DOI: 10.38061/idunas.1084748

IDUNAS	NATURAL & APPLIED SCIENCES JOURNAL	2022 Vol. 5 No. 1 (29-37)

Predicting Kidney Tumor Subtype from Ct Images Using Radiomics and Clinical Features

Research Article

Duygu Şirin[1], Albert Güveniş[1*]

[1]Boğaziçi University, Institute of Biomedical Engineering, Istanbul/Turkey

Author E-mails
guvenisboun.edu.tr
*Correspondance to: Albert Güveniş, Boğaziçi University, Institute of Biomedical Engineering, Istanbul/Turkey
DOI: 10.38061/idunas.1084748

Received: 13.03.2022; Accepted: 29.06.022

Abstract

Purpose: This study aims to evaluate the performance of machine learning methods in predicting the subtype (clear-cell vs. non-clear-cell) of kidney tumors using clinical patient and radiomics data from CT images.
Method: CT images of 192 malignant kidney tumor cases (142 clear-cell, 50 other) from TCIA's KiTS-19 Challenge were used in the study. There were several different tumor subtypes in the other group, most of them being chromophobe or papillary RCC. Patient clinical data were combined with the radiomic features extracted from CT images. Features were extracted from 3D images and all of the slices were included in the feature extraction process. Initial dataset consisted of 1157 features of which 1130 were radiomics and 27 were clinical. Features were selected using Kruskal Wallis – ANOVA test followed by Lasso Regression. After feature selection, 8 radiomic features remained. None of the clinical features were considered important for our model as a result. Training set classes were balanced using SMOTE. Training data with the selected features were used to train the Coarse Gaussian SVM and Subspace Discriminant classifiers.

Results: Coarse Gaussian SVM was faster compared to Subspace Discriminant with a training time of 0.47 sec and ~11000 obs/sec prediction speed. Training duration of Subspace Discriminant was 4.1 sec with ~960 obs/sec prediction speed. For Coarse Gaussian SVM, the AUC was found to be 0.86, while for Subspace Discrimination it was 0.85.

Conclusion: Both models produced promising results on classifying malignant tumors as ccRCC or non-ccRCC.

Keywords: Kidney Tumor, Clear-cell, Machine Learning, CT imaging.

1. INTRODUCTION

More than 400 000 patients are diagnosed with kidney cancer each year, more than 90% of them being renal cell carcinoma (RCC). RCC is known to be the most common type (75%) of kidney cancer as well having the highest mortality rate among genitourinary cancers. It has more than 10 histological and molecular subtypes and one of them is clear cell RCC (CCRCC) (Hsieh et al. 2017) .

Currently, a great effort is being out into the studies concerning how different kidney tumor morphology might affect the treatment process. Surgery, chemotherapy and targeted drugs are used for treatment and a variety of new, more effective drugs are continued to be developed. There has been a big improvement in the median survival of the disease in the past few years, thanks to the targeted drug development (Le and Hsieh 2018).

As for the diagnosis of renal cancer; laboratory tests, radiology and biopsy are used. Presently, biopsy is obligatory in order to deduce key information as whether the cancer is invasive, its grade, its stage and spread to lymph nodes. In addition to these, biopsy must be performed to identify specific proteins, genes and other factors which are unique to the tumor. These factors play an important role in prognosis prediction and in the construction of a treatment plan. They can also provide a clearer view on how to design a more effective drugs targeting specific intracellular pathways (Ökmen, Guvenis, and Uysal 2019).

Biopsy is a highly invasive diagnosis tool which carry a small risk of infection and bleeding. Moreover, for cancer cases, reaching a result might take several days. For these reasons, there is a need for obtaining the necessary information for diagnosis by using better tools.

Computed Tomography (CT) is widely used for clinical diagnosis, localization of pathology, observation of anatomical structure, surveillance of therapy evolution and planning the optimal treatment in cancer. CT is generally the first choice for imaging the evolution of renal tumor because it is more available than Magnetic Resonance Imaging (MRI) and it more useful than Ultrasound (US) Imaging (van Oostenbrugge, Fütterer, and Mulders 2018).

