Computational Engineering and Technology Innovations

www.ceti.reapress.com

Comput. Eng. Technol. Innov. Vol. 2, No. 3 (2025) 171–184.

Paper Type: Original Article

Quantum Bat Algorithm Using Ensemble Bayesian Neural Networks for Cad-Based Lung Cancer Prediction

Seyyed Ahmad Edalatpanah[1,*], Eisa Abdolmaleki[2], Victoria Nozick[3]

[1] Department of Applied Mathematics, Ayandegan Institute of Higher Education, Tonekabon, Iran; s.a.edalatpanah@aihe.ac.ir.

[2] Department of Mathematics, Tonekabon Branch, Islamic Azad University, Tonekabon, Iran; Abdolmaleki.eisa@gmail.com.

[3] Operations and Information Management Group, Aston Business School, Aston University, B4 7ET Birmingham, United Kingdom; victorianozick79@gmail.com.

Citation:

Received: 10 September 2025 *Revised: 24 November 2025* *Accepted: 03 April 2025*	Edalatpanah, S. A., Abdolmaleki, E., & Nozick, V. (2025). Quantum bat algorithm using ensemble bayesian neural networks for cad-based lung cancer prediction. *Computational Engineering and Technology Innovations, 2*(3), 171-184.

Abstract

Lung malignancies have raised global death rates, necessitating early therapies. This uses Computer Tomography (CT) pictures to identify lung cancer. The first stage is to automatically split Regions of Interest (ROI) when cancerous spots are discovered in order to increase diagnosis Accuracy (AC). When maximizing feature extractions, the Quantum inspired Bat Optimization Algorithm (QBOA) considers only relevant attributes. Ensemble Bayesian Neural Networks (EBNNs), which are used in medical applications to categorize data, may make good predictions while providing imprecise estimations. The LIDC-IDRI dataset was used in this work to verify the suggested approach, which is highly accurate in detecting and categorizing stages of lung cancer. Lung malignancies may be successfully identified using segmentations, optimum feature extractions, and ensemble-based classification, allowing for early intervention and improved patient care. Furthermore, the outcomes indicate how quickly and effectively the framework can detect and diagnose lung cancer. This unique technology offers a trustworthy and effective option for diagnosing early lung cancer, opening the path for improved clinical decision-making and patient outcomes.

Keywords: Computer tomography, Region of interest, Quantum-inspired bat optimization algorithm, Ensemble bayesian neural network.

1 | Introduction

Lung malignancies are rapidly growing illnesses defined by aberrant cell growths, making their early diagnosis critical to improve survival chances [1]. ML classifiers such as Decision Trees (DTs), Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and Naïve Bayes (NBs) are routinely used to reliably classify benign and malignant lung nodules using image processing. This work highlights many image processing approaches for lung cancer prognosis, including pre-processing (noise reductions and improvements), segmentation, feature extractions, binarizations, and picture acquisitions. Lung cancer

Corresponding Author: s.a.edalatpanah@aihe.ac.ir

https://doi.org/10.48314/ceti.vi.75

remains one of the major causes of cancer-related mortality worldwide, and early detection is widely recognized as a key factor for improving treatment effectiveness and patient survival. Recently, the rapid development of computational intelligence and machine learning has provided promising opportunities for enhancing lung cancer diagnosis and prediction through data-driven models. In particular, recent research indicates that machine learning-based techniques have expanded substantially in lung cancer studies, reflecting a growing trend toward developing intelligent screening and diagnostic frameworks. These approaches can support more efficient clinical decision-making by improving prediction Accuracy (AC) and reducing manual diagnostic burden, thereby contributing to more effective and cost-efficient healthcare systems [2].

In light of the previous empirical literature, Lung cancers [3] are known as one of the top causes of mortality globally and characterized by unchecked cell proliferation in lung tissue. The high prevalence of risk factors, such as smoking, and delayed diagnosis are the main causes of this high death rate. Although there is no way to completely avoid lung cancer, increasing patient survival rates requires early identification. Several classification methods, including Naive Bayes, SVMs, DTs, and Logistic Regression, are to be evaluated in this study. To identify the most effective method for early lung cancer diagnosis. By comparing these algorithms, the researchers hope to improve predicted AC and facilitate prompt intervention for at-risk patients. Lung cancer [4] is the leading cause of cancer-related fatalities globally, with smoking serving as the primary risk factor. Most lung tumors begin in the cells that border the airways. Using a lung cancer dataset from the UCI ML Repository, this work focuses on recognizing lung cancer using several classification algorithms. The study looks at how effectively algorithms such as NBs, neural networks, K-Nearest Neighbors (KNN), logistic regression, DTs, Random Forest, and NBs predict lung cancer. These methodologies are evaluated based on performance parameters such as ROC curve analysis, F1 score, Recall (RE), AC, and precision. Dimensionality reduction utilizing Principal Component Analysis (PCA), as well as visualization approaches such as clustering and classification trees, are being researched to increase diagnostic AC. Based on recent research, Lung cancers [5] are among the worst diseases, needing an early and accurate diagnosis. Artificial intelligence has considerably enhanced medical imaging analysis and sickness diagnosis by lowering the chance of human interpretation errors. This study provides a comprehensive review of studies completed over the last five years on the diagnosis of lung cancer and the classification of lung nodules. Fifty studies were chosen from significant databases including Science Direct, Scopus, Web of Science, and IEEE for thorough assessments. This work aims to outline recent developments, look at present challenges, and offer suggestions for further research for giving researchers thorough perspectives and access to crucial datasets.

Drawing from these loopholes, the goal of this work is to combine algorithmic sophisticated pre-processes, feature extractions, and classifications for predicting and classifying lung cancers in Computer Tomography (CT) images. Usage of Regions of Interest (ROI) based segmentations can precisely identify tumor locations resulting in improved diagnostic accuracies. Computational loads are lowered for efficiency and relevance of certain characteristics by Quantum inspired Bat Optimization Algorithm (QBOA) which enhance feature extractions. Ensemble Bayesian Neural Networks (EBNNs) are used in classifications, providing both uncertainty estimates and robust predictions. This approach is very helpful for medicinal applications. LIDC-IDRI dataset was used to validate the suggested method where exceptional accuracies in identifying and classifying stages of lung cancer were achieved. Thus, by providing precise and fast information, it encourages early diagnosis and improved patient treatment. The rest of this paper is organized as follows: A concise overview of research on preprocessing, feature extraction, and illness classification is given in Section 2. Section 3 outlines the recommended approach for the EBNNs. The findings of the experimental performance analysis are presented in Section 4. Section 5 concludes with a summary of the results.

