Article published online: 2024-12-11

Original Article

Predicting Renal Cell Carcinoma Subtypes and Fuhrman Grading Using Multiphasic CT-Based Texture Analysis and Machine Learning Techniques

Amit Gupta[1] Sanil Garg[1] Neel Yadav[1] Rohan Raju Dhanakshirur[2] Kshitiz Jain[3] Rishi Nayyar[4] Seema Kaushal[5] Chandan J. Das[1]

[1] Department of Radiodiagnosis and Interventional Radiology, All India Institute of Medical Sciences, New Delhi, India
[2] Amarnath and Shashi Khosla School of Information Technology, Indian Institute of Technology Delhi, New Delhi, India
[3] Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi, New Delhi, India
[4] Department of Urology, All India Institute of Medical Sciences, New Delhi, India
[5] Department of Pathology, All India Institute of Medical Sciences, New Delhi, India

Address for correspondence Chandan J. Das, MD, PhD, DNB, FRCP., Edin, Professor, Department of Radio-diagnosis and Interventional Radiology, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110029, India (e-mail: dascj@yahoo.com).

Indian J Radiol Imaging

Abstract

Objectives The aim of this study is to evaluate computed tomography texture analysis (CTTA) on multiphase CT scans for distinguishing clear cell renal cell carcinoma (ccRCC) from non-ccRCC and predicting Fuhrman's grade in ccRCC using open-source Python libraries.

Methods Conducted retrospectively, the study included 144 patients with RCCs (108 ccRCCs and 36 non-ccRCCs) who underwent preoperative multiphasic CT. Ninety ccRCCs were categorized into 71 low-grade and 19 high-grade ccRCCs. Tumor was marked on the largest axial tumor slice using "LabelMe" across different CT phases. First- and second-order texture features were computed using Python's scipy, numpy, and opencv libraries. Multivariable logistic regression analysis and machine learning (ML) models were used to evaluate CTTA parameters from different CT phases for RCC classification. The best ML model for distinguishing ccRCC and non-ccRCC was externally validated using data from the 2019 Kidney and Kidney Tumor Segmentation Challenge.

Results Entropy in the corticomedullary (CM) phase was the best individual parameter for distinguishing ccRCC from non-ccRCC with (F1 score: 0.83). The support vector machine (SVM) based ML model, incorporating CM phase features, performed the best, with an F1 score of 0.87. External validation for the same model yielded an accuracy of 0.82 and an F1 score of 0.81. ML models and individual texture parameters showed less accuracy for classifying

Keywords
- texture analysis
- machine learning
- renal cell carcinoma

DOI https://doi.org/10.1055/s-0044-1796639.
ISSN 0971-3026.

© 2024. Indian Radiological Association. All rights reserved.
This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (https://creativecommons.org/licenses/by-nc-nd/4.0/)

Thieme Medical and Scientific Publishers Pvt. Ltd., A-12, 2nd Floor, Sector 2, Noida-201301 UP, India

low- versus high-grade ccRCCs, with a maximum F1 score of 0.76 for the CM phase SVM model. Other CT phases yielded inferior results for both classification tasks.

Conclusion CTTA employing open-source Python tools is a viable tool for differentiating ccRCCs from non-ccRCCs and predicting ccRCC grade.

Introduction

Renal cell carcinoma (RCC) is a prevalent adult cancer characterized by multiple histological subtypes, with clear cell RCC (ccRCC) being the most common and aggressive form.[1–3] Accurate subtyping and grading of RCCs are vital for prognosis and individualized treatment, including novel therapies like tyrosine kinase and vascular endothelial growth factor inhibitors.[4–7] However, current diagnostic methods such as renal mass biopsy are invasive and often unreliable, while imaging techniques like multiphasic computed tomography (CT) and magnetic resonance imaging (MRI) suffer from high subjectivity and overlapping findings.[8–13]

Recent advancements in texture analysis offer a noninvasive and objective method to analyze tumor heterogeneity through CT texture analysis (CTTA), which has shown potential in predicting RCC subtypes and grades.[14–20] Despite promising results, the clinical utility of CTTA is hampered by variations in imaging protocols and software across studies, underscoring the need for a robust, generalizable model using open-source tools.[21,22] Machine learning (ML), a powerful branch of artificial intelligence, can significantly enhance CTTA by enabling the extraction and analysis of complex, high-dimensional data patterns not readily apparent to human observers.[23] This capability potentially allows for the detection of subtle diagnostic markers that improve the accuracy of RCC subtyping and grading.[18,19]

Therefore, in this study, we utilized open-source Python libraries to conduct CTTA for subtyping RCC and predicting its grade. We independently analyzed all available CT contrast phases—corticomedullary, nephrographic, and noncontrast images, as well as virtual noncontrast images obtained from dual-energy CT scans. Subsequently, we developed ML models based on different architectures using diverse texture parameters extracted from each of these contrast phases.

Methods

CT dataset retrieval: The institutional review board gave its clearance to conduct this retrospective observational study. Between January 2016 and November 2021, we searched our hospital's pathology report database for histopathological nephrectomy specimens with RCC. To categorize RCCs into ccRCCs and non-ccRCCs and to further grade ccRCCs using the Fuhrman classification system, these histopathological reports were used as reference standards. Abdominal CT images of these patients were collected from our hospital's picture archiving and communication system (PACS) using their hospital identification numbers. The study excluded patients who did not have a baseline preoperative scan in the PACS. In total, 144 RCC patients who had a preoperative abdomen CT performed at our hospital and subsequently underwent surgery were included in the analysis. Of the 144 RCCs that were included in the study, 36 (including 30 papillary and 6 chromophobe RCCs) were non-ccRCCs and 108 were ccRCCs. Seventy-one of the 108 ccRCCs were classified as low grade (Fuhrman grades I–II), whereas 19 were classified as high grade (Fuhrman grades III–IV). Eighteen ccRCCs were not graded in the pathology reports and were therefore not included in the analysis of the Fuhrman grade correlation. The inclusion and exclusion of the CT scans and the final number of analyzed CT scans in different phases are summarized in ►**Fig. 1**.

