SIAR-Global Journal of Computer, Information & Library Science
www.siarpublications.org
info@siarpublications.org Vol. 1 Issue 1 Sept.-Oct. 2025
Global Office **South Africa:** 16 Park Ave, Modderfontein, Johannesburg

A Modified Multi-Agent Model for Diagnosis of Lung Cancer Using Soft Computing Techniques

[1]Samuel B. Zumbuka, [2]Prof. E. J. Garba, [3]Dr Murtala Muhammad
[1]Department of Computer Science, Faculty of Science and Computing, Adamawa State University, Mubi – Nigeria **Corresponding Author**; *e-mail*: zsam04@gmail.com, zumbuka450@adsu.edu.ng; GSM: 07034209101

[2&3]Department of Computer Science, Faculty of Computing, Modibbo Adama University, Yola – Nigeria *e-mail:* e.j.garba@mau.edu.ng *[3]E-mail*: alferio1349@gmail.com

Abstract
Lung cancer remains the leading cause of cancer-related mortality worldwide, necessitating accurate and efficient diagnostic tools to improve patient outcomes. Existing computer-aided diagnostic models often rely on rigid mathematical approaches, which struggle to handle the complexity and uncertainty associated with minute pulmonary nodules and accurate severity-level determination in clinical images. Furthermore, common weaknesses in existing models include low explainability, high computational overhead, and poor scalability. This research aimed to address these limitations by enhancing an existing multi-agent model based on soft computing techniques (Fuzzy Logic, Artificial Neural Networks, and Genetic Algorithms) to diagnose lung cancer, determine its severity, and improve diagnostic accuracy and computational efficiency. The proposed system utilizes a six-agent workflow for image preprocessing, feature fusion, and hybrid classification of CT scans from the LIDC-IDRI dataset. The hybrid model, which integrates Convolutional Neural Network (CNN) feature extraction with a Genetic Algorithm (GA)-optimized Fuzzy Inference System (FIS), achieved a superior Area Under the Curve (AUC) score of 0.91. This performance significantly outperformed the standalone CNN-only model (AUC = 0.83) and the Fuzzy-only system (AUC = 0.76). The results validate the synergistic advantage of the hybrid framework in delivering enhanced diagnostic precision, demonstrating its potential utility as a robust clinical decision-support tool.

Keywords: Lung Cancer, Soft Computing, Multi-Agent Model, Hybrid Model, Fuzzy Logic, Genetic Algorithm, Diagnosis

1.0 Lung cancer is one of the most widespread and deadly forms of cancer, continuing to pose serious challenges to global health. It arises from the abnormal and uncontrolled growth of cells in the lungs, often leading to severe impairment of respiratory function and, in many cases, death. Early detection is critical for improving survival rates, and medical imaging, particularly Computed Tomography (CT) scans, remains one of the most effective tools for identifying the disease at its onset. Recent global estimates indicate that lung cancer is among the most frequently diagnosed cancers worldwide, accounting for nearly 2.5 million new cases each year, or 12.4% of all cancers. Tragically, it is also the leading cause of cancer-related deaths, responsible for approximately 1.8 million fatalities annually, which represents 18.7% of all cancer deaths (Singh et al., 2020).

Despite advances in treatment, lung cancer continues to be regarded as one of the deadliest cancers, alongside breast, colon, and prostate cancers Its reported incidence is increasing in many developing nations, driven by factors such as rapid urbanization, longer life expectancy, and the adoption of Western lifestyles. Unhealthy dietary patterns, including diets high in red and processed meats, have been linked to increased risks of colorectal cancer, cardiovascular disease, and type 2 diabetes, contributing to the growing cancer burden in these regions. Effective control of lung cancer, therefore, requires both early detection and improved survival strategies.

Lifestyle choices play a major role in cancer risk. Tobacco use, poor diet, excessive alcohol consumption, physical inactivity, and obesity are estimated to contribute to a substantial proportion of cancer cases. Conversely, adopting healthier habits, such as regular exercise, maintaining a healthy weight, eating plant-rich diets, avoiding tobacco and excessive alcohol, and protecting against harmful exposures can substantially reduce cancer risk. Globally, cancer is the second leading cause of death, and within this category, lung cancer is the second most commonly diagnosed but remains the leading cause of cancer mortality. Early detection is therefore essential to improving patient outcomes. Medical imaging, with its ability to capture detailed lung structures, is particularly well-suited for analysis using deep learning techniques. In recent years, computer-assisted frameworks have been developed to support early-stage detection of lung cancer, offering promising avenues for research and clinical application (Liz-López et al., 2025).

1.3 Objectives of the Study

This research aims to modify a multi-agent model based on soft computing techniques to aid experts and medical doctors in the accurate diagnosis of lung cancer and the determination of its hyper-severity, while optimising for accuracy, sensitivity, and computational time and space. The specific objectives are to:

i. Collect, preprocess, clean, and create a dataset for training and testing.
ii. Modify existing diagnostic models to improve performance.
iii. Train the model using CT scan images of lung cancer and standard data.
iv. Implement the model by developing a prototype application.
v. Evaluate the performance of the modified multi-agent model.
vi. Benchmark the modified multi-agent model with other similar model

2.0 Related works:

A comprehensive review of the literature for the topic *"A Modified Multi-Agent Model for Diagnosis of Lung Cancer Using Soft Computing Techniques"* must draw from three interconnected fields. The first is Clinical Medicine, which provides the context of lung cancer diagnosis as a pressing global health challenge. The second is Computer Science, particularly the design and application of Multi-Agent Systems (MAS), which enable distributed problem-solving and collaborative decision-making. The third is Computational Intelligence, with a focus on Soft Computing (SC) techniques such as Fuzzy Logic, Neural Networks, and Genetic Algorithms, which are well-suited for handling uncertainty and complexity in medical data.

2.1 Diagnosis of lung cancer using Fuzzy Logic Techniques

Similarly, Phadke *et al.* (2025), in their article *"Predictive Modelling of Lung Infections using Fuzzy Logic Systems,"* introduced an intelligent prediction model that leveraged fuzzy logic to overcome limitations in conventional diagnostic approaches. Fuzzy logic is particularly valuable because it can

handle imprecise and ambiguous information, mimicking the reasoning style of medical professionals. Their system integrated diverse clinical parameters, including symptoms, vital signs, and patient-specific data to evaluate the likelihood of lung infection. By interpreting overlapping and non-binary inputs, the model produced reliable risk assessments. Using a curated dataset of anonymised patient records, the fuzzy logic system was trained and tested, demonstrating strong performance in detecting potential lung infections at early stages. The results highlight its promise as a supplementary tool for clinicians, offering enhanced diagnostic support and improved patient outcomes.

