Enhanced Tuberculosis Detection in Chest X-Rays Using Optimal Feature Selection with Hybrid Binary Particle Swarm Optimization (HBPSO) and a Dual Classifier Framework Combining LSTM and Spiking Neural Networks

Gitesh S. Gujrathi[1], Dr. Mukesh Yadav[2]

[1]*Ph.D. Scholar, SAGE University, Indore, India*
[2]*Ph.D.Guide, HOD, Electronics & Telecommunication Department, SAGE University, Indore, India*
Corresponding Author: Gitesh S. Gujrathi, Email: gujrathi87@gmail.com

KEYWORDS
Bag of Words, Cartoon Texture Features, Chest X-Ray, Hybrid Binary Particle Swarm Optimization, Long Short-Term Memory, Multi-Scale Local Binary Patterns, Speeded Up Robust Features, Spiking Neural Networks.

ABSTRACT
Tuberculosis (TB) remains a significant global health challenge, particularly in developing countries, where rapid and accurate diagnosis is crucial for effective treatment. Recent advancements in machine learning and image processing have shown promise in enhancing the detection of TB from chest X-ray images. This research introduces a novel approach that integrates a hybrid feature selection method with a dual classifier framework to improve the accuracy of tuberculosis detection. Feature extraction is conducted using Multi-Scale Local Binary Patterns (MSLBP), Speeded-Up Robust Features (SURF) with Bag of Words (BoW), and Cartoon Texture Features, combining them into a comprehensive feature vector. The Hybrid Binary Particle Swarm Optimization (HBPSO) is employed for optimal feature selection, reducing redundancy and improving classification efficiency. Classification is performed using a dual framework involving a Long Short-Term Memory (LSTM) network and a Spiking Neural Network (SNN), which work in tandem to provide robust predictions. Majority voting is utilized to finalize predictions, ensuring high accuracy and adaptability across varying TB presentations. Extensive evaluation on publicly available datasets shows that this hybrid approach significantly outperforms existing methods, achieving a highest accuracy of 99.73%, along with superior precision, recall, and F1 scores.

1. Introduction

Tuberculosis (TB) remains one of the most pervasive infectious diseases worldwide, significantly affecting global health, particularly in underdeveloped regions where medical resources are scarce. Early detection of TB is crucial for reducing mortality rates, but existing diagnostic methods face challenges in accuracy, speed, and accessibility. Traditional diagnostic techniques, such as Sputum Smear Microscopy and Chest X-ray (CXR) imaging, are widely used; however, their effectiveness often depends on the availability of skilled professionals and robust medical infrastructure. Among these, CXR stands out as a valuable tool for detecting pulmonary TB due to its ability to reveal lung abnormalities associated with the disease. However, accurate interpretation of these images poses a challenge, necessitating automated approaches to enhance diagnostic efficiency and accuracy.

In recent years, computer-aided detection (CAD) systems have emerged as a promising solution for automating TB diagnosis through CXR analysis. These systems typically involve three key stages: feature extraction, feature selection, and classification. Advanced techniques for feature extraction can significantly enhance the detection process by capturing critical patterns and abnormalities in X-ray images that are indicative of TB.

This paper explores a novel TB detection framework using a combined feature set derived from three robust feature extraction techniques: Multi-Scale Local Binary Patterns (MSLBP), Speeded Up Robust Features (SURF) integrated with Bag of Words (BoW), and Cartoon Texture Features. The combined feature set leverages the strengths of each technique: MSLBP captures texture information at multiple scales, SURF with BoW efficiently encodes local keypoints into a compact representation, and Cartoon Texture Features distinguish between smooth and textured regions, which can help in identifying TB-related lung abnormalities such as consolidation, cavitation, or infiltrates. The comprehensive feature set formed through this combination allows for a more detailed and nuanced analysis of lung regions in CXR images, capturing both global and local characteristics.

To further refine the diagnostic process, feature selection is performed using Hybrid Binary Particle Swarm Optimization (HBPSO). Feature selection is critical for reducing the dimensionality of the feature set, removing

redundant or irrelevant features, and improving the computational efficiency of the classification process. HBPSO combines the traditional Particle Swarm Optimization (PSO) algorithm with binary search capabilities, optimizing the selection of features that most effectively contribute to TB detection. This approach ensures that only the most relevant features are retained, enhancing the overall accuracy and performance of the system.

In the final stage, classification is carried out using a Dual Classifier Framework that integrates two advanced neural network models: Long Short-Term Memory (LSTM) and Spiking Neural Networks (SNN). LSTM, known for its ability to model temporal dependencies, is adept at capturing sequential patterns in medical data, which can be beneficial in analyzing subtle variations in CXR images. SNN, on the other hand, mimics the behavior of biological neurons, processing information through spikes and making it highly efficient in handling spatiotemporal data. The combination of these two classifiers offers a powerful and flexible framework for TB detection.

The dual classifier system operates using a majority voting mechanism, where both LSTM and SNN independently make predictions, and the final classification is determined by the majority vote. In the case of a tie, the system defaults to the LSTM prediction, though this strategy can be adjusted based on specific requirements. This hybrid approach enhances classification accuracy and robustness, as it combines the strengths of both classifiers, allowing the system to adapt to varying conditions and disease presentations.

By integrating advanced feature extraction, optimized feature selection, and a sophisticated dual classifier framework, this study aims to develop a reliable and efficient system for automated TB detection using CXR images. The proposed framework has the potential to significantly improve early diagnosis, especially in resource-limited settings, contributing to global efforts in combating this deadly disease.

2. Literature Review

The COVID-19 pandemic has severely impacted global health and economies. With symptoms resembling other respiratory conditions like tuberculosis (TB) and pneumonia, diagnosing COVID-19 has become more complex. The authors of [1] proposed CDC Net, a CNN-based model to detect COVID-19 from chest X-ray images, aiming for early diagnosis. However, integrating radiology systems and ensuring the quality of X-ray films posed significant challenges, reducing the model's practical effectiveness in real-world settings.