2 | Literature Review

Several past studies have conducted on the prediction and classification of lung cancers. In this regard, this part of our study is dedicated to the studies in these fields.

Kasinathan and Jayakumar suggested cloud-integrated diagnostic tool that boosted lung tumor diagnosis while minimizing radiologists burdened and enhancing diagnostic precision. Using data from LIDC-IDRI and DICOM lung CT datasets, the system identified lung cancer stages using multilayer convolutional neural networks (M-CNNs) and segmented them using active contours. It outperformed other methods in terms of obtained AC, RE, and precision values and with an average classification AC of 98.6% and ranges of 97% to 99.1% for precision and RE values. These results show that Cloud-LTDSC is an accurate way to detect lung cancer [6].

Ma et al. [7] presented a versatile, comprehensible graph-based approach, to improve survival predictions in non-small cell lung cancers by combining clinical and multimodal data, including CT scans. SAGE convolution layers with self-attentions were used in constructions of unique multi-modal Fusion Graph Convolutional Networks (FGCNs) for enhanced convergences and more accurately match expected risks with actual outcomes. The model also used TopKPooling to remove noise and feature redundancy. Evaluations using publicly available datasets for the diagnosis of non-small cell lung cancers revealed that the model performed better, with a C-index of 0.76 in predictions, suggesting robustness and effectiveness.

Vishwa Kiran et al. [8] suggested lung cancer identifications with MLDS-LCDC model. The model used Gaussian Filters (GF) to pre-process CT images from lung cancer datasets. Subsequently, nodules were segmented using oriented FAST and rotated BRIEF ORB algorithms which extracted features based on Normalized Cuts (Ncuts) approach. Finally, Sunflower Optimization-based Wavelet Neural Networks (SFO-WNN) identified lung cancers where experimental results for classifications of lung cancers confirmed the model's diagnostic performances with values of 97.01% for RE, 98.64% for specificity (SP), and AC of 98.11% were better than alternative methods.

Long-Short Term Memory (LSTM) Networks, Support Vector Regressions (SVRs), and Back Propagations (BP) approaches were used by Tuncal et al. [9] to estimate lung cancer in data from eleven European countries dating back to 1970. The techniques predicted and forecasted illness occurrences based on patent histories, with SVRs outperforming others with more accurate projections, making them valuable for healthcare planning and regional resource distribution. predicted 192 observations, indicating their capacity to distinguish sick and non-cancerous cells.

In another study, the use of ML algorithms is recommended for discriminating benign and malignant lung tumors and predicting lung cancers. CNNs, SVMs, Random Forests (RFs), KNNs, Multi-Layer Perceptrons (MLPs), ANNs, and Entropy Degradation Method (EDM) were amongst the classification models evaluated on AC, RE, and SP in the study for lung cancer predictions. The findings revealed that on short datasets, CNNs achieved the highest AC of 96%, while EDM had the lowest AC of 77.8%, thus proving better efficiency of CNNs in detecting lung tumors when compared to other methods [10], [11].

Kumari et al. suggested lung cancer classifications and forecasts using bio-inspired methodologies such as ANNs, Genetic Algorithms (GA), Particle Swarm Optimizations (PSOs), and ant colony optimizations (ACOs). These solutions increase model accuracies by optimizing parameters, fine-tuning feature selections, and lowering dimensionalities. The algorithms were evaluated on many lung cancer datasets to improve predictions and help in early diagnostics. Comparative studies show that bio-inspired strategies improve classification AC and feature relevance, emphasizing their potential to increase lung cancer diagnosis through more efficient and reliable predictive modelling approaches [12].

3 | Proposed Methodology

In this paper, EBNNs approach is proposed to improve the performance of lung cancer predictions/classifications. The three main steps of the proposed method are preprocessing, feature extraction, and classification [13]. The general block diagram of the suggested system is shown in *Fig. 1*.

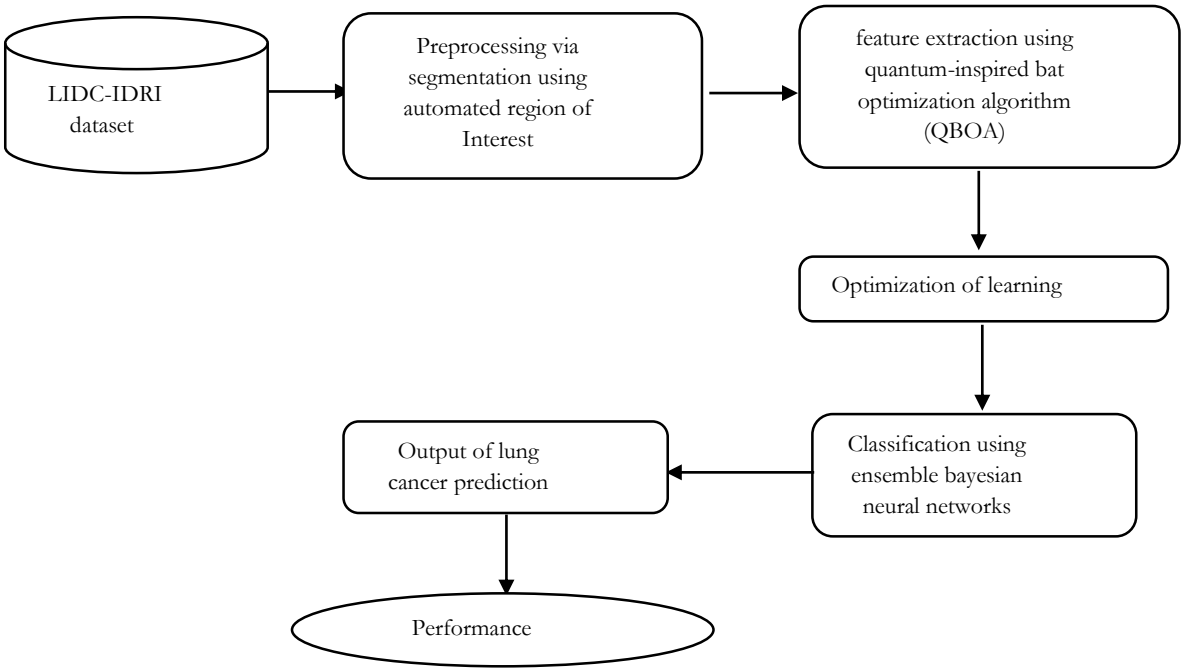

Fig. 1. Overall proposed flow diagram for lung cancer prediction.

