International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 13, No. 2, 2026, pp. 498-513.

ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

Hybrid Capsule Network with Graph Neural Networks for Accurate Feature Extraction and Tumor Localization in Lung Cancer Detection from CT Scans

Teena G. Nath [a], Dr.B. Chithra [b*], Dr.V. Kathiresan [c], Dr.B.M. Rajesh [d]

[a] *Research Scholar, Department of Computer Science, AVP College of Arts and Science, Tirupur, Tamil Nadu, India. E-mail: teenadale1885@gmail.com, Orcid: https://orcid.org/0009-0006-5490-359X*
[b*] *Assistant Professor, Department of Computer Science, AVP College of Arts and Science, Tirupur, Tamil Nadu, India. E-mail: chithra220@gmail.com, Orcid: https://orcid.org/0000-0002-1260-7470*
[c] *Associate Professor, Department of Computer Science, AVP College of Arts and Science, Tirupur, Tamil Nadu, India. E-mail: vkathirmca@gmail.com, Orcid: https://orcid.org/0000-0002-2799-1102*
[d] *Associate Professor, Department of Computer Science, Dr. N.G.P. Arts and Science College, Kalapatti, Coimbatore, Tamil Nadu, India. E-mail: rajenlight@gmail.com, Orcid: https://orcid.org/0000-0002-4207-0694*

Abstract

The model was evaluated on the LIDC-IDRI lung CT dataset using standard performance measures, including accuracy, precision, recall, F1-score, Matthews Correlation Coefficient (MCC), and Area Under the Receiver Operating Characteristic Curve (AUC). The experimental findings show that the proposed framework attains 83.33% accuracy, 85.00% precision, 80.95% recall, 82.93% F1-score, MCC of 0.6674, and AUC of 0.9045, outperforming baseline classifiers including Decision Tree, Voting Classifier, AdaBoost, and Stacking. Early detection of lung cancer is crucial for patient survival, but conventional Neural Networks (CNNs) often fail to detect and localize tumors due to their limited spatial and relational knowledge. To address these issues, this paper proposes a hybrid Capsule Network-Graph Neural Network (CapsNet-GNN) architecture. The Caps Net module extracts hierarchical and spatial information from CT scans of the lung, whereas the GNN module simulates inter-nodule relationships to improve tumor localization. The contributions of the Caps Net and GNN modules are also confirmed in ablation studies, which show that Caps Net and GNN complement each other in feature representation and relational learning. These findings provide statistical support for the soundness, generality, and clinical usefulness of the suggested framework, offer a dependable answer to automated lung cancer identification and tumor localization, and can be incorporated into clinical decision-support mechanisms.

Keywords: Lung Cancer Detection, Capsule Network, Graph Neural Network, CT Scan, Tumor Localization, ROC-AUC, MCC.

DOI: 10.71086/IAJSE/V13I2/IAJSE1390
Received: 30 March 2026/Revised: 06 May 2026/Accepted: 15 June 2026/Published: 30 June 2026

Introduction

Lung cancer is a very aggressive and deadly cancer known to be a major cause of cancer related mortality in the world each year. Regardless of the progress in medical imaging and treatment methods, the patient survival rate with lung cancer is still low, mainly because the cancer is diagnosed in late stages and the nature of Tumors is complicated (Bala & Kannan, 2025). Early diagnosis and proper localization of lung Tumors are thus important for improving clinical outcomes and enabling timely therapeutic intervention (Escarela et al., 2020). Computer tomography (CT) is a diagnostic imaging modality that has become a standard for screening pulmonary cancer, providing high-resolution visualization of pulmonary architecture and Tumor localization. However, manual interpretation of CT scans is time-consuming and interpretation itself is subjective and can be investigated by inter-observer variation, the reason why automated and intelligent computer-aided diagnostic systems are required (Aravind & Balasubramaniam, 2022).

Over the past several years, deep learning-based systems, particularly, Convolutional Neural Networks (CNNs), show remarkable performance in the medical image analysis, lung nodule detection, classification and segmentation. CNNs are highly effective at discovering discriminative features in large-scale imaging data, but they have intrinsic limitations in capturing spatial hierarchies and representing complex interactions between anatomical structures (Liu et al., 2024). The pooling operations and local receptive fields of CNNs tend to lose important spatial details, which are vital for accurate tumor localization, particularly when nodules are small, irregular, or overlap. Therefore, CNN-based models can have difficulty with extrapolation across heterogeneous tumor patterns and across different stages of the disease (Yang et al., 2023; Hudec et al., 2024).

To address these shortcomings, Capsule Networks (Caps Nets) have been proposed as a potential replacement for conventional CNN architectures (Nam et al., 2024). Caps Nets are used to encode hierarchical spatial relationships, allowing features to be encoded not only as features but also as their spatial attributes, such as orientation, position, and scale (Jiang et al., 2023). This feature makes Caps Nets especially well-suited for medical imaging applications in which structural and spatial data are crucial for proper diagnosis. However, Caps Nets also enhance local feature representation, but the global relationships and interactions between two or more Tumor regions in an image or a volumetric CT scan are not explicitly modeled.

Graph Neural Networks (GNNs) are gaining popularity recently because they are capable of modeling complex relational structures as graphs of nodes and edges (Alzaidi, 2024; Hussain et al., 2025). GNNs can be applied in medical imaging, and they are able to learn spatial correlations and characteristics of the immediate surroundings to make inferences regarding the correlations of the anatomical regions, which can further understand the distribution and interrelations of Tumor (Alam et al., 2024). GNNs provide an effective framework for improving Tumor localization and interpretability by combining graph-based reasoning with deep feature extraction (Nandhinieswari & Indumathi, 2024). Nevertheless, joint use of Caps Nets and GNNs for detecting and localizing lung cancer is a relatively unexplored field (Lau et al., 2020; Jalali et al., 2021).

These observations have inspired this research, and this paper proposes a new hybrid model that combines Capsule Networks with Graph Neural Networks to achieve resilient feature extraction and effective tumor localization during lung cancer diagnosis using CT scans (Alblehai et al., 2024). The proposed model combines the advantages of the spatial representation of CapsNets and the power of GNNs to address the weaknesses of the conventional CNN-based approaches (Yuan et al., 2023; Padmanaban et al., 2024). The method should have higher detection and localization rates and generalization as well, in a comprehensive experiment on multiple lung CT datasets (Faizi et al., 2025; Li et al., 2022). The study will be used to develop high-quality, interpretable, and clinically relevant deep learning systems to diagnose early lung cancer and automatically localize tumors (Asiya & Sugitha, 2025).

Key Contributions of the Research

- A new hybrid deep learning model combining Capsule Networks and Graph Neural Networks is proposed to improve lung cancer detection from CT scans.
- The model is a good demonstration of spatial relations between Tumor regions and leads to improved localization of Tumor regions as opposed to conventional CNN methods.
- The given method is more accurate and performs better across a range of Tumor sizes and stages, indicating its potential for practical clinical use.

This paper is organized in the following way: Section 1 contains the problem statement, that is, lung cancer detection based on CT scans and the purposes and reasons why this hybrid Capsule Network and Graph Neural Network system is proposed. Section 2 conducts a literature review and outlines the limitations of the current deep learning methods.

Section 3 explains the recommended methodology, including dataset preparation, model structure, and training. Section 4 shows the results of the experiment and performance analysis of the proposed model and the current approaches. In Section 5, the importance and implications of the results is discussed. Lastly, Section 6 provides a conclusion and outlines future research directions.

Literature Review

Feature Extraction

Stanciu et al., (2023) propose a model of lung nodule definition through the principles of deep learning involving complex convolutional neural networks (CNNs) on the CT images. The main idea of this model is to improve sensitivity and reduce false positives with the help of multi-scale feature extraction and streamlined training methods. With this method, significant improvements in the accuracy of detection are observed in contrast to the traditional methods. There is however a serious limitation that CNN models do not usually capture the intricate spatial correlations between nodules.

Litjens et al., (2017) also feature extraction using Graph Neural Networks (GNNs), which pays particular attention to the relational properties of the anatomical structures of medical images. This graph model illustrates the anatomy structures as the nodes and their relationship as edges so that the areas of the disease can be understood better in the context. Findings have established that GNNs are superior to the conventional CNNs in segmentation and classification. Nevertheless, the study notes that the incorporation of such graph-based models with feature extractors such as CapsNet may enhance the robustness of the spatial feature learning, therefore the accuracy of the tumor detection and classification.

