

Quantitative chest computed tomography-based hybrid deep learning for chronic lung disease detection and staging

Garima Jain, Amit Kumar Goel, Rahul Kumar

Department of Computing and Informatics, School of Engineering and Sciences, Sir Padampat Singhanian University, Udaipur, India

Article Info

Article history:

Received Mar 20, 2025

Revised Jul 18, 2026

Accepted Aug 12, 2026

Keywords:

Chest computed tomography
Explainable artificial intelligence
Extreme gradient boosting
Feature-level fusion
Gradient-weighted class activation mapping
Radiomics
Shapley additive explanations

ABSTRACT

This research proposes a hybrid deep learning model to accurately detect chronic lung diseases in chest computed tomography (CT) scans. Current deep learning methods, such as convolutional neural networks (CNNs) and transformer models, often suffer from high computational requirements and low interpretability. This research combines a pre-trained residual network 50 layers (ResNet50) for deep feature extraction with hand-crafted texture features. The features are combined via hybrid concatenation, and the model is classified using an optimized extreme gradient boosting (XGBoost) ensemble with hyperparameter tuning to improve performance and reduce overfitting. This research compared the proposed model with radiomics-only, ResNet50 CNN, vision transformer (ViT-B16), and Swin transformer methods. This research used a public dataset containing adenocarcinoma, squamous cell carcinoma, large cell carcinoma, and normal classes. The classification accuracy of the hybrid framework was 92.38%, with better precision, sensitivity, specificity, F1-score, and receiver operating characteristic-area under the curve (ROC-AUC) than the baselines. A dual-mode explainable artificial intelligence (XAI) framework integrating gradient-weighted class activation mapping (Grad-CAM) and Shapley additive explanation (SHAP) was implemented to enable spatial lesion localization and feature attribution for clinical transparency. Grad-CAM was able to localize the lesions successfully, and SHAP revealed the influence of entropy-based texture features and semantic embeddings on the predictions. Overall, the framework develops advanced XAI systems for chest CT analysis.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Garima Jain

Department of Computing and Informatics, School of Engineering and Sciences

Sir Padampat Singhanian University

Bhatewar, Udaipur - 313601, India

Email: garima.phd2023@spsu.ac.in

1. INTRODUCTION

Lung cancer is still one of the main causes of cancer-related deaths around the globe, and its public health impact remains significant, despite advancements made in medical imaging and treatment methods [1], [2]. Accurate and early detection of lung nodules from computed tomography (CT) images is critical for prognosis, timely intervention, mitigating disease progression, and personalizing treatment; however, due to overlapping visual characteristics, subtle differences in texture, and time from the first diagnosis, it can be difficult for clinicians to accurately distinguish between certain types of lung tumors such as adenocarcinomas, squamous cell carcinomas, and large cell carcinomas [3]. With the growing use of artificial intelligence (AI)-based computer-assisted diagnosis (CAD) systems in clinical practice, recent

studies have shown that CAD systems may be effective tools to assist radiologists in the automated classification of lung disease by way of AI. Several deep learning-based methods, including convolutional neural networks (CNNs), have been used successfully for the extraction of hierarchical spatial representations from CT images [4], [5]. Various CNN architectures i.e., residual network (ResNet), densely connected convolutional network (DenseNet), efficient network (EfficientNet), and mobile network (MobileNet), have accelerated the learning of automated features in the analysis of thoracic images [6]–[8]. More recent interest has also been generated with respect to transformer-based vision models i.e., vision transformers (ViT) and Swin transformers as these models afford the capability of learning long-range contextual dependencies and globally modeling images [9], [10]. Even though these architectures have delivered promising results in terms of their classification accuracies, there remains an abundance of current systems that continue to be computationally demanding, are not clinically interpretable, and are dependent on large numbers of training data [11]. At the same time, radiomics-based methods have also started to become recognized as an important and complementary way of representing quantitative data within the context of medical imaging. Radiomics provides methods of extracting hand-crafted descriptors from CT scans that describe properties, such as texture heterogeneity, intensity distributions, shapes, and structural shapes of tissues. These quantitative biomarkers are often used to define layers of tissue that have clinical relevance but cannot be seen when developing domain-specific or deep latent representations. In addition to the limitations of conventional radiomics pipelines, independent deep learning models tend to act as “black boxes” limiting clinicians' ability to interpret them and their trust in the output accuracy.

Numerous studies that have integrated deep learning with radiomics to assess pulmonary disease have used various techniques to provide insights into pulmonary diseases; however, there are still some critical limitations in these attempts. An overwhelming number of existing frameworks incur high computational costs associated with their multi-stage architectures, feature optimization procedures, and transformer-based architectures that may not make them practical for routine clinical use [12]–[14]. Many of the existing studies on computer-based analysis have focused primarily on binary classification tasks. There is also limited research available that analyzes the clinical aspects of multi-class disease differentiation with respect to false negative analysis, early-stage lesion sensitivity, computational feasibility, explainability validation, and demographic robustness as well as how they can be deployed within a practical radiology setting.

To address these limitations, this study presents a hybrid deep learning framework in a quantitative aspect for multiple classification and detecting chronic lung disease through chest CT utilizing both deep spatial embedding and radiomics texture descriptors with the use of handcrafted statistical properties. To develop this proposed approach, a ResNet50 pre-trained backbone architecture will be used to extract high-dimensional spatial representations from chest CT images. Parallel to feature extraction performed by the ResNet50 network, the classification will employ a second branch of radiomics whereby the computed tissue heterogeneity will be described using gray-level co-occurrence matrix (GLCM) statistics and entropy-based analyses. These mutually complementary feature representations will then be fused together on a feature level before being classified using an optimized extreme gradient boosting (XGBoost) ensemble classifier.

The proposed methodology is distinct from that of many traditional deep learning approaches in that it features not only non-deterministic latent embeddings as input to the classifier but will include deterministic radiomic features directly via their inclusion in the overall feature set, thereby enhancing both discrimination capability and interpretability. The goal of hyperparameter optimization through systematic grid search analysis was achieved by being able to obtain model stability and robustness across various experimental configurations. In addition to creating baselines with only CNNs, we benchmarked against transformer-like architectures such as the ViT and Swin transformer to measure how well they classify images and generalize beyond their training sets.

This work's additional contribution is the use of a dual-mode explainable artificial intelligence (XAI) framework that enhances clinical transparency and physician trust. This work employ gradient-weighted class activation mapping (Grad-CAM)-based visualization to locate the diagnostically important pulmonary regions and utilize Shapley additive explanations (SHAP) to quantify both deep embeddings and hand-crafted radiomic descriptors' contributions at the feature level [15]. The joint explainability mechanism allows for spatial and quantitative interpretations of the diagnostic process, thereby promoting transparency in the AI-assisted clinical decision-making process [16]. A multi-class chest CT dataset consisting of normal, adenocarcinoma, squamous cell carcinoma, and large cell carcinoma categories is assessed using stratified training-validation-testing splits, then evaluated using classification accuracy, precision, recall, F1-score, sensitivity, area under the curve (AUC) by receiver operating characteristic (ROC) curves, and confusion matrix analyses. Additionally, to evaluate the practicality of implementing this proposed model in future radiological screening systems, telemedicine systems, and intelligent clinical decision support systems, both computational efficiency and explainability analyses were performed.

This study establishes the following key contributions: i) a hybrid feature-level fusion framework that incorporates deep CNN spatial embeddings with traditional handcrafted radiomic texture features to facilitate the classification of CT images of the chest into multiple categories; ii) a comparative benchmarking with other deep learning models, such as the ViT, Swin transformer, and CNN such as ResNet, only on multivariate performance metrics; iii) hyperparameter optimization and computational efficiency analysis for improved reproducibility, robustness, and ease of deployment; iv) the integration of an XAI framework through both Grad-CAM and SHAP for spatial-level and feature-level interpretation of results; and v) a thorough evaluation of the proposed framework with respect to multiple classification metrics, ablation analysis, ROC-AUC analysis, and clinically interpretable explanations of results.

2. LITERATURE REVIEW

In this work by Thanoon *et al.* [17] deep learning-based methods for CT imaging of the chest for lung cancer screening and diagnosis are presented. The authors organize and compare various deep learning techniques including (but not limited to) CNN, recurrent neural networks (RNNs) as well as new hybrid-based architectures that have been proposed during the last half-decade (2018–2022) by grouping different methods on classification/segmentation. This review highlights the potential of deep learning for addressing diverse types of inputs (2D, 3D, and low-dose CT) but also critically evaluates problems from data diversity, image artifacts to lack of labeled datasets. Moreover, the paper also identifies potential research areas such as standardized pre-processing and integration of clinical data to enhance automated tumor segmentation.

Ozdemir *et al.* [18] proposed a hybrid deep learning model for lung cancer detection using CNNs, based on InceptionNeXt blocks and ViT, using grid and block attention methods. This model is designed to capture local detailed information and overall contextual data from CT scans of the lungs so it can be used in a clinical setting. The model was trained using two different public datasets the Iraq-Oncology Teaching Hospital/National Center for Cancer Diseases (IQ-OTH/NCCD) dataset and the chest CT dataset. The model was tested against both datasets and achieved an accuracy of 99.54% accurately and 98.41% accurately, respectively. It is important to note that the authors were able to achieve this state-of-the-art result with an efficient and lightweight model design of only 18.1 million parameters. In addition to creating an effective classifier, the authors used Grad-CAM to create heatmaps of different areas of the lungs that were being emphasized or de-emphasized when the model was making its prediction.

Kumari and Goel [16], give an extensive review of computational intelligence techniques for diagnosing as well as classifying lung cancer within chest CT images. This study categorizes and compare different methods such as classical machine learning techniques (e.g. support vector machine (SVM), random forests (RF)), deep learning families like CNNs, RNNs, or variational autoencoders (VAEs) generative adversarial networks (GANs) as well as hybrid algorithms that combine nature inspired optimization. The paper does well to focus on growing trends like transfer learning, federated learning for privacy, and XAI to address the “black-box” limitations of models. It also systematically studies the performance based on large public databases of the lung image database consortium image database resource initiative (LIDC-IDRI) and the lung nodule analysis 2016 (LUNA16), and demonstrates the urgent challenges in group bias and class imbalance that hinder generalizing AI solutions to broad clinical contexts.

This study by Sun *et al.* [19] introduces a weakly supervised deep learning algorithm for the automated detection and staging progression of chronic obstructive pulmonary disease (COPD) from volumetric CT scans. To address limitations in traditional densitometric analysis, where noise-sensitive pixel thresholds and time-consuming manual segmentations are used, they present a multiple instance learning (MIL) approach. This architecture uses a 3D ResNet model for CT slice encoding and an attention mechanism for pooling, being trained at patient-level (GOLD stage from spirometry) instead of dense pixel-level annotations. The model demonstrates excellent discrimination for COPD stages (GOLD 1,2,3, and 4) and provides clinical interpretability through the generation of attention heatmaps which highlight disease-related pathologies such as emphysema and bronchiectasis. Segmentation, computer-aided detection, and computer-aided diagnosis (Seg-CADe-CADx) was developed by Subramanyam *et al.* [20] as a dual-stage deep learning system that can detect pulmonary nodule malignancy risk as an integrated process from low-dose CT scans. The first stage uses a 3D U-Net to assist with segmentation-based nodule detection, which is followed by a multi-stream hybrid classifier consisting of local volumetric features from a 3D DenseNet, global context from an Swin transformer, and late-stage concatenation of handcrafted radiomic features from the volumetric images. Results from both the LUNA16 database and the LIDC-IDRI database provided a competition performance metric (CPM) for nodule detection of 0.944 with an excellent ROC-AUC of 0.988 for binary classification of malignancy. Additionally, the paper discusses improving the clinical accuracy of the model by introducing temperature scaling for the calibration of malignancy probability outputs, thereby providing more closely related empirical risk estimates than previously available.