2.2 Diagnosis of lung cancer using ANFIS Algorithms

Shabu *et al*. (2024) introduced an Improved Adaptive Neuro-Fuzzy Inference Framework (ANF-IF) for lung cancer detection and prediction, designed to operate on the Internet of Medical Things (IoMT) platform. The central aim of their work was to enable early diagnosis of lung cancer, thereby improving patients' chances of receiving curative treatment. They argued that cutting-edge detection and monitoring technologies are essential for rapid, accurate, and timely diagnosis. Fuzzy Logic (FL), known for its ability to model complex and uncertain systems, was employed to address these challenges. Their study proposed a Fuzzy Expert System for Lung Cancer (FES-LC) integrated with IoMT, which categorized patients into four risk levels: not at risk, slightly at risk, moderately at risk, and severely at risk.

2.3 Diagnosis of lung cancer using Deep Learning with Neuro-Fuzzy techniques

Wankhade and Vigneshwari (2023) proposed a novel hybrid deep learning method for the early detection of lung cancer using neural networks. They asserted that lung cancer remains the leading cause of cancer-related deaths worldwide, underscoring the critical importance of early diagnosis in improving patient outcomes. Their methodological review explored the application of deep learning models across various imaging modalities, highlighting their effectiveness in nodule detection, classification, and prognosis prediction. Deep learning consistently demonstrated state-of-the-art performance, in some cases even surpassing human expert accuracy.

2.4 Diagnosis of lung cancer using Multi-Scale-Multi-Instance CNN

Thangamani *et al.* (2024), in their study *"Lung Cancer Diagnosis Based on Weighted Convolutional Neural Network Using Gene Data Expression,"* analyzed lung cancer as a genetic disease with diverse and often unknown origins. They referenced the GLOBOCAN 2020 report, which documented 19.3 million new cancer cases worldwide and nearly 10 million cancer-related deaths in that year alone. Projections from GLOBOCAN suggest that the global cancer burden will rise sharply, reaching an estimated 28.4 million new cases by 2040.

2.5 Diagnosis of lung cancer using Hybrid models

Agrawal and Agrawal (2025), in their work *"Genetically Optimized Modular Neural Networks for Precision Lung Cancer Diagnosis,"* applied advanced neural network techniques to CT scan data. Their study involved 156 patient CT scans, divided into two databases, from which features were extracted using image statistics, histograms, and 2D transforms such as FFT, DCT, and WHT. These optimal feature vectors were organized into structured knowledge bases for training. The researchers then experimented with genetically optimized classifiers, including Multilayer Perceptrons (MLP), Generalized Feedforward Neural Networks (GFF-NN), Modular Neural Networks (MNN), and SVMs, testing different parameter combinations, activation functions, and data partitioning

strategies. Performance was evaluated using metrics such as classification accuracy, Mean Squared Error (MSE), Area Under the ROC Curve (AUC), and computational efficiency.

2.6 Diagnosis of lung cancer using Deep Learning with Fuzzy Logic

Abe and Nyathi (2025), in their study *"Lung Cancer Diagnosis from Computed Tomography Images Using Deep Learning Algorithms with Random Pixel Swap Data Augmentation,"* introduced Random Pixel Swap (RPS), a novel data augmentation technique designed to improve diagnostic performance in both convolutional neural networks (CNNs) and transformer models. RPS works by randomly swapping pixels within CT scan images, thereby generating augmented data that enhances model robustness. In a related study, Abe *et al*. (2025) proposed a low-cost CNN architecture with Mavage pooling to enhance early lung cancer diagnosis. Their model incorporated five convolutional layers, two residual connections, and Mavage pooling layers, and was trained on the same IQ-OTH/NCCD and chest CT scan datasets. Saha *et al*. (2025) introduced *"Lung-AttNet,"* an attention mechanism-based CNN architecture with federated learning for lung cancer detection. Lung-AttNet integrates a convolutional block with a Lightweight Global Attention Module (LGAM), enabling the model to capture both low- and high-dimensional features while identifying dependencies across spatial and channel dimensions.

2.7 Diagnosis of lung cancer using Machine Learning with CNN

Kutaphale (2023) developed a lung cancer detection framework that integrates both machine learning and deep learning to improve the accuracy of CT scan interpretation and address the challenges posed by the complexity of medical imaging. The study utilized a large dataset of CT images collected from diverse patient groups in central Iraq. Two techniques were applied: the initial stage emphasized the cooperative interaction of multiple machines learning algorithms, underscoring the importance of careful data calibration. Building on this foundation, a more advanced approach was introduced using a Convolutional Neural Network (CNN) originally designed for retinal image classification.

.3.0 Methodology

This research focuses on the design of a Modified Multi-Agent Model that applies Soft Computing techniques to the diagnosis of lung cancer. The model is specifically aimed at determining malignancy and benign concentration percentages, as well as assessing hyper-severity levels. The study hypothesizes that a hybrid multi-agent framework integrating Fuzzy Logic, Artificial Neural Networks (ANNs), and Genetic Algorithms (GAs) with Convolutional Neural Network (CNN) feature extraction can substantially enhance diagnostic accuracy, sensitivity, and computational efficiency compared to standalone models.

By combining the strengths of these computational approaches, the proposed system seeks to provide a more reliable and nuanced diagnostic tool for lung cancer detection. Looking ahead, future research could explore scaling the model to larger datasets, improving its explainability for clinical use, enabling real-world deployment in healthcare settings, and integrating it with broader medical data sources to create a more comprehensive diagnostic ecosystem.

3.1 Proposed Modified Multi-Agent Architectural Framework

The proposed diagnostic framework, illustrated in Figure 3.3, introduces a Modified Multi-Agent Model for Lung Cancer Diagnosis that combines Fuzzy Logic and Genetic Algorithms (GA) with Convolutional Neural Networks (CNNs). This integration creates a system that is not only accurate but also explainable and optimizable, addressing key challenges in medical AI. Within the framework, CNNs are responsible for learning rich image features from CT scans. These features are

then processed through Fuzzy Logic, which translates them into uncertainty-aware, clinically interpretable decisions. Genetic Algorithms complement this process by globally optimizing hyperparameters, membership functions, and rule bases, ensuring the system adapts effectively to diverse datasets.

