RESEARCH



The value of a dual-energy CT lodine map radiomics model for the prediction of collagen fiber content in the ccRCC tumor microenvironment

Zhongyuan Li^{1†}, Ning Wang^{2†}, Xue Bing², Yuhan Li³, Jian Yao², Ruobing Li⁴ and Aimei Ouyang^{2*}

Abstract

Background and purpose Renal cell carcinoma (RCC) is a heterogeneous group of cancers. The collagen fiber content in the tumor microenvironment of renal cancer has an important role in tumor progression and prognosis. A radiomics model was developed from dual-energy CT iodine maps to assess collagen fiber content in the tumor microenvironment of ccRCC.

Methods A total of 87 patients with ccRCC admitted to our hospital were included in this retrospective study. Among them, 59 cases contained large amounts of collagen fibers and 28 cases contained a small amount of collagen fibers. We established a radiomics model using preoperative dual-energy CT scan lodine map (IV) imaging to distinguish patients with multiple collagen fibers from those with few collagen fibers in the tumor microenvironment of ccRCC. We extracted features from dual-energy CT lodine map images to evaluate the effects of six classifiers, namely k-nearest neighbor (KNN), support vector machine (SVM), extreme gradient boosting (XGBoost), random forest (RF), logistic regression (LR), and decision tree (DT). The effects of the models built based on the dynamic and venous phases are also compared. Model performance was evaluated using quintuple cross-validation and area under the receiver operating characteristic curve (AUC). In addition, a clinical model was developed to assess the clinical factors affecting collagen fiber content.

Results Compared to KNN, SVM, and LR classifiers, RF, DT, and XGBoost classifiers trained with higher AUC values, with training sets of 0.997, 1.0, and 1.0, respectively. In the validation set, the highest AUC was found in the SVM classifier with a size of 0.722. In the comparative test of the active and intravenous phase models, the SVM classifier had the best effect with its validation set AUC of 0.698 and 0.741. In addition, there was a statistically significant effect of patient age and maximum tumor diameter on the collagen fiber content in the tumor microenvironment of kidney cancer.

Conclusion Radionics features based on preoperative dual-energy CT IV can be used to predict the amount of collagen fibers in the tumor microenvironment of renal cancer. This study better informs clinical prognosis and patient management. lodograms may add additional value to dual-energy CTs.

[†]Zhongyuan Li and Ning Wang are joint first authors.

*Correspondence: Aimei Ouyang 13370582510@163.com Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/ficenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords Renal cell carcinoma, Tumor microenvironment, Collagen fibers, Dual-energy CT, Radiomics; machine learning

Introduction

Renal cell carcinoma (RCC) is a heterogeneous group of cancers that includes many histological subtypes, of which clear cell histology is the most common subtype [1]. RCC is the seventh most common cancer-causing death worldwide, with 140,000 patients dying from this cancer each year [2-4]. The tumor microenvironment (TME) is the environment in which tumor cells grow and develop, which includes not only the tumor cells themselves but also surrounding cells such as immune cells, fibroblasts, and glial cells, as well as collagen fibers, interstitial cells, microvasculature and other biomolecules surrounding the tumor. The presence of these cells is critical to tumor development, treatment, and prognosis. In recent years, tremendous progress has been made in tumor cytology and molecular biology, and researchers have gained a deeper understanding of the relationship between tumors and their environment. These results not only help us understand tumorigenesis, development, and metastasis but also help us better diagnose, prevent and treat tumors. Tumor cells can interact with surrounding cells through multiple pathways, thus influencing the development and progression of cancer. In addition, nonmalignant cells in the tumor microenvironment play a key role in all stages of cancer development by stimulating cell proliferation. TME is a complex spatial network of interwoven extracellular matrix proteins, of which collagen is a major component of the ECM, and several studies have shown that abnormal aggregation of collagen fibers is highly correlated with tumor progression [5–7]. In addition, the fiber component also promotes the growth, infiltration, and distant metastasis of ccRCC cells and therefore plays a crucial role in the growth of ccRCC. Assessing the fiber content of the tumor microenvironment in ccRCC patients is essential for understanding tumor cell progression and prognostic treatment.

