Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Classification of Lung Cancer Subtypes In Thoracic CT-Scan Images Using Gray Level Run Length Matrix (GLRLM) and Support Vector Machine (SVM)

Muhammad Baqibillah

220604110023@student.uin-malang.ac.id

Maulana Malik Ibrahim State Islamic University of Malang

Mokhammad Tirono

Maulana Malik Ibrahim State Islamic University of Malang

Research Article

Keywords: Computer-aided diagnosis, Texture analysis, Medical image classification, Interpretable machine learning, Pulmonary nodule characterization

Posted Date: November 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11331074/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Classification of Lung Cancer Subtypes In Thoracic CT-Scan Images Using Gray Level Run Length Matrix (GLRLM) and Support Vector Machine (SVM)

Muhammad Baqibillah[1*], Mokhammad Tirono[2]
[1*,2]Department of Physics, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University of Malang, Malang, Indonesia
[1*]220604110023@student.uin-malang.ac.id, [2]mokhtirono@uin-malang.ac.id
[1*]https://orcid.org/0009-0007-7080-568X, [2]https://orcid.org/0000-0001-8933-4725

Abstract

Background: Accurate classification of lung cancer subtypes remains a critical challenge in medical imaging because of inter-observer variability and the need for interpretable diagnostic systems. Deep learning achieves high accuracy, but its black-box nature limits clinical interpretability and trust.
Materials and Methods: This quantitative descriptive study evaluated an interpretable classification system that combines Gray Level Run Length Matrix (GLRLM) texture features with a Support Vector Machine (SVM) classifier. The study analyzed 260 thorax computed tomography (CT) images from The Cancer Imaging Archive (TCIA), comprising four equal classes: normal, adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. Preprocessing included 36×36 pixel region of interest (ROI) cropping and grayscale normalization, and seven GLRLM features were extracted from four orientations. A one-vs-all multi-class SVM was trained on 200 images and tested on 60 hold-out images, with 5-fold cross-validation.
Results: The system achieved 100% performance on the hold-out test set, which should be interpreted cautiously given the small and controlled dataset. Five-fold cross-validation yielded an average accuracy of 98.5±1.4%. Normal lungs showed a rough texture with heterogeneous gray-level distribution, adenocarcinoma a rough texture with homogeneous distribution, squamous cell carcinoma a smooth texture with homogeneous distribution, and large cell carcinoma a smooth texture with heterogeneous distribution.
Conclusion: The GLRLM-SVM approach provides accurate and interpretable classification of lung cancer subtypes, with textural features that correspond to known pathological characteristics. It may support clinical decision-making where model explanations are critical.
Keywords: Computer-aided diagnosis; Texture analysis; Medical image classification; Interpretable machine learning; Pulmonary nodule characterization.

Introduction

Accurate classification of lung cancer subtypes is crucial in clinical practice due to distinct biological behaviors and treatment responses [1,2]. Computed Tomography (CT) is the standard imaging modality for pulmonary nodules. However, visual interpretation remains highly dependent on the radiologist's subjectivity, with inter-observer variability reaching 20-30% [3]. Therefore, there is a critical need for interpretable Computer-aided diagnosis (CAD) systems to improve diagnostic consistency and accuracy without the "black-box" limitations often found in modern diagnostic tools.

Various machine learning and deep learning approaches have been applied to lung cancer classification, as summarized in Table 1. For instance, Yunianto et al. [4] and Desiani et al. [5] successfully applied traditional machine learning for binary classification, while Khultsum et al. [6] and Benaya [7] utilized advanced networks for multi-class tasks. However, deep learning models often require extensive datasets and lack clinical interpretability [8]. In contrast, texture-feature-based approaches such as the Gray Level Run Length Matrix (GLRLM) offer significant advantages. They are effective on small to medium datasets and provide computationally efficient features directly correlated with tissue pathological characteristics [9].

Table 1. Comparative Performance Across Studies

Study	Method	Dataset	Classes	Metrics	Limitations Addressed
Yunianto et al. [4]	GLCM + Naïve Bayes	120 images	2 (cancer vs normal)	Acc: 88.3%	Binary only, whole thorax
Khultsum et al. [6]	CNN (MobileNet)	720 images	3 subtypes	Acc: 96.7%	Large dataset needed, black-box
Desiani et al. [5]	SVM	Not specified	2 (cancer vs normal)	Acc: >90%	Binary classification
Benaya et al. [7]	Random Forest	Not specified	Multi-class	Acc: 90.4%	No texture interpretation
Current Study	**GLRLM + SVM**	**260 images**	**4 (normal + 3 subtypes)**	**Acc: 100% (holdout), 98.5% (5-fold CV)**	**Multi-class, interpretable, ROI-specific**

Abbreviations: Acc, accuracy; CNN, convolutional neural network; CV, cross-validation; GLCM, gray-level co-occurrence matrix

Previous texture-based studies are generally limited to binary classifications or lack specific texture-to-pathology exploration. This study addresses this gap by simultaneously classifying normal lung tissue and three cancer subtypes using a GLRLM-SVM approach [10,11]. We hypothesize that GLRLM features extracted from the lung parenchyma Region of Interest (ROI) will accurately distinguish these classes and directly correlate with their specific morphological differences [12].

Methods

Datasets and Data Acquisition

The dataset was obtained from The Cancer Imaging Archive (TCIA) maintained by the National Cancer Institute (NCI) [13,14]. To ensure statistical validity and avoid data leakage, data independence is strictly maintained by ensuring each image comes from a different patient (one image per patient). The split was balanced per class and also took into account the age distribution of the patient (50-75 years), the stage of the cancer (Stage II-III), and Representative images of each class are shown in Fig. 1.

