Original Article

# The application of machine learning based on computed tomography images in the identification of renal tumors in children

**Honghao Song[1#], Xiaoqing Wang[2,3#], Hongwei Wang[1], Feng Guo[2], Rongde Wu[2], Wei Liu[2]**

[1]Department of Pediatric Surgery, Shandong Provincial Hospital, Shandong University, Jinan, China; [2]Department of Pediatric Surgery, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China; [3]Post-doctoral Research Station of Clinical Medicine, Liaocheng People's Hospital, Liaocheng, China

*Contributions:* (I) Conception and design: R Wu, W Liu; (II) Administrative support: R Wu, W Liu; (III) Provision of study materials or patients: R Wu, W Liu; (IV) Collection and assembly of data: H Song, X Wang, H Wang, F Guo; (V) Data analysis and interpretation: H Song, X Wang, H Wang, F Guo; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

*Correspondence to:* Wei Liu, MD. Department of Pediatric Surgery, Shandong Provincial Hospital Affiliated to Shandong First Medical University, 324 Jingwu Street, Jinan 250021, China. Email: lemontree1119@126.com.

**Background:** The clinical manifestations of Wilms tumor and non-Wilms tumor in children are similar, and the only way to confirm the diagnosis is by postoperative pathology. Computed tomography (CT) is one of the main methods for preoperative diagnosis of the two, but it is also difficult to distinguish because it is easily affected by the subjective influence and the experience of the radiologists.

**Methods:** The CT images of 82 children with renal tumors admitted to the Department of Pediatric Urology, Shandong Provincial Hospital from January 2011 to March 2022 were retrospectively analyzed. First, we drew the two-dimensional (2D) region of interest (ROI) of the largest cross-section on the corticomedullary phase (CMP) and nephrogenic phase (NP) images, and extracted seven types of 107 features in the ROI. Then, the texture features with similarity greater than 95% and repetition less than 90% were screened out, and the remaining texture features were further screened by analysis of variance (ANOVA) and recursive feature elimination (RFE). Finally, 15 texture feature were used to build the machine learning (ML) models. We used the synthetic minority oversampling technique (SMOTE) and 10-fold cross-validation to build ML models and verified them in the training, testing, and internal validation sets. The area under the receiver-operating characteristic curve (AUC) and calibration curve were used to evaluate the diagnostic performance.

**Results:** We collected 77 CMP and 81 NP images, which were randomly divided into the training set and the testing set according to the ratio of 7:3. In the internal validation of CMP, the Mean-PCC-ANOVA-5-AE pipeline model achieved the highest AUC 0.792 [95% confidence interval (CI): 0.653–0.930], and its accuracy (ACC), sensitivity (SEN), and specificity (SPE) were 0.833, 0.539 and 0.927, respectively. Correspondingly, in NP, the Mean-PCC-ANOVA-2-LR pipeline model achieved the highest AUC 0.655 (95% CI: 0.485–0.82) in the internal validation. The ACC, SEN, and SPE were 0.696, 0.539, and 0.744, respectively.

**Conclusions:** The ML models based on CT images have good diagnostic efficiency in differentiating Wilms tumors from non-Wilms tumors in children.

**Keywords:** Renal tumors; computed tomography (CT); machine learning (ML); children

Submitted Oct 14, 2023. Accepted for publication Jan 18, 2024. Published online Mar 11, 2024.
doi: 10.21037/tp-23-508
**View this article at:** https://dx.doi.org/10.21037/tp-23-508

## Introduction

### *Background*

Wilms tumor is the most common primary renal tumor in children, accounting for about 80% (1). With the exception of Wilms tumor, other renal tumors include clear cell sarcoma of the kidney (CCSK), malignant rhabdomyoma tumor of the kidney (MRTK), renal cell carcinoma (RCC), congenital mesoblastic nephroma (CMN), and multilocular cystic nephroma (2). With the development of medicine, precision medicine is more respected, and the preoperative clear pathological type of tumor is crucial for deciding on the treatment plans (3).

### *Rationale and knowledge gap*

The clinical manifestations of Wilms tumor and non-Wilms tumor are similar, and computed tomography (CT) is usually used to diagnose them. With the limitations of CT, it is sometimes difficult to identify between the two, and it is preferred to diagnose non-Wilms tumors as "Wilms tumors". A reliable preoperative diagnosis is helpful for the selection of treatment options for children, which is beneficial to the patients. The extensive integration of artificial intelligence (AI) with medical research can aid clinical decision-making (4,5). On this basis, AI can be combined with imaging to correct the influence of subjective factors on diagnostic accuracy. As a subset of AI, machine learning (ML) can be widely used in renal tumor characteristics analysis by identifying information that cannot be detected by the human eyes (6,7).

### *Objective*

Based on the aforementioned, we put forward the idea of building CT-based ML models to assist radiologists and clinicians in preoperative diagnosis and provide more objective information, thereby improving the accuracy of diagnosis with preoperative images. We present this article in accordance with the TRIPOD reporting checklist (available at https://tp.amegroups.com/article/view/10.21037/tp-23-508/rc).

**Highlight box**

**Key findings**

- Machine learning (ML) models are good at identifying non-Wilms tumors from Wilms tumors in children with relatively good results, which may perform better than human beings.
- Corticomedullary phase is better than nephrogenic phase (NP) in discrimination, but NP is less affected by various factors and is more stable.

**What is known and what is new?**

- Few researchers have investigated the use of ML to distinguish Wilms tumors from non-Wilms tumors.
- We developed models with good performance to identify non-Wilms tumors from Wilms tumors in children.

**What is the implication, and what should change now?**

- The most important is to expand the sample size.

## Methods

### *Patient selection*

From January 2011 to March 2022, 83 patients with Wilms tumors and 34 patients with non-Wilms tumors admitted to the Department of Pediatric Urology, Shandong Provincial Hospital were reviewed. The inclusion criteria were as follows: age ≤14 years, complete clinical data, definite pathological diagnosis, and complete preoperative imaging data [corticomedullary phase (CMP) and nephrogenic phase (NP)]. The exclusion criteria were the children who received chemotherapy or who were with combination of renal malformations, images with poor quality, or artifacts. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the institutional review board of Shandong Provincial Hospital Affiliated with Shandong First Medical University (SWYX: No. 2023-453) and individual consent for this retrospective analysis was waived.