CT protocol: Using a dual-source dual-energy 2 × 128 section multidetector CT scanner (Somatom Definition Flash, Siemens Healthineers, Germany), multiphase CT imaging was carried out. A noncontrast CT (NCCT) scan is the first step in the current standardized renal mass protocol used in our department. Corticomedullary (CM) and nephrographic (NP) phase images are then obtained at 25 to 30 and 60 to 70 seconds, respectively, following intravenous injection of 100 mL of iodinated contrast (Omnipaque 350, GE Healthcare) through a peripheral line at 3 to - mL/s. Delayed excretory phase images are obtained after 4 to 5 minutes in patients with suspicion of renal collecting system involvement. However, the same protocol was not followed for all RCC patients over all the previous years for which the CT scans were collected. Thus, for each patient, all the available phases of the CT were used for the study except the delayed phase images due to the availability of only a small number of scans with this phase (only for 10 of the 144 patients). Virtual noncontrast (VNC) images were obtained from the dual-energy scan acquisition wherever available. The retrieved CT scans were made anonymous by removing the Digital Imaging and Communications in Medicine (DICOM) metadata and reassigning them with a new unique identity (ID) number for the study.

Computed Tomography Texture Analysis Workflow

- *Manual tumor annotation:* To manually annotate tumors, we extracted the biggest tumor cross-section from each of the available CT phase images and saved it as a separate Joint Photographic Experts Group (JPEG) file for each patient. A single radiologist annotated the tumor on these

 © 2024. Indian Radiological Association. All rights reserved.

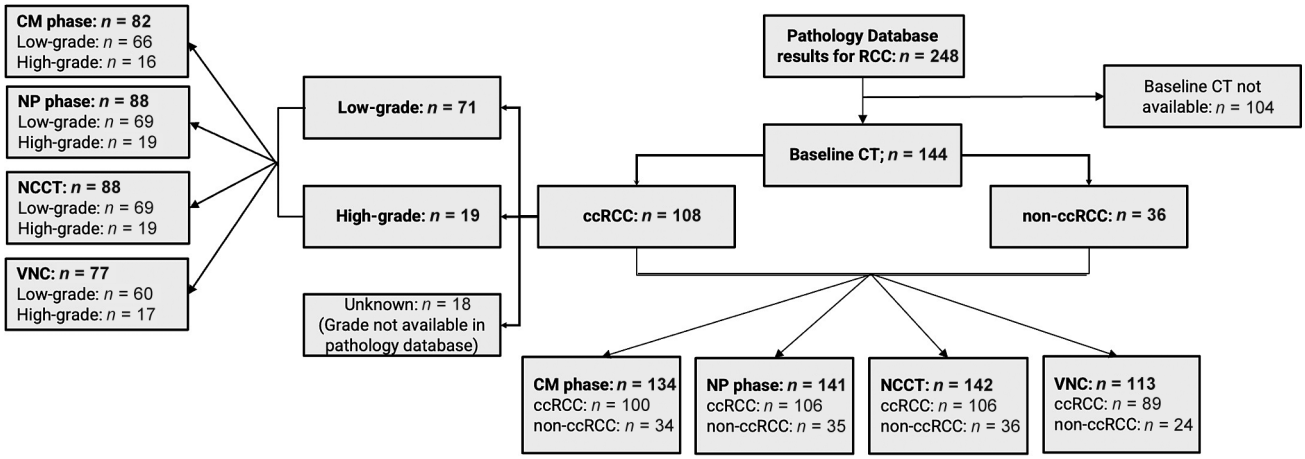

Fig. 1 Inclusion and exclusion of computed tomography scans in the study. ccRCC, clear cell renal cell carcinoma; CM, corticomedullary; CT, computed tomography; NCCT, noncontrast computed tomography; NP, nephrographic; VNC, virtual noncontrast.

CT images (A.G., with 9 years of experience in body imaging). Using the freely accessible user-friendly "LabelMe" software, tumor margins were manually marked in each exported JPEG file using the "Polygonal ROI" (region of interest) tool.[24] This annotation was done for all the available CT phases for each subject except for the delayed phase. For each case, the manual annotation was done only for the surgically resected primary renal tumor postimaging with available corresponding pathology report, regardless of the bilaterality or presence of tumor metastases. In the cases with more than one resected primary tumors, the largest measuring tumor was chosen for annotation. Next, for each image, the annotated data were saved as a JavaScript Object Notation (JSON) file. The associated label, the related instance of the tumor, the convex hull of the annotated tumor (in terms of the location of the tumor's boundary pixels in the picture), and other details about the image (like a slice of the CT scan) were all recorded in the JSON file. Based on these data, further features were extracted, and the tumor's features were analyzed.

- *Image preprocessing:* The CT images were prepared for subsequent texture analysis with adherence to standard practices in medical imaging by following the steps:
  1. *Creating a blank canvas:* A base image was created to match the dimensions and resolution specified in the JSON file, which contained relevant data about the CT images and tumor annotations.
  2. *Annotation image development:* Using the convex hull annotation, the tumor region was demarcated on the CT images. Pixels within the tumor were assigned a value of 255, while surrounding areas were set to 0, clearly isolating the tumor.
  3. *Image processing with OpenCV*[25]*:* The CT slice image under consideration (IUC) was processed using the OpenCV library, converting it into an 8-bit unsigned integer format for computational analysis.
  4. *Normalization:* The IUC was normalized to match the specifications in the JSON file, ensuring consistency in image dimensions and quality.
  5. *Mask image creation:* A bitwise AND operation was performed to combine the annotation image with the IUC, resulting in a "mask image" that exclusively contained tumor pixels.
  6. *Classification:* The mask image was categorized as either ccRCC or non-ccRCC based on the classifications in the JSON file. Only the mask image and its label were used for further analysis.

These steps allowed for precise isolation of the tumor regions, enabling targeted analysis of the tumor characteristics.

- *CT-based texture analysis:* ►**Fig. 2** visualizes the representative edge map, histogram of oriented gradients (HOG), and heat map of various contrast phases of a single CT slice of a patient with RCC. These visualizations highlight the unique changes in the appearance of these features, emphasizing the potential of texture parameters as distinct markers for CTTA.
  - *Analyzing texture using python libraries*: We employed powerful Python libraries—*scipy*, *numpy*, and *opencv*—to calculate specific texture parameters of the tumor. These libraries are sophisticated digital tools that help us extract meaningful patterns from the images.
  - *Processing the "mask image":* The "mask image," which contains only the tumor area, was read using the *opencv* library and analyzed as a matrix (a grid of numbers). This approach allows us to perform detailed mathematical calculations on the image.
  - *Calculating the first- and second-order texture features:* The first-order features, such as mean, variance, and standard deviation, were computed on this image using the *numpy* library. The second-order features such as energy, entropy, and smoothness were calculated using the *scipy* library.
- *Statistical analysis:* The generated texture parameters were grouped based on their corresponding labels (ccRCC vs. non-ccRCC). The entries of ccRCC were further

 © 2024. Indian Radiological Association. All rights reserved.

