Journal of Imaging Informatics in Medicine (2025) 39:1692–1714
https://doi.org/10.1007/s10278-025-01596-2

ORIGINAL PAPER

Classification of Renal Lesions by Leveraging Hybrid Features from CT Images Using Machine Learning Techniques

Ravinder Kaur[1] · **Sonam Khattar**[2] · **Sanjay Singla**[2]

Received: 16 January 2025 / Revised: 19 June 2025 / Accepted: 21 June 2025 / Published online: 14 July 2025

Abstract

Renal cancer is amid the several reasons of increasing mortality rates globally, which can be reduced by early detection and diagnosis. The classification of lesions is based mostly on their characteristics, which include varied shape and texture properties. Computed tomography (CT) imaging is a regularly used imaging modality for study of the renal soft tissues. Furthermore, a radiologist's ability to assess a corpus of CT images is limited, which can lead to misdiagnosis of kidney lesions, which might lead to cancer progression or unnecessary chemotherapy. To address these challenges, this study presents a machine learning technique based on a novel feature vector for the automated classification of renal lesions using a multi-model texture-based feature extraction. The proposed feature vector could serve as an integral component in improving the accuracy of a computer aided diagnosis (CAD) system for identifying the texture of renal lesion and can assist physicians in order to provide more precise lesion interpretation. In this work, the authors employed different texture models for the analysis of CT scans, in order to classify benign and malignant kidney lesions. Texture analysis is performed using features such as first-order statistics (FoS), spatial gray level co-occurrence matrix (SGLCM), Fourier power spectrum (FPS), statistical feature matrix (SFM), Law's texture energy measures (TEM), gray level difference statistics (GLDS), fractal, and neighborhood gray tone difference matrix (NGTDM). Multiple texture models were utilized to quantify the renal texture patterns, which used image texture analysis on a selected region of interest (ROI) from the renal lesions. In addition, dimensionality reduction is employed to discover the most discriminative features for categorization of benign and malignant lesions, and a unique feature vector based on correlation-based feature selection, information gain, and gain ratio is proposed. Different machine learning–based classifiers were employed to test the performance of the proposed features, out of which the random forest (RF) model outperforms all other techniques to distinguish benign from malignant tumors in terms of distinct performance evaluation metrics. The final feature set is evaluated using various machine learning classifiers, with the RF model achieving the highest performance. The proposed system is validated on a dataset of 50 subjects, achieving a classification accuracy of 95.8%, outperforming other conventional models.

Keywords Renal cancer · Analysis of texture · Lesion classification · Texture analysis · CT images · Machine learning · Feature selection

Introduction

Kidney or renal cancer is amid the most common forms of cancers, and as per the report published by the American Cancer Society (ACS), new kidney cancer cases and deaths are expected to reach 81,610 and 14,390 respectively, in 2024 [1]. The appropriate treatment of kidney cancer depends on early diagnosis [2, 3], and distinct types of imaging modalities such as ultrasound (US), computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) are commonly utilized for non-invasive diagnostic imaging procedures [4–6]. Among all

✉ Ravinder Kaur
ravinder.kaur7@yahoo.com

Sonam Khattar
sonam24jun@gmail.com

Sanjay Singla
Sanjay.e13538@cumail.in

[1] Thapar Institute of Engineering and Technology, Patiala, Punjab, India

[2] Department of Computer Science, Chandigarh University, Gharuan, Mohali, Punjab, India

imaging techniques, CT imaging generates high-resolution images and is considered a standard imaging modality for the classification of renal lesions as compared to other medical imaging modalities [7–14]. The categorization of renal lesions using CT images is more challenging due to more complexity and ambiguity, as well as the fact that the pixel values in medical images are somewhat equivalent. Further, the conventional manual analysis of voluminous medical data for the classification of renal lesions is tedious, laborious, and time-consuming due to the similarity in medical images. Moreover, traditional techniques of detecting kidney abnormalities using manual analysis are susceptible to inter or intrasubject variabilities. Thus, there is need to develop a computer aided diagnosis (CAD) system for classification of renal lesions using CT images, which can serve as a second opinion for doctors for performing lesion classification. Moreover, the computerized texture analysis of CT images is significant because it reveals important features that would otherwise be arduous to identify through visual inspection of CT scans. As it is clearly visible from Fig. 1, how the texture patterns for benign and malignant lesion varies, and this information about image texture could play a significant role in performing classification of renal lesions.

Over the last few decades, different machine learning and deep learning techniques have been employed for performing medical disease diagnosis, such as brain tumor segmentation, skin disease detection, Parkinson's disease detection, and many other disorders [15–17]. Thus, there is a wider scope to improve the diagnosis results for different medical issues by deploying computerized systems based on machine learning and deep learning. In case of renal lesion classification, the ability of the radiologist to perceive changes in image textural features and link them to pathological outcomes is often used as a visual criterion for recognizing kidney lesions. Despite widespread disagreement among trained radiologists on the interpretation of renal lesions using CT visual analysis, particularly in borderline scenarios, the reported diagnosis accuracy is around 79% for classification of renal lesions using CT visual analysis [18]. Due to the restricted accuracy of visual inspection, an impartial technique for lesion detection that relies on measurable texture analysis is necessary. As a result, computer-assisted medical image categorization is desirable in order to perform reliable image interpretation and aid doctors in disease diagnosis. As a result, there is a higher demand for accurate CAD systems that can simplify and speed up CT image interpretation. These technologies can be extremely helpful in diagnosing and treating patients. According to the literature, limited research work has been reported that explores the varied texturing models' capabilities for selecting best feature vector for renal lesion classification. In order to determine the accuracy and reliability of renal lesion classification, an effort has been made in this research work to suggest a classification approach employing a new feature vector by utilizing different feature selection techniques that include correlation-based feature selection (CFS), information gain, and gain ratio strategy. First, a fixed-size ROI is chosen in order to extract features utilizing texture models that have been explored in this study. Then, using the considered feature selection techniques, an ideal feature vector with different texture features is selected. As a result, the proposed feature was vectorized and fed into a different machine learning classifier to discriminate among cancerous and non-cancerous cases. Moreover, a crucial aspect of classification is choosing the best features. To that end, this research lays out a strategy for renal lesion categorization based on a novel set of hybrid features extracted from separate texture models.

The core research problem addressed in this study is the lack of an accurate and automated method for distinguishing benign and malignant renal lesions in CT scan images. Existing approaches often rely on limited features extracted from very few texture models, which can hinder clinical adoption. This study aims to solve this issue by proposing a hybrid feature-based technique that combines multiple texture analysis models and optimized feature selection strategies to enhance classification performance and support clinical decision-making. This study introduces a novel hybrid feature vector, optimized via advanced feature selection techniques, for improved classification of kidney lesions

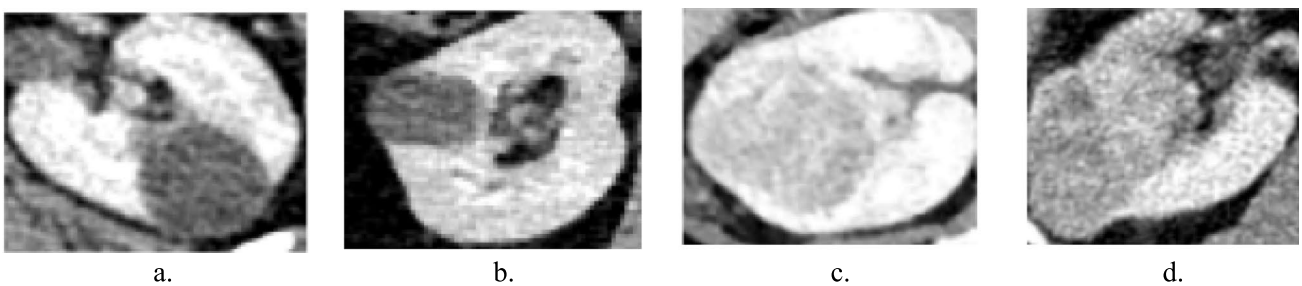

Fig. 1 Illustration of sample renal lesion images where **a**, **b** characterize the texture patterns of benign lesion whereas **c**, **d** represent the texture patterns of a malignant lesion

from CT images. The proposed feature vector could serve as an integral component in improving the accuracy of the computer-aided diagnosis (CAD) system for identifying the texture of renal lesions and can assist physicians to provide more precise lesion interpretation.

The rest of the article is structured as follows: the existing work reported in the same field has been mentioned in "Related Work". "Methodology" specifies the details related to the proposed methodologies and approaches that were used to conduct this research. "Experimental Results and Discussion" presents and discusses the outcome of the investigation as well as examination of the envisaged categorization technique in terms of distinct performance evaluation metrics. Lastly, "Clinical Relevance and Practical Application" concludes the outcomes of the conducted study.

Related Work

Different researchers have been working for recognition and identification of renal lesions in the previous decade, and they may be separated based on the retrieved characteristics and classifier utilized for categorization. Linguraru et al. [19] created a computerized diagnostic tool for kidney tumor categorization using support vector machine (SVM). They utilized histogram-based curvature features with intensity and enhancement descriptors to characterize hereditary renal lesions. Kidney lesions from CT images of 43 persons were used to test the tool's performance. Raman et al. [20] recently published a pilot study that used random forests to classify renal masses using computed tomography texture analysis (CTTA). In order to identify the renal cancer, CTTA assessed texture features which include entropy, mean gray-level intensity, the standard deviation, and skewness and kurtosis and further used a random forest (RF) model for lesion categorization. The study had one flaw: it only looked at four types of common renal masses. Other lesions, such as poor lipid angiomyolipoma, hyperdense cysts, and chromophobe RCCs, would almost certainly be justified in a bigger validation study. Furthermore, only four texture features were examined in this work, and the classification skills of multiple texture models were not examined. Liu et al. [21] suggested a computerized technique for detecting renal problems. Recently, QH et al. [22] proposed an ensemble classifier for the categorization of cystic renal lesions by employing radiomics features and reported an accuracy equivalent to 90.5%. However, the study was limited as they considered only first-order statistics, shape features, and second-order features. Further, Xu et al. [23] proposed a framework based on machine learning for the classification of renal lesions where they utilized both clinical and radiomic features. The investigation demonstrated that the use of machine learning models with CT scans and clinical data shows potential for accurately categorizing the likelihood of kidney cancer. Miskin et al. [24] also used machine learning methods based on CT scan textures to distinguish between benign and malignant cystic kidney tumors. The authors compared the performance or RF, logistic regression, and SVM, and out of the considered three classifiers, RF outperformed with reported specificity equivalent to 0.87. Han JH et al. [25] present a fully automated deep learning system using 3D U-Net and random forest classifiers for segmentation and classification of renal tumors on CT scans. Trained on 561 CT scans, it achieved a Dice coefficient of 0.83 for tumors >4 cm and 0.65 for those <4 cm. Classification accuracy reached 0.77 for RCC subtypes and 0.85 for benign vs. malignant tumors. According to the literature, limited research work has been reported that explores the varied texturing models' capabilities for selecting the best feature vector for renal lesion classification. In order to determine the accuracy and reliability of the renal lesion classification, an effort has been made in this research work to suggest a classification approach employing a new feature vector by utilizing some different feature selection techniques that include correlation-based feature selection (CFS), information gain, and gain ratio strategy. This study introduces a novel hybrid framework that integrates statistical texture analysis with random forest-based classification, optimized via advanced feature selection techniques, for improved classification of kidney lesions from CT images.

Methodology

Proposed Technique

The simulation model developed in this study integrates feature extraction, selection, and classification in a modular and automated workflow. The fundamental objective of this work is to suggest a set of hybrid features for performing computer-assisted diagnosis of kidney lesions using hidden texture information. Features derived from several texture models have been merged together to improve classification accuracy. It processes CT images by extracting texture features using multiple models (e.g., first-order statistics (FoS), spatial gray level co-occurrence matrix (SGLCM), gray level difference statistics (GLDS), Fractal), reduces dimensionality using three selection techniques (CFS, Information gain (IG), gain ratio (GR)), and evaluates classification performance using various machine learning algorithms.

Because subjective visual assessment is tedious and it takes a lot of time, the primitive aim of this study paper is as follows: (i) categorize renal lesions based on the categorization on features related to texture, (ii) evaluate the capability of several textural models in representing renal tissue, (iii) use the feature selection strategy for improving

the classifier's performance, and (iv) use numerous feature selection and classification techniques, for evaluating the performance of a proposed hybrid set of features. Fig. 2 displays the main stages of the proposed classification approach, which comprise operations such as segmentation of the region of interest, extraction of texture features, selection of dominating features, and classification, which are detailed in the following sub-sections.