The authors of [2] focused on using CNN networks to classify lung diseases, including TB and COVID-19, from chest X-ray images. Their approach utilized models like AlexNet, Darknet, and ResNet with Adam optimizers and 30 epochs for training. Although effective, their study highlighted the need for larger datasets to achieve faster and more accurate classifications, suggesting a limitation in its generalizability to more extensive, diverse data.

The authors of [3] explored TB detection using a backpropagation artificial neural network (ANN), achieving an accuracy of 81.82%. Despite the results, the backpropagation algorithm struggled to converge quickly, a common issue that reduces the overall efficiency and practicality of this technique for real-time TB diagnosis.

The authors of [4] applied multiple machine learning models such as logistic regression (LR), support vector machines (SVM), and random forest (RF) for TB detection, with SVM achieving the highest accuracy of 91.28%. However, these models showed limitations in scalability, particularly when applied to more extensive, varied datasets, making them less adaptable to large-scale medical applications.

The authors of [5] used Histogram of Oriented Gradients (HOG) for feature extraction and KNN/SVM for classification in TB detection. While their model reached 77.95% accuracy, its performance was significantly lower for positive HOG detections, with only 65.75% accuracy, indicating inefficiencies in handling certain types of TB image data.

The authors of [6] used SVM and KNN classifiers on chest CT scans to detect TB, with SVM performing better overall. However, their study was limited by a small dataset, which restricts the generalizability of the findings to larger populations and more complex TB cases.

The authors of [7] developed a deep learning model using ResNet, Xception, and EfficientNet to detect TB from chest X-ray images, using both original and segmented images. Although the study showed promising results, it lacked sufficient validation across diverse feature extraction techniques, limiting its robustness in real-world scenarios where feature variability is common.

The authors of [8] explored a hybrid deep learning approach combining Vision Transformer (ViTs) and EfficientNet models for TB recognition on MRI images, achieving success on the Shenzhen and Montgomery datasets. However, their method struggled with automatic parameter tuning for diverse datasets, making it challenging to apply broadly across various medical imaging data.

The authors of [9] enhanced TB detection from chest X-ray images using the SMOTE algorithm for class balancing, which improved deep learning model performance. Despite this, the method required the application of different data augmentation techniques for other lung diseases, limiting its effectiveness in diagnosing diseases with varying data distributions.

The authors of [10] analyzed lung caverns in CT images using a CNN model trained on ImageCLEF 2022 data. They employed a three-label prediction algorithm to detect lesions, but their method showed inconsistencies in detecting complex lung caverns, reducing its reliability for clinical applications in detecting early TB symptoms.

The authors of [11] used six deep learning models, including ResNet and DenseNet, on ECG data for COVID-19 detection. While they achieved high accuracy in detecting COVID-19, the study was limited in its application to other cardiovascular and respiratory diseases due to the narrow focus on ECG data.

The authors of [12] proposed a random forest-based segmentation and classification model for TB detection using sputum smear samples. While the model achieved high specificity (96.97%), its sensitivity (75.77%) was comparatively lower, pointing to challenges in achieving consistent pixel-level segmentation, which is crucial for accurate TB diagnosis.

The authors of [13] applied a deep neural network-based approach for TB detection, achieving recall and precision rates of 83.78% and 67.55%, respectively. However, the reliance on high-quality input images presented a significant drawback, as real-world medical images often vary in quality, potentially reducing the effectiveness of the model in practical applications.

The authors of [14] utilized the ImageCLEF dataset to develop a machine learning model for TB detection. Their model applied genetic algorithms for feature selection and used SVM for classification. Although the genetic algorithm optimized the model's performance, the large number of features extracted posed a risk of overfitting, which limited its generalizability to other datasets.

Research Gap: While considerable advancements have been made in using machine learning and deep learning techniques for TB detection, several gaps remain. Many of the models suffer from limited datasets, difficulties in feature extraction, and inefficient parameter tuning. Moreover, there is a lack of methods that effectively integrate multiple feature extraction techniques while addressing scalability and adaptability across varied datasets.

To address these gaps, our study introduces a comprehensive TB detection framework that combines advanced feature extraction techniques—Multi-Scale Local Binary Patterns (MSLBP), Speeded Up Robust Features (SURF) with Bag of Words (BoW), and Cartoon Texture Features—into a unified feature set. Feature selection is optimized using Hybrid Binary Particle Swarm Optimization (HBPSO), ensuring the most relevant features are selected. Classification is then carried out using a dual classifier framework consisting of Long Short-Term Memory (LSTM) networks and Spiking Neural Networks (SNN). This approach leverages both temporal dependencies and spike-based processing to enhance classification accuracy, particularly in varying conditions of TB manifestation, making the model more robust and adaptable than existing methods.

3. Proposed Methodology

The methodology proposed in this paper for Tuberculosis (TB) detection using chest X-ray images involves several critical steps, beginning with image acquisition and progressing through pre-processing, feature extraction, feature selection, and classification as shown in Figure 1. The approach combines advanced image processing techniques and machine learning algorithms to optimize the detection process and enhance accuracy. Following subsections describe the proposed work in detail.

3.1 Image Acquisition

The first step in the proposed methodology is image acquisition, where the relevant chest X-ray images are collected. The dataset utilized in this study is sourced from the publicly available datasets published in the

Kaggle Data Science Community. These datasets provide a large number of chest X-ray images that are labeled as either TB-positive or TB-negative, ensuring a diverse set of training data that captures a variety of TB manifestations in the lung.

The specific datasets used include the Shenzhen Hospital X-ray Set and the Montgomery County X-ray Set, which are well-known and widely used for TB detection research. These datasets contain high-resolution frontal chest X-ray images. The images are pre-labeled, with a clear distinction between patients with confirmed TB and healthy patients, which serves as the ground truth for subsequent training and evaluation stages.

