European Journal of Radiology 179 (2024) 111677

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Radiomics signature for automatic hydronephrosis detection in unenhanced Low-Dose CT

Judith Becker [a,1], Piotr Woźnicki [b,1], Josua A. Decker [a], Franka Risch [a], Ramona Wudy [a], David Kaufmann [a], Luca Canalini [a], Claudia Wollny [a], Christian Scheurig-Muenkler [a], Thomas Kroencke [a,c,*], Stefanie Bette [a,1], Florian Schwarz [d,e,1]

[a] Department of Diagnostic and Interventional Radiology, University Hospital Augsburg, Stenglinstr. 2, 86156 Augsburg, Germany
[b] Diagnostic and Interventional Radiology, University Hospital Würzburg, Josef-Schneider-Straße 2, 97080 Würzburg, Germany
[c] Centre for Advanced Analytics and Predictive Sciences (CAAPS), University of Augsburg, Universitätsstr. 2, 86159 Augsburg, Germany
[d] Centre for Diagnostic Imaging and Interventional Therapy, Donau-Isar-Klinikum, Perlasberger Straße 41, 94469 Deggendorf, Germany
[e] Medical Faculty, Ludwig Maximilian University Munich, Bavariaring 19, 80336 Munich, Germany

ARTICLE INFO

Keywords:
Automatic segmentation
Artificial intelligence
Radiomics
Deep learning
Hydronephrosis
Urinary retention

ABSTRACT

Purpose: To investigate the diagnostic performance of an automatic pipeline for detection of hydronephrosis on kidney's parenchyma on unenhanced low-dose CT of the abdomen.
Methods: This retrospective study included 95 patients with confirmed unilateral hydronephrosis in an unenhanced low-dose CT of the abdomen. Data were split into training (n = 67) and test (n = 28) cohorts. Both kidneys for each case were included in further analyses, whereas the kidney without hydronephrosis was used as control. Using the training cohort, we developed a pipeline consisting of a deep-learning model for automatic segmentation (a Convolutional Neural Network based on nnU-Net architecture) of the kidney's parenchyma and a radiomics classifier to detect hydronephrosis. The models were assessed using standard classification metrics, such as area under the ROC curve (AUC), sensitivity and specificity, as well as semantic segmentation metrics, including Dice coefficient and Jaccard index.
Results: Using manual segmentation of the kidney's parenchyma, hydronephrosis can be detected with an AUC of 0.84, a sensitivity of 75% and a specificity of 82%, a PPV of 81% and a NPV of 77%. Automatic kidney segmentation achieved a mean Dice score of 0.87 and 0.91 for the right and left kidney, respectively. Additionally, automatic segmentation achieved an AUC of 0.83, a sensitivity of 86%, specificity of 64%, PPV of 71%, and NPV of 82%.
Conclusion: Our proposed radiomics signature using automatic kidney's parenchyma segmentation allows for accurate hydronephrosis detection on unenhanced low-dose CT scans of the abdomen independently of widened renal pelvis. This method could be used in clinical routine to highlight hydronephrosis to radiologists as well as clinicians, especially in patients with concurrent parapelvic cysts and might reduce time and costs associated with diagnosing hydronephrosis.

1. Introduction

The number of hydronephrosis due to renal calculi is worldwide increasing [1–3]. Patients with urolithiasis usually present with flank pain and probably urinary abnormalities such as hematuria in the emergency departments [4]. Due to the increasing occurrence of this disease, the number of further investigations and treatment methods is rising, burdening the medical system [5]. The imaging method of choice in a patient suspicious to have urolithiasis is an ultrasound examination [6]. If the patient has signs of urinary retention in an ultrasound examination, this is highly suggestive for urolithiasis [7,8]. In most cases, an additional low-dose unenhanced CT scan is performed to evaluate the

Abbreviations: AUC, Area under the curve; CT, Computed Tomography; ROC, Receiver operating characteristic.

* Corresponding author at: Department of Diagnostic and Interventional Radiology and Neuroradiology, University Hospital Augsburg, Stenglinstr. 2, 86156 Augsburg, Germany.
E-mail address: thomas.kroencke@uk-augsburg.de (T. Kroencke).