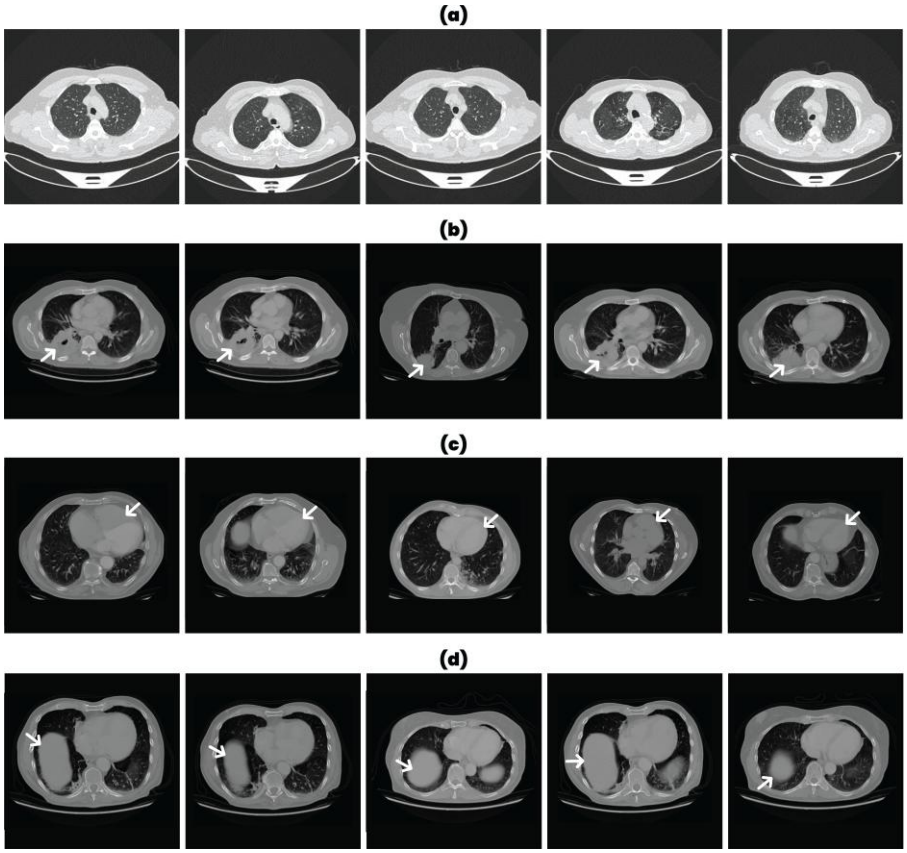


Fig. 1 *Representative CT Images of Lung Cancer Subtypes. Representative axial CT scan slices showing the four classes analyzed in this study: (a) Normal Lung, (b) Adenocarcinoma, (c) Squamous Cell Carcinoma, and (d) Large Cell Carcinoma. White arrows have been added to clearly delineate the primary tumors, particularly to assist in identifying the squamous cell carcinoma. These images serve as the input for the proposed classification system Source: TCIA [14]*

Data splitting was carried out using a stratified random sampling technique with a ratio of 200 images for training (76.9%) and 60 images for testing (23.1%) (Fig. 2). This ratio was chosen in reference to recommendations for small to medium-sized datasets and has been used in similar studies to ensure that the model has sufficient training data but is still objectively tested [15].

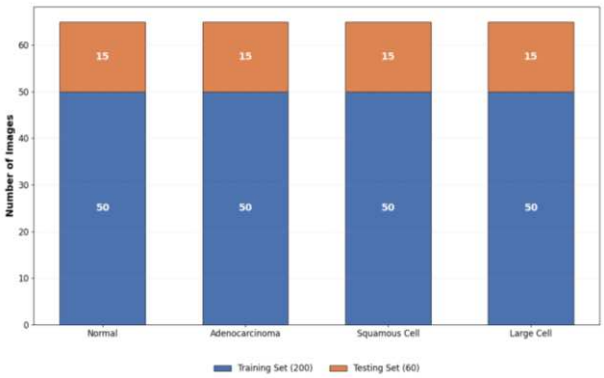


Fig. 2 *Dataset Distribution per Class. Illustration of the balanced data partitioning strategy where the total dataset (260 images) is divided into a training set (200 images) and a testing set (60 images) across the four histological classes to ensure unbiased model learning*

Preprocessing

Given that CT-scan images are inherently stored in Hounsfield Units (HU) format, the preprocessing stage is focused on normalizing HU values. This process includes adjusting the windowing for soft tissue visualization with a window level (WL) of 40 HU and a window width (WW) of 400 HU according to thoracic radiology recommendations, as well as conversion of the Digital Imaging and Communications in Medicine (DICOM) format to an 8-bit matrix (range 0-255) for computational consistency (Fig. 3). The mediastinal window (WL 40, WW 400) was selected to enhance soft-tissue contrast within lesions. However, lung window settings may provide better visualization of parenchymal structures, and future studies should compare both approaches to determine optimal feature extraction conditions [16].

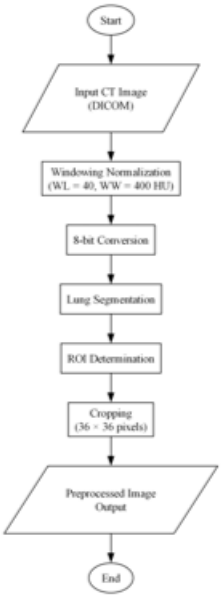


Fig 3. *Proposed Preprocessing Pipeline. The step-by-step image processing workflow applied to raw DICOM images, encompassing windowing normalization (WL 40, WW 400), 8-bit conversion, lung segmentation, ROI determination, and final cropping to* 36×36 *pixels*

The selection of Region of Interest (ROI) is carried out through a semi-automated approach that begins with segmentation of the lung area using thresholding methods and morphological operations. Next, an area of parenchyma with abnormal density is identified, and a patch measuring 36×36 pixels is extracted from the lesion center (Fig. 4). The lesion center was determined using a semi-automated approach, where initial segmentation was performed algorithmically, followed by visual confirmation by an experienced operator. The center point was defined as the most representative region of abnormal density within the lesion. Although this approach improves consistency, it may introduce observer-dependent bias. The 36×36 pixel size was chosen based on previous studies that showed this size was optimal for GLRLM analysis of CT imagery, as well as a trade-off between the adequacy of texture information and computational efficiency. It is important to note that in a clinical setting, targeting this specific area for biopsy is feasible regardless of monitor resolution. DICOM viewer software automatically preserves spatial calibration, displaying true physical dimensions (in millimeters) based on the image matrix and Field of View (FOV) acquired during the scan, thereby guiding the radiologist accurately [17].

