CERASUS JOURNAL OF MEDICINE

ORIGINAL ARTICLE

Classification of histological subtypes of non-small cell lung cancer using computerized tomography texture analysis

Tumay Bekci[1]
Merve Nur Tasdemir[1]
Esma Cinar[2]
Demet Sengul[2]
Eylem Karacay[3]
Sevval Arslan[3]
Sena Nur Cure[3]

1. *Giresun University Faculty of Medicine, Department of Radiology, Giresun, Türkiye.*
2. *Giresun University Faculty of Medicine, Department of Pathology, Giresun, Türkiye.*
3. *Giresun University Faculty of Medicine, Giresun, Türkiye.*

Corresponding Author: Merve Nur Tasdemir

Address:
Giresun University Faculty of Medicine, Department of Radiology, Giresun, Türkiye

Email: mervetsdmr@gmail.com

Received: 18 August 2024
Accepted: 31 August 2024
Published: 11 October 2024

Abstract

Objective: This study aimed to differentiate between the two main histological subtypes of non-small cell lung cancer using a non-invasive technique, computerized tomography texture analysis.

Method: We included 53 patients. All patients were histopathologically proven non-small cell lung cancer cases. All patients underwent thorax CT scans. In CT images, the differences present in the texture features of adenocarcinoma and squamous cell carcinoma, which are the two main histological subtypes of non-small cell lung cancer, were determined by the consensus of two radiologists for computerized tomography-based texture analysis.

Results: A total of 44 texture features were extracted, including 12 first-order features and 32 second-order features derived from gray-level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM), neighborhood gray-level different matrix (NGLDM), and gray-level zone length matrix (GLZLM) features in 51 CT images. None of the evaluated texture parameters were statistically significant. However, in patients with squamous cell lung cancer, the values of Intensity Histogram, NGTDM Complexity, and Intensity Based Robust Mean Absolute Deviation higher from adenocarcinoma patients and had the highest area under the curve in ROC analyses (0.727, 0.664, 0.666 respectively). We consider that the high values of these parameters in the squamous cell subtype, due to high intratumoral heterogeneity.

Conclusion: Intensity Histogram, NGTDM Complexity, and Intensity Based Robust Mean Absolute Deviation features can be used to differentiate between the subtypes of non-small cell lung cancer, adenocarcinoma and squamous cell carcinoma. These features were highly associated with the high intratumoral heterogeneity of squamous cell lung cancer.

Keywords: Artificial intelligence; CT; non-small cell lung cancer; radiomics; texture analysis

You may cite this article as: Bekci T, Tasdemir MN, Cinar E, et al. Classification of histological subtypes of non-small cell lung cancer using computerized tomography texture analysis. *Cerasus J Med.* 2024;1(3):168-172. doi: 10.70058/cjm.1535113

Introduction

Non-small cell lung cancer (NSCLC) is the leading cause of cancer-related deaths worldwide[1]. Non-small cell lung cancer has two main histological subtypes: adenocarcinoma and squamous cell carcinoma. These two subtypes exhibit different cellular structures and characteristics, thus requiring different treatment approaches [1, 2]. The accurate diagnosis of these subtypes is important for determining appropriate treatment options. Adenocarcinoma, typically develops in the peripheral lung tissue and is more commonly seen in non-smokers. Adenocarcinoma can often be surgically removed, but radiation and chemotherapy are also among the treatment options [2]. Additionally, targeted therapies are available for some cases of adenocarcinoma. Squamous cell carcinoma, on the other hand, generally develops in the central bronchi or large airways and is more frequently observed in smokers[2]. Surgery, radiation, and chemotherapy can be used in the treatment of squamous cell carcinoma. its nonspecific symptoms and the necessity of invasive procedures for diagnosing the specific subtypes of non-small cell lung cancer, which require different treatment approaches, complicates the process of diagnosis. Non-small cell lung cancer (NSCLC) are typically noticed in advanced stages, and the symptoms often vary depending on the size, location, and extent of tumor spread [3-5]. Because of this many patients receiving a diagnosis at advanced stages when treatment options are limited. Therefore, there is an urgent need for continued research efforts aimed at unraveling the complexities of NSCLC and developing innovative strategies for early detection, personalized treatment, and improved outcomes[4, 5]. Currently the invasive methods are preferred for the identification of NSCLC subtypes but they carry a high risk of complications such as pulmonary hemorrhage and pneumothorax. Therefore, the need for non-invasive methods that can use for the NSCLC diagnosis is increasing[4,5]. Texture analysis (TA) is an emerging technique that allows for the analysis of the distribution of pixel intensities and transforms digital medical images into mineable data by extracting quantitative features mathematically[6]. TA is a promising method, and the texture data obtained can be used in deep learning algorithms for diagnosis of NSCLC subtypes. NSCLC , in contrast , is a difficult process to diagnose and the fact that the subtypes of NSCLC can be differentiated by TA can provide early diagnosis and treatment of NSCLC. We think that TA may help differentiate adenocarcinoma and squamous cell carcinoma based on legion texture characteristics in CT. Accordingly, we aimed to investigate the feasibility and accuracy of TA for differentiating NSCLC's subtypes adenocarsinoma and squamous cell carcinoma on CT images [6-8].