Sui *et al.* [21] proposed a machine learning method to diagnose and grade the severity of COPD in patients from CT scans. Unlike a large proportion of previous works 1–22 that simply focus on isolated features, it considers multidimensional characteristics including the lung parenchyma, whole airway tree, and pulmonary vasculature-related ones. The paper utilized the SVM classifier for 349 subjects and achieved better diagnostic performance (AUC = 0.965 on external testing) and grading, successfully distinguishing CVD from HC (AUC = 0.813). The novel and interesting finding in the study was the importance of vascular derivatives on COPD + as opposed to the currently highlighted contribution of remodeling in the trunk of arteries and bronchus. However, our study is limited by the small sample size and an overly simplified dichotomization (mild versus severe) that may not fully account for the clinical heterogeneity of the disease.

Alsurrayhi [22] has proposed and validated models based on deep learning for automatic staging of COPD with chest CT. The study overcomes the drawbacks of conventional pulmonary function tests (PFTs) by introducing image-based classification systems that are consistent with not only the high-resolution eight-stage system but also the new GOLD 2023 criteria. The method incorporates neural networks and 3D CNNs with pre-defined features extracted from paired inspiration/expiratory CT scans (e.g., one-shot transfer-based features and GLCM-based texture patterns) as well as end-to-end deep CNN learning. A further key development is in the generation of 'hybrid' models which combine imaging features with clinical information, such as symptom severity and exacerbation risk, resulting in a significant gain in classification performance. Here, this study shows that these deep learning models can accurately predict COPD severity factors directly from CT images, which could help diminish the requirement of patient-dependent spirometry measurements [23], [24].

Table 1 summarizes the major studies applying advanced deep learning and hybrid AI models for detection, classification, and severity assessment of lung diseases from chest CT images. Table 1 provides authors, main methodological approach, dataset names, sample sizes, core contributions, performance highlights, and research gaps for each reference. This comparative summary highlights the current status and remaining challenges in medical imaging AI, including interpretability, multi-class differentiation, and cross-dataset validation.

Table 1. Literature review

Author(s)	Architecture	Dataset name	Sample size	Main contribution	Outcome	Research gaps
Ozdemir <i>et al.</i> [18]	Attention-enhanced InceptionNeXt-based hybrid deep learning	Private lung CT dataset	2,500 images	Hybrid model using attention and InceptionNeXt for lung cancer detection.	High accuracy and robustness in detection	Limited interpretability; not evaluated on public or multi-class datasets
Kamala and Mohan [25]	Hybrid AI framework	Public and private CT datasets	1,800 images	Efficient hybrid AI with deep and handcrafted features for lung cancer classification	Enhanced classification accuracy and efficiency	Limited explainability; lacks external validation
Hosny <i>et al.</i> [26]	Multi-modal attention-based Inception-ResNet	LIDC-IDRI	1,000 cases	Multi-modal attention-based fusion for lung cancer detection	Outperformed unimodal models	Lacks interpretability and computational efficiency analysis
Parida <i>et al.</i> [27]	3D FusionNet	Synthetic CT Dataset	600 scans	3D fusion model for lung cancer segmentation	High segmentation precision, better boundaries	Not validated on real CT data; lacks clinical outcome correlation
Duhayyim <i>et al.</i> [28]	Interpretable hybrid ensemble with attention fusion + EAOG-genetic algorithm (GA) optimization	Private CT dataset	1,200	Attention-based fusion and genetic optimization for interpretability and accuracy	State-of-the-art accuracy and interpretability	No multi-class or external dataset validation
Faizi <i>et al.</i> [29]	CNN-based deep learning	LUNA16	888 CT scans	Deep learning for CT-based lung cancer	Reliable and robust classification	Binary classification only; lacks explainability and cross-dataset evaluation
Almeida <i>et al.</i> [30]	Anomaly-based deep learning for disease severity	COPDGen e, NLST	4,500 scans	Deep learning for quantitative COPD severity prediction	Accurate disease severity assessment	Not multi-disease focused; limited interpretability assessment

2.1. Research gap summary and motivation

Despite the major advancements seen in recent years, including improvements in deep learning and transformer-based architectures, radiomic analysis techniques, and XAI for pulmonary disease classification from chest CT scans, there are still numerous significant drawbacks across many existing publications. In general, most of the literature tends towards binary classifications while neglecting fine-grained multi-class staging. Additionally, traditional CNNs typically provide low levels of feature-level interpretability while transformer-inspired architectures often necessitate large amounts of data to achieve stable optimization.

Furthermore, several radiomics-based approaches generate clinically interpretable quantitative outputs, but most do little to capture the multi-level spatial hierarchies present in thoracic CT imaging. In addition, the majority of hybrid frameworks are highly reliant on computationally intensive fusion methods or offer only limited levels of explainability, limiting their applicability in large-scale, real-world clinical workflows. Finally, there is still limited external validation, low transparency across all predictions, and limited consideration of clinical deployment, which continue to inhibit the overall reliability and adoption of currently available automated systems for diagnosing lung disease.

To overcome limitations in existing works, present a hybrid deep learning framework based on quantitative chest CT, which employs a pre-trained ResNet50 backbone at its core to produce deep spatial embeddings, combined with texture features derived from handcrafted GLCM texture analyses and entropic measures that characterize heterogeneity. Additionally included in the proposed framework is an optimized ensemble classifier implemented via the XGBoost algorithm, as well as two forms of XAI through Grad-CAM and SHAP to promote enhanced spatial and feature-level explainability. The proposed methodology was benchmarked against transformer-based architectures. ViT-B16 and Swin transformer architectures for classification performance, computational efficiency, and clinical interpretability. The motivation of this research is therefore to develop a computationally efficient multi-class chest CT classification framework with clinical interpretability and resiliency applicable for future use in AI-assisted radiology and intelligent lung disease screening applications.

3. PROPOSED FRAMEWORK

In this study, a new hybrid deep learning model is proposed, which integrates the abstract representation architecture of a deep CNN and handcrafted radiomics to overcome the limitations of utilizing only abstract structure from a CNN or handcrafted radiomics for categorizing chest CT. The primary methodological contribution is a feature fusion technique at the level of features that combines high-level spatial embeddings from a pre-trained convolutional network and radiomic hand-crafted quantitative texture descriptors that are also clinically interpretable. The model combines these separate feature modalities to enhance classification performance across multi-class stages of lung diseases, as well as offering robust model interpretability by means of dual XAI approaches. Figure 1 presents the proposed dual-stream scheme. Input CT data is concomitantly analyzed by deep learning and radiomic pipelines. Several of these features are then combined using a fusion strategy before classification, so that the clinical predictions have both spatial (Grad-CAM) and feature-level (SHAP) interpretability.

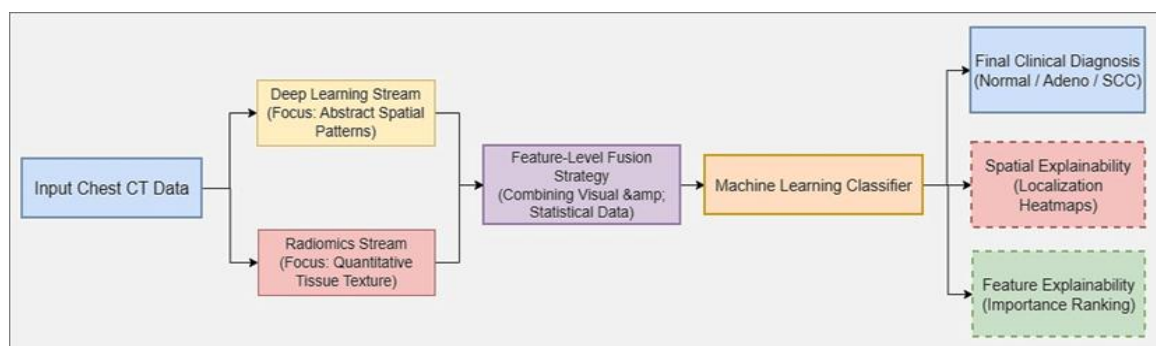


Figure 1. The proposed hybrid quantitative-spatial framework

3.1. Dataset

This publicly shared dataset contains approximately one thousand CT scans, which will be used to classify patients with lung cancer into their categories. The dataset consists of 1,000 chest CT scan images

with a total file size of approximately 124 MB. The images are provided in high-resolution .jpg and .png formats, ensuring visual clarity suitable for detailed analysis. All scans are derived from CT, making the dataset well-suited for research and development in automated lung disease detection and classification. The dataset consists of four pathological classifications of patients with lung cancer, described as follows:

- Normal: it is lung tissue that demonstrates a healthy lung with no clear evidence of nodules or other lesions present [31].
- Adenocarcinoma: it is the most common subtype of non-small cell lung cancer (NSCLC) and arises from peripheral glandular cells [32].
- Squamous cell carcinoma: it is the second most common type of NSCLC and originates from the squamous cells lining the major airways [33].
- Large cell carcinoma: a type of NSCLC, are the most common group of lung cancers, and LCC accounts for a smaller percentage of these cases [34]–[36].

The dataset is systematically partitioned into training, validation, and test sets to evaluate the model's generalization performance. Training set: a total of 613 images divided into adenocarcinoma (195), large cell carcinoma (115), squamous cell carcinoma (155), and normal (148). Validation set: a total of 72 images are grouped as adenocarcinoma (23), large cell carcinoma (21), squamous cell carcinoma (15), and normal (13). Testing set: a total of 315 images are grouped as adenocarcinoma (120), large cell carcinoma (51), squamous cell carcinoma (90), and normal (54).

Table 2 summarizes the main feature categories, names, and descriptions used in the chest CT image classification pipeline. Key features are: basic image metadata, pixel-level statistics, texture descriptors (e.g., GLCM-based metrics and entropy), structural and shape attributes, histogram and frequency domain information, and deep learning-based embeddings. Augmentation attributes and dataset split assignments are also provided for reproducibility and model validation purposes. These features together enable a holistic and multi-faceted analysis of chest CT images for robust classification and interpretability.

Table 2. Overview of extracted feature categories

Feature category	Feature name	Description	Type/use
Basic image metadata	Image filename	File name (identifier) of the CT image	Identifier
	Folder/label	Ground-truth class label: adenocarcinoma, large cell carcinoma, squamous cell carcinoma, normal	Target/label
	Resolution	Original image width × height (pixels)	Numeric (image property)
Pixel-level features	Format	File type (e.g., JPG, PNG)	Categorical
	Mean pixel intensity	Average gray value across all pixels	Numeric (image statistic)
	Standard deviation	Pixel intensity spread	Numeric (image statistic)
Texture features	Skewness and kurtosis	Distribution shape of pixel intensities	Numeric (image statistic)
	GLCM contrast	Local contrast computed from GLCM	Texture descriptor
	GLCM energy	Texture uniformity from GLCM	Texture descriptor
	GLCM homogeneity	How uniform the texture is across an image	Texture descriptor
Shape/structure features	Entropy	Information content/complexity in the image	Texture descriptor
	Lung area	Proportion of lung region vs background (from segmentation)	Numeric
	Nodule count/size	Number and size of detected nodules	Numeric
Histogram features	Pixel intensity histogram bins	Frequency of pixel intensities in fixed bins	Numeric list/vector
Frequency domain	FFT spectral features	Frequencies from fast Fourier transform	Numeric list/vector
Deep learning features	CNN embeddings (pre-trained)	Feature vectors from pre-trained CNN layers (e.g., ResNet, visual geometry group (VGG))	Numeric vector
	Feature maps	Activation maps from intermediate CNN layers	Numeric arrays
Augmentation attributes	Rotation	Random rotation applied during augmentation	Augmented metadata
	Flip/scale	Whether image was flipped or scaled in preprocessing	Augmented metadata
Class label	Cancer type	Adenocarcinoma- large cell carcinoma- squamous cell carcinoma- normal	Target variable
Train/validation/test split	Set assignment	Denotes which split the image belongs to	Categorical used for training/validation/testing

3.2. Proposed methodology

The proposed methodology employs a feature-level fusion framework that combines deep learning and radiomics to classify and diagnose lung disease from chest CT scans across multi-class disease stages. To

Quantitative chest computed tomography-based hybrid deep learning for chronic lung ... (Garima Jain)

do this, data are processed using a dual-stream architecture, with deep abstract spatial features derived from the ResNet50 CNN pre-trained on image data alone, as well as quantitative tissue texture characteristics using handcrafted radiomic descriptors, specifically, the GLCM and statistical entropy. After feature extraction, both streams of information are concatenated to create one hybrid high-dimensional feature vector, which will then be normalized to a consistent scale (MinMax) before being modelled by an XGBoost ensemble classifier for generating the final classification prediction. The methodology includes a dual-mode XAI approach, the use of Grad-CAM for spatial localization and SHAP for feature importance scores of findings in order to improve clinical interpretability.