Li et al., (2022) the combination of the handcrafted radiomic characteristics with the deep neural networks is useful in improving the classification performance, although it remains ineffective in the accurate representation of spatial and contextual dependencies in the tumor regions. The article highlights the necessity of then adding in spatial and hierarchical data to allow localization of tumors more precisely, a field that is not yet developed in existing deep learning models.

Abdullahi et al., (2025) propose a deep-learning-based pulmonary cancer-detection system, which adds data-augmentation and neural networks to enhance the precision of the detection. Nevertheless, even though the attention-based mechanisms are used to enhance feature learning, the framework does not explicitly emphasize the modeling of hierarchical and relational spatial features, which would potentially restrict its applicability in more sophisticated medical imaging tasks.

Segmentation

Yan et al., (2018) present a hybrid system of deep learning that integrates different types of neural networks to detect lung cancer using CT images. They use feature extraction and classification modules to enhance the accuracy of the diagnosis. The results of benchmarking indicate that the model is more robust and that it can be generalized with different tumor sizes. Nonetheless, the main limitation of this hybrid model is that it does not consider inter-tumor relations and spatial dependencies between tumor regions in different regions and this impacts the performance of the segmentation in more complex scenarios.

Gogoshin & Rodin, (2023) present a deep learning pipeline with stages of segmentation and classification with 3D-VNet and 3D-ResNet to find lung cancer in CT images. The model has a greater precision and greater nodule detection performance particularly during segmentation tasks. The joint-use of segmentation and classification models shown above has a future potential, yet the model lacks relational learning or capsule-based spatial representation, which would become highly beneficial in improving the accuracy of segmentation by significantly improving the description of spatial hierarchies and interdependences of tumors.

Veyrin-Forrer et al., (2024) offer a systematic review of the deep learning technique to detect, segment, and classify lung cancer on CT scans. Models further classified in the paper as CNNs, hybrid networks, and ensemble learning relate their advantages and disadvantages in segmentation activity. Another major issue noted in the review is the imbalance in data, model interpretability as well as cross-data generalization. The paper states that future studies should be dedicated to multimodal learning and explainable AI, which may provide the solution to these issues and make segmentation more effective by integrating several spatial and relational modeling methods.

Kumar et al., (2023) survey the emerging applications of the Graph Neural Networks (GNNs) in cancer and oncology studies. GNNs are effective especially in learning multimodal biomedical data which contains complex graph

connections. GNNs can be used to make intricate representations of space and relationship in lung cancer detection, which is important in segmentation tasks. Nevertheless, the article indicates that GNNs are hardly used together with Capsule Networks to localize tumors and extract hierarchical features that would greatly enhance the accuracy of segmentation in CT scans.

Methodology

Data Acquisition and Preprocessing

Publicly available datasets, such as LIDC-IDRI, NSCLC-Radiomics, and other lung cancer image repositories, are used to obtain high-resolution chest CT scans in the first stage (Nam et al., 2024). Raw CT images usually differ in intensity, resolution, and orientation and require preprocessing to normalize the data. The preprocessing stages will include intensity normalization to scale pixel values, denoising with a Gaussian or median filter, and lung segmentation to isolate areas of interest (ROIs) that may include nodules. Segmentation may be performed either using traditional thresholding or pre-trained segmentation models, such as U-Net. The model is also enhanced with random rotations, horizontal and vertical flips, scaling, and elastic deformations to enhance the generalization of the model. Also, the augmented dataset addresses class imbalance by providing a sufficient number of benign and malignant nodules. The result of this step is standardized, clean 2D or 3D lung CT ROIs that are available for feature extraction.

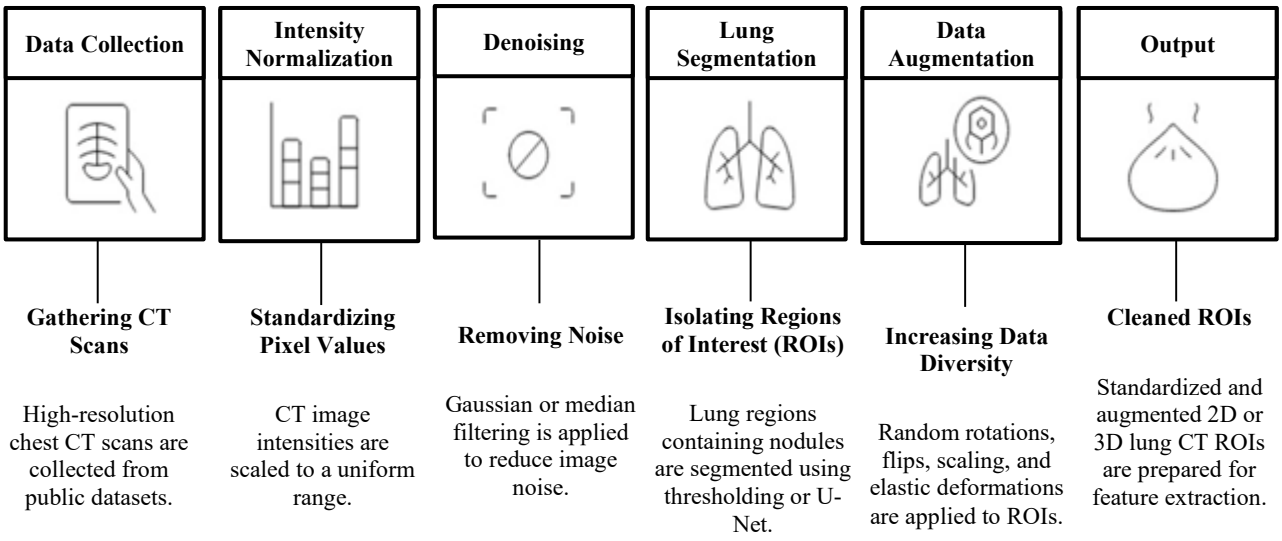

Fig. 1: Preprocessing Pipeline for Lung CT Image Preparation and ROI Generation

Fig. 1 shows the entire preprocessing workflow used to analyze CT images of the lungs, from data collection to clean regions of interest (ROIs). The high-resolution chest CT scans are initially collected from publicly available datasets, and the intensities are then normalized to equalize pixel values and eliminate interscan variation. This is then denoised either with a Gaussian or median filter to minimize noise without losing any critical structure. Next, the lung segmentation is conducted to isolate the lung with possible nodules and a thresholding-based or deep-learning-based algorithm, e.g., U-Net is run Data augmentation, such as rotations, flips, scaling, and elastic deformations are used on the segmented ROIs to improve the generalization of models and reduce the possibility of over fitting. The resultant standardized and enhanced ROIs are high-quality inputs to the downstream feature extraction and classification to provide robust and reliable model performance.

CNN Used in Proposed Model

CNNs (Convolutional Neural Networks) are selected in your proposed model as they have been demonstrated to be effective in the feature extraction tasks of medical images. CNNs are especially efficient in discovering discriminative information in a huge amount of data, e.g. CT images of the lungs. Hierarchical patterns are automatically learned by them on raw pixel data, making them capable of identifying a number of features including edges, textures, and shapes, which are essential in the detection of lung nodules. Such networks are also common in the detection of lung cancer because it can determine high dimensional imaging data and also performs efficiently in activities such as segmentation and classification.

Nonetheless, CNNs suffer some limitations including the fact that the pooling operations lead to loss of spatial relationships between features in addition to the fact that these operations are used to reduce the dimensionality and obtain high-level features. This is particularly a problem when small, irregular, or overlapping tumors are involved and it is necessary to retain the exact positional data of the tumor to be able to localize it accurately. Caps Nets can solve the spatial relationship issue by maintaining the orientation, position and hierarchy of features and GNNs can enhance the knowledge of inter-nodule relationships by capturing the spatial dependencies on a global basis. These networks collectively offer an all-round solution capable of increasing the accuracy, precision, and robustness of tumor detection and localization, to overcome the weaknesses of traditional CNN-based models.