[1] Equally contributing authors.

https://doi.org/10.1016/j.ejrad.2024.111677
Received 8 February 2024; Received in revised form 2 August 2024; Accepted 7 August 2024
Available online 9 August 2024
0720-048X/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

reason for urinary retention and the presence of renal calculi [6,9]. Urinary retention caused by urolithiasis might have an immediate treatment consequence for the patient. Therefore, it is important to answer this question sufficiently and with a short turnaround time on imaging methods, such as CT. In the past years, the computer-based recognition of texture patterns in medical image data sets ("radiomics") has become more important in radiology [10]. Radiomics can extract various quantitative values from the segmented tissue of interest, e.g. based on computed tomography data sets, which can be used for further evaluation [11]. Quantitative image analysis allows for voxel-wise extraction of data which are invisible to the human eye [10]. However, their use is not yet widespread in clinical routine due to complex and often time-consuming image postprocessing and the lack of generalizability and reproducibility. Recent studies, mainly in oncology, already indicate the potential of radiomics for tumor diagnosis, tumor prognosis and decision support [10,12–14]. Due to the rising prevalence of renal calculi, an automatic detection of hydronephrosis and calculi could fasten the report to clinicians and therefore accelerate individual patient's therapy [3,15]. Recently, algorithms for automatic kidney segmentation have been established [16]. With regard to urinary retention, radiomics might be able to detect changes in kidney texture including an increased fluid content in cases with hydronephrosis. Most radiomics studies for urolithiasis detection focused on the detection of calculi and prediction of stone composition [17–22].

Previous studies showed that radiomics can be used to differentiate phleboliths from ureteral calculi with a high degree of accuracy and to detect hydronephrosis and stone burden [17–19]. Homayounieh et al. performed automatic segmentation of the whole kidney, including renal parenchyma, renal pelvis and – if present – renal calculi. Segmentation of the entire kidney including the renal pelvis has limitations for automatic hydronephrosis detection and radiomics analysis: feature selection might mainly use the widened renal pelvis for detection of hydronephrosis as this is the most obvious discrimination to cases without hydronephrosis. However, might the algorithm be able to differ parapelvic cysts from hydronephrosis? This might almost be impossible in unenhanced low-dose CT [23].

Deep learning is based on artificial neural networks. In contrast to "traditional" machine learning where pre-defined image features are extracted, deep learning receives the whole image as input and combines feature extraction and decision-making [24]. In radiology, deep learning is mainly used in image segmentation, but also in abnormality detection and image reconstruction / enhancement [24].

A combined deep-learning based segmentation and radiomics-approach excluding the renal pelvis might exclude this potential bias and only focus on changes of the renal parenchyma in hydronephrosis.

To our current knowledge, automatic detection of hydronephrosis from the kidney parenchyma itself (without including renal pelvis), has not yet been investigated. Aim of this study was to determine whether automatic segmentation and radiomics feature extraction of kidney parenchyma itself can be used to detect hydronephrosis in a low-dose unenhanced CT scan.

2. Material and methods

2.1. Study design

This retrospective study was performed in accordance with the Declaration of Helsinki. The ethics committee of the Ludwig-Maximilians-University Munich approved this study (Protocol-Number: 20–1153) and waived the necessity of a written consent form for this study due to its retrospective and anonymized character.

2.2. Patients' population

This study included consecutive patients, who received an unenhanced low-dose CT of the abdomen at the University Hospital Augsburg between January 2019 and September 2019. Inclusion criteria comprised legal age, the evidence of urinary retention and urolithiasis. Patients with other causes of urinary retention (e.g. tumor-related obstruction), external material or poor imaging quality were excluded from this study. As reference standard for urinary retention (including grade and side of urinary retention), we used the radiological report (written by a board-certified radiologist).

2.3. Imaging protocol

Unenhanced low-dose CT of the abdomen was performed on a 128 slice dual-source Definition Flash CT scanner (Siemens Healthcare, Forchheim, Germany). Further settings were helical scan type with a tube voltage of 100 kV, automatic tube current modulation (CareDose 4D), a rotation time of 0.5 s, a pitch of 0.6 and a collimation of 128 x 0.6 mm. All scans were reconstructed in axial detection with a slice thickness and increment of 1 mm using iterative algorithm (I31f, Admire level 3, Siemens Healthineers) with a soft-tissue kernel.

2.4. Manual segmentation

CT scans were anonymized and exported from the picture archiving and communicating system (PACS). Both kidneys, the obstructed and the non-obstructed, in each patient were segmented. This allows for an individual comparison of the obstructed kidney with the non-obstructed one.

Sixty-seven patients (134 kidneys) were segmented manually by a board-certified radiologist using the open-source program 3D Slicer (v. 4.11) (https://www.slicer.org/). Segmentation was performed on axial reconstructions every few slices using automatic interpolation between them. Final manual segmentation was reviewed by the radiologist in all planes (axial, sagittal, coronal) to ensure correct segmentation. The renal pelvic system was excluded to avoid interfering factors such as parapelvic renal cysts. Fig. 1 shows an example of the manual segmentation of the kidney's parenchyma.

