International Journal of Computer Science Engineering Techniques (IJCSE) Volume 10 Issue 3 ,May-June 2026

ISSN : 2455-135X

https://www.ijcsejournal.org/ ***Type of Article*** *(Review Article)*

Hybrid Bio-Inspired and Artificial Intelligence Framework for Small Cell and Non-Small Cell Lung Cancer Detection and Classification

Abirami T[1], W. Rose Varuna[2]

[1]*Department of Information Technology, Bharathiar University, Tamil Nadu, India.*
[2]*Department of Information Technology, Bharathiar University, Tamil Nadu, India.*

[1]gntabirami@gmail.com

Received: Revised: Accepted: Published: (Size9)

Abstract - *Lung cancer is a serious disease and causes many deaths around the world. Finding it early is very important because it helps in better treatment and improving patient survival rates. Doctors usually use methods like CT scans, biopsies, and manual checking, but these methods can be slow and may not always be accurate. They also depend a lot on the doctor's experience and may miss cancer in the early stage. Artificial Intelligence (AI) helps improve diagnosis by automatically analyzing complex medical data. Deep Learning models can extract patterns from imaging and genomic data, thereby improving detection and classification performance. Bio-inspired algorithms, which are based on natural processes, identify the most relevant features and help enhance the model's performance. Bio-inspired algorithms include Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO), and swarm behaviors to optimize feature selection. This research work presents a hybrid approach that combines AI-based models along with bio-inspired optimization techniques for lung cancer detection and classification. When AI and bio-inspired algorithms are used together, they yield better, faster results. This paper explains how these methods work together to detection and classification of lung cancer. The results show that combining these methods improves accuracy and helps doctors make better decisions.*

Keywords - *Non-Small Cell Lung Cancer, Small Cell Lung Cancer, Bio-Inspired Algorithms, Artificial Intelligence Algorithms, Medical Imaging, Classification.*

1. Introduction

Lung cancer is the leading cause of death around the world. It affects millions of people every year. Lung Cancer is mainly classified into two major categories: Small Cell Lung Cancer and Non-Small Cell Lung Cancer. The most common type is Non-Small Cell Lung Cancer (NSCLC), which is the most prevalent, accounting for approximately 85% of all lung cancer cases. Detecting this disease at an early stage is very important because it can help doctors start treatment sooner and improve survival rates.

Traditional diagnostic methods, such as CT scans, laboratory tests, and tissue analysis, are valuable but require expert interpretation, can be time-consuming, and occasionally yield inaccurate results. As medical data grows in volume and complexity—with detailed CT images and thousands of genetic features—manual analysis becomes increasingly challenging.

Artificial Intelligence (AI) helps solve these problems by automatically learning patterns from data. Deep Learning (DL) models can identify cancer regions in images, while Machine Learning (ML) algorithms can classify tumor types. ML algorithms can classify cancer types and predict disease progression with high accuracy. Bio-inspired algorithms further improve these models by selecting important features and optimizing performance. Techniques like Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO) are widely used to optimize features and model parameters. This paper explains how these techniques work together to improve lung cancer detection. The study findings confirm that combining AI and bio-inspired techniques in developing reliable and efficient healthcare solutions.

2. Literature Survey

Mohammed Qaraad et al (2025) This study introduces an

ISSN : 2455-135X

enhanced optimization algorithm, iPSOgs, which improves search efficiency and convergence stability for lung cancer classification. The model integrates iPSOgs with XGBoost for simultaneous gene selection and classification of NSCLC subtypes. It achieved high performance with 95.8% accuracy, 98.7% AUC, and strong precision and recall values. The method also identified important genes such as DSG3, KRT5, and SPRR2E, which are relevant to lung cancer. Overall, the approach provides an effective and interpretable solution for accurate diagnosis and biomarker discovery in lung cancer [1].

Avigyan Roy et al (2025) This work proposes a lung cancer classification model that combines deep learning with a feature selection (FS) algorithm. A DenseNet121-based CNN is used to extract features from histopathological images, and an adaptive Genetic Algorithm (GA) is applied to optimize the selected features, significantly reducing the feature space while preserving high classification performance. Although the results are effective, the study suggests further improvements through better feature extraction models and more efficient feature selection methods [3].