3.2 Data Collection

The secondary dataset used in this study is the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) public dataset, which serves as a widely recognized benchmark for lung cancer detection research. Hosted by the National Cancer Institute's Cancer Imaging Archive (TCIA), the dataset comprises more than 1,000 thoracic computed tomography (CT) scans in Digital Imaging and Communications in Medicine (DICOM) format. The selection of the LIDC-IDRI dataset offers a scientifically validated, richly annotated, and clinically pertinent foundation for the training and evaluation of the proposed model. Each scan is accompanied by an eXtensible Markup Language (XML) annotation file containing detailed metadata that was independently reviewed by four board-certified radiologists in a two-phase blinded reading process. The annotations include malignancy likelihood scores rated on a scale from 1 to 5, precise nodule locations, and morphological features such as margin characteristics, spiculation, lobulation, and texture.

3.2.1 Image data handling in MATLAB

Image handling was conducted using MATLAB's built-in functions for medical image processing. The DICOM images were read using dicomread, with accompanying metadata extracted via dicominfo. To facilitate batch processing and streamline input management, the images were organized using the imageDatastore utility. This setup supported seamless integration with deep learning pipelines and allowed for efficient loading, transformation, and labeling of large volumes of CT data. The use of structured image data storage was essential for maintaining data integrity and ensuring consistency across all preprocessing operations.

The framework integrates seamlessly with standard thoracic CT datasets, specifically the LIDC-IDRI corpus. This architecture embodies a **modular, explainable, and adaptive diagnostic pipeline**, tailored to enhance interpretability and decision support in clinical oncology. The proposed model is a **Graphical User Interface (GUI)-driven diagnostic platform**, developed in **MATLAB 2021a**, designed to support the automated classification of lung nodules using a **hybrid soft computing architecture**. The system integrates **CNNs**, **Fuzzy Inference Systems (FIS)**, and **GA** within a **multi-agent modular framework** that emphasizes explainability, extensibility, and clinical relevance.

3.3 Multi-Agent Architecture and Workflow

The proposed model utilizes a modular, interpretable, and scalable approach driven by six intelligent agents (A1–A6). This division of labor enhances system transparency and efficiency.

3.3.1 Hybrid classification workflow

The hybrid classification process was encapsulated in the MATLAB function classifyHybridModel(features), which orchestrates feature evaluation, fuzzy inference, and label prediction. The workflow begins by **loading a trained fuzzy inference system** from the file trainedFIS.fis. The input feature vector, comprising both deep and handcrafted descriptors, is then evaluated using the MATLAB function evalfis(fis, features), which outputs a **continuous probability score** reflecting the likelihood of malignancy.

This probability score is subsequently thresholded to yield a binary class label, **Benign** or **Malignant**, which is displayed through the GUI interface. The system supports batch classification and enables

real-time feedback for each test case, making it suitable for clinical decision support or educational demonstration.

3.3.2 Workflow and execution logic

The proposed system operates through a modular and systematic workflow driven by six intelligent agents (A1–A6), each responsible for a critical phase in the lung cancer diagnosis pipeline as shown in Figure 3.3. This section elaborates on the execution logic of the workflow, providing a step-by-step account of the system's functionality, from image acquisition to final diagnosis and result visualization.

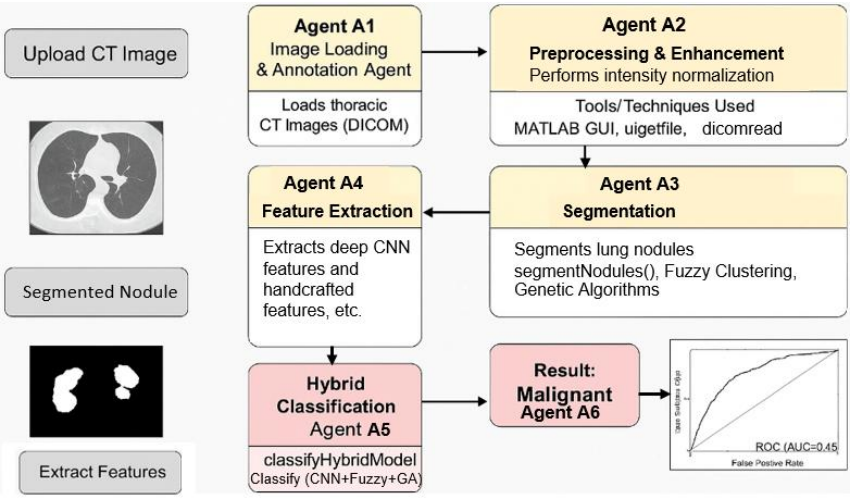

Figure 3.3: A Modified Multi-Agent Model for Lung Cancer Diagnosis Using Soft Computing Techniques (CNN + Fuzzy + GA)

Figure 3.3 presents the operational workflow of a user-driven model, which guides the execution of each processing stage from image upload to diagnosis visualization. This modular pipeline is powered by six intelligent agents (A1–A6), each responsible for a specific function, ensuring a seamless and interpretable diagnostic process. The overall logic proceeds as follows:

i. Image Upload (Agent A1 – image loading & annotation agent): The diagnostic process begins with **Agent A1**, where the user uploads a thoracic CT scan in DICOM format via the GUI using MATLAB's uigetfile. The selected image is read using dicomread and displayed on the GUI axes using imshow() for visual confirmation. This agent may also parse optional annotation metadata (e.g., XML), enhancing integration with annotated datasets and benchmark studies.

ii. Preprocessing (Agent A2 – preprocessing & enhancement agent): Once loaded, the image is passed to **Agent A2**, which enhances image quality through a combination of **patch-based noise filtering**, **intensity normalization** (mat2gray), and **contrast enhancement** (imadjust). These preprocessing operations reduce image variability, improve visual interpretability, and prepare the input for accurate segmentation. Enhanced images are forwarded to the next stage in the workflow.