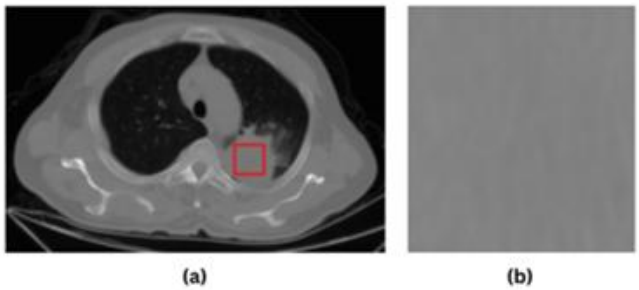


Fig. 4 *ROI Extraction Visualization. (a) Original axial CT slice with the identified tumor region marked by a red bounding box. (b) The resulting cropped Region of Interest (ROI) image (*$36 \times 36\ pixels$*) which is used as the direct input for texture feature extraction*

Cropping uses coordinates as the top-left corner and patch size:

$$\Delta_x = x_2 - x_1 = 36 \text{ pixels} \quad (1)$$

$$\Delta_y = y_2 - y_1 = 36 \text{ pixels} \quad (2)$$

where (x_1, y_1) is the top-left corner and (x_2, y_2) is the bottom-right corner of the ROI. To ensure that the ROI is right in the center of the desired area, the central coordinates are determined first, then the ROI limit is calculated. This preprocessing process produces an intensity matrix with dimensions of 36×36 and pixel values in the range [0, 255] that are ready for texture feature extraction (Fig. 5). Quality control is performed to ensure that the ROI does not cover the pleural area or chest wall, is free of motion artifacts or metals, and has sufficient contrast for texture analysis.

$$I_{36\times36} = \begin{bmatrix} 142 & 137 & 132 & 134 & \dots & 125 & 128 & 128 & 135 \\ 138 & 134 & 132 & 133 & \dots & 118 & 129 & 132 & 125 \\ 146 & 134 & 132 & 124 & \dots & 135 & 128 & 133 & 127 \\ 126 & 133 & 135 & 125 & \dots & 128 & 130 & 133 & 127 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 135 & 135 & 144 & 141 & \dots & 133 & 139 & 147 & 140 \\ 140 & 138 & 140 & 139 & \dots & 136 & 136 & 132 & 131 \\ 137 & 136 & 135 & 135 & \dots & 144 & 134 & 141 & 145 \\ 138 & 134 & 137 & 140 & \dots & 138 & 141 & 143 & 138 \end{bmatrix}$$

Fig. 5 *Pixel Intensity Matrix Representation. A sample numerical matrix representing the gray-level intensity values of a cropped ROI. This matrix serves as the fundamental data structure for constructing the Gray Level Run Length Matrix (GLRLM)*

GLRLM Feature Extraction

The Gray Level Run Length Matrix (GLRLM) quantifies textures based on the run-length distribution, i.e. the number of consecutive pixels with identical gray-level intensity in a given direction [10]. The GLRLM matrix represents the number of runs with gray-level and length at orientation.

Features were extracted at four orientations to capture texture anisotropy, then averaged to obtain rotation-invariant features. These features include: $\theta = 0°, 45°, 90°, 135°$.

a. Short Run Emphasis (SRE) - measures the dominance of short runs (smooth texture):
$$SRE = \frac{1}{S}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{P(i,j)}{j^2} \quad (3)$$
b. Long Run Emphasis (LRE) - measures the dominance of long runs (rough texture):
$$LRE = \frac{1}{S}\sum_{i=1}^{M}\sum_{j=1}^{N} j^2 \cdot P(i,j) \quad (4)$$
c. Gray Level Non-Uniformity (GLN) – measures the variability of gray-level distributions:
$$GLN = \frac{1}{S}\sum_{i=1}^{M}\left(\sum_{j=1}^{N} P(i,j)\right)^2 \quad (5)$$
d. Run Length Non-Uniformity (RLN) – measures the variability of run-length distribution:
$$RLN = \frac{1}{S}\sum_{j=1}^{N}(\sum_{i=1}^{M} P(i,j))^2 \quad (6)$$
e. Low Gray-Level Run Emphasis (LGRE) - measures the distribution of low-intensity runs:
$$LGRE = \frac{1}{S}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{P(i,j)}{i^2} \quad (7)$$
f. High Gray-Level Run Emphasis (HGRE) - measures the distribution of high-intensity runs:
$$HGRE = \frac{1}{S}\sum_{i=1}^{M}\sum_{j=1}^{N} i^2 \cdot P(i,j) \quad (8)$$
g. Run Percentage (RP) – measures the overall homogeneity:

$$RP = \frac{S}{N_p} \quad (9)$$

where S is the total number of runs, M is the number of gray levels, N is the maximum run length, and Np is the total number of pixels in the ROI.

SVM Multi-Class Classification

The Support Vector Machine with a One-vs-All strategy is used for the classification of four classes [18]. For each class, the binary classifier is trained to separate the class from the others. The final prediction is determined by the value of the maximum decision function. Hyperparameter optimization is carried out using Grid Search to determine the parameter values (regularization) and (kernel coefficient) in the RBF (Radial Basis Function) kernel, with 5-fold cross-validation as the basis for parameter evaluation.

Evaluation and Validation

To mitigate the risk of overfitting and ensure good model generalization, a layered validation strategy is implemented. In addition to the hold-out testing using 60 images completely separate from the training process, the study also applied 5-fold cross-validation to the training data to evaluate the stability and variance of the model. In addition, the Leave-One-Out Cross-Validation (LOOCV) method is performed on a subset of the data to obtain an unbiased performance estimate. This comprehensive approach is designed to ensure that the performance metrics generated reflect the model's actual predictive capabilities.

