



# Noninvasive Fuhrman grading of clear cell renal cell carcinoma using computed tomography radiomic features and machine learning

Mostafa Nazari<sup>1</sup> · Isaac Shiri<sup>2</sup> · Ghasem Hajianfar<sup>3</sup> · Niki Oveisi<sup>4</sup> · Hamid Abdollahi<sup>5</sup> · Mohammad Reza Deevband<sup>1</sup> · Mehrdad Oveisi<sup>3,6</sup> · Habib Zaidi<sup>2,7,8,9</sup>

Received: 27 September 2019 / Accepted: 5 March 2020 / Published online: 19 March 2020  
© Italian Society of Medical Radiology 2020

## Abstract

**Purpose** To identify optimal classification methods for computed tomography (CT) radiomics-based preoperative prediction of clear cell renal cell carcinoma (ccRCC) grade.

**Materials and methods** Seventy-one ccRCC patients (31 low grade and 40 high grade) were included in this study. Tumors were manually segmented on CT images followed by the application of three image preprocessing techniques (Laplacian of Gaussian, wavelet filter, and discretization of the intensity values) on delineated tumor volumes. Overall, 2530 radiomics features (tumor shape and size, intensity statistics, and texture) were extracted from each segmented tumor volume. Univariate analysis was performed to assess the association between each feature and the histological condition. Multivariate analysis involved the use of machine learning (ML) algorithms and the following three feature selection algorithms: the least absolute shrinkage and selection operator, Student's *t* test, and minimum Redundancy Maximum Relevance. These selected features were then used to construct three classification models (SVM, random forest, and logistic regression) to discriminate high from low-grade ccRCC at nephrectomy. Lastly, multivariate model performance was evaluated on the bootstrapped validation cohort using the area under the receiver operating characteristic curve (AUC) metric.

**Results** The univariate analysis demonstrated that among the different image sets, 128 bin-discretized images have statistically significant different texture parameters with a mean AUC of  $0.74 \pm 3$  ( $q$  value  $< 0.05$ ). The three ML-based classifiers showed proficient discrimination between high and low-grade ccRCC. The AUC was 0.78 for logistic regression, 0.62 for random forest, and 0.83 for the SVM model, respectively.

**Conclusion** CT radiomic features can be considered as a useful and promising noninvasive methodology for preoperative evaluation of ccRCC Fuhrman grades.

**Keywords** Computed tomography (CT) · Radiomics · Machine learning · Renal cell carcinoma · Fuhrman grading

✉ Isaac Shiri  
Isaac.Shiri@etu.unige.ch

✉ Mohammad Reza Deevband  
mdeevband@sbmu.ac.ir

<sup>1</sup> Department of Biomedical Engineering and Medical Physics, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>2</sup> Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211 Geneva 4, Switzerland

<sup>3</sup> Rajaie Cardiovascular Medical and Research Center, Iran University of Medical Science, Tehran, Iran

<sup>4</sup> School of Population and Public Health, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada

<sup>5</sup> Department of Radiologic Sciences and Medical Physics, Faculty of Allied Medicine, Kerman University, Kerman, Iran

<sup>6</sup> Department of Computer Science, University of British Columbia, Vancouver, BC, Canada

<sup>7</sup> Geneva University Neurocenter, Geneva University, Geneva, Switzerland

<sup>8</sup> Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen, Netherlands

<sup>9</sup> Department of Nuclear Medicine, University of Southern Denmark, Odense, Denmark

## Introduction

Renal cell carcinoma (RCC) is the seventh most common cancer worldwide, with a mortality rate of 140,000 per year [1]. The most common types of renal cancer cells are clear cells RCC (ccRCC), papillary RCC (pRCC), and chromophobe RCC (chRCC) [2, 3]. Approximately 70% of kidney cancers are made up of ccRCC, pRCC accounts for 10–15% of kidney cancers, whereas chRCC is the least common type with only 5% of kidney cancer cases [4]. ccRCC has a survival rate of less than 5 years and a higher risk of metastasis compared to pRCC and chRCC [5].

One of the most important tasks in cancer diagnosis and treatment is tumor staging and grading. Tumor grading defines the differentiation of tumor tissue cells relative to normal tissue cells. It is an indicator of how quickly a tumor is expected to grow and spread. The Fuhrman grading system is widely recognized among the clinical oncology community [6]. It is based on the assessment of the following cell nucleus characteristics: nuclear size, nuclear shape, and nucleolar prominence. Based on these assessments, the tumor will be classified into one of four different grades (I–IV). Grades I and II are considered as low-grade tumors with a favorable prognosis, while grades III and IV account for high-grade tumors commonly having unfavorable prognosis [7].

Currently, fine-needle aspiration (FNA) and imaging-guided biopsies are the gold-standard methods for preoperative kidney tumor grading. However, these techniques have some drawbacks including infection, bleeding, tumor cells spreading, and provide limited information regarding the whole tumor due to tumor heterogeneity. Because of intra-tumoral heterogeneity in ccRCC [8], biopsy underestimated the Fuhrman grade in 55% of the cases [9].