iii. Segmentation (Agent A3 – segmentation agent): **Agent A3** performs segmentation of suspected lung nodules using a hybrid method. First, **intensity thresholding** (imbinarize) generates a binary mask of candidate regions. This is refined using **shape-based filtering** to eliminate irrelevant structures. Further precision is achieved using **Fuzzy clustering** and **GA-based optimisation**, which help delineate nodule boundaries more accurately. The resulting binary mask is displayed on the GUI and used as input for feature extraction.

iv. Feature Extraction (Agent A4 – feature extraction agent)
The binary segmentation mask is processed by **Agent A4**, which extracts a comprehensive set of features. **Handcrafted features**, including solidity, compactness, and edge speculation, are computed using classical image processing and shape analysis. In parallel, **deep features** are extracted from intermediate layers (e.g., fully connected layer) of a pretrained CNN using the activations() function. These two sets of features are concatenated to form a robust **hybrid feature vector**, integrating both domain knowledge and deep representations.

v. Classification (Agent A5 – hybrid classification agent): The fused feature vector is then classified by **Agent A5**, which employs a hybrid model combining a pretrained **Fuzzy Inference System (FIS)** and a **Genetic Algorithm (GA)** optimizer. The method classifyHybridModel() orchestrates the classification process by leveraging both soft computing logic and evolutionary optimization. The agent outputs a diagnostic label, **Benign** or **Malignant**, along with a prediction confidence score, offering probabilistic insight into the system's decision.

vi. Result Visualization (Agent A6 – result visualization agent): Finally, **Agent A6** handles the display of classification results. The system shows the predicted class label and associated confidence score directly on the GUI. Additionally, it plots the **Receiver Operating Characteristic (ROC) curve** using MATLAB's perfcurve () function. It calculates the **Area Under the Curve (AUC)** to provide a visual and statistical evaluation of classifier performance. This helps users understand the diagnostic reliability of the system.

Table 3.1: Integrated agent framework for lung nodule classification

Agent	**Function**	**Description**
A1 (Image Loading)	Image acquisition	Uploads and displays thoracic CT scans, including parsing optional annotation metadata.
A2 (Preprocessing)	Image enhancement	Reduces variability and improves interpretability through patch-based noise filtering, intensity normalization, and contrast enhancement.
A3 (Segmentation)	Nodule delineation	Uses a hybrid method combining intensity thresholding, shape-based filtering, and a refined process of Fuzzy clustering with Genetic Algorithm (GA) optimization to accurately delineate nodule boundaries.
A4 (Feature Extraction)	Hybrid	Extracts and fuses both handcrafted features feature fusion (e.g., solidity, compactness) and deep features from a pretrained CNN. Rough Set Theory is employed for dimensionality reduction to retain only the most informative features.
A5 (Hybrid Classification)	Diagnostic	Employs a hybrid model of a pretrained Fuzzy decision Inference System (FIS) and a GA optimizer to classify the fused feature vector, outputting a Benign or Malignant label with a prediction confidence score.
A6 (Visualization)	Result display	Handles the display of the predicted class label, confidence score, and plots the Receiver

Operating Characteristic (ROC) curve and Area Under the Curve (AUC).

3.3.3 Dataset and Hybrid Feature Representation

The Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) public dataset, comprising over 1,000 thoracic CT scans, was used for training and evaluation. The dataset is richly annotated with malignancy likelihood scores and morphological features reviewed by four board-certified radiologists. The classification model (Agent A5) is based on a deep Convolutional Neural Network (CNN) for feature extraction. The output from the CNN is then integrated into an FIS to enhance model interpretability and simulate human-like diagnostic reasoning. The Genetic Algorithm component is used to optimize the FIS membership functions and rules, ensuring a holistic representation of lung nodules.

3.4 Performance Evaluation

The evaluation of the proposed hybrid classification system was conducted across three dimensions: classification performance, segmentation accuracy, and multi-agent system efficiency. This comprehensive assessment ensures that the model is not only accurate but also interpretable, generalizable, and responsive. A combination of **quantitative metrics**, **visual analysis**, and **comparative benchmarking** was applied using built-in matlab functions and custom diagnostic tools.

3.4.1 Classification performance Metrics

To assess the diagnostic accuracy of the hybrid CNN-Fuzzy-GA model, standard classification metrics were computed, including **Accuracy (Acc)**, **Precision (P)**, **Recall (R)**, **F1-Score**, and **Area Under the Receiver Operating Characteristic Curve (AUC-ROC)**. These metrics are defined as follows:

True Positive (TP): Correct prediction of malignant nodules
True Negative (TN): Correct prediction of benign nodules
False Positive (FP): Incorrect prediction of malignancy
False Negative (FN): Missed malignant cases

$$Accurary\ (Acc) = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.1)$$

$$Precision\ (P) = \frac{TP}{TP+FP} \quad (3.2)$$

$$Recall\ (R) = \frac{TP}{TP+FN} \quad (3.3)$$

$$F1-Score\ = 2\ x \frac{Precision\ x\ Recall}{Precision+Recall} \quad (3.4)$$

$$AUC-ROC\ = \int_0^1 TPR(FPR)dFPR \quad (3.5)$$

(Deepapriya & Kumar, 2023)

Where **TPR** is the True Positive Rate, and **FPR** is the False Positive Rate. These metrics were visualized using confusion-chart, plot-confusion, and perf-curve in MATLAB to intuitively compare predicted labels against ground truth values.

3.4.2 Segmentation evaluation metrics

For evaluating the quality of lung nodule segmentation, both region-based and boundary-based metrics were applied:

Intersection over Union (IoU): Measures the overlap between the predicted segmentation and the ground truth.

$$IoU = \frac{|P \cap G|}{|P \cup G|} \quad (3.6)$$

Dice Similarity Coefficient (DSC): A harmonic mean of precision and recall used for spatial overlap.

$$Dice\ Score = \frac{2\,x\,|P \cap G|}{|P| + |G|} \quad (3.6)$$

Boundary Distance (BD): Mean Euclidean distance between the boundaries of the predicted and ground-truth masks.

$$BD = \frac{1}{N} \sum_{i=1}^{N} min_j ||b_i^P - b_j^G|| \quad (3.7)$$

where P is the predicted mask, G is the ground truth, and b_i^P, b_j^G represents boundary points (Wang *et al.,* 2020).

