

# Multiphase Contrast-Enhanced CT-Based Machine Learning Models to Predict the Fuhrman Nuclear Grade of Clear Cell Renal Cell Carcinoma

This article was published in the following Dove Press journal:  
*Cancer Management and Research*

Shengsheng Lai<sup>1,\*</sup>

Lei Sun<sup>2,\*</sup>

Jialiang Wu<sup>3</sup>

Ruili Wei<sup>4</sup>

Shiwei Luo<sup>4</sup>

Wenshuang Ding<sup>5</sup>

Xilong Liu<sup>6</sup>

Ruimeng Yang<sup>4</sup>

Xin Zhen<sup>2</sup>

<sup>1</sup>School of Medical Equipment, Guangdong Food and Drug Vocational College, Guangzhou, Guangdong, 510520, People's Republic of China; <sup>2</sup>School of Biomedical Engineering, Southern Medical University, Guangzhou, Guangdong, 510515, People's Republic of China; <sup>3</sup>Department of Radiology, The University of Hong Kong Shenzhen Hospital, Shenzhen, Guangdong, 518000, People's Republic of China; <sup>4</sup>Department of Radiology, Guangzhou First People's Hospital, School of Medicine, South China University of Technology, Guangzhou, Guangdong, 510180, People's Republic of China; <sup>5</sup>Department of Pathology, Guangzhou First People's Hospital, School of Medicine, South China University of Technology, Guangzhou, Guangdong, 510180, People's Republic of China; <sup>6</sup>Department of Radiology, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, 510515, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Ruimeng Yang  
Department of Radiology, Guangzhou First People's Hospital, School of Medicine, South China University of Technology, Guangzhou, Guangdong, 510180, People's Republic of China  
Tel +86-20-81048873  
Email eyruimengyang@scut.edu.cn

Xin Zhen  
School of Biomedical Engineering, Southern Medical University, Guangzhou, Guangdong, 510515, People's Republic of China  
Tel +86-20-62789323  
Email xinzen@smu.edu.cn

**Objective:** To investigate the predictive performance of different machine learning models for the discrimination of low and high nuclear grade clear cell renal cell carcinoma (ccRCC) by using multiphase computed tomography (CT)-based radiomic features.

**Materials and Methods:** A total of 137 consecutive patients with pathologically proven ccRCC (including 96 low-grade [grade 1 or 2] and 41 high-grade [grade 3 or 4] ccRCC) from January 2011 to January 2019 were enrolled in this retrospective study. Target region of interest (ROI) delineation followed by texture extraction was performed on a representative slice with the largest section of the tumor on the four-phase (unenhanced phase [UP], corticomedullary phase [CMP], nephrographic phase [NP] and excretory phase [EP]) CT images. Fifteen concatenations of the four-phase features were fed into 176 classification models (built with 8 classifiers and 22 feature selection methods), the classification performances of the 2640 resultant discriminative models were compared, and the top-ranked features were analyzed.

**Results:** Image features extracted from the unenhanced phase (UP) CT images demonstrated a dominant classification performance over features from the other three phases. The discriminative model "Bagging + CMIM" achieved the highest classification AUC of 0.75. The top-ranked features from the UP included one shape-based feature and five first-order statistical features.

**Conclusion:** Image features extracted from the UP are more effective than other CT phases in differentiating low and high nuclear grade ccRCC based on machine learning-based classification modeling.

**Keywords:** clear cell renal cell carcinoma, Fuhrman nuclear grade, computed tomography, machine learning, classification

## Introduction

Renal cell carcinoma (RCC) is the most common primary malignant tumor of the kidney in adults, accounting for approximately 90–95% of renal tumors, with clear cell renal cell carcinoma (ccRCC) being the predominant subtype that portends a worse prognosis than other subtypes.<sup>1</sup> Histopathologic grade has been shown to be an independent prognostic factor of survival<sup>2–4</sup> and is critical in selecting individualized treatment strategies, such as radical nephrectomy (RN) for aggressive RCC cases, partial nephrectomy (PN) for non-aggressive cases, or more conservative ablation or active surveillance for selective patients who have extensive comorbidities or are unwilling to undergo surgery.<sup>5,6</sup> To date, several systems

have been employed to grade RCC, where the Fuhrman grading system<sup>7</sup> is the most widely used and highly recommended grading scheme.<sup>8</sup>

For evaluating the ccRCC Fuhrman nuclear grade, preoperative percutaneous biopsy of renal mass is controversial because although it is the optimal alternative due to the high spatial and temporal heterogeneity of ccRCC, it is an invasive approach and has potential sampling bias.<sup>9</sup> Several pioneer investigative efforts have been made to predict the Fuhrman grade of ccRCC by analyzing morphological or/and dynamic contrast-enhanced medical imaging features based on conventional CT, dual-energy spectral CT,<sup>10–15</sup> CT/MR perfusion<sup>16,17</sup> or diffusion-weighted imaging (DWI).<sup>18,19</sup> However, most of these aforementioned strategies are limited by the subjective preference of features, leading to unsatisfactory discriminative efficacy and model generalization capability.

Recently, radiomics has emerged as a promising noninvasive biomarker in assessing histopathological characterization and predicting biological aggressiveness in oncology. Several investigations have been reported in quantitatively analyzing CT-based radiomic features in an attempt to differentiate low and high Fuhrman nuclear grades.<sup>20–27</sup> However, these reviewed studies are confined by the limitations that either texture features were usually extracted from a single CT phase or classification modeling was built upon a randomly selected classifier. To the best of our knowledge, no comprehensive investigations have been reported in determining which phase(s), classifier(s) or their possible combinations could be more discriminative.