2.5. Automatic segmentation

A Convolutional Neural Network based on nnU-Net architecture was trained on the 67 manually segmented patients to automatically segment the kidneys in unenhanced CT images. NnU-Net is a deep learning-based segmentation method that automatically configures itself, including preprocessing, network architecture, training and post-processing and it was specifically developed for the biomedical imaging domain [25]. The training used manual segmentation masks as a ground truth to predict three classes: left kidney, right kidney, and background. CT images underwent resampling, z-score normalization, and various affine transformations to augment the dataset, including rotation, scaling, and mirroring. Training used the Nesterov momentum optimizer with a learning rate of 0.01, weight decay of 3*10–5, batch size of 2, and patch size of 128x128x128 pixels. It ran for 1000 epochs.

2.6. Radiomics feature extraction and analysis

This study complied with the guidelines of the CLEAR-checklist (Supplemental Table 1) [26]. For each kidney, 1256 radiomics features were extracted using the standard, IBSI-compliant python based library pyRadiomics [27]. Extraction parameters, including information on preprocessing, normalization, resampling and discretization are provided in Supplementary Table 2. Then, feature selection was performed with the least absolute shrinkage and selection operator (LASSO) regression, which limited the large number of features in the final model to ten key features [28]. Feature selection enhances the model's interpretability and performance by focusing on the most relevant data [29]. A logistic regression model was trained to detect hydronephrosis in the region of a single kidney. The model was developed and optimized in 5-

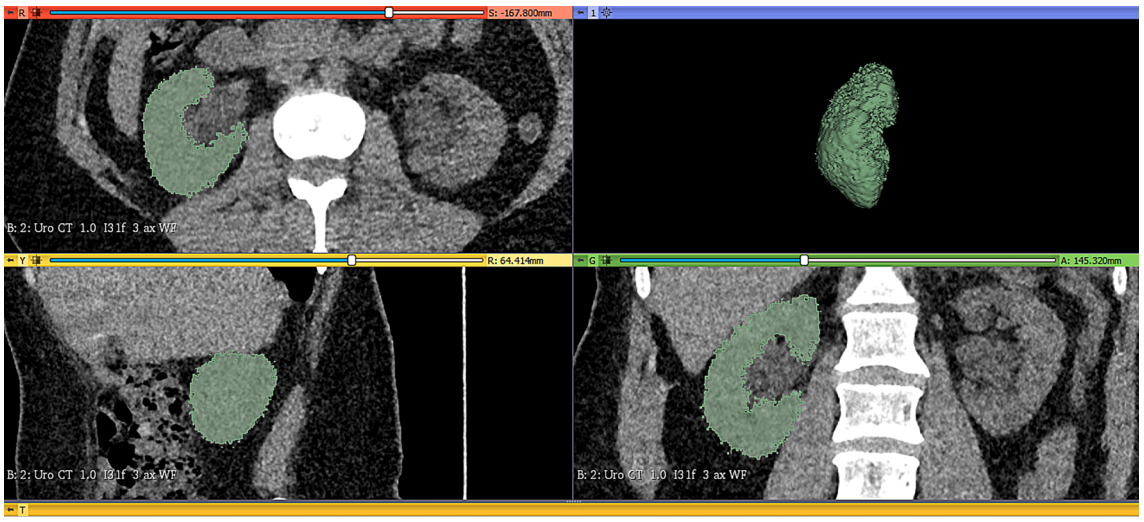

Fig. 1. Shows an example of the manual segmentation of the parenchyma of the right kidney in an unenhanced low-dose ct of the abdomen, performed by a board-certified radiologist. the renal pelvis was excluded from the segmentation.

fold cross-validation on the training dataset, and then evaluated on the test set. Separate models were trained for manual and automatic kidney segmentation masks. For model development, the AutoRadiomics framework was used [30]. The framework handles image preprocessing, feature extraction, feature selection, modeling, and evaluation. We utilized Bayesian optimization to select the optimal hyperparameters. All other parameters remained as a default configuration. Fig. 2 shows a visual representation of the study workflow.

2.7. Statistical analysis

Patients were randomly split into training (0.7) and test cohorts (0.3) in a stratified way. Training cohort was further split into 5 folds for cross-validation. The model was internally tested. Receiver operating characteristic (ROC) curves were generated for each independent variable and the area under the curve (AUC) was calculated. Diagnostic efficacy was compared using sensitivity and specificity at the optimal threshold selected with the Youden index and was reported with 95 % confidence intervals obtained with the bootstrap technique. Automatic segmentations were evaluated using voxel-based metrics: Dice coefficient, Jaccard index, precision, and recall. Holm-Bonferroni correction was applied to correct for multiple comparisons. A p-value < 0.05 was indicative of statistical significance. All statistical analyses were implemented in the programming language Python (version 3.8).