Feature Extraction Using Capsule Networks (Caps Nets)

The use of Capsule Networks is to extract spatially located and hierarchical features of the lung CT scans. In contrast to traditional CNNs, which are usually vulnerable to the loss of spatial associations through the pooling capability, Caps Nets do not lose the processing properties of detected features, such as their orientation, location, or hierarchy. The Caps Net structure is made up of primary capsules and upper-level capsules. Primary capsules capture low-level information like edges, textures and small patterns that are related to nodule morphology. These features are then combined in higher-level capsules that are used to express more complicated structures such as Tumor shape, size, and orientation. The attention-weighted aggregation between capsules ensures that relevant features are selectively forwarded to high-level capsules, which enhance the capacity of the network to differentiate between normal tissue and nodules with minor differences. This step generates rich feature vectors of each nodule, which represents local and global spatial features. The Caps Net output can be used to obtain a strong feature portrayal, which is very informative in the further relational modeling by conserving the hierarchical relationships.

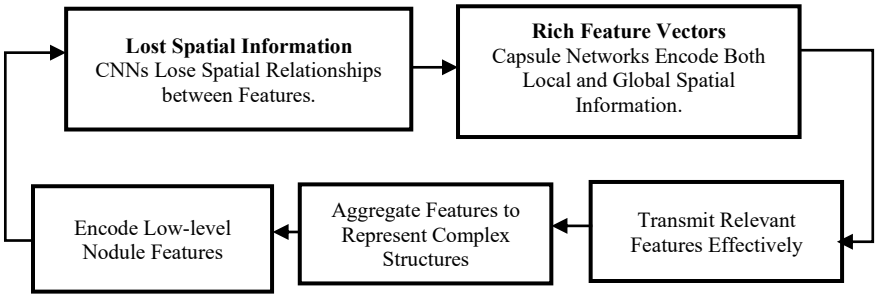

Fig. 2: Employing Capsule Networks to Preserve Spatial Relationships

Fig. 2 explains the reason why Capsule Networks should be adopted to overcome the shortcomings of typical CNN-based systems in medical image analysis. Classical CNNs can readily lose important spatial controls between features because of the pooling operations, so they may fail to represent delicate structures in the body. With the adoption of Capsule Networks, the low-level nodule characteristics are explicitly encoded and selectively aggregated to upper-level capsules, thus permitting the successful aggregation of the characteristics that translate into complex spatial hierarchies. This scheme enables the model to maintain part whole constraints and store both local and global space data in rich feature vectors. Capsule Networks are therefore effective in ensuring the efficient transmission of pertinent spatial features that lead to more discriminatory representations and enhanced robustness to the correct characterization of nodules of the lungs.

Relational Modeling and Classification Using Graph Neural Networks (GNNs)

These capsule-generated feature vectors are then mapped into a graph representation, each node corresponding to a detected nodule or a region of interest, and edges represent spatial or contextual connections between nodules. The interactions between Tumor regions can be learned in this formulation of the graph, thus this is essential in interpreting multi- nodule cases or how the geographical distribution of Tumors in the lungs is. The built graph is subjected to a Graph Neural Network (GNN). The GNN evolves the feature of nodes in the form of graph convolution layers, where every node combines the information of its neighbors, which essentially brings in the spatial dependency in a global scale. This enables the network to predict relationships amongst nodules that would be overlooked by conventional CNN pipelines. Following a series of graph convolution layers, a readout function is applied to aggregate information on the node level to generate global graph embeddings. Such embeddings are further classified by nodules (benign and malignant) and localization of Tumors. The proposed built-in Caps Net-GNN framework is not only more precise in detection but also more understandable, since spatial relationships between nodules are directly defined. This is a combination that is especially useful with nodules of complex shapes, overlapping Tumors and in cases of an early-stage lung cancer.

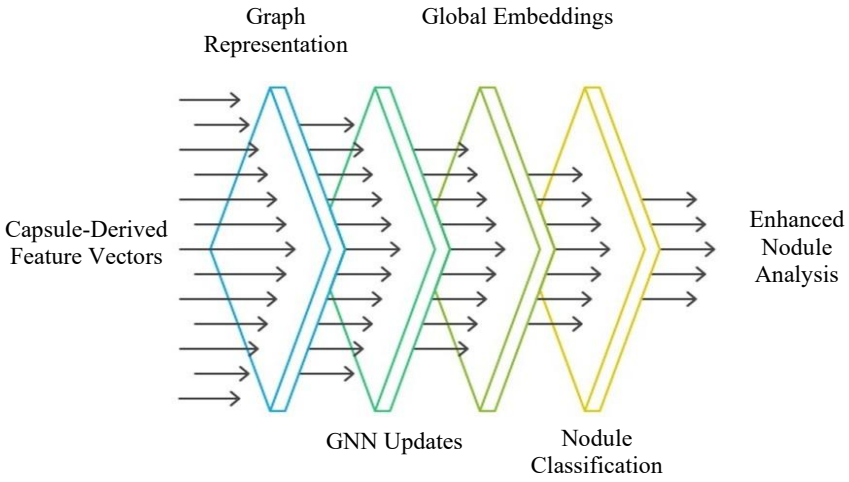

Fig. 3: Integration of Capsule-Derived Features with Graph Neural Networks for Enhanced Nodule Analysis

The fig. 3 shows the workflow of the suggested hybrid framework which incorporates Capsule Network-derived feature vectors and a Graph Neural Network (GNN) to analyze lung nodules in a comprehensive way. A set of rich features based on capsule features that describe the finer aspects of space and part-whole are initially obtained on segmented nodules in the lung. These characteristics are subsequently ranked into a graph-based representation of the nodules, with the nodes representing the nodules and the edges depicting their relational and contextual dependencies. By training GNN as an iterative process, both local and global embeddings are achieved, which is informed by the use of neighborhood information, bringing about the empowerment to model inter- nodule relationships effectively. The optimized global embeddings are then passed to the classification phase where the nodules are characterized better and the predictive accuracy is improved. It is this combination that has shown that the complementary capabilities of a capsule-based spatial encoding and a graph-based relational learning can be effectively used to produce strong and precise nodule analysis.

Capsule Aggregation and Feature Encoding in Capsule Networks

The routing between capsules determines how much a lower-level capsule i contributes to a higher-level capsule j

$$V_j = \text{squash}\left(\sum_i c_{ij}\, U_{j|i}\right) \qquad (1)$$

This equation (1) demonstrates how capsule aggregation allows lower-level capsules to contribute to higher-level capsules through attention-weighted coefficients, ensuring that hierarchical and spatial connections in the data are retained. Each lower-level capsule produces a prediction vector to the higher-level capsules, and the coupling coefficients identify the extent to which each prediction influences the higher-level capsule output. The weighted sum of these predictions is passed through a squash function, which rescales the vector length to represent feature presence probability while preserving orientation and instantiation parameters. This mechanism enables the network, in the context of lung CT scans, to remain sensitive to the shape, size, and orientation of nodules, including small, overlapping, or irregularly shaped cases, overcoming the spatial information loss associated with conventional CNN pooling operations. The attention-weighted aggregation further reduces false positives and enhances detection sensitivity by selectively reinforcing relevant spatial features, which is particularly effective for precise nodule characterization and tumor localization.