The underlying physical basis for utilizing texture analysis in renal lesion classification stems from the heterogeneity in tissue composition between benign and malignant lesions. Malignant renal tumors often exhibit irregular vascularization, necrosis, and cellular density variations, which manifest as complex and coarse textures in CT images. In contrast, benign lesions typically present with more homogeneous and smooth intensity patterns. Texture features such as contrast, entropy, and fractal dimensions quantitatively capture these variations in pixel intensity and spatial distribution, reflecting the physical differences in tissue microstructure. Therefore, the application of texture models provides a measurable, image-based biomarker that correlates with clinically relevant morphological characteristics, enabling more accurate lesion categorization.

ROI Extraction

To avoid any distortions, a region of interest of is manually picked from the lesion in the image's center. The small region of interest (ROI) size was chosen to take into account

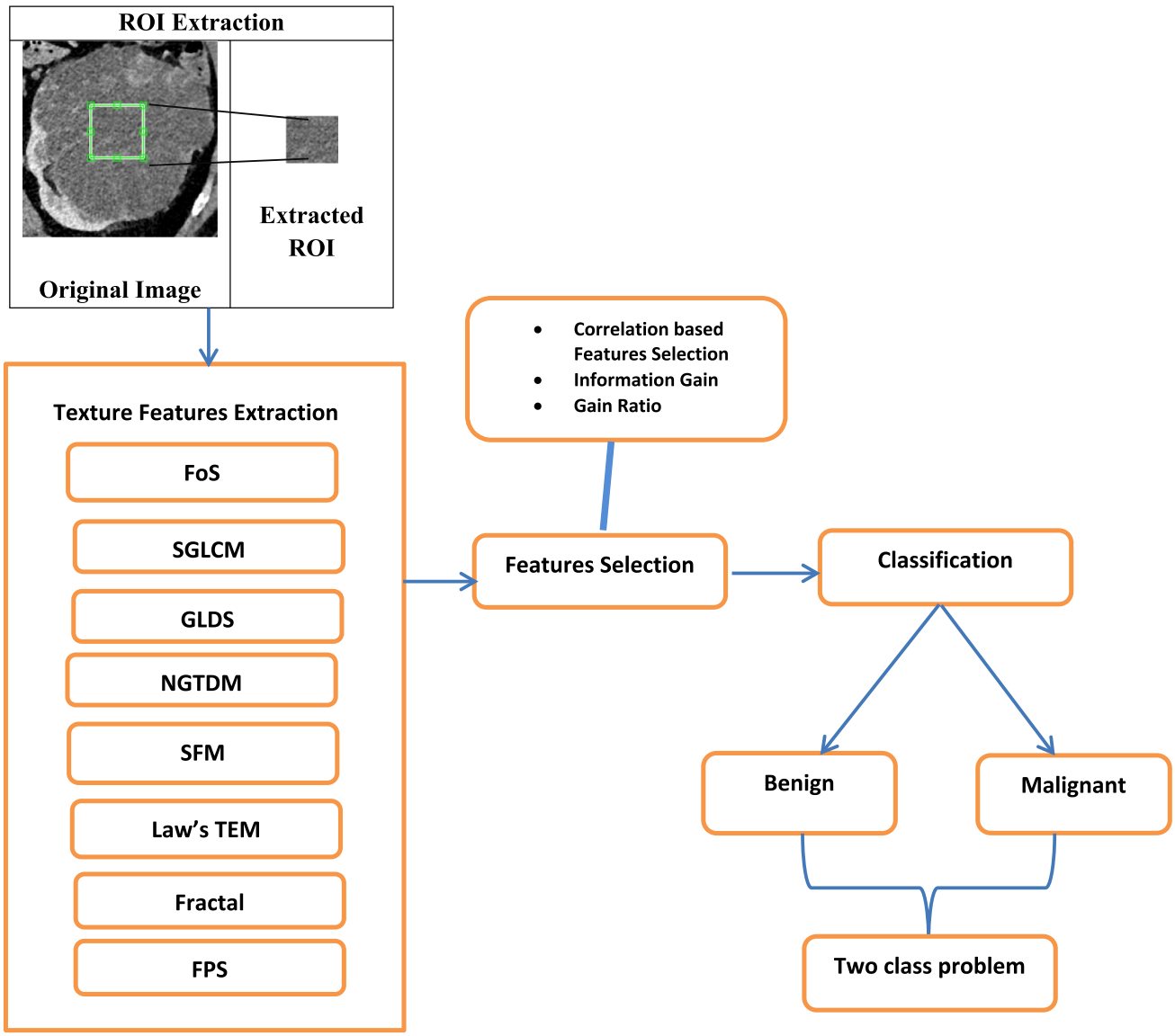

Fig. 2 Illustration of the proposed classification system's phases

small lesions, and it includes adequate space to capture diverse features from different texture models.

Feature Extraction

The image's texture elements aid in identifying essential details about the image's content. The gray level intensities of renal tissues differ between renal lesions that are benign and malignant. Texture analysis has been widely employed for the purpose of obtaining information regarding the texture of various medical images for classification [26, 27]. Because image texture offers data related to the spatial distribution of the values of the pixels as well as brightness level variations, these texture models can present picture content in a variety of ways, making it beneficial for diagnosing renal lesions. To extract information from a specific area of interest, in the course of this research, eight distinct varieties of well-known texture models were utilized that namely SGLCM [28], GLDS [29], FoS [30], Law's texture energy measures (TEM) [31], Fourier power spectrum (FPS) [32], statistical feature matrix (SFM) [33], fractal features [34], and neighborhood gray tone difference matrix (NGTDM) [34]. Overall, 44 features were recorded, all of which were normalized before performing classification. Distinct models for texture analysis along with corresponding features are listed in Table 1.

SGLCM SGLCM features are the commonly employed features for quantifying image textures. In matrix generated by SGLCM, the number of quantized gray levels, Ng, is equivalent to columns and rows in an image. To determine the frequency with which two gray levels exist in a clear correlation, co-occurrence matrices are utilized. This spatial association is a measured using the distance (D) and angular relationship θ among the nearby pixel values [28, 29]. Consider Ng signifies the quantity of quantized grayscale levels, and for a fixed value of D and θ, the size of the GLCM matrix is Ng × Ng and is defined as:

$P(x, y, D, \theta)$ = Count of pixels placed at pixel distance D and direction θ, such that the gray value of first pixel in the pixel set is x and that of second pixel is y. The most common choices for θ are 0°, 45°, 90°, and 135°, and their corresponding equations are mentioned as Eqs. 1 to 4 respectively. Thus,

$$P\left(x, y, D, 0^{\circ}\right) = \#\{((p, q), (r, s))\varepsilon\, I \text{ such that } p - r = 0, |q - s| = D, I(p, q) = x, l(r, s) = y\} \tag{1}$$

$$P\left(x, y, D, 45^{\circ}\right) = \#\{((p, q), (r, s))\varepsilon\, I \text{ such that } p - r = D, \quad q - s = -D, I(p, q) = x, l(r, s) = y\} \tag{2}$$

$$P\left(x, y, D, 90^{\circ}\right) = \#\{((p, q), (r, s))\varepsilon\, I \text{ such that } |p - r| = D, q - s = 0, I(p, q) = x, l(r, s) = y\} \tag{3}$$

$$P\left(x, y, D, 135^{\circ}\right) = \#\{((p, q), (r, s))\ \varepsilon\, I \text{ such that } p - r = D, q - s = D, I(p, q) = x, l(r, s) = y\} \tag{4}$$

Table 1 Shows various texture models and their obtained features

Model of Texture	Feature names that were extracted
SGLCM ($F_{1–13}$)	Second angular moment, inverse difference moment, correlation, contrast, entropy, squared sums, the sum of the averages, the sum of the variances, and the sum of the difference variance, difference entropy, and sum entropy correlation information measure 1, correlation information measure 2
GLDS ($F_{14–18}$)	Mean, energy, homogeneity, entropy, contrast
FoS ($F_{19–23}$)	Mean, median, variance, skewness, kurtosis
Law's TEM ($F_{24–29}$)	LL, LS, EE, ES, LE, SS
FPS ($F_{30–31}$)	Angular sum, radial sum
SFM ($F_{32–35}$)	Coarseness, periodicity, contrast, roughness
Fractal ($F_{36–39}$)	Hurst coefficient at different resolutions (H1, H2, H3, H4)
NGTDM ($F_{40–44}$)	Coarseness, contrast, busyness, complexity, and strength are some of the characteristics that can be found in a design

Here, # signifies how many components are in the set. Using a single co-occurrence matrix with a selected distance $D = 1$, the present research uses a set of 13 characteristics. For directions $\theta = 0°$, 45°, 90°, and 135°, four different values are computed for every attribute, and the mean value is used for more detailed analysis. The symbolizations utilized to estimate these attributes are mentioned as below:

Symbols	*Explanation*
$P(x,y)$	The likelihood that x and y gray levels combination occur in sequence
$p(x,y)$	$(x,y)th$ record in a normalized matrix $= \frac{P(x,y)}{R}$ where R is normalizing constant
$P_u(x)$	$= \sum_{x=1}^{Ng} P(x,y)$, attained by adding the rows of $p(x,y)$
N_g	Number of different grayscales that are present in the quantized image
$p_v(x)$	$= \sum_{y=1}^{Ng} p(x,y)$, obtained by adding the columns of $p(x,y)$
${}^{p}(u+v)^{(k)}$	$= \sum_{x=1}^{Ng}\sum_{y=1}^{Ng} p(x,y)$, $x+y=k$, $k=2, 3\ldots, 2N_g$
${}^{p}(u-v)^{(k)}$	$= \sum_{x=1}^{Ng}\sum_{y=1}^{Ng} p(x,y)$, $\lvert x-y\rvert = k$, $k = 0, 1, 2, 3\ldots\ldots\ldots, N_g - 1$

Here, μ_u, μ_v, σ_u, and σ_v represent the means and standard deviations of p_u and p_v.

The mathematical details of the calculated Haralick features are as follows [28]:

(a) Angular second moment: Texture homogeneity is measured by it. It is calculated by adding up all the squared values of SGLCM, as given in Eq. 5:

$$ASM = \sum_{x=1}^{N_g}\sum_{y=1}^{N_g}\left[p(x,y)\right]^2 \tag{5}$$

Due to the small number of gray levels in a homogeneous image, SGLCM has larger values of $P(x, y)$, resulting in large ASM values.

(b) Contrast: Over the whole image, it quantifies the intensity contrast differences between a certain pixel and its neighbors, calculated using Eq. 6:

$$\text{Contrast} = \sum_{n=0}^{N_g-1} n^2 \left(\sum_{x=1}^{N_g}\sum_{y=1}^{N_g} p(x,y)\cdot\delta(|x-y|-n) \right) \tag{6}$$

The objects in the image are clearly separated by high contrast levels.

(c) Correlation: It measures the correlation of a pixel with its neighboring pixels in the entire image, and computed using Eq. 7:

$$\text{Correlation} = \frac{\sum_{x=1}^{N_g}\sum_{y=1}^{N_g}\left(x-\mu_x\right)\left(y-\mu_y\right)p(x,y)}{\sigma_x\sigma_y} \tag{7}$$

(d) Inverse difference moment: This indicates that the elements' distributions in the GLCM matrix are fairly close to the major diagonal, as determined by Eq. 8. If the high values of the matrix are close to the major diagonal, then this property becomes quite valuable. Image homogeneity is indicated by an elevated value:

$$\text{Inverse Difference Moment} = \sum_{x=1}^{N_g}\sum_{y=1}^{N_g}\frac{p(x,y)}{1+(x-y)^2} \tag{8}$$

(e) Variance: It calculates the deviation of pixel values from the average value of SGLCM, computed using Eq. 9:

$$\text{Variance} = \sum_{x=1}^{N_g}(x-\mu)^2\, p_x(x) \tag{9}$$

(f) Sum average: The brightness of the image is measured using this feature, and its values are high for an image with high brightness. It can be calculated using Eq. 10:

$$\text{Sum Average} = \sum_{x=2}^{2N_g} x\, p_{u+v}(x) \tag{10}$$

(g) Entropy: This is used to measure the degree of randomness in a given image, as mentioned in Eq. 11. Its value is high for a homogenous image and low for the heterogeneous image:

$$\text{Entropy} = -\sum_{x=1}^{N_g}\sum_{y=1}^{N_g} p(x,y)\log p(x,y) \tag{11}$$

(h) Sum entropy: The utilization of this feature helps to measure the entropy of vector P_{u+v}, and it has low value as compared to entropy, estimated using Eq. 12:

$$\text{Sum Entropy} = -\sum_{x=2}^{2N_g} p_{u+v}(x)\log p_{u+v}(x) \tag{12}$$

(i) Sum variance: This has a low value for a highly contrasted image since it measures the dispersion of the P_{x+y} elements concerning the sum entropy. This can be calculated using expression given in Eq. 13:

$$\text{Sum Variance} = \sum_{x=2}^{2N_g} (x - \text{Sum Entropy})^2\, p_{u+v}(x) \tag{13}$$

(j) Difference entropy: This feature measures entropy vector P_{u-v} whose weight rises as one move away from the diagonal. Its value is computed using Eq. 14:

$$\text{Difference Entropy} = -\sum_{x=0}^{N_g-1} p_{|u-v|}(x)\log p_{|u-v|}(x) \tag{14}$$