3. Results

3.1. Patient population

A total of 272 patients with suspicion for hydronephrosis on unenhanced low-dose CT of the abdomen between 3rd of January 2019 and 13th of September 2019 were identified. 177 of the unenhanced low-dose CT of the abdomen were excluded because of missing evidence of an urolithiasis, hydronephrosis, external material or poor imaging quality (Fig. 3). 95 patients (training cohort: mean age 49 years (range 21–83), male 53/67 [79 %]; test cohort: mean age 50 years (range 27–79), male 21/28 [75 %]) were included in this study. Demographic as well as clinical characteristics of training and test cohorts are presented in Table 1. Hydronephrosis grade 1 and 2 were the most common among patients in this study, whereas grade 3 was rare and grade 4 was absent.

3.2. Automatic segmentation

The results of automatic kidney segmentation are shown in Table 2. In general, the automatic segmentation masks were highly precise, with mean Dice coefficient of 0.87 for the right kidney and 0.91 for the left kidney, when compared with manual segmentations. Fig. 4 shows examples of manual and automatic segmentation of the kidneys' parenchyma.

3.3. Hydronephrosis detection

The ROC curves for radiomics models using manual segmentations and automated segmentations are presented in Fig. 5. Confusion matrices for manual and automatic segmentations are shown in Table 3. We also performed a further analysis, showing the number of falsely predicted cases according to the grade of hydronephrosis (Table 3).

The numerical results of manual and automatic segmentations are summed up in Table 4. Detection of hydronephrosis using manual kidney segmentations achieved an AUC of 0.84, a sensitivity of 75 % (95 % CI 72–93), specificity of 82 % (95 % CI 66–96), PPV of 81 % (95 % CI 62–96) and a NPV of 77 % (95 % CI 60–90). Detection of hydronephrosis using automatic kidney segmentations achieved an AUC of 0.83, a sensitivity of 86 % (95 % CI 72–97), specificity of 64 % (95 % CI 46–82), a PPV of 71 % (95 % CI 55–86) and a NPV of 82 % (95 % CI 64–96).

We further analyzed the radiomics features used to detect hydronephrosis in the final model with automated segmentation. Table 5 provides details on the features of highest importance to the model, including their class, applied filter and the coefficient in the final model. Fig. 6 shows the differences in distribution of the most important features between kidneys with and without hydronephrosis. Additionally, we included waterfall plots showing the distribution of true and false predictions relative to the classification boundary as Supplementary **Figure S1**.

4. Discussion

Main goal of our study was to investigate whether automatic segmentation and radiomics feature extraction of the kidney parenchyma can be used to detect urinary retention with reliable results compared to the radiologists as gold standard.

A deep-learning-based automated segmentation and analysis of radiomics features of the kidney's parenchyma (excluding renal pelvis) can detect hydronephrosis with high accuracy and might be useful