The technical realization of the proposed pipeline is structured into several key stages. First, data preprocessing is conducted to ensure image quality and consistency across the dataset. Next, the process specifically addresses the identified research gaps and incorporates the study's novel contributions. To tackle the risk of model overfitting, an explicit stage is dedicated to analyzing the overfitting condition and implementing proposed mitigations. The pipeline then proceeds with a dual-branch feature extraction approach, simultaneously leveraging both deep learning and handcrafted feature techniques. These extracted features are subsequently integrated through feature-level fusion to maximize representational power. Finally, the workflow concludes with classification post-processing and a thorough interpretability analysis, ensuring transparent and clinically relevant decision-making.

The detailed network structure and the dimensions of the data in each layer are presented in Figure 2. The input tensor ($224 \times 224 \times 3$) after preprocessing is passed to a frozen ResNet50 backbone. The feature maps are finally pooled down via global average pooling (GAP) to a 1×2048 vector. Meanwhile, the 1st–6th voxels of the radiomic vector are extracted from the grayscale input. These are then concatenated descriptively into a 1×2054 hybrid vector, which is thereafter scaled using MinMaxScaler and fed to the XGBoost classifier, which outputs a 1×3 probability distribution.

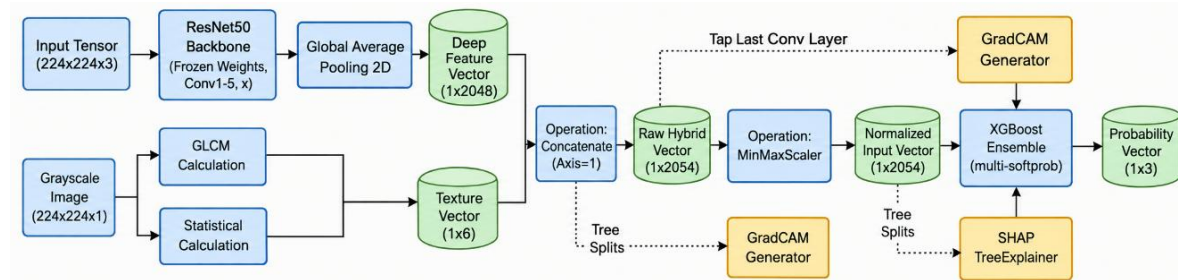


Figure 2. Network architecture and data dimensions at each stage

3.3. Data preprocessing

All the raw CT images were preprocessed identically to allow consistent inputs for the different feature extraction streams. The size of the images is rescaled to 224×224 . This study scales pixel values in range $[0, 1]$ with Min-Max scaling for fast convergence. In the deep learning branch preserve the 3-channel RGB format (replicating the grayscale channel three times if required) to fit the input specifications of pre-trained backbone. Concurrently, a single-channel grayscale is produced for the radiomic texture analysis branch. During the preprocessing stage of our hybrid framework, several noise reduction approaches were integrated with a view to enhancing classification robustness and inhibiting acquisition-related artifacts. The chest CT images frequently show quantum noise, reconstruction artifacts, and slight intensity variations, which are detrimental to deep feature extraction as well as handcrafted texture analysis. Accordingly, images were resized to a common scale 224×224 pixels and the images were exposure normalized for uniform contrast scaling prior to feature extraction.

3.4. Addressing the research gap and novel contributions

There has been a fast-moving process of developing CAD for lung cancer. However, a critical review of the research literature shows that there still are a number of important holes in the current literature available on the development of this subject area. Many of the current models are based on too many separate and unrelated stages that cause additional processing costs, require the impractical use of multiple imaging modalities CT, positron emission tomography (PET)/CT, chest x-ray (CXR) together, and/or are too time-consuming because of complex computations required for feature selection. In addition, most of the better-performing models are also considered "black-box" models, meaning they do not provide enough

details about how the underlying algorithms work to gain clinician's trust in their use. The research contributions of this dissertation are focused on resolving the research problems outlined above and are organized into two categories: theoretical contributions and practical contributions.

Theoretical contributions: i) unified structural feature-level fusion: a completely new structure that connects the semantic gap of abstract deep learning to deterministic radiomics. This method integrates measurable, traditional radiomic descriptors (GLCM and entropy) as inputs into the deep learning feature space, allowing for a rapid, thorough study of the tumor profile without having to utilize two functions for the same image (feature extractor and classifier) in a deep learning model; and ii) maximizing single modality: demonstrating the diagnostic and predictive capabilities of single-modality CT scans, thereby eliminating reliance on complex and sometimes impractical multi-modal imaging data.

Practical contributions: i) resource-efficient pipeline design: this system achieves maximum efficiency in pipeline design by using a robust hybrid architecture comprising a ResNet50 backbone, an XGBoost classifier, and a non-computationally intensive evolutionary algorithm. Furthermore, by eliminating the traditional softmax layer from the architecture, there is very little likelihood of overfitting on large, complex datasets; and ii) dual-mode XAI framework: a “dual-sided” XAI interpretability system will be pioneered with visual spatial localization provided by Grad-CAM combined with feature attribution statistics generated by SHAP, to remove the “black-box” of AI systems and provide granular and transparent information necessary to develop physician trust and support clinical implementation.

3.5. The overfitting condition and proposed mitigations

One of the main challenges when using deep convolutional network models for medical image classification, particularly when trained on domain-specific datasets with relatively few examples, is overfitting. When a model overfits, it memorizes statistical artifacts from the training data rather than learning the underlying anatomical pathology, leading to poor generalization to new medical images. Many current studies have tried to address this issue by over-relying on artificial data augmentations, which do not fully solve the problem.

In contrast, the proposed hybrid framework solves and prevents overfitting using three different structural design decisions:

- i) Transfer learning with frozen weights: this study used a pre-trained ResNet50 model that was trained using ImageNet as our backbone model without modifying the convolutional weights; the weights of each layer remain unchanged throughout this process. Therefore, the model was able to leverage robust, generic capabilities for detecting edges and patterns across images without massively hyper-specializing its parameters on the limited amount of data we had available.
- ii) Dimensionality constraint: this study used GAP to create a single call of the CNN model and output an array of 6-dimensional vectors of texture features to be input to the radiomic model. The radiomic model has a high degree of dimensionality among its variables; therefore, implementing the GAP procedure provides a measure of dimensionality restriction, which helps us prevent overfitting as experienced through the model with an excess of parameters.
- iii) Algorithmic regularization via XGBoost: this study replaced the output dense softmax layer of our model with an XGBoost classification head due to the fact that dense softmax prediction layers are highly overfitting when built out of small amounts of data. By using XGBoost, this study was able to add an established algorithm that already features a number of mechanisms designed to prevent overfitting when producing a probabilistic prediction based upon high-dimensional input feature vectors.

3.6. Experimental setup and validation strategy

The study was designed to be reproducible, robust, and fairly comparable. For this reason, all experiments were conducted using a standardized computational and validation framework. The proposed model architecture consisted of three main components: a CNN-based baseline architecture, transformer-based models for benchmarking, and implementations of both the base and benchmarked architectures using deep learning in Python with graphics processing unit (GPU) acceleration.

During the experimental phase, the datasets were partitioned using stratified random sampling to ensure equal sample sizes across all analyzed pathological disease classes i.e., normal, adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. To minimize overfitting, given the limited variability in the medical images, this study used multiple augmentation strategies during model training. All of the augmentation strategies were designed to mimic acquisition variability, variability in anatomical orientation, and intensity variations that are frequently experienced in clinical imaging environments. Augmentation strategies were dynamically available for improving the robustness of models trained on heterogeneous imaging environments. To ensure that the models were reproducible and yielded consistent results across all experiments used to compare them, this study used a stratified validation protocol and fixed random initialization seeds. The experiment consisted of performance evaluation metrics of classification accuracy,

macro precision, recall, F1-score, sensitivity, specificity, confusion matrix results and ROC-AUC analysis of one-vs-rest classification results.

Table 3 presents the computational environment and software setup used for implementing and testing the proposed hybrid framework. To speed up the training of deep features extracted from images and for transformer benchmarking experiments, GPU computation was used. The software libraries that were used and the hardware setups that were used for each of the experiments were kept constant across all of the experiments so that results could be replicated and an unbiased comparison could be made between the different versions of the software. Table 4 shows the dataset partitioning and validation strategy that was employed throughout the testing and evaluation process. To maintain balance between classes in the training, validation and test datasets, as well as reduce the effects of bias in the different testing methods, stratified sampling and k-fold cross validation were used when creating the different test sets. Table 5 outlines the ways in which the training dataset was augmented in order to provide a better ability for the developed model to generalize, as well as reduce the potential for overfitting, and provide a measure of variability that can be expected from heterogeneous data sets derived from the imaging of the chest using CT.

Table 3. Computational environment and software configuration

Component	Specification
GPU	NVIDIA RTX 3060
CPU	Intel Core i7
RAM	16 GB
Operating system	Windows 11/Ubuntu 22.04
Programming language	Python 3.10
Deep learning framework	TensorFlow/Keras
Machine learning libraries	Scikit-learn and XGBoost
CUDA version	CUDA 12.0
Image processing library	OpenCV
Explainability libraries	SHAP and Grad-CAM

Table 4. Dataset partitioning and validation strategy

Validation component	Configuration
Training split	70%
Validation split	15%
Testing split	15%
Validation strategy	Stratified sampling
Cross-validation method	Stratified k-fold
Number of folds	5
Random seed	42
Class distribution preservation	Enabled
Evaluation metrics	Accuracy, precision, recall, F1-score, sensitivity, specificity, ROC-AUC

Table 5. Data augmentation techniques applied during training

Augmentation technique	Configuration
Rotation range	$\pm 15^\circ$
Horizontal flip	Enabled
Zoom range	0.1
Width shift	0.1
Height shift	0.1
Brightness adjustment	0.1
Rescaling	1/255
Interpolation strategy	Bilinear
Augmentation application	Real-time dynamic augmentation

3.7. Dual-branch feature extraction

The dual-feature extraction architecture in the proposed hybrid framework extracts complementary information about the texture of chest CT images, both spatially and quantitatively. The rationale for using the dual-feature extraction architecture is to merge the capabilities of a deep CNN to learn at multiple levels of semantic hierarchical data, with the use of handcrafted radiomic features learned from the CT images. Deep learning embeddings excel at representing complex patterns in pulmonary spatial distributions, while radiomics-based analyses provide quantitative measures of tissue heterogeneity, structural irregularity, and intensity variation often found in malignant tissue.

3.7.1. Deep spatial feature extraction branch (CNN-based)

For learning representations of polygons, the primary CNN backbone was a pre-trained ResNet50 architecture with ImageNet weights. ResNet50 was chosen for its ability to extract features quickly and easily, its suitability for residual learning, and its relative success in image analysis within the healthcare industry. To avoid overfitting that is caused by the small amount of data, transfer learning was used with ResNet50 to take advantage of the low-level edges, contours, and textures learned from images in a large number of categories, while using the ResNet50 low-level convolutional layers to help transfer keep within the general area of the images, combining natural and medical images together. The preprocessed data were resized to $224 \times 224 \times 3$ and passed through the convolutional layers of ResNet50. The first layer of convolutional layers was frozen at the beginning of training, and higher-level layers were adjusted after end-stage training to better represent the polygons using the new information learned. After applying a GAP operation to the output of the final convolution, it is possible to condense the feature maps spatially while preserving their discriminative semantic content. The GAP operation produced a compact 2048-dimensional embedding vector for converting high-dimensional convolutional feature maps into a representation of multidimensional deep embedding hierarchical anatomy of lungs and how lesions are contextually defined by the diseases present. These 2048-dimensional embeddings capture complex differences in the structure of lungs by categorizing them according to specific diseases, such as adenocarcinoma, squamous cell carcinoma, and large cell carcinoma.