The purpose of this study was to investigate the discriminative power of different classification models built with assorted combinations of classifiers and feature selection approaches in differentiating low vs high Fuhrman grades. Features were extracted from unenhanced renal CT images and different postcontrast-enhanced CT phases, alone or in combination. The top-ranked features discriminating low vs high Fuhrman grade were extracted and analyzed.

## Materials and Methods

### Patients

This study was approved by the Ethics Committee of Guangzhou First People's Hospital and Nanfang Hospital and complied with the Declaration of Helsinki. The patient consent to review their medical records was waived because our study was a retrospective non-interventive study, which did not harm patients. The data were obtained through an

electronic search of the picture archiving and communication system (PACS) from January 2011 to January 2019. The inclusion criteria were as follows: 1) patients with pathologically proven ccRCC and with defined Fuhrman grade; and 2) patients who had undergone preoperative four-phase contrast-enhanced CT (CECT) scans. The exclusion criteria were as follows: 1) patients with purely cystic ccRCC; 2) ccRCC without Fuhrman grade; and 3) prominent artifacts on CT.

### Fuhrman Stage and Image Acquisition

To improve the reproducibility of pathologists and reduce the intra/inter-observer variability, the traditional 4-tier Fuhrman grading system was re-categorized into a simplified Fuhrman grading system with low-grade (grades 1 and 2) and high-grade (grades 3 and 4). Fuhrman grading was accomplished by a subspecialized genitourinary pathologist (W.S. Ding) with 8 years of diagnostic experience.

Preoperative CECT images were obtained on multiple scanners: Toshiba Aquilion One, Siemens Somatom Definition, GE HiSpeed 16 and Philips Brilliance 64. The acquisition parameters were as follows: tube voltage, 120–140 kV; automated tube current modulation and varied milliampere-second settings; and reconstruction slice thickness, 5 mm. All patients were injected with nonionic intravenous contrast material via the antecubital vein with mechanical power injectors according to their weight (1 mL/kg body weight, with a maximum of 150 mL). All patients underwent preoperative four-phase CT scans: Phase 1, unenhanced (UP); Phase 2, postcontrast corticomedullary phase (CMP), Phase 3, postcontrast nephrographic phase (NP); and Phase 4, postcontrast excretory phase (EP).

### Feature Extraction and Representation

All CT images retrieved from PACS were stored in anonymized DICOM format. ITK-SNAP software (<http://www.itksnap.org>) was used for the delineation of the target region of interest (ROI) on the CT slice with the largest cross-sectional area of the tumor in phases 1–4 for tumor segmentation. A lesion ROI was first delineated on the CMP and then applied to the other three phases with a slight adjustment to acquire the tailored ROIs for each phase. This procedure was conducted by two investigators (J. L. Wu and R. M. Yang with 4 and 15 years of experience in radiological diagnosis, respectively) together without prior knowledge of the lesion pathology.

Texture feature extraction was accomplished (with Pyradiomics<sup>28</sup>) on the delineated ROIs in each phase. The extracted features from each phase included 102 candidate features, which could be categorized into three subtypes, including shape features, first-order statistical features (histogram analysis), and second-order statistical features (or “texture features”). The features extracted from each of the four phases were termed  $F_{pha}^1$ ,  $F_{pha}^2$ ,  $F_{pha}^3$ , and  $F_{pha}^4$  (each with 102 features). The concatenation of features from any two phases were termed  $F_{pha}^{1,2}$ ,  $F_{pha}^{1,3}$ ,  $F_{pha}^{1,4}$ ,  $F_{pha}^{2,3}$ ,  $F_{pha}^{2,4}$ , and  $F_{pha}^{3,4}$  (each with 204 features), and the concatenation of features from any three phases were termed  $F_{pha}^{1,2,3}$ ,  $F_{pha}^{1,2,4}$ ,  $F_{pha}^{1,3,4}$ , and  $F_{pha}^{2,3,4}$  (each with 306 features). Similarly, the concatenated features from all four phases were termed  $F_{pha}^{1,2,3,4}$  (with 408 features). Discriminative capabilities were compared using the above 15 types of features as input for a specific discrimination model.

## Modeling and Comparisons

In this study, we built particular discrimination models based on a feature selection strategy from a pool of 22 feature selection methods, as well as a classifier from a pool of eight classification algorithms (refer to [Table S1](#) in the Supplement), thus resulting in a total of 176 different discrimination models. The 15 types of features were fed into each of the 176 discriminative models, resulting in 2640 (15 x 176) combinations in total to be compared. We evaluated each of these models with five-fold cross-validation, in each of which an optimal subset of features (a fixed size of 20 features was empirically selected to balance both sample size and feature numbers) was first estimated by a specific feature selection method, and these prescreened features were further fed into a classifier for discrimination modeling. To ease the data imbalance of the patient cohort, the synthetic minority over-sampling technique (SMOTE)<sup>29</sup> was used to oversample the minority high-grade ccRCC feature group by introducing synthetic feature samples. The discriminative powers of the

models were quantified by the area under the receiver operating characteristic (ROC) curve (AUC).