An emerging field, Radiomics, has the ability to provide many advantages in cancer imaging. It focuses on obtaining quantitative information from clinical images, which helps characterize the image phenotypes of the tumor in a more detailed way. Radiomics is concerned with reaching useful diagnostic, prognostic and predictive information.

One of the the main objectives of the presented work is to perform a comparison among existing machine learning methods for the classification of tumor histologic subtypes of renal cell carcinoma (RCC) patients. It also focuses on the question of which radiomic features and patient clinical data provide meaningful information about the histologic subtypes.

2. LITERATURE REVIEW

In a 2019 study (Han, Hwang, and Lee 2019), Convolutional Neural Networks were used to classify the tumor histologic subtypes of RCC cases. The data set included clear-cell, chromophobe and papillary subtypes. The model was fed with three-phase CT images and one slice from each phase was used. AUC values for differentiating clear-cell from non-clear-cell, papillary from non-papillary and chromophobe from non-chromophobe were; 0.93, 0.91 and 0.88 respectively.

Another paper by Kocak et al. focused on classifying the tumors as ccRCC or Non-ccRCC, as well as differentiation of ccRCC, pcRCC and chcRCC from each other. Artificial Neural Networks classifier predicted the subtypes as ccRCC or non-ccRCC with an AUC of 0.92, while the AUC of Support Vector Machine classifier was 0.79. Both of them performed worse in the three-class models (Kocak et al. 2018).

Zhnag et al. evaluated several models incorporating SVMs for classifying tumors as ccRCC or non-ccRCC, and chromophobe or papillary RCC. Slices with the largest cross-sectional area of the lesion from 3-phase CT images were used. Top 3 features were selected by Mann-Whitney U-tests, ROC curves and Pearson's correlation coefficient methods. An SVM with a nonlinear radial basis function kernel was

implemented. Best results were achieved using the corticomedullary phase images. AUC for ccRCC vs. non-ccRCC classification was 0.94 (Zhang et al. 2019).

Hoang et al., conducted a study which used random forest models for three classifications: oncocytoma vs. RCC, oncocytoma vs. ccRCC and papillary vs. ccRCC. Three consecutive slices containing the largest cross-sectional area from each of the four phases of MR images of 142 lesions from 41 patients were included. Pairwise Wilcoxon rank test, modified false discovery rate adjustment, Lasso regression were used for feature selection. ccRCC cases were distinguished from oncocytomas with an average accuracy of 77,9% (Hoang et al. 2018).

3. MATERIALS AND METHODS

Data Sets

CT images and patient clinical data from the Climb 4 Kidney Cancer-Kidney Tumor Segmentation Challenge (C4KC-KiTS) database (Clark et al. 2013; Heller et al. 2019) were acquired. 210 patients were included in the C4KC-KiTS database. In this study, 192 of the cases which had malignant tumors were used.

Image Pre-processing

Resampling, intensity normalization and gray level discretization were applied before starting the feature extraction process. Images had different slice thicknesses (0.5 mm to 5 mm) and different pixel sizes (0.438 mm to 1.04 mm). After reconstruction and resampling, 1 mm × 1 mm × 1 mm spatial resolution was achieved. Python software was used to perform resampling, and the new values of the resampled images were obtained by Cubic B-Spline interpolation method (Wang et al. 2011). Z-score normalization was used for the normalization of image intensity values. For gray level discretization, bin width was adjusted to be 0.01 on 3D Slicer software(Fedorov et al. 2012). Gray level discretization lessens the heterogeneity influences on the images, resulting from acquisition and reconstruction protocols(Larue et al. 2017).

Feature Extraction

Radiologic images carry relevant and significant clinical information (Tomaszewski et al. 2021). Feature extraction is an important step for finding the link between disease and image attributes, on the grounds that its enablement to obtain solid, quantitative representations.

Features were extracted using PyRadiomics extension on 3D Slicer. Three types of images were subject to feature extraction: original, Laplacian of Gaussian (LoG) and wavelet-transformed. Laplacian of Gaussian filter values were 2 mm, 4 mm, and 6 mm in order to explain patterns with various sizes.

