ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Academic article

Morphological Profiling of Lung Cancer Through Explainable Machine Learning

Karuna Tanthanawarakun[1], and Prompong Sugunnasil[2*]

[1] Faculty of Engineering, Chiang Mai University Chiang Mai, Chiang Mai, 50200, Thailand
[2] College of Arts, Media, and Technology, Chiang Mai University, Chiang Mai, 50200, Thailand
* Correspondence: prompong.sugunnasil@cmu.ac.th

Citation:
Tanthanawarakun, K.; Sugunnasil, P. Morphological profiling of lung cancer through explainable machine learning. *ASEAN J. Sci. Tech. Report.* **2025**, *28*(5), e257888. https://doi.org/10.55164/ajstr.v28i5.257888.

Article history:
Received: February 14, 2025
Revised: August 19, 2025
Accepted: September 1, 2025
Available online: September 14, 2025

Abstract: Lung cancer continues to be the predominant cause of cancer-related mortality globally, representing 11.6% of all cancer cases as reported by WHO, with approximately 2.2 million new diagnoses and 1.8 million fatalities recorded in 2020. For lung cancer diagnosis, Computed Tomography (CT) imaging serves as a critical tool in identifying both solid and subsolid ("ground glass") nodules. Although CT image segmentation has demonstrated significant clinical value, healthcare practitioners require a comprehensive understanding of the underlying algorithmic mechanisms to ensure diagnostic precision. This research investigates the application of interpretable machine learning methods for feature extraction from lung CT imaging data. We conduct a comparative analysis between transparent and opaque classification algorithms utilizing a comprehensive dataset comprising 1,229 normal and 1,010 abnormal pulmonary CT scans. The processed data undergoes evaluation using both interpretable models (including Logistic Regression, Decision Trees, and K-Nearest Neighbor) and black-box models (such as Multi-Layer Perceptrons, Convolutional Neural Networks, and Support Vector Machines). Our findings indicate that interpretable algorithms consistently outperform their black-box counterparts across multiple metrics. The evaluation framework incorporates accuracy, F1 score, precision, recall, computational efficiency, and resource utilization measurements. Results demonstrate exceptional classification accuracy for pulmonary malignancy detection while preserving explanatory capability, thereby providing clinicians with both practical and transparent diagnostic assistance. This investigation contributes to the development of accountable artificial intelligence systems for deployment in mission-critical healthcare environments.

Keywords: CT Scan image; lung cancer image; explainable AI; encoding; machine learning

1. Introduction

Lung cancer persists as one of the leading causes of cancer-related death in both men and women globally [1]. Two main categories of lung cancer are identifiable: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) Transformation from NSCLC to SCLC has been reported in clinical studies [2]. NSCLC represents approximately 85% of all lung cancer cases [3], encompassing various histological subtypes including adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. Alarmingly, 70% of patients

receive diagnoses at advanced stages, and 15% of cases remain undetected until reaching late stages [4]. The mortality incidence (MI), representing the ratio between fatalities and total diagnosed patients, reaches a staggering 80%.Consequently, developing precise detection techniques for lung cancer has become critically important.

Computed tomography scanning (CT scan) functions as a diagnostic tool that synthesizes multiple computerized views to create detailed internal visualizations of bodily structures. The scan yields valuable information regarding various pulmonary conditions, including pneumonia, cancer, vascular thrombosis, smoking-induced damage, and COVID-19 infection. With increasing reliance on CT scans, radiologists face mounting pressure to detect, analyze, and quantify these images in real-time despite resource constraints. Computerized techniques offer potential solutions to reduce resource demands, particularly in terms of time and personnel requirements, in lung cancer screening protocols. Various approaches, including Machine Learning [5], the Synthetic Minority Over-sampling Technique (SMOTE) [4], and Deep Learning [6], have been proposed to enhance image analysis processes.

Given that analytical outcomes directly impact human lives, simplistic binary classifications (such as benign or malignant) prove insufficient for justifying the utilization of information[7-8]. Despite achieving accurate results, the opaque nature of many machine learning algorithms remains problematic, with decision-making processes poorly understood. Consequently, medical applications of machine learning necessitate additional supporting evidence to validate methodological approaches. The interpretability and explainability of machine learning algorithms for CT scan image analysis, therefore, emerge as crucial drivers in medical image processing [9]. The benefits of explainable machine learning extend beyond mere accountability to include model refinement and improvement.

In this research, we propose a comprehensive comparative study between conventional machine learning approaches and explainable machine learning methodologies for feature extraction from lung CT images. Our primary objective is to classify images, determining whether specimens exhibit benign or malignant characteristics. The dataset comprises paired associations between lung CT images and corresponding feature vectors, with each vector element representing quantitative image information, such as the quantities of nodules of specific sizes in designated lung regions. Features undergo exhaustive generation based on predefined criteria, including nodule dimensions and anatomical location [10]. SCLC displays distinctive characteristics from NSCLC, notably exhibiting more aggressive growth patterns and a greater propensity for widespread metastasis [11]. CT-based lung cancer detection methodologies categorize findings according to cellular structure and anatomical location, providing detailed cross-sectional visualizations [12]. While CT scanning offers valuable diagnostic capabilities, including screening program support, three-dimensional imaging, and biopsy guidance, limitations persist regarding ionizing radiation exposure, false-positive results, and dependence on expert radiological interpretation [13].

Our research introduces an innovative data engineering technique designed to enhance classification efficiency for lung cancer diagnosis using CT scan images [14]. We employ explainable machine learning methods to capture distinctive imaging features and conduct rigorous comparative analysis of data encoding using interpretable classification algorithms [15]. Our findings consistently demonstrate that explainable approaches outperform non-interpretable alternatives, yielding superior diagnostic accuracy for pulmonary malignancies [16].

2. Materials and Methods

2.1 Interpretability and Explainable Machine Learning

The field of explainable machine learning addresses the critical need for algorithmic transparency, particularly in healthcare applications where understanding model decisions directly impacts patient care. While conventional machine learning often prioritizes performance metrics, explainability focuses on providing human-comprehensible justifications for model predictions [17]. Interpretability encompasses multiple dimensions, including algorithm transparency, feature relevance, and decision logic visualization. As noted by Doshi-Velez and Kim, interpretability in machine learning can be defined as "the ability to explain or present in understandable terms to a human" [9]. This definition emphasizes the human-centric nature of

explainability, recognizing that different stakeholders—such as clinicians, patients, and researchers—may require different types of explanations. For medical imaging applications, model interpretability facilitates trust establishment among healthcare professionals by revealing the reasoning behind diagnostic suggestions. This transparency becomes particularly crucial when algorithms influence treatment decisions with potentially life-altering consequences [8]. Our research employs interpretability methods that strike a balance between explanatory clarity and diagnostic accuracy, addressing the unique challenges posed by pulmonary nodule classification.

2.2 CT Image Feature Analysis

Computed tomography feature analysis involves the systematic extraction and quantification of imaging characteristics that correlate with pathological findings. In lung cancer assessment, these features capture crucial information about nodule properties, including size, shape, margin characteristics, and internal density patterns [10]. Pulmonary nodules in CT images typically present as approximately spherical opacities with diameters of less than 30mm. These can be classified based on their density profiles as solid, part-solid, or ground-glass opacities, each carrying different clinical implications regarding malignancy potential [11]. Feature extraction algorithms quantify these visual characteristics into numerical descriptors, making them suitable for algorithmic processing. Our methodology employs a comprehensive feature extraction approach that captures first-order statistics (intensity distribution measures), shape metrics (morphological descriptors), texture features (spatial pixel relationship patterns), and wavelet-based characteristics [4]. By transforming visual data into quantitative measurements, these techniques enable precise comparison between cases and facilitate reproducible assessment of subtle imaging findings that may elude visual detection.