### *Image acquisition and segmentation*

A plain scan was performed before the enhanced CT scan. The contrast agent applied was ioversol (350 mg/mL), and the injection dose was 1.5 mL/kg. The CMP was scanned for about 15–30 s after the injection of the contrast agent, and the NP was scanned for about 60–90 s.

Images with slice thicknesses of 3 and 5 mm and tube voltages of 70, 80, 100, and 120 kv were selected. Firstly, the DICOM file was obtained and input into the open-source 3D Slicer software (version 5.2.2) (8) based on the "Segment Editor" module to draw the two-dimensional (2D) region of interest (ROI) on CMP and NP, and the drawn ROI was output in NIfTI format for the next operation. ROIs were

constructed by manually drawing along the tumor edge on the largest cross-sectional plane and one image each was collected from the CMP or NP respectively. Author H.S. drew all ROIs independently, meanwhile, author X.W. used SPSS software (version 26.0.0.0) to select 10% of images randomly and then delineated the ROI using 3D Slicer software.

### *Feature extraction and models construction*

First-order statistic, shape-based 2D, gray level cooccurence matrix (GLCM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), neighbouring gray tone difference matrix (NGTDM), and gray level dependence matrix (GLDM) were extracted by the "Feature Extraction" module of the FAE (FeAture Explorer) software (version 0.5.4) (9). Additionally, some clinical features such as age and sex were also added to the feature selection.

Similarly, FAE software was also used to build models. All images were randomly divided into a training set and a testing set at a ratio of 7:3. In the process of building the ML models, we first used the synthetic minority oversampling technique (SMOTE) to balance the dataset, then used two different methods (Z-score and Mean) for feature matrix normalization, and then screened features with Pearson correlation coefficient (PCC) less than 0.95 for dimensionality reduction. Then we selected analysis of variance (ANOVA), and recursive feature elimination (RFE) to select 15 features, and then selected 10 classifiers for the final model construction. Finally, ten-fold cross-validation was used to determine the hyperparameter of the final models. The 10 classifiers were support vector machine (SVM), linear discriminant analysis (LDA), logistic regression (LR), Adaboost, Gaussian process, Auto-Encoder (AE), Random Forest (RF), logistic regression-lasso (LR-lasso), Decision tree, and Naïve Bayes, respectively (*Figure 1*).

### *Statistical analysis*

The area under the receiver-operating characteristic curve (AUC) and calibration curve were used to evaluate the diagnostic performance of the FAE software (version 0.5.4) (9).

The Shapiro-Wilk (S-W) test and Histogram results determined whether continuous variables were normally distributed. The rank sum test was used to compare nonnormal distributions and the Chi-squared test was used to compare categorical variables. The statistical data were implemented in SPSS software (version 26.0.0.0). Statistical significance was indicated by a P value of less than 0.05.

In addition, R software (version 4.2.0) was taken to write code for intraclass correlation coefficients (ICC) and selected variables with excellent reliability (≥0.90) (detailed coding is provided in Appendix 1).

## Results

### *Patient*

According to our inclusion and exclusion criteria, 19 non-Wilms tumor patients and 63 Wilms tumor patients were included in our research (*Figure 2*). Among the non-Wilms tumor children, the median age was 25.84 months, with 12 males and 7 females, and among the Wilms tumor children, the median age was 22.39 months, with 35 males and 28 females.

To expand the sample, we also selected images with slice thicknesses of 3 and 5 mm and tube voltages of 70, 80, 100, and 120 kv. There were no significant differences in CMP or NP (all P>0.05). The results are shown in *Tables 1,2*. There were no significant differences between the two groups (non-Wilms tumor and Wilms tumor, all P>0.05). The results are shown in *Tables 1,2*.

### *ML models*

A total of 600 pipelines were built in our ML models, including one method of balance data, two normalization methods, one method of dimensionality reduction, two feature selectors, ten classifiers, and fifteen features. Appendix 2 details all features with ICC greater than 0.90 on CMP and NP.

*Table 3* shows the data of the training and testing sets at a ratio of 7:3. Fifty-four patients were in the training set and we also selected another 23 cases as the testing set for the model of CMP. Sixty-two patients were in the training set and we also selected another 19 cases as the testing set for the model of NP.

### *The model of CMP*

In the model of CMP, 28 texture features with similarity greater than 95% and repetition less than 90% were screened, and 15 texture features were further screened by ANOVA and RFE for model construction. We found that the Mean-PCC-ANOVA-AE pipeline model based on five features achieved the highest AUC (0.792) in the