## Statistical Analysis

The randomized frequency matching method was employed for adjusting tumor sizes, specifically, the tumor sizes in the two groups were first, respectively, ranked in a descending order; then, a case with a particular tumor size was selected in one group, and its counterpart in the other group was randomly selected among those cases with a similar tumor size. Continuous variables are reported as the means ( $\pm$ standard deviations), and categorical variables are reported as numbers (proportions). The normality of the data distribution was assessed for continuous variables (Kolmogorov–Smirnov test for  $>50$  patients and Shapiro–Wilk test for  $\leq 50$  patients). Comparisons between groups were performed using the following statistical tests: the chi-square test for categorical variables, independent *t*-test for normally distributed continuous variables, and the Mann–Whitney *U*-test for nonnormally distributed continuous variables. Discriminative comparisons between the 15 types of features were conducted using the independent samples Kruskal–Wallis test with Bonferroni correction for adjusting the significance level in pairwise comparisons.

All statistical analyses were performed using SPSS version 20 (IBM Corporation, Armonk, NY, USA). A two-tailed significance level of  $p < 0.05$  was considered statistically significant.

## Results

### Demographics

This study comprised 96 low-grade [grade 1 (n=15, 10.9%) and 2 (n=81, 59.1%)] ccRCC patients and 41 high-grade [grade 3 (n=32, 23.4%) and 4 (n=9, 6.6%)] ccRCC patients who met the selection criteria and had complete imaging examinations. The patient characteristics of the two groups and the whole data set are provided in [Table 1](#). There were

**Table 1** Demographics and Characteristics of the Study Population

Characteristics	Low-Grade (n=96)	High-Grade (n=41)	p value
Patient age (mean $\pm$ SD, year)	53.9 $\pm$ 11.8	58.1 $\pm$ 12.9	0.062 <sup>a</sup>
Sex			0.976 <sup>b</sup>
Male (n/%)	70(72.9)	30(73.1)	
Female (n/%)	26(27.1)	11(26.9)	
Lesion diameter (mean $\pm$ SD, cm)	5.49 $\pm$ 2.38	5.80 $\pm$ 2.76	0.512 <sup>a</sup>

**Note:** The data are described as the mean  $\pm$  SD or frequency (%). <sup>a</sup>Independent *t*-test; <sup>b</sup>Chi-square test. A *p* value less than 0.05 was considered statistically significant.

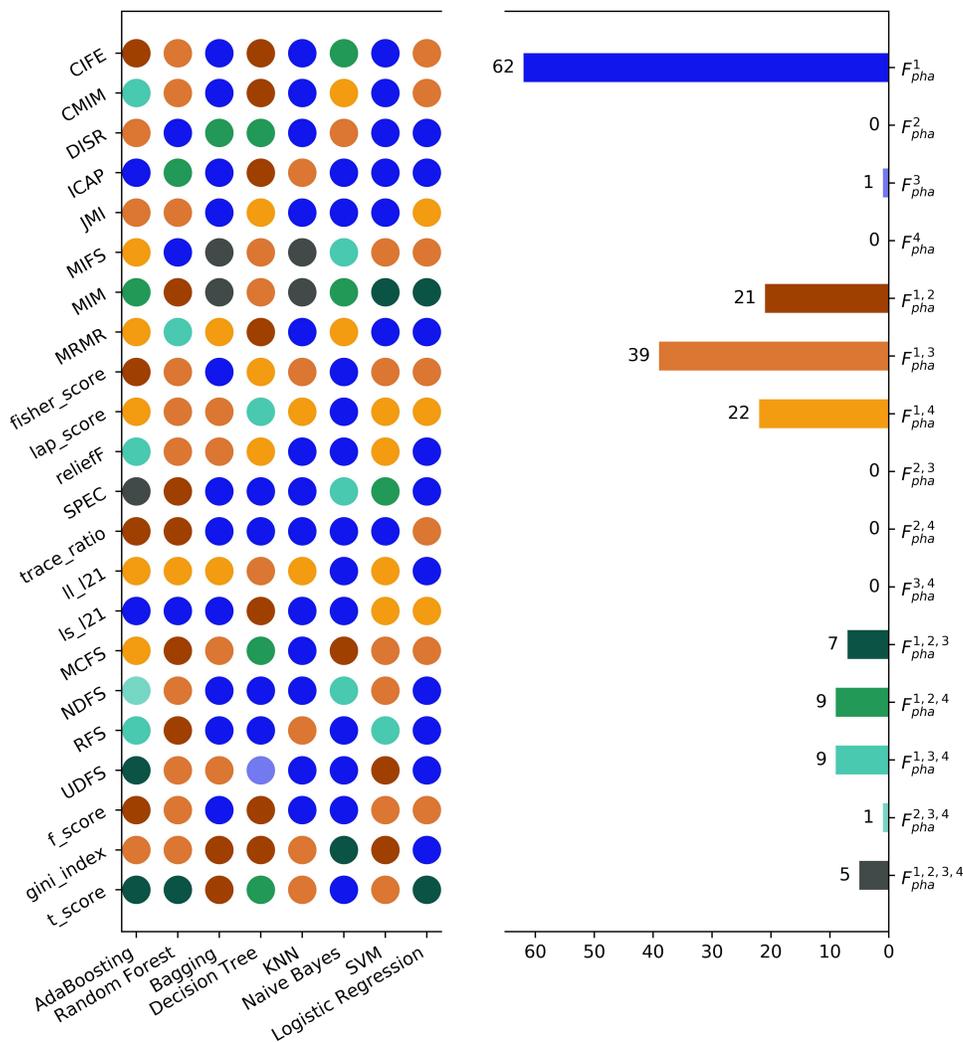
no significant differences in patient age, sex or lesion diameter between the two groups ( $p > 0.05$ ).