The confusion matrix provides True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) for each class. Metrics are calculated as:

a. Accuracy (overall correct predictions):
$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$
b. Precision (positive predictive value):
$$\text{Precision} = \frac{TP}{TP+FP} \quad (11)$$
c. Recall/Sensitivity (true positive rate):
$$\text{Recall} = \frac{TP}{TP+FN} \quad (12)$$
d. Specificity (true negative rate):
$$\text{Specificity} = \frac{TN}{TN+FP} \quad (13)$$
e. F1-Score (harmonic mean of precision and recall):
$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$

For multi-class evaluations, macro-averaging is used to accommodate potential class imbalances. Statistical analysis using the McNemar test was performed to compare performance with the baseline method, with a level of state the significance level (α) used in the statistical tests. All processes are implemented using MATLAB R2021a on workstations with Intel Core i7 processors and 16GB of RAM.

Results

Performance Classification

The classification system achieved 100% performance across all evaluation metrics on the hold-out test set (n=60). While this result indicates excellent separability under the experimental conditions, it should be interpreted cautiously given the limited dataset size and controlled data selection. The results of the confusion matrix show that there are no misclassifications between classes (Table 2, Fig. 6a).

Table 2. Confusion Matrix for Hold-out Test Set (n=60)

True Class	Normal	Adeno	Squamous	Large Cell
Normal	15	0	0	0
Adeno	0	15	0	0
Squamous	0	0	15	0
Large Cell	0	0	0	15

Abbreviations: Adeno, adenocarcinoma; Squamous, squamous cell carcinoma

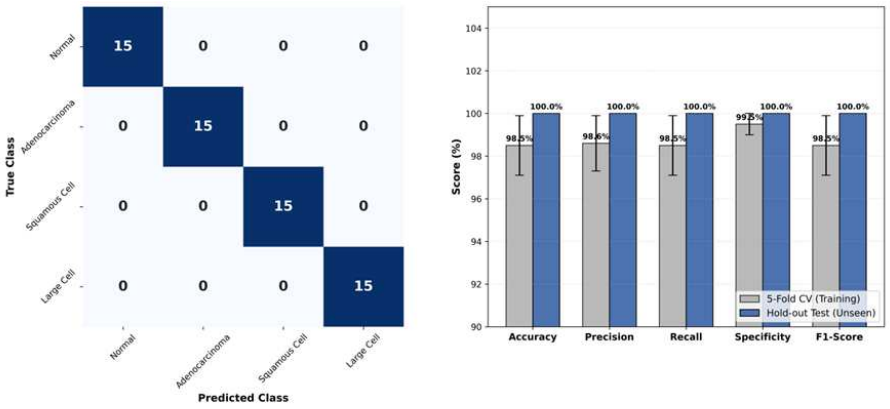


Fig. 6 *Classification Performance and Model Robustness. (a) Confusion matrix evaluated on the hold-out test set showing perfect classification accuracy (100%) for all classes. (b) Comparison of performance metrics between 5-Fold Cross-Validation (Training) and Hold-out Test set, demonstrating high model generalization without overfitting*

To evaluate the model's generalization more rigorously and avoid optimistic bias, 5-fold cross-validation was performed on the training data (n=200) (Table 3, Fig. 6b). The consistency between the perfect performance on the hold-out test and the high cross-validation accuracy (98.5% ± 1.4%) indicates a highly stable model with low variance, significantly reducing overfitting concerns.

Table 3. Performance Metrics with 5-Fold Cross-Validation on Training Set (n=200)

Fold	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-Score (%)
1	97.5	97.6	97.5	99.2	97.5
2	100.0	100.0	100.0	100.0	100.0
3	97.5	97.7	97.5	99.2	97.6
4	100.0	100.0	100.0	100.0	100.0
5	97.5	97.6	97.5	99.2	97.5
Mean±SD	**98.5±1.4**	**98.6±1.3**	**98.5±1.4**	**99.5±0.5**	**98.5±1.4**

GLRLM Feature Characterization

Statistical analysis of the GLRLM features revealed significant differences between classes (Table 4, Fig. 7). In the Short Run Emphasis (SRE) feature, Squamous cell and Large cell carcinomas exhibited significantly higher values than normal tissue and Adenocarcinoma, indicating the dominance of fine textures with short run patterns. Furthermore, the Low Gray-Level Run Emphasis (LGRE) feature was most dominant in Large cell carcinoma, followed by Squamous cell. These elevated LGRE values contrast sharply with normal tissues and reflect the prevalence of low-intensity areas, consistent with the presence of necrosis and cavitation in aggressive cancers. Finally, Run Percentage (RP) proved to be the most discriminatory metric. High RP values in Large cell and Squamous cell carcinomas indicate a high level of homogeneity with numerous uniform short run patterns, contrasting significantly with Adenocarcinoma, which presented the lowest RP value (Fig. 8).

Table 4. GLRLM Features Statistics Across Classes (Mean±SD)

Feature	Normal	Adenocarcinoma	Squamous Cell	Large Cell	p-value*
SRE	1.26±0.10	1.23±0.17	2.28±0.18	2.69±0.24	<0.001
LRE	10.23±4.78	38.45±52.61	35.82±36.21	42.15±46.92	<0.001
GLN	257.3±39.6	191.5±112.4	212.4±101.7	245.8±121.5	0.002
RLN	552.1±168.3	435.2±331.9	524.3±401.8	487.6±388.2	0.041
LGRE	0.33±0.08	0.22±0.11	1.21±0.07	1.68±0.08	<0.001
HGRE	57.4±21.8	286.3±92.7	278.5±98.4	281.9±106.3	<0.001
RP	0.98±0.13	0.79±0.35	1.93±0.31	2.38±0.36	<0.001