Materials and Methods

Radiomics workflow

The radiomics flow of this study included: (1) images acquisition, (2) image segmentation, (3) feature extraction, (4) data analysis. All the steps are shown in Fig. 1.

Patients

This retrospective study obtained approval from the institutional review board of our hospital, and written informed consent was waived [GEAH/KAEK-216]. Patients with histopathologically confirmed lung cancer who underwent CT examinations before biopsy procedure between January 2020 and April 2023 were identified from our hospital's database. Inclusion criteria comprised histopathologically confirmed lung cancer and preoperative CT examination. Exclusion criteria included CT images affected by motion artifacts.

CT Image Acquisition and Tumor Segmentation:

All CT images were obtained on a 128-Slice GE Revolution EVO CT Scan Machine with a breath-held helical acquisition of the entire thorax. CT parameters were as follows: tube voltage = 120 kVp; tube current = 150 mAs; detector collimation = 0.5 mm × 64; pitch = 0.625; rotation time = 0.5 s; reconstruction slice thickness = 1 mm; matrix = 512 × 512; field of view = 407 mm. All CT images were analyzed by two radiologists independently. Both radiologists were informed of the location of each lesion but were blinded to the pathological diagnosis. Tumor segmentation was performed manually by two radiologists with 11 and 4 years of experience in thoracal imaging using LIFEx software [www.lifexsoft.org] [9]. CT images were exported in Digital Imaging and Communications in Medicine (DICOM) format from the hospital database to LIFEx software. The region of interest encompassed the largest cross-sectional area of tumors in axial planes selected on CT images (Figure 1). All tumoral tissue, including necrosis, was included. Following tumor segmentation, texture feature extraction was performed.

Texture Feature Extraction:

Texture Analysis (TA) was conducted on two-dimensional images of segmented tumoral lesions on axial planes using LIFEx software. All CTs were