### Discriminative Capabilities of Different Feature Types

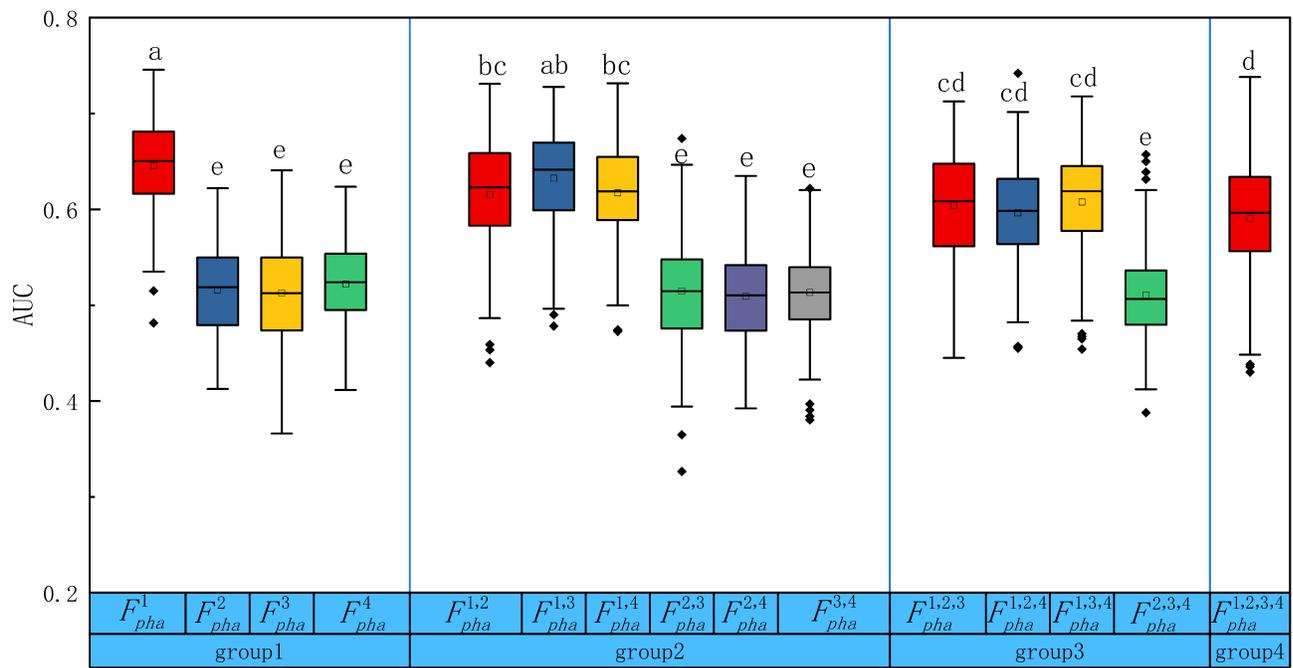
The 15 feature types were compared by being fed to each of the 176 discrimination models. The feature type with the highest AUC among the 15 feature types was marked with a dedicated color and shown as a color map, while the number of times this feature was ranked as the best feature type was also counted (Figure 1). We can see that Phase 1 ( $F_{pha}^1$ ) had the most (62 times) best-performing features, followed by  $F_{pha}^{1,3}$  (39 times). Other leading phases were  $F_{pha}^{1,4}$  (22 times) and  $F_{pha}^{1,2}$  (21 times). It seems that those combinations including features from Phase

1 ( $F_{pha}^1$ ) generally had a better discriminative performance, suggesting that Phase 1 was a superior feature representation for ccRCC Fuhrman grade classification.

The boxplot of the AUC distributions for all 15 feature types is illustrated in Figure 2. The superiority of Phase 1 over the other phases can also be verified by statistical comparisons. The models built on  $F_{pha}^1$  performed significantly better than the other 14 feature types, while the phase combinations including Phase 1, eg,  $F_{pha}^{1,2}$ ,  $F_{pha}^{1,3}$ , and  $F_{pha}^{1,4}$ , significantly outperformed the combinations without Phase 1, eg,  $F_{pha}^{2,3}$ ,  $F_{pha}^{2,4}$ ,  $F_{pha}^{3,4}$  and  $F_{pha}^{2,3,4}$ . Similarly,  $F_{pha}^{1,2,3}$ ,  $F_{pha}^{1,2,4}$ ,  $F_{pha}^{1,3,4}$  and  $F_{pha}^{1,2,3,4}$  achieved significantly better performance outcomes than  $F_{pha}^{2,3}$ ,  $F_{pha}^{2,4}$ ,  $F_{pha}^{3,4}$  and  $F_{pha}^{2,3,4}$ .



**Figure 1** Each colored circle represents a specific discriminative model (176 models in total) with different combinations of classifier and feature selection methods. A dedicated color is assigned to each circle with respect to the feature type that has the highest AUC value among all the 15 feature types for the corresponding discriminative model.