A number of noninvasive therapeutic strategies for RCC have been devised during the last decade, including radiofrequency ablation, cryoablation, and active surveillance [10, 11]. However, a proper criterion for patient management using these noninvasive/minimally invasive treatment methods is still lacking, as most patients are often treated surgically post-diagnosis [12, 13]. Therefore, it is desirable to recommend individualized treatment strategies, where radical approaches (e.g., surgery) are kept only for aggressive or high-grade ccRCC tumors (III, IV), whereas conservative management (e.g., active surveillance) is applied for low-grade (I and II) lesions [14]. To guide the decision-making process, an accurate noninvasive method for preoperative Fuhrman grading of renal cell carcinoma tumors is desirable. To this end, two promising approaches have been adopted in clinical settings, namely MRI-guided derivation of apparent diffusion coefficient (ADC) values [15] and CT-based semi-quantitative and quantitative techniques [16, 17].

Radiomics serve as the bridge between medical imaging and personalized medicine [18] and refers to the comprehensive quantification of tumor phenotype to uncover disease characteristics that fail to be revealed by the naked eye [19–21]. In fact, radiomics is a new era of science which faces many challenges, including image acquisition [22], reconstruction and processing [23, 24], and model development to provide robust and reproducible representations. Previous studies have shown that the radiomics signature is valuable for differentiating high-/low-grade ccRCC tumors [25–27]. This study aims to construct a radiomics feature-based machine learning model to predict the Fuhrman grade of ccRCC patients preoperatively.

## Materials and methods

Figure 1 presents the workflow followed in the current study.

### Patient population

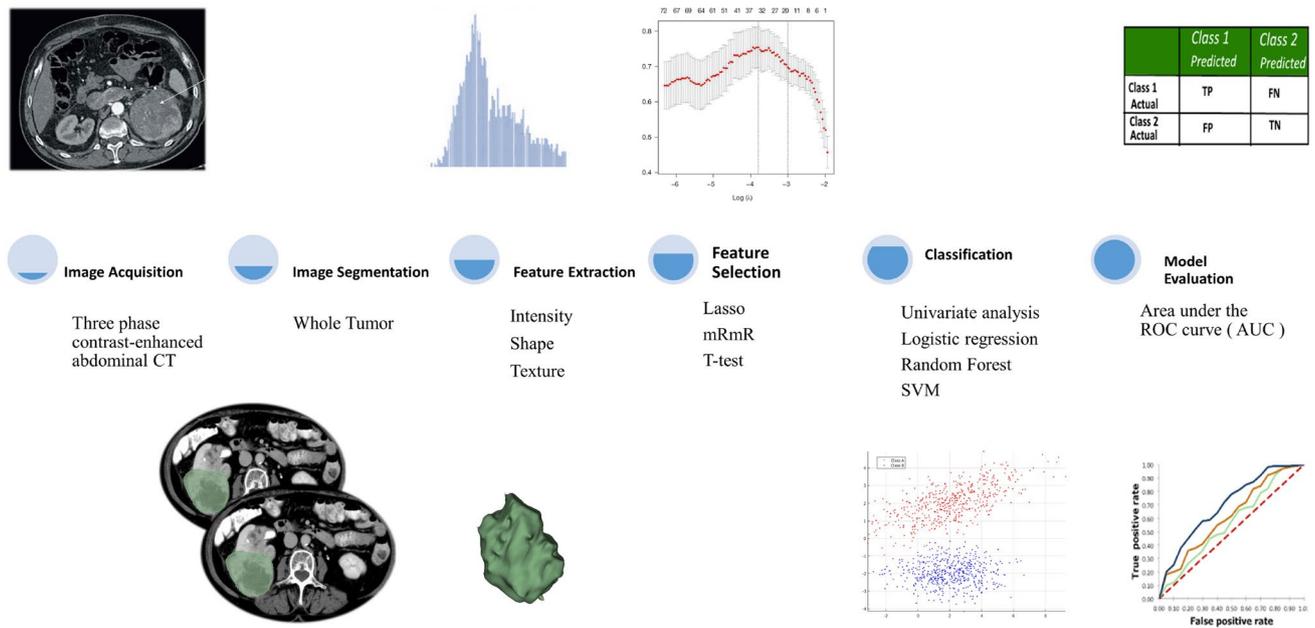
Two hundred and twenty-two clinical studies from the cancer image archive database [28] were included in the study protocol. The inclusion and exclusion criteria included simultaneous availability of CT images and pathological information, good CT image quality, and tumor visibility on CT images. This results in the inclusion of 71 patients in the current study. Table 1 provides the demographics of the patient population.

### CT scanning protocol

All patients underwent a three-phase CT scan, including (1) a routine unenhanced CT scan, (2) a corticomedullary phase (CMP) contrast-enhanced scan starting 40 s after injection of the contrast material, and (3) a nephrographic phase (NP) contrast-enhanced scan performed 70–90 s after intravenous injection of iodinated contrast material. The iodine content (300 mg/mL) was infused at an infusion rate of 3 mL/s at an infusion dose of 80–100 mL. All subjects were scanned on GE Healthcare and Siemens Healthineers CT scanners with a tube voltage of 120 kVp and a tube current of 150–300 mA using daily clinical reconstruction parameters.

### Tumor segmentation

In this study, manual volume of interest (VOI) segmentation was performed and verified by an experienced radiologist using the 3D slicer software package [29].