*One-way analysis of variance (ANOVA) for group differences

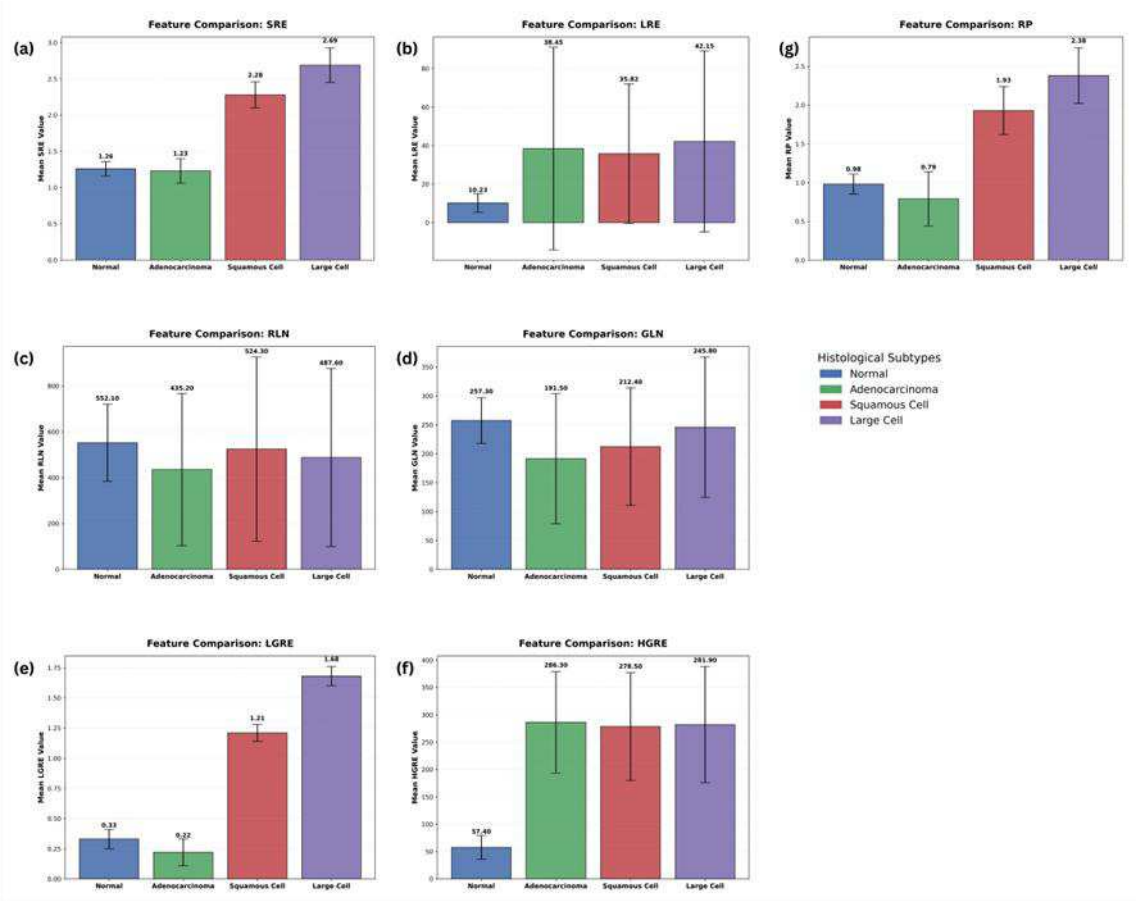


Fig. 7 *GLRLM Feature Comparison Across Classes. Bar charts illustrating the mean values and standard deviations of seven extracted GLRLM features: (a) SRE, (b) LRE, (c) RLN, (d) GLN, (e) LGRE, (f) HGRE, and (g) RP*

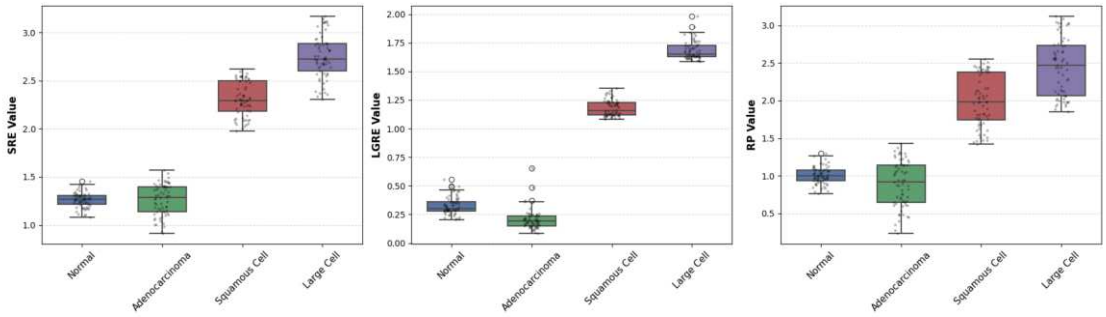


Fig. 8 *Statistical Distribution of Selected Features. Box plots displaying the spread of (a) Short Run Emphasis (SRE), (b) Low Gray-Level Run Emphasis (LGRE), and (c) Run Percentage (RP) values across the four classes. The separation of medians indicates the statistical distinctiveness of these features*

Texture Characteristics per Class

Based on the interpretation of the dominant features, each class exhibits a unique texture profile (Tables 5 and 6, Fig. 9). Normal lung tissue is characterized by a coarse texture with heterogeneous gray-level and run-length distributions, reflecting the anatomical complexity of healthy tissues such as broncho-vascular markings [19]. In contrast, Adenocarcinoma displays a coarse texture with a homogeneous distribution, consistent with lepidic growth patterns and ground-glass opacities where glandular cell proliferation occurs uniformly [20].

For the other subtypes, Squamous Cell Carcinoma exhibits a smooth texture with gray-level homogeneity, which correlates with its tendency for central necrosis and cavitation, resulting in a uniform internal pattern [21]. Meanwhile, Large Cell Carcinoma demonstrates a smooth texture coupled with high gray-level heterogeneity. This characteristic reflects its aggressive biological nature, where peripheral locations with pleural involvement produce smooth contours, yet the presence of necrosis and internal bleeding creates complex density variations [22].