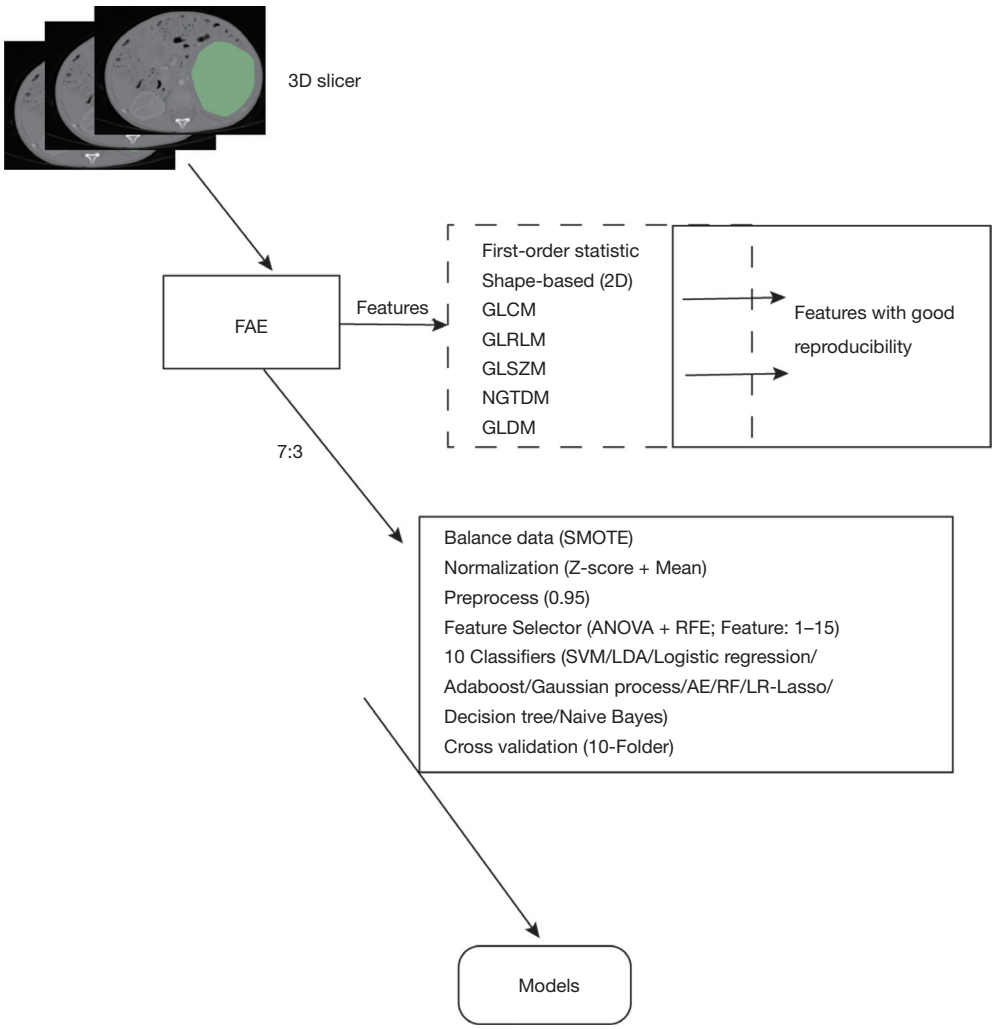


**Figure 1** The structure of the model. FAE, FeAture Explorer; GLCM, gray level cooccurence matrix; GLRLM, gray level run length matrix; GLSZM, gray level size zone matrix; NGTDM, neighbouring gray tone difference matrix; GLDM, gray level dependence matrix; SMOTE, synthetic minority oversampling technique; ANOVA, analysis of variance; RFE, recursive feature elimination; SVM, support vector machine; LDA, linear discriminant analysis; AE, Auto-Encoder; RF, Random Forest; LR-lasso, logistic regression-lasso.

validation set. The AUC of the model achieved 0.433 and 0.814 in the testing and training sets respectively (*Figure 3*). The five texture features that contributed the most to the CMP model were firstorder_90Percentile, firstorder_Interquartilerange, glcm_Sumentropy, glcm_Imc2, and firstorder_Uniformity (*Figure 4*).

### *The model of NP*

In the model of NP, 59 texture features with similarity greater than 95% and repetition less than 90% were removed, and 15 texture features were also used to construct the models. The Mean-PCC-ANOVA-LR pipeline model based on the two features resulted with the highest AUC in the validation set. The AUC could achieve 0.655. In the testing and training sets, the AUC of the model achieved 0.693 and 0.784, respectively (*Figure 5*). Shape_SurfaceVolumeRatio and ngtdm_Busyness contributed the most to the model (*Figure 6*).