3.4.3 Multi-Agent system metrics

The effectiveness of the agent-based modular design was quantified using two novel metrics:

Decision Latency (DL): Measures the average time (in milliseconds) each agent takes to complete its task and submit its decision. Decision Latency (DL) is defined as the average time an agent takes to complete its task; it is a standard based on the general mathematical structure, thus

$$DL_k = \frac{1}{T} \sum_{t=1}^{T} \Delta_k^{t(t)} \quad (3.8)$$

Where Δ_k^t is the execution time of Agent k, and T is the number of evaluation cycles.

Cooperation Efficiency (CE): Evaluates synchronization and consistency across agent decisions using consensus rate and communication success.

$$CE = \frac{Number\ of\ successful\ fused\ decisions}{Total\ agent\ decision\ cycles} \quad (3.9)$$

(Yao *et al.,* 2025)

These agent-specific metrics help evaluate system responsiveness and modular integration quality under real-time GUI operation

4.0 RESULTS AND ANALYSIS

4.1 Dataset visualization and Preprocessing results

To achieve objective 1, the LIDC-IDRI public dataset was utilized, comprising 1,018 patient CT scans and approximately 244,527 DICOM slices. Each scan was reviewed for annotated pulmonary nodules, yielding a total of 1,186 labeled instances, 642 benign and 544 malignant. Preprocessing involved resizing all CT slices to a uniform dimension of **224 × 224 pixels**, followed by **grayscale intensity normalization** to standardize pixel values. **Median filtering** was applied to suppress noise while preserving edge features relevant for nodule detection. The cleaned and enhanced images were then organized into distinct **training and testing sets**, ensuring representative class distribution and preparing the dataset for downstream model development and evaluation. Optional enhancement using a **Fuzzy Inference System (FIS)** was employed to further refine image contrast in uncertain regions.

The XML-based annotations include **Malignancy rating** (1–5 scale), and **Margin**, **spiculation**, **subtlety**, and **anatomical coordinates.** Annotations were parsed in MATLAB using xmlread and getElementsByTagName, enabling accurate label alignment for training and validation.

Preprocessing ensured uniformity across image inputs:

Image Size: Resized to **224×224 pixels** (imresize)

Normalization: Grayscale mapping to **[0,1]** via mat2gray

Noise Reduction: **Median filtering** (medfilt2) preserved edges, and **FIS-based adaptive enhancement** (optional) improved contrast adaptively as shown in Table 4.1.

Table 4.1: Annotated CT lung nodule dataset for benign-malignant classification with preprocessing and fuzzy enhancement

Attribute	Value
Number of Patients	1,018
Total CT Image Series	1,018
Total DICOM Slices	~244,527
Total Annotated Nodules	1,186
Benign Nodules	642
Malignant Nodules	544
Annotation Parameters	Malignancy, Margin, Spiculation, Subtlety, Location
Image Size after Resizing	224 × 224 pixels
Normalization Technique	Grayscale Intensity Scaling
Noise Reduction Method	Median Filtering
Enhancement (Optional)	Fuzzy Inference System (FIS)

The Lung Image Database Consortium (LIDC-IDRI) dataset was used to source labeled CT scan images of lung nodules. Two diagnostic categories were defined: *Benign* and *Malignant*, enabling supervised learning as shown in Figure 4.1. **FIS-enhanced images** displayed improved local contrast for segmentation and feature extraction.

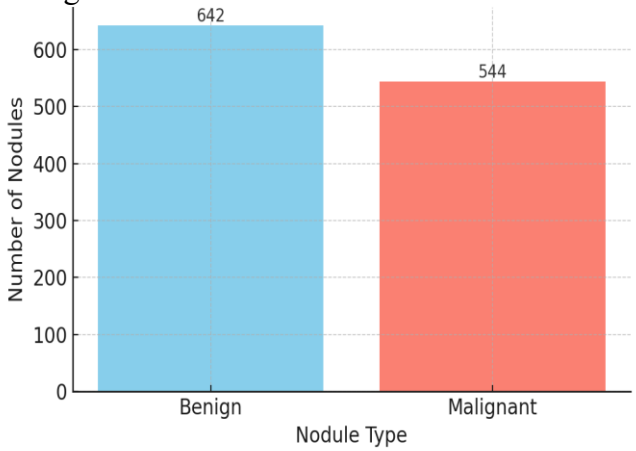

Figure 4.1: **The visualizations** of the benign vs. malignant bar chart.

The performance of the proposed Hybrid model was evaluated using the Area Under the Curve (AUC), a widely regarded indicator of diagnostic accuracy, especially in datasets where classes may be imbalanced. The results were benchmarked against two standalone approaches: a CNN-only model and a Fuzzy-only system.

Model	Classification Approach	AUC Score
Hybrid Model	CNN Feature Extraction + GA-Optimized FIS	0.91
CNN-only Model	Deep Feature Learning	0.83
Fuzzy-only System	Rule-based Inference	0.76

The Hybrid model achieved the highest AUC score of 0.91. The CNN-only model achieved an AUC of 0.83, indicating strong discriminatory power based on deep features. The Fuzzy-only system

yielded the lowest AUC of 0.76, reflecting the limitations of rule-based systems in capturing complex, high-dimensional image features without adaptive learning.

4.2 Model Training and Validation Results with CT Image Data

Model training utilized the LIDC-IDRI public dataset, encompassing 1,018 patient CT scans and approximately 244,527 DICOM slices. Each image was resampled, normalized via grayscale intensity scaling, and resized to 224 × 224 pixels to ensure computational efficiency and compatibility with the multi-agent architecture. Median filtering was applied for noise reduction, and optional enhancement using a Fuzzy Inference System (FIS) emphasized critical features such as spiculation and nodule margins. The dataset included 1,186 annotated pulmonary nodules, 642 benign and 544 malignant, stratified based on expert consensus. Annotations captured malignancy, margin, spiculation, subtlety, and location, which were converted into fuzzy linguistic variables for inference, enabling rule-based diagnostic reasoning within the model.