Algorithm 1: Hybrid Caps Net-GNN for Lung Nodule Detection and Localization

Input: Preprocessed lung CT images I = {I1, I2, ..., In}

Output: Nodule classification (benign/malignant) and Tumor localization

Step 1: Data Preprocessing

1.1 Load CT images from benchmark datasets (e.g., LIDC-IDRI)

1.2 Apply intensity normalization and denoising

1.3 Segment lungs to extract Regions of Interest (ROIs)

Step 2: Feature Extraction using Capsule Networks

2.1 Pass each ROI through a convolutional backbone to obtain deep spatial feature maps

2.2 Apply Primary Capsule layers to transform convolutional feature maps into vectorized capsule representations, preserving spatial and hierarchical information

2.3 Apply capsule transformation and feature aggregation to generate higher-level capsule feature vectors V_j*, encoding structured nodule representations*

2.4 Aggregate capsule outputs into a compact feature embedding for downstream relational modeling

Step 3: Graph Construction

3.1 Divide feature maps into spatial regions using region pooling (e.g., 4 × 4 grid)

3.2 Project each region feature vector into a graph node representation

3.3 Construct a graph $G = V + E$ *where:*

- *Each node represents a spatial region feature*
- *An additional node represents a mask-guided global nodule feature*

3.4 Connect nodes using k-Nearest Neighbor (kNN) edges based on feature similarity

Step 4: Relational Modelling with Graph Neural Networks

4.1 Initialize node features $h_j^{(o)}$ *using projected region and nodule features*

4.2 For each GNN layer l:

4.2.1 Update node features using neighborhood aggregation:

$$h_i^{(l+1)} = \sigma\left(\sum_{j \in N(i)} \frac{1}{\sqrt{D(i)}\sqrt{D(j)}} W^{(l)} h_j^{(l)}\right)$$

4.3 Compute refined relational node representations after multiple GNN layers

4.4 Fuse graph-based relational features with capsule embeddings

4.5 Apply fully connected layer + softmax for nodule classification

Step 5: Tumor Localization

5.1 Compute region-level importance scores from refined node features

5.2 Map region importance scores back to original CT spatial coordinates

5.3 Generate attention-based tumor localization heatmaps highlighting diagnostically relevant regions

Step 6: Output

6.1 Return nodule class labels (benign/malignant)

6.2 Return precise Tumor localization in CT images

End Algorithm.

The Algorithm 1 hybrid model, named CapsNet-GNN, consists of several components. First, the model preprocesses the lung CT image, which includes normalization, denoising, segmentation, and augmentation. Then, the regions of interest are fed into a convolutional backbone to obtain deep spatial feature maps, which are then used to obtain vectorized representations using the Primary Capsule layer. Higher-level capsule layers are used to transform the feature vectors and obtain nodule representations using feature transformation and attention-weighted aggregation. These representations are then normalized and stored. For graph construction, feature maps are split into a spatial grid of size 4×4 using region pooling. Each region is then projected to a graph node representation. An additional global nodule node is included using a mask, and all nodes are connected using k-Nearest Neighbor edges based on feature similarity. Neighborhood aggregation is then performed using a Graph Neural Network, which consists of two GNN layers. This captures local and global relational dependencies. Graph-based relational features are then fused with capsule-based representations and used to obtain nodule malignancy using fully connected classification. Region-level importance scores for tumor localization are computed based on refined node features, which are further transformed to original CT spatial coordinates and visualized using attention-based heatmaps. This framework successfully

incorporates local spatial feature extraction and global relational modeling, enabling robust lung cancer detection and localization tasks.

Experimental Results

Experimental Setup, Dataset, and Parameter Initialization

The tests were done on lung CT scan data on Python deep learning platform which utilized PyTorch. The proposed Caps Net -GNN model was tested on publicly available datasets like LIDC-IDRI, and the data was separated into training, validation, and testing parts. The preprocessing stages encompassed normalization, noise elimination, lung segmentation and augmentation of data to enhance the robustness of the model and alleviate bias on classes. The areas of interest which included lung nodules were extracted and resized and then fed into the network. The Capsule Network had the primary and higher-level layers of capsules, and the Graph Neural Network was constructed with many layers of graph convolutions that were used to model inter-nodule relationships. The learning rate, batch size, number of epochs and optimizer were also set and adjusted with the help of the validation data so that the training could be steady and the results would be maximized. Such an arrangement allowed doing the proper assessment of the suggested framework to detect lung cancer and localize Tumors.

Table 1: Experimental Setup and Parameter Initialization

Category	Parameter	Description / Value
Dataset	Dataset Name	LIDC-IDRI Lung CT Dataset
	Data Split	70% Training, 15% Validation, 15% Testing
	Image Size	224 × 224 pixels
	Preprocessing	Normalization, Noise Removal
	Data Augmentation	Rotation, Flipping, Scaling
Caps Net Parameters	Primary Capsules	32 Capsules
	Higher-Level Capsules	16 Capsules
	Routing Mechanism	Capsule transformation with feature aggregation
	Activation Function	Squash Function
GNN Parameters	Graph Layers	2 Graph Convolution Layers
	Node Features	Capsule Feature Vectors
	Edge Construction	k-Nearest Neighbor (kNN) based on cosine feature similarity
Training Parameters	Optimizer	Adam
	Learning Rate	0.001
	Batch Size	16
	Epochs	50
	Dropout Rate	0.5
	Loss Function	Cross-Entropy Loss
Hardware & Software	Framework	Python, TensorFlow, PyTorch
	Hardware	GPU-enabled workstation

Table 1 briefly describes the experimental setup and parameters adopted for testing the performance of the proposed approach using the Hybrid Capsule Network with Graph Neural Networks for lung cancer detection using CT scans. The LIDC-IDRI lung CT dataset was employed for the experiment and was divided into training, validation, and testing sets with a proportionate ratio of 70:15:15. pre-processing techniques like normalization, removal of noise, lung region extraction, and augmentation were employed to improve the overall performance and robustness of the model. The Capsule Network architecture employed a primary capsule and higher-level capsule layers for extracting the vector representation of the images using capsule transformation and squash functions. The Graph Neural Network module consists of two graph convolutional layers, which are applied to the region-based node representations obtained through spatial feature pooling. Edges are formed using a k-nearest neighbor strategy, which is based on cosine feature similarity, to model relations among the spatial regions of a CT image. The training of the model was carried out using the Adam optimizer with a learning rate of 0.001, a batch size of 16, and dropout regularization with a rate of 0.5, and the cross-entropy loss function to avoid overfitting. The experiments were conducted using the PyTorch deep learning library and a workstation with a GPU to facilitate efficient training and evaluation.

Performance Metrics

Accuracy, precision, recall, F1-score are considered to be the key performance metrics in evaluating the proposed model. The measures of accuracy consider the overall classification accuracy, whereas the measures of precision and recall consider the reliability and sensitivity of Tumor detection. F1-score offers a balanced measure of the precision and recall particularly when the data is imbalanced. To evaluate the localization of Tumors, the Dice coefficient is applied that evaluates the overlap between the predictive and ground truth regions. All these metrics will give an overall analysis of the proposed lung cancer detection framework. Deems the general accuracy of the model equation (2).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (2)$$

Precision indicates how many of the predicted positive cases are actually positive. High precision means fewer false alarms equation (3).

$$\text{Precision} = \frac{TP}{TP+FP} \qquad (3)$$

Recall measures how many actual positive cases were correctly detected. High recall ensures that arrhythmias are not missed equation (4).

$$\text{Recall} = \frac{TP}{TP+FN} \qquad (4)$$

F1-score is the harmonic mean of precision and recall, providing a single metric for imbalanced datasets equation (5).

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision}+\text{Recall}} \qquad (5)$$