(k) Difference variance: This is used to identify the deviation of vector elements of P_{x-y} from difference entropy by utilizing Eq. 15:

$$\text{Difference Variance} = \sum_{x=0}^{N_g-1} (x - \text{Difference Entropy})^2\, p_{|u-v|}(x) \tag{15}$$

(l) Information measure of correlation (IMC) 1: The value of this feature is low for a homogenous image, and its values are computed using expression mentioned in Eq. 16:

$$\text{IMC1} = \frac{H_{XY} - H_{XY1}}{\max\left(H_X, H_Y\right)} \tag{16}$$

Here, HX and HY represent entropies of p_x and p_y:

$$\text{HXY} = -\sum_{x=1}^{N_g}\sum_{y=1}^{N_g} p(x,y)\ \log\left(p(x,y)\right)$$

$$\text{HXY1} = -\sum_{x=1}^{N_g}\sum_{y=1}^{N_g} p(x,y)\ \log\left(p_u(x)p_v(y)\right)$$

$$\text{HXY2} = -\sum_{x=1}^{N_g}\sum_{y=1}^{N_g} p_u(x)p_v(y)\ \log\left(p_u(x)p_v(y).\right)$$

(xiii) IMC 2: The information measure for correlation 2 declines for images with higher energy levels, calculated using Eq. 17:

$$\text{IMC2} = \sqrt{1 - \exp\left[-2\left(H_{XY2} - H_{XY}\right)\right]} \tag{17}$$

GLDS The GLDS are estimated using the absolute differences among the pairs of gray to extract the five different texture measures, which include homogeneity, contrast, mean, energy, and entropy [29, 30]. Consider that image pixel values are random variables that can take discrete values $i = 0, 1, \ldots. N_g - 1$, where N_g is the number of gray levels. Let the pairs of gray levels be separated by a distance $\delta = (\Delta x, \Delta y)$. For a given displacement $\delta = (\Delta x, \Delta y)$, the difference image $f_\delta(x, y)$ is defined using Eq. 18:

$$f_\delta(x,y) = |f(x,y) - f(x+\Delta x, y+\Delta x)| \tag{18}$$

A gray level histogram is represented by the symbol $p\delta$, which stands for the probability density of (x, y) for m gray levels. A number of different texture features that were retrieved from p_δ are

(a) Homogeneity: This denotes the similarity in gray level intensities, calculated using Eq. 19:

$$Homogeneity = \frac{\sum p\delta(i)}{1+i} \tag{19}$$

(b) Contrast: The value of this feature is estimated using Eq. 20, as it identifies gray level intensity difference among neighboring pixels:

$$Contrast = \sum_{i}^{2} p_\delta(i) \tag{20}$$

(c) Mean: It is a measure of the mean intensity of grayscale levels in a specific region of a picture, defined by using Eq. 21:

$$Mean = \frac{1\sum_i p(i)}{m^\delta} \tag{21}$$

(d) Energy: This attribute stands for the magnitude of grayscale values; its value is estimated using Eq. 22:

$$Energy = \sum p_\delta(i)^2 \tag{22}$$

(e) Entropy: The dispersion of grayscale intensities in a specific region is quantified by the entropy and is calculated using Eq. 23:

$$Entropy = -\sum p_\delta(i)\log p_\delta(i) \tag{23}$$

FoS The first-order statistical features rely on the individual pixel values, and they do not consider any association with

their nearby pixels. The features are computed by employing the normalized histogram of a considered image. Considering into account the fact that the pixel values of the image are random variables that might take on discontinuous variations $i = 0, 1,\ldots.N_g - 1$, where N_g is the number of gray levels; the normalized histogram is defined using Eq. 24:

$$p(i) = \frac{\text{Number of pixels with intensity } i}{\text{Total number of pixels}}, \quad i = 0, 1, \ldots, N_g - 1 \tag{24}$$

The four first-order statistical features computed using image histogram includes mean, median, skewness, and kurtosis [30, 31].

(a) Mean: The average intensity level of an image can be determined by using the mean, calculated using Eq. 25:

$$\mu = \sum_{i=0}^{N_g-1} i \cdot p(i) \tag{25}$$

(b) Median: It is estimated by initially arranging the pixel values into ascending order and then exchanging the considered pixel value with the middle (median) value.

(c) Skewness: A histogram's skewness indicates how asymmetrical it is, and its value is calculated using Eq. 26:

$$\text{Skewness} = \mu_3 = \sigma^{-3} \sum_{i=0}^{N_g-1} (i - \mu)^3 \, p(i) \tag{26}$$

(d) Kurtosis: The dispersion of grayscale values around the mean is known as kurtosis, whose value is estimated using Eq. 27:

$$\text{Kurtosis} = \mu_4 = \sigma^{-4} \sum_{i=0}^{N_g-1} (i - \mu)^4 \, p(i) - 3 \tag{27}$$

For a Gaussian-shaped histogram, an integer constant of 3 is used in Eq. 27 to normalize $\mu 4$ to 0.

TEM This approach quantifies image texture by utilizing convolution with specific filters relying upon Laws masks, on a predetermined scale in order to quantify changes in surface texture [31, 32]. This involves the use of three initial vectors, mentioned in Eqs. 28, 29, and 30 that resemble the center-weighted local averaging (L5), edge detection (E5), and spot detection (S5):

$$L5 = [1 \quad 4 \quad 6 \quad 4 \quad 1] \tag{28}$$

$$E5 = [-1 \quad -2 \quad 0 \quad 2 \quad 1] \tag{29}$$

$$S5 = [-1 \quad 0 \quad 2 \quad 0 \quad -1] \tag{30}$$

Using the 1-D convolution kernels mentioned in Eqs. 28–30, by merging a horizontal 1-D kernel with a vertical 1-D kernel, nine separate 2-D kernels are produced, and different combination of obtained 3 × 3 kernel are mentioned as follows:

L5L5	*E5L5*	*S5L5*
L5E5	*E5E5*	*S5E5*
L5S5	*E5S5*	*S5S5*

Subsequently, after performing the convolution with the specific kernel, the texture energy measures are calculated by adding the pixel values in the local neighborhood. The digital image with a dimension of $N \times M$ is convoluted through each of these nine kernels, resulting in the generation of a nine grayscale images. The windowing process is carried out in a local neighborhood by utilizing Eq. 31:

$$NEW\ (\mathrm{x}, \mathrm{y})\Big) = \sum_{i=-7}^{7} \sum_{j=-7}^{7} |\mathrm{OLD}(x + i, y + i)| \tag{31}$$

Since L5L5 is the only convolutional kernel that has a zero mean, L5L5T is utilized as a normalizing image, and after performing normalization, the L5L5T image is deleted. The rotation-invariant features may be created by combining features that have similarities in both the horizontal and vertical axes with one another. This investigation makes use of the six TEM characteristics which includes LL, EE, and SS texture energies from L5 × L5, E5 × E5, and S5 × S5 kernels, respectively, and LE, ES, and LS represent mean texture energies from L5 × E5 and E5 × L5 kernels, E5 × S5 and S5 × E5 kernels, and L5 × S5 and S5 × L5 kernels, respectively. The image convolution is performed with Law's masks and parameters to quantify the ROI texture are acquired by calculating statistics from filtered image.

FPS This model provides the details about the texture directionality, coarseness, and texture contrast of the image. The spatial techniques are unable to describe repetitive global texture patterns, so discrete Fourier transform (DFT) is employed for quantifying texture [32, 33] which calculates the angular sum and radial sum of the DFT for texture analysis. The power spectrum in the frequency domain is used for calculating the FPS characteristics, represented by Eq. 32:

$$|F(u, v)|^2 = F(u, v) \cdot F^*(u, v) \tag{32}$$

Here, $F(u,v)$ represents the image Fourier transform and $F^*(u,v)$ denotes the complex conjugate of image Fourier transform. Further, in a polar coordinate system, the spectral features are expressed to form a function $S(r,\theta)$, where for every direction θ, $S(r,\theta)$ might be articulated as $S_\theta(r)$ and likewise, for every frequency r, $S(r,\theta)$ might be articulated as

$S_r(\theta)$. The values of S_θ and S_r are calculated using Eqs. 33 and 34, respectively:

$$S_\theta = \sum_{\theta=0}^{\pi} S_\theta(r) \tag{33}$$

$$S_r = \sum_{r=0}^{R_0} S_r(\theta) \tag{34}$$

where R_0 is the circle radius centered at the origin. Here, two distinct characteristic feature values S_r and S_θ are computed, representing the degree of texture orientation.

SFM These texture features highlight about the statistical characteristics of the pixels at different distances inside a specific image, to calculate features which include coarseness, periodicity, contrast, and roughness [30–32]. This model utilizes a matrix. M_{sfm} of dimensions $(L_r + 1) \times (2L_c + 1)$. The component located at position (i,j) in the matrix of elements signifies d, where $d=(j-Lc,i)$ in which an inter-sample spacing distance vector for $i = 0, 1,\ldots\ldots L_r$ and $j = 0, 1,\ldots\ldots 2L_c$. Here, L_r and L_c are the constants to decide the largest inter-spacing distance. Further, two such matrices are constructed namely contrast matrix (M_{con}) and the dissimilarity matrix (M_{dss}). A statistical feature, denoted as δ, provides a feature of an image I characterized by the inter-sample dispersion distance, represented by a vector $\delta = (\Delta x, \Delta y)$.

δcontrast and δdissimilarity are defined by employing Eqs. 35 and 36, respectively:

$$\delta\text{contrast} : CON(\delta) \equiv E\left\{ \left[I(x,y) - I(x+\Delta x, y+\Delta y)\right]^2 \right\} \tag{35}$$

$$\delta\text{dissimilarity} : DSS(\delta) \equiv E\left\{ \left|I(x,y) - I(x+\Delta x, y+\Delta y)\right| \right\} \tag{36}$$

The expectation operation is denoted by the letter E (.). In the image, the four texture features that were retrieved are as follows:

(a) Coarseness: It quantifies the prevalence of comparable grayscales in a specific region and serves as an indicator for primitive size; its value is computed by employing Eq. 37.

$$Coarseness = {}^{c}/_{m_{crs}} \tag{37}$$

Here, c represents the normalization factor, which is assumed to be equal to 100, and m_{crs} is the average of all the elements of the matrix (M_{dss}), which is computed by applying the expression given in Eq. 38:

$$m_{crs} = \sum_{(i,\, j \in N_r)} DSS(i,\, j)/n \tag{38}$$

Here, the set of displacement vectors is denoted by the symbol N_r, while the total number of elements in the matrix is denoted by the symbol n.

(a) Contrast: This feature measures the difference in gray level intensities for a specified neighborhood, estimated using Eq. 39:

$$Contrast = \left[\sum\nolimits_{(i,\, j \in N_r)}^{1/2} CON(i,\, j)/4 \right] \tag{39}$$

(b) Periodicity: This characteristic is a measurement of the frequency with which certain patterns appear in a specific picture, computed by employing Eq. 40:

$$Periodicity = \frac{\overline{M}_{dss} - M_{dss}}{\overline{M}_{dss}} \tag{40}$$

where $\overline{M}_{dss}$ *is* mean of all elements of M_{dss}.

(c) Roughness: It calculates the dissimilarity between the neighboring pixels, given by Eq. 41:

$$Roughness = \left(D_f^h + D_f^v\right)/2 \tag{41}$$

When the displacement vectors (0, Δy) and (Δx, 0) are taken into consideration, the fractal dimensions D_f^h *and* D_f^v are determined in the horizontal and vertical directions, respectively. Further, the D_f can be computed using the Eq. 42:

$$D_f = 3 - H \tag{42}$$

Here, H is estimated using a dissimilarity matrix as the (i, j + L_c)th element of dissimilarity matrix is $E\{|\Delta I|\}$ with $\delta = (j, i)$. The relationship between H and $E\{|\Delta I|\}$ is represented by using Eq. 43:

$$E\{|\Delta I|\} = C\|\delta\|^H \tag{43}$$

Table 2 Calculation of distinct parameters required for NGTDM feature estimation

i	ni	pi	si
1	6	0.375	13.35
2	2	0.125	2.00
3	4	0.25	2.63
4	0	0.00	0.00
5	4	0.25	10.075

Thus, it is necessary to calculate H in a separate manner for both the horizontal and vertical directions, using the dissimilarity matrix as described in Eq. 43 by taking into account the displacements of (Δx, 0) and (0, Δy) accordingly. D_f^h and D_f^v are computed by utilizing Eq. 42, and lastly, the value of roughness can be obtained by employing Eq. 41. This research study calculated the SFM features by taking $L_r = L_c = 4$.