**Fig. 1** Workflow adopted in the study protocol

**Table 1** Clinical characteristics of the patient population presenting with clear cell renal cell carcinoma

Characteristics	
Number of patients	71
Male	51 (71%)
Female	20 (29%)
Age (years)	60.3 ± 11.7
Low grade (I, II)	31
High grade (III, IV)	40

## Image preprocessing

Prior to feature extraction, the voxel size resampling method was applied on the images to create an isotropic dataset. This allowed comparisons between image data from different samples and scanners [30]. Laplacian of Gaussian (LoG), wavelet decomposition (WAV), and discretization into 32, 64, and 128 bins preprocessing were performed to generate a different set of features. For the LoG filter, different sigma values were used to extract fine, medium, and coarse features. Wavelet filtering yields 8 decompositions per level: all possible combinations of applying either high- (H) or a low- (L)-pass filtering in each of the three dimensions, including HHH, HHL, HLH, HLL, LHH, LHL, LLH, and LLL. The preprocessing steps (including discretization, LOG, and wavelet) were also performed on all intensity, histogram, and textural features.

## Feature extraction

Radiomic features were extracted using the PyRadiomics open-source python library [31]. The extracted features were then categorized into the following subgroups. Firstly, shape features depict the shape of the tumor volume and geometric properties, such as volume, maximum surface, tumor compactness, and sphericity. Furthermore, first-order statistic features describe the distribution of voxel intensities within tumor volumes, including mean, median, maximum, and minimum values of the voxel intensities. Second- and higher-order statistic features (known as textural features) are used to measure inter-relationships between voxel distributions within tumor volumes, reflecting changes in image space gray levels. These features include gray-level co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM), gray-level size-zone matrix (GLSZM), and gray-level dependence matrix (GLDM) features. Table 2 provides detailed information about the extracted features.

## Univariate analysis

For univariate analysis, early correlation tests between features were used to eliminate highly correlated features. Student's *t* tests were then used for comparisons between groups. To control the false discovery rate (FDR) in multiple hypothesis testing, the Benjamini–Hochberg (FDR) correction method was applied on the resulting *p* values to ultimately report *q*-values [32].

**Table 2** Summary of radiomic features used in this work

First-order statistics (FOS)	Gray-level co-occurrence matrix (GLCM)	Gray-level run length matrix (GLRLM)
Energy	Autocorrelation	Short-run emphasis (SRE)
Total energy	Joint average	Long-run emphasis (LRE)
Entropy	Cluster prominence	Gray-level non-uniformity (GLN)
Minimum	Cluster shade	Gray-level non-uniformity normalized (GLNN)
10th percentile	Cluster tendency	Run length non-uniformity (RLN)
90th percentile	Contrast	Run length non-uniformity normalized (RLNN)
Maximum	Correlation	Run percentage (RP)
Mean	Difference average	Gray-level variance (GLV)
Median	Difference entropy	Run variance (RV)
Interquartile range	Difference variance	Run entropy (RE)
Range	Joint energy	Low gray-level run emphasis (LGLRE)
Mean absolute deviation (MAD)	Joint entropy	High gray-level run emphasis (HGLRE)
Robust mean absolute deviation (rMAD)	Informal measure of correlation (IMC) 1	Short-run low gray-level emphasis (SRLGLE)
Root-mean-squared (RMS)	Informal measure of correlation (IMC) 2	Short-run high gray-level emphasis (SRHGLE)
SD	Inverse difference moment (IDM)	Long-run low gray-level emphasis (LRLGLE)
Skewness	Inverse difference moment normalized (IDMN)	16. Long-run high gray-level emphasis (LRHGLE)
Kurtosis	Inverse difference (ID)	Gray-level dependence matrix (GLDM)
Variance	Inverse difference normalized (IDN)	Small dependence emphasis (SDE)
19. Uniformity	Inverse variance	Large dependence emphasis (LDE)
	Maximum probability	Gray-level non-uniformity (GLN)
	Sum average	Dependence non-uniformity (DN)
	Sum entropy	Dependence non-uniformity normalized (DNN)
	23. Sum of squares	Gray-level variance (GLV)
Shape features	Gray-level size-zone matrix (GLSZM)	Dependence variance (DV)
Volume	Small area emphasis (SAE)	Dependence entropy (DE)
Surface area	Large area emphasis (LAE)	Low gray-level emphasis (LGLE)
Surface area to volume ratio	Gray-level non-uniformity (GLN)	High gray-level emphasis (HGLE)
Sphericity	Gray-level non-uniformity normalized (GLNN)	Small dependence low gray-level emphasis (SDLGLE)
Spherical disproportion	Size-zone non-uniformity (SZN)	Small dependence high gray-level emphasis (SDHGLE)
Maximum 3D diameter	Size-zone non-uniformity normalized (SZNN)	Large dependence low gray-level emphasis (LDLGLE)
Maximum 2D diameter (slice)	Zone percentage (ZP)	14. Large dependence high gray-level emphasis (LDHGLE)
Maximum 2D diameter (column)	Gray-level variance (GLV)	Neighboring gray tone difference matrix (NGTDM)
Maximum 2D diameter (row)	Zone variance (ZV)	1. Coarseness
Major axis	Zone entropy (ZE)	2. Contrast
Minor axis	Low gray-level zone emphasis (LGLZE)	3. Busyness
Least axis	High gray-level zone emphasis (HGLZE)	4. Complexity
Elongation	Small area low gray-level emphasis (SALGLE)	5. Strength
Flatness	Small area high gray-level emphasis (SAHGLE)	
	Large area low gray-level emphasis (LALGLE)	
	16. Large area high gray-level emphasis (LAHGLE)	

**Table 3** Feature selection and classification methods

Feature selection methods	Abbreviation	Classification methods	Abbreviation
Student <i>t</i> test		Logistic regression	LR
Minimum redundancy maximum relevance	MRMR	Random forest	RF
Least absolute shrinkage and selection operator	LASSO	Support vector machine	SVM

**Feature set preprocessing**

Due to the different range of the various radiomic features, without feature normalization, some features might exhibit a larger weight, while others might exhibit a lower weight. This depends on the distribution of feature values.