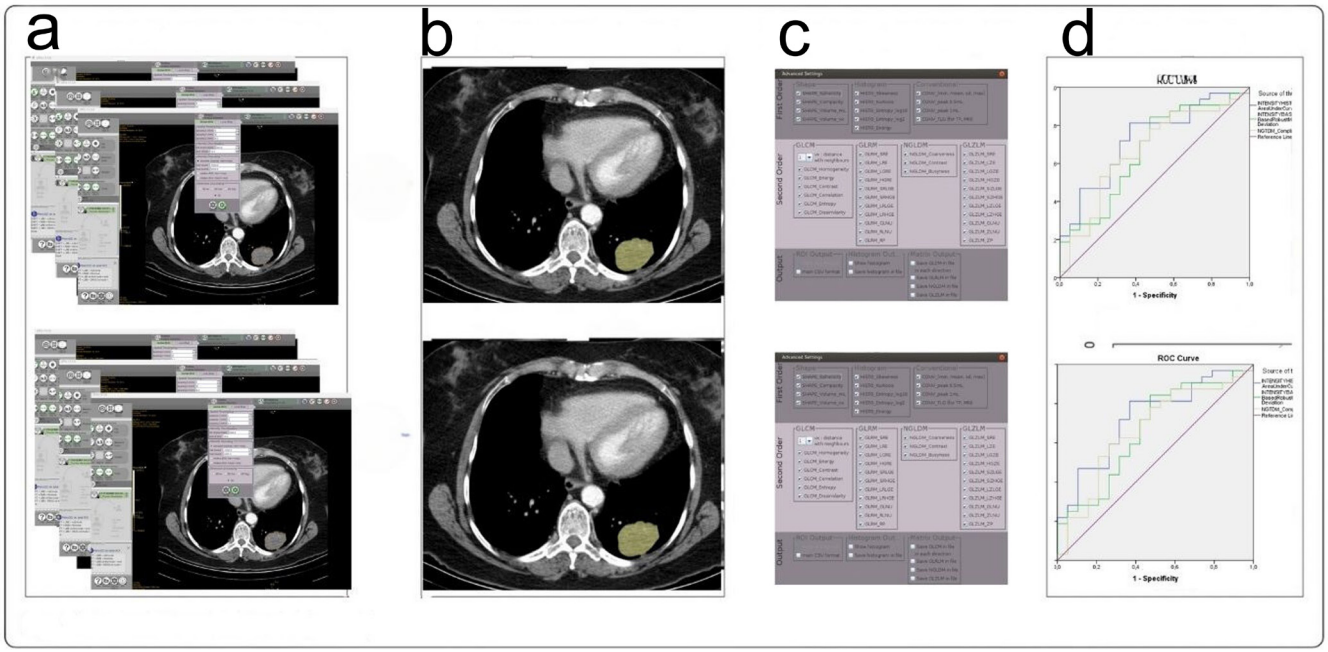

Figure 1. The radiomics flow of this study included: (a) images acquisition, (b) image segmentation, (c) feature extraction, (d) data analysis.

resampled to a voxel size of 1 × 1 × 3 mm (X spacing, Y spacing, Z spacing). A total of 44 texture features were extracted, including 12 first-order features and 32 second-order features derived from gray-level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM), neighborhood gray-level different matrix (NGLDM), and gray-level zone length matrix (GLZLM) features.

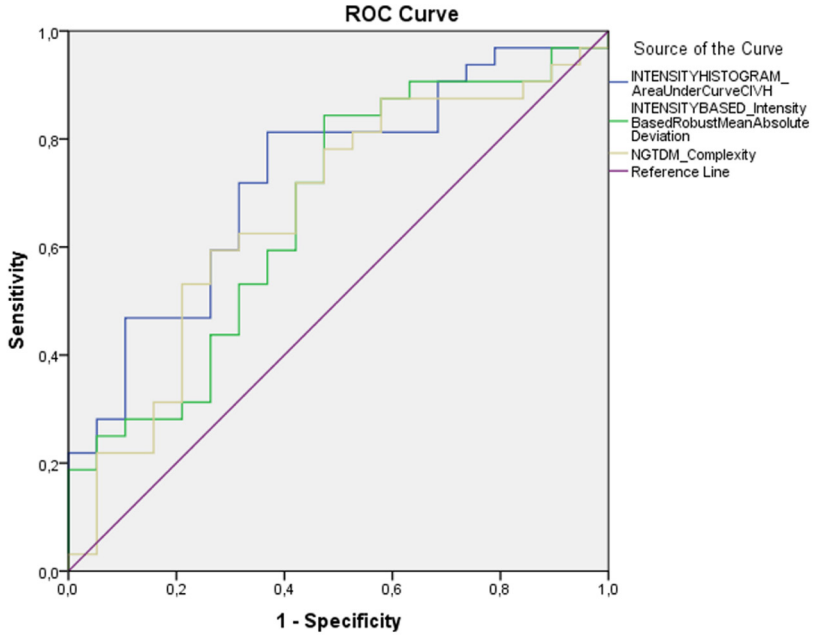

Figure 2. In patients with squamous cell lung cancer, the values of Intensity Histogram, NGTDM Complexity, and Intensity Based Robust Mean Absolute Deviation higher from adenocarcinoma patients and had the highest area under the curve in ROC analyses.