2.3 Detection of Lung Cancer Disease Based on CT Image

Lung cancer detection from CT images presents a complex diagnostic challenge, requiring the precise identification of potentially malignant pulmonary nodules against a background of typical anatomical structures. The process begins with image preprocessing and segmentation to isolate lung parenchyma from surrounding tissues [12]. Computer-aided detection systems have evolved significantly, progressing from basic rule-based algorithms to sophisticated deep learning architectures. Early detection methods demonstrated limited sensitivity, while contemporary approaches achieve substantially improved performance. As demonstrated by Li et al., automated systems utilizing deep features extracted from autoencoders can achieve accuracy levels of 75.01% with a sensitivity of 83.35% [18]. Research by Faisal and colleagues explored various machine learning approaches for nodule classification, finding that ensemble methods achieved impressive accuracy ratings exceeding 99% under controlled testing conditions [19]. Additionally, multilayer architectures employing convolutional networks with fusion-based classification have demonstrated detection rates above 99% with minimal false positives per image [20]. Our approach incorporates explainable feature engineering techniques that enhance diagnostic transparency while maintaining high classification performance. This methodology addresses the fundamental challenge of providing interpretable results that can meaningfully inform clinical decision-making beyond simple binary classification outputs.

2.4 Algorithmic Transparency in Clinical Decision Support

Machine learning algorithms in medical image analysis can be categorized along a spectrum of interpretability, from inherently transparent "white box" models to complex "black box" systems [21]. White box approaches—including linear regression, decision trees, and logistic regression—provide clear relationships between input features and predictions. Conversely, black box models, such as deep neural networks and complex ensembles, often achieve superior accuracy but operate through mechanisms that are resistant to straightforward human understanding. This tradeoff between transparency and performance presents significant challenges in clinical implementation. As demonstrated by studies utilizing the IQ-OTH/NCCD dataset, support vector machines with polynomial kernels achieved 89.89% classification accuracy, substantially outperforming linear alternatives (82.02%) [22]. This performance differential highlights the complexity of nodule classification tasks while emphasizing the need for sophisticated modeling approaches. The emerging field of Explainable AI (XAI) addresses transparency limitations through

techniques that illuminate decision processes, including feature importance ranking, partial dependence analysis, and model-agnostic interpretation methods [23]. Notably, comparative studies have shown that decision tree algorithms can achieve remarkable performance (93.24% accuracy) even without extensive preprocessing, highlighting the potential value of inherently interpretable models in clinical contexts [16].
Our methodology synthesizes these insights by implementing a balanced approach that enhances model transparency through feature engineering while preserving classification performance. This strategy aligns with the requirements for medical AI systems that provide not only accurate results but also explainable outputs that integrate meaningfully into clinical workflows.

2.5 Methods

This section outlines the specifics of the proposed methodology, detailing the processes and methods used in conducting the research. This encompasses the study's design.

Overview of the proposed architecture

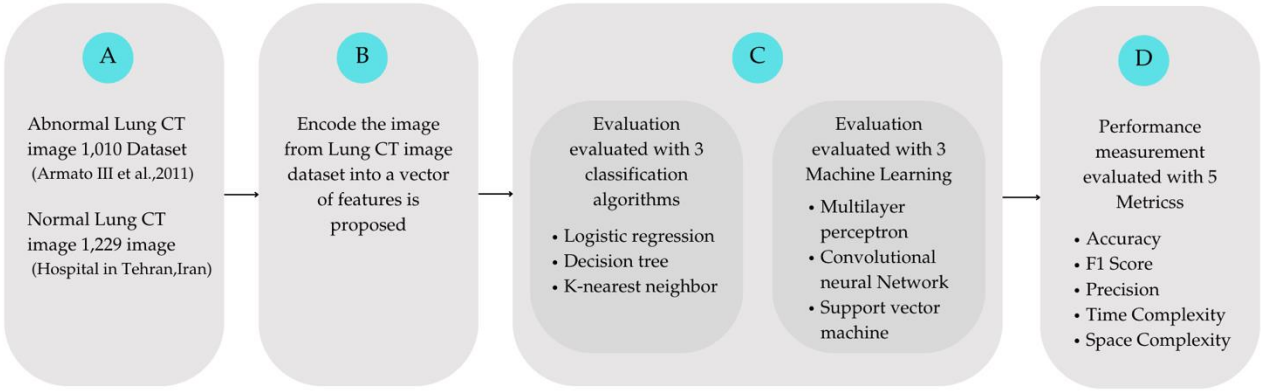

Figure 1. The overall architecture of the experiment.

Feature engineering plays a crucial role in analyzing lung cancer using CT scans. In this intricate process, we transform the raw data of the images into a structured format by converting salient features into pixel values that a computational model can interpret and understand. The first step involves identifying and extracting nodules, which are small, rounded opacities within the pulmonary tissue. Additionally, we isolate solid masses, which are denser, irregular growths that could indicate more advanced pathology. By distilling these complex visual cues into a set of features, we enable the algorithm to identify patterns and facilitate the accurate diagnosis of lung cancer. (Figure 1) presents an overview of the proposed architecture. (a) The input consists of a lung CT image dataset that includes normal and non-normal lung cancer CT images. We performed an image pre-processing to reduce biased information and prevent overfitting. The input consists of two types of images with the same format, which have been reshaped, resized, and reformatted. (b) This study proposes a process for encoding images from the lung CT image dataset into a vector of features. In the process of feature engineering for lung cancer detection using CT scans, we convert image-based features into quantifiable pixel values. This involves the meticulous extraction of both nodules and solid masses, transforming them into discernible features for further analysis (c) In addition, we compared the proposed image classification process with other baseline models using two types of evaluations: Explainable machine learning (i.e Logistic regression, Decision tree and K-nearest neighbor) and unexplainable algorithm (i.e. Multilayer perceptron, Convolutional Neural Network and Support vector machine). (d) The final step is performance measurement using five classification metrics: Accuracy, F1 Score, Precision, Time Complexity, and Space Complexity. We propose that the size and location of the nodule affect the outcome.

2.5.1 Collection of datasets

We collected the Lung Image Database Consortium Image Collection (LIDC-IDRI) [24] Figure 2) ,which consists of 244,527 diagnostic and lung cancer screening thoracic computed tomography (CT) scans fro 1,010

patients. Additionally, we collected normal CT images from large public Covid-19 (SARS-CoV-2) lung CT scan datasets, which include 1,229 negative cases (i.e., both normal and non-COVID images). The data is available as 512x512px PNG images and has been collected from real patients in the radiology centers of teaching hospitals in Tehran, Iran [25]. In conclusion, we have 1,224 cases for normal lung CT scans and 1,010 cases for lung with cancer.

2.5.2 Pre-Processing Data and Feature Extraction

This experiment requires data pre-processing because raw CT scan files are not directly applicable to the classification process. Since the images originate from different sources, we standardized them through reformatting, recoloring, and resizing to reduce potential bias in the dataset. This process helps ensure consistency across input images during the feature extraction phase. For the first dataset, which comprised non-normal images, we used DICOM files—an established standard for medical imaging that healthcare providers and equipment manufacturers have widely adopted. In this study, we focused exclusively on the lung region by removing unnecessary parts, such as the heart and background structures.

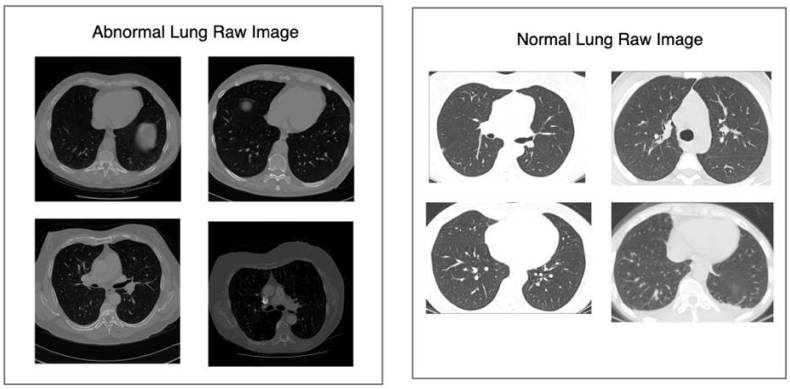

Figure 2. The example of the Abnormal Lung and Normal Lung Raw Dataset.