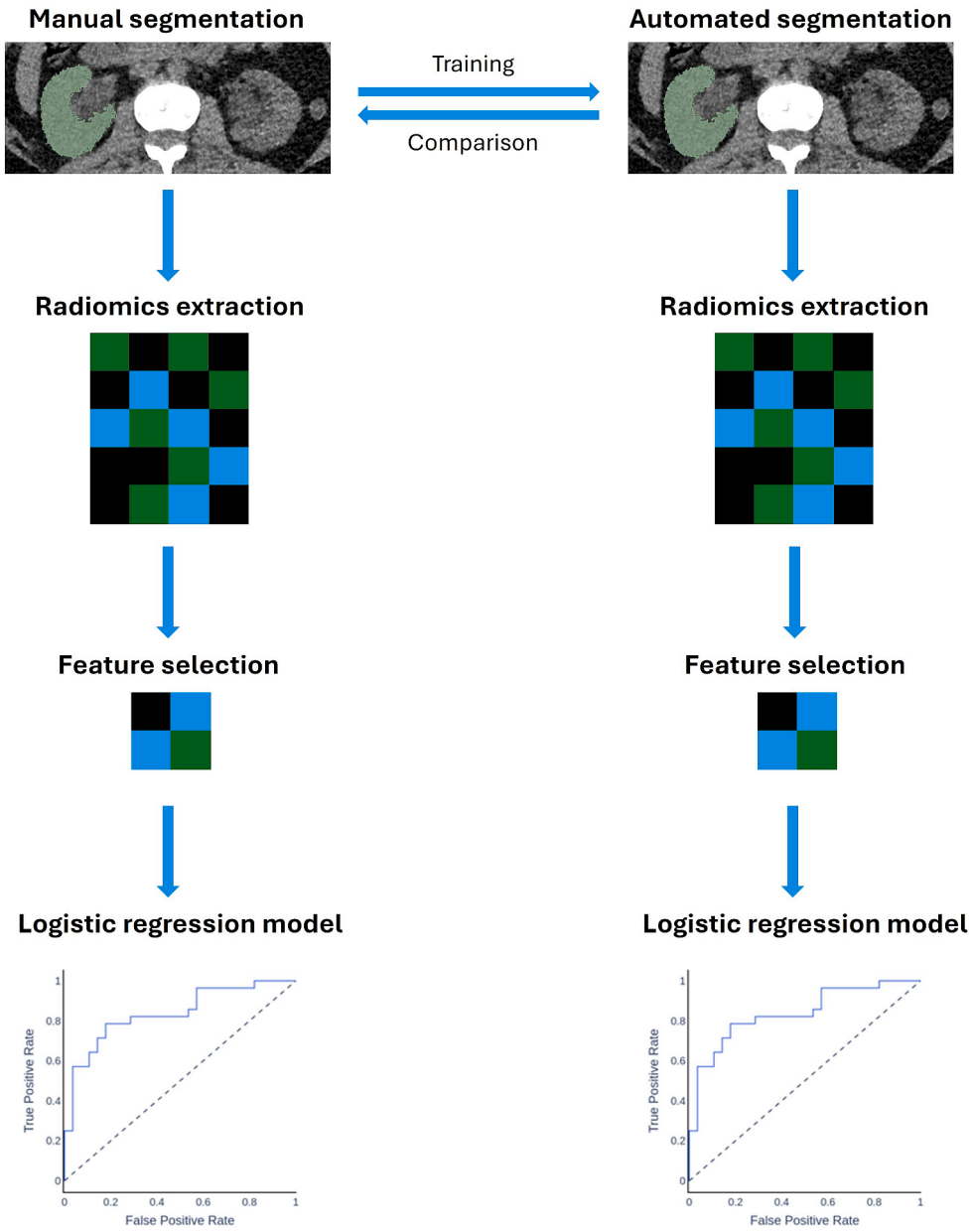

Fig. 2. Detailed description of the approach combining manual and automatic, deep-learning-based segmentation and radiomics extraction for detection of hydronephrosis.

especially in patients with present parapelvic cysts. The study suggests that textural changes of the renal parenchyma might be present in patients with hydronephrosis that can be detected using radiomics.

In patients presenting in the emergency department with acute flank pain and suspected diagnosis of urolithiasis after clinical examination and ultrasound, unenhanced low-dose CT of the abdomen is the gold standard to detect renal hydronephrosis and – if present – renal calculi [6,31].

A recent study also showed automated hydronephrosis detection on ultrasound images using a deep-learning-based algorithm [32]. Main advantages of ultrasound are the fast and wide availability as well as the lack of radiation exposure. However, detection of calculi is often limited using ultrasound.

The advantage of a CT scan in comparison to an ultrasound examination is the high sensitivity (94–100 %) and specificity (92–100 %) in detection of ureteral calculi and their location [9,33–36]. Nevertheless, the radiation exposure should always be considered especially in young patients and pregnant women.

In our study more men suffer from urolithiasis (training dataset 79 % and test dataset 73 %) compared to women. This gender difference has been published in several studies before, keeping in mind that the amount of women suffering from urolithiasis is recently increasing [2,37–39].

Using manual segmentation, hydronephrosis can be detected with an AUC of 0.83 [95 % CI: 0.72–0.93]. Although the present results may be promising, the possible inter-variability and very time-consuming workflow of manual segmentation must be addressed. Therefore, it is not possible to integrate the manual segmentation of the kidneys into everyday clinical practice. Some studies already indicated that manual segmentation is prone to failure, has a high variability and is too time-consuming [40–42]. Therefore, an automatic segmentation of the kidney's parenchyma is on the one hand investigator-independent and on the other hand less time-consuming. With the help of radiomics, more data can be extracted from one image than is ever possible for the human eye [10]. In our study, we created a deep-learning based algorithm for automatic segmentation after final manual segmentation of the kidneys.