Let the dataset be represented as:

$$D = \{(X_1, y_1), (X_2, y_2), \dots, (X_n, y_n)\} \quad (1)$$

Where X_{ni} denotes the i^{th} chest X-ray image, and $y_i \in \{0,1\}$ represents the corresponding label, where 0 indicates TB-negative and 1 indicates TB-positive.

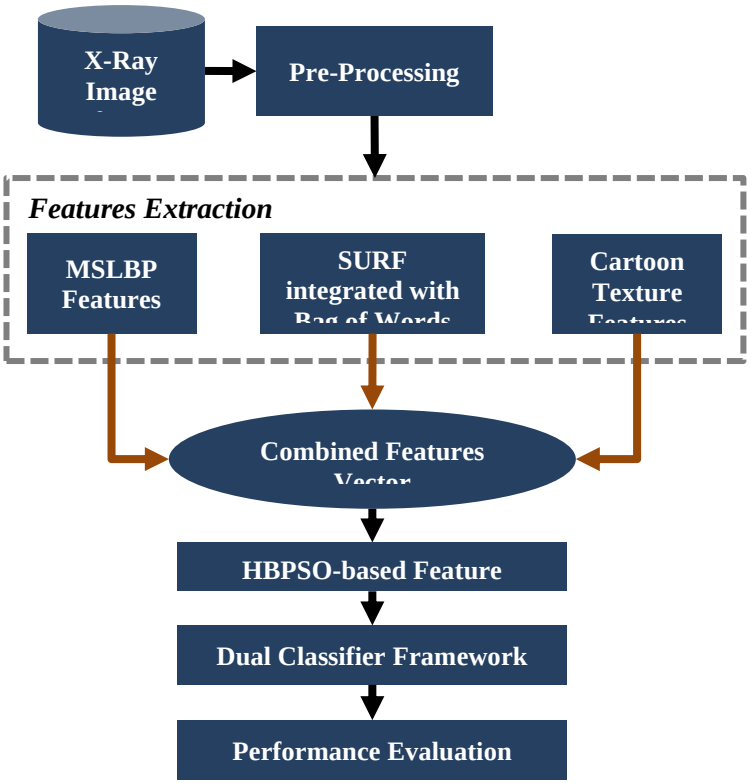

Figure 1: Block diagram of proposed method for tuberculosis diseases recognition

3.2 Pre-Processing

Pre-processing is a vital step to ensure that the images are in a standardized format, free from noise, and prepared for feature extraction. In this methodology, two primary pre-processing techniques are employed: normalization and resizing.

3.2.1 Normalization

Normalization is performed to ensure that pixel intensity values across all images fall within a consistent range, typically between 0 and 1. This helps eliminate variations in brightness and contrast across different X-ray images, allowing the feature extraction process to focus on the relevant patterns within the lung regions.

Normalization is mathematically expressed as:

$$X_i^{norm} = \frac{X_i - \min(X_i)}{\max(X_i) - \min(X_i)} \quad (2)$$

Where:

- X_i is the original image matrix containing pixel intensity values,
- $\min(X_i)$ and $\max(X_i)$ are the minimum and maximum pixel intensities in the image,
- X_i^{norm} represents the normalized image with pixel values scaled between 0 and 1.

This normalization process ensures that the images are consistent in terms of intensity, facilitating accurate feature extraction and comparison.

3.2.2 Resizing

Inconsistent image resolutions are addressed by resizing all chest X-ray images to a fixed dimension of 256×256 pixels. Resizing ensures that the input data is uniform, reducing computational complexity and standardizing the images for subsequent analysis.

Mathematically, resizing an image X_i from its original dimensions $W_i \times H_i$ to the target resolution $W \times H$ can be represented as:

$$X_i^{resized} = f_{resize}(X_i, W, H) \quad (3)$$

Where f_{resize} is the interpolation function used for resizing, typically applying bilinear or bicubic interpolation techniques. This method preserves the structural integrity of the lungs and TB-specific patterns while adjusting the image size.

Upon completing the pre-processing, the dataset is represented as:

$$D_{preprocessed} = \{(X_1^{resized}, y_1), (X_2^{resized}, y_2), \dots, (X_n^{resized}, y_n)\} \quad (4)$$

Where $X_i^{resized}$ is the resized and normalized version of each image, ready for feature extraction.

3.3 Feature Extraction

The feature extraction process aims to capture meaningful patterns from chest X-ray images that indicate the presence of tuberculosis (TB). The proposed methodology applies multiple feature extraction techniques, each contributing to a comprehensive representation of image characteristics. The extracted features are further combined to form a robust feature set, improving the accuracy and performance of the TB detection model. The following subsections describe the techniques used for feature extraction, starting with Multi-Scale Local Binary Patterns (MSLBP), SURF with Bag of Words (BoW), and Cartoon Texture Features, concluding with their combination into a single feature vector.

3.3.1 Multi-Scale Local Binary Patterns (MSLBP) Based Feature Extraction

Multi-Scale Local Binary Patterns (MSLBP) is an extension of the Local Binary Patterns (LBP) technique, designed to capture texture information at multiple scales. MSLBP enhances the detection of fine-grained textures and patterns in X-ray images, which are essential for identifying early-stage TB symptoms such as cavities, infiltrates, and consolidations in lung tissue.

Local Binary Patterns (LBP): In the standard LBP method, each pixel I_c in the image is compared to its surrounding neighbors. For a pixel located at coordinates (x, y), LBP is calculated by:

$$LBP(x, y) = \sum_{p=0}^{P-1} s(I_p - I_c) \cdot 2^p \quad (5)$$

Where I_p represents the intensity of a neighboring pixel, I_c is the intensity of the center pixel, $s(x)$ is a step function defined as:

$$s(x) = \begin{cases} 1 & if\ x \geq 0 \\ 0 & if\ x < 0 \end{cases} \quad (6)$$

The result is a binary number representing the local texture, which is converted into a decimal value.