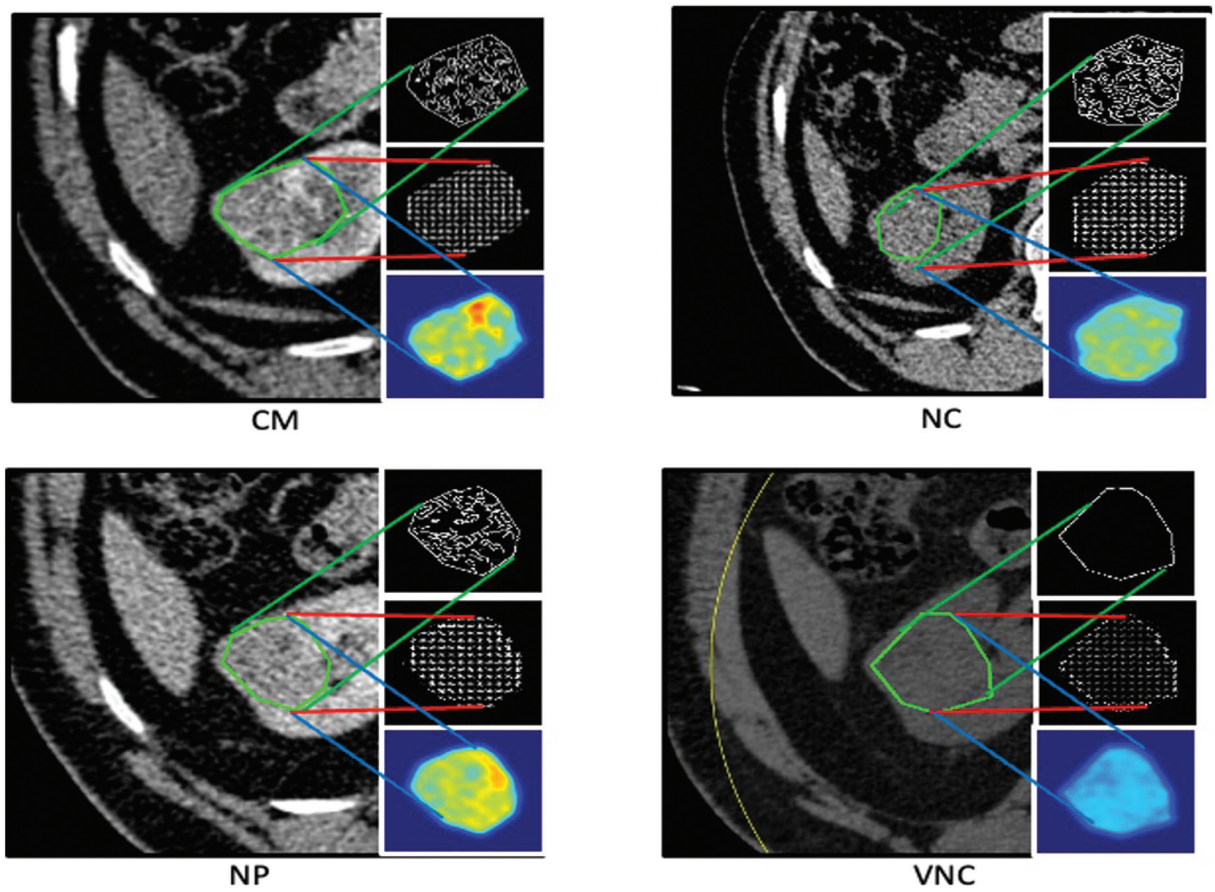

Fig. 2 The edge map, histogram of oriented gradients, and heat map of the annotated tumor on various contrast phases of a single computed tomography slice. CM, corticomedullary; NC, noncontrast; NP, nephrographic; VNC, virtual noncontrast.

divided into two subgroups, namely low- and high-grade ccRCCs, wherever available. These parameters were then analyzed using multivariable logistic regression analysis computed using the standard *matplotlib*, *seaborn*, and *scipy* libraries.[26,27] Various statistical measures like mean, median, mode, interquartile deviation, and standard deviation were computed for each of the texture parameters, using the *scipy* library. The corresponding *p*-values and F1 scores were recorded separately for both the classification tasks. F1 score refers to the harmonic mean of precision (the proportion of true positive results among all positive calls) and recall (the proportion of true positive results among all actual positives), providing a single score to gauge a test's accuracy in correctly identifying true cases, while minimizing false positives and false negatives. The *seaborn* and *matplotlib* libraries were used to plot the receiver operating characteristic (ROC) curves for the texture parameters for the classification of ccRCC versus non-ccRCCs as well as high- versus low-grade ccRCCs.

Machine Learning Models

We trained and compared various ML models, including support vector machine (SVM), random forest (RF), Naive Bayes (NB), Adaboost (AB), and logistic regression (LR), for RCC subtyping (ccRCC vs. non-ccRCC) and grade prediction (low-grade vs. high-grade) using texture parameters derived from each CT phase separately.[28]

- *Normalization and feature scaling*: Each texture parameter was normalized between 0 and 1 using a technique called min-max normalization. This step is crucial as it puts all parameters on an even scale, allowing each one to contribute equally to the model, regardless of their original scale or units.
- *Training of the model*: In constructing our ML models, we did not add any prior model. The prediction label for each of the samples in the test dataset was considered as another feature. We used a 70:20:10 split for training, validation, and test datasets, respectively. To address the issue of class imbalance, we employed a weighted loss function, with weights inversely proportional to the number of samples in each class, ensuring better representation of the minority class during training. We performed leave-one-out cross-validation (LOO-CV) within the training and validation sets. Specifically, for each LOO-CV iteration, we used four samples for validation (3 ccRCC and 1 non-ccRCC) and trained the model on the remaining samples. This process was repeated until every sample was used for validation at least once. We report the mean validation performance by averaging the outputs from each training iteration. The training process was

 © 2024. Indian Radiological Association. All rights reserved.

conducted using a standard ML library called *sklearn*, which provides tools for building and deploying ML models.

- *Statistical analysis*: The statistical analysis was repeated on the trained ML models, similar to each of the individual texture parameters. This repetition of analysis is an important step to validate and compare the performance of our ML models among the different architectures and with the initial texture parameter analysis. The study workflow is summarized in ►**Fig. 3**.