Performance Evaluation of the Capsnet-GNN

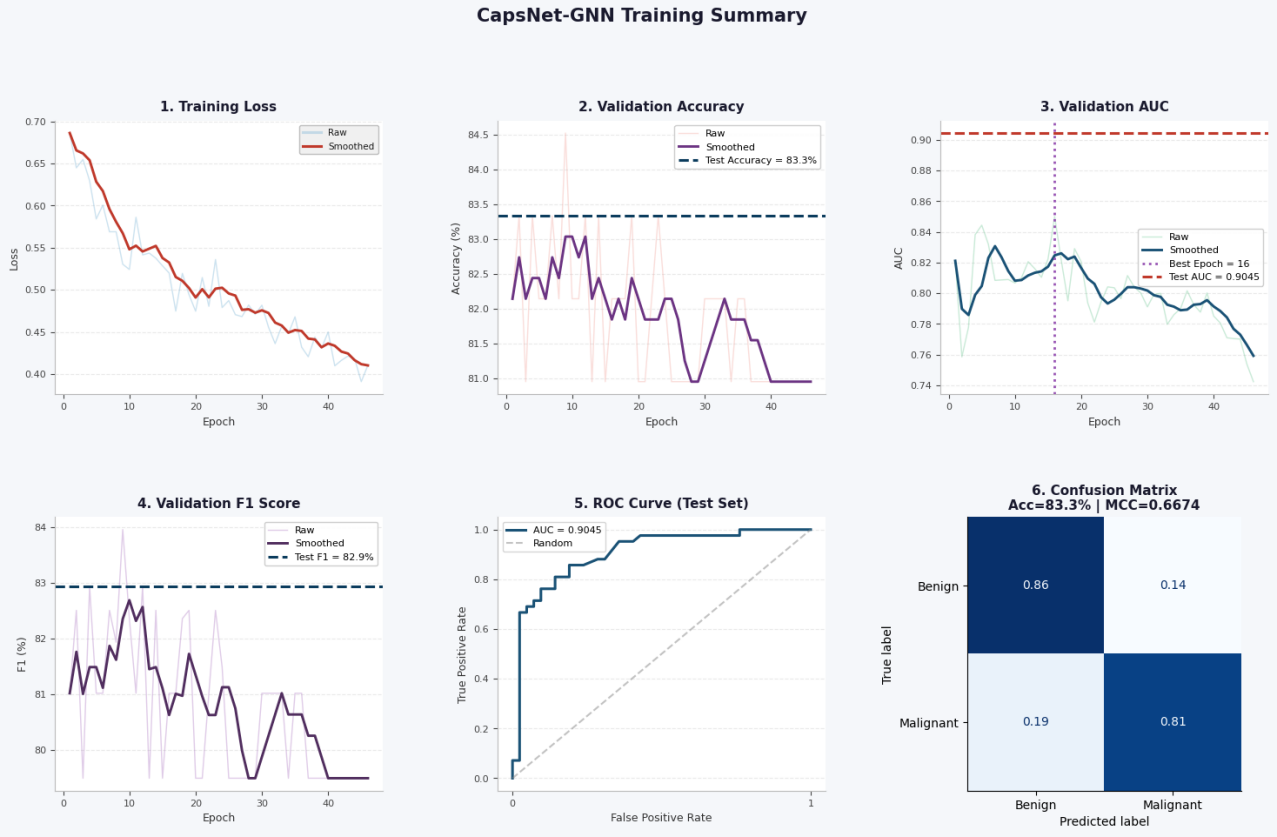

Fig. 4: Training and Performance Evaluation of the Proposed Hybrid CapsNet–GNN Model

Fig. 4 describes the training behavior and performance of the proposed CapsNet-GNN framework. In this figure, six panels have been created. In the first panel, the training loss curve indicates a downward trend, starting at 0.68 and ending at around 0.41. This indicates good convergence behavior with minimal overfitting. In the second and third panels, validation accuracy and F1-score curves show some variation during training. However, they eventually reach and maintain values of 83.3% and 82.9%, respectively. This behavior matches the final test results. In the fourth panel, validation AUC achieves its maximum at epoch 16 with a maximum of 0.8501. In contrast, the test AUC achieves a maximum of 0.9045. This indicates good discriminative ability. In the fifth panel, the ROC curve indicates good classification performance. This indicates that the model works well in classification. In addition, the model achieves better performance compared to the random baseline. In the sixth panel, the confusion matrix indicates good performance. In this case, 86% of the benign and 81% of the malignant nodules have been correctly classified. In addition, the MCC of this model is found to be 0.6674.

Per-Class Classification Analysis

Table 2: Per-Class Classification Performance of the Proposed CapsNet-GNN Model

Class	Precision	Recall	F1-Score
Benign	0.8182	0.8571	0.8372
Malignant	0.8500	0.8095	0.8293
Accuracy			**0.8333**
Macro Avg	0.8341	0.8333	0.8332
Weighted Avg	0.8341	0.8333	0.8332

Fig. 5 and table 2 shows a detailed performance analysis of the proposed CapsNet-GNN model using a set of four visualization windows in the class-wise performance metrics chart. The performance metrics indicate that, while the malignant class performs better in precision (0.85), the benign class performs better in recall (0.8571), reflecting the sensitivity of the model in recognizing true cases of benign nodules. The F1-Score comparison indicates that the proposed model performs equally well in classifying both types of nodules, as reflected in the F1-Scores of 0.8372 and 0.8293 for benign and malignant classes, respectively, and a macro average of 0.8332 for both classes combined. The precision-recall curve indicates a compromise in the performance of the proposed model, where higher precision in the malignant class is achieved at the cost of recall in the same class.