Fractal Features Fractals are utilized to define the object appearance and shape which possess the properties of scale-invariance and self-similarity. Fractal dimensions are employed to estimate the degree of surface roughness or boundary irregularities. In this model, Hurst coefficients, denoted by ($H(k)$), are calculated at distinct image resolutions, to describe texture surface. The large values of parameter $H(k)$ represent smooth texture surface, and the small value specifies a rough texture [33, 34]. This research work estimated the value of Hurst coefficient at two different resolutions: an original image with dimensions of $N \times M$ is referred to as $H1$, while an original image that has been scaled by a factor of 1/2 and is of dimensions $N/2 \times M/2$ is referred to as $H2$. The value of Hurst coefficient H, for an image I under consideration, can be computed using the relationship mentioned in Eq. 44:

$$E \mid \Delta l \mid = k(\Delta r)^H \tag{44}$$

Here, E(.) denotes the expectation operator, $\Delta I \equiv I(x_2, y_2) - I(x_1, y_1)$,

$\Delta r \equiv \left\| (x_2, y_2) - (x_1, y_1) \right\|$ is a spatial distance and $k = E(|\Delta l|)_{\Delta r=1}$

Hurst coefficients at two different resolutions are estimated using Eq. 45 to represent the image texture:

$$logE \mid \Delta l \mid = logk + Hlog(\Delta r) \tag{45}$$

NGTDM A neighboring gray tone difference matrix quantifies the difference between a gray value and the average gray value of its neighbors within distance δ [33, 34]. The sum of absolute differences for gray level i is stored in the matrix. Let Xgl be a set of segmented voxels and xgl(jx,jy,jz)∈Xgl be the gray level of a voxel at position (jx,jy,jz), then the average gray level of the neighborhood is calculated using Eq. 46:

$$\overline{A_i} = \overline{A_i}(j_x, j_y, j_z) = \frac{1}{W} \sum_{k_x=-\delta}^{\delta} \sum_{k_y=-\delta}^{\delta} \sum_{k_z=-\delta}^{\delta} x_{gl}(j_x + k_x, j_y + k_y, j_z + k_z) \tag{46}$$

where $(k_x, k_y, k_z) \neq (0,0,0)$ and $x_{gl}(j_x + k_x, j_y + k_y, j_z + k_z) \in x_{gl}$, Here, W is the number of voxels in the neighborhood that are also in Xgl. As a two-dimensional example, let the following matrix I, given as Eq. 47, represent a 4×4 image, having five discrete gray levels, but no voxels with gray level 4:

$$I = \begin{bmatrix} 1 & 2 & 5 & 2 \\ 3 & 5 & 1 & 3 \\ 1 & 3 & 5 & 5 \\ 3 & 1 & 1 & 1 \end{bmatrix} \tag{47}$$

The following NGTDM can be obtained using Table 2

6 pixels have gray level 1, therefore:

$$s1 = | 1 - 10/3 | + | 1 - 30/8 | + | 1 - 15/5 | + | 1 - 13/5 | + | 1 - 15/5 | + | 1 - 11/3 | = 13.35$$

For gray level 2, there are 2 pixels, therefore:

$$s2 = | 2 - 15/5 | + | 2 - 9/3 | = 2$$

Similar for gray values 3 and 5:

$$s3 = | 3 - 12/5 | + | 3 - 18/5 | + | 3 - 20/8 | + | 3 - 5/3 | = 3.03$$
$$s5 = | 5 - 14/5 | + | 5 - 18/5 | + | 5 - 20/8 | + | 5 - 11/5 | = 1\ 0.075$$

Consider Ni be the number of voxels in Xgl with gray level i, Nv,p be the total number of voxels in Xgl and equal to ∑ni (i.e., the number of voxels with a valid region; at least 1 neighbor). Nv,p ≤ Np, where Np is the total number of voxels in the ROI. Here, pi is the gray level probability and equal to ni/Nv:

$$\text{si} = \begin{cases} n_i \left| i - \overline{A_i} \right| & for\ n_i \neq 0 \\ 0 & for\ n_i = 0 \end{cases} \tag{48}$$

where si represents the sum of absolute differences for gray level i as mentioned in Eq. 48; Ng is the number of

Table 3 Discriminating features ranked by considered feature selection techniques

S. no.	Features	Ranks		
		CFS	Information Gain	Gain ratio
1	Median	1	1	2
2	Sum entropy	2	2	1
3	Strength	3	8	8
4	Hurst coefficient 3	4	4	7
5	Entropy	5	5	5
6	Angular sum	6	6	6
7	Radial sum	7	7	9
8	Kurtosis	8	3	10
9	Skewness	9	9	4
10	Coarseness	10	10	3

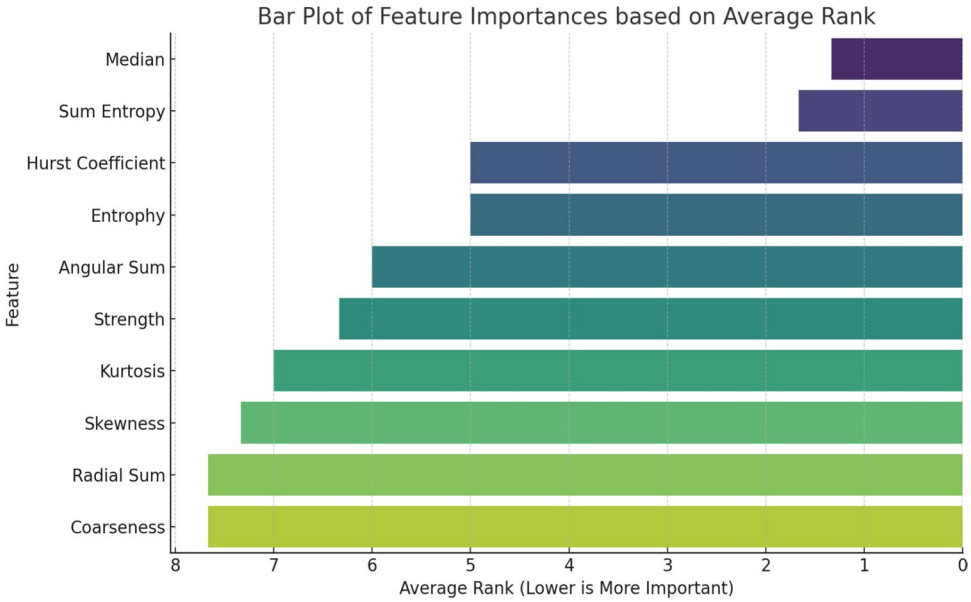

Fig. 3 Bar plot of feature based on average rank

discrete gray levels, and Ng,p is the number of gray levels where pi ≠ 0,

Feature Selection

Some of the 44 characteristics extracted from different texture models as given in Table 1 may provide less information for classification; thus, discriminative features must be chosen to improve classifier accuracy. The basic goal of feature selection methods is to find all favorable feature combinations for a given dataset so that a subset of features with high prognostic power may be identified. This can be performed by selecting the most important traits with the highest discriminating power and excluding the irrelevant or redundant ones [35]. In this study, three different techniques for selecting features have been utilized.

Correlation-Based Feature Selection (CFS) In this, features are chosen using a CFS approach [35], which is a convenient filtering technique that employs heuristic functions based on correlation to identify important picture features. The prominence of a subset of features is calculated using the CFS technique, which considers two factors: each image feature's predictive ability and the extent of redundancy associated with it. All subsets prefer feature subsets that have a better correlation with the class and lower inter-correlation. The CFS evaluation function for obtaining a feature subset is specified in Eq. 49:

$$Merit_e = \frac{M\overline{cij}}{\sqrt{M + M(M-1)\overline{cjj}}} \tag{49}$$

$Merit_e$ signifies merit of a feature subset e that comprises M features. The average feature–feature correlation is (cjj), while the average feature–feature inter-correlation is (cjj) (cij). The best initial search approach is used to study the search space, which involves frontward selection with exit condition of five consecutive completely extended non-improving subsets. Begin your exploratory search with a 0.0 merit empty set. Each subset's worth is estimated using a heuristic function, and the subset possessing elevated merit is informed. In this experimental study, the value corresponding to the best subset was reported as 0.482, and the feature subsets selected using the CFS approach are median, sum entropy, strength, hurst coefficient 3, entropy, angular sum, radial sum, kurtosis, skewness, and coarseness, and the ranks of the same are given in Table 3.

IG Information gain quantifies the amount of information or decrease in entropy (i.e., uncertainty of a random feature) for the projected class when considering the distribution of distinct features inside the associated class [36]. Considering set S_X having ith variable of sample x_i, a value of IG is measured by employing Eq. (50):

$$IG(S_X, x_i) = H(S_X) - \sum_{v \in \text{values}(x_i)} \frac{\left|S_{X_{x_i=v}}\right|}{|S_X|} H\left(S_{X_{x_i=v}}\right) \tag{50}$$

Here, $H(S)$ signifies entropy measured by utilizing Eq. 51 and $\frac{|S_{x_i=v}|}{|S_X|}$ is the proportion of the ith variable in the sample:

$$H(S) = -p_+(S)\log_2 p_+(S) - p_-(S)\log_2 p_-(S) \tag{51}$$

$p_+(S)$ indicates the likelihood of a sample in set S being either negative or positive.

In this case, we only keep the 10 features that have a score above 0.312 and ignore those with lower scores. Therefore, the discriminant features as per information gain score are median, sum entropy, strength, hurst coefficient 3, entropy, angular sum, radial sum, kurtosis, skewness, and coarseness, and the ranks of the same are given in Table 3.

Gain Ratio (GR) An improvement of information gain, GR eliminates bias caused by feature selection at every node. The set "Sx" is being considered, along with examples that represent i different classes [37]. To categorize a sample, the following data is required to get information using Eq. 52:

$$I(S_X, x_i) = -\sum p_i \log_2(p_i) \tag{52}$$

Here, p_i indicates the probability of a randomly selected subset. Now, the entropy is calculated by utilizing Eq. 53 for sample A, taking into account v different values:

$$H(S) = \sum I(S_X)\frac{S_i}{S} \tag{53}$$

Additionally, the gained information can be calculated using Eq. 54, and the obtained information can further be normalized using Eq. 55:

$$Gain(S) = I(S_X) - H(S) \tag{54}$$

$$SplitInfo_S(S_X) = -\sum \left(|S_i|/|S_X|\right)\log_2\left(|S_i|/|S_X|\right) \tag{55}$$

Lastly, Eq. 56 can be employed to estimate the final value of GR:

$$GR(S) = \frac{Gain(S)}{SplitInfo_A(S_X)} \tag{56}$$

Likewise IG, in this case as well, only those 10 features with a high score of 0.31 are retained, while those with scores lower than that are eliminated. Therefore, the discriminant features as per information gain score are median, sum entropy, strength, hurst coefficient 3, entropy, angular sum, radial sum, kurtosis, skewness, and coarseness, and the ranks of the same are given in Table 3.

To mitigate these limitations, we compared results across the three techniques and selected features that consistently ranked highly across multiple methods. Additionally, we applied cross-validation and feature importance analysis from ensemble models (e.g., random forest) to validate the selected features and reduce overfitting risk. This multi-pronged approach aimed to ensure a balanced selection that enhances model performance while maintaining generalizability. The feature importance analysis, based on average ranking across as shown in Fig. 3, correlation-based feature selection (CFS), information gaiN, and gain ratio, highlights that median and sum entropy are the most influential features in distinguishing benign from malignant renal lesions. These first-order statistical descriptors effectively capture the central intensity tendencies and the degree of randomness within lesion textures, which are critical for accurate classification. Features like hurst coefficient and entropy also rank highly, indicating the relevance of fractal characteristics and textural disorder in renal tissue. Mid-ranking features such as angular sum and strength reflect spatial frequency and structure, contributing moderately to the model's discriminative power. Conversely, features like skewness, radial sum, and coarseness are relatively less important, though they may still provide complementary value when combined with higher-ranked features. Overall, the ranking validates the effectiveness of integrating a diverse set of statistically robust features to enhance classification performance and support clinical decision-making

While the combination of correlation-based feature selection (CFS), information gain (IG), and gain ratio (GR) enhances the robustness of feature selection, each method has its own limitations. CFS favors features that are highly correlated with the target class and minimally correlated with each other, effectively reducing multicollinearity; however, it may discard relevant features that exhibit interdependence, potentially leading to information loss. Information gain, which ranks features based on their ability to reduce class entropy, tends to favor attributes with a large number of distinct values. This bias can result in overfitting, particularly when applied to categorical variables with high cardinality. Gain ratio, an improvement over IG, addresses this bias by normalizing with split information but may undervalue informative features that display low variability across samples. These limitations were considered when combining the outputs of all three techniques to select the most discriminative features, ensuring a balanced trade-off between relevance and redundancy.