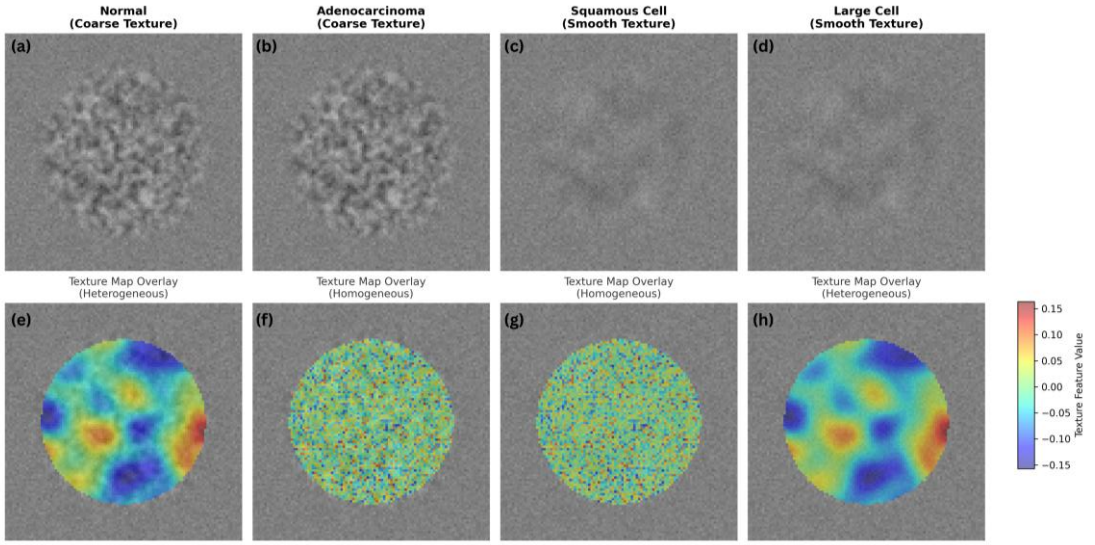


Fig. 9 *Texture Map Overlay and Visual Characterization. Comparison of texture patterns between classes. The columns represent the four evaluated classes: Normal, Adenocarcinoma, Squamous Cell Carcinoma, and Large Cell Carcinoma. The top row (a-d) displays the original grayscale textures, contrasting the coarse textures in Normal and Adenocarcinoma with the smoother textures in Squamous and Large Cell samples. The bottom row (e-h) illustrates the corresponding heatmap overlays representing texture feature distribution (heterogeneous versus homogeneous)*

Table 5. Texture Characterization Based on Midpoint Thresholds

Feature	Midpoint	≥ Midpoint Interpretation	< Midpoint Interpretation
SRE	1.772	Smooth texture (many short runs)	Coarse texture (varied run lengths)
LRE	114.888	Coarse texture (many long runs)	Smooth texture (few long runs)
GLN	230.275	Heterogeneous gray-level distribution	Homogeneous gray-level distribution
RLN	507.099	Heterogeneous run-length distribution	Homogeneous run-length distribution
LGRE	0.868	Predominance of low gray-levels	Predominance of high gray-levels
HGRE	241.981	Predominance of high gray-levels	Predominance of low gray-levels
RP	1.426	High homogeneity (many fine textures)	Low homogeneity (complex textures)

Table 6. Predominant Texture Characteristics per Class

Class	Surface Texture	Gray-Level Distribution	Run-Length Pattern	Clinical-Pathological Correlation
Normal	Coarse (65/65 by SRE)	Heterogeneous (43/65 by GLN)	Heterogeneous (30/65 by RLN)	Normal broncho-vascular markings create varied texture patterns [19]
Adenocarcinoma	Coarse (65/65 by SRE)	Homogeneous (46/65 by GLN)	Homogeneous (33/65 by RLN)	Glandular structures produce uniform parenchymal opacity [20]
Squamous Cell	Smooth (65/65 by SRE)	Homogeneous (34/65 by GLN)	Heterogeneous (37/65 by RLN)	Central necrosis and cavitation create smooth borders with internal heterogeneity [21]
Large Cell	Smooth (65/65 by SRE)	Heterogeneous (42/65 by GLN)	Homogeneous (34/65 by RLN)	Peripheral location with pleural involvement generates smooth margins yet heterogeneous internal density [22]

Feature Importance and Separability

Principal Component Analysis (PCA) confirmed a distinct separation between classes, with the first two principal components (PC1 and PC2) explaining 78.7% of the data variance (Fig. 10). Based on ANOVA F-statistics, the LGRE, RP, and SRE features were identified as having the highest discriminative power (Table 7). This aligns with the observation that the distribution of low-intensity and run patterns is highly effective in capturing morphological differences between lung cancer subtypes.

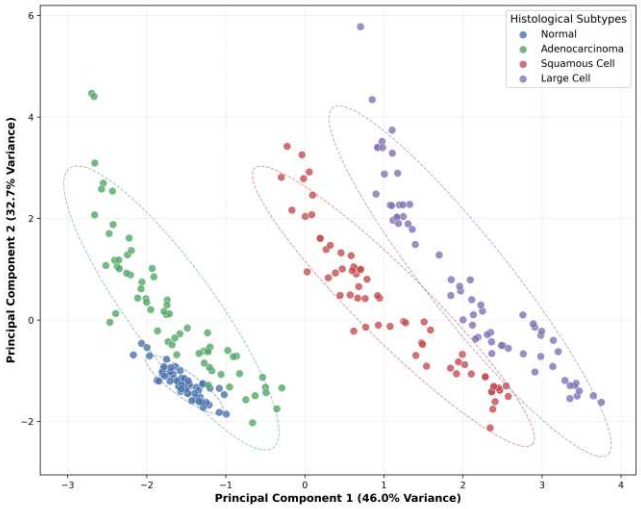


Fig. 10 *Principal Component Analysis (PCA) Visualization. A 2D scatter plot of the first two principal components (PC1 and PC2) explaining 78.7% of the total variance. The plot reveals distinct clustering of the four histological subtypes, validating the effectiveness of GLRLM features in separating the classes*