CT scan is commonly used as the primary imaging technique to detect kidney cancer. Conventional CT cannot accurately assess primary tumors due to its relatively low soft tissue resolution. Dual-energy CT imaging refers to the acquisition of data from two different energies of electrons through a single scan [8]. It has been pointed out that dual-source CT dual-energy scanning can solve the problem of conventional CT scans with more single data, and virtual flat-scan images as well as Iodine maps are obtained after corresponding software processing, and the advantage of this technique is that it can reduce motion artifacts and improve resolution [9]. Li et al. [10] improved the ability to identify tissue enhancement by comparing lesions' density and change characteristics on the IV when evaluating the contrast uptake of lesions. Ideograms allow for both qualitative analysis of iodine content and quantitative analysis of iodine concentration, and this post-processing technique provides useful information for differential diagnosis.

The concept of radiomics was formally introduced in 2012, and in recent years it has shown excellent performance in oncology applications. Currently, the application of radiomics in kidney cancer is focused on three aspects: (1) differentiation of benign and malignant renal tumors [11, 12]; (2) staging and grading [13, 14]; and (3) differentiation of different subtypes of kidney cancer [11, 15]. The application of radiomics with dual-energy CT IV imaging to assess the fibrous component in renal cell carcinoma is rare. This study aimed to develop radiomics models from dual-energy CT reconstructed Iodine map images for preoperative prediction of fibrous component content in patients with ccRCC, to improve the treatment and prognosis of patients with ccRCC. In this study, we investigated the feasibility of applying radiomics based on dual-energy CT iodine concentration images to assess the fibrous component content in renal cell carcinoma by combining clinical data (age, gender, maximum tumor diameter, nuclear grading, stage, presence of envelope and presence of necrotic areas).

Patients and methods Study cohort

This study was a retrospective study that was approved by the Ethics Committee of Shandong Qilu Hospital, whose ethics committee waived the requirement of informed consent. The study included 87 patients (60 males and 27 females, with a mean age of 57 years ± 10.94 years, ranging from 33 to 82 years) admitted to Shandong Qilu Hospital. Inclusion criteria were: (1) patients with pathologically confirmed RCC after partial or radical nephrectomy; (2) patients who obtained a complete dualenergy CT scan preoperatively. Exclusion criteria were: (1) patients with non-clear cell carcinoma; (2) patients receiving pre-operative radiotherapy and targeted therapy; and (3) patients with poor imaging due to other diseases including cardiac, hepatic, and renal insufficiency. We used the Radcloud Radiomics Platform (Huiying Medical Technology Co, https://mics.huiyihuiying.com/) to manage the imaging data, clinical data, and subsequent statistical analysis of radiomics. The validation

and training datasets were separated using a randomized method with a ratio of 3:7 and a random seed of 468 (Examples of groupings are shown in Table 1).

Clinical data

All patient data included in this study were collected from our hospital's integrated electronic medical record system. The clinical data included age, gender, maximum tumor diameter, nuclear grading, stage, envelope presence, and necrotic areas.

Study methods

CT scan method

The patient is flexed laterally, lying flat and resting, and is fixed in the middle of the scanning bed. The patient's toes are advanced laterally and the internal pressure line is positioned at the sternocarpal process and the horizontal line is positioned at the mid-axillary line. The contrast agent concentration (iohexol, 300 mg/ml) was $60 \sim 80$ ml, which was injected by a high-pressure syringe into the patient's forearm vein, with an average intravenous flow rate of about $3 \sim 5$ ml/s per minute. Then a double-phase enhancement scan of the dermatomedullary and parenchymal phases was performed in dual-energy mode. The delayed exposure time was adjustable to 30 and 80

Tab	le 1	Gro	uping	of	DT	С	lassifiers
-----	------	-----	-------	----	----	---	------------

Page 3 of 11

s, meanwhile, the tube flow voltage was adjustable to Sn100kVp and Sn150kVp. Adjustable tube flows a current of 130-180mAs and 80-90mAs with automatic delayed exposure system on.