Multi-Scale Extension: In MSLBP, the same LBP operation is applied over multiple scales to capture texture information at different resolutions. Let R represent the radius of the neighborhood, and P the number of neighbors. MSLBP operates over multiple scales by varying R and applying the LBP operator at each scale. The final MSLBP descriptor is the concatenation of the LBP histograms from each scale:

$$MSLBP = \left[H_{LBP}^{R_1}, H_{LBP}^{R_2}, \ldots, H_{LBP}^{R_n}\right] \quad (7)$$

Where $H_{LBP}^{R_i}$ is the LBP histogram at scale R_i, and n is the number of scales. This multi-scale approach enables capturing both fine and coarse texture patterns, improving the representation of TB-related abnormalities.

3.3.2 SURF (Speeded Up Robust Features) with Bag of Words (BoW)

SURF is a powerful feature extraction method known for its robustness to scale and rotation changes. It detects key points in the image and computes local descriptors that describe the surrounding region of each key point. The Bag of Words (BoW) model is applied to transform these descriptors into a global feature vector representing the entire image.

SURF Keypoint Detection and Descriptor Extraction: SURF detects key points based on the Hessian matrix approximation, which allows it to locate interest points efficiently. The Hessian matrix $H(x, \sigma)$ for a point $x = (x, y)$ in an image I is given by:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{yx}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (8)$$

Where L_{xx}, L_{xy}, L_{yx} and L_{yy} are the second-order partial derivatives of the Gaussian-smoothed image at scale σ. Key points are identified as local extrema in the determinant of the Hessian matrix. Once key points are detected, SURF computes descriptors for each point, which describe the intensity distribution in the surrounding region.

Bag of Words (BoW) Model: The BoW model is applied to quantize the SURF descriptors into a set of visual words. The process involves clustering the descriptors into k clusters using k-means clustering. Each cluster center represents a visual word in the vocabulary. For an image, the frequency of each visual word is counted, forming a histogram of visual words:

$$BoW = [f_1, f_2, \ldots, f_k] \quad (9)$$

Where f_i is the frequency of the i^{th} visual word. This histogram serves as the global feature vector for the image, capturing the distribution of key points across different regions of the image.

3.3.3 Cartoon Texture Features

Cartoon Texture Features aim to separate an image into its cartoon and texture components. This decomposition helps in isolating smooth regions (cartoon component) and fine details (texture component), which can improve the detection of TB-specific patterns in the lungs.

Cartoon-Texture Decomposition: The image I is modeled as the sum of two components:

$$I = I_{cartoon} + I_{texture} \quad (10)$$

Where $I_{cartoon}$ contains the smooth, piecewise-constant regions, and $I_{texture}$ contains the high-frequency details or fine textures. The decomposition is typically performed using a variation of the Rudin-Osher-Fatemi (ROF) model, which minimizes an energy functional:

$$E(I_{cartoon}, I_{texture}) = \int_{\Omega}^{1} |\nabla I_{cartoon}| + \lambda \int_{\Omega}^{1} |1 - I_{cartoon} - I_{texture}|^2 \quad (11)$$

Where $\nabla I_{cartoon}$ represents the gradient of the cartoon component, and λ is a regularization parameter controlling the trade-off between the cartoon and texture components. Features are then extracted from both $I_{cartoon}$ and $I_{texture}$, such as texture energy and smoothness characteristics, providing valuable information about the lung regions affected by TB.

3.3.4 Combined Feature Vector

The final step in the feature extraction process is to combine the features obtained from MSLBP, SURF with BoW, and Cartoon Texture features into a single, comprehensive feature vector. Let F_{MSLBP}, F_{SURF}, and $F_{Cartoon}$ denote the feature vectors extracted from MSLBP, SURF with BoW, and Cartoon Texture methods, respectively. The combined feature vector is represented as:

$$F_{combined} = [F_{MSLBP}, F_{SURF}, F_{Cartoon}] \quad (12)$$

This concatenated feature vector captures a rich set of information from the chest X-ray images, including texture, key point distribution, and structural components. The combined feature vector serves as the input to the subsequent feature selection and classification stages, enhancing the ability to distinguish between healthy and TB-affected lungs.

The use of multiple feature extraction techniques ensures that the model captures both global and local image characteristics, improving the robustness and accuracy of TB detection. This comprehensive feature set forms the foundation for feature selection and classification in the later stages of the proposed methodology.

3.4 Feature Selection using Hybrid Binary Particle Swarm Optimization (HBPSO)

After extracting the combined feature vector $F_{combined}$ from the feature extraction phase, the feature set may still contain redundant or irrelevant features, which can degrade the performance of the classifier. To address this, an HBPSO approach is employed for feature selection, which aims to optimize the selection of relevant features from the combined feature vector.

PSO is a population-based optimization algorithm inspired by the social behavior of birds flocking or fish schooling. Each potential solution, known as a particle, moves through the search space by adjusting its velocity based on its own experience and the experience of its neighbours. The PSO algorithm maintains a population of particles where each particle has:

- A position vector $X = [x_1, x_2, \dots, x_d]$, which represents a possible solution to the optimization problem (in this case, a feature subset).
- A velocity vector $V = [v_1, v_2, \dots, v_d]$, which dictates how the particle moves in the search space.
- A personal best position P_{best}, representing the best solution the particle has encountered.
- A global best position G_{best}, representing the best solution found by the entire swarm.

For feature selection, a binary version of PSO is used where each element of the position vector $x_i \in \{0,1\}$ represents whether the corresponding feature is selected (1) or not (0).

HBPSO Algorithm: In the proposed methodology, HBPSO enhances traditional Binary PSO by incorporating mutation strategies from genetic algorithms to avoid premature convergence and improve global search capabilities.