## Discussion

### *Key findings, strengths, and limitations*

CT is the main method for preoperative identification of

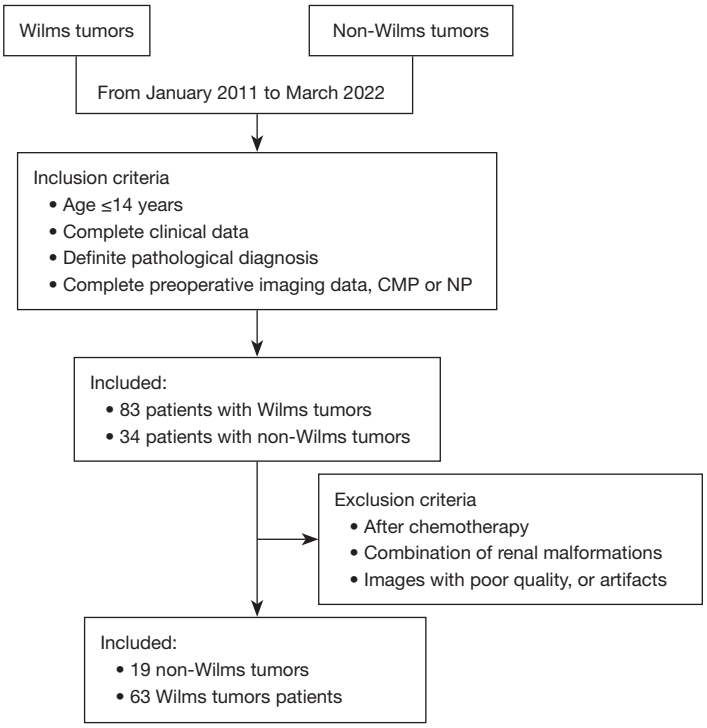


**Figure 2** Inclusion and exclusion criteria of patients. CMP, corticomedullary phase; NP, nephrogenic phase.

Wilms tumors and non-Wilms tumors (10). However, it is difficult to distinguish Wilms tumor from non-Wilms tumor through non-invasive examination methods before surgery. CT diagnosis is not only often limited by the experience of radiologists but also affected by the subjective thoughts of radiologists. For example, in our study, 26 of the 34 children with non-Wilms tumors were diagnosed preoperatively as Wilms tumors by CT diagnosis, three were diagnosed as renal tumors, and only three had the same postoperative pathological results. Similarly, among 63 children with a post-pathological diagnosis of Wilms tumor, 55 cases had a diagnosis consistent with the preoperative diagnosis of CT, three cases were misdiagnosed (one mesoblastic nephroma, one neurogenic tumor, and one angiomyolipoma), four cases were diagnosed as renal tumors, and one case was diagnosed as retroperitoneal tumors. So, the accuracy (ACC), sensitivity (SEN), and specificity (SPE) of radiologists were 0.592, 0.640, and 0.273 in our research. Then, is there a more objective preoperative non-invasive examination method that can avoid the influence by experience and subjective thoughts, while improving the accuracy of diagnosis? If the type of renal tumor in children can be diagnosed by imaging before surgery, children with renal tumors will benefit. There are differences in the treatment of pediatric renal tumors in different regions. For example, in North America [the Children's Oncology Group (COG)], it is advocated that tumor resection should be performed first, and subsequent treatment should be carried out after the pathology is confirmed. In Europe [the International Society of Paediatric Oncology (SIOP)], preoperative chemotherapy to reduce tumor staging and the risk of rupture is followed by operation. For children between the ages of 6 months and nine years, the standard chemotherapy regimen for Wilms tumor is generally adopted, and when non-Wilms tumor is suspected, the chemotherapy regimen is determined based on the results of biopsy. For children older than 10 years with an uncertain clinical diagnosis, the biopsy is generally performed first (11-13). Wilms tumors can be distinguished from non-Wilms tumors by non-invasive methods (imaging), which is beneficial both for the plan of direct resection of the tumor and the plan of further treatment after biopsy, yet biopsy may cause needle tract metastasis and increase the clinical stage.

Sharaby *et al.* (14) used a novel computer-aided prediction system to predict the preoperative chemotherapy susceptibility of Wilms tumor and this system was based on SVM. Wang *et al.* (15) also used a method based on ML to validate and compare the performance of dose reconstruction

**Table 1** Clinical features and the imaging features of CMP and NP

| Variables | Non-Wilms tumor | Wilms tumor | Value ($Z/\chi^2$) | P |
|---|---|---|---|---|
| Age (months) | 25.84 (8.84, 54.77) | 22.39 (8.81, 35.18) | Z=−0.385 | 0.7 |
| Gender | | | $\chi^2$=0.345 | 0.557 |
| Female | 7 | 28 | | |
| Male | 12 | 35 | | |
| Maximum diameter (cm) | | | | |
| ≥11 | 4 | 25 | $\chi^2$=2.216 | 0.317 |
| <11 | 15 | 38 | | |
| CMP | 18 | 59 | $\chi^2$=0 | 0.991 |
| NP | 19 | 62 | | |
| CMP slice (mm) | | | $\chi^2$=3.492 | 0.138 |
| 3 | 6 | 5 | | |
| 5 | 53 | 13 | | |
| CMP tube voltages (Kv) | | | $\chi^2$=0.772 | 0.856 |
| 70 | 7 | 2 | | |
| 80 | 7 | 2 | | |
| 100 | 22 | 5 | | |
| 120 | 23 | 9 | | |
| NP slice (mm) | | | $\chi^2$=2.484 | 0.247 |
| 3 | 5 | 4 | | |
| 5 | 57 | 15 | | |
| NP tube voltages (Kv) | | | $\chi^2$=1.317 | 0.725 |
| 70 | 7 | 3 | | |
| 80 | 8 | 2 | | |
| 100 | 24 | 5 | | |
| 120 | 23 | 9 | | |

CMP, corticomedullary phase; NP, nephrogenic phase.

methods in historical radiotherapy plans for Wilms tumor. Ma *et al.* (16) used ML to identify the stage of Wilms tumor in children and got an excellent performance. Therefore, the ML perhaps could be an excellent method to be applied to renal tumors in children, including determination of the tumor treatment regimen's response and the tumor's staging. Thus, although non-Wilms' tumor is a kind of highly heterogeneous disease because of its variety of types of tumors (2), ML models can overcome this and distinguish Wilms tumor from non-Wilms tumor and ML models have been widely used in adult kidney tumors (17-23). So, by constructing ML models, Wilms tumors can be better distinguished from non-Wilms tumors. To the best of our knowledge, few studies have applied ML to distinguish pediatric Wilms tumors from non-Wilms tumors. The possible reasons are as follows: (I) there are many kinds of non-Wilms' tumors, including CCSK, renal, RCC, MRTK, CMN, multilocular cystic nephroma, and other types. (II) There are no obvious specific diagnostic signs on CT, about Wilms' tumors, and it is difficult to distinguish them by the naked eye.

In our research, 82 patients, 77 images of CMP, and

**Table 2** The factors in CMP and NP

| Variables | CMP | NP | Value ($\chi^2$) | P |
|---|---|---|---|---|
| Age (months) | | | 1.611 | 0.204 |
| ≥26 | 35 | 45 | | |
| <26 | 42 | 36 | | |
| Gender | | | 0.002 | 0.964 |
| Female | 33 | 35 | | |
| Male | 44 | 46 | | |
| Side | | | 0.057 | 0.812 |
| Left | 48 | 49 | | |
| Right | 29 | 32 | | |
| Maximum diameter (cm) | | | 0.154 | 0.695 |
| ≥11 | 26 | 26 | | |
| <11 | 51 | 55 | | |

CMP, corticomedullary phase; NP, nephrogenic phase.

**Table 3** Training set and testing set on the CMP and NP ML model

| Phase | Tumor | Training | Testing |
|---|---|---|---|
| CMP | Wilms tumor | 41 | 18 |
| | Non-Wilms tumor | 13 | 5 |
| NP | Wilms tumor | 43 | 13 |
| | Non-Wilms tumor | 19 | 6 |

CMP, corticomedullary phase; NP, nephrogenic phase; ML, machine learning.