To overcome this issue, z-score normalization was applied to the feature values [33].

**Feature selection**

Three different feature selections methods were implemented in this framework (Table 3), namely enhanced

variable selection algorithms based on the least absolute shrinkage and selection operator methods [34], Student's *t* test, and the minimum redundancy maximum relevance (MRMR) algorithm.

### Multivariate machine learning classifier

The following three classifiers were implemented and compared (Table 3): logistic regression, random forest, and support vector machines (SVM).

### Model evaluation

The cross-validation (CV) technique was applied to tune the model parameters. Furthermore, bootstrapped datasets were used for model evaluation. The predictive power of all models was investigated using the area under the receiver operator characteristic (ROC) curve (AUC). All analyses and evaluations were performed using the R software for statistical computing (version 3.5.2).<sup>1</sup>

## Results

Considering the inclusion/exclusion criteria, 71 (31 low grade and 40 high grade) patients were enrolled in the study protocol. The mean age of low- and high-grade groups was 60.05 and 60.08 years, respectively. Overall, there were 51 male and 20 female subjects.

Univariate analysis demonstrated that among filtered and non-filtered images, only 128 bin-discretized images had a statistically significant difference in texture parameters with a mean AUC of  $0.74 \pm 3$  ( $q$  value  $< 0.05$ ). These features include Long-Run High Gray-Level Emphasis from GLRLM (AUC = 77,  $q$  value  $< 0.0005$ ), Cluster Tendency from GLCM (AUC = 72,  $q$  value  $< 0.002$ ), Contrast from NGTDM (AUC = 74,  $q$  value  $< 0.05$ ), and Dependence Non-Uniformity from GLDM (AUC = 72,  $q$  value  $< 0.05$  (Fig. 2).

Table 4 shows the AUC (95% CI) of three different ML-based classifiers. As shown in the table, there is a wide performance range, ranging from 0.5 to 0.86. Three different feature selection methods were applied prior to the implementation of each ML-based classifier to determine the best technique for this specific model. The results demonstrated that the LASSO method performed the best for logistic regression. Furthermore, Student's *t* test proved to be the best for random forest and SVM classifier models. The results of logistic regression suggested that 128 bin-discretized images and fine LoG features had the highest performance with a mean of AUC of 0.75. According to

the results, the predictive performance of the random forest model had a range of 0.48 to 0.67. Among these, wavelet-filtered images showed the lowest performance, whereas 128 bin-discretized images showed the highest performance. Among the three classifiers, SVM with Student's *t* test feature selection presented the best predictive performance. SVM with coarse LoG features demonstrated a mean AUC of 0.83 (Fig. 2).