The update rules for velocity and position in the binary search space are formulated as follows:

Velocity Update:

$$v_i(t+1) = wv_i(t) + c_1 r_1 [P_{best}(t) - x_i(t)] + c_2 r_2 [G_{best}(t) - x_i(t)] \quad (13)$$

Where:

- w is the inertia weight controlling the balance between exploration and exploitation.
- c_1 and c_2 are cognitive and social coefficients, respectively.
- r_1 and r_2 are random values drawn from a uniform distribution.

Position Update: The position update is based on a probabilistic sigmoid function applied to the velocity:

$$S(v_i(t)) = \frac{1}{1+e^{-v_i(t)}} \quad (14)$$

The position $x_i(t)$ is updated as follows:

$$x_i(t+1) = \begin{cases} 1 & if\ r_3 < S(v_i(t+1)) \\ 0 & otherwise \end{cases} \quad (15)$$

Where r_3 is a random value drawn from a uniform distribution.

The objective of HBPSO is to maximize the classification accuracy while minimizing the number of selected features. The fitness function f for each particle is formulated as:

$$f(X) = \alpha \cdot \text{Accuracy}(X) - \beta \cdot \frac{|X|}{d} \quad (16)$$

Where:

- $\text{Accuracy}(X)$ is the classification accuracy for the feature subset represented by X.
- $|X|$ is the number of selected features.
- α and β are weight factors controlling the trade-off between accuracy and the number of selected features.

The selected feature subset F_{HBPSO} obtained from HBPSO reduces the dimensionality of the combined feature vector $F_{combined}$, which improves computational efficiency and classification performance.

3.5 Classification

After feature selection using the HBPSO, the selected feature subset F_{HBPSO} is passed to a classification framework consisting of two advanced neural network architectures: the Long Short-Term Memory (LSTM) classifier and the Spiking Neural Network (SNN) classifier. The classification process aims to leverage the strengths of both models to improve prediction accuracy, with a final decision made by a Dual Classifier Framework.

3.5.1 Long Short-Term Memory (LSTM) Classifier

LSTM networks are a type of recurrent neural network (RNN) designed to handle sequential data. Unlike traditional RNNs, LSTMs can maintain long-term dependencies by using gating mechanisms, which allow the network to retain or forget information over time. This is particularly useful for the temporal processing of feature vectors, capturing sequential dependencies within the extracted feature set F_{HBPSO}.

LSTM Unit Structure: An LSTM unit consists of three gates:

- Forget Gate f_t
- Input Gate i_t
- Output Gate o_t

Each gate controls a different part of the information flow, ensuring relevant information is stored in memory and irrelevant information is discarded.

- *Forget Gate:* This gate controls how much of the previous cell state C_{t-1} is retained. It is computed as:

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, F_{HBPSO}] + b_f\big) \quad (17)$$

Where W_f is the weight matrix, h_{t-1} is the hidden state from the previous time step, and F_{HBPSO} is the current feature input.

- *Input Gate:* This gate determines how much of the new information from the current input should be stored in the cell state. It is calculated as:

$$i_t = \sigma(W_i \cdot [h_{t-1}, F_{HBPSO}] + b_i) \quad (18)$$

The candidate cell state $\tilde{C}_t$ is computed as:

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, F_{HBPSO}] + b_C) \quad (19)$$

- *Cell State Update:* The new cell state C_t is updated as a combination of the previous cell state and the new information:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (20)$$

- *Output Gate:* This gate controls the amount of information passed from the cell state to the next hidden state h_t. The output gate is calculated as:

$$o_t = \sigma(W_o \cdot [h_{t-1}, F_{HBPSO}] + b_o) \quad (21)$$

The final hidden state h_t is given by:

$$h_t = o_t \cdot \tanh(C_t) \quad (22)$$

Here, h_t represents the network's memory at time step t, encoding important sequential features from the input.

LSTM Classification: After processing the input through multiple LSTM layers, the final hidden state h_T at the last time step is passed through a fully connected layer with a softmax activation function for classification. The softmax function outputs the probability distribution over the possible classes y_{LSTM}:

$$y_{LSTM} = \text{softmax}(W_{out} \cdot h_T + b_{out}) \quad (23)$$

Where W_{out} and b_{out} are the weight and bias of the output layer.

3.5.2 Spiking Neural Network (SNN) Classifier

SNNs are biologically inspired networks that simulate the spiking behaviour of neurons. Unlike traditional neural networks, SNNs operate in continuous time and use discrete spikes to transmit information. This makes them well-suited for temporal data processing, which complements the temporal nature of the LSTM classifier.

Neuron Dynamics: The most commonly used neuron model in SNNs is the Leaky Integrate-and-Fire (LIF) model. The neuron's membrane potential $V_m(t)$ evolves according to the differential equation:

$$\tau_m = \frac{dV_m(t)}{dt} = -V_m(t) + I(t) \quad (24)$$

Where:

- τ_m is the membrane time constant.
- $V_m(t)$ is the membrane potential at time t.
- $I(t)$ is the input current, which in this case is derived from the feature vector F_{HBPSO}.

A neuron emits a spike whenever the membrane potential $V_m(t)$ exceeds a threshold V_{th}. After spiking, the membrane potential is reset to its resting value V_{rest}.

Spike Encoding and Decoding: The input feature vector F_{HBPSO} is first encoded into spike trains using a spike encoding mechanism, such as rate coding or temporal coding:

- *Rate Coding:* The firing rate of a neuron corresponds to the magnitude of the feature value. Higher values lead to a higher spike rate.
- *Temporal Coding:* The timing of individual spikes conveys information, with earlier spikes representing higher values.

The spikes are propagated through multiple layers of spiking neurons, with each layer applying its own dynamics. The output layer produces a spike train, which is decoded to obtain the classification result y_{SNN}. In this case, the spike count or timing of the first spike is used to decode the final output.