External validation: to assess the generalizability of our best performing ML model, we conducted an external validation. For this purpose, we utilized the public dataset containing CT scans from the 2019 Kidney and Kidney Tumor Segmentation Challenge (KiTS19).[29] The dataset also provides RCC histopathology type based on surgical specimens. To assess the ML models, a total of 40 cases each for ccRCC and non-ccRCC subtypes were randomly selected from the dataset. The CM phase CT scans were available for all of these cases, with NCCT scans available only in a small minority. Thus, we conducted the model testing and analysis using only the CM phase images in the public dataset. The tumor texture features were extracted using the same feature extraction technique as applied to the in-house dataset. The trained models were then tested for RCC subtyping (ccRCC vs. non-ccRCC) using these texture features from this external dataset. Since the dataset did not provide information about the pathological grading of ccRCC cases, the external validation of the models was performed only for classification of ccRCC versus non-ccRCC and not grade prediction of ccRCC.

Results

Patient demography and tumor characteristics: The mean age of the 108 people who had ccRCCs was 67.1 years (range: 51–72 years), and the average size of their cancer was 6.8 cm (standard deviation: ± 2.6 cm). There were 61 men and 47 women in the group. The average patient age for non-ccRCCs ($n = 36$) was 64.9 years (range: 48–69 years), with a male-to- female ratio of 19:17 and an average tumor size of 6.6 cm (standard deviation: ± 2.8 cm).

RCC Subtyping (ccRCC vs. Non-ccRCC)

- *Available CT phases:* For the included 144 patients (108 ccRCCs and 36 non-ccRCCs), various CT phases available were the following: NCCT ($n = 142$; 106 ccRCCs and 36 non-ccRCCs); CM phase ($n = 134$; 100 ccRCCs and 34 non-ccRCCs); NP phase ($n = 141$; 106 ccRCCs and 35 non-ccRCCs); and VNC ($n = 113$; 89 ccRCCs and 24 non-ccRCCs).
- *Texture parameters on each CT phase:* Entropy of the CM phase was observed to be the best performing individual parameter in classifying the ccRCC against non-ccRCC with an F1 score of 0.83, area under the curve (AUC) of 0.74, sensitivity of 0.81, and specificity of 0.95. Other individual texture parameters showed modest performance in classifying the ccRCC against non-ccRCC, with F1 scores ranging between 0.49 and 0.71 and AUC ranging between 0.43 and 0.68. The performances of various CTTA parameters derived from different CT phases for classification of ccRCC vs non-ccRCC are summarized in ►**Table 1**.
- *Performance of ML models based on each CT phase:* The various ML model architectures (SVM, RF, NB, AB, LR) considering all the texture features of the CM phase showed significant improvement in classifying the ccRCC against non-ccRCC with the best F1 score of 0.87, AUC of 0.79, sensitivity of 0.87, and specificity of 1.00, for the SVM model. The ML model architectures considering texture features of other phases showed modest performance with F1 scores ranging between 0.52 and 0.79 and

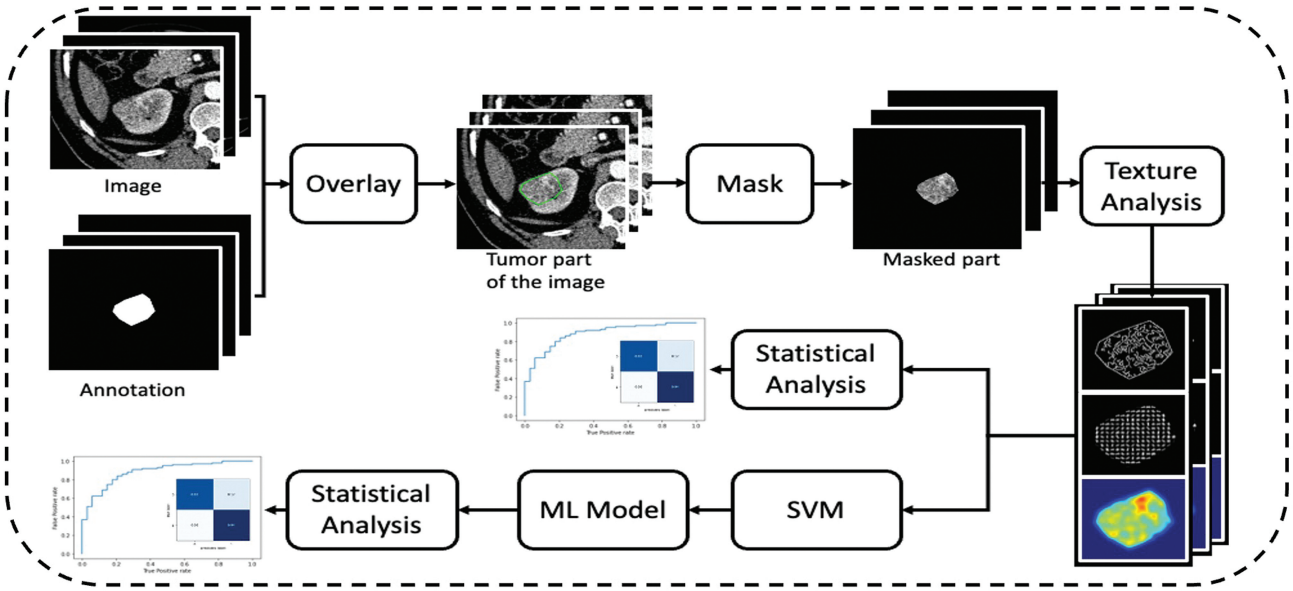

Fig. 3 Study workflow. (**A**) Manual annotation of the tumor by the radiologist. (**B**) Image preprocessing followed by texture analysis. ML, machine learning; SVM, support vector machine.

 © 2024. Indian Radiological Association. All rights reserved.

Table 1 Performance of CTTA parameters derived from various CT phases for RCC subtyping (ccRCC vs. non-ccRCC)

Phase	Feature	ccRCC		Non-ccRCC		*p*-Value	F1 score
		Median	Standard deviation	Median	Standard deviation		
CM phase ($n = 134$)	Mean	159.74	24.6	137.31	27.77	0.0006	0.71
	Variance	1,146.39	642.88	358.24	435.4	<0.0005	0.71
	Energy	1.61	1.34	1.41	2.09	0.3506	0.67
	Entropy	6.89	0.36	6.19	0.37	<0.0005	0.83
	Smoothness	0.99	0.02	0.98	0.03	0.0038	0.6
NP phase ($n = 141$)	Mean	156.83	20.4	145.51	21.36	0.0005	0.53
	Variance	623.5	387.1	381.93	308.07	0.0008	0.66
	Energy	1.51	1.29	1.69	1.93	0.2272	0.63
	Entropy	6.58	0.37	6.25	0.4	0.0012	0.59
	Smoothness	0.98	0.03	0.98	0.03	0.3132	0.53
NCCT ($n = 142$)	Mean	79.4	14.35	92.51	19.6	0.0098	0.44
	Variance	41.74	167.42	45.24	107.16	0.9783	0.65
	Energy	0.63	0.68	0.67	1.68	0.0178	0.66
	Entropy	4.71	0.57	4.76	0.65	0.4184	0.66
	Smoothness	0.97	0.03	0.97	0.03	0.7539	0.67
VNC ($n = 113$)	Mean	78.61	4.16	78.2	2.61	0.6423	0.63
	Variance	38.13	106.03	33.22	14.17	0.4465	0.59
	Energy	0.56	0.44	0.45	0.48	0.9884	0.49
	Entropy	4.67	0.24	4.55	0.23	0.0642	0.71
	Smoothness	0.97	0.03	0.97	0.02	0.9859	0.65

Abbreviations: ccRCC, clear cell renal cell carcinoma; CM, corticomedullary; CT, computed tomography; CTTA, computed tomography texture analysis; NCCT, noncontrast computed tomography; NP, nephrographic; RCC, renal cell carcinoma; VNC, virtual noncontrast.