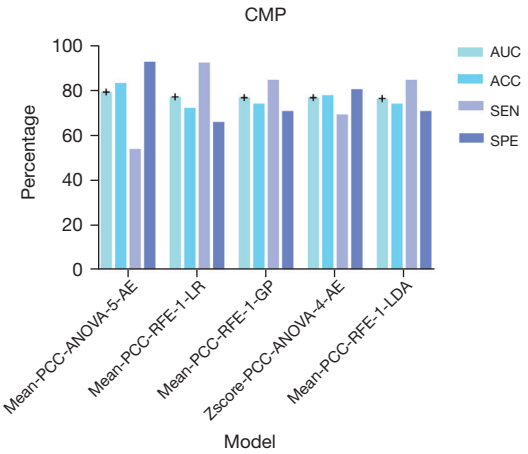


**Figure 3** The top five models in the CMP, according to the performance of AUC. CMP, corticomedullary phase; PCC, Pearson correlation coefficient; ANOVA, analysis of variance; RFE, recursive feature elimination; AE, Auto-Encoder; LDA, linear discriminant analysis; LR, logistic regression; GP, Gaussian process; AUC, the area under the receiver-operating characteristic curve; ACC, accuracy; SEN, sensitivity; SPE, specificity.

81 images of NP were included for the analysis. One of the strengths of our study is the inclusion of age in the construction of ML models. Age is one of the most common clinical indices for preoperative diagnosis (3) and the age distribution of each renal tumor in children is different (24,25). In the CMP model, we received the highest performance of AUC (0.792), ACC (0.833), SEN

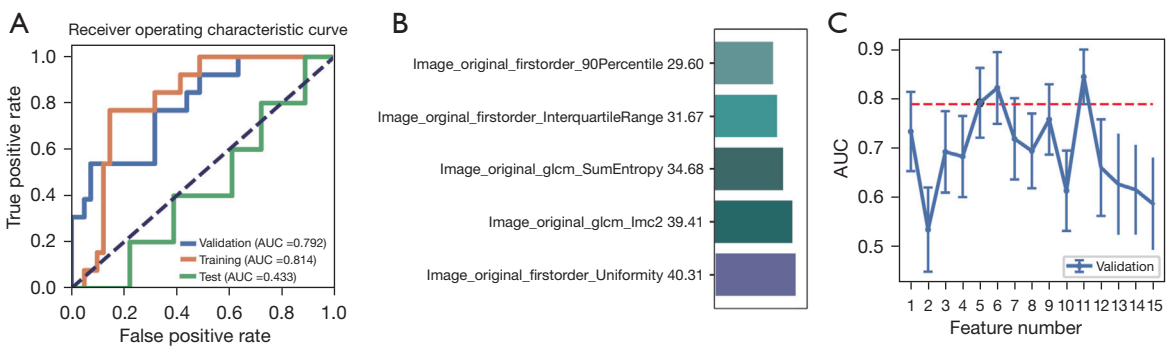


**Figure 4** The AUC (A), texture features (B), and feature number (C) in the model of CMP. The red dashed line indicates models that are within 1 standard error of the optimal model. AUC, the area under the receiver-operating characteristic curve; GLCM, gray level cooccurence matrix; CMP, corticomedullary phase.

(0.923), and SPE (0.927), where the highest AUC, ACC and SPE were generated by the Mean-PCC-ANOVA-5-AE pipeline model and the highest SEN was generated by the Mean-PCC-RFE-1-LR pipeline model. Compared to the performance of the CMP model (the AUC of training, testing, and validation sets were 0.814, 0.433, and 0.792 respectively), the NP model received a firm performance (the AUC of training, testing, and validation sets were 0.784, 0.693, and 0.655 respectively). The accuracy of CMP models ranged from 0.630 to 0.833, and all of our

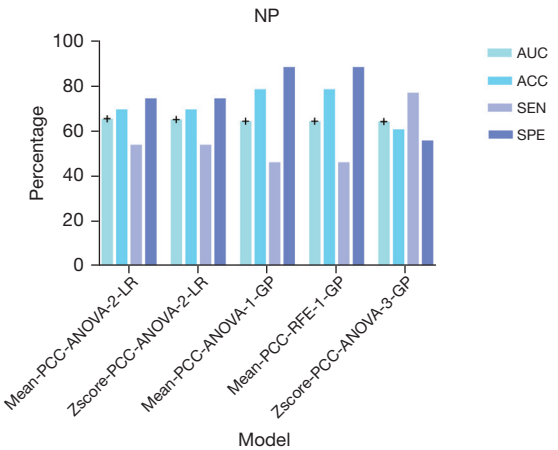


**Figure 5** The top five models in the NP, according to the performance of AUC. NP, nephrogenic phase; PCC, Pearson correlation coefficient; ANOVA, analysis of variance; RFE, recursive feature elimination; LR, logistic regression; GP, Gaussian process; AUC, the area under the receiver-operating characteristic curve; ACC, accuracy; SEN, sensitivity; SPE, specificity.

CMP models performed better than the accuracy (0.592) of preoperative image diagnosis. Different from the selected phase of Ma *et al.* (16,26), we selected CMP and NP as our targeted phases and received excellent results. In terms of building a model to distinguish Wilms tumors from non-Wilms tumors, Zhu *et al.* (3) built a deep learning (DL) model (ResNet34) to achieve this. In contrast to their study, we carefully drew the ROI along the edge of the tumor to avoid the influence of peritumoral renal parenchyma and fat. However, they chose 1.2 ROI to extract features. Fine manual drawing of ROI along the tumor margin may be beneficial to the final result of the models. Regardless of ML or DL, the performance of models is most closely related to the sample size. To improve the performance of our model, the sample size should be further increased in the future. In the ML model of CMP, the AUC of the testing set was 0.433. About this result, we did some analysis and the possible reasons are as follows: (I) although the CMP could not only identify the border of tumors distinctly but also could provide a piece of comprehensive information, the NP could better recognize peritumor renal parenchyma and peritumor perirenal fat. (II) Renal tumors in children are large and often accompanied by internal necrosis and cystic degeneration. In CMP, the CT value of the renal parenchyma was higher than that of the tumor (mean CT value: 120 *vs.* 45 HU), thus potentially masking the tumor.