3.5.3 Dual Classifier Framework

The classification task is finalized using a dual classifier framework, which integrates the predictions from both the LSTM and SNN classifiers. This framework capitalizes on the complementary strengths of the two models, where the LSTM is better at capturing long-term dependencies, and the SNN excels in processing temporal spike patterns.

Majority Voting Scheme: To combine the predictions from both classifiers, a majority voting scheme is employed:

- Let y_{LSTM} and y_{SNN} represent the class predictions from the LSTM and SNN classifiers, respectively.

- The final class prediction y_{final} is determined as follows:

$$y_{final} = \begin{cases} y_{LSTM}, & if\ y_{LSTM} = y_{SNN} \\ \arg\max(P_{LSTM}, P_{SNN}) & if\ y_{LSTM} \neq y_{SNN} \end{cases} \quad (25)$$

Where P_{LSTM} and P_{SNN} are the probabilities or confidence scores for the predicted classes from each model.

Final Decision: The final decision process prioritizes agreement between the two classifiers. In case of a disagreement, the classifier with the higher confidence score is given precedence. This combination of LSTM and SNN ensures robust classification performance, with the LSTM model handling sequential dependencies effectively and the SNN model leveraging the temporal characteristics of the data.

By employing this dual-classifier framework, the overall classification system benefits from the complementary strengths of each model, improving accuracy, reducing the error rate, and providing a more robust solution to the classification task.

Figure 2: Dual Classification Framework Integrating LSTM and SNN Outputs through an Ensemble Approach

4. Simulation and Results

4.1 Dataset

The dataset utilized in this research consists of chest radiograph images sourced from two publicly available datasets hosted on the KAGGLE platform [15-16]. These datasets, when combined, provide a total of 7,662 X-ray images, each representing anonymized patients. The primary motivation for merging these datasets is to increase the variability in the training data, thereby enhancing the generalization capabilities of the proposed model across diverse image sources during both validation and testing.

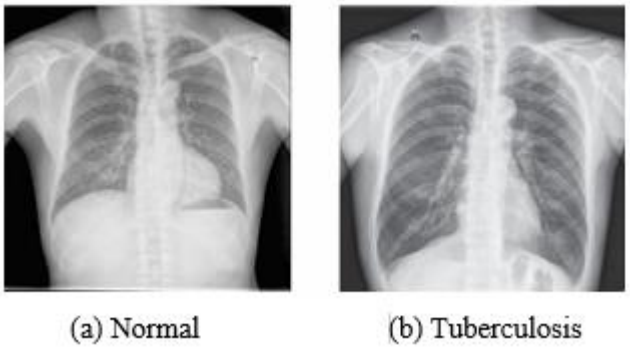

Figure 3: Images of a healthy patient and a patient with tuberculosis [16]

To ensure data quality, a thorough inspection of each image was conducted. Images exhibiting noise, such as illegible scans, presence of artifacts (e.g., superimposed text, boxes), incorrect positioning, or those that did not properly fit into the predefined categories, were discarded. After this cleaning process, the dataset was reduced to 5,748 valid images. These comprise 2,905 images of healthy individuals and 2,843 images of tuberculosis patients.

For the purpose of training and evaluation, the dataset was split into two subsets: 80% of the data was allocated to the training set, while the remaining 20% was reserved for testing. Furthermore, the training set was partitioned again, with 20% of it used for validation during the cross-validation phase of model training. This split ensures that the model is properly evaluated on unseen data, improving its robustness and reducing

overfitting. Figure 3 shows sample images from the dataset.

4.2 Results

The results, as illustrated in Figures 4 to 8, demonstrate that integrating hybrid features with feature selection significantly improves tuberculosis detection in chest X-rays. Notably, the dual classifier framework (Figure 8) combining LSTM and SNN achieves the highest accuracy, surpassing individual classifiers.

Additionally, both LSTM and SNN classifiers (Figures 4-7) benefit from feature selection, resulting in reduced misclassification and more robust detection performance. These findings underscore the effectiveness of combining advanced feature extraction with optimized classifiers for accurate TB detection.

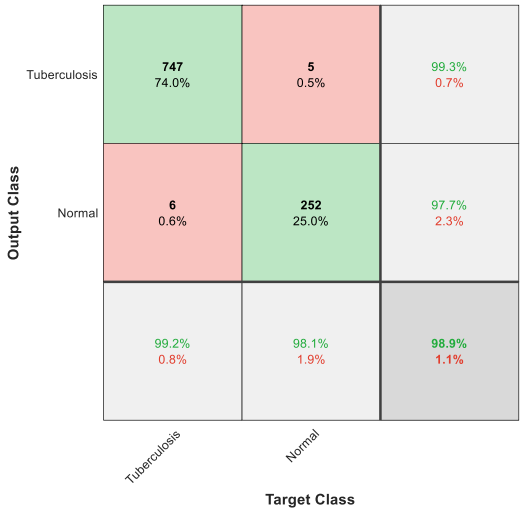

Figure 4: Confusion matrix plot for Tuberculosis Detection in Chest X-Rays using Hybrid Features and SNN classifier

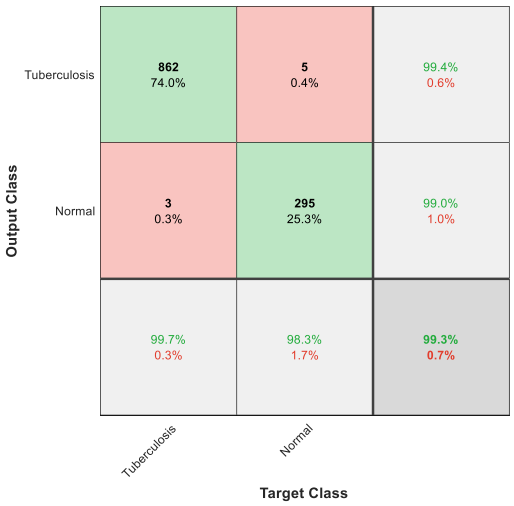

Figure 5: Confusion matrix plot for Tuberculosis Detection in Chest X-Rays using Hybrid Features and LSTM classifier