3.1 | Input Dataset Collection

In this paper, the LIDC-IDRI dataset uses a robust lung CT image resource from the U.S. National Cancer Institute, containing over 1,000 chest CT scans, each annotated by four radiologists with detailed labels for lung nodule boundaries, malignancy, and type [14]. This dataset, ideal for advancing lung nodule detection, offers extensive labelling for training models. For the experiment, 6,002 images were utilized, dividing 4,733 into the training set and 1,269 into the test set, maintaining a strict 4:1 ratio and ensuring non-overlapping sets. To standardize data, pixel values in DICOM files were converted to Hounsfield Units (HU) using DICOM attributes for rescale slope (m) and Intercept (b), ensuring uniform data for reliable clinical interpretation is represented in *Fig. 2*.

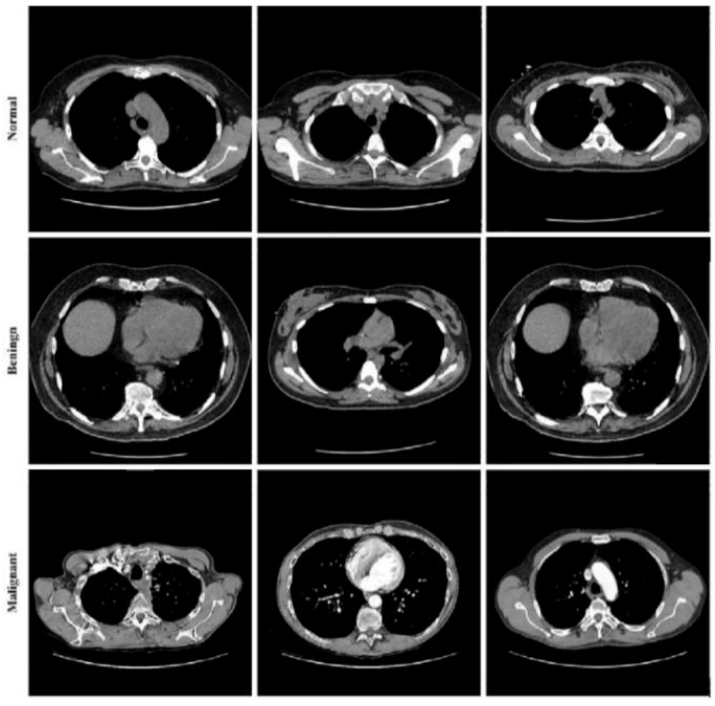

Fig. 2. Sample images from the LIDC-IDRI dataset.

3.2 | Preprocessing via Segmentation Using Automated Regions of Interest

In this paper, automated ROI segmentation for lung cancer prediction focuses on isolating lung nodules in CT scans, enhancing model AC and computational efficiency [15]. Techniques like as U-Net and Mask R-CNN are widely used to remove unnecessary backgrounds and extract exact boundaries from medical images, assisting in the correct diagnosis of cancer. Net encoders-decoders use convolution layers to gather features during the encoding phase and up samples to restore spatial information during the decoding phase. They are appropriate for fine-grained segmentations because they exclude connections while retaining a significant amount of spatial information. Using annotated CT scans with tumor borders in pixels, U-Nets were trained to detect lung malignancies. This enables networks to concentrate on possibly cancerous areas, improving prediction AC. Mask R-CNNs [16], which are enhanced variants of Faster R-CNNs, use mask branches for pixel-level segmentations. It can accurately identify many nodules with precise boundaries, even when tumor shapes vary. This application creates bounding boxes and applies segmentation masks to nodules to enhance data visibility and interpretability for radiologists.

Mask R-CNN's segmentations use composite loss functions that combine losses of mask accuracies, bounding box regressions, and classifications based on the following equation:

$$\text{Total Loss} = L_{cls} + L_{box} + L_{mask}.$$

The approach for both U-Net and Mask R-CNN [17] comprises importing annotated CT images, preprocessing, mask generation, and post-processing to reduce artifacts, allowing models to focus only on key lung regions. Automated ROI segmentation decreases False Positives (FP), increases diagnostic AC, and provides possible advancements in clinical lung cancer detection (see *Fig. 3*).

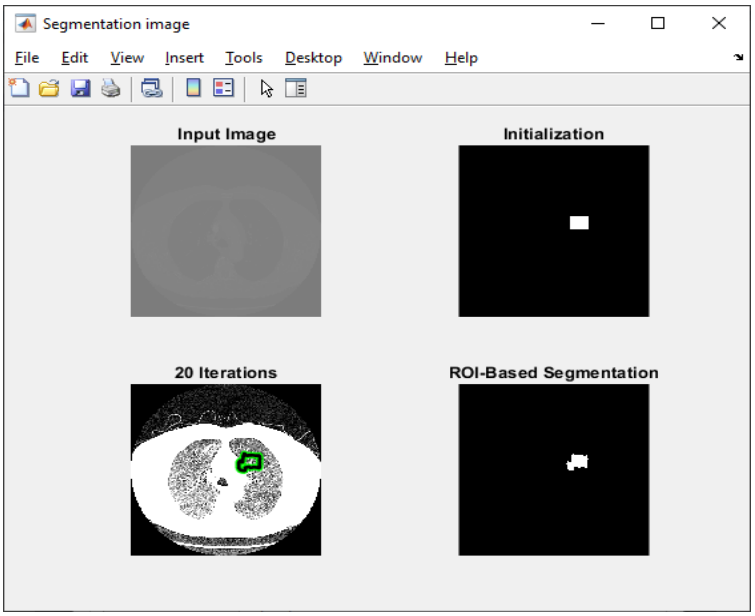

Fig. 3. Segmentation process using automated ROI.

3.3 | Feature Extraction using Quantum inspired Bat Optimization Algorithm

The QBOA combines quantum computing ideas with the Bat Algorithm to improve feature extraction, particularly for high-dimensional data [18]. QBOA enhances the search process by using notions like as superposition and entanglement, allowing for more accurate exploration of feature spaces and eliminating local optima difficulties. First, a population of bat agents is created, each of which represents a potential feature subset that is captured as binary vectors. The technique evaluates each subgroup's fitness, typically using classification AC as a performance metric *(Fig. 4)*.