Table 7. Feature Importance Ranking

Rank	Feature	F-statistic	p-value	Discriminative Power
1	LGRE	1847.3	<0.001	Very High
2	RP	1523.6	<0.001	Very High
3	SRE	1289.4	<0.001	Very High
4	HGRE	782.5	<0.001	High
5	LRE	215.7	<0.001	Moderate
6	GLN	87.3	0.002	Moderate
7	RLN	34.6	0.041	Low

Discussion

Clinical-Pathological Interpretation

The correspondence between GLRLM features and pathological characteristics provides strong clinical validation for these findings [23]. In Adenocarcinoma, a homogeneous rough texture corresponds to the predominance of acinar, papillary, or lepidic histological patterns that maintain the alveolar architecture. Histopathological studies show that the lepidic-predominant subtype retains the pulmonary architecture, resulting in a uniform ground-glass opacity on CT, where gray-level homogeneity correlates with uniform cellularity and minimal necrosis in the early stages [24].

This is in contrast to Squamous Cell Carcinoma, where a homogeneous smooth texture is consistent with a central location and the presence of extensive cavitation. The high LGRE and RP values in this case reflect a low-density necrotic nucleus surrounded by viable tumor tissue [21]. Previous studies have also noted different enhancement patterns due to peripheral vascularization and central necrosis [25]. As for Large Cell Carcinoma, the smooth but heterogeneous texture accurately reflects poorly differentiated histology [22]. Heterogeneity (high GLN) correlates with areas of necrosis, bleeding, and mixed solid-cystic components [25]. Peripheral locations with pleural involvement result in smooth outer margins (high SRE), while chaotic internal architecture creates heterogeneity.

Comparison with Previous Studies and Advantages

Although direct comparisons with previous studies have limitations due to differences in datasets, the position of this study is strengthened by several aspects of novelty. In contrast to previous studies which were mostly binary, this study classified normal and three subtypes of cancer simultaneously. In addition, the use of specific analysis on the ROI of the pulmonary parenchyma improves the signal-to-noise ratio for texture features, as well as provides a more in-depth clinical interpretation than just accuracy metrics [4,6].

Regarding the performance of classifications that achieve perfect accuracy in hold-out test data, this can be explained by several factors. First, the tight selection of ROI on representative areas ensures a strong texture signal. Second, GLRLM features are specifically engineered to capture distinctive histopathological differences between classes. Third, the characteristics of the expert-curated NBIA dataset minimize ambiguous cases that are often the source of misclassification. The results of cross-validation resulting in an average accuracy of 98.5% ± 1.4% confirm that the model has high stability, although external validation in a larger independent cohort is still needed to test generalizations in more

varied real-world clinical cases. Although perfect classification performance was observed in the hold-out test set, this outcome is likely influenced by the relatively small and well-curated dataset. In real-world clinical settings, variability in imaging protocols, tumor heterogeneity, and early-stage lesions may reduce model performance. Therefore, these results should be interpreted as a proof-of-concept rather than definitive clinical performance.

Compared to the often black-box approach of Deep Learning, the GLRLM-SVM method offers a significant advantage in terms of clinical interpretability. Each feature of GLRLM has a clear physical correspondence, so clinicians can understand the basis of the model's decisions and improve diagnostic confidence [12]. In addition, this method shows high data efficiency, being able to achieve competitive performance with only 260 images, in contrast to artificial neural networks which generally require very large datasets [8]. From a computational perspective, the efficiency of training and inference time allows the integration of these methods into real-time clinical workflows without the need for specialized hardware (graphics processing units, GPUs), making it a viable solution for implementation in resource-constrained environments. However, while SVM demonstrates strong performance on small to medium-sized datasets, its scalability to larger datasets is limited due to increased computational complexity and training time. Therefore, for large-scale clinical deployment involving extensive imaging databases, alternative approaches such as deep learning models may be more suitable, provided that interpretability challenges are addressed.

Clinical Implications, Limitations, and Future Directions

The GLRLM-SVM system has the potential to be integrated as a second-reader to assist radiologists in subtype classification, which can reduce inter-observer variability [3]. Accurate identification of subtypes is crucial for therapy planning, given the different responses to treatment, such as the contraindications of bevacizumab for Squamous cell carcinoma patients [26]. In addition, texture-based predictions can guide the selection of biopsy sites, especially in heterogeneous tumors, to ensure representative sampling. Since standard CT-guided percutaneous biopsies typically require a minimum lesion size of 8 to 10 mm, the spatial calibration of our ROI into true physical dimensions ensures that radiologists can confidently target these clinically actionable lesions. Importantly, this approach is not intended to replace histopathological examination, which remains the gold standard for lung cancer diagnosis. Instead, the system may serve as a decision-support tool to assist radiologists and guide biopsy site selection, particularly in heterogeneous tumors.

However, the study has some limitations, including a relatively small sample size for robust validation, the use of a single modality (CT only), and the majority of cases are in advanced stages (II–III), and early-stage (stage I) cancers were not included. This limitation may affect the model's applicability in early detection scenarios. A detailed summary of these limitations, along with our proposed solutions and future research directions, is provided in Table 8. Future research directions should include external validation in multicenter cohorts, development of 3D texture analysis for volumetric data, as well as multi-modal integration with positron emission tomography (PET) data or genomic markers [23]. The combination of GLRLM features with deep learning architectures is also a promising opportunity to leverage the advantages of interpretability and representational learning simultaneously.

