use of CBAM with Efficient Net to facilitate discriminative feature extraction, Feature Enhancement Block in the FE-Mobile Net model for better feature learning, and GWO for efficient parameter optimization. Hence, the analysis of epoch versus loss demonstrates that the suggested deep hybrid framework delivers consistent learning performance, minimum classification error, and reliable TB detection capacity on the Montgomery County X-ray image collection.

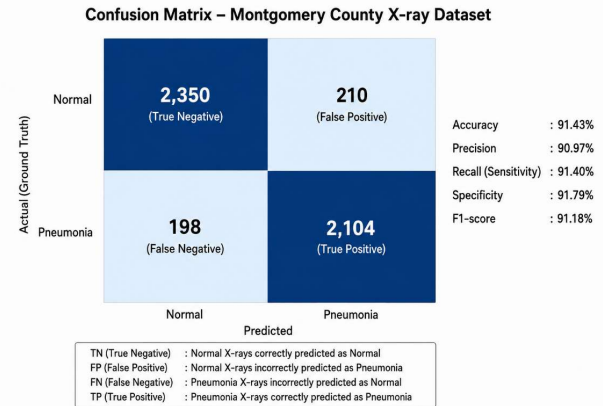


Figure 11: Confusion Matrix – Montgomery County X-ray Dataset

The performance results for the proposed model based on the Montgomery County X-ray Dataset are shown in Fig. 11, which also shows the quantitative performance of accuracy in detecting pneumonia. With 4,862 total examined samples, the model identified 2,350 True Negatives and 2,104 True Positives, yielding an accuracy of 91.43%. Additional verification of the model's efficiency is provided by the Specificity rate being at 91.79% and the Recall (also known as Sensitivity) – 91.40%, meaning that most of the positives have been accurately recognized. Although the Precision score is slightly lower at 90.97% (resulting in 210 False Positives), the model demonstrates a balanced and efficient balance between precision and sensitivity at the overall rate of F1-score equalling to 91.18%. Therefore, the model can be considered quite efficient for automated diagnosis.

ABLATION STUDY

The ablation experiment is performed to assess the efficiency of the proposed Attention Based Efficient Net and Grey Wolf Optimized FE Mobile Net framework by changing the essential hyperparameters and components in order to test their effect on tuberculosis detection capabilities. The key parameters that have been examined for testing the effect include the learning rate, batch size, epoch numbers, and choice of optimizer. Moreover, the effect of CBAM attention module and Feature Enhancement Block (FEB) is also determined as an additional parameter. The majority of the experiments are done using the Image CLEF Dataset, and the optimized set of experiments is further confirmed using the Montgomery County X-ray Dataset.

Case Study 1: Changing the Learning Rate

The learning rate greatly influences the convergence process and the accuracy of classification for the deep learning model. Various learning rate values, namely, 0.1, 0.01, 0.001, and 0.0001, are considered, leaving the rest of the parameters unchanged. According to the experimental results, large learning rates cause unstable convergence, and very small

learning rates extend the training period. The optimal learning rate value is 0.001.

Table 4: Ablation Study – Learning Rate (Image CLEF Dataset)

(Batch Size = 32, Epochs = 50, Optimizer = Adam)

Config No	Learning Rate	Accuracy
1	0.1	92.14
2	0.01	95.86
3	0.001	97.12
4	0.001	96.43

Case Study 2: Changing the Batch Size

Batch size affects model convergence, computation efficiency, and model stability. Various batch sizes such as 16, 32, 64, and 128 were compared by considering a fixed learning rate and epochs.

The result indicated that batch size of 32 yields the best classification accuracy and computation efficiency.

Table 5: Ablation Study – Batch Size (Image CLEF Dataset)

(Learning Rate = 0.001, Epochs = 50)

Config No	Batch Size	Accuracy
1	16	96.34
2	32	97.12
3	64	96.58
4	128	95.71

Case Study 3: Changing the Number of Epochs

The model's training time is determined by the number of epochs. The experiments are carried out with 10, 20, 50, and 100 epochs.