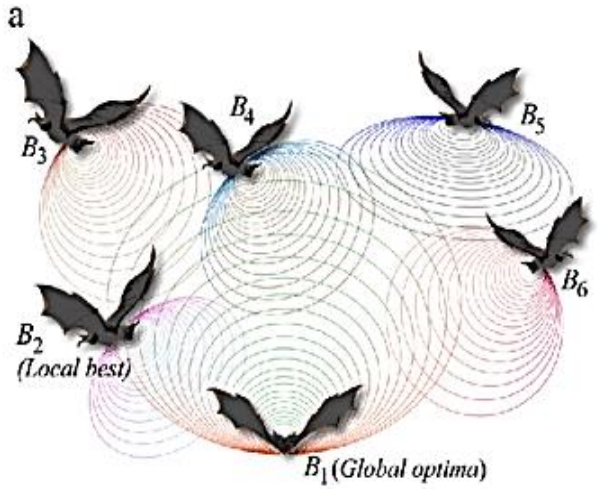

Fig. 4. The schematic diagram of the bat algorithm.

Fitness functions are generally dependent on classifying model accuracies of specified features and for classifiers like DTs may be defined as:

$$F_{accuracy} = \frac{\text{Correctly Classified Instances}}{\text{Total Instances}}.$$

3.3.1|Algorithm 1: Quantum inspired bat optimization algorithm

The process of this QBOA algorithm is as follows:

I. Initialize population: Initialize bat agents $B = \{b_1 + b_2 + b_3\}$ as binary feature subsets.

II. Evaluate fitness: Calculate fitness for each bat $F(b_i)$ based on classification AC using selected features.

III. Quantum-inspired bat movement: Update bat positions using quantum-inspired principles:

$$b_i(t+1) = b_i(t) + \alpha.\left(b_{best} - b_i(t)\right) + \beta.v_i(t).$$

IV. Update velocity: Update velocity using:

$$v_i(t+1) = \omega\,.v_i(t) + c_1.r_1.\left(b_{best} - b_i(t)\right) + c_2.r_2.\left(b_{global} - b_i(t)\right).$$

V. Generate new solutions: Refine solutions using quantum-inspired entanglements and explorations.

VI. Use $F(b_i)$. For evaluating fitness of new solutions.

VII. Acceptance criteria: If new solutions have higher fitness, update current solutions.

VIII. Terminate algorithm: Converge or achieve stopping conditions for selecting best feature subsets F_{opt} .

IX. Return optimal feature subsets: Return F_{opt}, feature subsets with high classification accuracies.

In [19], usages of Quantum-inspired methods like super positions and entanglements enhanced explorations and exploitations while identifying relevant feature subsets. The method's site switches and integrations of local best responses with global studies until optimal findings were achieved resulted in better feature selections.

3.3.2|Classifications Using Ensemble Bayesian Neural Networks

EBNNs used in this study [20] improved classification accuracies by integrating ensemble learning and Bayesian inferences. EBNNs train several Bayesian Neural Networks (BNNs) independently, using weights

as distribution probabilities in place of fixed values where this probabilistic approach cause prediction uncertainties for BNNs.

Hence, EBNNs are especially useful in high-risk applications like medical diagnostics, where trust on models is important. EBNNs generate singular outputs by combining and occasionally averaging predictions of multiple BNNs. The ensemble strategy prevents data over fits and enhances models' robustness by increasing variances with last categories having greatest overall likelihoodness. EBNNs enhance decision-making reliability and give insights into predictions by examining output variations. The combination of ensemble techniques with Bayesian inferences in EBNNs is effective for medical diagnosis and prognosis when reliable, intelligible results are required. EBNNs enhance classification problems by combining the advantages of ensemble learning with Bayesian techniques, notably in industries such as medical diagnosis and treatment where uncertainty is high. Combining ensemble learning with BNNs improves prediction AC [21], which is very relevant in medical imaging applications. By treating network weights as distributions rather than fixed values, BNNs can account for projected uncertainty and unpredictability when making critical decisions. Algorithmic EBNNs are described below.

I. Basic BNNs: in BNNs with inputs X and outputs Y, likelihood $p(Y \mid X,W)$ measures probabilities of Y for given X and network weights. The purpose is to compute posterior distributions of weights.

$$p(W|X,Y) = \frac{p(Y|X,W)p(W)}{p(Y|X)}.$$

Because finding the precise posterior may be computationally intensive, approximation techniques such as Markov Chain Monte Carlo and variational inference are widely utilized. In this situation, the marginal probability is $Y \mid X$ (), and the weights prior is (W).

II. Ensemble of BNNs: in EBNN, several BNNs are trained using different initializations or training subsets to build an ensemble. Each BNN m in the ensemble makes a prediction $f_m(x, W_m)$ given input x and weight distribution W_m. The ensemble forecast $\hat{y}$ represents the average of these predictions:

$$\hat{y} = \frac{1}{M}\sum_{m=1}^{M} f_m\,(x, W_m),$$

where M are counts of BNNs in ensembles and these aggregations reduce data over fits and improve robustness.

III. Uncertainty estimation: EBNN also quantifies prediction uncertainty, important for applications like treatment prognosis. The variance across ensemble predictions captures this uncertainty:

$$\text{Uncertainty} = \frac{1}{M}\sum_{m=1}^{M} (f_m(x,\ W_m) - \hat{y})^2.$$

When model confidence is crucial, the outcome is a reliable forecast with uncertainty measures that helps with informed decision-making. The above-maintained processed are shown as following (*Fig. 5*):

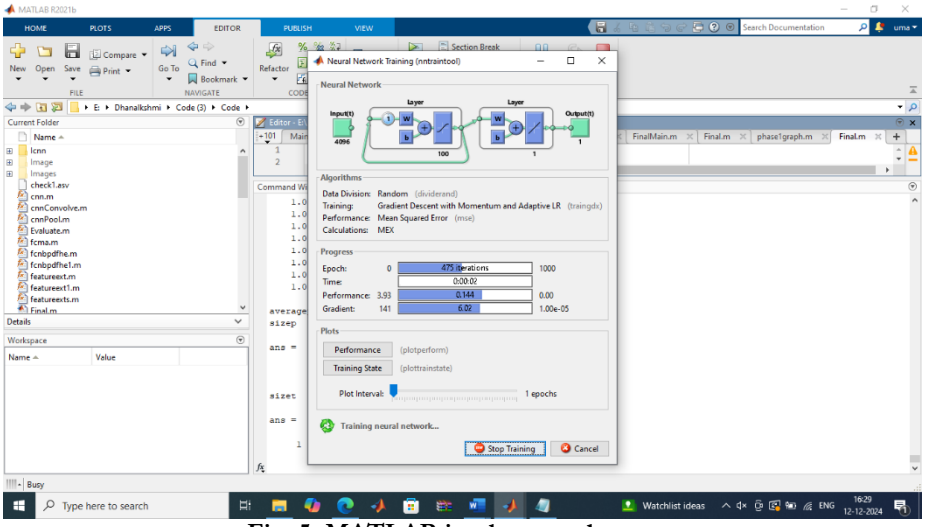

Fig. 5. MATLAB implemented screens.