Pathological evaluation

Samples are taken from the tumor site of the wax block. The tissue was cut into sections of 4 μ m thickness, and paraffin-sealed slices were made. Then patient sections were stained for Masson histopathology. Pathologist with 5 years of experience performing observations under a light microscope. Based on the pathological findings, they were divided into two groups: a large amount of collagen fibers and a small amount of collagen fibers. At the hotspot for collagen fibers, a small amount was designated if it covered less than half of the visual field, while a large amount referred to coverage greater than this threshold (Fig. 1).

Tumor segmentation

The images were uploaded to the Radcloud Radiomics Platform (http://mics.radcloud.cn/), and two radiologists with more than 1–2 years of experience manually outlined the entire lesion layer by layer on all consecutive levels of the tumor, and then radiologists with more

Clinical characteristics	Training (<i>n</i> =60)	Validation (n=27)	Chi-square value	P-Value
Mean age (year)	56.55	58.37	0.506	0.479
Gender (n)			0.036	0.529
Male	41	19		
Female	19	8		
Amounts of collagen fibers (n)			0.242	0.399
Large	41	17		
Small	19	10		
Mean maximum tumor diameter (cm)	4.688	4.459	0.224	0.637
Nuclear grading (n)			0.645	0.293
High	13	8		
Low	47	19		
Tumor stage (n)			5.175	0.159
I	36	21		
II	6	3		
III	9	3		
IV	9	0		
Envelope (n)			0.074	0.483
Presence	33	14		
Absence	27	13		
Necrotic areas (n)			< 0.001	0.601
Presence	40	18		
Absence	20	9		



Fig. 1 Masson stained pathological sections [The collagen fiber composition is shown after Masson staining (magnification × 200), and the blue area is collagen fiber.] and corresponding CT iodogram images. Negatives: A1 ~ 2 a 56-year-old male with biopsy-confirmed renal carcinoma clear cell carcinoma a large amount of collagen fibers; B1 ~ 2 a 63-year-old female with biopsy-confirmed renal carcinoma clear cell carcinoma with a small amount of collagen fibers

than 5 years of experience reviewed all the contours, and if the discrepancy was \geq 5%, the senior radiologist determined that the tumor boundary as [16]. After a week, 30 cases were randomly selected, and the two radiologists repeated the segmentation) to assess the inter- and intra- class correlation coefficients (ICCs). ICC > 0.75 was regarded as favorable robustness and reproducibility. We did a physician consistency check and found that there was good consistency in the division of the ROIs between the two physicians.

Feature extraction and selection

Two thousand eight hundred eighteen quantitative imaging features were extracted from CT images by the Radcloud platform (http://radcloud.cn/), including 1409 features each in the arterial and venous phases. These features can be divided into three groups. The first category (first-order statistics) includes 252 descriptors that quantify the distribution of voxel intensities in CT images using commonly used basic metrics. The second category (shape and size-based features) includes 28 3D features that reflect the region's shape and size. Based on the gray run length and gray co-occurrence texture matrix calculations, 1050 texture features that can quantify the heterogeneous differences of the region are grouped into the third group (texture features). In summary, a large number of image features can be computed. However, all these extracted features may not be useful for a specific task. Therefore, downscaling and selecting taskspecific features for optimal performance are necessary steps. Feature selection methods that reduce redundant features include variance thresholding (variance threshold = 0.8), SelectKBest, and the least absolute shrinkage selection operator (LASSO). Since the threshold value of the variance threshold method is 0.8, the feature values with variance less than 0.8 are removed. The select best method is a univariate feature selection method that analyzes the relationship between features and classification results using *p*-values, as all features with *p*-values less than 0.05 are used. The L1 regularizer is used as the cost function in the LASSO model, the cross-validation error value is 3, and the maximum number of iterations is 1000.

Machine learning model building

After feature characterization, a total of 2818 features were identified as significantly related to the topic. Based on the selected features, the radiomics score (Rad-score) was computed individually for each patient through a mathematical formula that incorporates a selection of *n* radiomics signatures. The Rad-score formula was derived as follows: Rad-score= $\alpha + n \sum_{i=1}^{n} \beta_i X_i$, and α represents the intercept, β_i denotes the value of radiomics feature; X_i represents the corresponding coefficient.