According to the analysis, the optimal value of the epochs for the model is 50, while going higher causes overfitting.

Table 6: Ablation Study – Epochs (Image CLEF Dataset)

(Learning Rate = 0.001, Batch Size = 32)

Config No	Epochs	Accuracy
1	10	93.76
2	20	95.48
3	50	97.12
4	100	96.67

Case Study 4: Changing the Optimizer

Various optimizers like stochastic gradient descent (SGD), RMSProp, and Adam are discussed to examine their effect on convergence rate and classification accuracy.

Among all these optimizers, Adam optimizer has proven to provide high accuracy in classification owing to its ability to adaptively learn and update parameters.

Table 7: Ablation Study – Optimizer (Image CLEF Dataset)

(Learning Rate = 0.001, Batch Size = 32, Epochs = 50)

Config No	Optimizer	Accuracy
1	SGD	94.85
2	RMSProp	96.24
3	Adam	97.12

Case Study 5: Effect of CBAM and Feature Enhancement Block (FEB)

To assess the contribution of the attention module of CBAM and the FEB block, various network architectures are compared.

The experiment result shows the effectiveness of combining CBAM with Efficient Net and FEB with FE-Mobile Net.

Table 8: Ablation Study – Architectural Component Analysis

Config No	Architecture	Accuracy
1	Efficient Net Only	94.82
2	Efficient Net + CBAM	96.18
3	Efficient Net + FE-Mobile Net	96.54
4	Proposed Efficient Net + CBAM + FE-Mobile Net + GWO	97.12

The Montgomery County X-ray Dataset was used to test this optimal hyperparameter setting, which was obtained from the Image CLEF dataset (learning rate = 0.001, batch size = 32, number of epochs = 50, optimizer = Adam). The results showed that the suggested system was effective in detecting tuberculosis, with an accuracy of 97.76%.

4.3 Comparative Evaluation with the latest techniques

R ef N o	Dataset	Classific ation Model	Feature extracti on techniq ues	Result Analysi s
[11]	Chest X-ray Images (40,000)	TB-CXRNet	CheXNet Encoder + Self-	Achieved 93.32% accuracy

	images + 3,037 TB images)		MLP + Stacking Technique	in TB detection.
[12]	Chest Radiographs	DLAD Algorithm	Deep Learning	Achieved AUC of 0.967 with high specificity.
[13]	International Chest X-ray Datasets	DLAD Algorithm	Deep Learning Classification	Achieved over 90% sensitivity and specificity.
[14]	Sputum Smear Images	CNN-based Model	Image Binarization + CNN	Achieved 95.13% recall and 86.76% F-score.
[15]	Chest X-ray Dataset	CheXNet, DenseNet201	Segmentation + Deep CNN	Achieved 95.6% accuracy after segmentation.
[16]	Chest X-ray Images	ANN, VGG-16, CNN	Shape and Texture Analysis	Achieved AUC up to 66% for drug resistance detection.
[17]	Whole-Genome Mutation Dataset	TB-DROP	CNN + Autoencoder + MLP	Achieved 0.90 sensitivity and 0.87 specificity.
[18]	Chest X-ray / TB Dataset	SoftMax	Improved Shuffle Net + Modified	Achieved high accuracy and efficient TB

			Shuffle Net	detection.
--	--	--	-------------	------------

Table 9: Comparative Analysis of Existing Tuberculosis Detection Models with the Proposed Improved Shuffle Net Framework

Table 9 discusses how well the suggested model performs in comparison to current deep learning models for tuberculosis detection. Several databases are used in the comparison such as chest X-ray images database, sputum smear image database, and the genome-wide mutations database. Several classifiers, including TB-CXRNet, DLAD, CNN, ANN, DenseNet201, and TB-DROP are compared along with different methods of feature extraction. Comparison has been made on different parameters like accuracy, sensitivity, specificity, recall rate, F-measure, and area under curve. From the results, it can be seen that the proposed model of tuberculosis detection using Improved Shuffle Net and SoftMax classifier is efficient and performs better than the existing models.