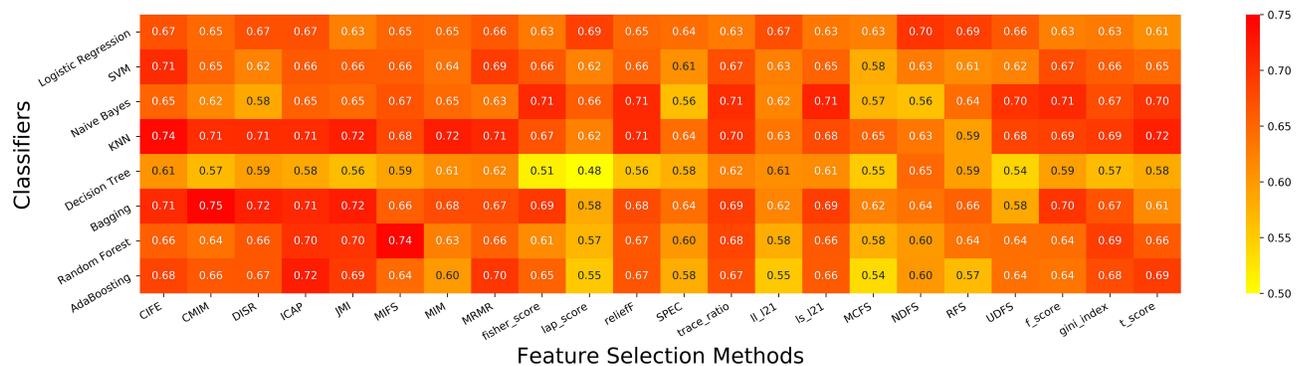
**Figure 2** Boxplots of the AUC distributions for all the 15 feature types. The boxes run from the 25th percentile to the 75th percentile; the two ends of the whiskers represent the 5% and 95% percentiles of the data, and the horizontal line and the square in the box represent the median and mean values, respectively. The diamonds represent outliers. The letters above each box indicate whether a statistically significant difference (by Kruskal–Wallis test with Bonferroni correction) exists between any two feature types. No common letters indicate that the two feature types are significantly different.

### Key Feature Analysis

The AUC values achieved by all 176 discriminative models using features from Phase1  $F^1_{pha}$  are shown as a heatmap in Figure 3. The highest AUC of 0.75 was achieved by the discriminative model with the combination of the classifier “Bagging” and the feature selection method “CMIM”. In the five-fold cross-validation, the feature selection method embedded in each discriminative model selected 20 key features for further classification. We counted the number of times each feature in  $F^1_{pha}$  was selected as the top-20

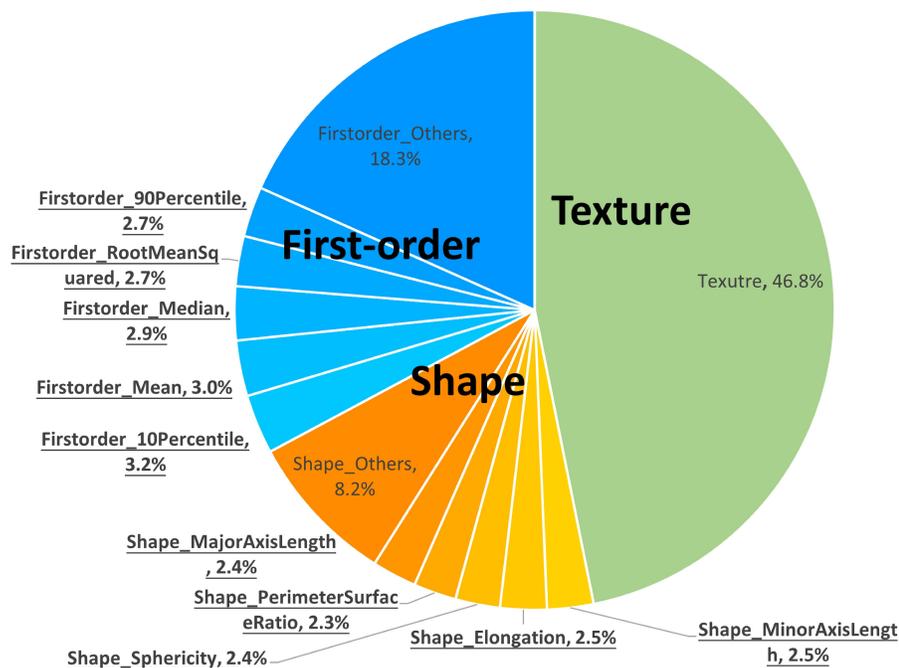
features (only for those models with AUCs >0.6), and the corresponding percentage is summarized and shown as a pie chart in Figure 4. The top-10 most frequently selected features in  $F^1_{pha}$  are highlighted in Figure 4, including five shape features and five first-order statistical features. No texture features were among the top-10 features.

The top-10 most frequently selected features are summarized in Table 2. All the top-5 first-order statistical features (four with  $p < 10^{-3}$  and one with  $p < 10^{-2}$ ) and one shape feature (elongation, with  $p = 0.023$ ) were statistically



**Figure 3** A heatmap representation of the AUC values obtained by the 176 discriminative models ( $F_{pha}^1$  as feature input) built with different combinations of classifiers and feature selection methods.

## Number of times (%) being selected as the Top-20 features



**Figure 4** A pie chart showing the number of times (%) the features in  $F_{\phi}^I$  were selected into the top-20 features in the five-fold cross-validation of all discriminative models with AUCs >0.6. The most frequent 10 features are highlighted with underlined font.

significant features. We estimated the capability of using the mean of the mean feature values of the two groups (ie, “M” in Table 2) as the threshold to differentiate the two groups. It was observed that the top-5 first-order statistical features, ie, the 10th percentile, the 90th percentile, the mean, the median and the root mean squared, demonstrated discriminative capabilities in which ~60% of the high-grade group (~40% of the low-grade group) had larger feature values, while comparatively ~60% of the low-grade group (~40% of the high-grade group) had smaller feature values.

## Discussion

The originality of this multicenter retrospective study is that it comprehensively explores the optimal source of CT images from multiple CT phases (unenhanced phase and three dynamic enhanced phases) and the most ideal discriminative scheme among different classification models for the differentiation of low and high Fuhrman nuclear grade ccRCC. After thorough comparisons, the discriminative model composed of “CMIM” and “Bagging” yielded satisfactory performance with radiomics features

extracted from the unenhanced phase (UP). Additionally, the shape- and first-order-based features were found to be more prominent over the texture features.