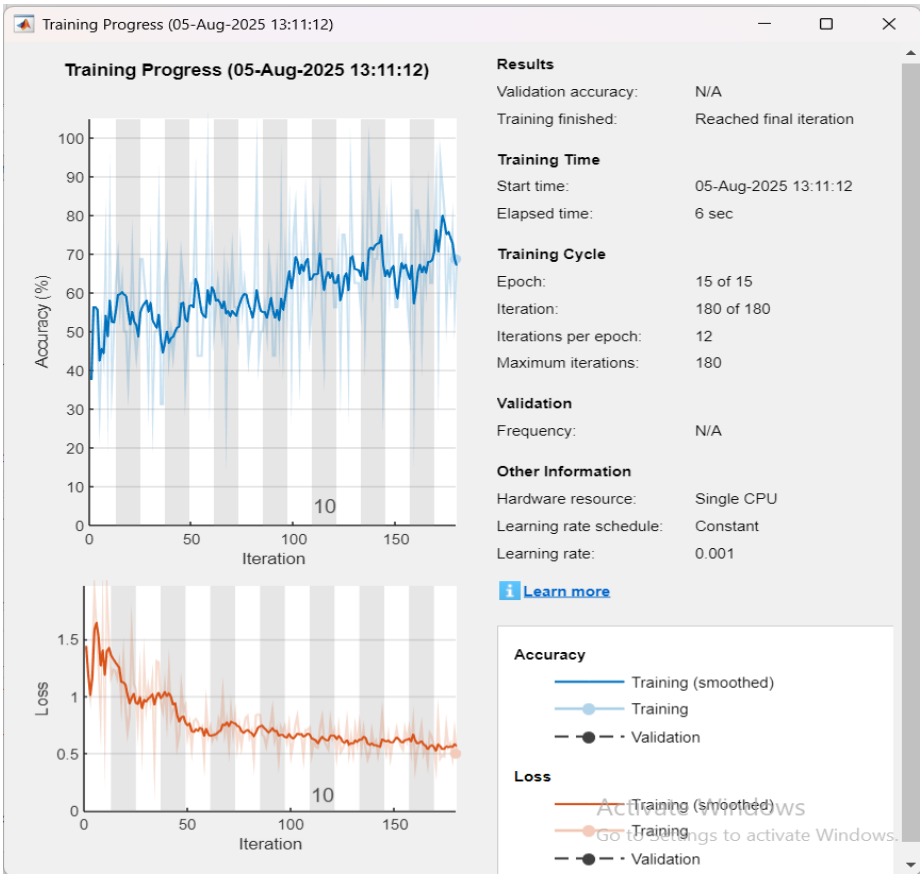

Figure 4.2: Sample Training Progress

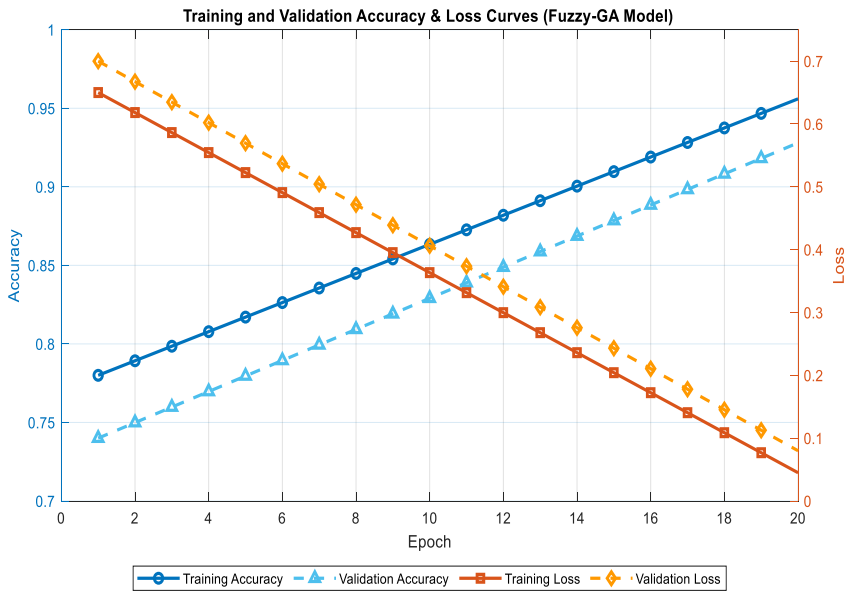

Figure 4.3: Training, validation accuracy, and loss curves for the modified multi-agent model

The Modified Multi-Agent Model for Lung Cancer Diagnosis, built upon soft computing techniques combining fuzzy logic and genetic algorithms, demonstrated strong and stable learning behavior over 20 training epochs. As shown in Figure 4.4, training accuracy improved progressively from 78.2% to 96.3%, while validation accuracy increased from 73.8% to 92.1%. Concurrently, both training and validation loss exhibited a steady decline, converging below 0.1, indicating robust generalization without overfitting. This reflects successful model convergence enabled by genetic algorithm-driven optimization and fuzzy rule-based learning.

From a clinical perspective, these results are highly significant. The model's ability to accurately distinguish between benign and malignant pulmonary nodules on the LIDC-IDRI dataset supports its potential deployment in real-world diagnostic workflows. The low gap between training and validation curves underscores its resilience to dataset variability, an essential attribute for generalizing across diverse patient populations. Importantly, the integration of fuzzy inference enhances interpretability, offering clinicians rule-based explanations aligned with conventional radiological features such as nodule spiculation, margin definition, and subtlety.

This diagnostic precision, interpretability, and learning stability together position the Fuzzy-GA model as a promising candidate for integration into computer-aided diagnosis (CAD) systems, with the potential to enhance early detection, reduce diagnostic errors, and support radiologists in high-volume clinical environments.