After all radiomic features were extracted, certain patient clinical data were added to get a combined data set. Clinical data included information such as age, sex, body mass index; as well as presence of several diseases, alcohol and tobacco use. Afterwards, the combined data were split into training and test sets as 85% and 15% respectively. As a result, 162 training cases included 128 clear-cell RCC and 34 other histologic subtypes (chromophobe, papillary, clear-cell papillary, multilocular cystic, urothelial, wilms). Further, 15 of the test cases were clear-cell and the remaining 15 were other (chromophobe, papillary, clear-cell papillary).

Feature Selection

Feature selection process was executed on Matlab (R2021b) software. Kruskal-Wallis (KW) test was conducted as the first step of feature selection. KW compares the medians of the groups of data to

determine if the samples come from the same population. In this case, each feature was tested for its ability to differentiate between the data classified as clear-cell and other with p = 0.05. Only 111 features among the 1157 were decided to be relevant. Moreover, none of the patient clinical features were selected.

Secondly, least absolute shrinkage and selection Operator (LASSO) was used at the next phase for selecting features. Lasso is an improved version of ordinary least squares estimates in regression analysis combining subset selection and ridge regression (Tibshiranit 1996). It causes some regression coefficients to shrink and set some of them to zero. At the end, coefficients belonging to the less important features become zero. Lambda with the minimum standard error was chosen to obtain the optimal set of coefficients. Subsequently, 8 features were selected as the most relevant for our model.

Model Training and Evaluation

Coarse Gaussian Support Vector Machine and Subspace Discriminant classifiers were trained with the selected features in the Classification Learner App of Matlab. SVM classifier aims to find the optimal hyperplane in the N-dimensional space that distinctly classifies the data points. The optimal hyperplane can be described as the most distant of all possible ones to both classes. The data points closest to this hyperplane are defined as support vectors. In case simple hyperplanes do not show sufficient separation performance. Hence, kernels are reproduced which use several functions. In this study, the SVM classifier with a Gaussian (radial basis fuction) kernel was used. Box constraint level was 1 and kernel scale was chosen as 11 for the classifier. 5-fold cross validation was used in the training process.

Discriminant classifiers assume that different classes have different Gaussian distributions. Their objective is to classify the data points while minimizing the classification cost. Ensemble learning combines several classifiers to improve the prediction performance. Each learner, discriminant classifier, is trained using a random subset of features among the selected ones. At the end, the best model is introduced. The model included 30 learners with 4 subspaces and the training was performed using 5-fold cross validation. Previously reserved test data set was used to test the model performances. Accuracy and area under the curve (AUC) for both models were calculated for evaluation.

4. RESULTS

Feature Extraction and Selection

Radiologic images carry relevant and significant clinical information (Tomaszewski et al. 2021). Feature extraction is an important step for finding the link between disease and image attributes, on the grounds that its enablement to obtain solid, quantitative representations.

Features were extracted using PyRadiomics extension on 3D Slicer. Three types of images were subject to feature extraction: original, Laplacian of Gaussian (LoG) and wavelet-transformed. Laplacian of Gaussian filter values were 2 mm, 4 mm, and 6 mm in order to explain patterns with various sizes.

After all radiomic features were extracted, certain patient clinical data were added to get a combined data set. Clinical data included information such as age, sex, body mass index; as well as presence of several diseases, alcohol and tobacco use. Afterwards, the combined data were split into training and test sets as 85% and 15% respectively. As a result, 162 training cases included 128 clear-cell RCC and 34 other histologic subtypes (chromophobe, papillary, clear-cell papillary, multilocular cystic, urothelial, wilms). Further, 15 of the test cases were clear-cell and the remaining 15 were other (chromophobe, papillary, clear-cell papillary).

In addition to the 27 clinical features which can be seen on Table 1, 1130 radiomic features were extracted from the CT images, adding up to 1157 features in total. Among the radiomic features; 744 were from wavelet-transformed images with 8 distinct filters and 6 classes of features (first-order, gray level

dependence matrix (gldm), gray level co-occurrence matrix (glcm), gray level run-length matrix (glrlm), gray level size zone matrix (glszm) and neighboring gray tone difference matrix (ngtdm)). Laplacian of Gaussian (LoG) filtered images produced 279 features, while 107 features were extracted from original images.