Several supervised learning classifiers are available for classification analysis, which create models that attempt to separate or predict data related to outcomes or phenotypes (e.g., patient outcomes or responses). In this study, six classifiers, k-nearest neighbor (KNN), support vector machine (SVM), extreme gradient boost (XGBoost), random forest (RF), logistic regression (LR), and decision tree (DT), were used to construct radiomics-based models, and validation methods were used to improve the validity of the models and to compare the variability and superiority of the kinetic-venous phase models.

Statistical methods

Statistical analysis was performed on the Radcloud platform. In the training and validation datasets, we used subject operating characteristic (ROC) curves, or area under the curve (AUC), to assess the performance of the predictions. Four metrics, P [accuracy=true positives/ (true positives+false positives)], R [check-all rate=true positives/(true positives+false negatives)], f1-score [f1-score=P*R*2/ (P+R)], and support (total number of test sets), were employed in this study to assess how well the classifiers performed.

SPSS 25.0 (IBM) was used to analyze the above data statistically, and differences were considered statistically significant at P < 0.05. The clinical data were analyzed by chi-square test and multiparametric regression analysis. All statistical tests were performed using one-sided and two-sided tests.

Results

Clinical characteristics

A total of 87 patients (60 males as well as 27 females; mean age 57 years \pm 10.94; age range 33 to 82 years) were included in this study, and the number of patients with multiple collagen fibers was 59 (68.0% of cases) and 28 (32.0% of cases) with small amount collagen fibers. The maximum diameter of the tumor was obtained on the transverse axis of the image images, and its size ranged from 1.2 cm to 11.0 cm. In the training cohort, there were

significant differences (P < 0.05) in patient age and tumor maximum diameter between the content of multiple collagen fibers and the content of small amount of collagen fibers (Tables 2 and 3), and the rest of the gender, nuclear grading, tumor stage, presence or absence of envelope and presence or absence of necrotic areas were not significant (P > 0.05) (Table 4).

Classification results

Radiomic features with good intra- and interobserver reproducibility (ICC > 0.75) were selected from all 2818. Then we filtered 892 features from the features using the variance threshold method, then 11 features using the best selection method K. Finally, 7 optimal features were screened by the LASSO algorithm (Table 5 and Fig. 2). Using the same screening method, 444 and 445 features were screened out of 1409 features in the arterial and venous phases, respectively, and the best method K was selected to screen 3 and 4 features, respectively, and the optimal features were both screened as 3 using the LASSO algorithm (Table 6).

Figure 3 shows the classification results and the corresponding ROC curves of the six classifiers for the different feature sets extracted from the images. Compared to KNN, SVM, and LR classifiers, the RF, DT, and XGBoost classifiers trained with higher AUC values of 0.997, 1.0, and 1.0 for the training set, respectively. In

Table 3Relationship between patient age and maximum tumordiameter and collagen fiber content

Model	Model fitting conditions	Likelihood ratio test				
	-2 log likelihood of the simplified model	chi-square test	Degree of freedom	Significance		
Intercept distance only	110.172					
The final	6.945	103.227	75	.017		

Table 2 Relationship between patient age, maximum tumor diameter, and collagen fiber content

Effect	Model fitting conditions	Likelihood ratio test		
	-2 log likelihood of the simplified model	chi-square test	Degree of freedom	Significance
Intercept	6.945 ^a	.000	0	
Age	54.938	47.993	29	.015
Maximum tumor diameter	69.986	63.041	41	.015

The chi-square statistic is the difference in -2 log likelihood between the final model and the simplified model. The simplified model is formed by omitting an effect from the final model. The original assumption is that all parameters of the effect are 0

^a Since omitting this effect does not increase the degrees of freedom, this simplified model is equivalent to the final model

			0		
Clinical Characteristics	Large (n = 59)	Small (n = 28)	Total (n=87)	Chi-square value	P-Value
Gender				0.421	0.346
Male	38	21	59		
Female	20	8	28		
Nuclear Grading				1.130	0.215
High	16	5	21		
Low	42	24	66		
Tumor Stage				0.629	0.890
I	36	21	57		
II	6	3	9		
III	9	3	12		
IV	6	3	9		
Envelope				0.023	0.530
Presence	31	15	46		
Absence	27	14	41		
Necrotic Areas				0.026	0.528
Presence	39	19	58		
Absence	19	10	29		