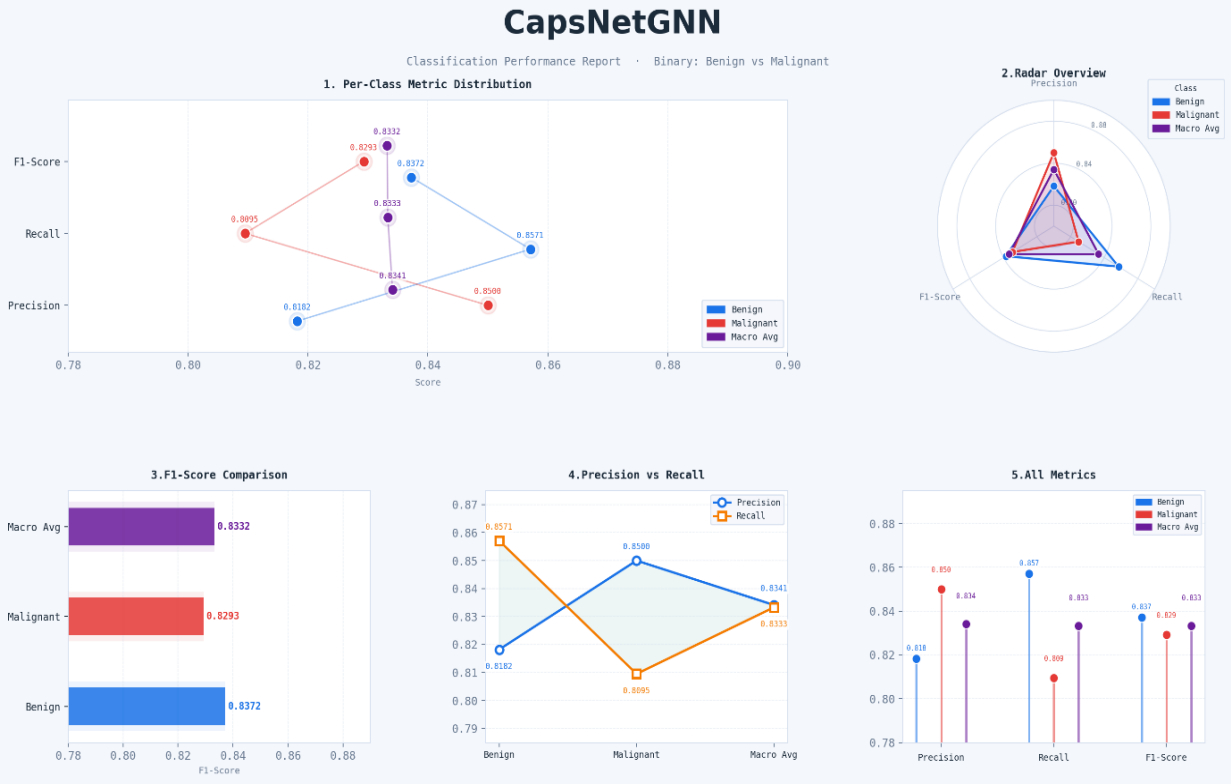

Fig. 5: CapsNet-GNN Per-Class Classification Performance Report

Tumor Localization Results

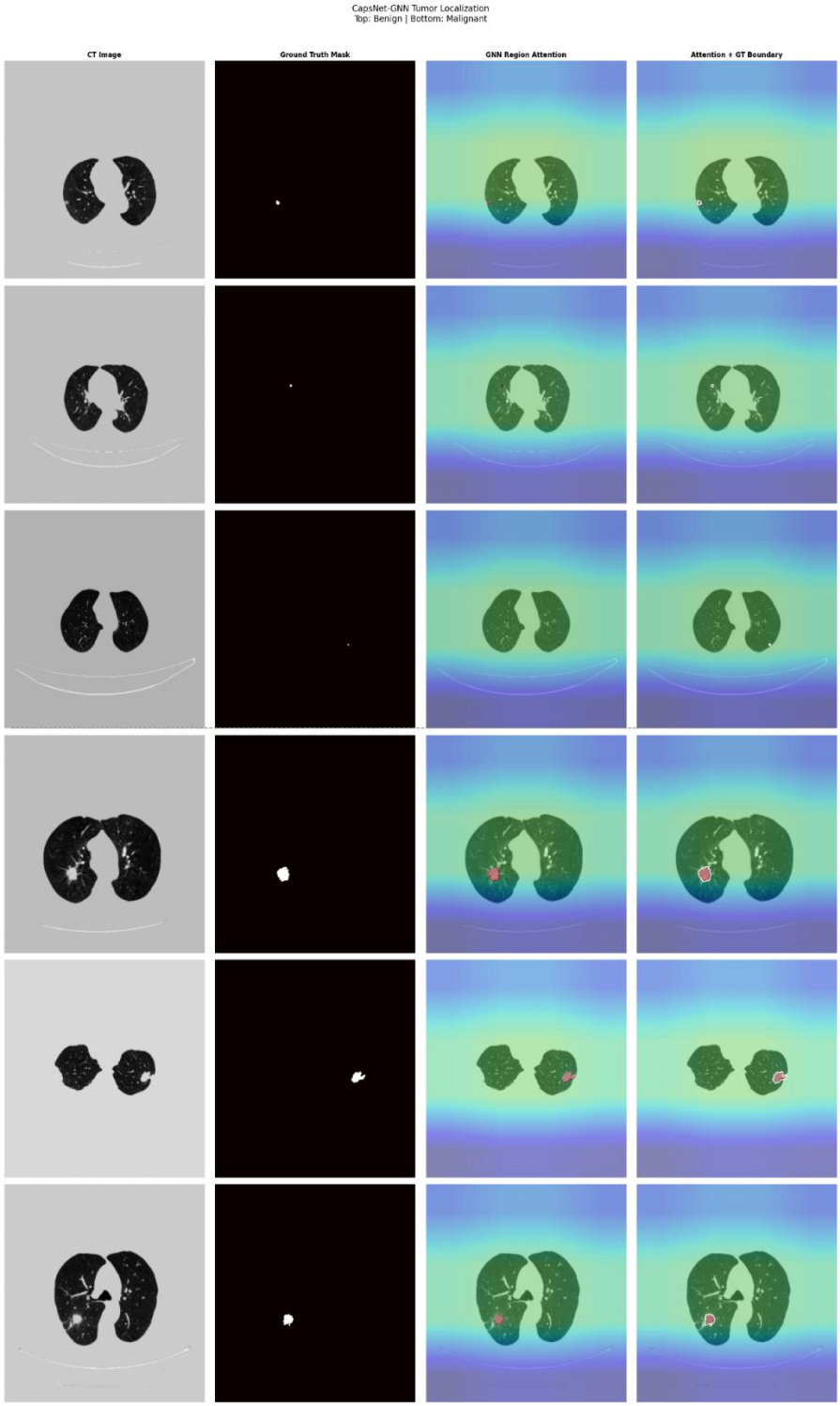

Fig. 6: CapsNet-GNN Tumor Localization: Benign (Top) vs. Malignant (Bottom) Cases

Fig. 6 presents the tumor localization outputs of the proposed CapsNet-GNN framework across six representative CT samples, with the top three rows corresponding to benign cases and the bottom three to malignant cases. Each row displays the original CT image, ground truth mask, GNN region-based localization map, and the localization result overlaid with ground truth boundary contours. For benign cases, the ground truth masks reveal minimal or near-absent

nodule regions, and the GNN localization maps distribute activation broadly across the lung field, reflecting the absence of a focused tumor region. In contrast, malignant cases exhibit clearly visible nodule masks, and the GNN localization maps concentrate high-intensity responses precisely around the tumor regions, closely aligning with the ground truth boundaries visible in the fourth column. This spatial correspondence between predicted localization and ground truth confirms that the relational modeling of the GNN effectively identifies and delineates diagnostically relevant tumor regions, demonstrating the framework capability for clinically interpretable and spatially precise tumor localization in lung CT scans.

Comparative Analysis with Existing Methods

Table 3: Comparative Performance of CapsNet-GNN Against Baseline Classifiers

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Decision Tree (Vinay et al., 2011)	75.9	67.8	75.9	69.0
Voting Classifier (Vinay et al., 2011)	79.0	67.9	75.5	69.0
AdaBoost (Vinay et al., 2011)	83.0	59.8	77.4	67.5
Stacking (Vinay et al., 2011)	81.0	61.3	75.5	66.0
CapsNet-GNN	**83.33**	**85.00**	**80.95**	**82.93**

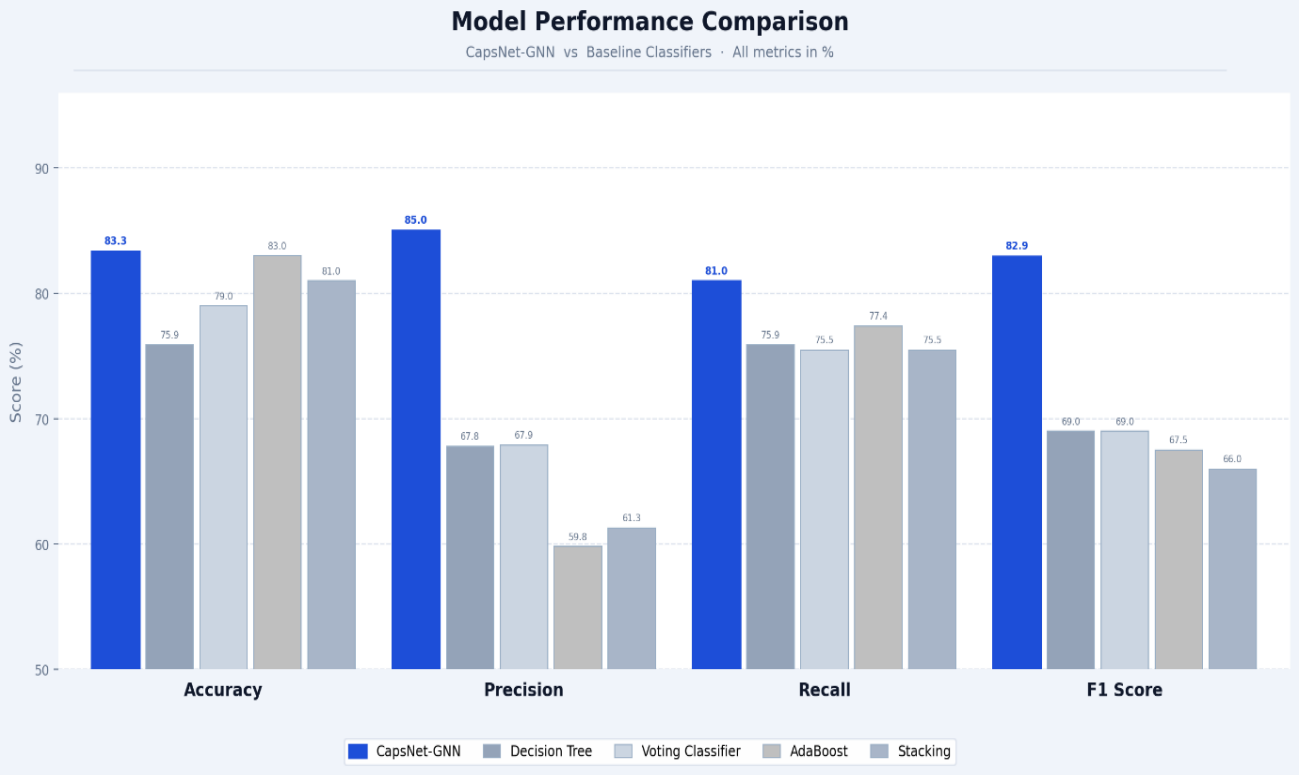

Fig. 7: Comparative Model Performance Analysis: CapsNet-GNN vs. Baseline Classifiers

Fig. 7 and table 3 shows a grouped bar chart that compares the proposed CapsNet-GNN model with four different baseline classifiers on the basis of accuracy, precision, recall, and F1 score. From the chart, it can be concluded that the proposed model has outperformed all the baseline classifiers on all the given parameters. Most importantly, the proposed model has recorded a precision of 85.0%, compared to the relatively poor precision of AdaBoost (59.8%) and Stacking (61.3%). Although the AdaBoost classifier has recorded almost the same accuracy as the proposed model, i.e., 83.0%, the F1 score of the AdaBoost classifier, i.e., 67.5%, indicates that the precision and recall values of the AdaBoost classifier are not well balanced, unlike the proposed model, which has recorded an F1 score of 82.93%.