AUC scores ranging between 0.49 and 0.71. The performances of various ML architectures based on texture parameters from different CT phases for classification of ccRCC versus non-ccRCC are summarized in ►**Table 2**. The ROC curve and confusion matrix for the "entropy" of the CM phase and the best performing SVM model generated using the texture features of the CM phase for classification of ccRCC versus ccRCC are shown in ►**Fig. 4**.

- *External validation results for RCC subtyping:* The best performing ML model on the internal dataset, that is, the SVM model on the CM phase, was selected for testing on the public dataset. The model achieved an accuracy of 0.82, F1 score of 0.81, *sensitivity* of 0.81, and specificity of 0.82 on the external dataset for classification of ccRCC versus non-ccRCC.

Fuhrman Grade Prediction for ccRCCs (Low vs. High Grade)

- *Available CT phases:* For the total of 90 RCCs (71 low grade and 19 high grade) with available Fuhrman grading, various CT phases available were the following: NCCT ($n = 88$; 69 low grade and 19 high grade); CM phase ($n = 82$; 66 low *grade* and 16 high grade); NP phase ($n = 88$; 69 low grade and 19 high grade); and VNC ($n = 77$; 60 low grade and 17 high grade).
- *Texture parameters on each CT phase:* The entropy of the CM phase was observed to be the best performing individual parameter in classifying the low-grade ccRCC against high-grade ccRCC with a modest F1 score of 0.67, AUC of 0.62, sensitivity of 0.71, and specificity of 0.65. Other individual texture parameters showed poor performance in classifying the low-grade ccRCC against high-grade ccRCC, with F1 scores ranging between 0.41 and 0.59 and AUC ranging between 0.38 and 0.53. The performances of various CTTA parameters derived from different CT phases for classification of low- versus high-grade ccRCC are summarized in ►**Table 3**.
- *Performance of ML models based on each CT phase:* The SVM model considering all the texture features of the CM phase showed some *improvement* in classifying the high-grade ccRCC against low-grade ccRCC with an F1 score of 0.76, AUC of 0.65, sensitivity of 0.85, and specificity of 0.69. The SVM models considering texture features of other phases showed poor performance, with F1 scores ranging between 0.6 and 0.64 and AUC scores ranging between 0.39 and 0.61. The performances of various SVM-based models based on texture parameters from different CT phases for classification of high- versus low-grade ccRCC are summarized in ►**Table 4**. The ROC curve and confusion matrix for the "entropy" of the CM phase and ML model generated using the texture features of the CM

 © 2024. Indian Radiological Association. All rights reserved.

Table 2 Performance of various ML model architectures based on texture parameters from various CT phases for RCC subtyping (ccRCC vs. non-ccRCC)

	Phase	ccRCC		Non-ccRCC		*p*-Value	F1 score
		Median	Standard deviation	Median	Standard deviation		
SVM	CM phase ($n = 134$)	0.92	0.17	0.42	0.27	<0.0005	0.87
	NP phase ($n = 141$)	0.82	0.12	0.66	0.17	<0.0005	0.79
	NCCT ($n = 142$)	0.78	0.09	0.74	0.16	<0.0005	0.66
	VNC ($n = 113$)	0.79	0.06	0.76	0.07	0.05116	0.52
Random forest	CM phase ($n = 134$)	0.13	0.14	0.60	0.26	<0.0005	0.81
	NP phase ($n = 141$)	0.19	0.09	0.41	0.19	<0.0005	0.76
	NCCT ($n = 142$)	0.21	0.06	0.35	0.14	<0.0005	0.70
	VNC ($n = 113$)	0.18	0.04	0.31	0.12	<0.0005	0.64
Naive Bayes	CM phase ($n = 134$)	0.84	0.28	0.26	0.36	<0.0005	0.80
	NP phase ($n = 141$)	0.75	0.27	0.48	0.33	<0.0005	0.76
	NCCT ($n = 142$)	0.87	0.23	0.74	0.35	0.011	0.67
	VNC ($n = 113$)	0.36	0.31	0.27	0.26	0.216	0.53
Adaboost	CM phase ($n = 134$)	1.0	0.32	0.0	0.416	<0.005	0.79
	NP phase ($n = 141$)	1.0	0.42	0.0	0.300	<0.005	0.70
	NCCT ($n = 142$)	0.47	0.49	0.0	0.09	<0.005	0.64
	VNC ($n = 113$)	1.0	0.49	0.2	0.32	<0.005	0.60
Logistic regression	CM phase ($n = 134$)	1.0	0.492	0.0	0.237	<0.005	0.77
	NP phase ($n = 141$)	0.0	0.400	0.0	0.13	0.109	0.73
	NCCT ($n = 142$)	0.13	0.34	0.02	0.165	0.012	0.69
	VNC ($n = 113$)	0.89	0.500	0.200	0.300	<0.005	0.64

Abbreviations: ccRCC, clear cell renal cell carcinoma; CM, corticomedullary; CT, computed tomography; ML, machine learning; NCCT, noncontrast computed tomography; NP, nephrographic; VNC, virtual noncontrast.