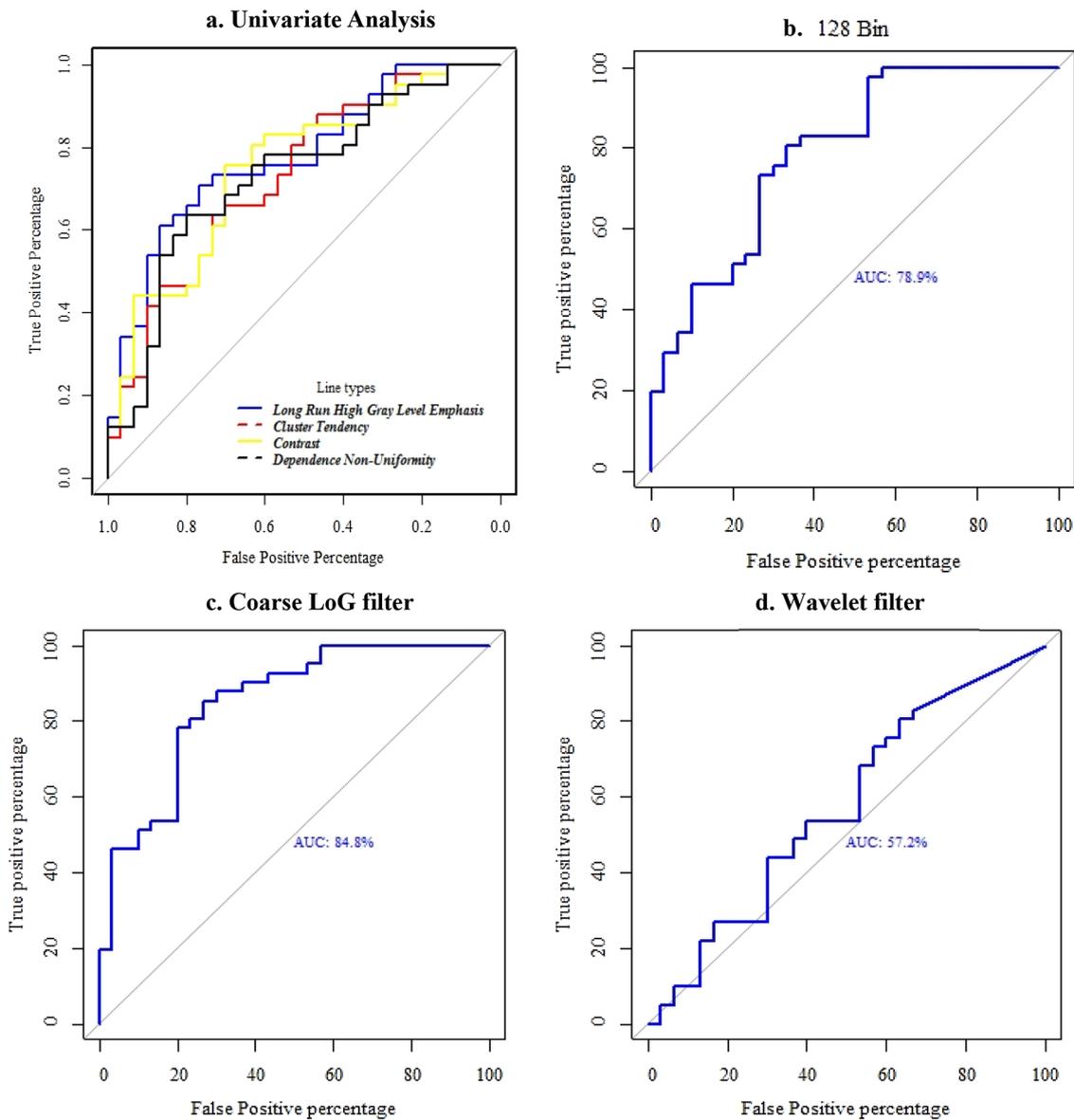
## Discussion

There is a significant association between Fuhrman grade and patient's prognosis [35, 36]. Among the different therapeutic strategies, radical surgery remains the main and most effective treatment approach to cure RCC patients. Radiofrequency ablation and active surveillance could be considered alternative treatment options for small renal masses and low risk small renal mass [37, 38]. Therefore, the preoperative assessment of RCC tumor aggressiveness plays a key role in optimal treatment planning and selection of appropriate follow-up regimens [11]. Various strategies were proposed for preoperative noninvasive prediction of ccRCC Fuhrman grade. MRI-derived ADC values are known to be an indicator of tumor activity. Several studies have assessed the utility of ADC in distinguishing low- and high-grade clear cell RCC [15, 39]. These studies showed that MRI has an acceptable predictive accuracy in the preoperative detection of high-grade RCC (AUC = 0.80) [40]. However, MRI is not as widely available as CT and a wide range of ADC values for ccRCC have been reported in the literature [41, 42]. Therefore, their robustness and repeatability need to be further validated. Conversely, CT-based semiquantitative and quantitative studies have attempted to classify low- and high-grade ccRCC [16, 17]. These studies showed that CT is a promising method for classification of low- and high-grade ccRCC.

The radiomics approach converts medical images into quantitative, high-dimensional, and mineable features enabling to predict tumor status. However, the abundance of predictive modeling techniques implies a selection process to choose the most appropriate one for predicting tumor status. Given that a number of previous radiomic studies [17, 25] for Fuhrman grade prediction did not include shape features in their analyses, this study combined shape features and texture features to differentiate low and high grades of ccRCC. It was observed that shape features cannot be ignored from multivariate machine learning models.

Univariate analysis of the extracted radiomic features demonstrated that among filtered and non-filtered images, only the 128 bin-discretized images showed statistically significant texture parameters. In a similar univariate analysis, Feng et al. [26] analyzed CT texture parameters

<sup>1</sup> <https://www.r-project.org>.



**Fig. 2** Area under receiver operating characteristic curve for discrimination between high- and low-grade ccRCC. **a** Univariate analysis of best predictor, **b** LR model with 128 bin discretization, **c** SVM model

with coarse LoG filter, **d** RF model with wavelet filter. AUC: area under receiver operating characteristic curve, LR: logistic regression, SVM: support vector machine, RF: random forest

and found effective quantitative parameters to evaluate the heterogeneity of ccRCC. After applying the LoG filter, they reported that only entropy had a statistically significant difference after FDR correction in all image phases. In this work, four features showed statistically significant differences between two groups. These features include Long-Run High Gray-level Emphasis from GLRLM, cluster tendency from GLCM, Contrast from NGTDM, and Dependence Non-Uniformity from GLDM matrix. Among these features, the Long-Run High Gray-level Emphasis demonstrated the highest AUC (AUC = 77,  $q$  value < 0.0005).

The first machine learning model applied in this study was logistic regression. This is a machine learning classification algorithm used to predict the class probability of a categorical dependent variable. It was observed that among the three different feature selection methods, the best results for the logistic regression model was obtained when using the LASSO algorithm. These results suggest that the AUC logistic regression model is approximately similar to results obtained in previous studies. Ding et al. [17] used a texture-score-based logistic regression model on a training cohort resulting in an AUC of 0.878. When predictive models were applied on the validation cohort,