Ablation Study and Model Analysis

Table 4: Ablation Study - Component Progression of the CapsNet-GNN Framework

Metric	TinyCNN	CapsNet	CapsNet-GNN
Accuracy	55.9524	60.7143	83.33
F1 Score	44.7761	52.1739	82.93
AUC	0.4912	0.5859	0.9045
MCC	0.1302	0.2294	0.6674

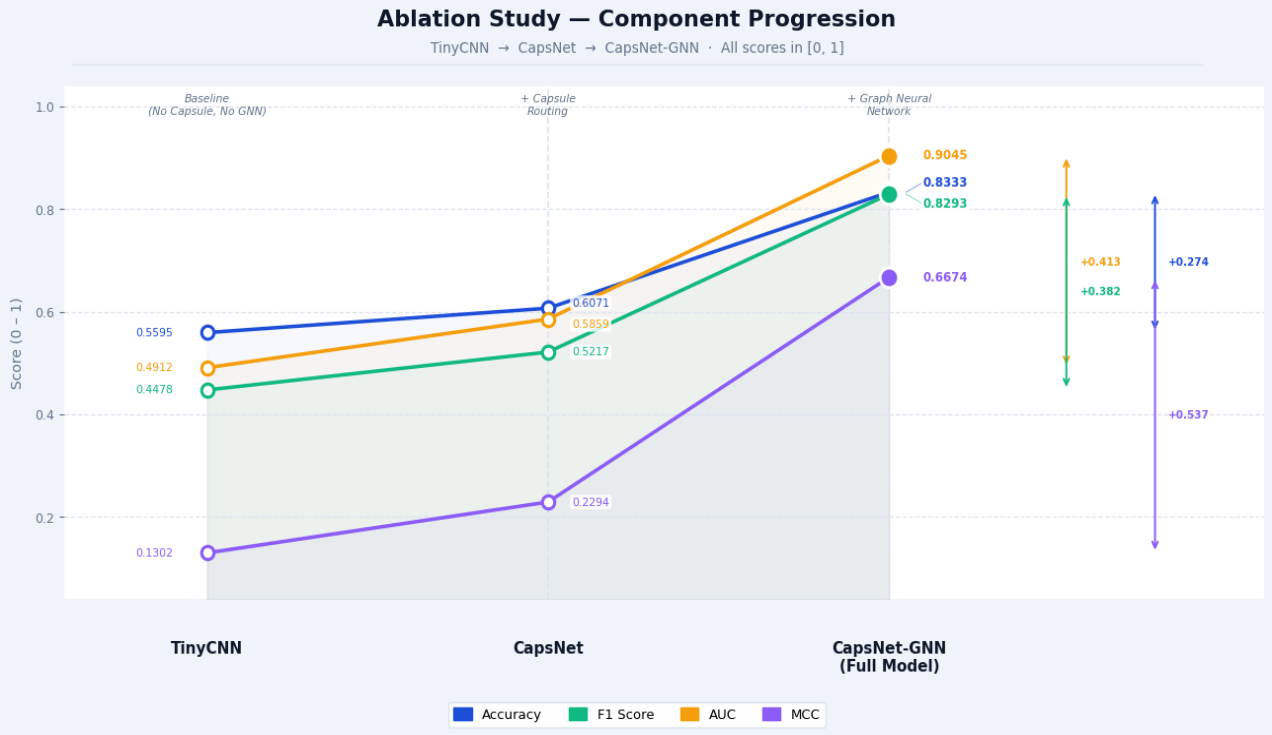

Fig. 8: Ablation Study - Component Progression from Tiny CNN to CapsNet-GNN

Fig. 8 and table 4 demonstrates the performance improvement obtained by gradually adding the capsule routing and GNN components to the baseline model. As expected, the baseline model, Tiny CNN, shows a dismal performance in all evaluation parameters, with an AUC of merely 0.4912 and an MCC of 0.1302, indicating a nearly random performance. Adding the capsule routing component to the model, as represented by the CapsNet model, shows a moderate improvement in performance, with an accuracy of 60.71% and an AUC of 0.5859, confirming the effectiveness of spatial and hierarchical feature encoding. Significant improvements are observed after adding the GNN component to the model, with a performance improvement of +22.62% in accuracy, +0.413 in AUC, +0.382 in F1-score, and +0.537 in MCC, leading to the best-performing model, CapsNet-GNN, with an accuracy of 83.33% and an AUC of 0.9045, confirming that the components are individually essential and, together, indispensable for effective lung nodule classification.

Discussion

The results obtained from the experiment have validated that the hybrid model based on the CapsNet-GNN architecture is better than traditional baseline classifiers in the identification and localization of lung cancer. The hybrid model based on Graph Neural Networks and Capsule Networks is very effective in understanding the formation of complex tumor formation, as validated through the classification accuracy of 83.33%, F1 score of 82.93%, AUC score of 0.9045, and MCC score of 0.6674. Ablation studies have also validated the importance of the model components: the importance of Capsule Networks in improving spatial feature representation and the importance of Graph Networks in improving relational learning.

The findings also emphasize the essence of capturing local and global spatial information. Although Capsule Networks do not greatly harm part-whole relations in lung nodules, Graph Neural Networks augment the contextual knowledge of modeling inter-nodule relations. This two-fold solution is used to counter a major weakness of the classical CNNs that tend to ignore spatial hierarchies and do not highlight interactions between two or more Tumor regions. The hybrid framework therefore minimizes the cases that are misclassified which are the difficult cases like the small, irregular or adjacent Tumors and also ensures uniform performance of the hybrid framework on the various Tumor sizes, shapes and good quality CT scans.

The results obtained from the experiment have validated that the hybrid model based on the CapsNet-GNN architecture is better than traditional baseline classifiers in the identification and localization of lung cancer. The hybrid model based on Graph Neural Networks and Capsule Networks is very effective in understanding the formation of complex tumor formation, as validated through the classification accuracy of 83.33%, F1 score of 82.93%, AUC score of 0.9045, and MCC score of 0.6674. Ablation studies have also validated the importance of the model components: the importance of Capsule Networks in improving spatial feature representation and the importance of Graph Networks in improving relational learning.

Altogether, the suggested Caps NetGNN framework is credible, interpretable, and generalizable in lung cancer identification and localization. The hierarchical feature extraction combined with relational modeling makes the model overcame the restrictions of the current deep learning models and show high possibilities of integration into clinical decision-support systems. These results formed a basis of future studies, such as expansion to multi-modal imaging, combining with the features of radiomics, and real-time implementation in clinical practice.

Conclusion and Future Work

In this paper, a new hybrid model of Capsule Network-Graph Neural Network, named CapsNet-GNN, is proposed that can effectively classify and localize lung cancer. The performance of the suggested model is proved through the experimental results, which show that the model outperforms other classifiers such as Decision Tree, Voting Classifier, AdaBoost, and Stacking classifiers in all performance metrics, with an accuracy of 83.33%, precision of 85.00%, recall of 80.95%, F1-score of 82.93%, AUC of 0.9045, and a value of 0.6674 for the Matthews Correlation Coefficient. This model is effective in classifying and localizing the lung cancer due to the effective feature extraction of the CapsNet model and the utilization of the GNN model to capture the relational dependencies. In addition, the performance of the model is consistent in all the experiments, which proves that the model is effective in classifying and localizing the lung cancer, irrespective of the size, shape, and image quality of the CT image.

The suggested framework demonstrates considerable advancements, multiple additions can contribute to its better functioning and usage. First, multi-modal imaging data including PET-CT and MRI should be incorporated to offer more specific information on early-stage Tumor and enhance the generalization of the model. Second, the use of radiomics qualities and deep learning representations could be employed to improve interpretability and explainable AI of clinical decision support. Third, the model can be incorporated into automated clinical processes in real-time through the deployment of optimally efficient computational-wise implementation. Lastly, the multi-class Tumor classification, metastasis detection and other thoracic diseases could be extended to an area of more clinical use. In future studies the proposed hybrid CapNet-GNN framework will be developed into future research to investigate how PET scans can be incorporated into CT scans. PET-CT fusion will also enable the model to take advantage of both structural and functional data so as to better detect tumors at an early stage and better determine the behavior of the tumor. The multi-modal imaging will also aid in the localization difficulties of tumors as well as their classification and metastasis. In addition, the addition of radiomics characteristics and dynamic processing of PET-CT data will increase understanding and clinical usability of the system, which may promote automated diagnostic systems in clinics.