3.3.3 | Algorithm 2: Ensemble bayesian neural networks algorithm

The following steps indicate the process of EBNN Algorithm:

I. Initialize data: Fill in X with features and labels Y

II. Initialize the BNNs: Set up M BNNs with weight-priors.

III. Train BNNs: Estimate p(W/X, Y) for each BNN using the posterior distribution.

IV. Make predictions: For fresh data, predict f_m (x,W_m) from each BNN.

V. Aggregate predictions: The average forecast throughout the ensemble.

VI. Estimating uncertainty: Calculate the variance of the ensemble forecasts.

VII. Classify: Assign classes based on the highest likelihood.

Because of its ensemble approach, which yields both precise predictions and quantified uncertainty, the EBNN [22] model performs exceptionally well when handling complex data. EBNN increases the robustness and confidence of medical diagnostic applications by calculating variance and averaging predictions from BNNs. Because of its probabilistic, ensemble approach, EBNN is very useful for applications like disease detection and classification that require reliable, comprehensible findings.

4 | Experimental Result

This section discusses performance evaluations of the model using both current and planned methods [23]. The suggested system was built using MATLAB on an Intel i3 CPU with a 3.0 GHz processor, 8 GB of RAM, and 1TB of hard drive storage. On the available datasets, the system's simulation efficiency was contrasted with that of more modern conventional systems and other well-known optimization methods, like QBOA. These optimization methods were chosen because they may be used to solve complex problems, such medical image classification, where model performance is highly dependent on parameter adjustments. The results of these comparisons are presented to show how the recommended method increases processing efficiency and AC. The findings aim to demonstrate the superiority of the new system in handling challenges like class imbalance, feature extraction, and generalisation over existing optimization-based models. These datasets [24] serve as a crucial benchmark for developing advanced lung cancer prediction models. The results of the performance comparisons between our proposed hybrid model and existing methods are summarised in *Table 1*[1].

[1] From the following link, the LIDC-IDRI Challenge datasets are collected: https://www.kaggle.com/datasets/washingtongold/lidcidri30

Table 1. Results of performance comparison.

Metrics	Methods				
	CNN	**PSO-CNN**	**ACO-CNN**	**BAT-CNN**	**QBOA-CNN**
SP	83.75	88.5	99.75	99.5	99.85
Precision	77.50	88.5	92	94	94.5
RE	64.38	79.88	93.64	99.38	99.9
F-Score	69.71	82.20	96.67	97.53	98
G-Mean	74.69	82.99	96.41	97.39	98.1
AC	67.38	84.19	96.5	97..44	98.25

4.1 | Recall

RE (also known as True Positive (TP) rate) quantifies the model's capacity to accurately detect every instance of lung cancer in the context of lung cancer prediction [25]. This is essential to preventing the omission of actual instances of lung cancer, which is particularly critical in a disease like cancer where early identification greatly enhances prognoses.

$$\mathrm{SE} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}},$$

where:

I. TP: Correctly identified cases of lung cancer.

II. False Negative (FN): Cases of lung cancer that were incorrectly identified as non-cancerous.

In *Fig. 6*, the RE performance of various models for lung cancer prediction is analyzed. The QBOA-CNN model demonstrates the highest RE at 99.8%, highlighting its superior predictive AC. BAT-CNN and ACO-CNN models closely follow with sensitivities of 99.38%, respectively. PSO-CNN shows moderate performance with 79.88% RE, while the standard CNN model achieves the lowest RE at 75%. These results underline the effectiveness of hybrid optimization-based models, such as QBOA-CNN, in accurately identifying TP cases of lung cancer, outperforming traditional CNN approaches and other optimization-based algorithms.

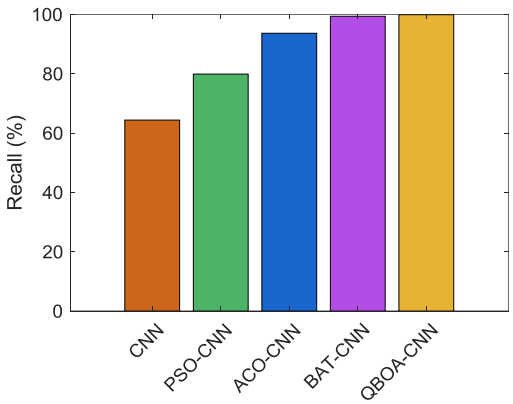

Fig. 6. Recall.

4.2 | Specificity

SP in the context of lung cancer prediction [26] refers to the model's capacity to accurately detect true negatives, or situations in which lung cancer is not present. This is particularly helpful in lowering FP, which can put patients through needless stress and result in more unneeded diagnostic testing.

$$\mathrm{SP} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}},$$

where:

I. TN: Correctly identified non-cancerous cases.

II. FP: Non-cancerous cases incorrectly identified as lung cancer.

The SP performance of many models for predicting lung cancer is examined in *Fig. 7*. With the highest SP of 99.85%, the QBOA-CNN model demonstrates its exceptional capacity to accurately detect negative instances. Following closely behind with specificities of 99.75% and 99.5%, respectively, are the ACO-CNN and BAT-CNN models. The basic CNN model gets the lowest SP at 83.75%, while the PSO-CNN model reaches a moderate SP of 88.5%. These findings demonstrate how well hybrid optimization-based models, such ACO-CNN and QBOA-CNN, reduce FP.