In terms of radiomics analysis-based predictions of the Fuhrman grade of ccRCC, most of the previous radiomics-based studies focused on features extracted from a single dynamic enhanced CT phase<sup>20,23</sup> or multiple postcontrast CT phases without including the UP and the excretory phase [EP].<sup>22,25,26</sup> Only two previous investigations included the UP, eg, Lin et al<sup>21</sup> reported that they employed a machine learning model based on three-phase [precontrast phase (PCP, equal to UP in our study), CMP and NP] CT images and claimed to achieve superior diagnostic performance to those based on single-phase CT images in differentiating low- from high-grade ccRCC. Kocak’s study<sup>24</sup> acknowledged the role of the UP, and they stated that machine learning-based unenhanced CT texture analysis could be a promising noninvasive method with favorable accuracy. Our finding is consistent with the previous study, such that the UP phase seems to be more suitable for texture analysis in differentiating the nuclear

**Table 2** The Top-10 Most Frequently Selected Features in  $F^1_{plur}$ . Their Corresponding Mean ( $\pm$ SD) Values (or Median (IQR)) and the P-values Between Low and High Fuhrman Grades

Feature Category	Top-Ranked Features	High-Grade	Low-Grade	P-value	M <sup>1a</sup>	(<M   >M) <sup>1b</sup>
Shape-based (n=5)	Elongation (7th)	0.80 (0.57–0.96)	0.83 (0.55–0.98)	0.023 <sup>d</sup>	0.81	High (58.5%   41.5%) Low (36.5%   63.5%)
	Major Axis Length (10th)	87.63 (25.36–219.06)	82.98 (25.15–190.27)	0.633 <sup>d</sup>	85.31	High (51.2%   48.8%) Low (53.1%   46.9%)
	Minor Axis Length (6th)	67.42 (20.96–153.33)	67.93 (22.77–149.78)	0.960 <sup>d</sup>	67.67	High (51.2%   48.8%) Low (53.1%   46.9%)
	Perimeter Surface Ratio (9th)	0.07 (0.02–0.19)	0.07 (0.03–0.17)	0.808 <sup>d</sup>	0.07	High (68.3%   31.7%) Low (63.5%   36.5%)
	Sphericity (8th)	0.91 (0.81–0.94)	0.92 (0.80–0.96)	0.176 <sup>d</sup>	0.92	High (39.0%   61.0%) Low (32.3%   67.7%)
	10th Percentile (1st)	20.97±6.13	17.24±7.11	0.0042 <sup>c</sup>	19.11	High (36.6%   63.4%) Low (59.4%   40.6%)
	90th Percentile (5th)	47.74±5.36	43.15±8.59	<0.001 <sup>c</sup>	45.45	High (31.7%   68.3%) Low (63.5%   36.5%)
	Mean (2nd)	34.35±5.02	30.21±7.66	<0.001 <sup>c</sup>	32.28	High (29.3%   70.7%) Low (58.3.5%   41.7%)
	Median (3th)	34.57±5.26	30.29±7.79	<0.001 <sup>c</sup>	32.43	High (29.3%   70.7%) Low (58.3.5%   41.7%)
	Root Mean Squared (4rd)	36.21±4.69	32.05±7.51	<0.001 <sup>c</sup>	34.13	High (26.8%   73.2%) Low (60.4%   39.6%)

**Notes:** <sup>1a</sup>M: the mean of the mean feature value of the high-grade group and the mean feature value of the low-grade group. <sup>1b</sup>Percentage of patients in the high and low Fuhrman grade groups with feature values less than or greater than the “M” value. <sup>c</sup>Independent t-test. <sup>d</sup>Mann–Whitney U-test.

grade. This may be ascribed to the theoretical ground that there is no direct association between the Fuhrman nuclear grade and microscopic tumor vascularity.

In addition, several previous studies<sup>25,30</sup> have demonstrated that tumor size was associated with the Fuhrman grade of ccRCC such that larger tumors tend to be scored with higher grades. Tumor size is readily accessible in clinics; however, small (<4 cm) ccRCCs are still occasionally detected in our routine clinical practices. In this study, we intentionally excluded this factor for unbiased evaluation while recruiting the validation patient cohort by matching the sizes (as well as other factors such as age and sex, see Table 1) of the two groups, since we sought to confirm whether hidden image patterns other than the tumor size exists and can be utilized for nuclear grades stratification. It would be more meaningful to predict the ccRCCs grades regardless of its tumor size but purely depending on the CT radiomics features barely captured by the naked eye. This strict inclusion criterion for patient enrollment might partially explain the relatively lower accuracies witnessed in the current study when compared with those reported results<sup>25,26</sup> where tumor sizes were not paired.

It was also interesting to find that the top-10 features were shape-based features (n=5) and first-order features (n=5). The shape-based features were consistent in all four different CT phases, and only the elongation feature exhibited a statistically meaningful difference ( $p=0.023$ ) between the two groups, where the low-grade group showed higher values than the high-grade group, indicating that tumors in the low-grade group tended to be more circle-like (non-elongated). For the first-order features, all five top-10 features, ie, the 10th percentile, 90th percentile, mean, median, and root mean squared, in the high-grade group had significantly higher values (all  $p<10^{-3}$ ). This finding suggested that the CT HUs, notably the high-intensity HUs (eg, the 90th percentile, with  $p<10^{-4}$ ), were critical in differentiating low-grade vs high-grade HUs. Interestingly, this macroscopic phenotype coincided with the microscopic nuclear grading based on the nucleus size and prominence of the high-density nucleolus in the tumor cells.<sup>7</sup> Nuclear grading also depends on the assessment of the morphological nuclear characteristics; however, at the macro scale, this microscopic heterogeneity might not necessarily be reflected in CT morphologically but instead is more likely to be presented as CT image intensity fluctuations.