## Conclusions

In order to improve the clinical diagnosis by non-invasive

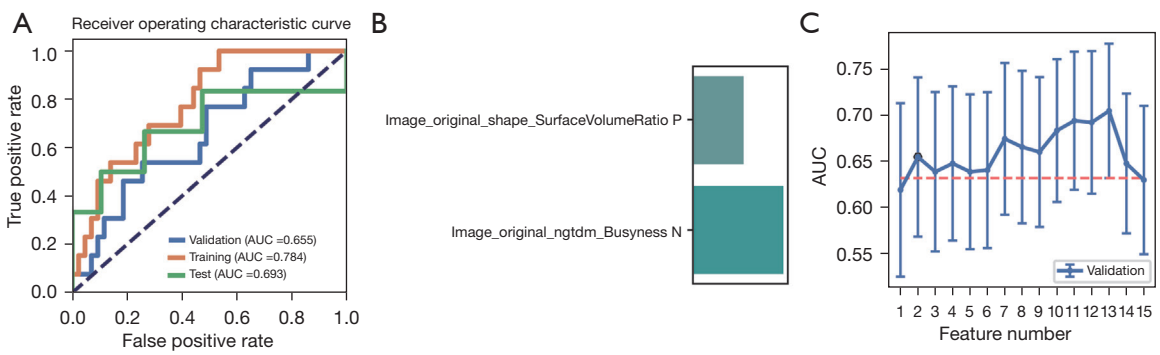


**Figure 6** The AUC (A), texture features (B), and feature number (C) in the model of NP. The red dashed line indicates models that are within 1 standard error of the optimal model. AUC, the area under the receiver-operating characteristic curve; NGTDM, neighbouring gray tone difference matrix; NP, nephrogenic phase.

imaging methods, we presented ML models with excellent results for identifying non-Wilms tumors from Wilms tumors in children with excellent results. Each phase, whether it is CMP or NP, has its own advantages and disadvantages. CMP is better than NP in discrimination, but NP is less affected by various factors and more stable. Perhaps, our research can promote the integration of pediatric renal tumors and ML, thereby further promoting the application of AI in pediatric tumors.

## Acknowledgments

We would like to thank all involved participants in this study.

*Funding:* This work was supported by grants from the Youth Science Foundation of Shandong First Medical University (No. 202201-066 to X.W.), the China Postdoctoral Science Foundation (No. 2023M741507 to X.W.), the Science and Technology Development Plan Project of Shandong Province, China (No. 2018GSF118209 to W.L.), and the Shandong Provincial Natural Science Foundation (Nos. ZR2017MH091, and ZR2022MH095 to R.W., and No. ZR2023MH030 to W.L.).

## Footnote

*Reporting Checklist:* The authors have completed the TRIPOD reporting checklist. Available at https://tp.amegroups.com/article/view/10.21037/tp-23-508/rc

*Data Sharing Statement:* Available at https://tp.amegroups.com/article/view/10.21037/tp-23-508/dss

*Peer Review File:* Available at https://tp.amegroups.com/article/view/10.21037/tp-23-508/prf

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at https://tp.amegroups.com/article/view/10.21037/tp-23-508/coif). X.W. reports grants from the Youth Science Foundation of Shandong First Medical University (No. 202201-066) and the China Postdoctoral Science Foundation (No. 2023M741507). W.L. reports funding from the Science and Technology Development Plan Project of Shandong Province, China (No. 2018GSF118209) and the Shandong Provincial Natural Science Foundation (No. ZR2023MH030). R.W. reports funding from the Shandong Provincial Natural Science Foundation (Nos. ZR2017MH091, and ZR2022MH095). The other authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the institutional review board of Shandong Provincial Hospital Affiliated with Shandong First Medical University (SWYX: No. 2023-453) and individual consent for this retrospective analysis was waived.



## References

1. Qureshi SS, Bhagat M, Verma K, et al. Incidence, treatment, and outcomes of primary and recurrent Non-Wilms renal tumors in children: Report of 109 patients treated at a single institution. J Pediatr Urol 2020;16:475.e1-9.
2. Broecker B. Non-Wilms' renal tumors in children. Urol Clin North Am 2000;27:463-9, ix.
3. Zhu Y, Li H, Huang Y, et al. CT-based identification of pediatric non-Wilms tumors using convolutional neural networks at a single center. Pediatr Res 2023;94:1104-10.
4. van Timmeren JE, Cester D, Tanadini-Lang S, et al. Radiomics in medical imaging-"how-to" guide and critical reflection. Insights Imaging 2020;11:91.
5. Xu HL, Gong TT, Liu FH, et al. Artificial intelligence performance in image-based ovarian cancer identification: A systematic review and meta-analysis. EClinicalMedicine 2022;53:101662.
6. Kocak B, Kus EA, Yardimci AH, et al. Machine Learning in Radiomic Renal Mass Characterization: Fundamentals, Applications, Challenges, and Future Directions. AJR Am J Roentgenol 2020;215:920-8.
7. de Leon AD, Kapur P, Pedrosa I. Radiomics in Kidney Cancer: MR Imaging. Magn Reson Imaging Clin N Am

 