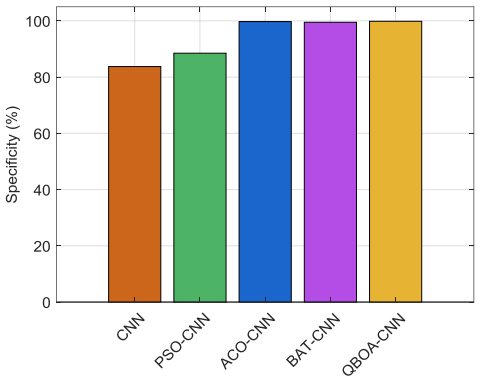

Fig. 7. Specificity.

4.3 | Precision (P)

In the case of lung cancer prediction, precision quantifies the AC of the positive predictions [27]. By determining the percentage of cases that are really diagnosed as lung cancer, this statistic aids in evaluating the model's confidence in its positive predictions.'

$$P = \frac{TP}{TP + FP}.$$

The AC findings for several models show that it is possible to reduce erroneous positive predictions (*Fig. 8*). With 94.5% precision, the QBOA-CNN model outperforms the BAT-CNN model, which comes in second with 94% and the ACO-CNN model with 92%. With a precision of 88.5%, the PSO-CNN model performs moderately, whereas the normal CNN model has the lowest precision (77.5%). These results underline the effectiveness of optimization-based CNN methods, particularly QBOA-CNN, in improving predictive reliability.

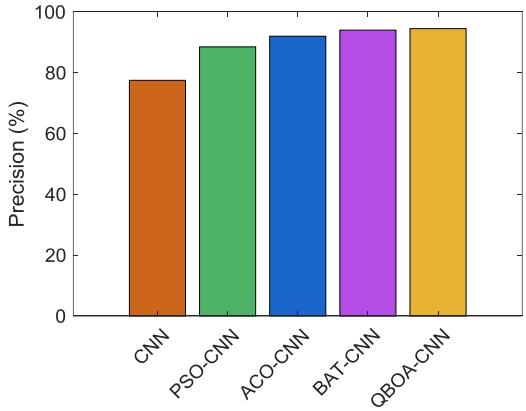

Fig. 8. Precision.

4.4 | Recall

Regarding the prediction of lung cancer [28], RE denotes the model's capacity to accurately detect every instance of TP lung cancer. For the majority of lung cancer cases, if not all of them, to be identified, high RE is crucial.

$$\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}.$$

In *Fig. 9*, the RE performance of various models for lung cancer prediction is analyzed. The QBOA-CNN model demonstrates the highest RE at 99.9%, highlighting its superior predictive AC in identifying TP cases. The BAT-CNN model follows closely with 99.38%, while the ACO-CNN model achieves 93.64% RE. The basic CNN model obtains the lowest RE at 64.38%, while the PSO-CNN model performs moderately at 79.88%. According to these results, hybrid models, like QBOA-CNN and BAT-CNN, perform better than traditional CNN models.

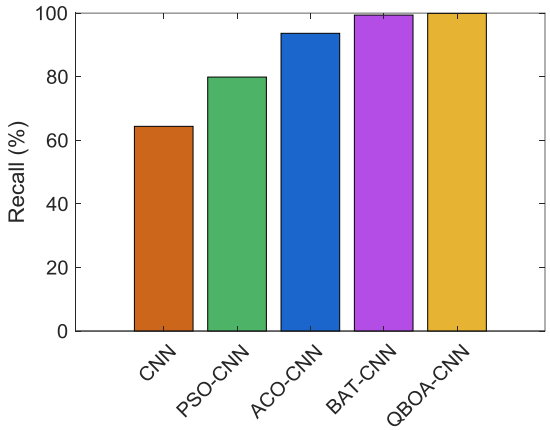

Fig. 9. Recall.

4.5 | F-Score (FS)

F-Measure (sometimes called the F1-score) combines precision and RE to balance the trade-off between recognizing all positive cases and reducing FP in the context of lung cancer prediction [29]. In the context of lung cancer prediction, the F-Measure helps evaluate a model's ability to accurately detect cancer cases while lowering incorrect cancer predictions.

$$\mathrm{FM} = 2 \times \frac{\mathrm{P} \times \mathrm{R}}{\mathrm{P} + \mathrm{R}}.$$

Fig. 10, F-Score comparisons show balancing of AC and RE values where QBOA-CNN model performs best with an F-Score of 98%, followed by BAT-CNN (97.53%) and ACO-CNN (96.67%). The PSO-CNN model does rather well at 82.2%, whereas the simple CNN model gets the lowest F-Score at 69.71%. These findings demonstrate the potential for balanced and dependable prediction performance from hybrid models like QBOA-CNN. This illustrates their ability to use state-of-the-art computer technology to properly forecast lung cancer.

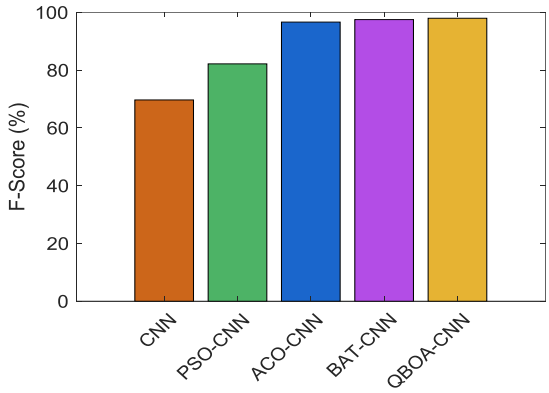

Fig. 10. F-Score.

4.6 | G Mean

G-Means (GMs) are particularly helpful when working with unbalanced datasets like comparing lung cancer cases with non-cancerous cases [30], they strike balances between REs and specificities in the context of lung cancer predictions. A high G-Mean ensures that the model detects lung cancer while accurately identifying non-cancer patients.

$$\mathrm{GM} = \sqrt{\mathrm{SE} \times \mathrm{SP}}.$$

Fig. 11 evaluates the balanced AC of different models by displaying their G-Mean performance. With its highest G-Mean of 98.1%, the QBOA-CNN model has an amazing ability to balance sensitivity and SP. The BAT-CNN and ACO-CNN models come next, with G-Means of 97.39% and 96.41%, respectively. The basic CNN model has the lowest G-Mean (74.69%), while the PSO-CNN model has a moderate G-Mean of 82.99%. These results demonstrate how hybrid optimization strategies, such as QBOA-CNN, improve prediction AC.