Karthika M S et al (2024) This research presents a lung cancer classification approach for high-dimensional microarray data using dimensionality reduction and feature selection techniques. Methods such as Detrended Fluctuation Analysis (DFA) and Elephant Herd Optimization (EHO) were applied to reduce and optimize features. Among various classifiers tested, the Linear SVM combined with DFA and EHO achieved a maximum accuracy of 92.26%, emphasizing the critical role of feature selection in improving classification performance [4].

Muhammad Asim Saleem et al (2023) This work focuses on proposing a deep learning model optimized using the Sooty Tern Optimization Algorithm (SHOA) for accurate detection of Non-small cell lung cancer (NSCLC). The method uses Otsu segmentation for nodule extraction, Local Binary Patterns (LBP) for feature extraction, and CNN-GRU classifiers for classification. The proposed framework achieved a high accuracy of 98.32%, surpassing existing methods and demonstrating the effectiveness of bio-inspired optimization in improving diagnostic performance [5].

Kunpeng Li et al (2024) This study determines a deep learning model called Lung Adenocarcinoma Convolutional Neural Network (LATCNN) for accurate prediction of lung adenocarcinoma. The model employs a hybrid feature selection strategy, which begins with Fast Correlation-Based Filter (FCBF) to remove irrelevant features, and the approach applies the k-means-SMOTE technique to address class imbalance. An enhanced Particle Swarm Optimization (PSO) algorithm with fast-decay dynamic inertia weights, combined with a Classification and Regression Tree (CART) as the fitness function, is then used to further eliminate redundant

features. For classification, an attention-based Convolutional Neural Network (atCNN) is developed, which integrates an attention mechanism to emphasize the most significant gene features. Experimental results achieving high performance, accuracy, recall, F1 score, and MCC values of 99.70%, 99.33%, 99.98%, and 98.67%, existing models showing strong potential for advancing lung adenocarcinoma diagnosis and treatment [9].

3. Proposed Methodology

The proposed PSO+XGBoost system focuses on enhancing lung cancer detection by integrating advanced Artificial Intelligence techniques with bio-inspired optimization algorithms. This hybrid approach combines deep learning for extraction features with nature-inspired methods for optimal feature selection, resulting in enhanced diagnostic performance and efficiency. The existing systems focus only on single-modal data and show reduced optimization efficiency, which affects early-stage lung cancer detection performance. The overall workflow is structured into multiple interconnected stages, beginning with Data Collection and moving through Preprocessing, Feature Extraction, Feature Selection, Classification, Multimodal Integration, and Output Generation.

3.1 Data Collection

The process begins with gathering from multiple medical sources to ensure diversity and robustness in the dataset. This includes Computed Tomography (CT) scan images, histopathological images, patient clinical records such as age, smoking history, and symptoms, as well as genetic or genomic data where available. The inclusion of multimodal data allows the system to capture both visual and biological characteristics of lung cancer, thereby improving the model's capability to generalize across diverse patient conditions and disease variations.

3.2 Data Preprocessing

The collected data is preprocessed and organized for analysis. Medical images often contain noise, artifacts, and variations in size or intensity, which can negatively affect model performance. Therefore, image preprocessing methods, including noise removal, resizing, normalization, and contrast enhancement, are employed to standardize the input. For non-image data, preprocessing includes managing missing values, encoding categorical variables, and normalization to ensure uniform scaling. This step is crucial for eliminating inconsistencies and preparing the dataset for reliable analysis.

ISSN : 2455-135X

3.3 Feature Extraction

Feature extraction is carried out using deep learning techniques, especially Convolutional Neural Networks (CNNs). CNNs automatically learn hierarchical feature representations from images, and identifying critical patterns involves the shape, size, texture, and boundary characteristics of lung nodules. Unlike traditional methods, this approach removes the dependency on manual feature engineering and enables the system to capture complex, non-linear relationships within the data. These extracted features form a high-dimensional representation of the input data.