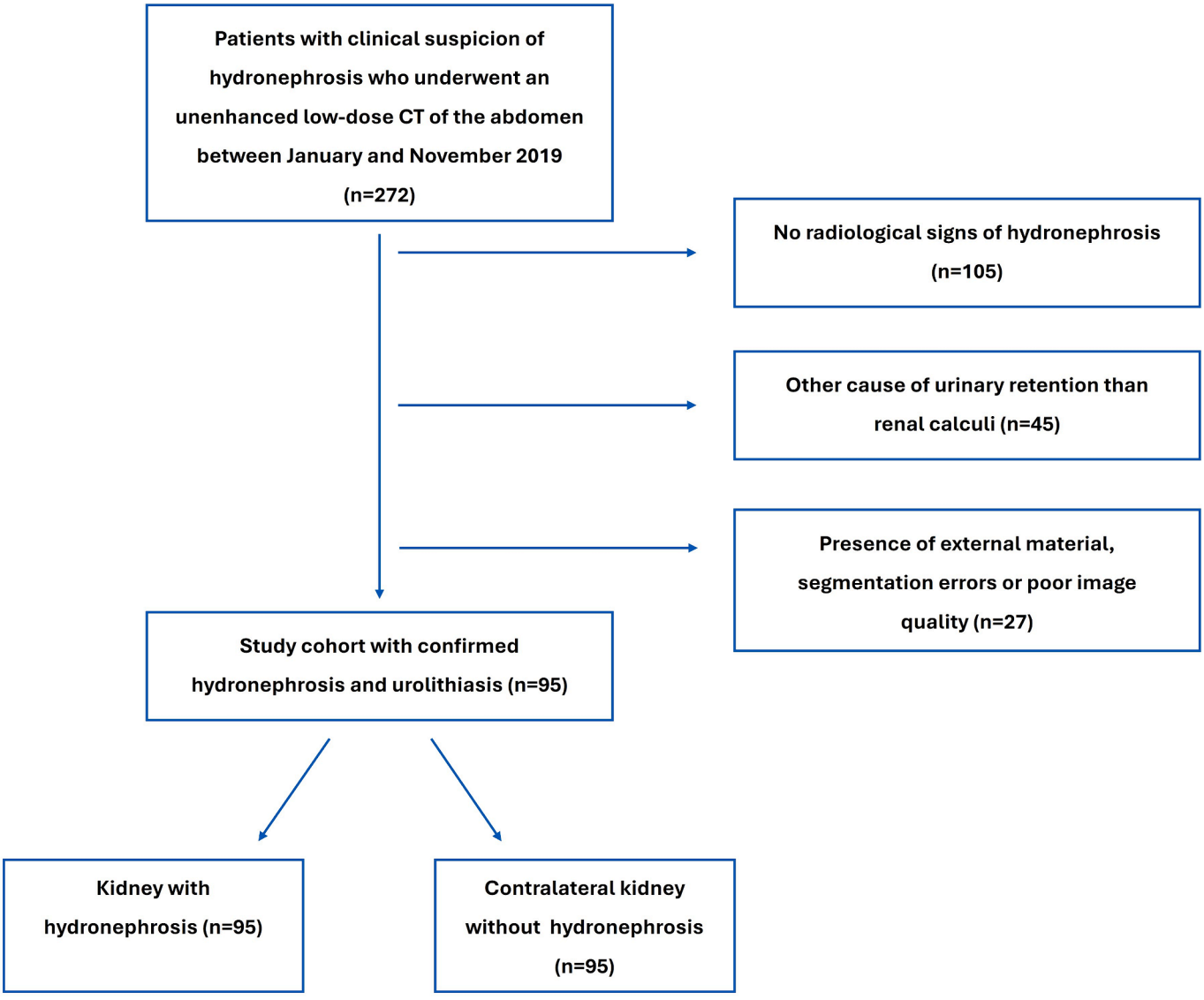

Fig. 3. Study flowchart for inclusion and exclusion criteria.

Table 1
Demographic and clinical characteristics of training and test cohorts.

	Training cohort (n = 67)	Test cohort (n = 28)
Men	53 (79 %)	21 (75 %)
Women	14 (21 %)	7 (25 %)
Mean age	49 ± 15	50 ± 16
Stone location:		
- renal	3 (4 %)	0
- ureter	61 (91 %)	25 (89 %)
- bladder	3 (4 %)	3 (11 %)
Stone size:		
- <4mm	27 (40 %)	11 (39 %)
- 4–9,9mm	38 (57 %)	16 (57 %)
- >=10 mm	2 (3 %)	1 (4 %)
Hydronephrosis grade:		
- 1	27 (40 %)	13 (46 %)
- 2	30 (45 %)	14 (50 %)
- 3	10 (15 %)	1 (4 %)
- 4	0	0
Side:		
- left	33 (49 %)	10 (36 %)
- right	34 (51 %)	18 (64 %)

Table 2
Results of automatic kidney segmentation in the test set.

	Dice	Jaccard	Precision	Recall
Right kidney	0.87 ± 0.18	0.80 ± 0.20	0.86 ± 0.17	0.93 ± 0.14
Left kidney	0.91 ± 0.06	0.85 ± 0.10	0.88 ± 0.08	0.95 ± 0.07

Metrics for comparison between automatic and manual kidney segmentations.

The automatic segmentation of the kidney showed good results compared to the manual segmentation. Our deep-learning model was able to automatically segment the parenchyma of the kidneys with a good concordance to manual segmentation (Dice score right: 0.87 and left 0.91, Jaccard right: 0.8 and left: 0.85). The high accuracy of automatic segmentation of the kidneys was shown in other previous studies with comparable and partially even better results [43,44]. For example Da Cruz et al. showed that automatic segmentation of the kidneys (including the pelvis) using artificial intelligence and a deep learning algorithm is possible with a high sensitivity (97.42 %) and high specificity (99.94 %) [45]. Due to the fact that there are now several possible automatic segmentation programs for the kidneys, it will be important to evaluate more precisely which is the most reliable method.