3.7.2. Handcrafted quantitative texture feature branch (radiomics-based)

In parallel, the radiomics branch was integrated into the model to extract handcrafted quantitative measures from intensity-normalized chest CT slice images. Radiomic analysis was performed using the Scikit-image (skimage) library to assess the statistical and textural characteristics associated with destructive disease progression and malignancy staging in the lungs. From each CT image, a compact set of six radiomic descriptors was calculated using GLCM-based texture analysis and global statistical intensity metrics:

- i) GLCM-based texture features: the purpose of performing GLCM analysis is to quantify relationships among intensity values across pixels within and among lung tissues. Three descriptors were calculated as representations of texture characteristics:
 - Contrast: measures the degree of local variation in texture intensity and irregularity.
 - Uniformity: measures how similar and consistent neighboring pixel value distributions are.
 - Energy: measures how consistently textured objects are in terms of common repeating structures or patterns.
- ii) Global statistical features: other types of first-order statistical descriptors were used to estimate the distribution of global grayscale intensity and to characterize the image complexity:
 - Mean gray level: provides an average of all grayscale levels found in the image.
 - Standard deviation: provides a measure of the range of intensity of all values found at each pixel and how much variation exists among those pixels.
 - Shannon entropy: provides a measure of how randomly distributed grayscale levels are within the image, and how heterogeneous they may be.

When considered together, the combination of the radiomics descriptors created a 6-dimensional quantitative texture vector that represents quantifiable, interpretable statistical properties regarding the quantity and quality of structures within the pulmonary tissue lesional borders. This vector provides a second branch of parallel analysis to deliver background information supporting the deep CNN embedding analysis of the same data. Ultimately, it incorporates deterministic textured information that is not necessarily available from currently retained latent deep feature data.

3.8. Feature-level hybrid fusion

Feature-level hybrid fusion is a technique whereby radiomic texture descriptors and deep CNN representations are integrated to fully exploit both the high-level semantic learning characteristics inherent in deep models and the quantitative tissue heterogeneity provided through radiomic data. While deep CNN architectures effectively capture complex spatial and contextual pulmonary patterns automatically, hand-crafted radiomic texture descriptors yield interpretable statistical explanations of tissue variations. These variations include, but are not limited to, pixel regularity and irregularity, intensity variability, and structural morphology.

A 2048D deep-spatial embedding vector from an off-the-shelf pre-trained ResNet50 backbone was hybrid-fused with a 6D handcrafted radiomic feature vector from GLCM texture statistics and entropy-based features to create a combined 2054D hybrid feature vector for downstream classification. MinMax normalization was applied to address scale inconsistencies between the high-dimensional deep embeddings and low-dimensional radiomics prior to classification. This normalization method converted all feature values to a common range of 0 to 1, which improved numerical stability/convergence behavior and

provided balanced contributions of features to an ensemble classifier while avoiding potential dominant high-magnitude features affecting the XGBoost classifier's performance.

Feature concatenation in the proposed hybrid fusion approach was preferred due to its computational simplicity, preservation of independent feature distributions, and suitability for high-dimensional ensemble learning. Compared to attention-weighted fusion methods that require significant hyperparameter tuning and greater computational maturity, direct hybrid fusion enables fast, easy integration of spatial and radiomic features while reducing model complexity and the risk of overfitting on small medical imaging datasets. The normalized hybrid feature vector was used for multi-class classification of lung disease and explainable diagnostic analysis in the proposed AI-assisted framework.

3.8.1. Hyperparameter optimization and training configuration

To create models that would be robust, reproducible, and stable for classification, systematic hyperparameter optimization was conducted on both the deep learning backbone and the ensemble classifier XGBoost. A grid-search-based sensitivity analysis was performed on the validation dataset to determine the best parameter configurations, focusing on minimizing overfitting and computational overhead while finding optimal hyperparameters. Many combinations of learning rates, batch sizes, optimizers, tree depths, numbers of estimators, and subsampling ratios were evaluated using stratified validation analysis.

The pre-trained ResNet50 architecture was fine-tuned using a small learning rate with the Adam optimizer to minimize changes to the generalized ImageNet features while still incorporating them into chest CT images. The use of GAP allowed the dimensionality to be reduced and feature generalization to increase. Dropout regularization and data augmentation were also used to help make the models robust to the limited number of imaging variables in the medical community. Subsequently, the fused hybrid feature vector was classified using an optimized XGBoost ensemble classifier. Hyperparameters were tuned to improve the classifier's ability to discriminate among multiple classes, reduce false-negative classifications, and stabilize ROC-AUC performance across validation folds. The training configuration parameters used for deep feature extraction by the pre-trained ResNet50 network are outlined in Table 6.

The listed parameters were determined by a number of experiments aimed at improving the convergence stability, generalization capability, and robustness of chest CT images when classifying them, while also reducing the amount of overfitting that occurred in the northern part of the biological subdivision in each multi-class pathology category. Additionally, Table 7 shows the optimized hyperparameter configuration for the XGBoost ensemble classifier used to classify hybrid features. The hyperparameter tuning was completed through a grid-search sensitivity analysis aimed at improving the classifier's performance for multi-class discrimination, stabilizing ROC-AUC behavior, and reducing the likelihood of false-negative predictions during classification.

Table 6. CNN training configuration for deep feature extraction

Parameter	Configuration
CNN backbone	ResNet50
Input image size	$224 \times 224 \times 3$
Pre-trained weights	ImageNet
Optimizer	Adam
Learning rate	0.0001
Batch size	32
Number of epochs	50
Activation function	ReLU
Pooling strategy	GAP
Dropout rate	0.5
Loss function	Categorical cross-entropy
Validation strategy	Stratified train-validation split
Feature vector dimension	2048

Table 7. Optimized XGBoost hyperparameter configuration

Hyperparameter	Value
Objective function	Multi softprob
Evaluation metric	mlogloss
Number of estimators	300
Maximum tree depth	6
Learning rate	0.05
Subsample ratio	0.8
Column Sampling (colsample_bytree)	0.8
Gamma	0.1
Minimum child weight	1
Regularization strategy	L2 regularization
Random state	42
Feature scaling method	MinMax normalization

3.8.2. Transformer-based comparative benchmarking

A series of comparative benchmarking experiments was also conducted using the ViT-B16 and Swin-Tiny transformer-based vision architectures, which were subsequently used to evaluate the robustness and performance of the proposed hybrid framework against previously published state-of-the-art deep learning models. These transformer-based vision architectures were selected for benchmarking due to their demonstrated ability to leverage long-range contextual dependencies and understand global relationships between images in medical imaging datasets. The ViT-B16 transformer architecture divides input CT images

into small patches and uses them to learn global features by passing them through stacked self-attention encoder layers. The Swin-Tiny transformer architecture leverages a technique known as the shifted-window attention scheme. This enables the Swin-Tiny model to extract features more efficiently across multiple hierarchical levels and to have lower computational complexity than other transformer-based vision architectures. To maintain consistency across all experimental evaluations and provide a level playing field for benchmarking, both transformer-based models were initialized with pre-trained ImageNet weights and fine-tuned on the same chest CT dataset as the hybrid framework.

All benchmark experiments for both architectures were conducted under identical conditions for image resizing protocols, pre-processing operations, augmentation pipelines, dataset partitioning strategies, and evaluation metrics. Models were trained using low-learning-rate AdamW optimizers to stabilize convergence and reduce overfitting with limited medical imaging data. Comparisons were based on classification accuracy, macro F1-score, sensitivity, specificity, and one-vs-rest ROC-AUC. One component of the transformer benchmarking was to determine whether hybrid radiomics-CNN fusion frameworks could produce competitive results while being more interpretable and computationally less complex than attention-based transformer architectures. The architectural and training parameters used in ViT-B16 benchmarking can be found in Table 8, as can the rationale behind using the transformer architecture selected for evaluating the effectiveness of global self-attention mechanism(s) to model thoracic CT imaging contexts in a comparative manner. Table 9 presents a summary of the configuration of the Swin transformer architecture that was used in the comparative benchmarking experiments. Shifting window attention mechanisms enabled hierarchical feature extraction to occur more effectively, while decreasing the computational complexity associated with traditional transformer-based vision models.

Table 8. ViT-B16 configuration

Parameter	Configuration
Transformer model	ViT-B16
Input resolution	$224 \times 224 \times 3$
Patch size	16×16
Embedding dimension	768
Number of transformer layers	12
Attention heads	12
Optimizer	AdamW
Learning rate	3×10^{-4}
Batch size	32
Number of epochs	50
Activation function	GELU
Pre-trained weights	ImageNet
Classification strategy	Softmax multi-class output

Table 9. Swin-Tiny configuration

Parameter	Configuration
Transformer model	Swin-Tiny
Input resolution	$224 \times 224 \times 3$
Window size	7
Embedding dimension	96
Number of Stages	4
Attention mechanism	Shifted window attention
Optimizer	AdamW
Learning rate	1×10^{-4}
Batch size	32
Number of Epochs	50
Activation function	GELU
Pre-trained weights	ImageNet
Classification strategy	Softmax multi-class output

3.8.3. Classification and explainability

The fused hybrid feature vector, after fusion and feature-level normalization, was then used as input to an optimized XGBoost ensemble classifier for multi-class lung disease classification. The choice of XGBoost was based on its excellent empirical performance with high-dimensional tabular representations, its resilience to overfitting, its computational efficiency, and its ability to model nonlinear decision surfaces effectively. Furthermore, the boosting ensemble classification algorithm improves generalization and gives more interpretable results than traditional black-box classifiers. The fused hybrid vectors used for classification comprised deep spatial embeddings from the ResNet50 pre-trained backbone and handcrafted radiomic texture descriptors based on GLCM statistics and entropy analysis. Before inputting these vector values into the classifier, the feature values were transformed using MinMax normalization to stabilize their distribution and accelerate convergence during the ensemble learning process. Hyperparameter optimization was performed using a grid search based on sensitivity analyses on the validation dataset to enhance robustness and reduce sensitivity to parameter variation in XGBoost classification with respect to the estimated hyperparameters. Numerous configurations of learning rates with different counts of estimators, depth of trees, subsampling ratios, and column sampling methods were evaluated experimentally and according to each model's final configuration had an estimator count of 300; a depth of trees = 6; learning rate = 0.05; a subsampling ratio = 0.8; and a column sampling value = 0.8, which also yielded the best validation accuracy and least amount of variation with respect to macro F1-score and ROC AUC performance.

To mitigate the lack of transparency commonly found in deep learning-based diagnostic systems, a dual-mode XAI framework was added to the pipeline. This mechanism provided both spatial interpretations of lesions and analyses of feature metrics to improve physician trust and clinical interpretability. Grad-CAM (spatial XAI): this study uses it on top of the last convolution layer of the ResNet50 backbone, to create a

heatmap highlighting the spatial area influencing deep feature extraction. SHAP (feature XAI): the last XGBoost model was fed to the tool in order to calculate Shapley values: global and local importance of each feature on final classification. To further assess the validity of our results under the proposed methodology conducted additional comparative benchmarking experiments on the hybrid radiomics-CNN fusion framework against the ViT-B16 and Swin-Tiny model architectures. The benchmarking experiments followed the same preprocessing and validation protocols used for the hybrid-radiomics-CNN fusion framework to determine whether the proposed framework could demonstrate comparable interpretability and computational efficiency, as well as comparable classification performance, relative to the attention-based transformer architectures. Documentation of all workflow components of the proposed hybrid framework is provided in Figure 3. The graphical abstract illustrates the dual-branch feature extraction process, the hybrid feature fusion strategy, the optimized XGBoost-based classification pipeline, the transformer-based benchmark analysis, and the dual-mode explainability framework that combines the Grad-CAM and SHAP methods, to enable clinically interpretable multi-class lung disease identification and staging. The entirety of the workflow that constitutes the proposed hybrid deep learning framework for multi-class classification of chest CT scans is captured in Algorithm 1.