3.4 Feature Selection (Bio-Inspired Optimization)

Feature Selection: not all extracted features contribute equally to classification performance. To address this, bio-inspired optimization algorithms are utilized for feature selection. Methods such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO) are used to identify the most relevant subset of features. These algorithms imitate natural processes, including evolution, swarm intelligence, and collective behavior, to search for optimal solutions. By removing redundant along with irrelevant features, this step reduces dimensionality, decreases computational complexity, and enhances model accuracy.

3.5 Classification

The optimized feature set is then used in the classification stage, where machine learning and deep learning classifiers are applied to detect and categorize lung cancer. Models such as CNN-based classifiers, Support Vector Machines (SVM), and XGBoost are utilized depending on the data type and problem complexity. The system is capable of distinguishing between cancerous and non-cancerous cases, as well as identifying specific types of lung cancer. This classification process is designed to be highly accurate and reliable, supporting early diagnosis.

3.6 Multimodal Integration

Multimodal data integration, where different types of data are combined to provide a more comprehensive analysis. For instance, imaging data can be fused with clinical records or genetic information to improve prediction performance. This holistic approach enables the system to consider multiple perspectives of the disease, leading to more informed and precise outcomes.

3.7 Output Generation

The system produces an output that includes the detection result, classification of cancer type, and a confidence score indicating the reliability of the prediction. These puts are presented in a user-friendly format to assist medical professionals in decision-making. By providing both diagnostic insights and confidence levels, the system acts as a supportive tool for clinicians, potentially improving early detection rates and patient outcomes. Figure 1 shows the Workflow diagram of the Proposed Methodology.

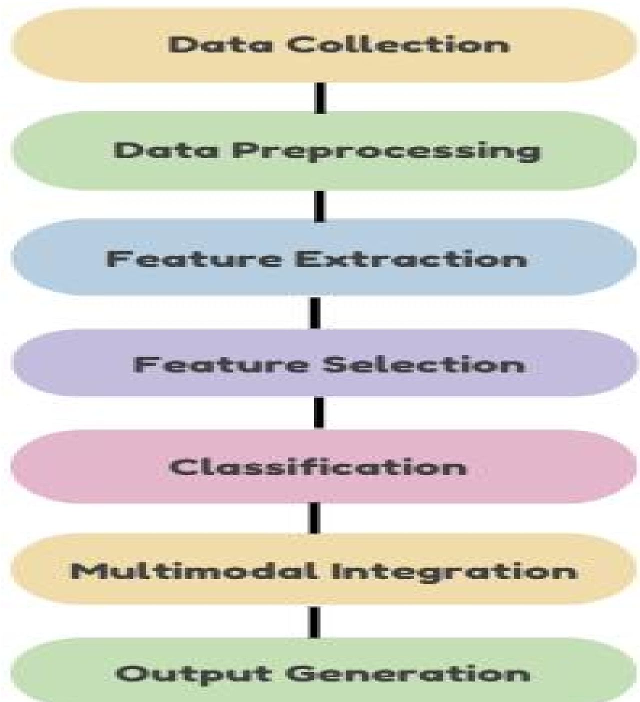


Fig.1 represents the Workflow of the proposed methodology

4. Results and Discussion

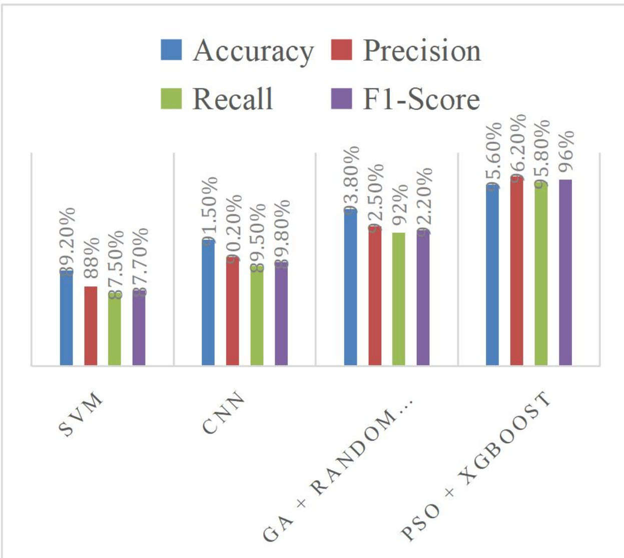
The proposed hybrid framework integrating Artificial Intelligence (AI) and bio-inspired optimization algorithms was evaluated on a multimodal dataset comprising CT scan images, histopathological data, and clinical attributes. The model significantly improves lung cancer detection and classification performance compared with traditional approaches. Among the PSO techniques performed the most efficient search capability and faster convergence are achieved. Overall, the study confirms that combining AI with bio-inspired techniques improves diagnostic accuracy and supports early detection of lung cancer. Table 1 shows the different algorithms employed for lung cancer detection. Figure 4.1 graph represents the classification of lung cancer using various algorithms.