**Table 4** Classifiers performance with different preprocessing techniques

	AUC (95% CI)		
	LR	RF	SVM
Original image	0.68 ± 0.08	0.62 ± 0.07	0.76 ± 0.07
32_bin	0.73 ± 0.09	0.56 ± 0.06	0.70 ± 0.08
64_bin	0.70 ± 0.07	0.55 ± 0.06	0.65 ± 0.07
128_bin	0.75 ± 0.08	0.60 ± 0.08	0.77 ± 0.08
LoG			
LoG_sigma.0.5	0.69 ± 0.08	0.53 ± 0.06	0.62 ± 0.06
LoG_sigma.1.0	0.74 ± 0.11	0.53 ± 0.02	0.64 ± 0.04
LoG_sigma.1.5	0.65 ± 0.09	0.56 ± 0.02	0.72 ± 0.07
LoG_sigma.2.0	0.68 ± 0.09	0.55 ± 0.06	0.74 ± 0.08
LoG_sigma.2.5	0.73 ± 0.11	0.54 ± 0.05	0.76 ± 0.07
LoG_sigma.3.0	0.74 ± 0.10	0.55 ± 0.09	0.77 ± 0.06
LoG_sigma.3.5	0.62 ± 0.06	0.57 ± 0.04	0.79 ± 0.08
LoG_sigma.4.0	0.65 ± 0.09	0.60 ± 0.12	0.81 ± 0.06
LoG_sigma.4.5	0.62 ± 0.06	0.62 ± 0.06	0.83 ± 0.08
LoG_sigma.5.0	0.70 ± 0.08	0.56 ± 0.05	0.78 ± 0.06
WAVELET			
Wav_HHL	0.62 ± 0.05	0.58 ± 0.04	0.71 ± 0.07
Wav_HLH	0.65 ± 0.06	0.55 ± 0.08	0.62 ± 0.06
Wav_LHH	0.67 ± 0.08	0.53 ± 0.03	0.65 ± 0.07
Wav_HLL	0.62 ± 0.06	0.57 ± 0.05	0.76 ± 0.08
Wav_LLH	0.68 ± 0.07	0.56 ± 0.06	0.59 ± 0.08
Wav_LHL	0.62 ± 0.06	0.53 ± 0.05	0.65 ± 0.08
Wav_LLL	0.62 ± 0.06	0.55 ± 0.08	0.75 ± 0.10

good results were still obtained (AUC > 0.670). Shu et al. [43] extracted radiomic features from corticomedullary (CMP) and nephrographic phases (NP) of CT images of 161 and 99 patients diagnosed with low- and high-grade ccRCCs. They constructed logistic regression classification models to discriminate between high- and low-grade ccRCC. Application of the model on CMP and NP showed an AUC of 0.766 (95% CI 0.709–0.816) and 0.818 (95% CI 0.765–0.838), respectively. Random forest, an ensemble learning method consisting of a collection of decision trees, was an alternative machine learning model investigated in this work. It uses a weighted average of these trees for the final decision [44], commonly resulting in a good outcome for a large range of data, but is susceptible to overfitting. In this work, applying the random forest model on the dataset yielded unsatisfactory results (Table 4). SVM, which creates a decision boundary between two classes to enable the prediction of labels from one or more feature vectors, was the best-performing classifier. After applying the SVM model on filtered and unfiltered images, the best classification result was obtained when coarse LoG features were used with a mean AUC of 0.81. LoG filtering is an advanced image filtering technique combining Laplacian filtering and Gaussian

filtering. In a similar single-center retrospective study [25], the performance of quantitative CT texture analysis combined with different ML-based classifiers was evaluated for discriminating low- and high-grade ccRCC. Despite differences in the procedure followed, they also determined that the highest predictive performance is achieved by an SVM classifier. In summary, both studies support each other with a common conclusion that CT texture analysis is a useful and promising noninvasive method to predict the Fuhrman grades of ccRCCs preoperatively. In the current study, we applied machine learning algorithms on CT radiomic features to predict noninvasively Fuhrman grades of ccRCCs, demonstrating promising results. The noninvasive identification of ccRCC grading could help in defining appropriate treatments, especially for patients with small mass and could potentially serve as an alternative for FNA in renal cancer.

This work bears a number of limitations. (1) This was a retrospective study using a small sample size with no external validation data. As such, we used the bootstrap technique (estimate statistics on a population by sampling a dataset with replacement) to evaluate our models. Future studies exploiting the current model could involve a large size of external validation set emanating from a multicenter study protocol. (2) Since the tumor boundary was drawn manually, interference of the volume effect cannot be completely avoided. Yet, the current study involved image segmentation performed and evaluated by one human observer. Future studies should consider the effect of volume segmentation to provide a repeatable study for clinical multicentric studies.

## Conclusion

The results of this study show that CT-based SVM classifier with t test features selection could be a useful and promising noninvasive approach for the prediction of low and high Fuhrman nuclear-grade ccRCCs. Moreover, the results demonstrated that 128 bin-discretized preprocessing is an effective method under these conditions.

**Acknowledgments** This work was supported by the Shahid Beheshti University of Medical Sciences under Grant Number 388 and the Swiss National Science Foundation under Grant SNRF 320030\_176052.

## Compliance with ethical standards

**Conflict of interest** The authors have no conflict of interest to declare.

**Ethical standards** This article does not contain any studies with human participants or animals performed by any of the authors.

**Informed consent** Informed consent is not needed for this study.

**Research involving human participants and/or animals** All human subject studies were downloaded from The Cancer Imaging Archive

(TCIA), an open-access database of medical images for cancer research. The site is funded by the National Cancer Institute's Cancer Imaging Program, and the contract operated by the University of Arkansas for Medical Sciences.