The CT scan lung images in this dataset exhibit multi-dimensional. Our approach involves initially converting these images from a 3D format to a 2D representation and subsequently performing a precise cropping operation to extract the lung region. During the lung region extraction process, we calculated the area encompassing most of the lung region, ensuring that it encapsulates the most relevant portion of the complete image. Subsequently, we reformatted the entire image into a PNG file as part of the preprocessing procedure. The second dataset comprises normal lung images, consisting of 1,229 negative cases. These normal lung images are stored as 2D PNG files with dimensions of 415x345 pixels, which differ from the size of the non-normal lung images. To maintain uniformity with the non-normal lung images, we conducted resizing and zoom-out operations. Following this, we utilized the OpenCV (cv2) library for preprocessing and executed cropping operations to isolate the lung region accurately.

2.5.2.1 Data Pre-processing

To ensure consistency, interpretability, and computational efficiency, two distinct pre-processing pipelines were employed in this study. These pipelines were tailored to support the different experiments. The aim was to transform raw lung CT scan images, both normal and abnormal, into structured feature sets suitable for training and evaluating machine learning models. In the first experimental approach, CT scan images were transformed into raw pixel-level vectors, preserving the grayscale intensity values directly from the image arrays. Non-Normal Dataset: DICOM files were subjected to lung segmentation using masking techniques to isolate lung regions see (Figure 3). Each image slice was then binarized, converted to grayscale, and flattened into one-dimensional pixel vectors(Figure 4). These vectors were exported as CSV files for model input.

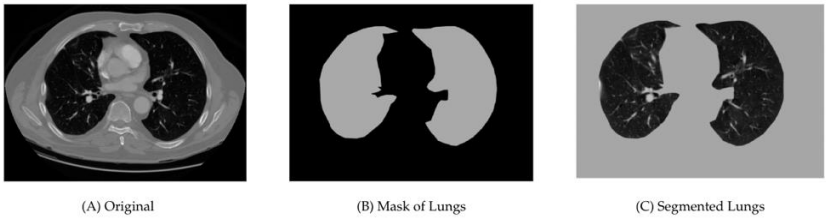

Figure 3. The Example of Abnormal Lung and Normal Lung Masking Process.

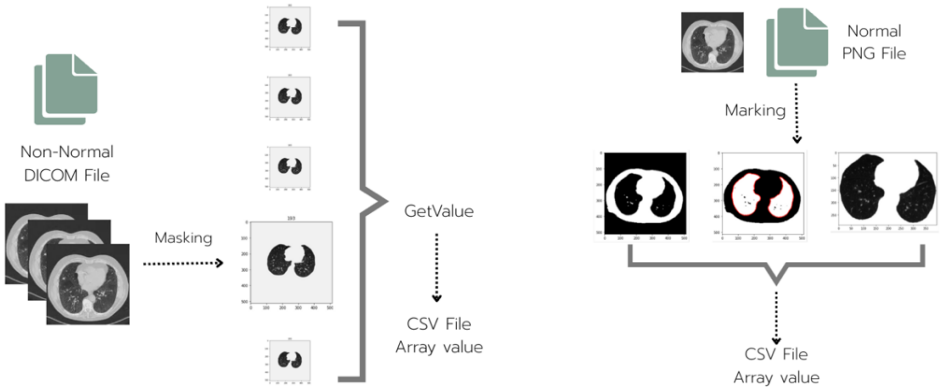

Figure 4. AbNormal Lung and Normal Feature Annotation and Vector Encoding Process.

The normal dataset is composed of PNG image files underwent manual region-of-interest (ROI) annotation to isolate lung regions. The annotated images were processed in the same manner as the non-normal dataset, ensuring consistency in vector representation (Figure 4).

The resulting feature vectors served as direct input for both explainable models (Logistic Regression, Decision Tree, KNN) and traditional machine learning models (CNN, MLP, SVM). The second dataset design employed an analytical approach that incorporated advanced morphological feature engineering techniques to extract measurable attributes from pulmonary radiographs. The methodology involved a series of systematic procedures executed in sequential order. Employing the OpenCV (cv2) computational library, comprehensive contour analysis was conducted to identify and differentiate various pulmonary structures. This analytical procedure facilitated the detection of potential malignant and benign nodules, parenchymal tissue characteristics, vascular elements, and nodule candidates. The implementation of advanced computer vision algorithms enabled the precise delineation of these structures based on their radiographic density, morphology, and spatial relationships.

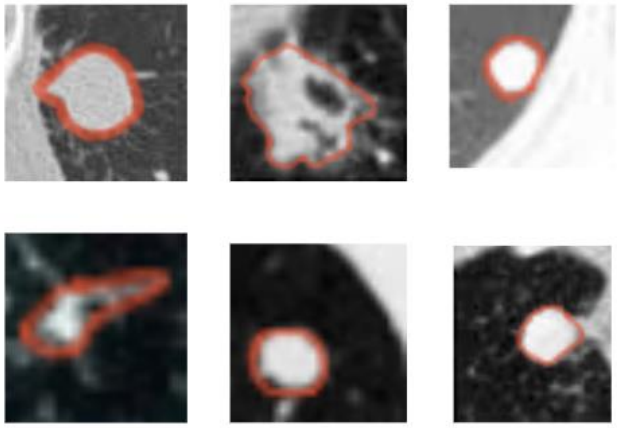

Figure 5.The Example of lung nodules.

Pulmonary radiographs were partitioned into four anatomically distinct quadrants, as illustrated in (Fig5). This segmentation strategy permitted localized analysis of morphological characteristics, enhancing the spatial specificity of feature extraction.

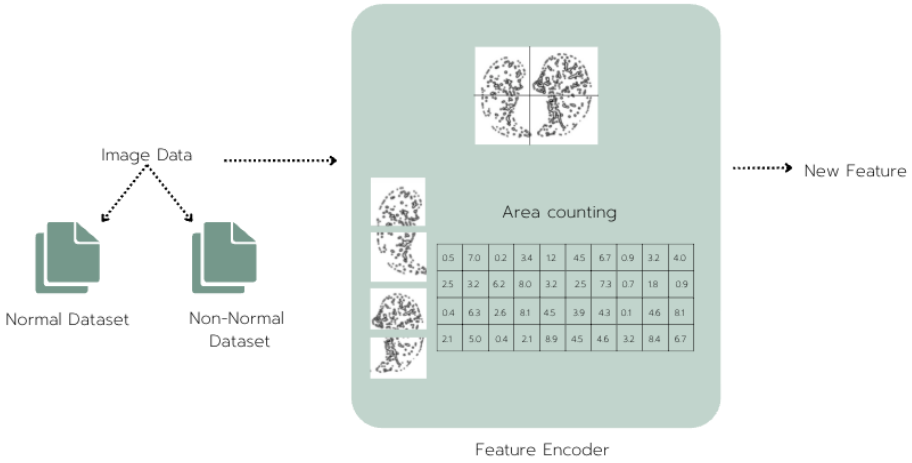

Figure 6. Shows the contributing zone of the CT scan image.

Following the visualization and statistical analysis of contour area distributions (Figure 7), optimal threshold parameters were established within the range of 0 to 70,000 spatial units. These thresholds were implemented with systematic increments of 5 units across the entire range, establishing discrete categorical boundaries for subsequent feature quantification.

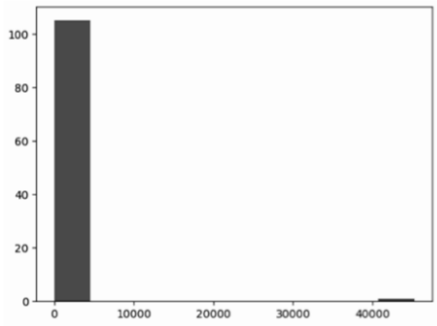

Figure 7. Graphical representation of contour result data distribution across the analyzed pulmonary specimens.

A comprehensive feature mapping framework was developed wherein contour areas were enumerated within each predefined threshold range. This methodical approach facilitated the transformation of complex morphological characteristics into a structured, high-dimensional feature space, as illustrated in (Figure 6). The features extracted from both experimental pipelines—the raw vector dataset (Experiment 1) and the morphologically encoded dataset (Experiment 2)—were utilized in the classification phase to evaluate model performance. Specifically, both feature sets were applied to two categories of machine learning models. Explainable Models, including Logistic Regression, Decision Tree, and K-Nearest Neighbor (KNN), which provide transparent decision-making processes suitable for clinical interpretation. Traditional or Non-Interpretable Models, such as Multi-layer Perceptron (MLP), Support Vector Machine (SVM), and Convolutional Neural Network (CNN), which leverage more complex computational architectures for pattern recognition.