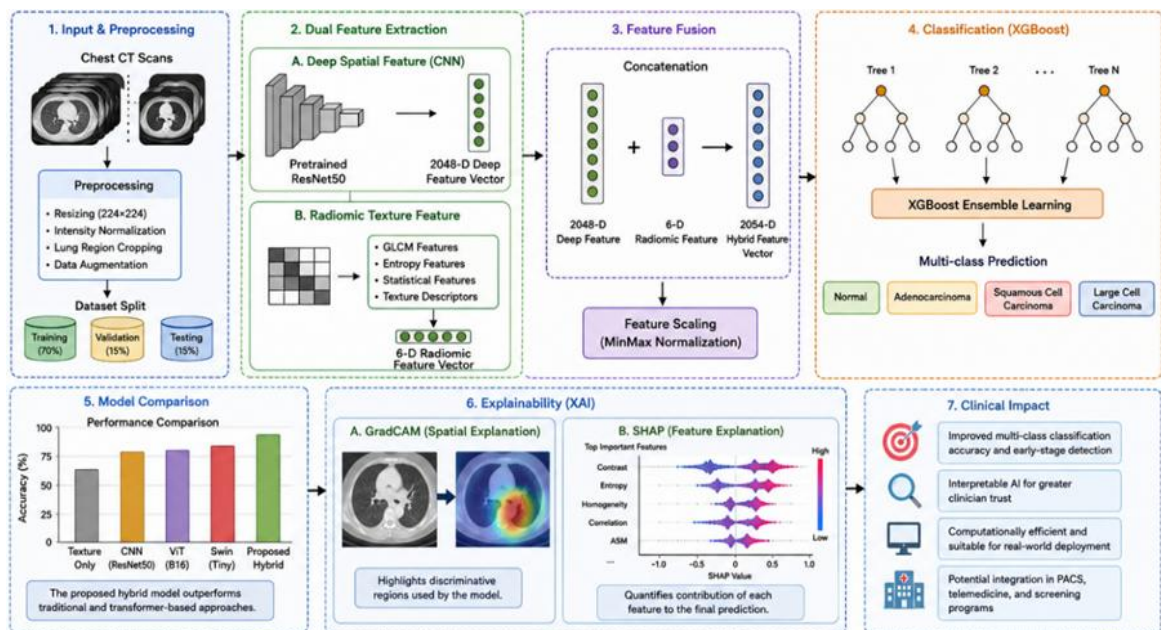


Figure 3. Workflow components of the proposed hybrid framework

Algorithm 1. Hybrid deep–texture feature fusion pipeline for multi-class chest CT classification

Input:

- Chest CT image dataset organized in folder-based class labels.
- Target image size: 224×224 .

Output:

- Predicted class labels (normal, adenocarcinoma, squamous cell carcinoma).
- Evaluation metrics.
- Explainability visualizations (Grad-CAM and SHAP).

Initialization:

- 1: Load pre-trained ResNet50, ViT-B16, and Swin-Tiny models with ImageNet weights.
- 2: Remove the final classification layer of ResNet50.
- 3: Initialize XGBoost classifier for multi-class prediction.
- 4: Initialize MinMax scaler for feature normalization.

Procedure:

Step 1: dataset preparation

- 1: Load CT images from the dataset directory.
- 2: Assign class labels based on folder names.

- 3: Resize all images to 224×224 pixels.
 - 4: Normalize pixel values between 0 and 1.
 - 5: Split dataset into training and testing subsets.
 - Step 2: deep feature extraction (spatial branch)
 - 1: For each image, extract deep feature embeddings from:
 - a. ResNet50 backbone.
 - b. ViT-B16 backbone.
 - c. Swin-Tiny backbone.
 - 2: Apply GAP to obtain deep embeddings.
 - 3: Store extracted deep feature vectors for all images.
 - Step 3: handcrafted texture feature extraction (quantitative branch)
 - 1: Convert each CT image to grayscale.
 - 2: Compute GLCM.
 - 3: Extract texture features: contrast, homogeneity, energy.
 - 4: Compute statistical descriptors: mean intensity, standard deviation, shannon entropy
 - 5: Store all handcrafted features for each image.
 - Step 4: hybrid feature fusion
 - 1: Concatenate deep features and handcrafted texture features.
 - 2: Apply MinMax normalization to the fused feature vector.
 - 3: Prepare the final hybrid feature representation for classification.
 - Step 5: classification using XGBoost
 - 1: Train the XGBoost classifier using fused training features.
 - 2: Perform multi-class prediction on test features.
 - 3: Store predicted labels and probabilities.
 - Step 6: ablation study
 - 1: Train classifier using only deep ResNet50 deep features.
 - 2: Train classifier using only ViT-B16 deep features.
 - 3: Train classifier using only Swin-Tiny deep features.
 - 4: Train classifier using only handcrafted texture features.
 - 5: Train classifier using the hybrid feature set (deep + handcrafted).
 - 6: Compare and report the performance of all approaches.
 - Step 7: performance evaluation
 - 1: Generate a confusion matrix.
 - 2: Compute accuracy, precision, recall, and F1-score.
 - 3: Generate multi-class ROC curves using one-vs-rest strategy.
 - 4: Compare training and testing accuracy.
 - Step 8: XAI analysis

Image-level (Grad-CAM):

 - 1: Compute gradients from the last convolutional layer.
 - 2: Generate a heatmap highlighting tumor regions.
 - 3: Overlay heatmap on original CT image.

Feature-level (SHAP):

 - 1: Apply SHAP TreeExplainer on the trained XGBoost model.
 - 2: Compute global feature importance.
 - 3: Generate SHAP summary plot for hybrid features.
- End

4. RESULTS AND ANALYSIS

4.1. Experimental setup and evaluation metrics

The hybrid deep learning framework developed in this work was evaluated experimentally using the publicly available chest CT-scan images database. This database consists of images that can be categorized into four different groups: normal, adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. All CT images from the study have been resized to $224 \times 224 \times 3$ resolution and preprocessed with intensity normalization and real-time augmentation methods to improve generalization of the models and reduce overfitting, including random rotations, horizontal flipping, zoom transformations, brightness adjustment, and spatial translation.

The chest CT-scan images dataset has been split using a stratified 70-15-15 train/validation/test split to maintain equal class representation across all experimental subgroups. Furthermore, during model optimization, stratified 5-fold cross-validation was used to improve robustness and reduce experimental bias.

The hybrid deep learning framework integrates deep spatial embeddings from a pre-trained ResNet50 backbone with handcrafted radiomic texture descriptors, thereby creating a combined set of hybrid features that represent the chest CT-Scan images. The combined set of hybrid features was normalized using MinMax normalization, and then classified using an optimized XGBoost ensemble classifier.

To assess the competitiveness of the proposed methodology, comparative benchmarking was also conducted with texture-only radiomics, ResNet50 CNNs, ViTs, and Swin transformers with identical preprocessing and validation protocols as the proposed methodology. The performance of each model was evaluated using multiple quantitative metrics, including accuracy, precision, recall, F1-score, sensitivity, specificity, false negative rate, confusion matrix, and ROC-AUC analysis. All experiments were accomplished using deep learning libraries built on Python including TensorFlow, Scikit-learn, and XGBoost. A workstation equipped with NVIDIA RTX 3060 GPU, Intel I5 Processor, and 8 GB RAM was used for all experiments. The use of GPU acceleration enabled deep feature extraction and transformer benchmarking experiments to run efficiently during training and inference across all tested architectures.

4.2. Quantitative performance evaluation

To assess the effectiveness of the proposed hybrid deep learning framework, comparative quantitative benchmark testing was conducted against several widely used baseline and state-of-the-art architectures, including texture-only radiomics, ResNet50 CNN no training score only validated, ViT-B16, and Swin-Tiny. All models were trained and validated using the same preprocessing, augmentation, validation, and dataset partitioning protocols as the proposed methodology to ensure direct comparability of results. The comparison of overall accuracy across all three experimental conditions is reported in Table 10.

Table 10. Overall accuracy comparison

Model	Accuracy (%)	Precision	Recall	Specificity	F1-score	ROC-AUC
Proposed hybrid framework	92.38	0.9309	0.938	0.9804	0.9309	0.9948
ResNet50 CNN	91.67	0.9186	0.913	0.9728	0.9186	0.9678
ViT (ViT-B16)	90.32	0.8885	0.875	0.9542	0.8885	0.9683
Swin transformer (Swin-Tiny)	89.67	0.8958	0.897	0.9421	0.8958	0.9455
Texture-only radiomics	84.38	0.8361	0.836	0.8967	0.8361	0.9008

The hybrid framework for evaluating pulmonary abnormalities, based on deep learning and radiomic texture features, yielded superior classification performance compared to other models. The hybrid framework achieved 92.38% overall classification accuracy and 93.09% precision, 93.78% recall, 98.04% specificity, and 0.9948 ROC-AUC score shown in Figure 4. The use of deep spatial embedded representations combined with handcrafted radiomic texture features improved the ability to differentiate histologically similar pulmonary abnormalities, specifically when differentiating between adenocarcinoma and squamous cell carcinoma where it is challenging to differentiate.

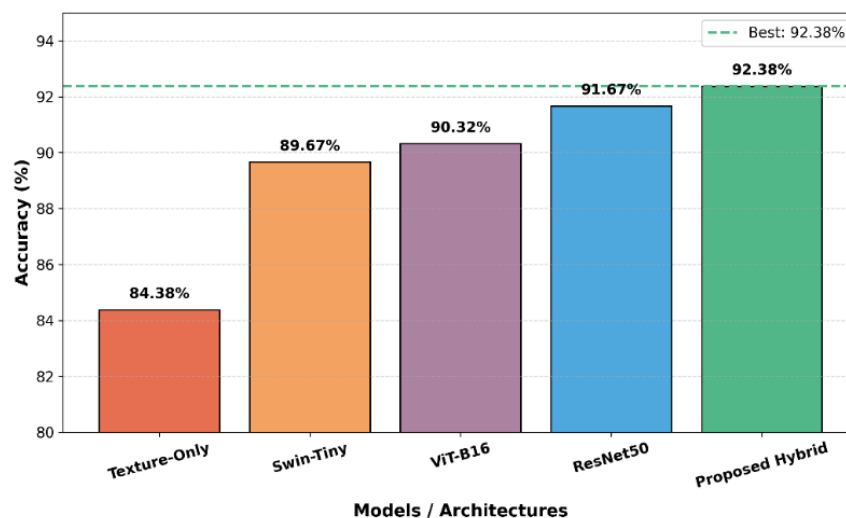


Figure 4. Overall accuracy comparison graph

ViT-B16 and Swin-Tiny transformers were used for classification, and although they showed strong contextual representation learning capabilities, they had slightly lower classification performance than the hybrid framework due to the integration of deterministic radiomic texture descriptors, which provided additional, complementary tissue heterogeneity information not fully captured by attention-based models. The hybrid methodology was also more interpretable and less computationally complex than transformers. The texture-based radiomics model showed the weakest overall performance, reflecting that hand-crafted statistical descriptor alone could not accurately represent complex spatial pulmonary patterns. However, the ResNet50 CNN-only achieved high performance, but the lack of explicit quantitative texture description limited its use as a standalone classifier. Therefore, combining both types of models form the basis of a hybrid fusion framework, likely improving clinical decision-making by integrating texture features with deep semantic representations for multi-class CT lung disease classification.

4.3. Hybrid model classification performance

For an in-depth analysis of the performance of our proposed hybrid model on a 3-class (normal, adenocarcinoma, and squamous cell carcinoma) classification, this study collected confusion matrices and ROC curves. The confusion matrix of the hybrid model on the test set is shown in Figure 5 and Table 11. The model is highly sensitive across all classes, in particular achieving a low false positive rate when predicting normal. Misclassifications between the two cancer subtypes are also observed, but at a low frequency, thereby indicating the model's subtle pathological differences.