Statistical Analysis:

IBM SPSS V23 was used for statistical analyses. Normality distributions of quantitative parameters were assessed using the Shapiro-Wilk test. The Mann Whitney U test compared data not conforming to normal distributions. Receiver Operator Characteristic (ROC) analysis evaluated diagnostic test performance, and sensitivity and specificity were assessed. Data are presented as mean [95% CI]. A p value < 0.05 was considered statistically significant.

Results

Patient Characteristics:

53 of 51 patients (14 females and 37 males; mean age, 64 years; maximum age, 85; minimum age, 37;) affected by histologically confirmed NSCLC were retrospectively enrolled. 2 patients with motion artifacts excluded from the study. 32 patients had lung adenocarcinoma and 19 had squamous cell lung adenocarcinoma.

Texture Features:

None of the evaluated texture parameters were statistically significant. However, in patients with squamous cell lung cancer, the values

of Intensity Histogram, NGTDM Complexity, and Intensity Based Robust Mean Absolute Deviation higher from adenocarcinoma patients and had the highest area under the curve in ROC analyses (0.727, 0.664, 0.666 respectively) (Figure 2).

Discussion

In our study assessing the feasibility of using radiomics-based texture parameters for subtype recognition of non-small cell lung cancers, including adenocarcinoma and squamous cell carcinoma, we found that in squamous cell carcinomas, the values of Intensity histogram, NGTDM Complexity, and Intensity Based Robust Mean Absolute Deviation were high, with the highest area under the curve in ROC analyses (0.727, 0.664, 0.666 respectively).

It is known that squamous cell lung cancer is the tumor type with the highest genetic heterogeneity among non-small cell lung cancers. Squamous cell lung cancer exhibits heterogeneity associated with its genetic structure. Common genetic alterations include mutations in the TP53 gene, frequently observed in squamous cell lung cancer, contributing to tumor heterogeneity and resistance to treatment [1-4]. Fang et al. [5] reported that p53 is the most frequently mutated gene; KRAS, EGFR, MLL3, and STK11 the most frequently mutated genes in adenocarcinomas; PI3KCA, SOX2, CDK2, P63, and FGFR1 the most frequently mutated genes in squamous cell carcinomas; and RB1, MLL2, SMO, and PI3KCA the most frequently mutated genes in small cell lung cancer. Copy number variations (CNVs) are also prevalent in squamous cell lung cancer, indicating the presence of diverse subpopulations of tumor cells and contributing to intratumoral heterogeneity[1-4]. Short tandem repeats (STRs) may also display heterogeneous distribution in squamous cell lung cancer, with variations in length or number contributing to genetic diversity among tumor cells. Furthermore, epigenetic changes, such as DNA methylation, histone modifications, and non-coding RNAs, can contribute to heterogeneity by regulating gene expression differently in distinct subgroups of tumor cells [8-9]. Clonal evolution, characterized by genetic changes over time and the emergence of various subclones, further contributes to intratumoral heterogeneity in squamous cell lung cancer. These genetic alterations and clonal evolution serve as the basis for heterogeneity in squamous cell lung cancer, influencing its biological behavior and response to treatment[1-4]. In this context, it has been demonstrated that texture analysis parameters can be effectively used to indicate intratumoral heterogeneity. In our study, we believe that the parameters with the highest area under the ROC curve, namely Intensity histogram, NGTDM Complexity, and Intensity Based Robust Mean Absolute Deviation, are effective parameters for demonstrating heterogeneity. We consider that the high values of these parameters in the squamous cell subtype, with high intratumoral heterogeneity, are consistent with the texture features represented by these parameters

NGTDM is a feature set that measures the relationships between grayscale pixels in an image. Complexity denotes the complexity of the NGTDM matrix, reflecting the complexity of relationships between grayscale pixels. Elevated NGTDM Complexity values may indicate high complexity in the relationships between different grayscale pixels, thereby expressing intratumoral heterogeneity, as the complexity of relationships between grayscale pixels may vary across different regions[8-10].