Table 1. List of clinical features

Feature Name	**Feature Name**
gender	malignant_lymphoma
body_mass_index	localized_solid_tumor
myocardial_infarction	metastatic_solid_tumor
congestive_heart_failure	moderate_to_severe_liver_disease
peripheral_vascular_disease	smoking_history_never_smoked
cerebrovascular_disease	smoking_history_previous_smoker
copd	smoking_history_current_smoker
connective_tissue_disease	chewing_tobacco_use_never_or_not_in_last_3mo
peptic_ulcer_disease	chewing_tobacco_use_quit_in_last_3mo
uncomplicated_diabetes_mellitus	alcohol_use_two_or_less_daily
diabetes_mellitus_with_end_organ_damage	alcohol_use_never_or_not_in_last_3mo
chronic_kidney_disease	alcohol_use_more_than_two_daily
hemiplegia_from_stroke	radiographic_size
leukemia	

Table 2. Selected features for the models

Image Type	**Feature Name**
Original	First Order - Interquartile Range
Original	GLCM - IDN
Log filtered (sigma: 2 mm)	3D GLRLM – Long Run Emphasis
Log filtered (sigma: 2 mm)	3D GLRLM – Run Variance
Log filtered (sigma: 2 mm)	3D GLDM - Dependence Variance
Log filtered (sigma: 4 mm)	3D First Order - Kurtosis
Log filtered (sigma: 6 mm)	3D First Order - 90^{th} Percentile
Log filtered (sigma: 6 mm)	3D First Order - Maximum

As a result of the Kruskal-Wallis test, 111 features were eliminated as they were not significant for the problem in question. The process was followed by Lasso regression to detect the most useful features, which left 8 of them (see Table 2) to be used in the classification models. These included: First Order Interquartile Range, GLCM Inverse Difference Normalized and GLRLM Run Variance. Prior to model training, new instances belonging to the minority class were created synthetic minority oversampling technique (SMOTE) to balance the training set. Ultimately, both classes consisted of 128 cases and the models were trained with a total number of 256 cases.

Performance Evaluation

Coarse Gaussian SVM was faster compared to Subspace Discriminant with a training time of 0.47 sec and ~11000 obs/sec (observations per second) prediction speed. Training duration of Subspace Discriminant was 4.1 sec with ~960 obs/sec prediction speed. For Coarse Gaussian SVM; validation accuracy was 67,6% while the accuracy of test was 80%, with and AUC of 0.86. Similarly, Subspace Discriminant had 68,8% validation accuracy and 80% test accuracy; AUC was 0.85. Fig. 1 shows the confusion matrices of the two models. The recieving operatör characteristic (ROC) curves can be seen on Fig. 2.

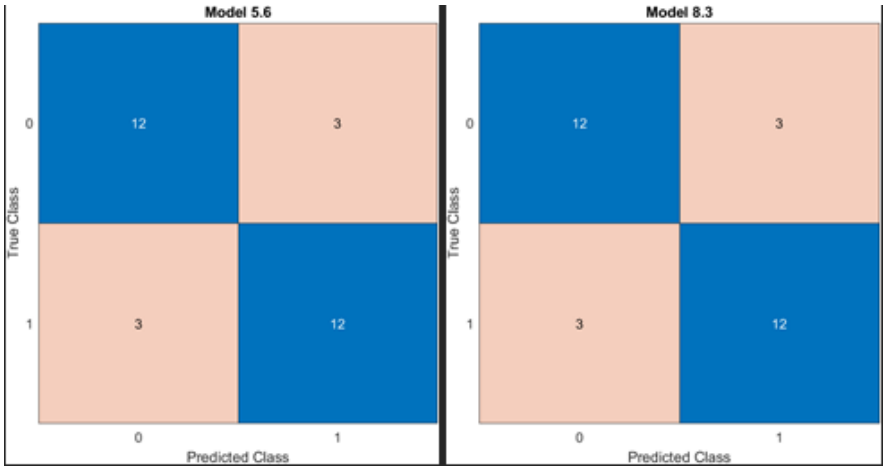

Figure 1. Confusion matrices for Coarse Gaussian SVM and Subspace Discriminant on test data set.