Explainable Machine Learning (XAI): Explainable Artificial Intelligence (XAI) refers to transparent models that have an intrinsic architecture that satisfies at least one of the three transparency dimensions discussed in a previous section.

1.1) Decision tree

Decision trees belong to the family of supervised learning algorithms that organize decision sequences into intuitive tree-like structures (Figure 8). The architecture begins with the root node, positioned at the apex of the tree, which encompasses the entire training dataset and serves as the starting point for analysis [15]. As the algorithm progresses, it evaluates internal nodes where specific attributes are selected based on their ability to effectively separate data into more uniform subgroups according to predetermined criteria. The process then advances to data partitioning, where the algorithm divides information based on specific thresholds or categories of the selected feature. This division creates distinct pathways representing different decision outcomes. The algorithm continues this pattern recursively, generating new decision points and branches as it navigates deeper into the tree hierarchy, until it reaches termination conditions such as maximum depth or minimum improvement thresholds. At the tree's endpoints, leaf nodes represent final classification outcomes where predictions are assigned to new data instances. In classification applications, these terminal nodes contain specific category labels that are determined by the unique sequence of decisions encountered while traversing from the root to the respective leaf position.

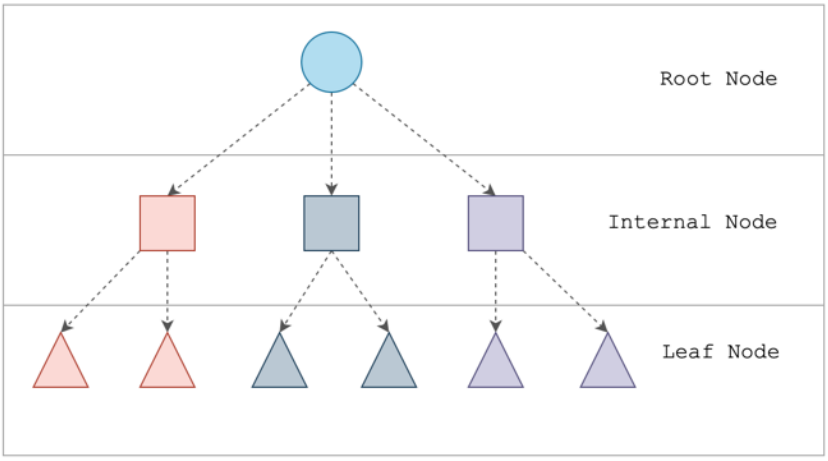

Figure 8. Structure of the Decision tree machine learning model.

Recent research has focused on evaluating various machine learning approaches for early detection of pulmonary malignancies [15]. A comparative analysis examined multiple classification algorithms, including probabilistic methods (Naïve Bayes), boundary-based approaches (Support Vector Machine), connectionist systems (Neural Networks and Multi-Layer Perceptron), hierarchical models (Decision Tree), and gradient-enhanced variants (Gradient-boosted Decision Tree). Additionally, ensemble methods combining multiple algorithms through Random Forest and Majority Voting techniques were assessed, with particular attention to combinations that integrate the capabilities of MLP, GBT, and SVM. Among all methodologies evaluated, the Gradient Boosted Tree demonstrated superior performance characteristics, achieving a classification accuracy of 90% across testing scenarios. This finding highlights the potential value of boosted decision tree architectures in clinical diagnostic applications, particularly for the early identification of lung cancer, where accurate classification directly impacts treatment planning and patient outcomes.

1.2) K-nearest Neighbors

The K-nearest Neighbors (KNN) algorithm functions as a supervised learning classifier that utilizes proximity relationships to determine the classification of data points. While capable of addressing both regression and classification problems, KNN primarily serves as a classification method, operating on the fundamental principle that similar data instances naturally cluster together in feature space [26] (Figure 9). In their research, the investigators implement an enhanced k-Nearest-Neighbors approach supplemented with genetic algorithm optimization for feature selection. This strategic combination effectively reduces dimensionality while simultaneously improving classifier performance. The methodology targets an explicitly accurate diagnosis of disease progression stages in lung cancer patients. To maximize the effectiveness of their

proposed system, the research team conducts systematic experimentation to identify the optimal value for the k parameter—the number of neighboring points considered during classification. Through rigorous testing on a comprehensive lung cancer dataset, their approach demonstrates superior accuracy in disease stage determination compared to conventional methods. The genetic algorithm component plays a crucial role in identifying the most diagnostically relevant features while eliminating redundant or noisy variables that could potentially degrade classification performance. This feature selection process not only improves accuracy but also enhances computational efficiency by reducing the dimensionality of the feature space, allowing for more rapid diagnosis without sacrificing precision. The methodology's success in accurately classifying different disease stages highlights its potential clinical value, particularly in settings where timely and accurate staging directly impacts treatment planning and patient management strategies.

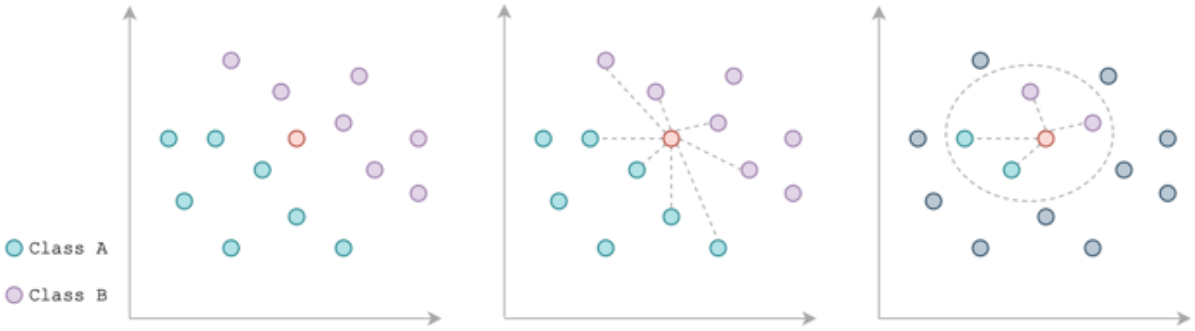

Figure 9. Structure of K-nearest Neighbors (KNN) machine learning model.

1.3) Logistic Regression

Logistic Regression represents a versatile statistical methodology widely applied in classification scenarios and predictive modeling. This approach effectively estimates occurrence probabilities for specific events within datasets containing multiple independent variables. When applied to binary classification problems—such as differentiating between non-COVID and COVID cases—the model generates probability outputs constrained between 0 and 1. A fundamental aspect of this technique involves the logit transformation, which converts odds ratios (the success probability relative to the failure probability) into a more mathematically tractable form. Research by Chen et al. [27] examined various classification methods for lung cancer prediction, with Logistic Regression selected as one of the key comparative algorithms. Their investigation utilized comprehensive datasets from the UCI Machine Learning Repository, incorporating a diverse set of predictive attributes. These features encompassed demographic factors (Age, Gender), environmental exposures (Air Pollution, Dust Allergy, Occupational Hazards), lifestyle variables (Alcohol Use, Balanced Diet, Obesity, Smoking, Passive Smoking), genetic predisposition (Genetic Risk, Chronic Lung Disease), and symptom indicators (Chest Pain, Hemoptysis, Fatigue, Weight Loss, Dyspnea, Wheezing, Dysphagia, Digital Clubbing, Respiratory Infections, Persistent Cough, and Sleep Disturbances). Within their analytical framework, outcome classifications employed a three-tier labeling system: malignant tumors received designation '2', benign tumors were labeled '1', and tumor-free healthy individuals were assigned '0'. Performance evaluation revealed that Logistic Regression achieved 66.7% accuracy in correctly classifying cases across these categories. Notably, while Logistic Regression demonstrated reasonable predictive capability, Support Vector Machine (SVM) algorithms consistently outperformed other classification methodologies within the comparative analysis. This performance differential highlights the complexity of lung cancer prediction and suggests that while Logistic Regression offers interpretability advantages, more sophisticated algorithms may provide enhanced predictive power for complex medical classification tasks.