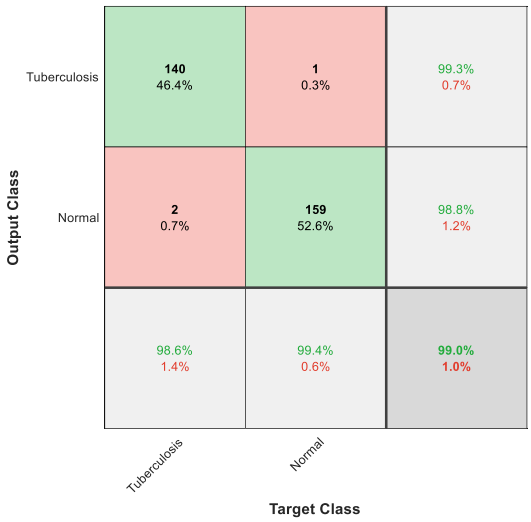

Figure 6: Confusion matrix plot for Tuberculosis Detection in Chest X-Rays using Hybrid Features, Feature Selection and SNN classifier

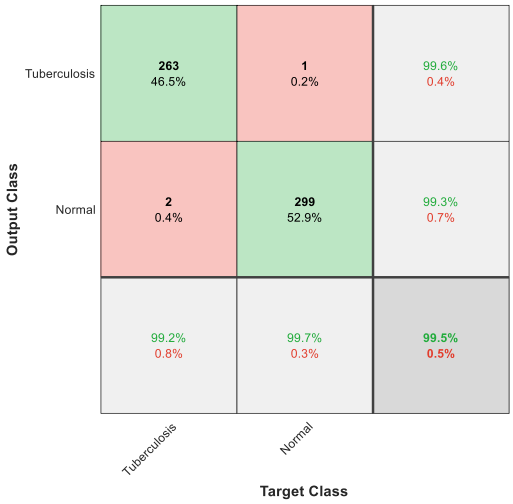

Figure 7: Confusion matrix plot for Tuberculosis Detection in Chest X-Rays using Hybrid Features, Feature Selection and LSTM classifier

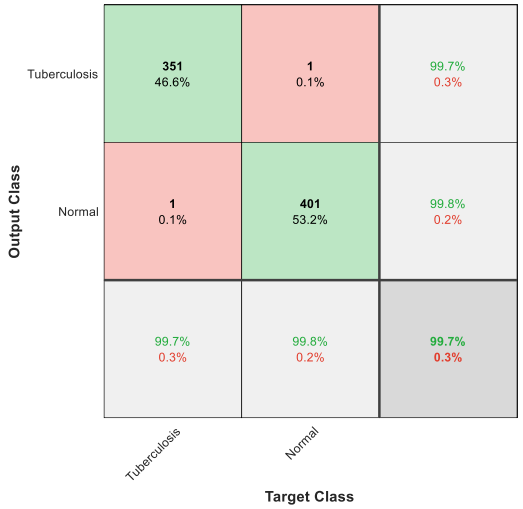

Figure 8: Confusion matrix plot for Tuberculosis Detection in Chest X-Rays using Hybrid Features, Feature Selection and dual classifier

Table 1: Performance Comparison of Tuberculosis Detection Methods Using Hybrid Features and Different Classifiers

Parameters	Hybrid Features + SNN	Hybrid Features + LSTM	Hybrid Features + Feature Selection + SNN	Hybrid Features + Feature Selection + LSTM	Hybrid Features + Feature Selection + Dual Classifier
Accuracy	0.9772	0.9828	0.9901	0.9947	0.9973
Error Rate	0.0228	0.0172	0.0099	0.0053	0.0027
Sensitivity	0.9779	0.9829	0.9859	0.9925	0.9972
Specificity	0.9924	0.9943	0.9938	0.9967	0.9975
Precision	0.9777	0.9829	0.9929	0.9962	0.9972
False Positive Rate	0.0076	0.0057	0.0062	0.0033	0.0025
F1-Score	0.9777	0.9829	0.9894	0.9943	0.9972
Matthews Correlation Coefficient (MCC)	0.9701	0.9772	0.9801	0.9893	0.9947
Kappa Statistics	0.9393	0.9542	0.9801	0.9893	0.9947

Table 1 compares the performance of various tuberculosis (TB) detection methods using hybrid features and different classifiers. It shows improvements across different evaluation metrics, with the dual classifier achieving the highest accuracy of 99.73%.

Table 2 compares previous TB detection research with the proposed hybrid feature-based methods. The proposed methods show superior performance across different evaluation metrics, with the dual classifier outperforming the rest.

Table 2: Comparative Results with Previous Research Works and Proposed Hybrid Methods

Citation	Method Used	Accuracy	Precision	Sensitivity (Recall)	Specificity	F1-Score
Authors of [3]	Backpropagation ANN for TB Detection	81.82%	--	--	--	--
Authors of [4]	SVM for TB Detection	91.28%	--	--	--	--
Authors of [5]	HOG + KNN/SVM for TB Detection	77.95%	--	--	--	--
	Positive HOG Detections	65.75%	--	--	--	--
Authors of [12]	Random Forest-based Segmentation & Classification for TB	--	--	75.77%	96.97%	--
Authors of [13]	Deep Neural Network for TB Detection	--	67.55%	83.78%	--	--
Proposed Method	**Hybrid Features + Feature Selection + SNN**	**99.01%**	**99.29%**	**98.59%**	**99.38%**	**98.94%**
Proposed Method	**Hybrid Features + Feature Selection + LSTM**	**99.47%**	**99.62%**	**99.25%**	**99.67%**	**99.43%**
Proposed Method	**Hybrid Features + Feature Selection + Dual Classifier**	**99.73%**	**99.72%**	**99.72%**	**99.75%**	**99.72%**