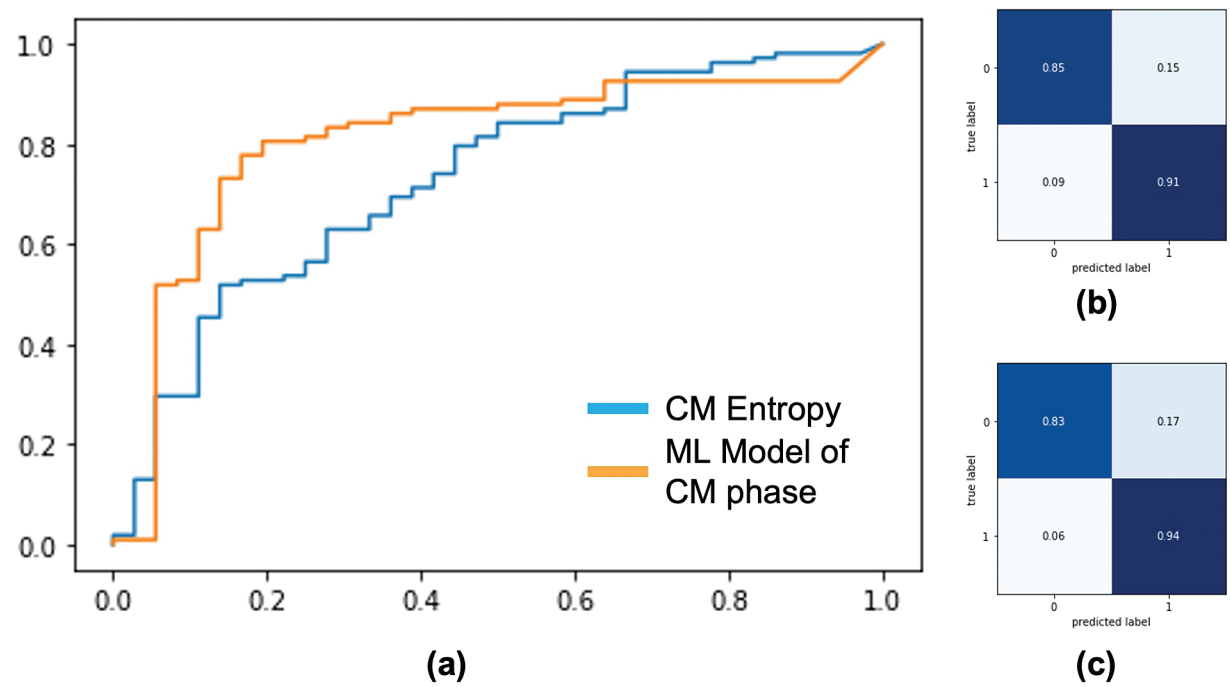

Fig. 4 Performance of the entropy of the corticomedullary (CM) phase and machine learning (ML) model generated using all the texture parameters of the CM phase to classify clear cell renal cell carcinoma (ccRCC) against non-ccRCC. (**A**) Receiver operating characteristic curve. (**B**) Confusion matrix of the ML model. (**C**) Confusion matrix of the entropy of the CM phase.

 © 2024. Indian Radiological Association. All rights reserved.

Table 3 Performance of CTTA parameters derived from various CT phases for ccRCC Fuhrman grade prediction (low grade vs. high grade)

Phase	Feature	High-grade ccRCC		Low-grade ccRCC		*p*-Value	F1 score
		Median	Standard deviation	Median	Standard deviation		
CM phase ($n = 82$)	Mean	161.98	19.34	163.03	25.6	0.64	0.58
	Variance	976	610.83	1,192.8	648.18	0.2	0.56
	Energy	2.67	1.42	1.34	1.24	0	0.49
	Entropy	6.84	0.34	6.9	0.33	0.22	0.67
	Smoothness	0.99	0.03	0.99	0.02	0.7	0.41
NP phase ($n = 88$)	Mean	149.73	21.38	159.27	20.48	0.53	0.5
	Variance	451.06	415.71	685.07	379.66	0.2	0.55
	Energy	2.11	1.05	1.36	1.32	0.18	0.54
	Entropy	6.42	0.38	6.66	0.38	0.09	0.6
	Smoothness	0.99	0.02	0.98	0.03	0.47	0.59
NCCT ($n = 88$)	Mean	81.4	10.59	78.61	13.27	0.79	0.52
	Variance	38.92	30.66	42.36	115.4	0.24	0.56
	Energy	0.76	0.45	0.5	0.69	0.18	0.58
	Entropy	4.68	0.34	4.74	0.55	0.28	0.57
	Smoothness	0.97	0.02	0.96	0.04	0.28	0.52
VNC ($n = 77$)	Mean	81.31	3.44	77.83	4.37	0.05	0.52
	Variance	38.65	17.21	38.2	13.72	0.94	0.51
	Energy	0.71	0.41	0.48	0.4	0.01	0.55
	Entropy	4.68	0.27	4.67	0.22	0.9	0.54
	Smoothness	0.97	0.02	0.97	0.04	0.32	0.51

Abbreviations: ccRCC, clear cell renal cell carcinoma; CM, corticomedullary; CT, computed tomography; CTTA, computed tomography texture analysis; NCCT, noncontrast computed tomography; NP, nephrographic; SVM, support vector machine; VNC, virtual noncontrast.

phase for the classification of high- ccRCC versus low-grade ccRCC are shown in ►**Fig. 5**. However, we were unable to successfully train other ML models (RF, NB, AB, LR), under the same conditions for grade prediction of ccRCC. These models likely require a larger amount of data to accurately capture patterns and prevent overfitting, making them less suitable for small datasets in this context.

Discussion

The present study shows that texture parameters generated from CTTA using open-source Python modules perform well for both RCC subtyping and grading of ccRCC, with the entropy values on CM phase CT images achieving the highest performance. In addition, ML-based prediction models based on the CTTA parameters further enhance the results for both the classification tasks.

Entropy in the CM phase, reflecting pixel irregularity and texture complexity, was the most distinguishing feature for ccRCC subtype and grade prediction.[14] These results correspond with the standard technique of reading contrast-enhanced CT for RCC characterization—where increased tumor heterogeneity is regarded to favor a diagnosis of ccRCC and point toward a higher-grade cancer. However, qualitative evaluations of

Table 4 Performance of SVM ML models based on texture parameters from various CT phases for ccRCC Fuhrman grade prediction (low grade vs. high grade)

Phase	High-grade ccRCC		Low-grade ccRCC		*p*-Value	F1 score
	Median	Standard deviation	Median	Standard deviation		
CM phase ($n = 82$)	0.77	0.23	0.87	0.13	0.03	0.76
NP phase ($n = 88$)	0.74	0.15	0.84	0.12	0.05	0.64
NCCT ($n = 88$)	0.73	0.11	0.79	0.07	0.01	0.6
VNC ($n = 77$)	0.78	0.07	0.79	0.05	0.15	0.61

Abbreviations: ccRCC, clear cell renal cell carcinoma; CM, corticomedullary; CT, computed tomography; ML, machine learning; NCCT, noncontrast computed tomography; NP, nephrographic; SVM, support vector machine; VNC, virtual noncontrast.