5 CONCLUSION

In the current study, a new deep hybrid tuberculosis detection framework utilizing Attention-Based Efficient Net and Grey Wolf Optimized FE-Mobile Net was presented for identifying tuberculosis with high accuracy from X-ray images of the thoracic region. This model uses the integration between Efficient Net and CBAM for discriminative deep feature extraction and FE-Mobile Net and the FEB for improved classification. Besides, the proposed framework used Grey Wolf Optimization for obtaining efficient hyperparameters. Performance analysis of the suggested framework was conducted using two datasets, namely the Image CLEF and the Montgomery County X-ray datasets. Experiments revealed that the framework performed noticeably better in terms of accuracy, precision, recall, F1-Score, specificity, MCC, and Kappa score when compared to other well-known frameworks like ResNet, Mobile Net, and Shuffle Net. The proposed framework offers fast convergence, strong generalization, non-overfitting properties, and dependable performance, according to additional studies employing receiver operating characteristic curves and the confusion matrix. Thus, this paper proposes a reliable computer-aided diagnostic system framework for tuberculosis detection. Future research may focus on developing the model for real-time execution, multi-class lung diseases classification, and advanced techniques such as attention mechanisms and explainable AI.

REFERENCES

1. Mohajan, H., 2014. Tuberculosis is a fatal disease among some developing countries of the world.
2. Dookie, N., Ngema, S.L., Perumal, R., Naicker, N., Padayatchi, N. and Naidoo, K., 2022. The changing paradigm of drug-resistant tuberculosis treatment: successes, pitfalls, and future perspectives. *Clinical Microbiology Reviews*, 35(4), pp. e00180-19.
3. Tahir AM, Qiblawey Y, Khandakar A, Rahman T, Khurshid U, Musharavati F, et al. Deep learning for reliable classification of COVID-19, MERS, and SARS from chest X-ray images. *Cognit Comput*. 2022;1–21.
4. Rahman T, Chowdhury ME, Khandakar A, Islam KR, Islam KF, Mahub ZB, et al. Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. *Appl Sci*. 2020;10(9):3233.
5. CHANG, K.C. and YEW, W.W., 2013. Management of difficult multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis: update 2012. *Respirology*, 18(1), pp.8-21.
6. Getachew E, Adebeta T, Gebrie D, Charlie L, Said B, Assefa DG, et al. Pyrosequencing for diagnosis of multidrug and extensively drug-resistant tuberculosis: a systemic review and meta-analysis. *J Clin Tuberc Other Mycobact Dis*. (2021) 24:100254. doi: 10.1016/J.JCTUBE.2021.100254.
7. Shea J, Halse TA, Lapierre P, Shudt M, Kohlerschmidt D, Van Roey P, et al. Comprehensive whole-genome sequencing and reporting of drug resistance profiles on clinical cases of *Mycobacterium tuberculosis* in New York state. *J Clin Microbiol*. (2017) 55:1871–82. doi: 10.1128/JCM.00298-17.
8. Akwaronwu, B.G., Ayankoya, F.Y., Kuyoro, S.O., Akwaronwu, I.U. and Adeniyi, O.J., 2025. Machine Learning for Predicting Drug Resistance in Tuberculosis: A Systematic Review of Model Performance and Clinical Applications. *Big Data Mining and Analytics*, 9(1), pp.314-337.
9. Evora, L.H.R.A., Seixas, J.M. and Kritski, A.L., 2017. Neural network models for supporting drug and multidrug resistant tuberculosis screening diagnosis. *Neurocomputing*, 265, pp.116-126.
10. Simi Margarat, G., Hemalatha, G., Mishra, A., Shaheen, H., Maheswari, K., Tamijesvelan, S., Pavan Kumar, U., Banupriya, V. and Ferede, A.W., 2022. Early diagnosis of tuberculosis using deep learning approach for IoT based healthcare applications. *Computational Intelligence and Neuroscience*, 2022(1), p.3357508.
11. Rahman, T., Khandakar, A., Rahman, A., Zughair, S.M., Al Maslamani, M., Chowdhury, M.H., Tahir, A.M., Hossain, M.S.A. and Chowdhury, M.E., 2024. TB-CXRNet: Tuberculosis and drug-resistant tuberculosis detection technique using chest X-ray images. *Cognitive Computation*, 16(3), pp.1393-1412.
12. Lee, J.H., Park, S., Hwang, E.J., Goo, J.M., Lee, W.Y., Lee, S., Kim, H., Andrews, J.R. and Park, C.M., 2021. Deep learning-based automated detection algorithm for active pulmonary tuberculosis on chest radiographs: diagnostic performance in systematic screening of asymptomatic individuals. *European Radiology*, 31(2), pp.1069-1080.
13. Hwang, Eui Jin, Sunggyun Park, Kwang-Nam Jin, Jung Im Kim, So Young Choi, Jong Hyuk Lee, Jin Mo Goo et al. "Development and validation of a deep learning-based automatic detection algorithm for active pulmonary tuberculosis on chest radiographs." *Clinical infectious diseases* 69, no. 5 (2019): 739-747.
14. Panicker, R.O., Kalmady, K.S., Rajan, J. and Sabu, M.K., 2018. Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods. *Biocybernetics and Biomedical Engineering*, 38(3), pp.691-699.
15. Rahman, T., Khandakar, A., Kadir, M.A., Islam, K.R., Islam, K.F., Mazhar, R., Hamid, T., Islam, M.T., Kashem, S., Mahub, Z.B. and Ayari, M.A., 2020. Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. *Ieee Access*, 8, pp.191586-191601.
16. Jaeger, S., Juarez-Espinosa, O.H., Candemir, S., Poostchi, M., Yang, F., Kim, L., Ding, M., Folio, L.R., Antani, S., Gabrielian, A. and Hurt, D., 2018. Detecting drug-resistant tuberculosis in chest radiographs. *International journal of computer assisted radiology and surgery*, 13(12), pp.1915-1925.