Table 8. Summary of Study Limitations and Proposed Future Directions

Limitation	Description	Proposed Solution / Future Direction
Small Sample Size	260 images may not capture full population variance	External validation with multi-center, larger independent cohorts
Single Modality	Relies solely on CT scan images	Multi-modal integration combining CT with PET data or genomic markers
Advanced Stages Bias	Majority of cases are Stage II-III	Inclusion of early-stage (Stage I) pulmonary nodules in future training datasets
2D Texture Analysis	Evaluates only a single axial slice (2D)	Development of 3D GLRLM texture analysis for volumetric tumor data
ROI Extraction	Requires operator intervention, which may introduce bias	Integration of fully automated deep learning-based segmentation prior to feature extraction

Conclusion

This study demonstrated that the classification system based on the Gray Level Run Length Matrix (GLRLM) and Support Vector Machine (SVM) was able to identify lung cancer subtypes with high performance on a controlled dataset of CT-scan thorax images. Detailed texture characterization reveals specific patterns for each subtype that correlate with histopathological characteristics: Normal with heterogeneous rough texture, Adenocarcinoma with homogeneous rough texture, Squamous Cell Carcinoma with homogeneous smooth texture, and Large Cell Carcinoma with a heterogeneous smooth texture. This approach offers advantages in terms of clinical interpretability, data efficiency, and computational feasibility over deep learning methods, making it a promising alternative or complement to clinical decision support systems, especially in resource-constrained environments. However, this model should be considered as a decision-support tool rather than a replacement for histopathological diagnosis, and further validation on larger, heterogeneous datasets is required.

Acknowledgements

Not applicable

References

1. Bray F, Laversanne M, Sung H, et al (2024) Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 74:229-263. https://doi.org/10.3322/caac.21834
2. Travis WD, Brambilla E, Nicholson AG, et al (2015) The 2015 World Health Organization classification of lung tumors. J Thorac Oncol 10:1243-1260.
3. Gould MK, Donington J, Lynch WR, et al (2013) Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 143:e93S-e120S.
4. Yunianto M, Soeparmi S, Cari C, et al (2021) Classification of lung cancer using Naïve Bayes with variation of filters and feature extraction using gray level co-occurrence matrix (GLCM). Indones J Appl Phys 11:256-267.
5. Desiani A, Indra Maiyanti S, Andriani Y, et al (2023) Comparison of lung cancer disease classification using Support Vector Machine and K-Nearest Neighbor. Jurnal PROCESSOR 18:54-62.
6. Khultsum U, Sarasati F, Taufik G (2022) Application of MobileNet method for classification of lung cancer disease images. JURIKOM 9:1366-1373.
7. Benaya D (2024) Implementation of Random Forest in lung cancer classification. JOINTER 5:27-31.
8. Litjens G, Kooi T, Bejnordi BE, et al (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60-88.
9. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278:563-577.
10. Galloway MM (1975) Texture analysis using gray level run lengths. Comput Graph Image Process 4:172-179.
11. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273-297.
12. Linardatos P, Papastefanopoulos V, Kotsiantis S (2021) Explainable AI: a review of machine learning interpretability methods. Entropy 23:18.
13. Clark K, Vendt B, Smith K, et al (2013) The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 26:1045-1057.
14. Li P, Wang S, Li T, et al (2020) A Large-Scale CT and PET/CT Dataset for Lung Cancer Diagnosis (Version 5) [Dataset]. The Cancer Imaging Archive. https://doi.org/10.7937/TCIA.2020.NNC2-0461
15. Vabalas A, Gowen E, Poliakoff E, Casson AJ (2019) Machine learning algorithm validation with a limited sample size. PLoS One 14:e0224365.
16. Webb WR, Müller NL, Naidich DP (2014) High-Resolution CT of the Lung. 5th edn. Wolters Kluwer, Philadelphia.
17. Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Trans Syst Man Cybern SMC-3:610-621.
18. Hsu CW, Lin CJ (2002) A comparison of methods for multiclass support vector machines. IEEE Trans Neural Netw 13:415-425.
19. Hansell DM, Bankier AA, MacMahon H, et al (2008) Fleischner Society: glossary of terms for thoracic imaging. Radiology 246:697-722.
20. Travis WD, Brambilla E, Noguchi M, et al (2011) International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol 6:244-285.
21. Lewis JW, Pearlberg JL, Beute GH, et al (1990) Can computed tomography of the chest stage lung cancer? Yes and no. Ann Thorac Surg 49:591-595.
22. Kim HY, Shim YM, Lee KS, et al (2007) Persistent pulmonary nodular ground-glass opacity at thin-section CT: histopathologic comparisons. Radiology 245:267-275.
23. Ganeshan B, Goh V, Mandeville HC, et al (2013) Non-small cell lung cancer: histopathologic correlates for texture parameters at CT. Radiology 266:326-336.
24. Austin JH, Garg K, Aberle D, et al (2013) Radiologic implications of the 2011 classification of adenocarcinoma of the lung. Radiology 266:62-71.
25. Ganeshan B, Panayiotou E, Burnand K, et al (2012) Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. Eur Radiol 22:796-802.
26. Sandler A, Gray R, Perry MC, et al (2006) Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. N Engl J Med 355:2542-2550.

Statements & Declarations

Funding

This research was supported by personal funds.

Competing Interests

The authors have no competing interests to declare that are relevant to the content of this article.

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Muhammad Baqibillah. The first draft of the manuscript was written by Muhammad Baqibillah, and Mokhammad Tirono commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Ethics approval

Ethical approval was waived by the institutional ethics committee because this study utilized fully anonymized secondary data obtained from a publicly accessible repository, The Cancer Imaging Archive (TCIA). As the dataset does not contain identifiable patient information, informed consent was not required. All procedures performed were in accordance with the ethical standards of the institutional research committee and with the 1964 Declaration of Helsinki and its later amendments.

Consent to participate

Not applicable. The study used only anonymized secondary data from a publicly accessible repository.

Consent to publish

Not applicable. No potentially identifiable human images, personal details, or clinical video data are presented in this study.

Data availability

The imaging dataset analyzed during the current study is available in The Cancer Imaging Archive (TCIA) repository at https://doi.org/10.7937/TCIA.2020.NNC2-0461 [14]

Code availability

The custom MATLAB code used during the current study is available from the corresponding author on reasonable request.