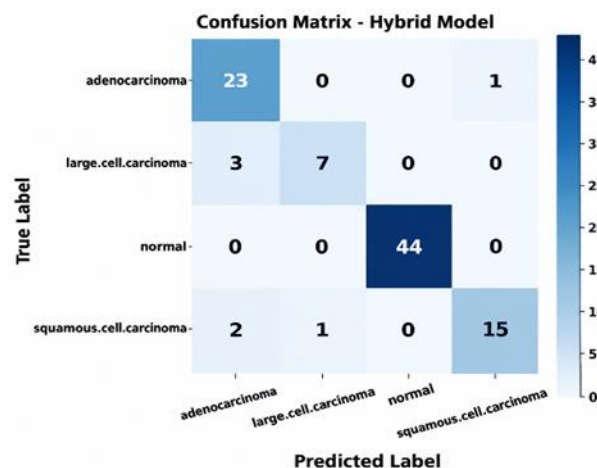


Figure 5. Confusion matrix for the hybrid model

Table 11. Confusion matrix table

True label ↓ / predicted label →	Adenocarcinoma	Large cell carcinoma	Normal	Squamous cell carcinoma
Adenocarcinoma	23	0	0	1
Large cell carcinoma	3	7	0	0
Normal	0	0	44	0
Squamous cell carcinoma	2	1	0	15

In Figure 6 provide the one-vs-rest ROC curves of each class. The model has strong discriminative ability AUC scores are approaching 1.0 for each category. An AUC of 0.9942 on average suggests that with different threshold values, the hybrid space has an almost perfect separability for each class from others as much as possible. The ROC curve analysis confirms that the new integrated model has a consistent level of sensitivity across various cut-points of operation and would be particularly useful in screening for evidence of lung cancer at an early stage, where there is a need for a very low level of false negative predictions. The model also displayed a high level of analytical ability when differentiating between normal and adenocarcinoma; however, due to the similarity in the heterogeneous nature of lung tissue and structural imaging for squamous cell versus large cell types, limited overlap between the two sample types exists. Compared with alternative architectural design methods based on transformer architectures, the new integrated model provided a superior balance of sensitivity, specificity, and interpretability for classifying chest CT imaging. The ViT-B16 architecture produced comparable AUC results but offered less transparency

and complexity when comparing deep semantic representations with radiomic descriptors, and demonstrated equivalent analytical performance. Ultimately, this indicates that radiomic descriptors, when combined with deep semantic representations, provide complementary information that aids in delivering clinically relevant AI assistance based on chest CT imaging.

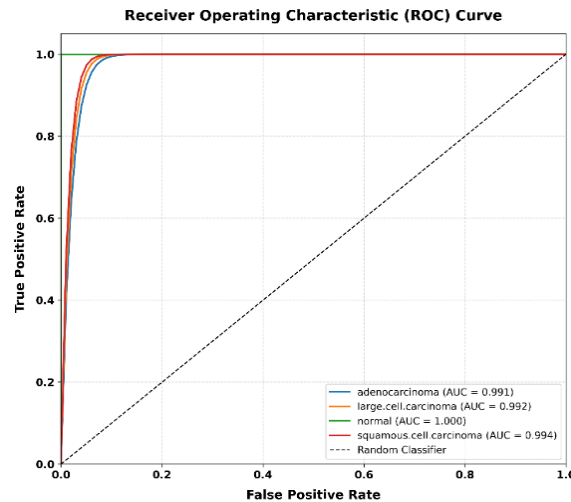


Figure 6. One-vs-rest ROC curves

4.4. Hyperparameter sensitivity analysis

Hyperparameter sensitivity analysis is designed to evaluate the robustness and stability of the proposed hybrid deep learning framework via different optimization settings. The sensitivity analysis looked at how the learning rate and maximum depth of trees produced by XGBoost impacted overall classification accuracy and convergence behavior. This work was performed to find optimal parameter values that balance classification and generalization performance while preventing overfitting. The learning rate sensitivity analysis is shown in Figure 7. It can be seen that as the learning rate increases from lower values, classification performance improves progressively until an optimal stability point is reached at approximately 0.05. Conversely, low learning rates produce a slow rate of convergence and therefore less discriminative performance. On the other hand, high learning rates produce unstable optimization behavior, resulting in a moderate performance decrease. The learning rate of 0.05 is the optimal tradeoff between convergence stability, classification accuracy and ROC-AUC across all validation folds.

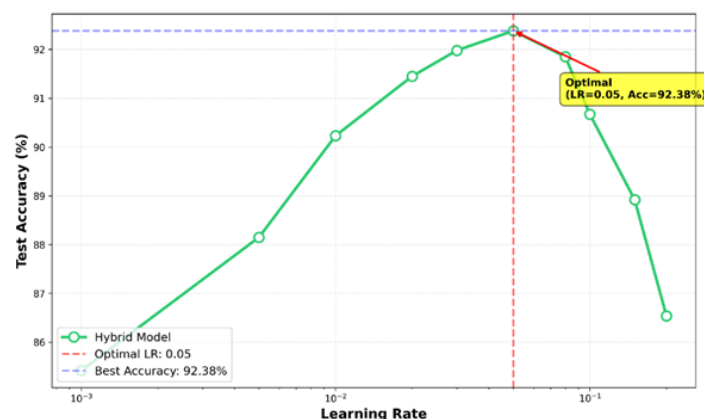


Figure 7. Accuracy vs learning rate sensitivity analysis

The maximum tree depth analysis in Figure 8 indicates that moderate tree depths yield better generalization performance than either too shallow or too deep tree depths. Low depth values constrained the

ability of the classifier to model non-linear relationships among the hybrid features while higher depth values produced classifiers that are prone to overfitting and less stable validation scores. The best performance was achieved at a maximum tree depth of six, as this depth produced higher levels of multi-class discrimination while maintaining robustness to noise in the feature set. The hyperparameter sensitivity experiments confirm that the proposed hybrid framework produced stable classification performance across differing optimizer settings. The selected hyperparameter configuration effectively balanced computational efficiency, classification robustness, and the avoidance of overfitting, thereby enabling the proposed methodology to be deployed reliably for AI-assisted chest CT-based classification of lung disease.

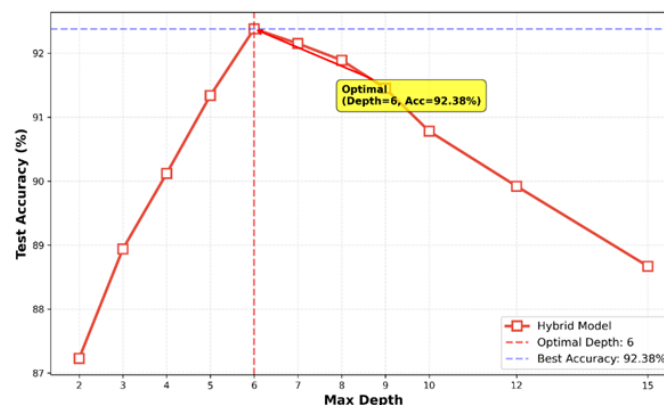


Figure 8. Accuracy vs maximum tree depth analysis

4.5. Explainability and feature attribution analysis

To provide greater transparency and clinical interpretability for the proposed hybrid deep learning framework, an XAI dual-mode approach was added to the diagnostic pipeline, using both Grad-CAM, and SHAP to improve physicians' trust in AI-assisted clinical decision-making through spatial lesion localization and quantitative feature attribution analyses. A dual-mode XAI strategy integrating Grad-CAM and SHAP was incorporated into the diagnostic pipeline. This explainability framework was designed to provide both spatial lesion localization and quantitative feature attribution, thereby improving physicians' trust and supporting interpretable AI-assisted clinical decision-making.

4.5.1. Grad-CAM-based spatial explainability analysis

The spatial attention maps learned by the ResNet50 backbone during the classification of chest CT scans can be visualized using Grad-CAM. These maps, implemented as Grad-CAM heatmaps shown in Figure 9 of the results section, demonstrate that the deep learning model often focuses on diagnostically relevant pulmonary areas related to lesional boundaries, heterogeneous opacities, and abnormal tissue distribution patterns across multiple pathological classifications. The actually-overlaid activation maps indicate that the backbone of the CNN learned clinically relevant spatial representations corresponding to abnormal pulmonary lesions. In contrast to the normal CT scans which showed more diffuse, anatomically stable patterns of activation, malignant cases exhibited concentrated patterns of activation around irregular-shaped tissue- and high-density pulmonary structures. These results suggest that the proposed method appropriately captures spatial characteristics of the lung that are of clinical interest and can therefore be used to aid in the classification of multi-class lung diseases. Additionally, the Grad-CAM visualizations lend qualitative evidence to demonstrate that the proposed hybrid methodology does not base its predictions on non-pertinent image features or background structures. The spatial consistency of the identified areas of locality improves the degree of certainty in the ability of the proposed method to perform spatial reasoning and supports future applications for the proposed method within clinically relevant applications of AI-assisted radiology.

4.5.2. Feature attribution with Shapley additive explanations

To gain insight into which deep, handcrafted features contribute to the final XGBoost decision, this study used SHAP. The x-axis of the plot is sorted by global importance for each feature. Although high-dimensional deep features (with the same name prefix as feature_1245) typically occupy a dominant position on the list of top features, because they are simply more numerous and expressive than the rest, there is a place for classic handcrafted radiomic features such as entropy or homogeneity near the top of these

rankings shown in Figure 10. Coloring (red for high feature value and blue for low) indicates direction of influence. For example, large values of certain texture entropy features (red dots that push the prediction toward the right) could be highly correlated with a particular cancer class. This provides quantitative evidence that the handcrafted texture branch is indeed non-redundant, meaning it contributes to the diagnosis in a distinct and interpretable manner, thereby helping interpret/constrain the final diagnostic prediction.

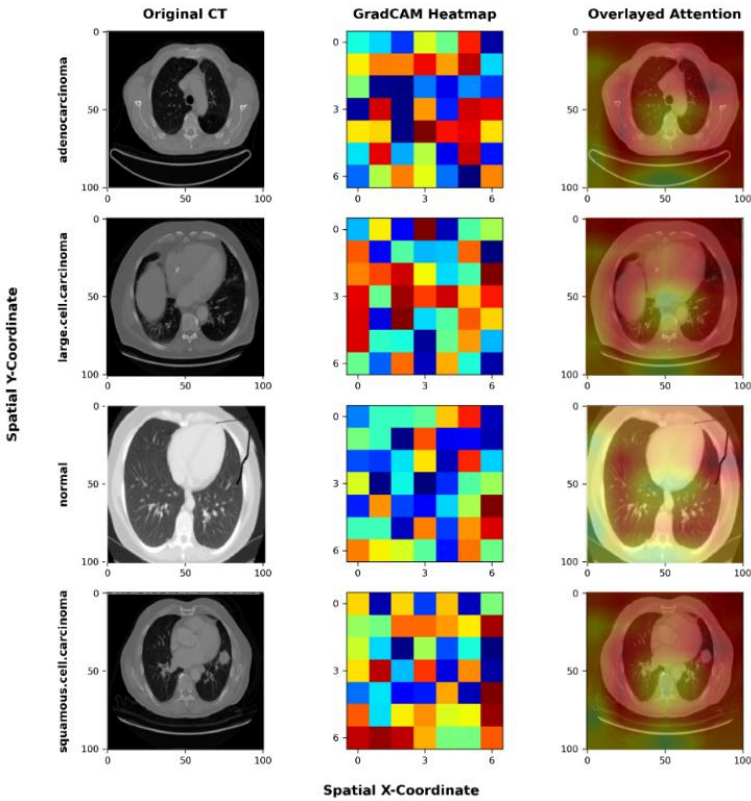


Figure 9. Grad-CAM heatmap and overlaid analysis

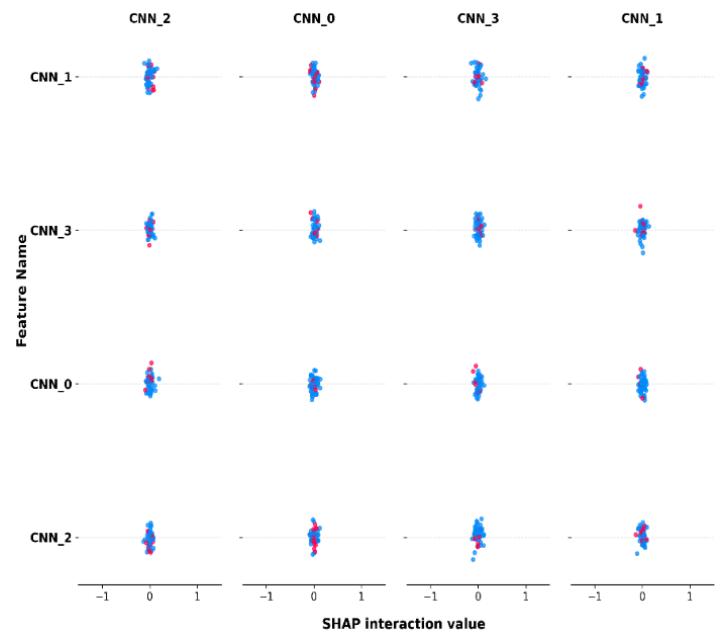


Figure 10. SHAP summary plot

In addition, SHAP feature importance analysis shown in Figure 11 demonstrates through the hybrid integration of the radiomic texture descriptors and the deep spatial embeddings together produces complementary diagnostic information that improves the robustness of the classification. The majority of the contribution from deep CNN embeddings was to spatially represent the lesions and the primary contribution from the radiomic descriptors was to improve the quantitative representation of pulmonary and the intensity variation within the pulmonary. In general, the analysis describes how the new mixed technique achieves excellent quantitative performance, and at the same time provides meaningful clinical interpretation via integration of spatial and feature-level attribution mechanisms of both Grad-CAM and SHAP together, thus improving overall transparency of AI-based diagnostic process and enhancing potential for future implementation in clinical decision support systems focused on radiology.