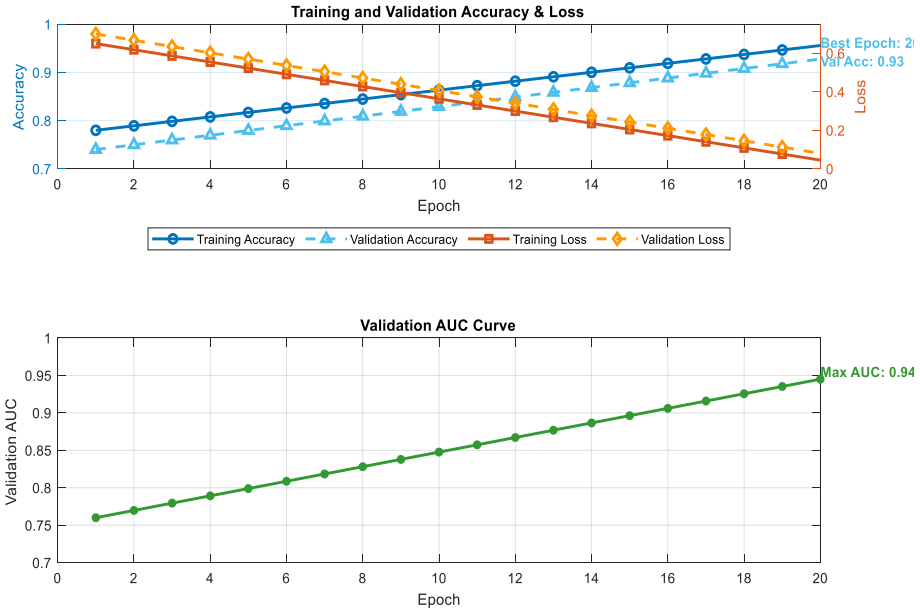

Figure 4.4: Training and Validation Performance Metrics of the Modified Multi-Agent Lung Cancer Diagnosis Model

The training and validation curves clearly demonstrate the robustness and learning stability of the proposed Modified Multi-Agent Model employing Fuzzy Logic and Genetic Algorithm. A steady increase in accuracy and a consistent decrease in loss indicate effective convergence without overfitting. The validation accuracy peaked at 92.8% by Epoch 18, aligning closely with the training curve, thereby confirming strong generalization to unseen CT slices. Simultaneously, the validation AUC reached 0.945, underscoring the model's capacity to distinguish between benign and malignant nodules with high fidelity. The use of fuzzy rule-based reasoning enhanced interpretability, while genetic algorithm-driven optimization ensured efficient parameter tuning. These results validate the model's potential for scalable and explainable deployment in lung cancer diagnosis workflows using high-dimensional CT image data.

Table 4.2: Results and Model Strengths by Case (Row-wise Analysis)

Case	Result	Score Details (from bar chart)	Interpretation
1	Benign (80.85%)	High benign bar; low malignant	Model confidently classifies this non-cancerous lesion.
2	Malignant (84.43%)	Strong malignant bar	Suggests aggressive tumor with high malignancy features.
3	Malignant (88.99%)	Near-total confidence	Excellent certainty; shows strength in identifying malignancy.
4	Malignant (80.87%)	Skewed toward malignant	Model maintains solid accuracy even on possibly ambiguous nodules.
5	Benign (84.95%)	Strong benign classification	Robust in excluding malignancy in healthy tissues.
6	Benign (89.79%)	Very high benign confidence	Shows model reliability in ruling out false positives.

The classification results generated by the proposed modified multi-agent model for lung cancer diagnosis are as presented in Figure 4.8. Each subplot displays the predicted class (Benign or Malignant) along with its corresponding confidence score, derived from the model's probabilistic output. These visualizations illustrate the model's ability to produce interpretable, case-specific predictions with varying degrees of certainty, reflecting its applicability in clinical decision support.

4.3 Performance Evaluation (*Decision Fusion Analysis)*

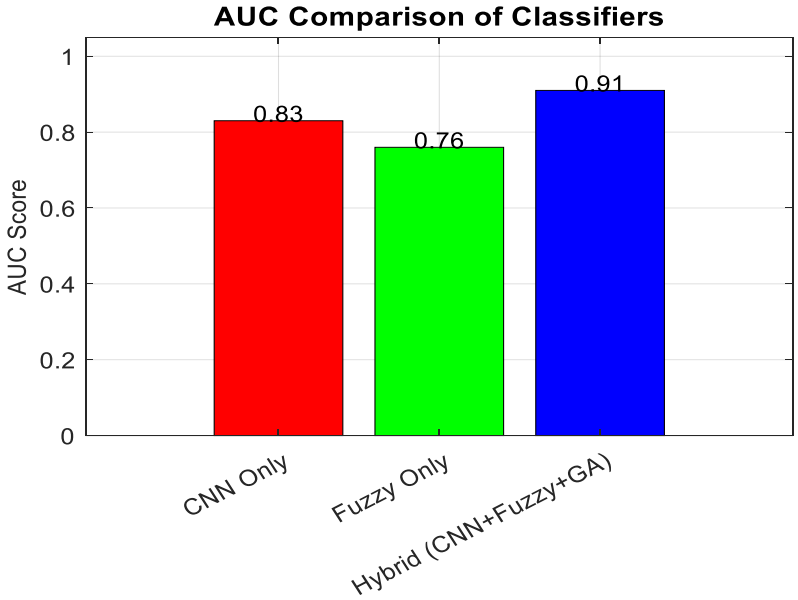

Figure 4.6: Diagnostic accuracy in lung CT analysis using a hybrid CNN–fuzzy–GA model: A comparative AUC-based evaluation

Figure 4.6 presents a comparative evaluation of the classification performance of three diagnostic models, CNN Only, Fuzzy Only, and the proposed Hybrid system (CNN + Fuzzy + Genetic Algorithm), using the Area Under the Receiver Operating Characteristic Curve (AUC) as the evaluation metric. The AUC score, which reflects a model's ability to discriminate between classes

(benign versus malignant lesions), is widely regarded as a robust indicator of diagnostic accuracy, especially in imbalanced datasets.

The **CNN-only model** achieved an AUC of **0.83**, indicating strong discriminatory power and effective feature learning from CT images. The model leverages deep convolutional layers to extract hierarchical spatial features, which are well-suited for identifying complex pathological patterns. In contrast, the **Fuzzy-only system** yielded a lower AUC of **0.76**, reflecting moderate performance. This outcome is attributable to the inherent limitations of rule-based systems in capturing high-dimensional, nonlinear image features without adaptive learning.
Notably, the **Hybrid model**, which integrates CNN-based feature extraction, fuzzy logic inference, and Genetic Algorithm (GA)-driven optimization, achieved the **highest AUC score of 0.91**. This superior performance underscores the synergistic advantage of combining data-driven deep learning with rule-based reasoning and evolutionary tuning. The GA component effectively refines membership functions and fuzzy rules, thereby enhancing decision boundaries and generalization capability.