Table 4 Relationship between patients' clinical characteristics and collagen fibre content (p-value) (n = 87)

Table 5 Description of the selected radiomics features with their associated feature group and filter

Radiomic feature	Radiomic class	Filter	Coefficient
LargeDependenceEmphasis	gldm	V IMDI_original	-0.921
DependenceEntropy	gldm	V IMDI_wavelet-LLL	-0.382
LargeDependenceEmphasis	gldm	V IMDI_wavelet-LLL	-0.112
InterquartileRange	firstorder	V IMDI_squareroot	-0.083
DependenceVariance	gldm	V IMDI_original	-0.024
DependenceVariance	gldm	V IMDI_squareroot	0.112
DependenceEntropy	gldm	V IMDI_logarithm	0.322

Label: GLDM = Gray Level Dependence Matrix

the validation set, the highest AUC was found in the SVM classifier with a size of 0.722, while the RF, DT, and XGBoost classifiers also achieved good data with AUCs of 0.639, 0.694, and 0.639, respectively. In the dynamic and venous phase model comparison tests, the validation set AUCs were 0.698 and 0.741 for the SVM classifier, and the LR classifier The validation set AUCs were 0.432 and 0.500, respectively, so the AUCs of the models based on the venous phase was higher (Figs. 4 and 5).

We summarize the metrics of the six classifiers, precision, recall, f1 score, and support, in Tables 7 and 8 for the arterial and venous phases and all images, respectively. From the features extracted on all images, the more effective classifiers were RF, DT, and XGBoost, whose four categories of metrics in the training set were 0.90, 0.95, 0.92, 19, 1.00, 1.00, 1.00 19, and 1.00, 0.95, 0.97, 19, respectively. Except for this, the more effective classifiers for the model in both the arterial and venous phases were SVM.



Fig. 2 A 892 features were selected from 2818 using the variance threshold method (variance threshold = 0.8). B 11 features were further selected using the best method Select K. C ~ E 7 features corresponding to the best alpha values were selected using the Lasso model

Discussion

In this study, we developed a radiomics model from dualenergy CT Iodine map images for assessing collagen fiber content in the tumor microenvironment of kidney cancer. The results showed that there was an association between the radiomics features in renal dual-energy CT imaging and the collagen fiber content in the tumor microenvironment of kidney cancer, especially when trained using the SVM classifier, the model showed good predictive performance, and AUC is 0.722. When the arterial and venous phase model was compared, the venous phase model was more effective. Meanwhile, the clinical model developed in this study showed a significant correlation between collagen fiber content in the tumor microenvironment of kidney cancer and age and maximum tumor diameter.

In acquiring impact images of ccRCC patients, dualenergy CT was chosen for this study. Dual-energy CT takes advantage of the different X-ray attenuation values of different substances at different energies, such as iodine (Z=53) and calcium (Z=20), and acquires and analyzes images at different energies according to the different slopes of changes in attenuation values of these two substances. Currently, dual-energy CT has been applied in a large number of studies, and Han Bo [17] applied dual-source CT to the study of renal occupancies and found that dual-source CT increased the mean CT value. The results were in good agreement with pathology and significantly reduced radiation dose. Besides, in this study, the arterial and venous phase models were established separately and compared, and then it was concluded that the venous phase model has a higher AUC. Its corresponding classifier is more effective in the cases selected in this study. As a result, it is more convincing in assessing the collagen fiber content in the tumor microenvironment of kidney cancer.