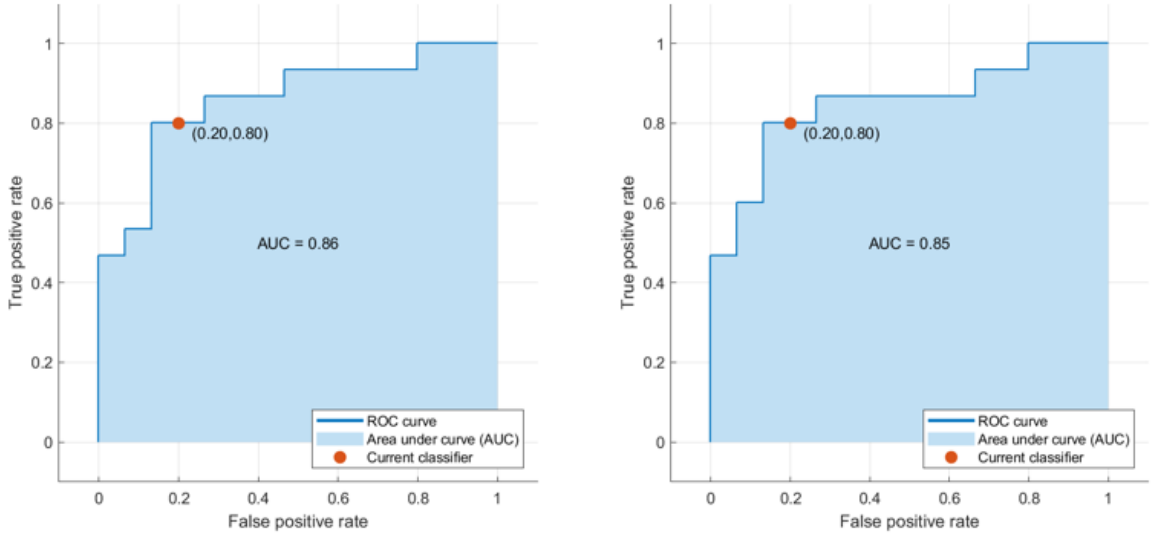

Figure 2. ROC curves of classification models on the test dataset.

5. DISCUSSION

This study investigates the usefulness of machine learning algorithms for malignant kidney tumor histologic subtype classification. In consideration of the performance evaluation, both models demonstrated promising results when classifying the tumors as clear-cell RCC or non-clear-cell RCC. Nonetheless, Coarse Gaussian SVM might be slightly more preferable because of its training and prediction speed.

Our methodology produced similar results as other studies focusing on the similar questions. Therefore, we can deduce that machine learning in radiomics is a viable method for determining the histologic tumor subtype of renal tumors. However, our study differs from others with the data source which was used, as well as other dimensions such as having a high number of cases. Dissimilar to the studies of Kocak B.et al., Zhang G. et al., Hoang et al. and Han et al., this study used all slices of the CT images as an input to the models. Furthermore, we tested if the inclusion of patient clinical data would be useful. Our study found that the specific clinical data included did not have an impact on the classification.

In the future, improved models might be constructed by the addition of blood and urine biomarkers as clinical features. Increasing the size of the data set to achieve better representation of other histologic subtypes can also be considered in order to answer different classification problems.

6. CONCLUSION

We proposed two different models bases on machine learning algorithms to label the malignant tumor cases as ccRCC or non-ccRCC using relevant radiomic features extracted from renal CT images. Both models produced similar results which can be considered as encouraging. These types of classifiers were considered for the first time. This work supports the objective of having a fast and non-invasive technique in the diagnosis process of RCC patients; specifically for deciding the tumor histologic subtype.

7. REFERENCES

1. Clark, Kenneth, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle, Lawrence Tarbox, and Fred Prior. 2013. “The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository.” Journal of Digital Imaging 26(6):1045–57. doi: 10.1007/s10278-013-9622-7.

2. Fedorov, Andriy, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean Christophe Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, Fiona Fennessy, Milan Sonka, John Buatti, Stephen Aylward, James v. Miller, Steve Pieper, and Ron Kikinis. 2012. “3D Slicer as an Image Computing Platform for the Quantitative Imaging Network.” Magnetic Resonance Imaging 30(9):1323–41. doi: 10.1016/j.mri.2012.05.001.