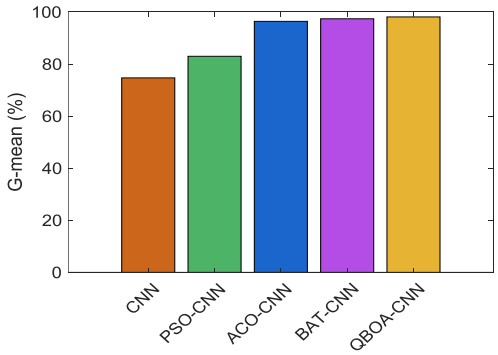

Fig. 11. G-Mean.

4.7 | Accuracy

AC, implies models' accurate identifications of both positive (patients with lung cancers) and negative (patients without lung cancers) instances are measured by crucial parameters in predictions [31]. High accuracies indicate that models can predict both classes, a necessity for proper diagnosis.

$$\mathrm{AC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}.$$

Fig. 12 shows the AC results for a variety of models. The QBOA-CNN model's exceptional prediction power is shown in its best AC of 98.25%. The AC of the ACO-CNN model is 96.5%, while the BAT-CNN model

comes in second at 97.44%. The PSO-CNN model has a high AC rate (84.19%), whereas the normal CNN model has the lowest (67.38%). These studies show how to increase prediction AC with optimization-based models like QBOA-CNN and BAT-CNN.

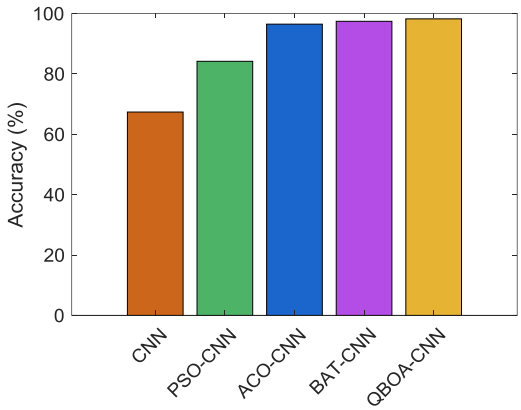

Fig. 12. Accuracy.

5 | Conclusion

This research presents a comprehensive system for lung cancer prediction and classification that combines advanced preprocessing techniques with a robust ensemble learning model. The method uses automated ROA segmentation to accurately detect tumor regions, resulting in much higher diagnosis AC. QBOA enhances feature selection by removing just the most critical information and reducing computational complexity. The EBNNs enhances classification and provides more reliable predictions by including uncertainty estimates, both of which are necessary for making smart medical decisions. The framework outperforms on the LIDC-IDRI dataset, with 98.25% AC, 99.8% RE, 99.85% SP, 94.5% precision, 99.9% RE, 98% F-Measure, and 98.1% G-Mean. These findings indicate how quickly and effectively the framework can detect and diagnose lung cancer. This unique technology offers a trustworthy and effective option for diagnosing early lung cancer, opening the path for improved clinical decision-making and patient outcomes. The method sets a standard for accurate and reliable lung cancer diagnosis in modern oncology by merging complex algorithms with ensemble techniques.

Authors' Contributions

The author solely conducted the research and prepared the manuscript and has approved its final version.

Data Availability

The data are available from the corresponding author upon reasonable request.

Funding

This work was carried out without financial support from any public, commercial, or non-profit organizations.

Conflict of Interest

There are no competing interests to declare.

Consent for Publication

The author confirms consent for the publication of this work

Ethics Approval and Consent to Participate

This article does not include experiments involving humans or animals.

References

[1] Ahmed, B. T. (2019). Lung cancer prediction and detection using image processing mechanisms: An overview. *Signal and image processing letters*, *1*(3), 20–31. http://dx.doi.org/10.31763/simple.v1i3.11

[2] Zhang, W., Zhuang, D., Wei, W., Li, Y., Ma, L., Du, H., ..., & & Li, X. (2025). Advancements and future trends in machine learning for lung cancer: A comprehensive bibliometric analysis. *Clinical and translational oncology*, *27*, 1–12. https://doi.org/10.1007/s12094-025-03945-7

[3] Radhika, P. R., Nair, R. A. S., & Veena, G. (2019). A comparative study of lung cancer detection using machine learning algorithms. *2019 IEEE international conference on electrical, computer and communication technologies (ICECCT)* (pp. 1–4). IEEE. https://doi.org/10.1109/ICECCT.2019.8869001

[4] Bharati, S., Podder, P., Mondal, R., Mahmood, A., & Raihan-Al-Masud, M. (2020). Comparative performance analysis of different classification algorithm for the purpose of prediction of lung cancer. *Intelligent systems design and applications* (pp. 447–457). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-16660-1_44%0A%0A

[5] Jassim, M. M., & Jaber, M. M. (2022). Systematic review for lung cancer detection and lung nodule classification: Taxonomy, challenges, and recommendation future works. *Journal of intelligent systems*, *31*(1), 944–964. https://doi.org/10.1515/jisys-2022-0062

[6] Kasinathan, G., & Jayakumar, S. (2022). Cloud-based lung tumor detection and stage classification using deep learning techniques. *Biomed research international*, *2022*(1), 4185835. https://doi.org/10.1155/2022/4185835

[7] Ma, X., Ning, F., Xu, X., Shan, J., Li, H., Tian, X., & Li, S. (2024). Survival prediction for non-small cell lung cancer based on multimodal fusion and deep learning. *IEEE access*, *12*, 123236–123249. https://doi.org/10.1109/ACCESS.2024.3453930

[8] Vishwa Kiran, S., Kaur, I., Thangaraj, K., Saveetha, V., Kingsy Grace, R., & Arulkumar, N. (2023). Machine learning with data science-enabled lung cancer diagnosis and classification using computed tomography images. *International journal of image and graphics*, *23*(3), 2240002. https://doi.org/10.1142/S0219467822400022

[9] Tuncal, K., Sekeroglu, B., & Ozkan, C. (2020). Lung cancer incidence prediction using machine learning algorithms. *Journal of advances in information technology*, *11*(2), 91–96. https://doi.org/10.12720/jait.11.2.91-96

[10] Mitra, S., Majumder, A. B., & Saha, T. (2023). An observation and analysis the role of convolutional neural network towards lung cancer prediction. *Baghdad science journal*, *20*(6), 2568–2592. https://dx.doi.org/10.21123/bsj.2023.9029