2019;27:1-13.
8. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 2012;30:1323-41.
9. Song Y, Zhang J, Zhang YD, et al. FeAture Explorer (FAE): A tool for developing and comparing radiomics models. PLoS One 2020;15:e0237587.
10. Krishna S, Murray CA, McInnes MD, et al. CT imaging of solid renal masses: pitfalls and solutions. Clin Radiol 2017;72:708-21.
11. Wilms Tumor and Other Childhood Kidney Tumors Treatment (PDQ®): Health Professional Version. 2002.
12. Watson T, Oostveen M, Rogers H, et al. The role of imaging in the initial investigation of paediatric renal tumours. Lancet Child Adolesc Health 2020;4:232-41.
13. Brisse HJ, de la Monneraye Y, Cardoen L, et al. From Wilms to kidney tumors: which ones require a biopsy? Pediatr Radiol 2020;50:1049-51.
14. Sharaby I, Alksas A, Nashat A, et al. Prediction of Wilms' Tumor Susceptibility to Preoperative Chemotherapy Using a Novel Computer-Aided Prediction System. Diagnostics (Basel) 2023;13:486.
15. Wang Z, Virgolin M, Balgobind BV, et al. Validation and Comparison of Radiograph-Based Organ Dose Reconstruction Approaches for Wilms Tumor Radiation Treatment Plans. Adv Radiat Oncol 2022;7:101015.
16. Ma XH, Shu L, Jia X, et al. Machine Learning-Based CT Radiomics Method for Identifying the Stage of Wilms Tumor in Children. Front Pediatr 2022;10:873035.
17. Demirjian NL, Varghese BA, Cen SY, et al. CT-based radiomics stratification of tumor grade and TNM stage of clear cell renal cell carcinoma. Eur Radiol 2022;32:2552-63.
18. 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;29:1153-63.
19. Kocak B, Durmaz ES, Ates E, et al. 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;212:W132-9.
20. Hodgdon T, McInnes MD, Schieda N, et al. Can Quantitative CT Texture Analysis be Used to Differentiate Fat-poor Renal Angiomyolipoma from Renal Cell Carcinoma on Unenhanced CT Images? Radiology 2015;276:787-96.
21. Feng Z, Rong P, Cao P, et al. Machine learning-based quantitative texture analysis of CT images of small renal masses: Differentiation of angiomyolipoma without visible fat from renal cell carcinoma. Eur Radiol 2018;28:1625-33.
22. Wu K, Wu P, Yang K, et al. A comprehensive texture feature analysis framework of renal cell carcinoma: pathological, prognostic, and genomic evaluation based on CT images. Eur Radiol 2022;32:2255-65.
23. Nazari M, Shiri I, Zaidi H. Radiomics-based machine learning model to predict risk of death within 5-years in clear cell renal cell carcinoma patients. Comput Biol Med 2021;129:104135.
24. Ahmed HU, Arya M, Levitt G, et al. Part I: Primary malignant non-Wilms' renal tumours in children. Lancet Oncol 2007;8:730-7.
25. Wilms tumour. Nat Rev Dis Primers 2021;7:76.
26. Ma XH, Yang J, Jia X, et al. Preoperative radiomic signature based on CT images for noninvasive evaluation of localized nephroblastoma in pediatric patients. Front Oncol 2023;13:1122210.

**Cite this article as:** Song H, Wang X, Wang H, Guo F, Wu R, Liu W. The application of machine learning based on computed tomography images in the identification of renal tumors in children. Transl Pediatr 2024;13(3):417-426. doi: 10.21037/tp-23-508

Supplementary

## Appendix 1 Detailed coding for intraclass correlation coefficients

```
> install.packages("psych")
library(psych)
setwd("~/Desktop/Fanomics")
> getwd()
> df_icc1 <- read.csv("icc1.csv")
> df_icc2 <- read.csv("icc2.csv")
>  n <- dim(df_icc1)[1]
> p <- dim(df_icc1)[2]
> df_intra <- rbind(df_icc1,df_icc2)
icc_all_intra <- apply(df_intra, 2, function(x) ICC(x = data.frame(x[1:n], x[(n+1):(2*n)]),lmer = F)$results[1,2])
df_icc_all_intra <- data.frame(feature = names(df_intra),icc = icc_all_intra, row.names = 1:ncol(df_intra))
write.csv(df_icc_all_intra,"df_icc_all_intra.csv")
```