2) Traditional Machine Learning

Unexplainable machine learning refers to a model that lacks transparency or interpretability. This kind of machine learning model is complex and challenging to understand.

2.1) Multi-layer Perceptron

The Multilayer Perceptron (MLP), as described by Tamarindo et al. in 2018, emerges as a crucial component within the realm of feed-forward neural networks. Comprising three distinct layers – an input layer, a hidden layer, and an output layer – the MLP's architecture is eloquently depicted in (Figure 10). The primary role of the input layer is to receive and process the incoming signals. On the other hand, the output layer assumes responsibility for pivotal functions such as classification and prediction. Notably, MLP's innovative prowess lies in its capacity to accommodate an infinite sequence of hidden layers, discretely interposed between the input and output layers. This network structure, characterized by a unidirectional flow of data from the input to the output layer, closely aligns with the feed-forward network paradigm. The training regimen used to equip the nodes within the MLP with the requisite knowledge is underpinned by the backpropagation learning technique. Through this method, MLPs refine their model parameters, thereby enhancing their capability to generalize and make accurate predictions. Notably, MLPs exhibit a remarkable capacity to address problems that defy linear separability, rendering them exceptionally adept at approximating a wide spectrum of continuous functions.

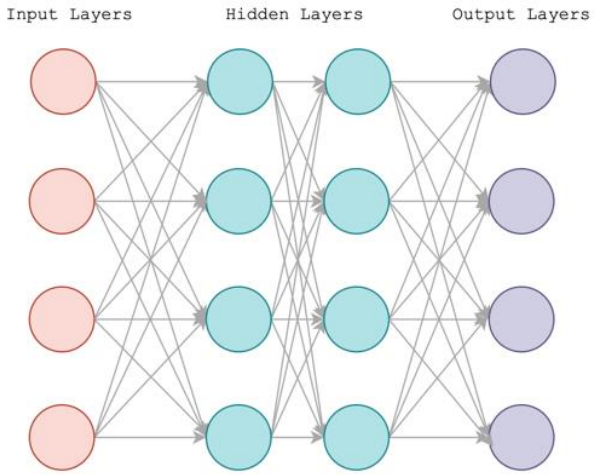

Figure 10. Structure of a Multi-layer perceptron (MLP) machine learning model

This study presented a practical methodology for the detection and classification of lung cancer in CT scan images [5]. The approach involved extracting texture features, including statistical features, which were then fed into various classification algorithms. These algorithms included k-nearest neighbors, support vector machines, decision trees, multinomial naive Bayes, stochastic gradient descent, random forests, and multi-layer perceptrons (MLPs). The analysis outcomes revealed that the MLP classifier achieved the highest accuracy, reaching an impressive rate of 88.55%, surpassing the performance of the other classifiers examined in the study.

2.2) Convolutional Neural Network

Convolutional Neural Networks distinguish themselves from traditional neural architectures through their exceptional ability to process structured grid data, particularly images and audio signals (Figure 11). Their specialized structure comprises three principal layer types. The convolutional layer serves as the network's foundation, utilizing sliding filter matrices to detect local patterns through the generation of feature maps. Pooling layers perform downsampling operations, reducing spatial dimensions while preserving essential information and providing translation invariance. Fully connected layers serve as the final processing stage, integrating the detected features to inform classification decisions. This hierarchical organization creates a progressive feature extraction pipeline. The initial convolutional layers identify fundamental visual elements—

detecting basic attributes such as color boundaries, edges, and simple textures. As information propagates through intermediate layers, the network begins recognizing more complex patterns and structural components of increasing abstraction. The deeper layers gradually assemble these components into recognizable object parts, with each successive layer capturing increasingly sophisticated representations. This architectural progression enables the network to transform raw pixel data into meaningful semantic concepts, ultimately converging at the fully connected layer where comprehensive image understanding occurs and final classification decisions are made. This hierarchical feature extraction mirrors how human visual processing combines elemental visual information into coherent object recognition, but accomplishes this through learned mathematical operations rather than biological processes.

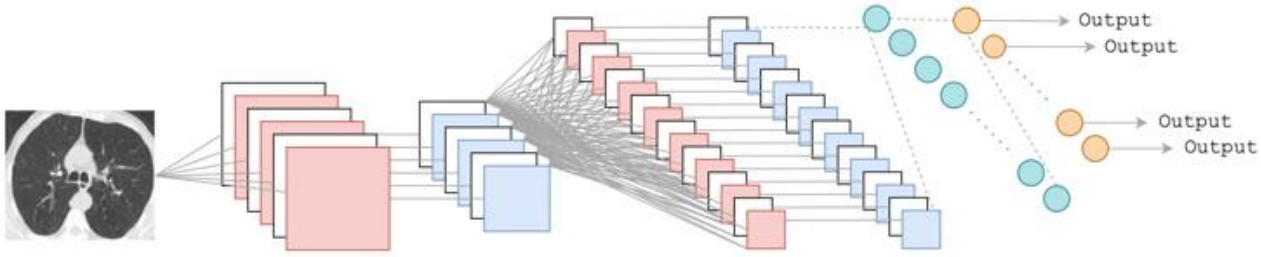

Figure 11. Convolutional Neural Network (CNN) machine learning model.

The research by Cai et al. [28] addresses the critical challenge of developing effective computerized detection models for COVID-19 diagnosis using CT imaging, with particular focus on overcoming the limitations of small dataset sizes. The investigators propose an innovative architectural approach—the stacked autoencoder detector model—which demonstrates significant performance improvements across multiple evaluation metrics. Their methodology employs a two-phase implementation strategy. Initially, they construct four specialized autoencoder modules that form the foundation layers of their detection system. These autoencoders are specifically engineered to extract high-dimensional, information-rich features from thoracic CT images that might otherwise remain undetected through conventional approaches. In the second phase, these individual autoencoder components are integrated through a cascading arrangement, with their outputs connected to a fully-connected dense layer and ultimately to a softmax classification layer. This creates a comprehensive end-to-end detection pipeline capable of distinguishing COVID-19 manifestations from normal pulmonary imaging. Performance evaluation reveals remarkable effectiveness, particularly noteworthy given the limited amount of available training data. The system achieves exceptional results across all primary assessment metrics: 94.7% overall accuracy, 96.54% precision (indicating minimal false positives), 94.1% recall (demonstrating high sensitivity to positive cases), and a balanced F1-score of 94.8%.

2.3) Support vector machine

The Support Vector Machine (SVM) represents a sophisticated classification and regression methodology, distinguished by its ability to optimize decision boundaries while controlling model complexity. This technique particularly excels when processing high-dimensional feature spaces containing thousands of predictor variables, making it exceptionally versatile across diverse application domains. The wide-ranging applicability of SVM extends to numerous fields, including customer analytics, computer vision applications, biological data processing, natural language understanding, network security, protein analysis, and audio recognition systems, demonstrating its remarkable flexibility and effectiveness. At its core, SVM operates by transforming input data into higher-dimensional feature spaces where separation between classes becomes mathematically tractable. The algorithm identifies optimal hyperplanes that maximize the margin between different categories, creating robust decision boundaries that generalize effectively to new data. Through kernel transformations, SVMs can represent complex nonlinear relationships while maintaining computational efficiency, enabling powerful classification capabilities even with challenging datasets. For example, consider the following (Figure 12), in which the data points are categorized into two distinct groups.