We found that no texture features were high-scored as the top-10 features in this study, although previous investigations have shown potential in some texture features,

eg, entropy was claimed to be a good biomarker in differentiating low-grade vs high-grade.<sup>22,23,26</sup> However, these results were not conclusive but instead contradictory in these pioneer studies, eg, Shu et al<sup>26</sup> and Deng et al<sup>23</sup> both found that high entropy is associated with high-grade renal cancer; in contrast, Feng et al<sup>22</sup> claimed that the entropy of low-grade tumors is higher than that of high-grade tumors. We speculate that this inconsistency may be attributed to various internal heterogenous histological components (eg, intratumoral necrosis and vasculature) of ccRCC that might be disproportionately distributed between the low- and high-grade ccRCC cases from the abovementioned studies.

There were several limitations that need to be acknowledged for this retrospective study. First, the enrolled patient cohort was relatively small, and an independent testing set was not available. We did not reserve an independent testing set from the whole patient cohort because of the small sample size. Subdividing the patient data into a training/validation set and an independent testing set would further deteriorate the data scarcity problem since fewer data would be reserved for training. This study investigated which phasic CT features are more suitable for nuclear grading modeling; though building a classification model with possibly high classification accuracy was not our intention, evaluation on a larger patient scale is of greater interest and would definitely further confirm the findings here. Second, the radiomics features were only extracted from a representative axial slice rather than from the three-dimensional (3D) whole tumor. The 3D volumetric analysis may theoretically include more tumor information; however, it is also time-consuming, and manual identification of all volumetric tumor slices may potentially increase the measurement errors. Third, Fuhrman grading was employed in this study instead of the WHO/International Society of Urologic Pathology (ISUP) grading system, which is currently widely used and has a stricter magnification requirement for pathological image quality, eg, requiring 400× magnification for scoring Grade 1 and 2. However, some of the collected digital pathological images in this study were only with low magnification, and their corresponding original histological sections were either lost during patient data sources preservation and relocation, or were not accessible since the patient data were collected from the other institution. We have reevaluated all the ccRCC cases that had pathological

images qualified for both the Fuhrman and WHO/ISUP systems and found consistent grading results for most (129/137, 94%) of the cases.

## Conclusion

In conclusion, the current comprehensive study demonstrated the feasibility of applying machine learning-based quantitative CT texture analysis in the differential diagnosis of low-grade from high-grade ccRCC. Of all four dynamic enhanced CT phases, radiomics features extracted from the UP contribute most to differentiation, and one shape-based and five first-order-based features of the top-10 features from the UP may be used as potential quantitative imaging biomarkers.

## Acknowledgment(s)

We gratefully acknowledge all the members of Guangzhou First People's Hospital, for continuous assistance.

## Funding

This study has received funding from the National Natural Science Foundation of China (81971574, 81874216), the National Key Research and Development Program of China (2017YFC0112900), the Natural Science Foundation of Guangdong Province, P.R. China (2018A030313282), the Fundamental Research Funds for the Central Universities, SCUT (2018MS23), the Guangzhou Science and Technology Project, P.R. China (202002030268), Medical Science and Technology Research Project of Guangdong Province (A2019465).

## Disclosure

The authors report no conflicts of interest in this work.