Classification

A prognostic model for renal lesions classification is created using a different classifier. This machine learning methodology integrates a number of self-contained prediction algorithms with decision points, such as classification and regression trees. Machine learning classifiers may provide the benefits of high generalization and adaptability to big

medical datasets, as well as speedy training and prediction. These machine learning algorithms are fed feature data together with the known labels based on the pathological findings or radiographical appearance of each lesion. For the relevant classification problem, an ideal tuning of parameters was done through rigorous inquiry, and five-fold cross-validation was used to offer classification results. Once the features have been retrieved, they are given scores and passed on to classifiers so that they can be classified as either normal or abnormal cases. An approach to decision-making that involves identifying features from images and assigning labels to those features is called classification. Classification without supervision and supervised classification are the two primary categories that it falls under during this process. A classification that is unsupervised is one that relies on the results without any pre-defined (human) input used. On the other hand, supervised classification depends on input that has been pre-defined by experts. Here is a brief description of the classifiers that were utilized in this research:

1. Naive Bayes (NB): When it comes to the area of healthcare diagnostics through the use of artificial intelligence, naive Bayes comes under the category of basic probabilistic classification. This classification method operates based on the Bayes' theorem and makes a strong assumption that the features are independent of one another. Since NB assigns a label to instances of a problem in vectors of features and labels derived from a finite set, it is necessary for a learning problem to have several linear parameters for it to function properly. In addition to this, it necessitates a limited amount of training data in order to estimate the key parameters for classification [38]. In addition to this, it is included in the category of conditional probabilities, such as a vector $x=(x_1,\ldots\ldots,x_n)$ with n features is assigned probabilities $t(CC_k|x_1\ldots\ldots,x_n)$ for "k" possible outcomes or CC_k classes. Therefore, following Bayes' theorem, the formulation of conditional probability is as follows (Eq. 57):

$$t\left(CC_k|x\right)=\frac{t\left(CC_k\right)t\left(x|CC_k\right)}{t(x)} \tag{57}$$

It can be seen from Eq. 57 that denominator is not dependent on CC and is thus effectively constant. Additionally, the combined likelihood model's numerator is the same as its denominator $t(CC_k,x_1,\ldots\ldots,x_n)$ which can be written as given in Eq. 58 using the chain rule for applications of conditional probability [34].

$$\begin{aligned} t\left(CC_k,x_1,\ldots\ldots\ldots,x_n\right) &= t\left(x_1,\ldots\ldots\ldots,x_n,CC_k\right)\\ &= t\left(x_1|x_2,\ldots\ldots\ldots,x_n,CC_k\right)t\left(x_2,\ldots\ldots\ldots,x_n,CC_k\right)\\ &=\cdots\\ &=\cdots\\ &= t\left(x_1|x_2,\ldots\ldots\ldots,x_n,CC_k\right)t\left(x_2|x_3,\ldots\ldots\ldots,x_n,CC_k\right)\ldots\ldots\ldots t\left(x_{n-1}|x_n,CC_k\right)t\left(x_n|CC_k\right)t\left(CC_k\right) \end{aligned} \tag{58}$$

Now, take into consideration the fact that each feature x_i is assumed to be conditionally independent of the other feature x_j for $j\neq i$, in the category. CC_k means that $t(x_i|x_{i+1},\ldots\ldots\ldots,x_n,CC_k)=t(x_i|CC_k)$. Thus, the joint model is given in Eq. 59:

$$\begin{aligned} t\left(CC_k|x_1,\ldots\ldots\ldots,x_n\right) &= at\left(CC_k,x_1,\ldots\ldots x_n\right)\\ &= t\left(CC_k\right)t\left(x_1|CC_k\right)t\left(x_2|CC_k\right)t\left(x_3|CC_k\right)\\ &= t\left(CC_k\right)\prod_{i=1}^{n}t\left(x_i|CC_k\right), \end{aligned} \tag{59}$$

Here, α indicates proportionality. Also, under the independent assumptions, conditional probability distribution over class variable C is given in Eq. (60):

$$t\left(CC_k|x_1,\ldots\ldots\ldots,x_n\right)=\frac{1}{Z}t\left(CC_k\right)\prod_{i=1}^{n}t\left(x_i|CC_k\right) \tag{60}$$

Here, $Z=t(x)=\sum_k t\left(C_k\right)t\left(x|C_k\right)$ is the scaling factor dependent on $x_1,\ldots\ldots,x_n$.

Finally, the Bayes' classifier that marks a label $\hat{y}=C_k$ for some $k\in\{1,\ldots\ldots,K\}$ is calculated using Eq. 61:

$$\hat{y}=\operatorname{argmax}t\left(C_k\right)\prod_{i=1}^{n}t\left(x_i|C_k\right) \tag{61}$$

2. Support vector machine (SVM): Support vector machines make use of an ideal linear hyperplane, as a means of partitioning a dataset into a feature space. Finding the optimal hyperplane involves finding the maximum margin between two sets. Therefore, the support vectors, which serve as training patterns of borders, determine the final hyperplane. First, it performs a nonlinear mapping of the input vector into a high-dimensional feature space that is concealed from both the output and the input. This is one of its primary processes. The second step involves the selection of the best hyperplane design for feature separation [39]. Therefore, the following are the steps to create a support vector machine:

Step 1: Define hyperplane acting as decision surface using Eq. 62:

$$\sum_{i=1}^{N} \alpha_i d_i K\left(x, x_i\right) = 0 \tag{62}$$

where $K(x, x_i) = \varphi^T(x)\varphi(x_i)$ signifies the products of the internal vector in the feature space using input pattern x_i and vector x for the i^{th} feature with dimension m_o. It is referred to as an internal kernel product, where

$$w = \sum_{i=1}^{N} \alpha_i d_i \varphi\left(x_i\right) \tag{63}$$

$$\varphi(x) = \left[\varphi_o(x), \varphi_1(x), \ldots\ldots, \varphi_{m_1}(x)\right]^T \tag{64}$$

Here, $\varphi_j(x) \quad for \quad j = 1,,, m_1$ represents non-linear transformations and $\varphi_o(x) = 1$ for every x, while w_o signifies the bias b. α_i and d_i represent the Lagrange coefficient and target output.

Step 2: The Mercer's theorem must be satisfied in order for the kernel $K(x, x_i)$ mentioned in Eq. 65 to perform its purpose, which is selected as polynomial

$$K\left(x, x_i\right) = \left(1 + x^T x_i\right) \tag{65}$$

Step 3: The Lagrange multipliers $\{\alpha_i\}$ for $i = 1, \ldots\ldots 10$ in order to optimize the function $Q(\alpha)$, abbreviated as $\alpha_{0,i}$ is given as

$$Q(\alpha) \sum_{i=1}^{N} = \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j K\left(x_i, x_j\right) \tag{66}$$

Further, it follows the constraints given in Eqs. 68 and 69:

$$\sum_{i=1}^{N} \alpha_i d_i = 0 \tag{68}$$

$$0 \le \alpha_i \le C \quad for \quad i = 1, 2, \ldots\ldots, N \tag{69}$$

Step 4: To finish, the vector w_o with linear weight signifies that the optimal values of Lagrange multipliers are given in Eq. 70 as

$$w_o \sum_{i=1}^{N} \alpha_{0,i} d_i \varphi\left(x_i\right) \tag{70}$$

Here, $\varphi(x_i)$ signifies the image in the feature space due to x_i. On the other hand, w_o indicates the ideal bias b_o.

3. Random forest (RF): Random forest is an ensemble method that computes by constructing a multitude of decision trees during training and outputs the class based on the mean or mode of the classes. The decision tree is a widely used method in machine learning that combines the ability to learn with the capability to serve as a ready-made mechanism for data mining [40]. The training algorithm for RF utilizes the common approach of bagging, also known as bootstrap aggregation, to facilitate the learning process of the individual trees [36]. Hence, for a given training set $X = x_1, \ldots\ldots, x_n$ with labels $Y = y_1, \ldots\ldots, y_n$, the iterative bagging involves randomly selecting an instance by interchanging the training set and fitting trees to the samples, as shown; thus, for $b = 1, \ldots\ldots, B$:

(a) Sample with interchanging n training example X_b, Y_b for X, Y.
(b) Train a classifier tree f_b for X_b, Y_b.

After training of $x^{'}$ samples, predictions are built by computing the mean of individual classifier tree predictions on $x^{'}$ and are represented as mentioned in Eq. 71:

$$\hat{f} = \frac{1}{B} \sum_{b=1}^{B} f_b\left(x'\right) \tag{71}$$

This further leads in reduction of the variance and sensitivity without enlarging the bias, as long as the trees are non-interacting. Additionally, for uncertainty estimate, predictions can be formed using the standard deviation derived from multiple individual predictions using Eq. 72:

$$\sigma = \sqrt{\frac{\sum_{b=1}^{B} \left(f_b(x') - \hat{f}\right)^2}{B - 1}} \tag{72}$$

K Nearest Neighbor (k-NN):

Nearest neighbor is the common approach used for the recognition of statistical patterns which form finite partition of sample space for any observation x' classified into jth class [41]. For a point x' in the d^{th} dimensional space of feature, function $f_x : \mathbb{R}^d \longrightarrow \mathbb{R}$ is computed based on Euclidean distance $f_x(x) = \|x - x^{'}\|$. Further, the entire training set comprising X samples are ordered with respect to x' corresponding to function $r'_x : \{1, \ldots\ldots, N\} \longrightarrow \{1, \ldots\ldots N\}$ which reorders the indexing of N training points. Thus, the function is recursively defined using Eq. 73:

$$\begin{aligned} r_{x'}(1) &= \arg\min_i f_{x'}\left(x_i\right) \quad with\ i \in \{1, \ldots\ldots N\} \\ r_{x'}(j) &= \arg\min_i f_{x'}\left(x_i\right) \quad with\ i \in \{1, \ldots\ldots N\} \\ i \ne r_{x'}(1) \ldots\ldots\ldots, i \ne r_{x'}(j-1)\ &for \quad j = 2, \ldots\ldots, N \end{aligned} \tag{73}$$

Table 4 Details of parameter settings used for the considered classifiers

S. No.	Classifier	Details of parameter settings utilized
1	Naive Bayes	With kernel estimation function
2	SVM	Sequential minimal optimization with c = 1.0, epsilon = 1.0 e^{-12}, tolerance = 0.01
3	Random forest	With 100 iterations
4	K-NN	Using the Euclidean distance function and by taking 2 as value for k
5	Neural network	MLP with learning rate = 0.4, momentum = 0.3, and validation threshold = 25

Here, $x_{r_{x'}(j)}$ is point on set X in position j from $x^{'}$ known as jth nearest neighbor, while $f_{x'}\left(x_{r_{x'}(j)}\right) = \left\|x_{r_{x'}(j)} - x'\right\|$ is the distance from $x^{'}$, which can be stated as $j < k \Rightarrow f_{x'}\left(x_{r_{x'}(j)}\right) \leq f_{x'}\left(x_{r_{x'}(k)}\right)$. Hence, the decision rule for k-NN classifier to perform binary classification is stated using the rule mentioned in Eq. 74:

$$kNN(x) = \text{sign}\left(\sum_{i=1}^{k} y_{r_x}(i)\right) \tag{74}$$

Here, class labels in the ith nearest sample for the training set are represented by the symbol $y_{r_x(i)} \in \{-1, +1\}$.

4. Neural network (NN): A computational method that is based on the structure and operation of biological neurons is called a neural network. It contains information from NN structure so network learns from input and output. Nonlinear statistical data with complicated input-output relationships is often modeled using this paradigm [41]. A multilayer perceptron (MLP) has three interconnected layers: initial input, primary hidden, and final output. Each node or neuron uses a non-linear activation function and supervised learning to train the classifier. Sigmoids are the activation functions that are utilized, and they are represented using Eq. 75:

$$y\left(v_i\right) = \tanh\left(v_i\right) \; and \; y\left(v_i\right) = \left(1 + e^{-v_i}\right)^{-1} \tag{75}$$

One of the activations that is utilized in this context is the hyperbolic tangent, whose range is set between −1 and 1, while the other activation is the logistic function, with the range varying from 0 to 1. Here, y_i indicates ith node output and v_i denotes weighted summation of input connections. As a result of the fact that MLP is a highly connected design, every single node in a single layer is connected to the nodes of the layer below it by means of weight w_{ij}.

Lastly, the learning process is carried out by adjusting the weights based on the corrections that are designed to reduce the amount of inaccuracy in the overall output using Eq. 76:

$$\varepsilon(n) = \frac{1}{2}\sum_{j} e_j^2(n) \tag{76}$$

Additionally, the change in weight during gradient descent can be expressed using Eq. 78:

:

$$-\frac{\partial \varepsilon(n)}{\partial v_j(n)} = e_j(n)\phi'\left(v_j(n)\right) \tag{78}$$

The study of the change in weights for hiding layers, on the other hand, is rather complicated; hence, the derivative can be written using Eq. 79:

$$-\frac{\partial \varepsilon(n)}{\partial v_j(n)} = \phi'\left(v_j(n)\right)\sum_{k} -\frac{\partial \varepsilon(n)}{\partial v_k(n)} w_{kj}(n) \tag{79}$$

Here, $\phi^{'}$ signifies the activation function. The initiation parameters that were utilized for classification are demonstrated in Table 4.