## References

- Capitanio U, Montorsi F (2016) Renal cancer. *Lancet* 387(10021):894–906
- Srigley JR, Delahunt B, Eble JN, Egevad L, Epstein JI, Grignon D, Hes O, Moch H, Montironi R, Tickoo SK (2013) The International Society of Urological Pathology (ISUP) Vancouver classification of renal neoplasia. *Am J Surg Pathol* 37(10):1469–1489
- Marconi L, Dabestani S, Lam TB, Hofmann F, Stewart F, Norrie J, Bex A, Bensalah K, Canfield SE, Hora M (2016) Systematic review and meta-analysis of diagnostic accuracy of percutaneous renal tumour biopsy. *Eur Urol Oncol* 69(4):660–673
- Muglia VF, Prando A (2015) Renal cell carcinoma: histological classification and correlation with imaging findings. *Radiol Bras* 48(3):166–174
- Delahunt B, Chevillet JC, Martignoni G, Humphrey PA, Magi-Galluzzi C, McKenney J, Egevad L, Algaba F, Moch H, Grignon DJ (2013) The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters. *Am J Surg Pathol* 37(10):1490–1504
- Lohse CM, Blute ML, Zincke H, Weaver AL, Chevillet JC (2002) Comparison of standardized and nonstandardized nuclear grade of renal cell carcinoma to predict outcome among 2,042 patients. *Am J Clin Pathol* 118(6):877–886
- Delahunt B (2009) Advances and controversies in grading and staging of renal cell carcinoma. *Mod Pathol* 22(S2):S24
- Tomaszewski JJ, Uzzo RG, Smaldone MC (2014) Heterogeneity and renal mass biopsy: a review of its role and reliability. *Cancer Biol Med* 11(3):162
- Blumenfeld AJ, Guru K, Fuchs GJ, Kim HL (2010) Percutaneous biopsy of renal cell carcinoma underestimates nuclear grade. *Urology* 76(3):610–613
- Kutikov A, Kunkle DA, Uzzo RG (2009) Focal therapy for kidney cancer: a systematic review. *Curr Opin Urol* 19(2):148–153
- Ljungberg B, Albiges L, Abu-Ghanem Y, Bensalah K, Dabestani S, Fernández-Pello S, Giles RH, Hofmann F, Hora M, Kuczyk MA (2019) European association of urology guidelines on renal cell carcinoma: the 2019 update. *Eur Urol* 75(5):799–810
- Volpe A, Panzarella T, Rendon RA, Haider MA, Kondylis FI, Jewett MA (2004) The natural history of incidentally detected small renal masses. *Cancer* 100(4):738–745
- Bratslavsky G, Kirkali Z (2010) The changing face of renal-cell carcinoma. *J Endourol* 24(5):753–757
- Kunkle DA, Egleston BL, Uzzo RG (2008) Excise, ablate or observe: the small renal mass dilemma—a meta-analysis and review. *J Urol* 179(4):1227–1234
- Yoshida R, Yoshizako T, Hisatoshi A, Mori H, Tamaki Y, Ishikawa N, Kitagaki H (2017) The additional utility of apparent diffusion coefficient values of clear-cell renal cell carcinoma for predicting metastasis during clinical staging. *Acta Radiol Open* 6(1):2058460116687174
- Ishigami K, Leite LV, Pakalniskis MG, Lee DK, Holanda DG, Kuehn DM (2014) Tumor grade of clear cell renal cell carcinoma assessed by contrast-enhanced computed tomography. *Springer-Plus* 3(1):694
- Ding J, Xing Z, Jiang Z, Chen J, Pan L, Qiu J, Xing W (2018) CT-based radiomic model predicts high grade of clear cell renal cell carcinoma. *Eur J Radiol* 103:51–56
- Lambin P, Leijenaar RT, Deist TM, Peerlings J, De Jong EE, Van Timmeren J, Sanduleanu S, Larue RT, Even AJ, Jochems A (2017) Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 14(12):749
- Shiri I, Maleki H, Hajianfar G, Abdollahi H, Ashrafinia S, Hatt M, Oveisi M, Rahmim A (2019) Next generation radiogenomics sequencing for prediction of EGFR and KRAS mutation status in NSCLC patients using multimodal imaging and machine learning approaches. *arXiv preprint arXiv:190702121*
- Hajianfar G, Shiri I, Maleki H, Oveisi N, Haghparast A, Abdollahi H, Oveisi M (2019) Noninvasive O6 Methylguanine-DNA Methyltransferase status prediction in glioblastoma multiforme cancer using magnetic resonance imaging radiomics features: univariate and multivariate radiogenomics analysis. *World Neurosurg* 132:140–161
- Abdollahi H, Mahdavi SR, Shiri I, Mofid B, Bakhshandeh M, Rahmani K (2019) Magnetic resonance imaging radiomic feature analysis of radiation-induced femoral head changes in prostate cancer radiotherapy. *J Cancer Res Ther* 15(8):11
- Abdollahi H, Shiri I, Heydari M (2019) Medical imaging technologists in radiomics era: an alice in wonderland problem. *Iran J Public Health* 48(1):184
- Shiri I, Rahmim A, Ghaffarian P, Geramifar P, Abdollahi H, Bitarafan-Rajabi A (2017) The impact of image reconstruction settings on 18F-FDG PET radiomic features: multi-scanner phantom and patient studies. *Eur Radiol* 27(11):4498–4509
- Shiri I, Ghafarian P, Geramifar P, Leung KH-Y, Ghelichoghli M, Oveisi M, Rahmim A, Ay MR (2019) Direct attenuation correction of brain PET images using only emission data via a deep convolutional encoder-decoder (Deep-DAC). *Eur Radiol* 29(12):6867–6879
- Bektas CT, Kocak B, Yardimci AH, Turkcanoglu MH, Yucetas U, Koca SB, Erdim C, Kilickesmez O (2019) Clear cell renal cell carcinoma: machine learning-based quantitative computed tomography texture analysis for prediction of Fuhrman nuclear grade. *Eur Radiol* 29(3):1153–1163
- Feng Z, Shen Q, Li Y, Hu Z (2019) CT texture analysis: a potential tool for predicting the Fuhrman grade of clear-cell renal carcinoma. *Cancer Imaging* 19(1):6
- Shu J, Wen D, Xi Y, Xia Y, Cai Z, Xu W, Meng X, Liu B, Yin H (2020) Clear cell renal cell carcinoma: machine learning-based computed tomography radiomics analysis for the prediction of WHO/ISUP grade. *Eur J Radiol* 121:108738 **in press**
- Akin O, Elnajjar P, Heller M, Jarosz R, Erickson B, Kirk S, Filipini J (2016) Radiology data from the cancer genome atlas kidney renal clear cell: carcinoma [TCGA-KIRC] collection. The Cancer Imaging Archive Website: <https://wiki.cancerimagingarchive.net/display/Public/TCGA-KIRC>
- Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, Buatti J, Aylward S, Miller JV, Pieper S, Kikinis R (2012) 3D slicer as an image computing platform for the quantitative imaging network. *Magn Reson Imaging* 30(9):1323–1341
- Shafiq-ul-Hassan M, Zhang GG, Latifi K, Ullah G, Hunt DC, Balagurunathan Y, Abdallah MA, Schabath MB, Goldgof DG, Mackin D (2017) Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med Phys* 44(3):1050–1062
- Van Griethuysen JJ, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RG, Fillion-Robin J-C, Pieper S, Aerts HJ (2017) Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 77(21):e104–e107
- Haynes W (2013) Benjamini–hochberg method. *Encyclopedia of systems biology*, pp 78–78
- Kickingeder P, Götz M, Muschelli J, Wick A, Neuberger U, Shinohara RT, Sill M, Nowosielski M, Schlemmer H-P, Radbruch A (2016) Large-scale radiomic profiling of recurrent glioblastoma