 © 2024. Indian Radiological Association. All rights reserved.

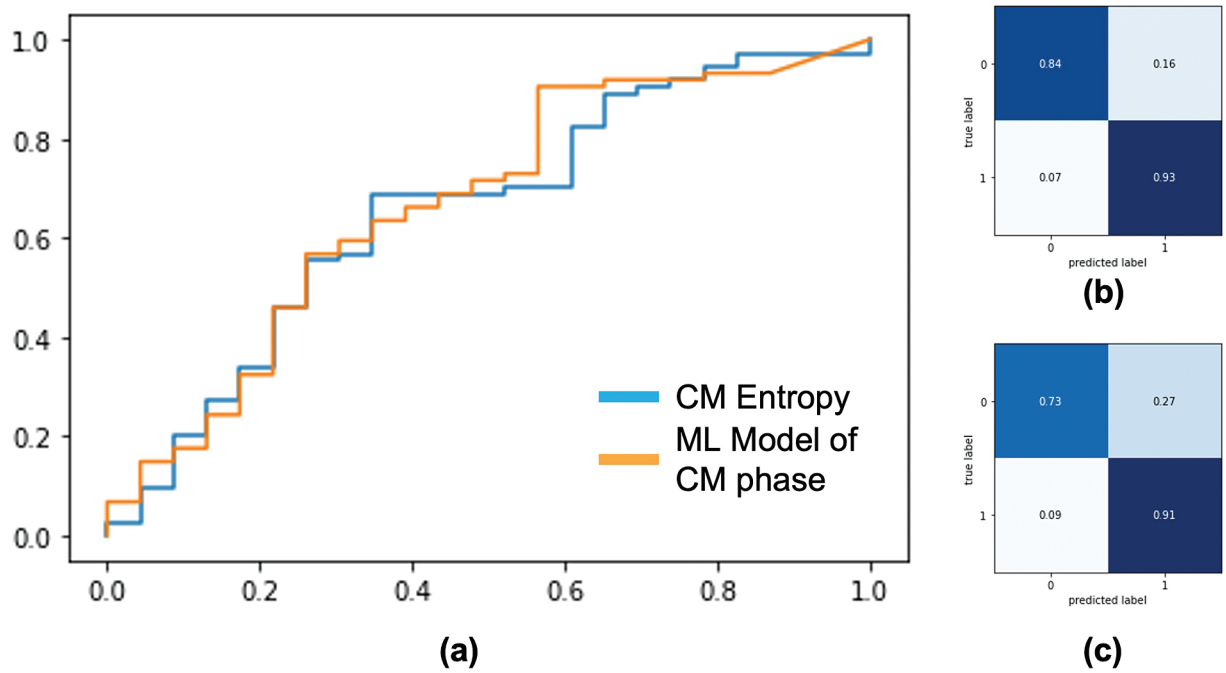

Fig. 5 Performance of the entropy of the corticomedullary (CM) phase and machine learning (ML) model generated using all the texture parameters of the CM phase to classify high-grade clear cell renal cell carcinoma (ccRCC) against low-grade ccRCC. (**a**) Receiver operating characteristic curve. (**b**) Confusion matrix of the ML model. (**c**) Confusion matrix of the entropy of the CM phase.

heterogeneity are extremely individualized and vulnerable to biases.[10,12] Therefore, objective measures such as entropy can help in the consistent and accurate evaluation of cancer complexity, leading to a better characterization of RCC. The potential of entropy values being able to classify RCC subtypes as well as ccRCC grades in the same patient cohort is a significant finding from our study. Our results are in line with the previous studies investigating the use of CTTA in RCC characterization.[15,17,30] However, a major point to note is that most of these studies used either specific software (like MaZda and pyRadiomics) or in-house algorithms for the extraction of texture parameters, and only a few compared the different CT phases for CTTA. As noted by Doshi et al, there is a significant variation in the way different software packages compute and process image data, and this limits the generalizability and clinical reliability of texture parameters.[22] However, unlike most studies that use proprietary software, we employed open-source Python libraries, ensuring greater cost-effectiveness and reproducibility. Additonally, we did not use third-order texture features, as first- and second-order features provided sufficient clinically relevant insights, while being less computationally intensive and more interpretable.[21]

Of the various ML model architectures tested in this study, the SVM approach exhibited the best performance in both the classification tasks (i.e., ccRCC vs. non-ccRCC and high- vs. low-grade ccRCC). Budai et al reported the best performance of their ML model on the CM phase for differentiation between ccRCC and non-ccRCC, with an AUC of 0.87 for the internal test set and AUC of 0.83 on the external validation set.[18] Our ML prediction model also fared the best on the CM phase with an AUC of 0.79. There was, however, only average success with ML-based grade prediction for ccRCC, with an AUC of 0.65.

CTTA using open-source tools offers a cost-effective and scalable solution for improving RCC characterization accuracy. By providing objective measures like entropy and ML models, CTTA can help radiologists reduce subjective biases, particularly in RCC subtyping and ccRCC grading, complementing qualitative evaluations for more consistent diagnoses and better treatment planning. Open-source Python tools ensure adaptability, transparency, and easier integration into clinical practice. The SVM model achieved an accuracy of 0.82 and an F1 score of 0.81 on external data, closely matching the in-house F1 score of 0.87. This minor variation highlights the model's robustness and validates its applicability in real-world settings with diverse data. Our study has a few drawbacks. There was highly variable availability of postcontrast phases for patients due to the study's retrospective nature and collection of scans over a long period. Thus, analysis of combination of texture features from various CT phases could not be performed. Only one observer performed the tumor annotation; thus, interobserver agreement could not be derived. Interobserver variability could impact the reproducibility and reliability of the tumor annotations, potentially introducing bias in the texture analysis results. This limitation may affect the generalizability of the findings, as consistency across different observers is crucial for broader clinical applicability. A single axial tumor slice was used for analysis instead of the three-dimensional (3D) volumetric approach due to time and labor constraints. Exclusively operated RCC cases were analyzed, which may have caused a selection bias toward larger and higher-grade tumors. Additionally, the small number of ccRCC cases with Fuhrman grade data restricted the sample size, limiting ML model exploration beyond SVM for grade prediction and external validation. Future studies should aim for multicenter

 © 2024. Indian Radiological Association. All rights reserved.

data collection, use a 3D volumetric approach based on automated segmentation tools, and include multiple observers to improve generalizability and model performance.

Conclusion

CTTA employing open-source Python tools is practical and offers promising results for differentiating ccRCCs from non-ccRCCs and for grade prediction of ccRCCs, with the greatest results being seen with CM phase CT acquisition. Furthermore, we created a ML model using common SVM algorithms that performed even better for CTTA-based RCC characterization, underscoring the value of integrating advanced computational techniques in medical imaging analysis.