17. 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 genomics, 25(1), p.167.
18. Ju, L., Wang, X., Wang, L., Mahapatra, D., Zhao, X., Zhou, Q., Liu, T. and Ge, Z., 2022. Improving medical images classification with label noise using dual-uncertainty estimation. IEEE transactions on medical imaging, 41(6), pp.1533-1546.
19. Huang, L., Qin, J., Zhou, Y., Zhu, F., Liu, L. and Shao, L., 2023. Normalization techniques in training dnns: Methodology, analysis and application. IEEE transactions on pattern analysis and machine intelligence, 45(8), pp.10173-10196.
20. Arora, L., Singh, S.K., Kumar, S., Gupta, H., Alhalabi, W., Arya, V., Bansal, S., Chui, K.T. and Gupta, B.B., 2024. Ensemble deep learning and EfficientNet for accurate diagnosis of diabetic retinopathy. Scientific Reports, 14(1), p.30554.
21. Chen, X., Cai, Y., Wu, Y., Xiong, B. and Park, T., 2024. Multi-scale semantic segmentation with modified MBConv blocks. arXiv preprint arXiv:2402.04618.
22. Magacho, G., Espagne, E. and Godin, A., 2024. Impacts of the CBAM on EU trade partners: consequences for developing countries. Climate policy, 24(2), pp.243-259.
23. Li, Y., Yuan, G., Wen, Y., Hu, J., Evangelidis, G., Tulyakov, S., Wang, Y. and Ren, J., 2022. Efficientformer: Vision transformers at mobilenet speed. Advances in neural information processing systems, 35, pp.12934-12949.
24. Makhadmeh, S.N., Al-Betar, M.A., Doush, I.A., Awadallah, M.A., Kassaymeh, S., Mirjalili, S. and Zitar, R.A., 2023. Recent advances in Grey Wolf Optimizer, its versions and applications. IEEE access, 12, pp.22991-23028.
25. Stahler, D.R. and Cassidy, K.A., 2025. Flexibility in group formation and maintenance in grey wolf societies: lessons from Yellowstone. Animal Behaviour, p.123369.