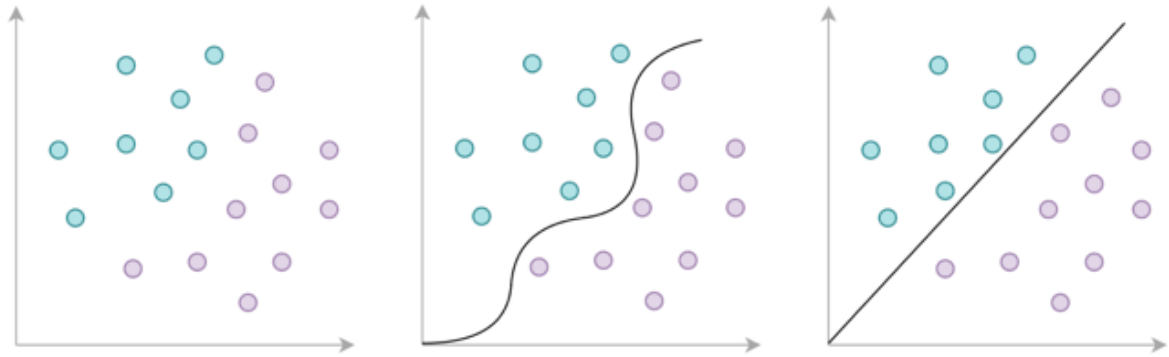

Figure 12. Support vector machine (SVM) machine learning model.

Research by Rahman et al. [14] demonstrates the effectiveness of SVM in medical imaging through an innovative approach for pulmonary nodule analysis in CT scans. Their methodology employs a multi-stage process, starting with the precise segmentation of lungs using active contour models. This initial delineation is followed by specialized masking techniques that transform connected nodules into isolated structures, making them suitable for classification. The critical detection phase employs an SVM classifier leveraging a combination of two-dimensional stochastic features and three-dimensional anatomical characteristics. This hybrid feature approach enables the system to accurately identify subtle nodule patterns while maintaining contextual awareness of surrounding anatomical structures. Following detection, the system applies active contour modeling to extract precise nodule boundaries, facilitating volumetric and morphological analysis. This classification framework successfully differentiates between nodules connected to pulmonary structures (such as vessel attachments or bronchial wall nodules) and isolated parenchymal nodules. The complete system achieves 89% detection accuracy while maintaining minimal false positive rates, demonstrating SVM's capability to handle complex medical image classification tasks with high precision.

Performance evaluation metrics

Performance evaluation metrics were applied to measure the effectiveness of a machine learning algorithm. Five classification metrics were utilized, including accuracy, F1 score, precision, time complexity, and space complexity. These metrics are crucial for evaluating the accuracy, efficiency, and efficacy of a model or system, and for identifying areas that require improvement. **Accuracy**, which measures the proportion of correct predictions made by the model or algorithm, is one of the most used performance evaluation metrics in machine learning. However, it may not be applicable for multiclass classification problems where the output has more than two categories. The F1 score, which balances precision and recall, is a flexible metric for dealing with imbalanced datasets.

$$\frac{TP + TN}{TP + FP + TN + FN}$$

F1 score is a metric commonly used in machine learning and statistics to evaluate the performance of a classification model, particularly when dealing with imbalanced datasets. It is a single numerical value that combines precision and recall, providing a balanced measure of a model's accuracy.

$$\frac{2TP}{2TP + FP + FN}$$

Precision and recall are two other important performance evaluation metrics that measure the proportion of true positive predictions among all positive predictions and the proportion of true positive

predictions among all actual positive instances, respectively. These metrics are particularly Fundamental in applications where false positives or false negatives can have serious consequences, such as medical diagnosis.

$$\frac{TP}{TP + FP}$$

Time complexity and space complexity are measures of the computational resources an algorithm requires during execution. Time complexity quantifies how the runtime increases relative to the input size, typically expressed using Big O notation. This mathematical representation establishes an upper bound on growth rates, enabling performance comparisons between different algorithms. Space complexity analyzes memory utilization patterns, assessing both primary algorithm storage needs and auxiliary memory requirements. For data-intensive applications such as medical image processing, space complexity becomes particularly critical when handling large datasets like high-resolution CT scans. When evaluating machine learning models for clinical applications, these metrics take on special significance. While complex models may achieve superior classification accuracy, their computational demands might render them impractical for real-time diagnostic settings. Algorithms demonstrating favorable time and space complexity characteristics offer advantages in resource-constrained environments where processing efficiency directly impacts clinical workflow. The balance between algorithmic complexity and computational efficiency represents a key consideration when developing diagnostic systems, particularly those intended for deployment in settings with limited computational resources or where rapid results are essential for patient care.

3. Results and Discussion

3.1 Experimental Setup

In this study, we utilized the Lung Image Database Consortium Image Collection (LIDC-IDRI), as reported by Armato, III. et al. [24], which comprises 244,527 thoracic computed tomography (CT) scans for diagnostic and lung cancer screening purposes from 1,010 patients. We also incorporated normal CT scans from a publicly accessible COVID-19 (SARS-CoV-2) lung CT dataset, which includes 1,229 negative cases, encompassing both normal and non-COVID-19 images. These datasets provide images in a 512x512px PNG format, sourced from actual patients at the radiology departments of academic hospitals in Tehran, as documented by Aria, M. et al. [25]. The final dataset includes 1,224 normal lung CT scans and 1,010 lung cancer cases.

3.2 Experiment Result

Table 1. Performance results of traditional machine learning models with normal and non-normal lung image datasets.

Model	Accuracy	Precision	Recall	F1 Score	Time Complexity (min)	Time Complexity (MB)
MLP	0.55	0.27	0.50	0.35	206.63	4811.45
CNN	0.96	0.97	0.96	0.96	3.40	5628.08
SVM	0.95	0.96	0.95	0.95	108.20	9767.71

In this section, we present two distinct solutions. The first solution pertains to the image classification results, where we report the outcomes of our efforts in distinguishing individuals as having either normal lung conditions or lung cancer, based on image analysis in (Table 1). The second solution focuses on the results of our feature engineering endeavors, which also aim to identify individuals in the same manner, differentiating between normal and cancerous cases, specifically those related to the lungs. We implemented on Google Colaboratory notebooks, which allocated an A100 GPU. In this experiment, various performance metrics were employed to assess the effectiveness of the machine learning model. A total of two experimental runs were conducted, each consisting of three iterations, with the results subsequently averaged for inclusion

in this report. Initially, the model was trained and tested using both normal and non-normal lung image datasets. The results have been presented in (Table 2). The experiment yielded compelling results. The Convolutional Neural Network (CNN) classification model emerged as the standout performer, achieving an impressive accuracy rate of 0.95. This results in effectiveness in accurately classifying lung images as either normal or those with lung cancer. On the other hand, when it comes to computational efficiency, in terms of both runtime complexity and space complexity, the Logistic Regression classification model outperforms the others. This model demonstrated an advantageous balance between accuracy and resource utilization, making it a strong candidate for applications where efficiency is a critical consideration. In essence, our findings emphasize the importance of selecting the right tool for the task at hand. While the CNN excelled in accuracy, the Logistic Regression model showcased superior efficiency, highlighting the need to strike a balance between precision and computational resources in practical applications.

Table 2. Performance comparison of different machine learning models with normal and non-normal lung image datasets.

Model	Accuracy	Precision	Recall	F1 Score	Time Complexity (min)	Space Complexity (MB)
Logistic Regression	0.91	0.92	0.90	0.91	0.71	3859.11
Decision Tree	0.94	0.95	0.94	0.94	22.6	5268.47
K-Nearest Neighbor	0.94	0.95	0.94	0.94	160.57	7793.39
MLP	0.55	0.27	0.50	0.35	206.63	4811.45
CNN	0.96	0.97	0.96	0.96	3.40	5628.08
SVM	0.95	0.96	0.95	0.95	108.20	9767.71

In the second experimental phase, we utilized a morphological feature engineering dataset for training and testing, employing the same machine learning model as in the first phase. The outcomes of these experiments are summarized in (Table 3). Notably, the highest accuracy rate achieved was 0.88, and this was attained by both the Multilayer Perceptron and Support Vector Machine classification models. These results underscore the effectiveness of these models in handling morphological feature engineering data for lung cancer classification. Similar to the first experimental phase, when considering time complexity and space complexity, the Logistic Regression classification model demonstrated superior efficiency. This finding reaffirms the Logistic Regression model as a compelling choice for applications where computational resources need to be optimized while maintaining high classification accuracy.

Table 3. Performance comparison of different machine learning models with the Feature Engineering Encoder dataset.