Homayounieh et al. proved that with an automated segmentation

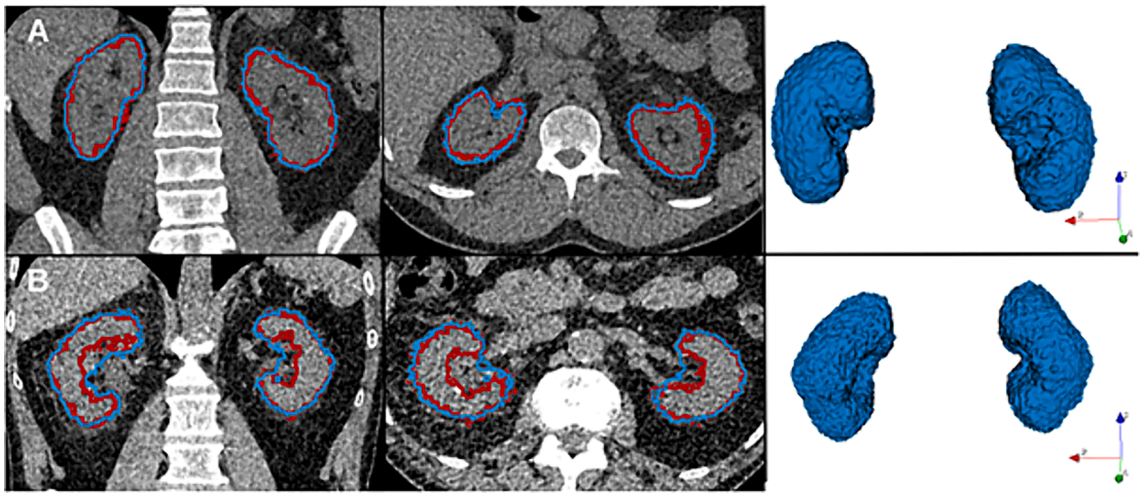

Fig. 4. Examples of automatic segmentation (blue contour) compared with the reference manual volumetry (red contour) in two patients with hydronephrosis. Panel A (top) presents a case with hydronephrosis on the left side, panel B (bottom) with hydronephrosis on the right side. This figure highlights the high concordance between manual and automatic segmentation.

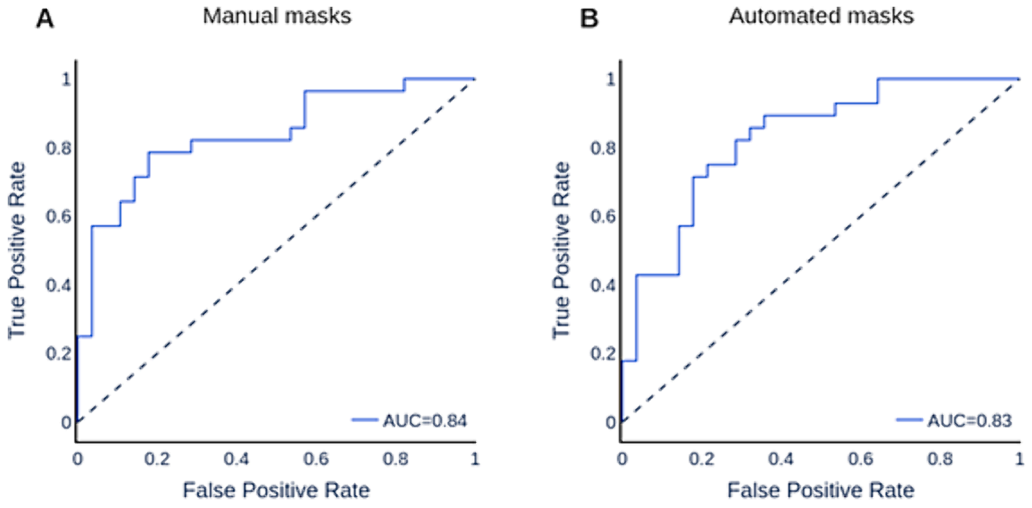

Fig. 5. ROC curves for radiomics models using manual segmentations (panel A, AUC=0.84) and automated segmentations (panel B, AUC=0.83) for detection of hydronephrosis using segmentations and radiomics features.

Table 3
Confusion matrices for automatic and manual segmentation.