References

Escarela, G., Rodríguez, C. E., & Núñez-Antonio, G. (2020). Copula modeling of receiver operating characteristic and predictiveness curves. *Statistics in Medicine*, *39*(28), 4252-4266. https://doi.org/10.1002/sim.8723

Bala, N. M., & Kannan, K. S. (2025). Hybrid Deep Learning Model Based Lung Cancer Prediction and Classification with OTSU Segmentation Method. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, *16*(2), 75-94. https://doi.org/10.58346/JOWUA.2025.I2.005

Liu, Y., Yao, S., Wang, X., Chen, J., & Li, X. (2024). MD-UNet: a medical image segmentation network based on mixed depthwise convolution. *Medical & Biological Engineering & Computing*, *62*(4), 1201-1212. https://doi.org/10.1007/s11517-023-03005-8

Asiya., & Sugitha, N. (2025). Ensuring Data Integrity and Security in AI-Based Lung Cancer Nodule Classification and Detection in CT Imaging Systems. *Journal of Internet Services and Information Security, 15*(1), 361-370. https://doi.org/10.58346/JISIS.2025.I1.023

Yang, D., Zhao, H., Yu, K., & Geng, L. (2023). NAUNet: lightweight retinal vessel segmentation network with nested connections and efficient attention. *Multimedia Tools and Applications*, *82*(16), 25357-25379. https://doi.org/10.1007/s11042-022-14319-4

Alzaidi, E. R. (2024). Optimization of deep learning models to predict lung cancer using chest X-ray images. *International Academic Journal of Science and Engineering*, *11*(1), 351-361. https://doi.org/10.9756/IAJSE/V11I1/IAJSE1140

Jiang, Y., Dong, J., Cheng, T., Zhang, Y., Lin, X., & Liang, J. (2023). iU-Net: a hybrid structured network with a novel feature fusion approach for medical image segmentation. *BioData Mining*, *16*(1), 5. https://doi.org/10.1186/s13040-023-00320-6

Nandhinieswari, S., & Indumathi, A. (2024). Bilevel optimized recursive feature eliminator for cervical cancer feature selection process. *Archives for Technical Sciences, 2*(31), 311–328. https://doi.org/10.70102/afts.2024.1631.311

Hussain, T., Shouno, H., Hussain, A., Hussain, D., Ismail, M., Mir, T. H., ... & Akhy, S. A. (2025). EFFResNet-ViT: A fusion-based convolutional and vision transformer model for explainable medical image classification. *IEEE Access, 13,* 54040-54068. https://doi.org/10.1109/ACCESS.2025.3554184

Alam, T., Yeh, W. C., Hsu, F. R., Wei-Chung, S., Singh, A. R., Hassan, T., ... & Hussain, T. (2024). An integrated approach using yolov8 and resnet, seresnet & vision transformer (vit) algorithms based on roi fracture prediction in x-ray images of the elbow. *Current Medical Imaging*, *20*(1), e15734056309890. https://doi.org/10.2174/0115734056309890240912054616

Lau, S. L., Chong, E. K., Yang, X., & Wang, X. (2020). Automated pavement crack segmentation using u-net-based convolutional neural network. *IEEE Access*, *8*, 114892-114899. https://doi.org/10.1109/ACCESS.2020.3003638

Jalali, Y., Fateh, M., Rezvani, M., Abolghasemi, V., & Anisi, M. H. (2021). ResBCDU-Net: A deep learning framework for lung CT image segmentation. *Sensors*, *21*(1), 268. https://doi.org/10.3390/s21010268

Alblehai, F., Mahmoud, M. H., Akhtar, M. S., Shaker, A., & Mohamed, A. A. (2024). Analysing biosensor clinical pathogen information using mayfly optimized convolute neural network approach. *Expert Systems*, *41*(6), e13027. https://doi.org/10.1111/exsy.13027

Yuan, C., Song, S., Yang, J., Sun, Y., Yang, B., & Xu, L. (2023). Pulmonary arteries segmentation from CT images using PA-Net with attention module and contour loss. *Medical Physics*, *50*(8), 4887-4898. https://doi.org/10.1002/mp.16265

Padmanaban, H., Rajarajan, G., & Nagarajan, S. (2024). Xception Taylor Cascade Neuro Network based infection level identification of tuberculosis using sputum images. *Intelligent Decision Technologies*, *18*(2), 799-824. https://doi.org/10.3233/IDT-240395

Faizi, M. K., Qiang, Y., Shagar, M. M. B., Wei, Y., Qiao, Y., Zhao, J., & Urrehman, Z. (2025). Graph neural network model using radiomics for lung CT image segmentation. *Scientific Reports*, *15*(1), 34148. https://doi.org/10.1038/s41598-025-12141-0

Li, M. M., Huang, K., & Zitnik, M. (2022). Graph representation learning in biomedicine and healthcare. *Nature biomedical engineering*, *6*(12), 1353-1369. https://doi.org/10.1038/s41551-022-00942-x

Stanciu, I. M., Parosanu, A. I., Orlov-Slavu, C., Iaciu, I. C., Popa, A. M., Olaru, C. M., ... & Nitipir, C. (2023). Mechanisms of resistance to CDK4/6 inhibitors and predictive biomarkers of response in HR+/HER2-metastatic breast cancer—a review of the literature. *Diagnostics*, *13*(5), 987. https://doi.org/10.3390/diagnostics13050987

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, *42*, 60-88. https://doi.org/10.1016/j.media.2017.07.005

Yan, C., Hu, J., & Zhang, C. (2018). Deep transformer: A framework for 2D text image rectification from planar transformations. *Neurocomputing*, *289*, 32-43. https://doi.org/10.1016/j.neucom.2018.02.015

Li, R., Xiao, C., Huang, Y., Hassan, H., & Huang, B. (2022). Deep learning applications in computed tomography images for pulmonary nodule detection and diagnosis: A review. *Diagnostics*, *12*(2), 298. https://doi.org/10.3390/diagnostics12020298

Abdullahi, K., Ramakrishnan, K., & Ali, A. B. (2025). Deep learning techniques for lung cancer diagnosis with computed tomography imaging: A systematic review for detection, segmentation, and classification. *Information*, *16*(6), 451. https://doi.org/10.3390/info16060451

Gogoshin, G., & Rodin, A. S. (2023). Graph neural networks in cancer and oncology research: emerging and future trends. *Cancers*, *15*(24), 5858. https://doi.org/10.3390/cancers15245858

Veyrin-Forrer, L., Kamal, A., Duffner, S., Plantevit, M., & Robardet, C. (2024). On GNN explainability with activation rules. *Data Mining and Knowledge Discovery*, *38*(5), 3227-3261. https://doi.org/10.1007/s10618-022-00870-z

Kumar, A., Agrawal, R. K., & Joseph, L. (2023). IterMiUnet: A lightweight architecture for automatic blood vessel segmentation. *Multimedia Tools and Applications*, *82*(28), 43207-43231. https://doi.org/10.1007/s11042-023-15433-7

Aravind, R. V., & Balasubramaniam, P. (2022). Dwell-time-dependent memory based state feedback controller design for switched Takagi–Sugeno fuzzy nonlinear systems. *Information Sciences*, *609*, 838-854. https://doi.org/10.1016/j.ins.2022.07.099

Hudec, M., Vučetić, M., & Barčáková, N. (2024). An extended proximity relation and quantified aggregation for designing robust fuzzy query engine. *Knowledge-Based Systems*, *290*, 111574. https://doi.org/10.1016/j.knosys.2024.111574

Nam, D., Panina, A., & Pak, A. (2024). Lung cancer segmentation dataset with Lung-RADS class. *Mendeley Data*, *1*. https://doi.org/10.17632/5rr22hgzwr.1

Vinay, K., Rao, A., & Kumar, G. H. (2011, December). Computerized analysis of classification of lung nodules and comparison between homogeneous and heterogeneous ensemble of classifier model. In *2011 third national conference on computer vision, pattern recognition, image processing and graphics* (pp. 231-234). IEEE. https://doi.org/10.1109/NCVPRIPG.2011.56