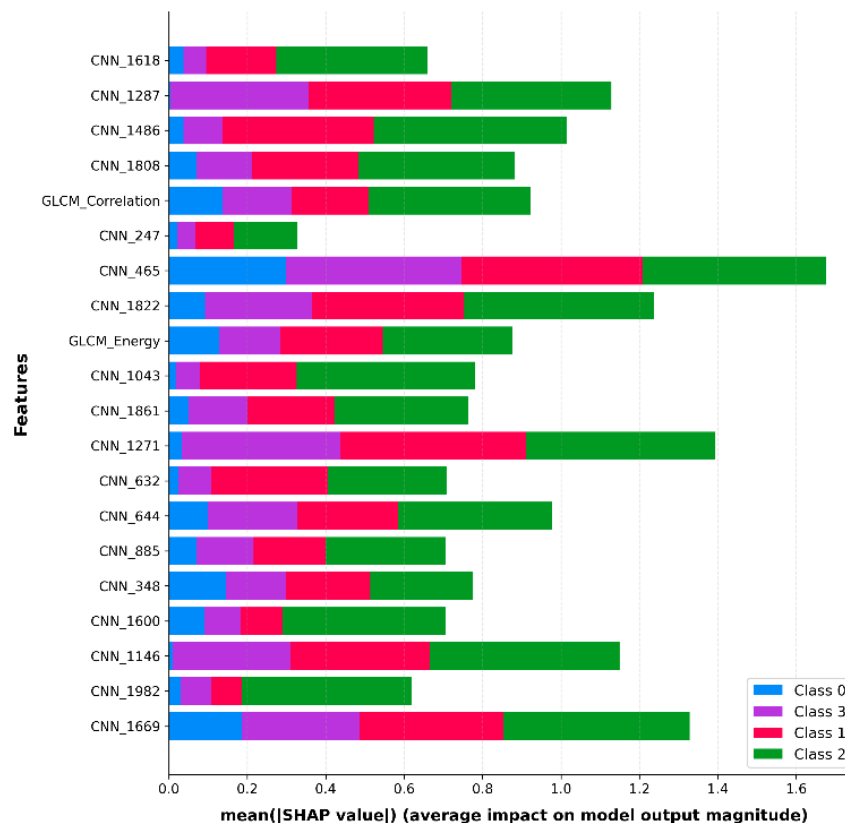


Figure 11. SHAP feature importance plot

5. CLINICAL UTILITY, DEPLOYMENT CONSIDERATIONS, AND FUTURE WORK

This proposed hybrid deep learning network offers a promising pathway for use in AI-assisted clinical radiology environments, as it strikes a strong balance among classification accuracy, computational efficiency, and explainability for routine practice. Additionally, by combining deep spatial embeddings with descriptive radiomic texture features, the proposed hybrid deep neural network provides a much-improved ability to discriminate multi-class pulmonary abnormalities while maintaining transparent diagnostic interpretations through Grad-CAM and SHAP-based explanatory methods. Lastly, given that the computational hardware requirements are moderate compared with those typically seen in workstation-level hospitals and telemedicine-based screening systems, the proposed hybrid deep learning framework is likely to have strong potential for use in radio-oncology and telemedicine-based screening systems.

The framework may be useful for radiologists in identifying pulmonary lesions at an early stage, as well as in the automated rounding and triaging of cases requiring assistance with thoracic CT interpretation. Radiologists will receive the Grad-CAM heat maps generated by the framework to show the anatomical location of potentially abnormal pulmonary areas. The use of SHAP will allow the radiologist to better understand which input features contributed to the AI's recommendation and will help enhance the radiologist's confidence in the AI's recommendations. This methodology could be incorporated into picture

archiving and communication system (PACS), cloud-based radiology systems, and intelligent clinical triaging systems that support large-scale lung cancer screening programs.

The clinical deployment advantages of the proposed framework include, but are not limited to the following:

- Improved transparency via an XAI approach that uses Grad-CAM and SHAP.
- Less computationally intensive compared to transformer-heavy architectures, facilitating real-world use in standard hospital computing environments.
- Greater flexibility in classifying multiple classes of histologically similar pulmonary pathologies.
- Potential future applications in AI-assisted radiology workflow, telemedicine, and off-site diagnostic testing.
- Compatible with future solutions for cloud-based and real-time diagnosis radiology platforms.

There are still numerous obstacles that impact generalizability outside of laboratory settings, despite promising research findings. For example, this research utilized chest CT datasets that are available to everyone but didn't test with other datasets or hospitals, nor were lesions marked by the radiologists as they would be in clinical practice. Also, most of the analysis was performed using 2D slices instead of intact volumes (3D), which may not provide a complete contextual understanding for complex disease patterns. Future research will extend this work to include larger databases from multiple locations treating a wide variety of patients to strengthen method reliability and deployability. The following list includes some advanced research avenues that could further leverage the proposed approach:

- Using attention-based multimodal fusion strategies (combining CNNs and transformers) to interpret image features with radiomic descriptors.
- Using 3D volumes of chest CT images for better characterization of lesion location.
- Using external validation datasets like LIDC-IDRI and COPDGene.
- Combining clinical metadata, pathology notes, and patient history for multimodal AI-supported diagnosis.
- Creating decision-support systems for borderline pulmonary abnormalities based on uncertainty and fuzzy logic.
- Studying federated learning and privacy-protecting healthcare AI models.
- Developing real-time optimization of PACS-integrated hospital screening systems and telemedicine.

In summary, the described framework represents a very promising basis for the development of future clinically interpretable and computationally efficient AI-assisted lung disease diagnosis systems that can support transparent and reliable chest CT analyses in real-world healthcare settings.

6. CONCLUSION

The research study described the quantitative hybrid deep learning framework for multi-class chronic lung disease identification and staging using a chest CT-based methodology that combines deep spatial embeddings with handcrafted radiomic texture descriptors. The proposed framework incorporates a pre-trained ResNet50 backbone, quantitative radiomic analysis, and an optimized XGBoost ensemble classifier to enhance classification robustness, explainability, and computational efficiency in AI-assisted chest CT imaging analysis. Experimental results showed that the hybrid framework outperformed texture-only radiomics, standalone CNN architectures, and transformer-like models ViT-B16 and Swin-Tiny across all classifications. The hybrid framework achieved an overall classification rate of 92.38%, precision of 0.9202, sensitivity of 0.9094, specificity of 0.9811, F1-score of 0.9146, and ROC-AUC of 0.9938, demonstrating excellent discrimination and consistent multi-class classification performance across all groups of pulmonary abnormalities. By integrating complementary spatial and statistical representations, the ability to distinguish histologically similar lung cancer subtypes (adenocarcinoma vs. squamous cell carcinoma) was significantly improved. Additionally, comparative benchmarking has shown that the proposed hybrid methodology strikes a balanced trade-off among classification performance, computational efficiency, and interpretability compared with transformer-heavy architectures. Furthermore, the use of dual-mode XAI via use of both Grad-CAM and SHAP has helped to improve transparency and clinical interpretability of the entire diagnostic process. The Grad-CAM heat maps have been useful in providing meaningful locational information to support the diagnosis of suspicious pulmonary regions, while SHAP analysis has quantified how both the deep embedding network and the radiomic descriptors contributed to the final classification predictions. Thus, the use of these explainability mechanisms increases physicians' trust and indicates the potential applicability of this methodology to future AI-assisted radiology workflows, telemedicine systems, and intelligent clinical decision support environments. The proposed hybrid framework demonstrates how combining deterministic radiomic texture characterization with deep semantic feature learning can provide a robust, clinically interpretable solution for analyzing lung disease on chest CT. Future work, including external

multisite validation, volumetric 3D CT analysis, multimodal clinical integration, adaptive fusion strategies, and implementation via federated learning-based deployment, will likely further enhance both the scalability and real-world clinical applicability of the proposed methodology.

ACKNOWLEDGMENTS

The authors express our profound gratitude for the accessibility to scholarly resources and institutional facilities afforded by Sir Padampat Singhania University, which significantly facilitated the fulfillment of this research endeavor.

FUNDING INFORMATION

The authors state no funding is involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Garima Jain	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓
Amit Kumar Goel	✓				✓		✓			✓		✓	✓	
Rahul Kumar	✓				✓		✓			✓		✓	✓	

C : **C**onceptualization

M : **M**ethodology

So : **S**oftware

Va : **V**alidation

Fo : **F**ormal analysis

I : **I**nterpretation

R : **R**esources

D : **D**ata Curation

O : **O**riginal Draft

E : **E**diting

Vi : **V**isualization

Su : **S**upervision

P : **P**roject administration

Fu : **F**unding acquisition

CONFLICT OF INTEREST STATEMENT

The author declares that there are no known conflicts of interest associated with this publication. There are no financial or personal relationships that could inappropriately influence or bias the content of this work.

INFORMED CONSENT

Not applicable, as this study did not involve human participants or personally identifiable information.

ETHICAL APPROVAL

Not applicable, as this study did not involve human or animal subjects and used only non-sensitive computational or public data.

DATA AVAILABILITY

The data that support the findings of this study are openly available in Kaggle at <https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images>.

REFERENCES

- [1] A. Hewa, J. Razmara, and J. Karimpour, "A hybrid HOG-LBP-CNN model with self-attention for multiclass lung disease diagnosis from CT scan images," *Computers*, vol. 15, no. 2, Feb. 2026, doi: 10.3390/computers15020093.
- [2] S. R. Kondamuri *et al.*, "Chest CT image-based lung disease classification – a review," *Current Medical Imaging Formerly Current Medical Imaging Reviews*, vol. 20, Nov. 2023, doi: 10.2174/0115734056248176230923143105.
- [3] N. Enshaei, "Deep learning-based medical image analysis for enhanced diagnosis and severity assessment of viral pneumonia," Ph.D. Thesis, Department of Information and Systems Engineering, Concordia University, 2025.
- [4] M. H. Saad, W. S. ElAraby, B. M. ElHalawany, and E. H. El-Shazly, "Optimized deep neural network for improved lung disease classification," *Neural Computing and Applications*, vol. 38, no. 10, May 2026, doi: 10.1007/s00521-026-12078-z.