Model	Accuracy	Precision	Recall	F1 Score	Time Complexity (min)	Space Complexity (MB)
Logistic Regression	0.86	0.87	0.87	0.86	0.04	176.80
Decision Tree	0.87	0.88	0.88	0.87	0.48	241.45
K-Nearest Neighbor	0.87	0.88	0.88	0.87	3.21	494.99
MLP	0.88	0.89	0.89	0.88	21.99	220.83
CNN	0.82	0.83	0.83	0.82	0.16	257.79
SVM	0.88	0.89	0.87	0.87	7.15	447.40

Our analysis revealed two distinct dataset types. Models trained on standardized non-normal and normal images achieved higher accuracy compared to those trained on feature-encoded datasets. However, while differences in accuracy were apparent, precision and F1-score values showed only marginal variation.

Significantly, the use of standardized images substantially reduced computational demands, with both time and memory complexity significantly lowered. This finding highlights the potential for considerable efficiency gains in lung cancer classification tasks. When comparing explainable machine learning models with traditional non-interpretable approaches, we observed broadly comparable performance. The choice of algorithm—whether Decision Tree, Logistic Regression, K-Nearest Neighbor, CNN, or SVM—did not substantially alter the overall diagnostic effectiveness when applied to the respective datasets. To assess whether performance differences were statistically significant, paired t-tests were conducted across cross-validation folds. CNN achieved significantly higher accuracy compared to Logistic Regression and Decision Tree ($p < 0.05$). In contrast, no significant difference was observed between CNN and SVM ($p = 0.18$). Explainable models (Decision Tree, Logistic Regression, and KNN) demonstrated statistically comparable performance to non-interpretable models on feature-engineered datasets, as evidenced by overlapping 95% confidence intervals. These results suggest that small variations in performance do not undermine the interpretability advantage of an explainable app. Overall, our findings underscore the importance of considering both accuracy and resource efficiency in medical image classification. While certain datasets may achieve marginally higher accuracy, the performance gap is often not statistically significant. Prioritizing computationally efficient and interpretable methods is therefore essential for real-world clinical deployment, where resource constraints and the need for transparent decision-making are critical factors.

3.3 Discussion

Unlike previous studies that primarily prioritized accuracy, our approach introduces a novel morphological feature engineering pipeline that reduces computational burden while encoding radiologically interpretable features. This dual benefit bridges the gap between AI accuracy and clinical usability. Our comparative analysis highlights the trade-offs between accuracy, efficiency, and interpretability. While CNN achieved the highest accuracy (0.96), explainable models such as Decision Trees and K-Nearest Neighbors produced comparable results (0.94), with only a marginal difference of ~2%. Significantly, morphological feature engineering reduced processing time and memory usage by up to 95% while maintaining clinically acceptable accuracy (up to 0.88). By framing predictions in terms of radiological features such as nodule size, opacity, and distribution, explainable models not only preserved diagnostic reliability but also enhanced transparency and clinical trust. Overall, these findings suggest that explainable models, combined with feature engineering, offer a practical and scalable solution for real-world deployment, supporting radiologists in their diagnostic workflows rather than replacing them.

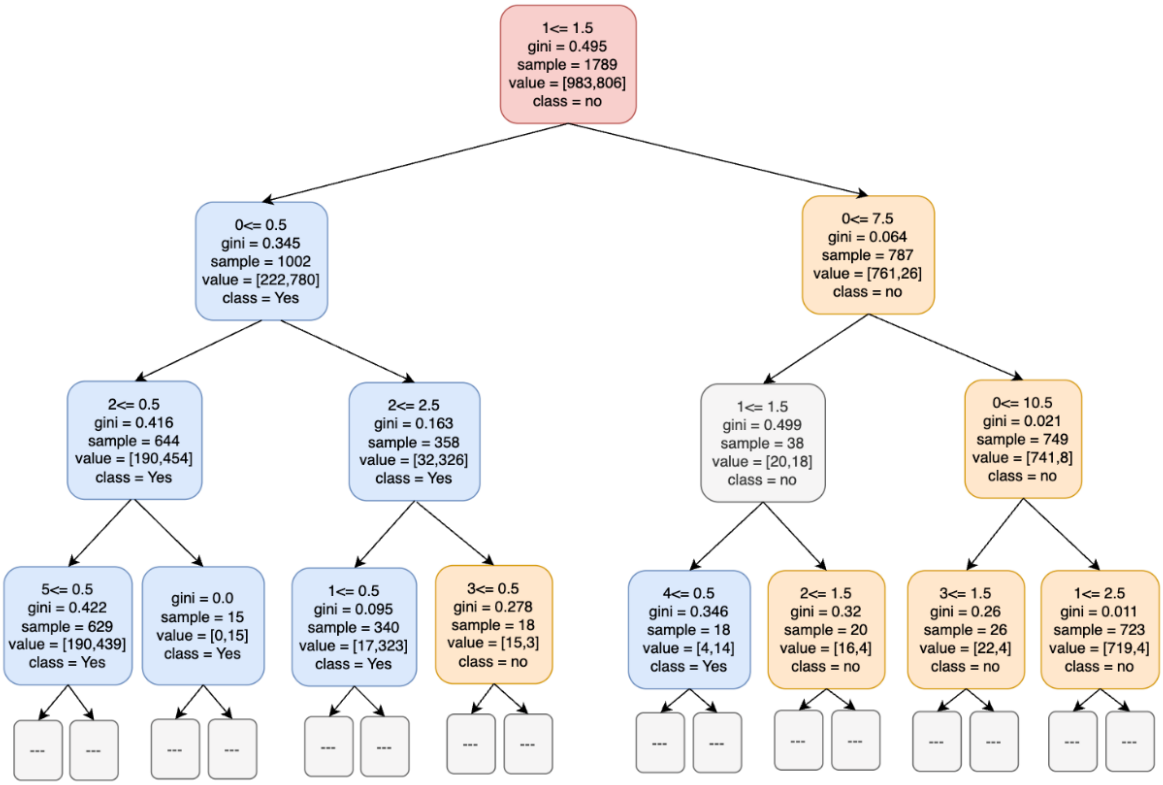

Figure 13. Decision Tree classifier for lung cancer detection.

The tree structure illustrates how radiological features (e.g., nodule size, opacity) are used in decision rules, with internal nodes representing feature thresholds and leaf nodes indicating class predictions ("cancer" or "normal") shown in (Figure 13). The gini impurity metric quantifies class purity at each node, where lower values denote stronger separation. For example, one path shows that if feature_1 ≤ 0.5 and feature_3 > 1.0, the model predicts "lung cancer" with high confidence. This visualization clarifies the reasoning process, enabling clinicians to trace decision paths and verify which features most influenced each prediction. Such transparency is critical for integrating AI into clinical workflows, where interpretability and accountability are essential.

4. Conclusions

This research introduced a novel application of explainable machine learning methodologies for feature encoding from lung cancer CT imaging data. By systematically comparing explainable algorithms with black-box approaches, our framework demonstrated reliable effectiveness in identifying malignant pulmonary nodules. Our results show that explainable models can achieve diagnostic performance comparable to conventional methods, while offering the added advantage of interpretability. A key contribution of this study lies in the morphological feature engineering pipeline, which substantially reduced computational demands in terms of both processing time and memory usage. This efficiency gain represents a meaningful step toward practical deployment in clinical environments. The broader implications extend to real-world healthcare, where rapid, accurate, and interpretable decision support is essential. By aligning with radiologists' diagnostic reasoning, explainable models can enhance trust and usability in clinical workflows, thereby supporting early detection of lung malignancies and contributing to improved patient outcomes. Looking ahead, this methodology holds promise for broader medical imaging tasks, including oncology, dermatology, neurology, and other radiological applications. Future research should refine feature extraction techniques, validate performance in diverse patient populations, and explore hybrid approaches that combine feature engineering with deep learning. Explainable machine learning models integrated with morphological

feature engineering offer a compelling alternative to black-box approaches. Although CNNs deliver slightly higher accuracy, explainable models provide transparency, accountability, and clinical trust—attributes essential for the responsible integration of AI in healthcare. By balancing accuracy, efficiency, and interpretability, explainable AI can serve as a practical decision-support tool, ultimately laying the foundation for trustworthy and sustainable clinical adoption.