In recent years, researchers have found in studies of liver, oesophageal, breast, and kidney cancers that radiomics has proven to be an effective tool for assessing information within the tumor microenvironment [18, 19]. At present, studies using imaging histology to assess the immune microenvironment and genes in kidney cancer have emerged. For example, Lianghong Jiao et al. [20] used imaging histology to explore the clinical and immune features of ccRCC associated with IL-23 expression levels and to develop a preoperative prediction

	Indicators	KNN	SVM	XGBoost	RF	LR	DT
	Precision	0.58/0.44	0.3/0.5	1/0.42	0.95/0.62	0.41/0.23	1/0.31
	Sensitivity	0.58/0.44	0.47/0.78	0.95/0.56	1/0.56	0.58/0.33	1/0.56
Arterial phase	F1-score	0.58/0.44	0.37/0.61	0.97/0.48	0.97/0.59	0.48/0.27	1/0.4
	Support	19/9	19/9	19/9	19/9	19/9	19/9
venous phase	Precision	1/0.67	0.86/0.62	1/0.57	0.95/0.33	0/0	1/0.5
	Sensitivity	1/0.22	0.95/0.56	0.95/0.44	1/0.22	0/0	1/0.33
	F1-score	1/0.33	0.9/0.59	0.97/0.5	0.97/0.27	0/0	1/0.4
	Support	19/9	19/9	19/9	19/9	19/9	19/9

Table 6 The results of Precision, Sensitivity, F1-score, and Support in training/validation cohorts of the arterial phase and venous phase



Fig. 3 ROC curves of six methods (KNN, SVM, XGBoost, RF, LR and DT). A1 ~ F1 ROC curves for the training set in a small amount of collagen fiber group; A2 ~ F2 ROC curve of the validation set in a small amount of collagen fiber group

model based on contrast CT scans. Based on the presence of differentially expressed metabolism-related prognostic genes and immune-related components, Wang Yi et al. [21] initially distinguished two distinct metabolic subtypes (C1 and C2 subtypes) and immune subtypes (I1 and I2 subtypes). Gao et al. [22] successfully classified patients into different subtypes based on gene expression levels in the tumor microenvironment (TME), and novel prognostic radiogenetics biomarkers correlated well with the immune-related gene expression status of ccRCC patients and could successfully stratify the survival status of patients in the TCGA database. However, it is rare to study the collagen fiber content of the tumor microenvironment using radiomics. Therefore, in this study,



Fig. 4 ROC curves of six methods in the Arterial phase (KNN, SVM, XGBoost, RF, LR and DT). A1 ~ F1 ROC curves for the training set in a small amount of collagen fiber group; A2 ~ F2 ROC curve of the validation set in a small amount of collagen fiber group



Fig. 5 ROC curves of six methods in the venous phase (KNN, SVM, XGBoost, RF, LR and DT). A1~F1 ROC curves for the training set in a small amount of collagen fiber group; A2~F2 ROC curve of the validation set in a small amount of collagen fiber group

it is a valuable attempt for us that we used radiomics to assess the content of collagen fibers in the tumor microenvironment of kidney cancer and in the future, further evaluation of the complex information of the tumor microenvironment is more necessary to make better clinical decisions in the era of precision medicine.

This study assessed collagen fiber levels in the tumor microenvironment of kidney cancer using radiomics. Collagen is the most important component of the ECM and the most abundant protein in human tissues, with 28 unique isoforms having been identified [23–25]. Collagen fibers are located in the extracellular matrix and have important roles in tissue scaffolding, cell adhesion, cell migration, angiogenesis, tissue morphogenesis, and tissue recovery. Emerging collagen fibers can directly establish an invasive pathway for matrix metalloproteinase resistance to promote metastasis, and their density can facilitate macromolecular transport to alter renal cancer cell metabolism, inhibit transformed immune cell function, and promote gene expression [7]. In addition, collagen fibers stimulate fibroblast production and cross-linking between fibers, while at the same time increasing tissue fibrosis and stiffness to promote invasion and metastasis of kidney cancer cells. Also in the kidney, collagen plays a key function in branching morphogenesis, a process that involves the invasion of epithelial buds and tubes into the

Table 7 Description of the selected radiomics features with their associated feature group and filter	
--	--

	Radiomic feature	Radiomic class	Filter	Coefficient
	HighGrayLevelZoneEmphasis	glszm	wavelet-HLL	0.049
Arterial	HighGrayLevelZoneEmphasis	glszm	wavelet-LHL	0.049
phase	SmallAreaHighGrayLevelEmphasis	glszm	wavelet-HLL	0.074
	ZoneEntropy	glszm	wavelet-HHL	-0.065
venous	ZoneEntropy	glszm	wavelet-LLH	0.092
phaod	RunVariance	glrlm	original	0.142