Experimental Results and Discussion

This section explains the experimental platform that was utilized during the assessment of presented texture models, as well as the detailed results and findings of different experimental setups. The proposed technique for categorization of renal lesions has been tested on CT datasets acquired

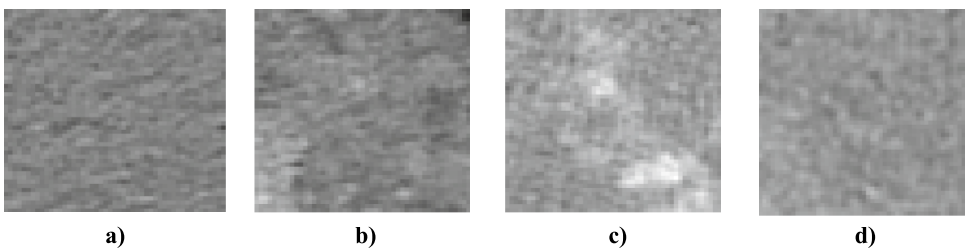

Fig. 4 Some examples of ROI considered in this study where **a**, **b** represent benign lesion **c**, **d** are malignant lesions

from 50 subjects using different performance assessment metrics. For further performance evaluation, five-fold cross validation was utilized, with four-fifths of the relevant dataset being used for training and the remaining one-fifth being used for testing. To eliminate any sampling bias, the same procedure was done five times with each fold picked at random. The accuracy obtained for numerous rounds is averaged to calculate total classifier accuracy.

Dataset Details

The experimentation dataset was made out of CT scans received from PGIMER Chandigarh, India. On a Windows 10 system with a CPU speed of 2.3 GHz and 6 GB of memory, the proposed technique was tested using MATLAB. This study looked at CT images acquired from 50 patients. The 128-slice SIEMENS CT-scanner with slice thickness varying from 0.6 to 1 mm is used to capture abdominal CT images in non-contrast-enhanced arterial, venous, portal venous, and delayed phase. Fig. 4 depicts some of the image examples considered in this study. The goal of this study was to determine the applicability of the texture models under consideration by evaluating their capabilities on a variety of images.

Metrics for Performance Assessment

The proposed method's competence in categorization of renal lesions is assessed using different metrics which are discussed in this section. The following are some of the terms used to calculate them:

True positive (TP): a lesion that has been classified as cancerous but is not actually cancerous.

True negative (TN): a benign lesion that is not genuinely benign.

False positive (FP): a benign lesion that has been classified as cancerous.

False negative (FN): a benign lesion that turns out to be cancerous.

Sensitivity Sensitivity relates to a technique of characterization of capacity to identify a large proportion for true positive instances, or people with the condition, and it is calculated using Eq. 80. All positives are projected as positives with 100% sensitivity, which indicates malignant lesions are classified as malignant:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{80}$$

Specificity This relates to a characterization technique's capacity to detect a higher number of negative cases, i.e., individuals who do not have disease or benign lesions are categorized as benign, and it is calculated using Eq. 81:

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{81}$$

Accuracy It is used to assess a classifier's ability to categorize a large number of images, and it is calculated using Eq. 82. A high accuracy value indicates that the classifier is doing well:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{82}$$

Roc-AUC The area under the receiver-operating characteristic (ROC-AUC) is a statistic used to evaluate the classifier capability. AUC measures an experiment's precision; a value near 1 indicates more specificity and sensitivity.

F score: The F score is a weighted average of recall and precision that can be used to assess classification accuracy.

The F-score is a number between one and zero, with one being the best and zero being the worst.

Results and Discussion

This section examines the classification results obtained by combining the features from various texture models for categorizing benign and malignant renal lesions to establish each model's capabilities. A total of 44 features were extracted individual texture models' classification results are shown in Table 1. Each texture model contributes for classification of renal lesions, although each one of them has their associated limitations and benefits. The FoS texture model works by ignoring image pixel interaction and merely predicting the attributes of different pixel values. The SGLCM model, on the other hand, evaluates image attributes using second-order statistics that look at

Table 5 Performance comparison of distinct classifiers by considering all features (no. of features = 44)

Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)	ROC-AUC	F-score (%)
K-NN	83.33	89.7	88.5	0.840	0.834
NN	88.88	90.4	58.6	0.789	0.801
SVM	86.66	88.2	80.4	0.717	0.856
RF	90.88	92.4	90.7	0.889	0.889
Naïve bayes	76.38	85.7	85.5	0.764	0.764

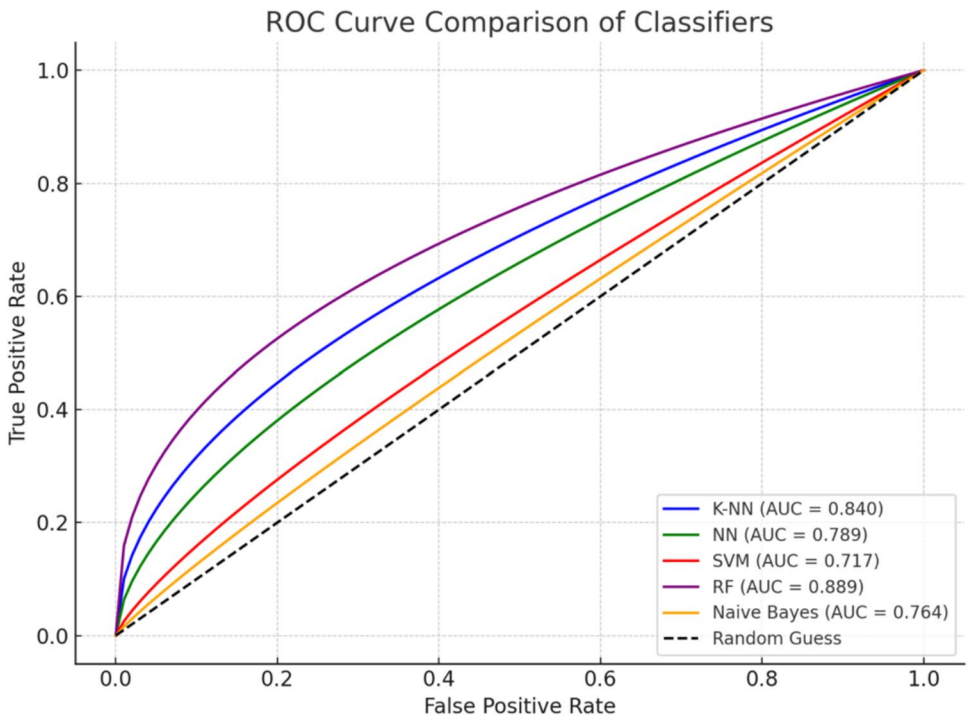

Fig. 5 ROC curve comparison of different machine learning classifiers with 44 features

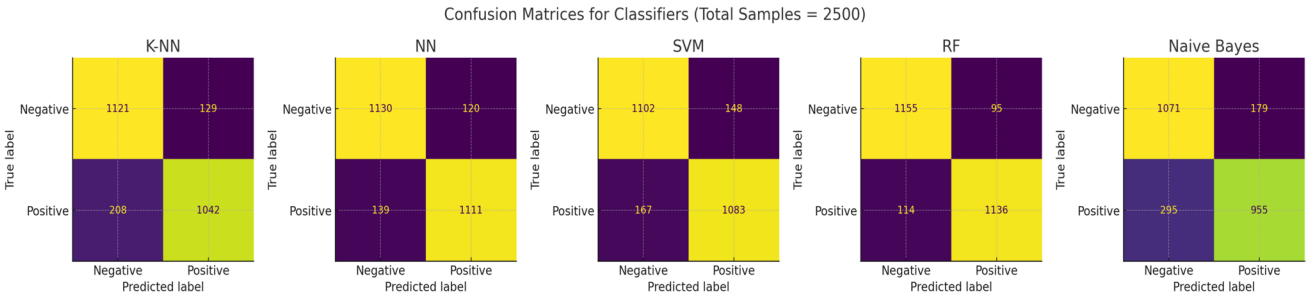

Fig. 6 Confusion matrix of different machine learning classifiers with 40 features

Table 6 Performance comparison of distinct classifiers by considering most discriminant features (no. of features = 10)

Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)	ROC-AUC	F-score (%)
K-NN	90.66	91.7	92.1	0.917	0.916
NN	91.44	92.4	70.2	0.851	0.844
SVM	90.22	91.2	88.4	0.874	0.872
RF	94.44	93.4	95.8	0.951	0.944
Naive Bayes	89.65	89.7	90.7	0.897	0.791

the combined probability distribution of pixel pairings. In terms of tonal variations, the malignant lesion differs from the benign lesion, and the SGLCM model has been found to perform better in determining the difference between benign and malignant lesions by evaluating tonal variations. Furthermore, the FPS model takes into account the energy distribution in the image Fourier transform's power spectrum, which evaluates texture directionality and coarseness of texture patterns. Other texture models' performance results are unsatisfactory because they may be unable to evaluate texture features for the photos in question efficiently. Furthermore, different feature selection techniques have been employed to choose the most discriminating features for dimensionality reduction. While comparing to all features obtained from different texture models, the selected set of 10 features mentioned

in Table 3 produces better classification performance when examined, according to empirical evidence acquired after executing numerous trials. Moreover, in comparison to the separate texture models, the performance of the classifier improves when all the features derived from distinct texture models are combined to generate a feature vector consisting of 44 features. Combining all texture features, alternately, increases dimensionality of the feature vector, which increases the computational rate. To overcome this issue, the authors used three feature selection techniques, namely CFS, IG, and GR, to choose the best discriminating characteristics to be utilized to categorize malignant and benign renal lesions in order to reduce feature vector dimensionality. Table 3 shows the set of selected feature vectors by different feature selection techniques from a collection of 44 features, and this proposed feature vector has demonstrated good classification results. The dimensionally reduced feature vector was passed to the considered classifiers for classification. The classifiers were trained and validated using training and testing data using the fivefold cross-validation technique. For classifier generation, the training set has 10 selected features and a class label, whereas the testing set only contains the feature vector and no class label. To determine various performance measures, the outcomes of the classifier model for the test dataset were compared to the real class labels. The initial feature vector and proposed feature vector's performance were assessed using the various assessment metrics stated in the preceding section, including sensitivity, accuracy, specificity, ROC-AUC, and F-score. Table 5 shows the performance of different classifiers using all 44 features for classification in terms of various evaluation metrics.

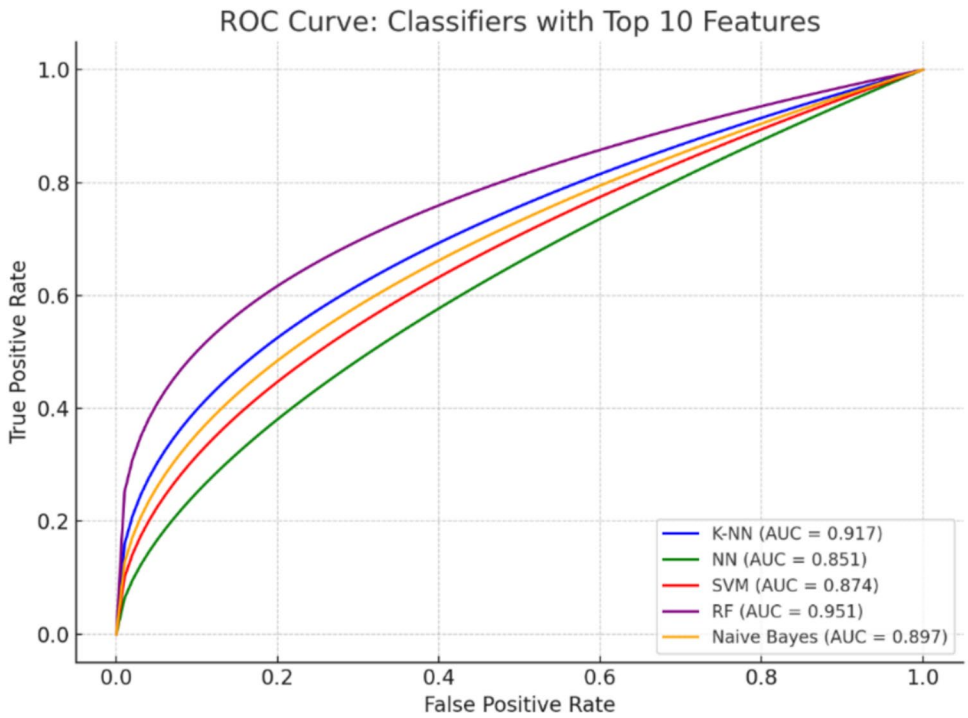

Fig. 7 ROC curve comparison of different machine learning classifiers with 10 features

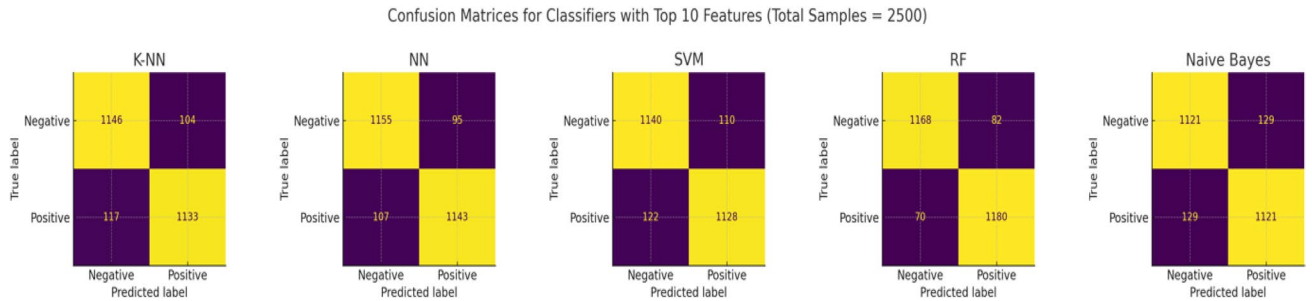

Fig. 8 Confusion matrix of different machine learning classifier with 10 features

Additionally, the performance of five classifiers namely K-nearest neighbors (K-NN), neural network (NN), support vector machine (SVM), random forest (RF), and naive Bayes was evaluated using ROC-AUC curves as shown in Fig. 5 and confusion matrices on a dataset comprising 2500 samples (1250 positive and 1250 negative) as shown in Fig. 6. The ROC curve provides a graphical representation of the trade-off between the true positive rate and the false positive rate across various thresholds. As illustrated in the ROC plot, random forest demonstrated the highest area under the curve (AUC = 0.889), indicating superior discriminative power. K-NN and NN followed with AUC scores of 0.840 and 0.789, respectively. SVM showed moderate performance with an AUC of 0.717, while naive Bayes yielded the lowest AUC of 0.764, suggesting limited classification effectiveness.