Ethical Approval

Ethical approval was obtained from the institutional review board. The need for patient consent was waived off by the institutional review board due to the retrospective nature of the analysis.

Funding

None.

Conflict of Interest

None declared.

References

1 Moch H, Cubilla AL, Humphrey PA, Reuter VE, Ulbright TM. The 2016 WHO classification of tumours of the urinary system and male genital organs-part A: renal, penile, and testicular tumours. Eur Urol 2016;70(01):93–105
2 Capitanio U, Cloutier V, Zini L, et al. A critical assessment of the prognostic value of clear cell, papillary and chromophobe histological subtypes in renal cell carcinoma: a population-based study. BJU Int 2009;103(11):1496–1500
3 Patard JJ, Leray E, Rioux-Leclercq N, et al. Prognostic value of histologic subtypes in renal cell carcinoma: a multicenter experience. J Clin Oncol 2005;23(12):2763–2771
4 Lee CH, Motzer R. Combination VEGFR/immune checkpoint inhibitor therapy: a promising new treatment for renal cell carcinoma. Br J Cancer 2018;119(08):911–912
5 Atkins MB, Tannir NM. Current and emerging therapies for first-line treatment of metastatic clear cell renal cell carcinoma. Cancer Treat Rev 2018;70:127–137
6 Fuhrman SA, Lasky LC, Limas C. Prognostic significance of morphologic parameters in renal cell carcinoma. Am J Surg Pathol 1982;6(07):655–663
7 Ljungberg B, Albiges L, Abu-Ghanem Y, et al. European Association of Urology Guidelines on Renal Cell Carcinoma: the 2019 update. Eur Urol 2019;75(05):799–810
8 Abel EJ, Carrasco A, Culp SH, et al. Limitations of preoperative biopsy in patients with metastatic renal cell carcinoma: comparison to surgical pathology in 405 cases. BJU Int 2012;110(11):1742–1746
9 Abel EJ, Culp SH, Matin SF, et al. Percutaneous biopsy of primary tumor in metastatic renal cell carcinoma to predict high risk pathological features: comparison with nephrectomy assessment. J Urol 2010;184(05):1877–1881
10 Egbert ND, Caoili EM, Cohan RH, et al. Differentiation of papillary renal cell carcinoma subtypes on CT and MRI. AJR Am J Roentgenol 2013;201(02):347–355
11 Karlo CA, Di Paolo PL, Chaim J, et al. Radiogenomics of clear cell renal cell carcinoma: associations between CT imaging features and mutations. Radiology 2014;270(02):464–471
12 Young JR, Margolis D, Sauk S, Pantuck AJ, Sayre J, Raman SS. Clear cell renal cell carcinoma: discrimination from other renal cell carcinoma subtypes and oncocytoma at multiphasic multidetector CT. Radiology 2013;267(02):444–453
13 Wang P, Pei X, Yin XP, et al. Radiomics models based on enhanced computed tomography to distinguish clear cell from non-clear cell renal cell carcinomas. Sci Rep 2021;11(01):13729
14 Miles KA, Ganeshan B, Hayball MP. CT texture analysis using the filtration-histogram method: what do the measurements mean? Cancer Imaging 2013;13(03):400–406
15 Deng Y, Soule E, Samuel A, et al. CT texture analysis in the differentiation of major renal cell carcinoma subtypes and correlation with Fuhrman grade. Eur Radiol 2019;29(12):6922–6929
16 Ding J, Xing Z, Jiang Z, et al. CT-based radiomic model predicts high grade of clear cell renal cell carcinoma. Eur J Radiol 2018;103:51–56
17 Feng Z, Shen Q, Li Y, Hu Z. CT texture analysis: a potential tool for predicting the Fuhrman grade of clear-cell renal carcinoma. Cancer Imaging 2019;19(01):6
18 Budai BK, Stollmayer R, Rónaszéki AD, et al. Radiomics analysis of contrast-enhanced CT scans can distinguish between clear cell and non-clear cell renal cell carcinoma in different imaging protocols. Front Med (Lausanne) 2022;9:974485
19 Kocak B, Yardimci AH, Bektas CT, et al. Textural differences between renal cell carcinoma subtypes: Machine learning-based quantitative computed tomography texture analysis with independent external validation. Eur J Radiol 2018;107:149–157
20 Han D, Yu Y, Yu N, et al. Prediction models for clear cell renal cell carcinoma ISUP/WHO grade: comparison between CT radiomics and conventional contrast-enhanced CT. Br J Radiol 2020;93(1114):20200131
21 Lubner MG, Smith AD, Sandrasegaran K, Sahani DV, Pickhardt PJ. CT texture analysis: definitions, applications, biologic correlates, and challenges. Radiographics 2017;37(05):1483–1503
22 Doshi AM, Tong A, Davenport MS, et al. Assessment of renal cell carcinoma by texture analysis in clinical practice: a six-site, six-platform analysis of reliability. AJR Am J Roentgenol 2021;217(05):1132–1140
23 Rana M, Bhushan M. Machine learning and deep learning approach for medical image analysis: diagnosis to detection. Multimedia Tools Appl 2022:1–39
24 Russell BC, Torralba A, Murphy KP, Freeman WT. LabelMe: a database and web-based tool for image annotation. Int J Comput Vis 2008;77(01):157–173
25 OpenCV. Open Source Computer Vision Library. Accessed November 18, 2024 at: https://opencv.org/
26 Tosi S. Matplotlib for Python Developers. Packt; 2023. Accessed October 8, 2024 at: https://www.packtpub.com/product/matplotlib-for-python-developers/9781847197900
27 Waskom ML. seaborn: statistical data visualization. J Open Source Softw 2021;6(60):3021
28 Hearst MA, Dumais ST, Osman E, Platt J, Scholkopf B. Support vector machines. IEEE Intell Syst Their Appl 1998;13(04):18–28
29 Heller N, Sathianathen N, Kalapara A, et al. Data from C4KC-KiTS [Data set]. The Cancer Imaging Archive; 2019. Accessed October 8, 2024 at: https://doi.org/10.7937/TCIA.2019.IX49E8NX
30 Lubner MG, Stabo N, Abel EJ, Del Rio AM, Pickhardt PJ. CT textural analysis of large primary renal cell carcinomas: pretreatment tumor heterogeneity correlates with histologic findings and clinical outcomes. AJR Am J Roentgenol 2016;207(01):96–105

 © 2024. Indian Radiological Association. All rights reserved.