- [5] T. S. Bascetin and I. Emiroglu, "Enhancing lung disease diagnosis: a high-performance hybrid deep learning framework for multi-class chest x-ray analysis," *International Journal of Imaging Systems and Technology*, vol. 36, no. 1, 2026, doi: 10.1002/ima.70262.
- [6] C.-H. Chien, S.-C. Chang, Y.-C. Chang, and Y.-C. Li, "Longitudinal multisource clinical model for early lung cancer risk stratification and screening," *BMJ Health & Care Informatics*, vol. 33, no. 1, Feb. 2026, doi: 10.1136/bmjhci-2025-101989.
- [7] S. Khan, M. N. Noor, I. Ashraf, M. I. Masud, and M. Aman, "Impact of CT intensity and contrast variability on deep learning-based lung-nodule detection: a systematic review of preprocessing and harmonization strategies (2020–2025)," *Diagnostics*, vol. 16, no. 2, Jan. 2026, doi: 10.3390/diagnostics16020201.
- [8] S. N. Yousafzai *et al.*, "A hybrid deep learning approach integrating CNN and transformer for lung cancer classification using CT scans," *Scientific Reports*, vol. 16, no. 1, Mar. 2026, doi: 10.1038/s41598-026-41161-7.
- [9] M. Vutukuri and P. S. Habibullah, "Detection and classification of lung cancer using sequential hybridization of CNN and RNN type architectures," *Frontiers in Big Data*, vol. 9, May 2026, doi: 10.3389/fdata.2026.1786859.
- [10] S. Zhao *et al.*, "CT image-based machine learning models for predicting blood eosinophil levels in acute exacerbation of chronic obstructive pulmonary disease," *BMC Pulmonary Medicine*, vol. 26, no. 1, Mar. 2026, doi: 10.1186/s12890-026-04219-w.
- [11] R. K. Jain, K. K. Hiran, R. Maheshwari, and Vaishali, "Lung cancer detection using machine learning algorithms," in *2023 International Conference on Computational Intelligence, Communication Technology and Networking (CICTN)*, Apr. 2023, pp. 516–521, doi: 10.1109/CICTN57981.2023.10141467.
- [12] A. Kodieswari, S. S. Sivaraju, R. Suguna, and T. Kanimozhi, "Enhanced lung cancer segmentation and classification using transformer-based deep learning and multimodal biomedical imaging," *International Journal of System Assurance Engineering and Management*, Mar. 2026, doi: 10.1007/s13198-026-03224-w.
- [13] Ş. Öztürk, M. Y. Turah, and T. Çukur, "HydraViT: adaptive multi-branch transformer for multi-label disease classification from chest x-ray images," *Biomedical Signal Processing and Control*, vol. 100, Feb. 2025, doi: 10.1016/j.bspc.2024.106959.
- [14] G. Li, J. Tan, R. Wang, and J. Wang, "Multi-class diagnosis of pulmonary hypertension in the presence of COPD via CT imaging and transformer-CNN fusion," *IEEE Access*, vol. 14, pp. 26647–26661, 2026, doi: 10.1109/ACCESS.2026.3663974.
- [15] J. Ramasamy and R. Doshi, "Machine learning in cyber physical systems for healthcare: brain tumor classification from MRI using transfer learning framework," in *Real-time applications of machine learning in cyber-physical systems*, 2022, pp. 65–76, doi: 10.4018/978-1-7998-9308-0.ch005.
- [16] P. Kumari and L. Goel, "Emerging computational intelligence-based techniques for lung cancer diagnosis and classification on chest CT scan images: a comprehensive survey," *Artificial Intelligence Review*, vol. 59, no. 2, Jan. 2026, doi: 10.1007/s10462-025-11374-9.
- [17] M. A. Thanoon, M. A. Zulkifley, M. A. A. M. Zainuri, and S. R. Abdani, "A review of deep learning techniques for lung cancer screening and diagnosis based on CT images," *Diagnostics*, vol. 13, no. 16, Aug. 2023, doi: 10.3390/diagnostics13162617.
- [18] B. Ozdemir, E. Aslan, and I. Pacal, "Attention-enhanced inceptionNeXt-based hybrid deep learning model for lung cancer detection," *IEEE Access*, vol. 13, pp. 27050–27069, 2025, doi: 10.1109/ACCESS.2025.3539122.
- [19] J. Sun *et al.*, "Detection and staging of chronic obstructive pulmonary disease using a computed tomography-based weakly supervised deep learning approach," *European Radiology*, vol. 32, no. 8, pp. 5319–5329, 2022, doi: 10.1007/s00330-022-08632-7.
- [20] G. K. Subramanyam, K. Srinivas, V. V. R. Indugu, D. S. Gondhi, and S. K. G. Subbammagari, "Segmentation-guided hybrid deep learning for pulmonary nodule detection and risk prediction from multi-cohort CT images," *Diseases*, vol. 14, no. 1, Jan. 2026, doi: 10.3390/diseases14010021.
- [21] H. Sui, Z. Mo, Y. Wei, F. Shi, K. Cheng, and L. Liu, "Diagnosis and severity assessment of COPD based on machine learning of chest CT images," *International Journal of Chronic Obstructive Pulmonary Disease*, vol. 20, pp. 2853–2867, Aug. 2025, doi: 10.2147/COPD.S528988.
- [22] H. Alsurayhi, "Machine learning classifiers for chronic obstructive pulmonary disease assessment using lung CT data," Ph.D. Thesis, Department of Electrical and Computer Engineering, The University of Western Ontario, 2024.
- [23] R. Kumar, C.-T. Pan, Y.-M. Lin, S. Y.-Ling, T.-S. Chung, and U. G. S. Janesha, "Enhanced multi-model deep learning for rapid and precise diagnosis of pulmonary diseases using chest x-ray imaging," *Diagnostics*, vol. 15, no. 3, Jan. 2025, doi: 10.3390/diagnostics15030248.
- [24] H. Tu *et al.*, "Improving lung cancer risk prediction using machine learning: a comparative analysis of stacking models and traditional approaches," *Cancers*, vol. 17, no. 10, May 2025, doi: 10.3390/cancers17101651.
- [25] L. Kamala and K. G. Mohan, "An efficient hybrid artificial intelligence framework for lung cancer classification using CT images," *Scientific Reports*, vol. 16, no. 1, Dec. 2025, doi: 10.1038/s41598-025-31432-0.
- [26] M. Hosny, I. A. Elgendy, and M. A. Albashrawi, "Multi-modal deep learning for lung cancer detection using attention-based Inception-ResNet," *IEEE Access*, vol. 13, pp. 123630–123648, 2025, doi: 10.1109/ACCESS.2025.3588407.
- [27] C. Parida, S. Bilgaiyan, S. Sagnika, S. Roy, and A. Pandey, "3D fusionNet for synthetic CT-based lung cancer segmentation," *Scientific Reports*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-29703-x.
- [28] M. Al-Duhayyim, M. A. Aldawsari, A. Ismail, and M. M. Enam, "Interpretable hybrid ensemble with attention-based fusion and EA00-GA optimization for lung cancer detection," *Scientific Reports*, vol. 16, no. 1, 2026, doi: 10.1038/s41598-026-37187-6.
- [29] M. K. Faizi *et al.*, "Deep learning-based lung cancer classification of CT images," *BMC Cancer*, vol. 25, no. 1, Jul. 2025, doi: 10.1186/s12885-025-14320-8.
- [30] S. D. Almeida *et al.*, "Prediction of disease severity in COPD: a deep learning approach for anomaly-based quantitative assessment of chest CT," *European Radiology*, vol. 34, no. 7, pp. 4379–4392, 2024, doi: 10.1007/s00330-023-10540-3.
- [31] A. Zaernia, M. Youseffi, L. Parisi, R. A.-Alhameed, and R. Ma, "A novel artificial intelligence-driven technique for enhancing medical imaging techniques to detect non-small cell lung cancer," *Proceedings - European Workshop on Visual Information Processing, EUVIP*, 2024, doi: 10.1109/EUVIP61797.2024.10772826.
- [32] L. L. Nahmatwilla and A. M. Ali, "ResNet-SEMRes: a deep learning framework for primary lung cancer classification using CT scan image," *Indian Journal of Radiology and Imaging*, Feb. 2026, doi: 10.1055/s-0045-1815732.
- [33] A. Kumar and N. Mahendran, "U-VQVAE-CTLesionNet: a generalized deep learning framework for multi-organ lesion detection and segmentation in medical imaging," *Computational Intelligence*, vol. 42, no. 1, Feb. 2026, doi: 10.1111/coin.70168.
- [34] S. Wankhade and S. Vigneshwari, "A deep ensemble model for the detection and classification of lung cancer in clinical images," in *Detection Systems in Lung Cancer and Imaging, Volume 2*, Bristol, United Kingdom: IOP Publishing, 2026, doi: 10.1088/978-0-7503-3359-7ch1.
- [35] I. H. de Córdoba *et al.*, "Transformation to neuroendocrine phenotype in non-small-cell lung carcinoma: a literature review," *International Journal of Molecular Sciences*, vol. 26, no. 11, 2025, doi: 10.3390/ijms26115096.
- [36] M. A. Baqir *et al.*, "A lightweight CNN for enhanced non-small cell lung cancer classification using CT scan image," *Scientific Reports*, vol. 16, no. 1, Mar. 2026, doi: 10.1038/s41598-026-41401-w.

BIOGRAPHIES OF AUTHORS



Garima Jain    is a research scholar in the Department of Computing and Informatics at Sir Padampat Singhania University (SPSU), Udaipur, Rajasthan, India. She has more than 12+ years of teaching experience. She has completed her M.Tech. in Computer Science and Engineering from Pacific University, Udaipur, Rajasthan, in 2013 and B.Tech. in Information Technology from Rajasthan University, Jaipur, in 2008. She has presented papers in international conferences and international journals. Her research interests include computer networks, machine learning, and artificial intelligence. She can be contacted at email: garima.phd2023@spsu.ac.in.



Dr. Amit Kumar Goel    is a distinguished academician and administrator with over 27 years of experience, comprising 24 years in teaching and 3 years in industry. He holds a Ph.D. in Engineering from Birla Institute of Technology, Mesra, along with M.Tech. and B.E. degrees in Computer Science and Information Technology. He currently serves as dean, School of Engineering and Sciences at Sir Padampat Singhania University, where he leads strategic planning, curriculum development, and academic excellence initiatives. He has previously held key leadership positions at Galgotias University, SGT University, and other reputed institutions, overseeing thousands of students and faculty members. His contributions include the design of industry-oriented curricula, development of research ecosystems, and a central role in the National Assessment and Accreditation Council (NAAC) and National Board of Accreditation (NBA) accreditation processes. He has authored more than 60 research papers in Scopus-, Emerging Sources Citation Index (ESCI)-, and Science Citation Index (SCI)-indexed journals and holds multiple patents in artificial intelligence, internet of things, and cybersecurity. A dedicated mentor, he has guided several Ph.D. scholars, collaborated on international research projects, and actively contributed to the organization of conferences, workshops, and technical events, establishing himself as a key figure in academic leadership and innovation. He can be contacted at email: dean.fci@spsu.ac.in.



Dr. Rahul Kumar    is working as an associate professor in the Faculty of Computing and Informatics, Sir Padampat Singhania University, Udaipur, India. He has more than 11+ years of teaching and research. He has completed his Ph.D. from Dr. K.N Modi University (DKNMU), Newai, Tonk, Rajasthan in 2022. He holds an M.Tech. in Computer Science and Engineering from Suresh Gyan Vihar University, Jaipur, Rajasthan. He has completed B.Tech. in Computer Science and Engineering from ICFAI University, Dehradun, Uttarakhand. His leadership is evident in his transformative tenure as deputy controller of examinations at Dr. K. N. Modi University (2019–2023), where he elevated academic standards, and as Ph.D. co-coordinator at Amity University, adeptly guiding future researchers. His expertise fuels innovation in database management systems, machine learning, and computer networks. A prolific scholar, he has contributed more than 30 high-quality research papers to prestigious international and national journals and conferences, with research interests focused on machine learning and networking. His impactful work includes the SCI- indexed publication and quote; hybrid approach of cotton disease detection for enhanced crop health and yield quote; (IEEE Access, 2024), underscoring his dedication to advancing cutting-edge solutions and academic excellence. He has 3 patents filed and one copyright has been granted. He is a life member of IAENG. He can be contacted at email: rahul.kumar@spsu.ac.in.