[11] Taher, F., Prakash, N., Shaffie, A., Soliman, A., & El-Baz, A. (2021). An overview of lung cancer classification algorithms and their performances. *IAENG international journal of computer science*, *48*(4), 1–7. https://www.iaeng.org/IJCS/issues_v48/issue_4/IJCS_48_4_19.pdf

[12] Kumari, J., Sinha, S., & Singh, L. (2025). Advancing lung cancer diagnosis with bio-inspired algorithms: A comprehensive assessment. *Multimedia tools and applications*, *84*(18), 18919–18941. https://doi.org/10.1007/s11042-024-19795-4

[13] Sakib, S. N. (2022). *Research proposal: Lung cancer prediction and classification using machine learning models, preprint*. https://doi.org/10.22541/au.165151786.62381525/v1

[14] Shao, Y., Wang, M., Mai, J., Fu, X., Li, M., Zheng, J., … & Ji, H. (2022). LIDP: A lung image dataset with pathological information for lung cancer screening. *Medical image computing and computer assisted intervention -- MICCAI 2022* (pp. 770–779). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-16437-8_74%0A%0A

[15] Suresh, S., & Mohan, S. (2020). ROI-based feature learning for efficient true positive prediction using convolutional neural network for lung cancer diagnosis. *Neural computing and applications*, *32*(20), 15989–16009. https://doi.org/10.1007/s00521-020-04787-w

[16] Aharonu, M., & Ramasamy, L. (2024). A multi-model deep learning framework and algorithms for survival rate prediction of lung cancer subtypes with region of interest using histopathology imagery. *IEEE access*, *12*, 155309–155329. https://doi.org/10.1109/ACCESS.2024.3484495

[17] Usman, M., Lee, B. D., Byon, S. S., Kim, S. H., Lee, B. I., & Shin, Y. G. (2020). Volumetric lung nodule segmentation using adaptive ROI with multi-view residual learning. *Scientific reports*, *10*(1), 12839. https://doi.org/10.1038/s41598-020-69817-y

[18] Priyadharshini, P., & Zoraida, B. S. E. (2021). Bat-inspired metaheuristic convolutional neural network algorithms for CAD-based lung cancer prediction. *Journal of applied science and engineering*, *24*(1), 65–71. http://jase.tku.edu.tw/articles/jase-202102-24-1-0008.pdf

[19] Tbaishat, D., Tubishat, M., Makhadmeh, S. N., & Alomari, O. A. (2025). A hybrid bat and grey wolf optimizer for gene selection in cancer classification. *Knowledge and information systems*, *67*(1), 455–495. https://doi.org/10.1007/s10115-024-02225-0

[20] Hussain Ali, Y., Chinnaperumal, S., Marappan, R., Raju, S. K., Sadiq, A. T., Farhan, A. K., & Srinivasan, P. (2023). Multi-layered non-local bayes model for lung cancer early diagnosis prediction with the internet of medical things. *Bioengineering*, *10*(2), 138. https://doi.org/10.3390/bioengineering10020138

[21] Admane, M., & Patil, S. (2022). Modeling lung cancer diagnosis using bayesian network inference. *2022 international conference on smart generation computing, communication and networking (smart Gencon)* (pp. 1–16). IEEE. http://ftp.mi.fu-berlin.de/bkeller/MatLab/Examples/bayes/lungbayesdemo.html

[22] Zhang, J., & Zhang, H. (2023). A bayesian neural network model based on ct images for staging non-small cell lung cancer. *2023 8th international conference on computer and communication systems (ICCCS)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICCCS57501.2023.10150484

[23] Ling, B., Liao, X., Huang, Y., Liang, L., Jiang, Y., Pang, Y., & Qi, G. (2020). Dentification of prognostic markers of lung cancer through bioinformatics analysis and in vitro experiments. *International journal of oncology*, *56*(1), 193–205. https://doi.org/10.3892/ijo.2019.4926

[24] Abdullah, D. M., Abdulazeez, A. M., & Sallow, A. B. (2021). Lung cancer prediction and classification based on correlation selection method using machine learning techniques. *Qubahan academic journal*, *1*(2), 141–149. https://doi.org/10.48161/qaj.v1n2a58

[25] Nageswaran, S., Arunkumar, G., Bisht, A. K., Mewada, S., Kumar, J. S., Jawarneh, M., & Asenso, E. (2022). Lung cancer classification and prediction using machine learning and image processing. *BioMed research international*, *2022*(1), 1755460. https://doi.org/10.1155/2022/1755460

[26] Berk, Ş., Kaya, S., Akkol, E. K., & Bardakçı, H. (2022). A comprehensive and current review on the role of flavonoids in lung cancer–Experimental and theoretical approaches. *Phytomedicine*, *98*, 153938. https://doi.org/10.1016/j.phymed.2022.153938

[27] Qureshi, R., Basit, S. A., Shamsi, J. A., Fan, X., Nawaz, M., Yan, H., & Alam, T. (2022). Machine learning based personalized drug response prediction for lung cancer patients. *Scientific reports*, *12*(1), 18935. https://doi.org/10.1038/s41598-022-23649-0

[28] Pavithran, M. S., Saranyaraj, D., & Chakrabortty, A. (2025). Lung cancer risk prediction using augmented machine learning pipelines with explainable AI. *Frontiers in artificial intelligence*, *8*, 1602775. https://doi.org/10.3389/frai.2025.1602775

[29] ALzubi, J. A., Bharathikannan, B., Tanwar, S., Manikandan, R., Khanna, A., & Thaventhiran, C. (2019). Boosted neural network ensemble classification for lung cancer disease diagnosis. *Applied soft computing*, *80*, 579–591. https://doi.org/10.1016/j.asoc.2019.04.031

[30] Pradhan, K., Chawla, P., & Rawat, S. (2023). A deep learning-based approach for detection of lung cancer using self adaptive sea lion optimization algorithm (SA-SLnO). *Journal of ambient intelligence and humanized computing*, *14*(9), 12933–12947. https://doi.org/10.1007/s12652-022-04118-y

[31] Deepapriya, B. S., Kumar, P., Nandakumar, G., Gnanavel, S., Padmanaban, R., Anbarasan, A. K., & Meena, K. (2023). Performance evaluation of deep learning techniques for lung cancer prediction. *Soft computing*, *27*(13), 9191–9198. https://doi.org/10.1007/s00500-023-08313-7