## References

- Leibovich BC, Lohse CM, Crispin PL, et al. Histological subtype is an independent predictor of outcome for patients with renal cell carcinoma. *J Urol*. 2010;183(4):1309–1316. doi:10.1016/j.juro.2009.12.035
- Rioux-Leclercq N, Karakiewicz PI, Trinh QD, et al. Prognostic ability of simplified nuclear grading of renal cell carcinoma. *Am Cancer Soc*. 2007;109(5):868–874. doi:10.1002/cncr.22463
- Novara G, Martignoni G, Artibani W, Ficarra V. Grading systems in renal cell carcinoma. *J Urol*. 2007;177(2):430–436. doi:10.1016/j.juro.2006.09.034
- Zisman A, Pantuck AJ, Dorey F, et al. Improved prognostication of renal cell carcinoma using an integrated staging system. *J Clin Oncol*. 2001;19(6):1649–1657. doi:10.1200/JCO.2001.19.6.1649
- Ljungberg B, Bensalah K, Canfield S, et al. EAU guidelines on renal cell carcinoma: 2014 update. *Eur Urol*. 2015;67(5):913–924. doi:10.1016/j.eururo.2015.01.005
- Silverman SG, Israel GM, Trinh QD. Incompletely characterized incidental renal masses: emerging data support conservative management. *Radiology*. 2015;275(1):28–42. doi:10.1148/radiol.14141144
- Fuhrman SA, Lasky LC, Limas C. Prognostic significance of morphologic parameters in renal cell carcinoma. *Am J Surg Pathol*. 1982;6(7):655–664. doi:10.1097/00000478-198210000-00007
- Rini BI, Campbell SC, Escudier B. Renal cell carcinoma. *Lancet*. 2009;373(9669):1119–1132. doi:10.1016/S0140-6736(09)60229-4
- Lane BR, Samplaski MK, Herts BR, Zhou M, Novick AC, Campbell SC. Renal mass biopsy—a renaissance? *J Urol*. 2008;179(1):20–27. doi:10.1016/j.juro.2007.08.124
- Coy H, Young JR, Pantuck AJ, et al. Association of tumor grade, enhancement on multiphasic CT and microvessel density in patients with clear cell renal cell carcinoma. *Abdom Radiol*. 2019;44(1):180–189. doi:10.1007/s00261-018-1688-8
- Zhu YH, Wang X, Zhang J, Chen YH, Kong W, Huang YR. Low enhancement on multiphase contrast-enhanced CT images: an independent predictor of the presence of high tumor grade of clear cell renal cell carcinoma. *AJR Am J Roentgenol*. 2014;203(3):W295–300. doi:10.2214/AJR.13.12297
- Wei J, Zhao J, Zhang X, et al. Analysis of dual energy spectral CT and pathological grading of clear cell renal cell carcinoma (ccRCC). *PLoS One*. 2018;13(5):e0195699. doi:10.1371/journal.pone.0195699
- Zhang X, Wang Y, Yang L, et al. Delayed enhancement of the peritumoral cortex in clear cell renal cell carcinoma: correlation with Fuhrman grade. *Clin Radiol*. 2018;73(11):982.e1–982.e7. doi:10.1016/j.crad.2018.06.010
- Oh S, Sung DJ, Yang KS, et al. Correlation of CT imaging features and tumor size with Fuhrman grade of clear cell renal cell carcinoma. *Acta Radiol*. 2017;58(3):376–384. doi:10.1177/0284185116649795
- Choi SY, Sung DJ, Yang KS, et al. Small (<4 cm) clear cell renal cell carcinoma: correlation between CT findings and histologic grade. *Abdom Radiol*. 2016;41(6):1160–1169. doi:10.1007/s00261-016-0732-9
- Chen C, Kang Q, Wei Q, et al. Correlation between CT perfusion parameters and Fuhrman grade in pT1b renal cell carcinoma. *Abdom Radiol (NY)*. 2017;42(5):1464–1471. doi:10.1007/s00261-016-1009-z
- Palmowski M, Schifferdecker I, Zwick S, et al. Tumor perfusion assessed by dynamic contrast-enhanced MRI correlates to the grading of renal cell carcinoma: initial results. *Eur J Radiol*. 2010;74(3):e176–80. doi:10.1016/j.ejrad.2009.05.042
- Woo S, Suh CH, Kim SY, Cho JY, Kim SH. Diagnostic performance of DWI for differentiating high- from low-grade clear cell renal cell carcinoma: a systematic review and meta-analysis. *AJR Am J Roentgenol*. 2017;209(6):W374–W381. doi:10.2214/AJR.17.18283
- Mytsyk Y, Dutka I, Borys Y, et al. Renal cell carcinoma: applicability of the apparent coefficient of the diffusion-weighted estimated by MRI for improving their differential diagnosis, histologic subtyping, and differentiation grade. *Int Urol Nephrol*. 2017;49(2):215–224. doi:10.1007/s11255-016-1460-3
- Bektas CT, Kocak B, Yardimci AH, et al. Clear cell renal cell carcinoma: machine learning-based quantitative computed tomography texture analysis for prediction of Fuhrman nuclear grade. *Eur Radiol*. 2019;29(3):1153–1163. doi:10.1007/s00330-018-5698-2
- Lin F, Cui E, Lei Y, Luo L. CT-based machine learning model to predict the Fuhrman nuclear grade of clear cell renal cell carcinoma. *Abdom Radiol*. 2019;44(7):2528–2534. doi:10.1007/s00261-019-01992-7
- 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(1):6. doi:10.1186/s40644-019-0195-7
- 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. doi:10.1007/s00330-019-06260-2

24. Kocak B, Durmaz ES, Ates E, Kaya OK, Kilickesmez O. Unenhanced CT texture analysis of clear cell renal cell carcinomas: a machine learning-based study for predicting histopathologic nuclear grade. *AJR Am J Roentgenol.* 2019;W1–W8. doi:10.2214/AJR.18.20742
25. 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. doi:10.1016/j.ejrad.2018.04.013
26. Shu J, Tang Y, Cui J, et al. Clear cell renal cell carcinoma: CT-based radiomics features for the prediction of Fuhrman grade. *Eur J Radiol.* 2018;109:8–12. doi:10.1016/j.ejrad.2018.10.005
27. Huhdanpaa H, Hwang D, Cen S, et al. CT prediction of the Fuhrman grade of clear cell renal cell carcinoma (RCC): towards the development of computer-assisted diagnostic method. *Abdom Imaging.* 2015;40(8):3168–3174. doi:10.1007/s00261-015-0531-8
28. van Griethuysen J, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017;77(21):e104–e107. doi:10.1158/0008-5472.CAN-17-0339
29. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res.* 2002;16(16):321–357. doi:10.1613/jair.953
30. Ishigami K, Leite LV, Pakalniskis MG, Lee DK, Holanda DG, Kuehn DM. Tumor grade of clear cell renal cell carcinoma assessed by contrast-enhanced computed tomography. *Springerplus.* 2014;3(1):694. doi:10.1186/2193-1801-3-694

## Cancer Management and Research

Dovepress

### Publish your work in this journal

Cancer Management and Research is an international, peer-reviewed open access journal focusing on cancer research and the optimal use of preventative and integrated treatment interventions to achieve improved outcomes, enhanced survival and quality of life for the cancer patient.

The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/cancer-management-and-research-journal>