- identifies an imaging predictor for stratifying anti-angiogenic treatment response. *Clin Cancer Res* 22(23):5765–5771
34. Guo P, Zeng F, Hu X, Zhang D, Zhu S, Deng Y, Hao Y (2015) Improved variable selection algorithm using a LASSO-type penalty, with an application to assessing hepatitis B infection relevant factors in community residents. *PLoS ONE* 10(7):e0134151
  35. Tian K, Rubadue CA, Lin DI, Veta M, Pyle ME, Irshad H, Heng YJ (2019) Automated clear cell renal carcinoma grade classification with prognostic significance. *PLoS ONE* 14(10):e0222641
  36. Ficarra V, Righetti R, Martignoni G, D'Amico A, Piloni S, Rubilotta E, Malossini G, Mobilio G (2001) Prognostic value of renal cell carcinoma nuclear grading: multivariate analysis of 333 cases. *Urol Int* 67(2):130–134
  37. Vetterlein MW, Jindal T, Becker A, Regier M, Kluth LA, Tilki D, Chun FK-H (2016) Small renal masses in the elderly: contemporary treatment approaches and comparative oncological outcomes of nonsurgical and surgical strategies. *Investig Clin Urol* 57(4):231–239
  38. Jewett MA, Mattar K, Basiuk J, Morash CG, Pautler SE, Siemens DR, Tanguay S, Rendon RA, Gleave ME, Drachenberg DE (2011) Active surveillance of small renal masses: progression patterns of early stage kidney cancer. *Eur Urol* 60(1):39–44
  39. Rosenkrantz AB, Niver BE, Fitzgerald EF, Babb JS, Chandarana H, Melamed J (2010) Utility of the apparent diffusion coefficient for distinguishing clear cell renal cell carcinoma of low and high nuclear grade. *AJR Am J Roentgenol* 195(5):W344–W351
  40. Maruyama M, Yoshizako T, Uchida K, Araki H, Tamaki Y, Ishikawa N, Shiina H, Kitagaki H (2015) Comparison of utility of tumor size and apparent diffusion coefficient for differentiation of low-and high-grade clear-cell renal cell carcinoma. *Acta Radiol* 56(2):250–256
  41. Zhang J, Mazaheri Tehrani Y, Wang L, Ishill NM, Schwartz LH, Hricak H (2008) Renal masses: characterization with diffusion-weighted MR imaging: a preliminary experience. *Radiology* 247(2):458–464
  42. Goyal A, Sharma R, Bhalla AS, Gamanagatti S, Seth A, Iyer VK, Das P (2012) Diffusion-weighted MRI in renal cell carcinoma: a surrogate marker for predicting nuclear grade and histological subtype. *Acta Radiol* 53(3):349–358
  43. Shu J, Tang Y, Cui J, Yang R, Meng X, Cai Z, Zhang J, Xu W, Wen D, Yin H (2018) Clear cell renal cell carcinoma: CT-based radiomics features for the prediction of Fuhrman grade. *Eur J Radiol* 109:8–12
  44. Breiman L (2001) Random forests. *Mach Learn* 45:5–32

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.