Additionally, confusion matrices were generated to examine each model's classification ability in terms of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) as shown in Fig. 6. Random forest achieved the best results with 1136 TP and 1155 TN, reflecting its high sensitivity (90.88%) and specificity (92.4%). K-NN and NN also performed well, although NN had slightly higher false negatives. SVM displayed balanced performance but incurred more misclassifications compared to RF and K-NN. Naive Bayes, in contrast, showed the poorest outcomes with a higher number of FP (179) and FN (295), corroborating its lower AUC. The ROC curve also includes a reference diagonal known as the random guess line (AUC = 0.5) which represents a model with no discriminative capability. All classifiers evaluated in this study performed well above this baseline,

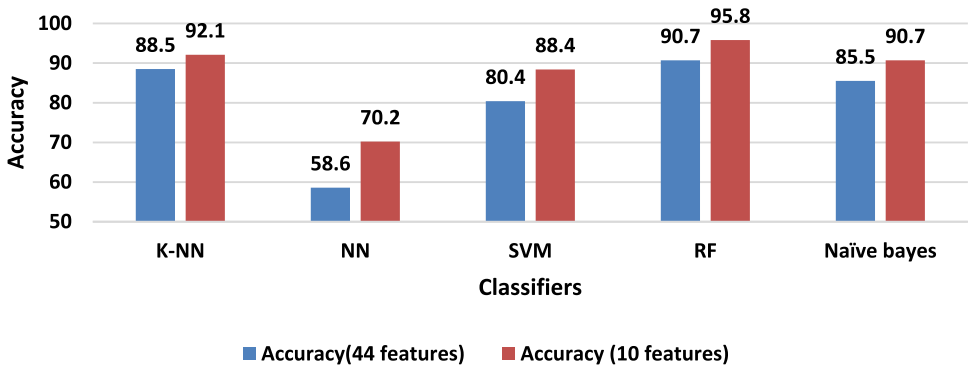

Fig. 9 Comparison of different classifiers in terms of accuracy using different feature sets

Fig. 10 Comparison plots of sensitivity and specificity for distinct classification techniques

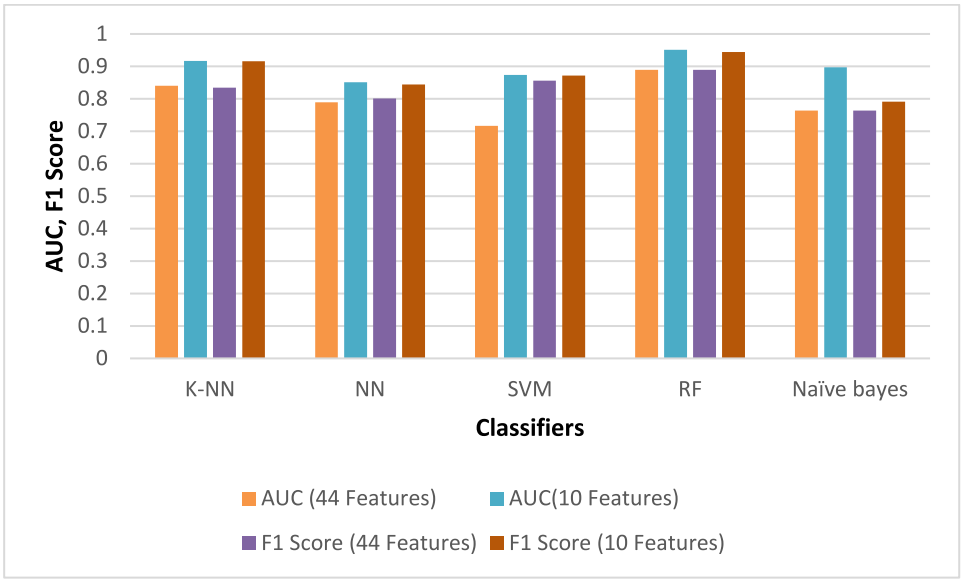

Fig. 11 Comparison plots of AUC and F1 Score for distinct classification techniques

Table 7 Comparison with existing studies on classification of renal lesions

Reference	Feature vector	Classifier	Accuracy
Linguraru et al. [19]	Histogram curvature features	Random forest	90.5%
SP Raman et al. [20]	Entropy, standard deviation, skewness and kurtosis	Random forest	91.2%
QH et al. [22]	First-order and second-order statistics	Ensemble model based on decision trees	90.5%
Xu et al. [23]	FoS, GLCM, GLRM	Random forest	93.2%
Han JH et al. [25]	Radiological (HU) and radiomic features	Random forest	85.0%
Miskin et al. [24]	Mean, SD, entropy, skewness, and kurtosis	Random forest	89.2%
Proposed method	Selected 10 features	Random forest	95.8%

indicating learned patterns beyond random chance. These results clearly establish Random forest as the most robust and reliable model for the given dataset.

Further, the classification performance of five machine learning models, including K-NN, neural network (NN), support vector machine (SVM), random forest (RF), and naive Bayes, was re-evaluated using the 10 most discriminant features as reported in Table 6. Moreover, the ROC curves, shown in Fig. 7, illustrate improved true positive vs. false positive trade-offs for each classifier. Random forest achieved the highest AUC (0.951), signifying excellent overall discriminative power. K-NN (AUC = 0.917), SVM (AUC = 0.874), and naive Bayes (AUC = 0.897) also demonstrated strong performance. The NN model, despite showing high sensitivity (91.44%) and specificity (92.4%), yielded a relatively lower AUC (0.851), possibly due to inconsistent performance at different thresholds.

The confusion matrices as shown in Fig. 8 further elucidate classifier behavior across a dataset of 2500 samples (1250 positive and 1250 negative cases). Random forest outperformed others with 1181 true positives and 1168 true negatives, reflecting its superior sensitivity (94.44%) and specificity (93.4%). K-NN and SVM also showed robust results with high TP and TN values, indicating balanced classification. Naive Bayes, while yielding relatively high TP and TN counts, had slightly more false predictions compared to RF and K-NN. The NN classifier recorded 1143 TPs and 1155 TNs, yet had performance limitations likely due to overfitting or threshold sensitivity. Overall, the use of top-ranked features significantly enhanced the performance of all classifiers, with random forest emerging as the most reliable model.

Random Forest algorithm has the ability to handle high-dimensional data and mitigate overfitting through ensemble learning, which likely contributed to its superior performance in our study. On the other hand, SVM struggled with scalability and required intensive parameter tuning, while naive Bayes made strong independence assumptions that did not hold true for our feature set, leading to suboptimal performance. The highest accuracy among all the machine learning classifiers was obtained by RF classifier equivalent to 95.8%, as shown in Fig. 9 by employing a discriminating feature set consisting of 10 features.

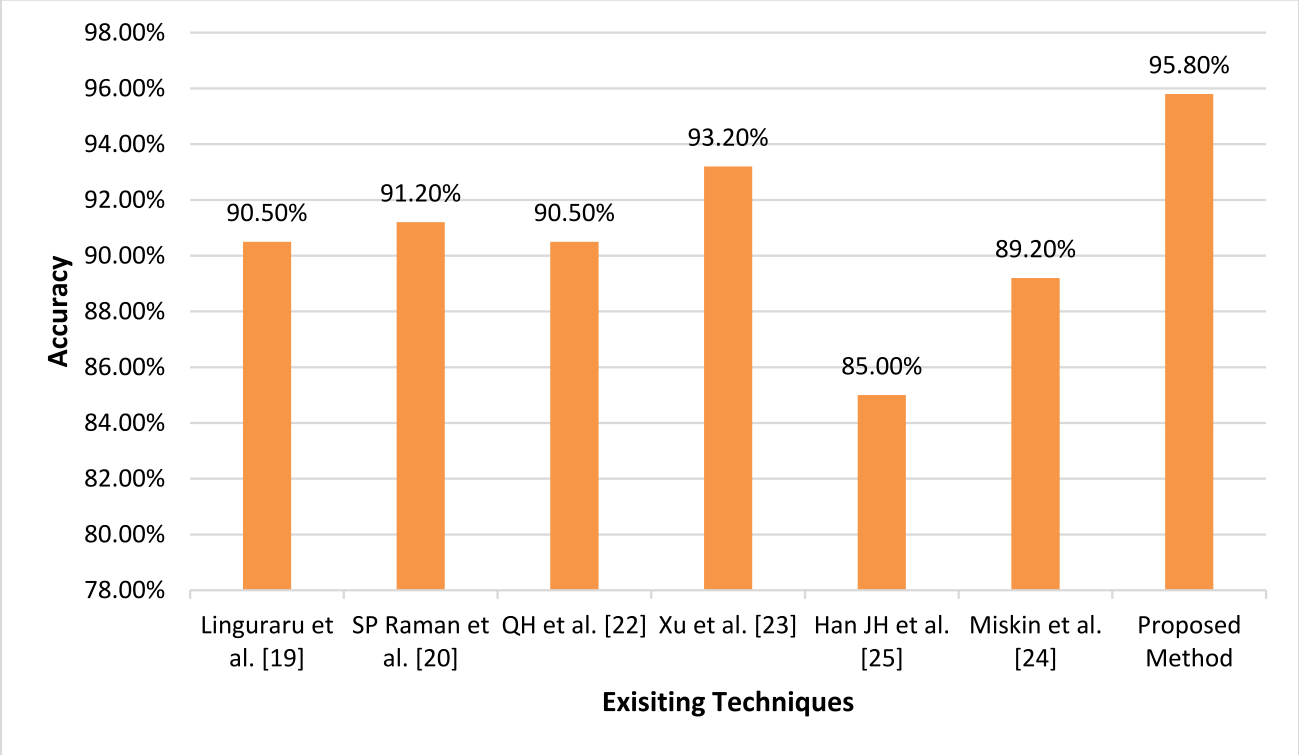

Fig. 12 Comparison of proposed technique with exiting approaches in terms of accuracy

Further, Fig. 10 illustrates the comparison of sensitivity and specificity values by considering both feature vectors consisting of 44 features and 10 features, and it is clearly evident that feature vector consisting of 10 features performs better as compared to others. The F-score and AUC results were also produced to evaluate classification performance as shown in Fig. 11, providing the weighted average of recall and precision, and the highest value of F-score and AUC was reported by RF model by considering the selected 10, making it more suitable for real-time classification tasks of renal lesions

The favorable values of sensitivity and specificity of the proposed feature set combination are 94.4% and 93.4% reported by employing a random forest classifier, respectively, and estimating the likelihood of the negative and positive labels being correct. Thus, the observed results show that the suggested feature vector consisting of 10 features performs better in the categorization of renal lesions from CT images in terms of all evaluated metrics. Further, the comparative analysis of the proposed methodology with the existing approaches for categorizing renal lesions has been highlighted in Table 7 which shows how the proposed classification technique performs better as compared to other state-of-the-art methods. Thus, based on experimental analysis, it can be found that reducing the number of features and creating a new hybrid optimal feature improved the performance of classification. Also, among the classifiers used, namely k-NN, NN, SVM, RF, and naive Bayes, the performance of RF is found to outperform others with accuracy of 95.8%. The performance parameters were analyzed using the same dataset, yielding reliable results. However, the suggested approach is more accurate in categorizing renal lesions than other current research on the same database, as evidenced by performing different experimental trials and also by the values given in Table 7 and Fig. 12.