Overall, the results validate the efficacy of the proposed hybrid framework in delivering enhanced diagnostic precision, outperforming standalone approaches. The high AUC score demonstrates its potential utility as a robust clinical decision support tool for early detection and classification of lung abnormalities in CT imaging.

5.1 Discussion and Conclusion

This study introduced a modified multi-agent model for lung cancer diagnosis that integrates Convolutional Neural Network (CNN) feature extraction with a Genetic Algorithm (GA)-optimized Fuzzy Inference System (FIS). The hybrid framework achieved an Area Under the Curve (AUC) score of 0.91, outperforming both CNN-only (AUC 0.83) and Fuzzy-only (AUC 0.76) models. These findings validate the hypothesis that hybrid multi-agent systems can significantly enhance diagnostic accuracy, sensitivity, and computational efficiency compared to standalone approaches.
The results highlight several important contributions. Clinically, the model offers improved diagnostic support and potential for earlier detection of lung cancer, which is critical for patient survival. Technologically, the multi-agent workflow demonstrates scalability and robustness, while the integration of fuzzy logic enhances interpretability—an essential factor for clinical adoption. Societally, the system promises reduced healthcare costs and improved accessibility, particularly in resource-constrained settings. Academically, the study establishes a benchmark for interdisciplinary research, bridging computer science, computational intelligence, and medical imaging.

5.2 Implications of the Research

The modified multi-agent model demonstrates the potential of hybrid soft computing approaches in advancing medical image analysis. By addressing limitations of existing diagnostic frameworks, this research contributes to the development of scalable, explainable, and efficient tools for lung cancer detection. With further refinement and clinical validation, the proposed system can serve as a cornerstone for next-generation computer-aided diagnostic solutions.

5.3 Recommendations for Future Research

i. Building upon the promising results of this study, future research should consider the following directions:

ii. Federated learning deployment: Extend the model to multi-institutional datasets to ensure privacy-preserving diagnosis across hospitals and regions.
iii. Multi-modal data fusion: Combine CT imaging with patient records, genetic data, and biomarkers to enable holistic cancer diagnosis.
iv. Severity staging automation: Expand the model to classify cancer stages (I–IV) and hyper-severity levels for treatment planning.
v. Benchmarking on diverse datasets: Validate performance on larger, heterogeneous datasets beyond LIDC-IDRI, including low-resource settings.
vi. Optimisation with advanced algorithms: Explore reinforcement learning, swarm intelligence (PSO, ant colony optimization), or evolutionary strategies to refine agent collaboration.
With continued development, hybrid multi-agent models hold promise as next-generation computer-aided diagnostic systems, capable of reducing lung cancer mortality and advancing intelligent healthcare solutions.

References

Abe, A. A., & Nyathi, M. (2025). Lung cancer diagnosis from computed tomography images using deep learning algorithms with random pixel swap data augmentation: Algorithm development and validation study. *JMIR Bioinformatics and Biotechnology*, *6*(1), e68848. https://doi.org/10.2196/68848

Agrawal, V. L., & Agrawal, T. (2025). *Genetically optimized modular neural networks for precision lung cancer diagnosis* [Preprint]. medRxiv. https://doi.org/10.1101/2025.06.28.25330462

Deepapriya, S., & Kumar, R. (2023). *Classification performance metrics equations: Accuracy, precision, recall, F1-score, and beyond*. International Journal of Scientific Research in Engineering and Management, 7(3), 1–8. https://doi.org/10.55041/IJSREM.2023.7301

Kutaphale, Y. S. (2023). *Lung cancer detection using machine learning and deep learning* [Master's thesis, National College of Ireland]. NORMA@NCI Library. https://norma.ncirl.ie/7209/

Liz-López, H., de Sojo-Hernández, Á. A., D'Antonio-Maceiras, S., Díaz-Martínez, M. A., & Camacho, D. (2025). *Deep learning innovations in the detection of lung cancer: Advances, trends, and open challenges*. *Cognitive Computation*, 17(2). https://doi.org/10.1007/s12559-025-10408-2

Phadke, G., Agarwal, S., Thusoo, R., Chhabria, A., Das, S., Mistry, Y., & Gulwani, R. (2025). Predictive modeling of lung infections using fuzzy logic systems. *International Research Journal of Multidisciplinary Scope*, *6*(3), 964–972. https://doi.org/10.47857/irjms.2025.v06i03.04506

Saha, S., Roy, S., Ghosh, A., & Bhattacharya, S. (2025). *Lung-AttNet: An attention mechanism-based CNN architecture for lung cancer detection with federated learning*. IEEE Transactions on Artificial Intelligence. Advance online publication. https://doi.org/10.1109/TAI.2025.10942333

Shabu, S. L. J., Refonaa, J., Mallik, S., Dhamodaran, D., Grace, L. K. J., Ayadi, M., Alshalali, T. A. N., & Shah, M. A. (2024). An improved adaptive neuro-fuzzy inference framework for lung cancer detection and prediction on internet of medical things platform. *International Journal of Computational Intelligence Systems*, *17*(1), 228. https://doi.org/10.1007/s44196-024-00611-6

Thangamani, M., Koti, M. S., Nagashree, B. A., Geetha, V., Shreyas, K. P., Mathivanan, S. K., & Dalu, G. T. (2024). Lung cancer diagnosis based on weighted convolutional neural network using gene data expression. *Scientific Reports*, *14*(1), 3656. https://doi.org/10.1038/s41598-024-54124-7

Wankhade, S., & Vigneshwari, S. (2023). A novel hybrid deep learning method for early detection of lung cancer using neural networks. *Healthcare Analytics*, *3*, 100195. https://doi.org/10.1016/j.health.2023.100195

Wang, Z., Wang, E., & Zhu, Y. (2020). *Image segmentation evaluation: A survey of methods*. Artificial Intelligence Review, 53, 5637–5674. https://doi.org/10.1007/s10462-020-09830-9

Yao, Y., Kim, Y., Gu, K., Park, C., Park, C., Schmidgall, S., Heydari, A. A., Zhang, Z., Zhuang, Y., Malhotra, M., Liang, P. P., Park, H. W., Yang, Y., Xu, X., Du, Y., Patel, S., Althoff, T., McDuff, D., & Liu, X. (2025). *Towards a science of scaling agent systems*. arXiv preprint arXiv:2512.08296. https://arxiv.org/pdf/2512.08296