Table 1 Comparison of Lung Cancer in different algorithms

ISSN : 2455-135X

Algorithms	Accuracy	Precision	Recall	F1-Score
SVM	89.2%	88%	87.5%	87.7%
CNN	91.5%	90.2%	89.5%	89.8%
GA + Random Forest	93.8%	92.5%	92%	92.2%
PSO + XGBoost	95.6%	96.2%	95.8%	96%

Fig 4.1



Diagnosis of Various Lung Cancer

5. Conclusion

This study proposed an approach that used deep learning techniques for feature extraction and bio-inspired methods, involving Genetic Algorithm (GA), Ant Colony Optimization (ACO), and Particle Swarm Optimization (PSO), for optimal feature selection. The findings showed that the hybrid model significantly outperformed traditional methods, achieving higher accuracy and improved reliability. The use of multimodal data further enhanced early-stage detection and classification performance by combining imaging, clinical, and histopathological information. Overall, the study confirms that integrating AI and bio-inspired algorithms can serve as a powerful decision-support system, enhancing diagnostic accuracy and supporting the early detection of lung cancer.

6. Future Work

Future research can concentrate on improving the model by utilizing larger and more diverse datasets to enhance robustness and generalization. In addition, advanced deep learning techniques, including transformer-based methods, can be explored to improve feature learning. Integrating real-time clinical data and applying explainable AI methods will improve practical usability and trust. Additionally, combining multiple bio-inspired algorithms and developing lightweight techniques can further improve performance and support deployment in real-world healthcare settings.

References

[1] Qaraad, Mohammed, et al. "An optimized computational framework for non-small cell lung cancer subtype classification and biomarker discovery." Artificial Intelligence Review (2025).
[2] Tortora, Matteo, et al. "Deep reinforcement learning for fractionated radiotherapy in non-small cell lung carcinoma." Artificial Intelligence in Medicine 119 (2021): 102137.
[3] Roy, Avigyan, et al. "Adaptive genetic algorithm based deep feature selector for cancer detection in lung histopathological images: A. Roy et al." Scientific Reports 15.1 (2025): 4803.
[4] Karthika, M. S., Harikumar Rajaguru, and Ajin R. Nair. "Performance enhancement of classifiers through Bio inspired feature selection methods for early detection of lung cancer from microarray genes." Heliyon 10.16 (2024).
[5] Saleem, Muhammad Asim, et al. "Sooty tern optimization algorithm-based deep learning model for diagnosing NSCLC tumours." Sensors 23.4 (2023): 2147.
[6] Rao, Kokkila Gadda Koteswara, and Lalitha Kumari Pappala. "Early-Stage Multimodal Lung Cancer Detection Using Attention-Based Feature Fusion." IEEE Access (2026).
[7] Sandhiya, R., and R. Sasikumar. "Swallow Winged Kite Optimization with Shuffle Attention Xtreme Gradient Boost Network for lung cancer detection using CT image." European Spine Journal (2025): 1-28.
[8] Mohamed, Tehnan IA, et al. "A novel feature selection algorithm for identifying hub genes in lung cancer." Scientific Reports 13.1 (2023): 21671.
[9] Li, Kunpeng, et al. "Lung adenocarcinoma identification based on hybrid feature selections and attentional convolutional neural networks." Mathem. Biosci. Eng 21 (2024): 2991-3015.