5. Acknowledgements

The authors would like to thank the Department of Computer Science at Chiang Mai University for providing computational resources and the radiology departments of the teaching hospitals in Tehran for their assistance with the CT scan datasets.

Author Contributions: Conceptualization, K.T. and P.S.; methodology, K.T.; software, K.T.; validation, K.T. and P.S.; formal analysis, K.T.; investigation, K.T.; resources, P.S.; data curation, K.T.; writing—original draft preparation, K.T.; writing—review and editing, P.S.; visualization, K.T.; supervision, P.S.; project administration, P.S. All authors have read and agreed to the published version of the manuscript.

References

[1] Takenaka, M.; Hanagiri, T.; Shinohara, S.; Kuwata, T.; Chikaishi, Y.; Oka, S.; Tanaka, F. The prognostic significance of HER2 overexpression in non-small cell lung cancer. *Abbreviated Journal Name* **2011**.

[2] Oser, M. G.; Niederst, M. J.; Sequist, L. V.; Engelman, J. A. Molecular drivers and cells of origin in the transformation from non-small-cell lung cancer to small-cell lung cancer. *Lancet Oncol.* **2015**, *16*(4), E165-E172. https://doi.org/10.1016/S1470-2045(14)71180-5

[3] de Castro, J.; Rodríguez, M. C.; Martínez-Zorzano, V. S.; Sánchez-Rodríguez, P.; Sánchez-Yagüe, J. Erythrocyte fatty acids as potential biomarkers in the diagnosis of advanced lung adenocarcinoma, lung squamous cell carcinoma, and small cell lung cancer. *Abbreviated Journal Name* **2014**, *142*(1), 111-120. https://doi.org/10.1309/AJCP1QUQQLLT8BLI

[4] Yu, L.; Tao, G.; Zhu, L.; Wang, G.; Li, Z.; Ye, J.; Chen, Q. Prediction of pathologic stage in non-small cell lung cancer using machine learning algorithm based on CT image feature analysis. *Abbreviated Journal Name* **2019**. https://doi.org/10.1186/s12885-019-5646-9

[5] Singh, G. A. P.; Gupta, P. K. Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans. *Abbreviated Journal Name* **2019**.

[6] Kadir, T.; Gleeson, F. Lung cancer prediction using machine learning and advanced imaging techniques. *Abbreviated Journal Name* **2018**, *7*(3). https://doi.org/10.21037/tlcr.2018.05.15

[7] Elish, M. C. The stakes of uncertainty: developing and integrating machine learning in clinical care. *Abbreviated Journal Name* **2018**. https://doi.org/10.1111/1559-8918.2018.01213

[8] Tjoa, E.; Guan, C. A survey on explainable artificial intelligence (XAI): Toward medical XAI. *IEEE Transactions on Neural Networks and Learning Systems* **2020**, *32*(11). https://doi.org/10.1109/TNNLS.2020.3027314

[9] Doshi-Velez, F.; Kim, B. Towards a rigorous science of interpretable machine learning. *Abbreviated Journal Name* **2017**.

[10] Balagurunathan, Y.; Kumar, V.; Gu, Y.; Kim, J.; Wang, H.; Liu, Y.; Goldgof, D. B.; Hall, L. O.; Korn, R.; Zhao, B.; Schwartz, L. H.; Basu, S.; Eschrich, S.; Gatenby, R. A.; Gillies, R. J. Test–Retest Reproducibility Analysis of Lung CT Image Features. *Med. Phys.* **2014**, *41*(5), 2405–2427.

[11] Paing, M. P.; Choomchuay, S. Ground glass opacity (GGO) nodules detection from lung CT scans. *Abbreviated Journal Name* **2017**. https://doi.org/10.1109/ISESD.2017.8253338

[12] Ramalho, G. L. B.; Rebouças Filho, P. P.; Medeiros, F. N. S. D.; Cortez, P. C. Lung disease detection using feature extraction and extreme learning machines. *Abbreviated Journal Name* **2014**. https://doi.org/10.1590/rbeb.2014.019

[13] Abdillah, B.; Bustamam, A.; Sarwinda, D. Image processing based detection of lung cancer on CT scan images. *Abbreviated Journal Name* **2017**. https://doi.org/10.1088/1742-6596/893/1/012063

[14] Keshani, M.; Azimifar, Z.; Tajeripour, F.; Boostani, R. Lung nodule segmentation and recognition using SVM classifier and active contour modeling: A complete intelligent system. *Abbreviated Journal Name* **2013**, *43*(4), 287-300. https://doi.org/10.1016/j.compbiomed.2012.12.004

[15] Faisal, M. I.; Bashir, S.; Khan, Z. S.; Khan, F. H. An evaluation of machine learning classifiers and ensembles for early stage prediction of lung cancer. *Abbreviated Journal Name* **2018**. https://doi.org/10.1109/ICEEST.2018.8643311

[16] Günaydin, Ö.; Günay, M.; Şengel, Ö. Comparison of Lung Cancer Detection Algorithms. In *Proceedings of the 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, April 2019; IEEE, **2019**; pp 1–4. https://doi.org/10.1109/EBBT.2019.8741826

[17] Tonekaboni, S.; Joshi, S.; McCradden, M. D.; Goldenberg, A. What clinicians want: contextualizing explainable machine learning for clinical end use. *Abbreviated Journal Name* **2019**.

[18] Kumar, D.; Wong, A.; Clausi, D. A. Lung nodule classification using deep features in CT images. *Abbreviated Journal Name* **2015**. https://doi.org/10.1109/CRV.2015.25

[19] Thallam, C.; Peruboyina, A.; Raju, S. S. T.; Sampath, N. Early stage lung cancer prediction using various machine learning techniques. In *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*; IEEE: City, Country, **2020**; pp 1285–1292. https://doi.org/10.1109/ICECA49313.2020.9297576

[20] Li, X.; Shen, L.; Xie, X.; Huang, S.; Xie, Z.; Hong, X.; Yu, J. Multi-resolution convolutional networks for chest X-ray radiograph-based lung nodule detection. *Artificial Intelligence in Medicine* **2020**, *103*, 101744. https://doi.org/10.1016/j.artmed.2019.101744

[21] Daisy, T. White Box vs. Black Box Algorithms in Machine Learning. ActiveState, 2023. https://www.activestate.com/ (accessed 2023-07-09).

[22] Kareem, H. F.; AL-Husieny, M. S.; Mohsen, F. Y.; Khalil, E. A.; Hassan, Z. S. Evaluation of SVM Performance in the Detection of Lung Cancer in Marked CT Scan Dataset. *Indonesian J. Electr. Eng. Comput. Sci.* **2021**, *21*(3), 1731. https://doi.org/10.11591/ijeecs.v21.i3.pp1731-1738

[23] Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; Pedreschi, D. A Survey of Methods for Explaining Black Box Models. *ACM Comput. Surv.* **2018**, *51*(5), 1–42. https://doi.org/10.1145/3236009

[24] Armato III, S. G.; McLennan, G.; Bidaut, L.; McNitt-Gray, M. F.; Meyer, C. R.; Reeves, A. P.; Clarke, L. P. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Abbreviated Journal Name* **2011**.

[25] Aria, M.; Ghaderzadeh, M.; Asadi, F.; Jafari, R. Lung CT Scan Dataset from Tehran Teaching Hospitals. Mendeley Data, 2021. https://doi.org/10.17632/hn6vr7r5cm.1 (accessed 2023-07-09).

[26] Maleki, N.; Zeinali, Y.; Niaki, S. T. A. A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection. *Abbreviated Journal Name* **2021**, *164*, 113981. https://doi.org/10.1016/j.eswa.2020.113981

[27] Radhika, P. R.; Nair, R. A.; Veena, G. A comparative study of lung cancer detection using machine learning algorithms. *Abbreviated Journal Name* **2019**.

[28] Li, D.; Fu, Z.; Xu, J. Stacked-autoencoder-based model for COVID-19 diagnosis on CT images. *Abbreviated Journal Name* **2021**, *51*, 2805-2817. https://doi.org/10.1007/s10489-020-02002-w