Label: GLSZM = Gray-Level Size Zone Matrix

Label: GLRLM = Gray Level Run Length Matrix, GLSZM = Gray-Level Size Zone Matrix

Table 8	The results of	f Precision, Sensitivity,	F1-score, and	Support in tra	aining cohorts and '	validation cohorts
---------	----------------	---------------------------	---------------	----------------	----------------------	--------------------

	Indicators	KNN	SVM	XGBoost	RF	LR	DT
	Precision	0.520	0.520	1.000	0.900	0.460	1.000
terining only of	Sensitivity	0.790	0.740	0.950	0.950	0.630	1.000
training cohort	F1-score	0.620	0.610	0.970	0.920	0.530	1.000
	Support	19	19	19	19	19	19
validation cohort	Precision	0.360	0.530	0.500	0.500	0.110	0.620
	Sensitivity	0.560	1.000	0.330	0.440	0.110	0.560
	F1-score	0.430	0.690	0.400	0.470	0.110	0.590
	Support	9	9	9	9	9	9

surrounding extracellular matrix-rich mesenchyme [26]. Therefore, preoperative assessment of collagen fibers in the tumor microenvironment of kidney cancer patients by non-invasive means is an important study for clinicians, and dual-energy CT IV radiomics by assessing and monitoring tumor characteristics (e.g., temporal and spatial heterogeneity) can achieve in-depth interpretation of information on the tumor microenvironment of kidney cancer and be used for clinical diagnosis and prognosis.

There are some limitations of the study. First of all, since this is a single-center study with a small

sample size, more research is required to assess the model and findings using a larger data set. Second, rather than three-dimensional CT scans, the radiomics model uses two-dimensional images, necessitating further evaluation of its performance using threedimensional data. Additionally, user dependence and variability may be introduced since tumor segmentation is done manually. In future work, fully automated segmentation may become a reality. Finally, this work is a retrospective study, so selection bias cannot be completely avoided.

Conclusion

In conclusion, preoperative models based on radiomics features of dual-energy CT IV can predict collagen fiber content in the tumor microenvironment of kidney cancer, Radiomics models, as opposed to the conventional visual evaluation of images, can provide a tool to help assessing the collagen fiber content, better informing clinical prognosis and patient management. Further evaluation of our findings on a large dataset will be necessary for future work.

Acknowledgements

We thank for all study participants who were enrolled in this study. We thank Radcloud Radiomics Platform for its help with research.

Authors' contributions

Zhongyuan Li wrote the main manuscript text. Ning Wang prepared Fig. 1. Xue Bing prepared Figs. 2, 3, 4, 5 and tables. Yuhan Li, Jian Yao, Ruobing Li were involved in the experimental procedure. Aimei Ouyang Provided guidance. All authors reviewed the manuscript.

Funding

This article is supported by Jinan Science and Technology Plan (Post subsidy) Project (No. 202019036). The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

This study was a retrospective study and was approved by the ethics committee of Central Hospital Affiliated to Shandong First Medical University. The requirement for informed consent was waived by the ethics committee of Central Hospital Affiliated to Shandong First Medical University. Authors confirm that all methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests The authors declare no competing interests.

Author details

¹School of Medical Imaging, Weifang Medical University, No. 7166, Baotong West Street, Weifang, Shandong 261053, P. R. China. ²Department of Radiology, Central Hospital Affiliated to Shandong First Medical University, No. 105 Jiefang Road, Jinan, Shandong 250013, P. R. China. ³Department of Radiology, Longkou Traditional Chinese Medical Hospital, Yantai, Shandong 265700, P. R. China. ⁴Shandong First Medical University, No. 105 Jiefang Road, Jinan, Shandong 250013, P. R. China.

Received: 18 January 2023 Accepted: 16 October 2023 Published online: 15 November 2023

References

2018;68(6):394-424.

- 1. Motzer RJ, Jonasch E, Boyle S, et al. NCCN Guidelines Insights: Kidney Cancer, Version 1.2021. J Natl Compr Canc Netw. 2020;18(9):1160–70.
- Capitanio U, Montorsi F. Renal cancer. Lancet. 2016;387(10021):894–906.
 Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin.

- Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209–49.
- Caron JM, Ames JJ, Contois L, et al. Inhibition of ovarian tumor growth by targeting the HU177 cryptic collagen epitope. Am J Pathol. 2016;186(6):1649–61.
- 6. Chen Y, Terajima M, Yang Y, et al. Lysyl hydroxylase 2 induces a collagen cross-link switch in tumor stroma. J Clin Invest. 2015;125(3):1147–62.
- Li ZM, Qiu WF, Zhang JM, et al. Advances in the study of collagen changes in the tumor microenvironment and tumor metastasis. Chongqing Medicine. 2018;47(16):2206–9.
- Ning SY, Shang NJ, Jiang ZY, et al. Application of dual-source CT dualenergy scanning in the diagnosis of benign and malignant pulmonary lesions. Modern Oncol. 2021;29(09):1593–6.
- Chi YE. A comparative study of the diagnostic value of enhanced CT enhanced MRI and ultrasonography in the diagnosis of occupying renal lesions. M M I Bimonthly. 2017;26(05):1293–5.
- Li Y, Xue R, Lai Q, et al. Endoscope-assisted resection of nonneoplastic space-occupying lesion in oral and maxillofacial areas. Sci Rep. 2017;7(1):16920.
- Coy H, Hsieh K, Wu W, et al. Deep learning and radiomics: the utility of Google TensorFlow[™] Inception in classifying clear cell renal cell carcinoma and oncocytoma on multiphasic CT. Abdom Radiol (NY). 2019;44(6):2009–20.
- 12. Zhou L, Zhang Z, Chen YC, Zhao ZY, Yin XD, Jiang HB. A deep learningbased radiomics model for differentiating benign and malignant renal tumors. Transl Oncol. 2019;12(2):292–300.
- 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;103:51–6.
- Shu J, Wen D, Xi Y, et al. Clear cell renal cell carcinoma: Machine learningbased computed tomography radiomics analysis for the prediction of WHO/ISUP grade. Eur J Radiol. 2019;121: 108738.
- Yin Q, Hung SC, Rathmell WK, et al. Integrative radiomics expression predicts molecular subtypes of primary clear cell renal cell carcinoma. Clin Radiol. 2018;73(9):782–91.
- Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48(4):441–6.
- 17. Han B. CT diagnosis analysis of renal occupied lesions. China Modern Doctor. 2017;55(19):96–8.
- Moon SH, Kim J, Joung JG, et al. Correlations between metabolic texture features, genetic heterogeneity, and mutation burden in patients with lung cancer. Eur J Nucl Med Mol Imaging. 2019;46(2):446–54.
- Kim JH, Ko ES, Lim Y, et al. Breast cancer heterogeneity: MR imaging texture analysis and survival outcomes. Radiology. 2017;282(3):665–75.
- Jiao L, Wang X, Ying Y, Yang K. The prognosis and immune microenvironment related factor interleukin-23 in clear cell renal cell carcinoma: a radiological investigation. Urol J. 2022;19(5):363–70.
- 21. Wang Y, Zheng XD, Zhu GQ, et al. Crosstalk between metabolism and immune activity reveals four subtypes with therapeutic implications in clear cell renal cell carcinoma. Front Immunol. 2022;13: 861328.
- 22. Gao J, Ye F, Han F, Jiang H, Zhang J. A radiogenomics biomarker based on immunological heterogeneity for non-invasive prognosis of renal clear cell carcinoma. Front Immunol. 2022;13: 956679.
- Myllyharju J, Kivirikko KI. Collagens, modifying enzymes and their mutations in humans, flies, and worms. Trends Genet. 2004;20(1):33–43.
- 24. Ricard-Blum S. The collagen family. Cold Spring Harb Perspect Biol. 2011;3(1): a004978.
- Mouw JK, Ou G, Weaver VM. Extracellular matrix assembly: a multiscale deconstruction. Nat Rev Mol Cell Biol. 2014;15(12):771–85.
- 26. Walker C, Mojares E, Del Río HA. Role of extracellular matrix in development and cancer progression. Int J Mol Sci. 2018;19(10):3028.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.