Intensity histogram is a graphical representation illustrating the distribution of pixel intensities within an image. Typically, the intensity values of pixels are displayed in a histogram, which indicates the frequency of intensity values, i.e., the number of pixels at each intensity level. A high intensity histogram in a tumor may indicate variations in intensity among different regions, which can be associated with intratumoral heterogeneity. Robust Mean Absolute Deviation term is a measure of the average absolute difference of each value from the mean of the values in a dataset. The term ‘robust’ indicates resistance to outliers, meaning that the influence of outliers is minimized [8-10]. Thus, this feature can be used to measure the variance or distribution of a specific point or region based on the intensity of an image or dataset. This feature is particularly important in fields such as image processing or analysis, as it can help measure changes in intensity within a specific region. The association of high values of these parameters with high intratumoral heterogeneity in squamous cell lung carcinomas may be logical. This is because these parameters can quantitatively assess intratumoral heterogeneity by measuring the complexity of intensity or grayscale pixel relationships between different regions. These findings indicate that intratumoral heterogeneity is high in squamous cell lung carcinomas and that the parameters used to analyze this heterogeneity are clinically relevant. The use of these parameters can play a critical role in areas such as better

patient classification, prediction of treatment response, and prognosis determination.

Although our study has some limitations likely limited by the small size of the patient cohort, resulting in the inability to identify a statistically significant texture parameter we think that this study provides important findings that will guide future research. Particularly, with larger patient cohorts and analyses supported by clinical data, it will be possible to better understand the impact of these parameters on clinical outcomes. Additionally, the potential of these parameters in personalized treatment strategies should be evaluated.

In conclusion, the use of texture parameters as an indicator of intratumoral heterogeneity with CT texture analysis in squamous cell lung carcinomas is an important step in the diagnosis and development of more effective and personalized treatment approaches.

Funding: There is no institution or person supporting this study.

Conflict of Interest: None of the authors have a conflict of interest.

Authors' contribution: Concept: M.N.T, Design: T.B, E.C, Data Collection or Processing: T.B, E.C, D.S, Literature Search: M.N.T, D.S, E.K, S.A, S.N.C, Writing: T.B.

Ethical Declaration: Ethics approval for the study was obtained from the Non-Interventional Clinical Research Ethics Committee of Giresun Training and Research Hospital with decision number 23.10.2023/24.

References

1. Megyesfalvi Z, Gay CM, Popper H, et al. Clinical insights into small cell lung cancer: Tumor heterogeneity, diagnosis, therapy, and future directions. *CA Cancer J Clin*. 2023;73(6):620-652.
2. Chen Z, Fillmore CM, Hammerman PS, Kim CF, Wong KK. Non-small-cell lung cancers: a heterogeneous set of diseases. *Nat Rev Cancer*. 2015 Apr;15(4):247].
3. de Sousa VML, Carvalho L. Heterogeneity in Lung Cancer. *Pathobiology*. 2018;85(1-2):96-107.
4. Wu F, Fan J, He Y, et al. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat Commun*. 2021;12(1):2540.
5. Fang B, Mehran RJ, Heymach JV, Swisher SG. Predictive biomarkers in precision medicine and drug development against lung cancer. *Chin J Cancer*. 2015;34(7):295-309.
6. Phillips I, Ajaz M, Ezhil V, et al. Clinical applications of textural analysis in non-small cell lung cancer. *Br J Radiol*. 2018;91(1081):20170267.
7. Suga M, Nishii R, Miwa K, et al. Differentiation between non-small cell lung cancer and radiation pneumonitis after carbon-ion radiotherapy by ^{18}F-FDG PET/CT texture analysis. *Sci Rep*. 2021;11(1):11509.
8. Chen S, Harmon S, Perk T, et al. Using neighborhood gray tone difference matrix texture features on dual time point PET/CT images to differentiate malignant from benign FDG-avid solitary pulmonary nodules. *Cancer Imaging*. 2019;19(1):56.
9. Nioche C, Orlhac F, Boughdad S, et al. LIFEx: A Freeware for Radiomic Feature Calculation in Multimodality Imaging to Accelerate Advances in the Characterization of Tumor Heterogeneity. *Cancer Res*. 2018;78(16):4786-4789.
10. Bekci T, Cakir IM, Aslan S. Differentiation of affected and nonaffected ovaries in ovarian torsion with magnetic resonance imaging texture analysis. *Rev Assoc Med Bras (1992)*. 2022;68(5):641-646.