5. Conclusion

This research presents a robust and effective methodology for the detection of tuberculosis in chest X-ray images, leveraging a hybrid feature extraction and selection process alongside a dual classifier framework. By integrating features such as MSLBP, SURF with BoW, and Cartoon Texture Features, and optimizing them through HBPSO, this approach ensures both comprehensive and refined feature representation. The combination of LSTM and SNN in the dual classifier framework enhances classification performance by utilizing both temporal sequence learning and spike-based processing, leading to more accurate and flexible detection. Experimental results demonstrate that the proposed method achieves a highest accuracy of 99.73% compared to traditional models. The presented framework offers a valuable contribution to the computer-aided diagnosis of TB and could potentially be extended to the detection of other lung diseases in future work.

References

[1] M H. Malik, T. Anees, M. Din, and A. Naeem, "CDC_Net: Multi-Classification Convolutional Neural Network Model for Detection of COVID-19, Pneumothorax, Pneumonia, Lung Cancer, and Tuberculosis using Chest X-rays", Multimedia Tools and Applications, Vol.82, No.9, pp.13855-13880, 2023.

[2] S.H. Karaddi and L.D. Sharma, "Automated Multi-Class Classification of Lung Diseases From CXR-Images using Pre-Trained Convolutional Neural Networks", Expert Systems with Applications, Vol.211, p.118650, 2023.

[3] V. Lestari, H. Mawengkang, and Z. Situmorang, "Artificial Neural Network Backpropagation Method to Predict Tuberculosis Cases", Sinkron: jurnal dan penelitian teknik informatika, Vol.7, No.1, pp.35-47, 2023.

[4] H.W. Gichuhi, M. Magumba, M. Kumar, and R.W. Mayega, "A Machine Learning Approach to Explore Individual Risk

Factors for Tuberculosis Treatment Non-Adherence In Mukono District", PLOS Global Public Health, Vol.3, No.7, p.e0001466, 2023.

[5] A.R. Lubis, S. Prayudani, Y. Fatmi, and Y.Y. Lase, "Detection of HOG Features on Tuberculosis X-Ray Results Using SVM and KNN", In: Proc. of 2nd International Conf. On Innovative and Creative Information Technology (ICITech), IEEE, pp.25-29, 2021.

[6] T. Shakya, R.B. Jeyavathana, and P.K. Kumar, "Improved Accuracy in Automatic Detection of Tuberculosis Disease from Lung CT images using Support Vector Machine Classifier over K-Nearest Neighbours Classifier", In: Proc. of International Conf. on Cyber Resilience (ICCR), IEEE, pp.1-5, 2022.

[7] S.I. Nafisah and G. Muhammad, "Tuberculosis detection in chest radiograph using convolutional neural network architecture and explainable artificial intelligence", Neural Computing and Applications, Vol.36, No.1, pp.111-131, 2024.

[8] L.B. Ammar, K. Gasmi, and I.B. Ltaifa, "ViT-TB: ensemble learning based ViT model for tuberculosis recognition", Cybernetics and Systems, Vol.55, No.3, pp.634-653, 2024.

[9] L. Venkataramana, D.V.V. Prasad, S. Saraswathi, C.M. Mithumary, R. Karthikeyan, and N. Monika, "Classification of COVID-19 from tuberculosis and pneumonia using deep learning techniques", Medical & Biological Engineering & Computing, Vol.60, No.9, pp.2681-2691, 2022.

[10] T. Asakawa, R. Tsuneda, K. Shimizu, T. Komoda, and M. Aono, "Caverns Detection and Caverns Report in Tuberculosis: lesion detection based on image using YOLO-V3 and median based multi-label multi-class classification using SRGAN", In CLEF (Working Notes), pp.1347-1354, 2022.

[11] T. Rahman, A. Akinbi, M.E. Chowdhury, T.A. Rashid, A. Şengür, A. Khandakar, K.R. Islam, and A.M. Ismael, "COV-ECGNET: COVID-19 detection using ECG trace images with deep convolutional neural network", Health Information Science and Systems, Vol.10, No.1, pp.1-16, 2022.

[12] S. Ayas and M. Ekinci, "Random forest-based tuberculosis bacteria classification in images of ZN-stained sputum smear samples", Signal, Image and Video Processing, Vol.8, pp.49-61, 2023.

[13] M. Singh, G.V. Pujar, S.A. Kumar, M. Bhagyalalitha, H.S. Akshatha, B. Abuhaija, A.R. Alsoud, L. Abualigah, N.M. Beeraka, and A.H. Gandomi, "Evolution of machine learning in tuberculosis diagnosis: a review of deep learning-based medical applications", Electronics, Vol.11, No.17, p.2634, 2022.

[14] O. Hrizi, K. Gasmi, I. Ben Ltaifa, H. Alshammari, H. Karamti, M. Krichen, L. Ben Ammar, and M.A. Mahmood, "Tuberculosis disease diagnosis based on an optimized machine learning model", Journal of Healthcare Engineering, Vol.2022, No.1, p.8950243, 2022.

[15] T. Rahman, A. Khandakar, M.A. Kadir, K.R. Islam, K.F. Islam, R. Mazhar, T. Hamid, M.T. Islam, S. Kashem, Z.B. Mahbub, and M.A. Ayari, "Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization", IEEE Access, Vol.8, pp.191586-191601, 2020.

[16] N. Singh and S. Hamde, "Tuberculosis detection using shape and texture features of chest X-rays". In: Innovations in Electronics and Communication Engineering: Proc. of the 7th ICIECE, Springer Singapore, pp.43-50, 2019.