3. Han, Seokmin, Sung il Hwang, and Hak Jong Lee. 2019. “The Classification of Renal Cancer in 3-Phase CT Images Using a Deep Learning Method.” Journal of Digital Imaging 32(4):638–43. doi: 10.1007/s10278-019-00230-2.

4. Heller, N., et al. 2019. "Data from C4KC-KiTS" The Cancer Imaging Archive. doi: 10.7937/TCIA.2019.IX49E8NX

5. Hoang, Uyen N., S. Mojdeh Mirmomen, Osorio Meirelles, Jianhua Yao, Maria Merino, Adam Metwalli, W. Marston Linehan, and Ashkan A. Malayeri. 2018. “Assessment of Multiphasic Contrast-Enhanced Mr Textures in Differentiating

Small Renal Mass Subtypes." Abdominal Radiology 43(12):3400–3409. doi: 10.1007/s00261-018-1625-x.

6. Hsieh, James J., Mark P. Purdue, Sabina Signoretti, Charles Swanton, Laurence Albiges, Manuela Schmidinger, Daniel Y. Heng, James Larkin, and Vincenzo Ficarra. 2017. "Renal Cell Carcinoma." Nature Reviews Disease Primers 3. doi: 10.1038/nrdp.2017.9.

7. Kocak, Burak, Aytul Hande Yardimci, Ceyda Turan Bektas, Mehmet Hamza Turkcanoglu, Cagri Erdim, Ugur Yucetas, Sevim Baykal Koca, and Ozgur Kilickesmez. 2018. "Textural Differences between Renal Cell Carcinoma Subtypes: Machine Learning-Based Quantitative Computed Tomography Texture Analysis with Independent External Validation." European Journal of Radiology 107:149–57. doi: 10.1016/j.ejrad.2018.08.014.

8. Larue, Ruben T. H. M., Janna E. van Timmeren, Evelyn E. C. de Jong, Giacomo Feliciani, Ralph T. H. Leijenaar, Wendy M. J. Schreurs, Meindert N. Sosef, Frank H. P. J. Raat, Frans H. R. van der Zande, Marco Das, Wouter van Elmpt, and Philippe Lambin. 2017. "Influence of Gray Level Discretization on Radiomic Feature Stability for Different CT Scanners, Tube Currents and Slice Thicknesses: A Comprehensive Phantom Study." Acta Oncologica 56(11):1544–53. doi: 10.1080/0284186X.2017.1351624.

9. Le, Valerie H., and James J. Hsieh. 2018. "Genomics and Genetics of Clear Cell Renal Cell Carcinoma: A Mini-Review." Journal of Translational Genetics and Genomics. doi: 10.20517/jtgg.2018.28.

10. Ökmen, Harika Beste, Albert Guvenis, and Hadi Uysal. 2019. "Predicting the Polybromo-1 (Pbrm1) Mutation of a Clear Cell Renal Cell Carcinoma Using Computed Tomography Images and Knn Classification with Random Subspace." Pp. 30–34 in Vibroengineering Procedia. Vol. 26. JVE International.

11. Van Oostenbrugge, Tim J., Jurgen J. Fütterer, and Peter F. A. Mulders. 2018. "Diagnostic Imaging for Solid Renal Tumors: A Pictorial Review." Kidney Cancer 2(2):79–93.

12. Tibshiranit, Robert. 1996. Regression Shrinkage and Selection via the Lasso. Vol. 58.

13. Tomaszewski, Michal R., Gillies, Robert J. 2021. "The Biological Meaning of Radiomic Features", Radiology Vol. 298, No. 3. doi: 10.1148/radiol.2021202553

14. Wang, Z., Wang, K., An, S. 2011. "Cubic B-spline interpolation and realization." International Conference on Information Computing and Applications pp. 82-89.

15. Zhang, G. M. Y., B. Shi, H. D. Xue, B. Ganeshan, H. Sun, and Z. Y. Jin. 2019. "Can Quantitative CT Texture Analysis Be Used to Differentiate Subtypes of Renal Cell Carcinoma?" Clinical Radiology 74(4):287–94. doi: 10.1016/j.crad.2018.11.009.