The proposed feature vector for the classification of renal lesion offers several significant advantages over traditional approaches. Firstly, by integrating diverse texture models, it captures both global and local spatial patterns, enhancing the characterization of renal lesion tissue heterogeneity. Secondly, the use of advanced feature selection methods—correlation-based feature selection, information gain, and gain ratio—enables the identification of the most discriminative features, thereby reducing computational complexity while maintaining high diagnostic accuracy. Thirdly, the random forest classifier, due to its ensemble nature, provides robust and generalizable performance even with limited data, achieving a notable accuracy of 95.8%. Lastly, the hybrid feature design employs less computational power by reducing the feature dimensionality, making the model more suitable for clinical decision support and real-time diagnostic applications. Further, reducing dimensionality in classification tasks offers several significant

advantages. It helps minimize overfitting by eliminating irrelevant or redundant features, allowing the model to generalize better to unseen data. This reduction also leads to improved model performance, as many classification algorithms operate more efficiently with fewer, more meaningful features. Additionally, it lowers computational cost by reducing memory usage and speeding up both training and inference times.

Clinical Relevance and Practical Application

The proposed hybrid feature-based classification framework has significant potential for integration into a computer-aided diagnosis (CAD) tool in real-world clinical settings. By leveraging texture features from CT images and using widely accepted machine learning algorithms, the model can serve as a reliable CAD tool to assist radiologists in distinguishing benign from malignant renal lesions. Its relatively low computational complexity and compatibility with standard DICOM image formats make it feasible for deployment within existing radiology workflows. Furthermore, the use of meaningful features enhances clinical trust, which is essential for adoption in healthcare environments. With further validation using larger and diverse patient datasets, the system could be embedded into PACS (Picture Archiving and Communication Systems) or radiology software to provide a valuable second opinion during routine kidney cancer screening or diagnosis.

Conclusion

This study introduces a machine learning–based technique based on novel hybrid feature vector for classifying benign and malignant renal lesions. The suggested feature vector combines multiple texture models—FoS, GLDS, SGLCM, NGTDM, Law's TEM, FPS, SFM, and Fractal—to capture diverse structural and statistical characteristics of renal tissue. Unlike previous studies that rely on limited or single-type feature representations, the proposed method leverages a hybrid feature vector, by extracting the information from eight different texture models for renal lesion classification. To enhance model efficiency and reduce computational complexity, three robust feature selection techniques—information gain, gain ratio, and correlation-based selection—were applied to derive a compact set of the most discriminative 10 features. The experimental results demonstrated that by employing reduced hybrid feature set consisting of 10 features for texture analysis; classification performance has been substantially improved. The accuracy of the suggested technique is reported to be 95.8% by using Random Forest Classifier, which is higher as compared to the existing technique. The results highlight the potential of the proposed feature vector in supporting CAD systems for renal cancer. By effectively combining multiple texture descriptors and machine learning, the model demonstrates promising utility as a second-opinion tool for radiologists, subject to further clinical validation.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Ravinder Kaur, Sonam Khattar, and Sanjay Singla. The first draft of the manuscript was written by Ravinder Kaur, and all authors gave their suggestions on previous versions of the manuscript. All authors read and approved the final manuscript.

Declarations

Ethics Approval This research study has been conducted on images, and no ethical approval is required.

Competing Interests The authors declare no competing interests.

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA: a cancer journal for clinicians 2024 Jan 1;74(1).
2. Islami F, Baeker Bispo J, Lee H, Wiese D, Yabroff KR, Bandi P, Sloan K, Patel AV, Daniels EC, Kamal AH, Guerra CE. American Cancer Society's report on the status of cancer disparities in the United States, 2023. CA: A Cancer Journal for Clinicians 2024; 74(2):136-66.
3. Kiri S, Ryba T. Cancer, metastasis, and the epigenome. Molecular Cancer 2024; 23(1):154.
4. Patel VV, Yadav AR. A review on kidney tumor segmentation and detection using different artificial intelligence algorithms. In AIP Conference Proceedings 2024 (Vol. 3107, No. 1), AIP Publishing.
5. Distante A, Marandino L, Bertolo R, Ingels A, Pavan N, Pecoraro A, Marchioni M, Carbonara U, Erdem S, Amparore D, Campi R. Artificial intelligence in renal cell carcinoma histopathology: Current applications and future perspectives. Diagnostics 2023; 13(13):2294.
6. Wang X, Zhu Z. Context understanding in computer vision: A survey. Computer Vision and Image Understanding 2023;229:103646.
7. Yasuda Y, Zhang JH, Attawettayanon W, Rathi N, Wilkins L, Roversi G, Zhang A, Accioly JP, Shah S, Munoz-Lopez C, Palacios DA. Comprehensive management of renal masses in solitary kidneys. European Urology Oncology 2023;6(1):84-94.
8. Hora M, Albiges L, Bedke J, Campi R, Capitanio U, Giles RH, Ljungberg B, Marconi L, Klatte T, Volpe A, Abu-Ghanem Y. European Association of Urology guidelines panel on renal cell carcinoma update on the new World Health Organization classification of kidney tumors 2022: the urologist's point of view. Eur Urol 2023; 83(2):97-100.
9. Jiang P, Ali SN, Peta A, Arada RB, Brevik A, Xie L, Okhunov Z, Clayman RV, Landman J. A review of the recommendations and strength of evidence for clinical practice guidelines on the management of small renal masses. Journal of Endourology 2023; 37(8):903-13.
10. Williamson SR, Taneja K, Cheng L. Renal cell carcinoma staging: pitfalls, challenges, and updates. Histopathology 2019; 74(1):18-30.
11. Hussain S, Mubeen I, Ullah N, Shah SS, Khan BA, Zahoor M, Ullah R, Khan FA, Sultan MA. Modern diagnostic imaging

technique applications and risk factors in the medical field: a review. BioMed research international 2022;2022(1):5164970.
12. Kaur R, Juneja M. A Survey of Different Imaging Modalities for Renal Cancer. Indian Journal of Science and Technology 2016;9(44).
13. Kaur R, Juneja M, Mandal AK. Computer-aided diagnosis of renal lesions in CT images: a comprehensive survey and future prospects. Computers & Electrical Engineering 2019;77:423-34.
14. Kaur R, Juneja M. A survey of kidney segmentation techniques in CT images. Current Medical Imaging 2018;14(2):238-50.
15. Hossain E, Hossain MS, Hossain MS, Al Jannat S, Huda M, Alsharif S, Faragallah OS, Eid M, Rashed AN. Brain Tumor Auto-Segmentation on Multimodal Imaging Modalities Using Deep Neural Network. Computers, Materials & Continua 2022;72(3).
16. Inthiyaz S, Altahan BR, Ahammad SH, Rajesh V, Kalangi RR, Smirani LK, Hossain MA, Rashed AN. Skin disease detection using deep learning. Advances in Engineering Software 2023; 175:103361.
17. Soundararajan R, Prabu AV, Routray S, Malla PP, Ray AK, Palai G, Faragallah OS, Baz M, Abualnaja MM, Eid MM, Rashed AN. Deeply trained real-time body sensor networks for analyzing the symptoms of Parkinson's disease. IEEE Access 2022; 10:63403-21.
18. Sun HY, Kim JH, Hwang J, Hong SS, Doo SW, Yang WJ, Cho YJ, Song YS. Diagnostic accuracy of contrast-enhanced computed tomography and contrast-enhanced magnetic resonance imaging of small renal masses in real practice: sensitivity and specificity according to subjective radiologic interpretation. World journal of surgical oncology 2016; 14(1):260.
19. MG Linguraru, S. Wang, F. Shah, R. Gautam, J. Peterson, WM. Linehan, and RM Summers, Automated noninvasive classification of renal cancer on multiphase CT. Medical physics, 38(2011), pp. 5738-5746.
20. SP Raman, Y Chen, JL Schroeder, P Huang, and EK Fishman, CT Texture Analysis of Renal Masses: Pilot Study Using Random Forest Classification for Prediction of Pathology. Academic radiology 21(2014), pp. 1587-1596.
21. J. Liu, S. Wang, MG Linguraru, J. Yao and RM Summers, Computer-aided detection of exophytic renal lesions on non-contrast CT images. Medical image analysis, 19(2015), pp.15-29.
22. He QH, Feng JJ, Lv FJ, Jiang Q, Xiao MZ. Deep learning and radiomic feature-based blending ensemble classifier for malignancy risk prediction in cystic renal lesions. Insights into Imaging 2023;14(1):6.
23. Xu J, He X, Shao W, Bian J, Terry R. Classification of Benign and Malignant Renal Tumors Based on CT Scans and Clinical Data Using Machine Learning Methods. In Informatics 2023 (Vol. 10, No. 3, p. 55). MDPI.
24. Miskin N, Qin L, Silverman SG, Shinagare AB. Differentiating benign from malignant cystic renal masses: a feasibility study of computed tomography texture-based machine learning algorithms. Journal of computer assisted tomography 2023; 47(3):376-81.
25. Han JH, Kim BW, Kim TM, Ko JY, Choi SJ, Kang M, Kim SY, Cho JY, Ku JH, Kwak C, Kim YG. Fully automated segmentation and classification of renal tumors on CT scans via machine learning. BMC cancer 2025; 25(1):173.
26. Kaur R, Juneja M, Mandal AK. Machine learning based quantitative texture analysis of CT images for diagnosis of renal lesions. Biomedical Signal Processing and Control 2021;64:102311.
27. Feng Z, Rong P, Cao P, Zhou Q, Zhu W, Yan Z, Liu Q, Wang W. Machine learning-based quantitative texture analysis of CT images of small renal masses: differentiation of angiomyolipoma without visible fat from renal cell carcinoma. European radiology 2018; 28:1625-33.
28. Sarvestani ZM, Jamali J, Taghizadeh M, Dindarloo MH. A novel machine learning approach on texture analysis for automatic breast microcalcification diagnosis classification of mammogram images. Journal of Cancer Research and Clinical Oncology 2023; 149(9):6151-70.
29. Chen A, Karwoski RA, Gierada DS, Bartholmai BJ, Koo CW. Quantitative CT analysis of diffuse lung disease. Radiographics 2020 Jan;40(1):28-43.
30. Ramola A, Shakya AK, Van Pham D. Study of statistical methods for texture analysis and their modern evolutions. Engineering Reports 2020; 2(4):e12149.
31. Yap FY, Varghese BA, Cen SY, Hwang DH, Lei X, Desai B, Lau C, Yang LL, Fullenkamp AJ, Hajian S, Rivas M. Shape and texture-based radiomics signature on CT effectively discriminates benign from malignant renal masses. European radiology 2021; 31:1011-21.
32. Chandra TB, Verma K, Singh BK, Jain D, Netam SS. Automatic detection of tuberculosis related abnormalities in Chest X-ray images using hierarchical feature extraction scheme. Expert Systems with Applications 2020 15;158:113514.
33. Ghalati MK, Nunes A, Ferreira H, Serranho P, Bernardes R. Texture analysis and its applications in biomedical imaging: A survey. IEEE Reviews in Biomedical Engineering 2021 27;15:222-46.
34. Varghese BA, Fields BK, Hwang DH, Duddalwar VA, Matcuk Jr. GR, Cen SY. Spatial assessments in texture analysis: what the radiologist needs to know. Frontiers in Radiology 2023 24;3:1240544.
35. Pudjihartono N, Fadason T, Kempa-Liehr AW, O'Sullivan JM. A review of feature selection methods for machine learning-based disease risk prediction. Frontiers in Bioinformatics 2022 27;2:927312.
36. Uğuz H. A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm. Knowledge-Based Systems 2011 ;24(7):1024-32.
37. Pasha SJ, Mohamed ES. Advanced hybrid ensemble gain ratio feature selection model using machine learning for enhanced disease risk prediction. Informatics in Medicine Unlocked 2022 1;32:101064.
38. Jetti CR, Shaik R, Shaik S, Sanagapalli S. Disease prediction using Naïve Bayes—machine learning algorithm. International Journal of Science & Healthcare Research 2021.
39. Almansour NA, Syed HF, Khayat NR, Altheeb RK, Juri RE, Alhiyafi J, Alrashed S, Olatunji SO. Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. Computers in biology and medicine 2019;109:101-11.
40. Swarupa AN, Sree VH, Nookambika S, Kishore YK, Teja UR. Disease prediction: smart disease prediction system using random forest algorithm. In 2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT) 2021 (pp. 48-51). IEEE.
41. Uddin S, Haque I, Lu H, Moni MA, Gide E. Comparative performance analysis of K-nearest neighbor (KNN) algorithm and its different variants for disease prediction. Scientific Reports 2022;12(1):6256.