https://dx.doi.org/10.21037/tp-23-508

## Appendix 2 All features with intraclass correlation coefficients greater than 0.90 on CMP and NP

| CMP | | NP | |
|---|---|---|---|
| Feature | ICC | Feature | ICC |
| image_original_shape_SurfaceVolumeRatio | 0.999703 | image_original_ngtdm_Strength | 0.997886 |
| image_original_ngtdm_Coarseness | 0.999602 | image_original_shape_SurfaceVolumeRatio | 0.997055 |
| image_original_gldm_SmallDependenceEmphasis | 0.999219 | image_original_glszm_ZoneVariance | 0.994225 |
| image_original_firstorder_RobustMeanAbsoluteDeviation | 0.998798 | image_original_glszm_LargeAreaEmphasis | 0.994144 |
| image_original_glszm_ZonePercentage | 0.998405 | image_original_glszm_SizeZoneNonUniformityNormalized | 0.993853 |
| image_original_firstorder_MeanAbsoluteDeviation | 0.998284 | image_original_glrlm_LongRunLowGrayLevelEmphasis | 0.990296 |
| image_original_glcm_Contrast | 0.99777 | image_original_glcm_ClusterShade | 0.988236 |
| image_original_glcm_Imc1 | 0.997564 | image_original_gldm_DependenceNonUniformityNormalized | 0.987703 |
| image_original_firstorder_90Percentile | 0.997399 | image_original_glrlm_RunVariance | 0.987175 |
| image_original_glcm_DifferenceEntropy | 0.997288 | image_original_glrlm_LongRunEmphasis | 0.98707 |
| image_original_glcm_SumSquares | 0.997225 | image_original_gldm_LargeDependenceEmphasis | 0.985449 |
| image_original_glcm_ClusterTendency | 0.997185 | image_original_glcm_DifferenceEntropy | 0.983747 |
| image_original_glrlm_GrayLevelNonUniformityNormalized | 0.997132 | image_original_glrlm_RunPercentage | 0.983314 |
| image_original_glcm_DifferenceAverage | 0.996764 | image_original_glcm_DifferenceVariance | 0.982974 |
| image_original_gldm_DependenceNonUniformity | 0.996724 | image_original_gldm_GrayLevelNonUniformity | 0.981646 |
| image_original_firstorder_Energy | 0.996709 | image_original_glcm_InverseVariance | 0.981574 |
| image_original_gldm_GrayLevelVariance | 0.996695 | image_original_glcm_Id | 0.980907 |
| image_original_glrlm_RunLengthNonUniformityNormalized | 0.996597 | image_original_glcm_Idm | 0.980738 |
| image_original_glcm_DifferenceVariance | 0.996488 | image_original_glszm_LargeAreaLowGrayLevelEmphasis | 0.980694 |
| image_original_glcm_Idm | 0.996361 | image_original_glcm_DifferenceAverage | 0.980483 |
| image_original_firstorder_RootMeanSquared | 0.996295 | image_original_glrlm_RunLengthNonUniformityNormalized | 0.98037 |
| image_original_glcm_Id | 0.996225 | image_original_ngtdm_Coarseness | 0.980166 |
| image_original_firstorder_Variance | 0.99615 | image_original_glcm_Contrast | 0.97914 |
| image_original_firstorder_Entropy | 0.996053 | image_original_glszm_SmallAreaEmphasis | 0.97773 |
| image_original_firstorder_Uniformity | 0.995976 | image_original_ngtdm_Busyness | 0.974418 |
| image_original_glrlm_RunPercentage | 0.99597 | image_original_gldm_DependenceNonUniformity | 0.966119 |
| image_original_firstorder_Median | 0.995948 | image_original_shape_MinorAxisLength | 0.965445 |
| image_original_glcm_JointEntropy | 0.995881 | image_original_glszm_LargeAreaHighGrayLevelEmphasis | 0.964691 |
| image_original_glrlm_LongRunEmphasis | 0.995749 | image_original_glrlm_LongRunHighGrayLevelEmphasis | 0.961037 |
| image_original_glcm_InverseVariance | 0.995528 | image_original_gldm_DependenceVariance | 0.958119 |
| image_original_gldm_GrayLevelNonUniformity | 0.995504 | image_original_gldm_SmallDependenceEmphasis | 0.957753 |
| image_original_gldm_LargeDependenceEmphasis | 0.995439 | image_original_glszm_GrayLevelNonUniformity | 0.955541 |
| image_original_glcm_Correlation | 0.99543 | image_original_firstorder_Minimum | 0.939263 |
| image_original_shape_Maximum2DDiameterColumn | 0.995421 | image_original_ngtdm_Complexity | 0.938753 |
| image_original_glcm_SumEntropy | 0.995352 | image_original_glrlm_GrayLevelNonUniformity | 0.935427 |
| image_original_glrlm_ShortRunEmphasis | 0.994974 | image_original_shape_Sphericity | 0.935179 |
| image_original_glcm_MCC | 0.994897 | image_original_glcm_MaximumProbability | 0.934252 |
| image_original_firstorder_Mean | 0.99483 | image_original_glszm_LowGrayLevelZoneEmphasis | 0.933156 |
| image_original_glcm_JointEnergy | 0.994762 | image_original_glcm_ClusterProminence | 0.933095 |
| image_original_firstorder_TotalEnergy | 0.99472 | image_original_glrlm_ShortRunEmphasis | 0.932651 |
| image_original_glrlm_LongRunLowGrayLevelEmphasis | 0.994709 | image_original_glcm_Idn | 0.927744 |
| image_original_firstorder_InterquartileRange | 0.994634 | image_original_shape_SurfaceArea | 0.923885 |
| image_original_glrlm_RunVariance | 0.994585 | image_original_firstorder_10Percentile | 0.923572 |
| image_original_glcm_Imc2 | 0.994256 | image_original_shape_VoxelVolume | 0.921104 |
| image_original_glrlm_RunEntropy | 0.994007 | image_original_shape_MeshVolume | 0.921026 |
| image_original_gldm_DependenceEntropy | 0.993948 | image_original_shape_Maximum2DDiameterRow | 0.916918 |
| image_original_gldm_DependenceVariance | 0.993479 | image_original_glszm_ZonePercentage | 0.913083 |
| image_original_glrlm_GrayLevelNonUniformity | 0.993425 | image_original_glrlm_ShortRunLowGrayLevelEmphasis | 0.901776 |
| image_original_gldm_DependenceNonUniformityNormalized | 0.993417 | | |
| image_original_shape_Sphericity | 0.993237 | | |
| image_original_shape_VoxelVolume | 0.990466 | | |
| image_original_shape_MeshVolume | 0.990441 | | |
| image_original_shape_MajorAxisLength | 0.990134 | | |
| image_original_glrlm_RunLengthNonUniformity | 0.990073 | | |
| image_original_shape_SurfaceArea | 0.990066 | | |
| image_original_firstorder_10Percentile | 0.989598 | | |
| image_original_shape_Maximum2DDiameterSlice | 0.989024 | | |
| image_original_shape_Maximum3DDiameter | 0.989024 | | |
| image_original_glcm_MaximumProbability | 0.989006 | | |
| image_original_glrlm_LongRunHighGrayLevelEmphasis | 0.988914 | | |
| image_original_glszm_GrayLevelNonUniformityNormalized | 0.987143 | | |
| image_original_shape_MinorAxisLength | 0.987114 | | |
| image_original_shape_Maximum2DDiameterRow | 0.986042 | | |
| image_original_glszm_SizeZoneNonUniformityNormalized | 0.982644 | | |
| image_original_glszm_SizeZoneNonUniformity | 0.982329 | | |
| image_original_glszm_GrayLevelNonUniformity | 0.981122 | | |
| image_original_glszm_LargeAreaLowGrayLevelEmphasis | 0.977235 | | |
| image_original_glszm_ZoneEntropy | 0.970264 | | |
| image_original_glrlm_GrayLevelVariance | 0.965288 | | |
| image_original_glszm_SmallAreaEmphasis | 0.96069 | | |
| image_original_glszm_LargeAreaEmphasis | 0.950659 | | |
| image_original_glszm_ZoneVariance | 0.94956 | | |
| image_original_firstorder_Skewness | 0.946556 | | |
| image_original_glszm_LargeAreaHighGrayLevelEmphasis | 0.941172 | | |
| image_original_glszm_SmallAreaLowGrayLevelEmphasis | 0.925727 | | |
| image_original_ngtdm_Busyness | 0.916888 | | |
| image_original_gldm_LargeDependenceLowGrayLevelEmphasis | 0.910933 | | |
| image_original_firstorder_Minimum | 0.908471 | | |
| image_original_glcm_ClusterShade | 0.901755 | | |


https://dx.doi.org/10.21037/tp-23-508