Manual segmentations			
		True	
		Positive	**Negative**
Predicted	**Positive**	22	6
	Negative	5	23
		2/5 HN grade I	
		3/5 HN grade II	
Automatic segmentations			
		True	
		Positive	**Negative**
Predicted	**Positive**	25	3
	Negative	10	18
		6/10 HN grade I	
		4/10 HN grade II	

HN: hydronephrosis.

program hydronephrosis can be detected with a high AUC (0.99) [23]. In comparison to the present study, they included the kidney and the renal pelvis in the automatic segmentation while patients with large renal cysts (>3 cm) were excluded. Additionally, they only included patients who had a calculus in the renal pelvis [23]. To avoid the possibility that inclusion of the renal pelvis itself, calculus in the renal pelvis, congenital ampullary pelvis systems or peripelvic/parapelvic renal cysts may influence the parameters determined by radiomics, we decided to only segmentate the kidney's parenchyma itself. Our aim was to reduce the susceptibility of parameters to interference with renal cysts, especially parapelvic and peripelvic renal cysts, or ampullary renal pelvis systems because these features can be falsely attributed with urinary retention. Since the present study shows that the automated segmentation of the renal parenchyma in comparison with the manual segmentation is possible with a good concordance (Dice score right 0.87 and left 0.91) we assume a better informative value and a higher accuracy with regard to urinary retention by excluding the pelvic system. After using the automated segmentation of the kidney's parenchyma, the automated detection of hydronephrosis was possible with a sensitivity and specificity of 86 % and 64 %, respectively. We therefore assume that the use of radiomics can highlight low-dose unenhanced CT scans with

Table 4
Comparison of diagnostic performance using ROC analysis and binary classification metrics (95% confidence intervals provided in brackets).

	AUC	Sensitivity	Specificity	PPV	NPV
Radiomics (manual masks)	0.84 [0.72–0.93]	0.75 [0.59–0.90]	0.82 [0.66–0.96]	0.81 [0.62–0.96]	0.77 [0.60–0.90]
Radiomics (automatic masks)	0.83 [0.72–0.93]	0.86 [0.72–0.97]	0.64 [0.46–0.82]	0.71 [0.55–0.86]	0.82 [0.64–0.96]

AUC: area under the ROC curve, PPV: positive predictive value, NPV: negative predictive value.

Table 5
Ten most important radiomics features in the radiomics signature using automated segmentations.

No.	Feature class	Feature	Filter applied	coefficient[1]
1	First order	90Percentile	wavelet-LLL	−1.90
2	GLCM	Imc2	–	−1.64
3	GLDM	Dependence Variance	LoG (σ = 5 mm)	−1.14
4	GLSZM	Low Gray Level Zone Emphasis	wavelet-LLH	1.11
5	GLDM	Small Dependence Low Gray Level Emphasis	wavelet-HHH	−1.07
6	shape	Sphericity	–	−1.0
7	GLSZM	Large Area Low Gray Level Emphasis	wavelet-HHL	0.98
8	First order	Median	wavelet-HHH	0.96
9	GLSZM	Zone Entropy	wavelet-HHH	0.96
10	First order	Skewness	LoG (σ = 5 mm)	−0.96

Ten features selected after least absolute shrinkage and selection operator (LASSO) regression. GLCM: Gray Level Co-occurrence Matrix, GLDM: Gray Level Dependence Matrix, GLSZM: Gray Level Size Zone Matrix, wavelet filters: L=low, H=high, LoG: Laplacian of Gaussian.
[1] coefficient in the logistic regression model.

suspicion of hydronephrosis to the radiologists and can be prioritized in reporting. The high sensitivity of the automated segmentation is of great value in a screening method, while the radiologists can add more specificity (82 %). This may lead to a faster treatment of the patient. Automated segmentation, which achieved good concordance with manual results (Dice scores: 0.87 right, 0.91 left), eliminates inter-operator variability. Despite lower specificity, it presents a viable alternative, offering high sensitivity and consistent results, making it particularly useful in high-throughput screening environments where rapid and reliable detection of hydronephrosis is critical in clinical routine. Additionally, it significantly reduces the time required for analysis.

Our radiomics model detected hydronephrosis in renal parenchyma only, excluding the renal pelvis. These results suggest that there might be changes in the parenchyma itself that can be detected by textural analyses. Especially in patients with parapelvic cysts, the segmentation of both, the renal pelvis and parenchyma, might introduce a bias misinterpreting cysts as hydronephrosis. This developed algorithm detects hydronephrosis independently of widened renal pelvis / cysts and might therefore be applied also in patients with concurrent parapelvic cysts.

This automatic segmentation approach, combined with radiomics analysis, might support radiologists in clinical routine, especially in times of high throughput, and might also act as a second reader. In the rapidly growing era of AI-based image analysis, this might be another tool in daily routine in the future. As this algorithm does not depend on widening of the renal pelvis, it can be used also in patients with parapelvic cysts. With an AUC of 0.83 and 0.84 respectively, this algorithm

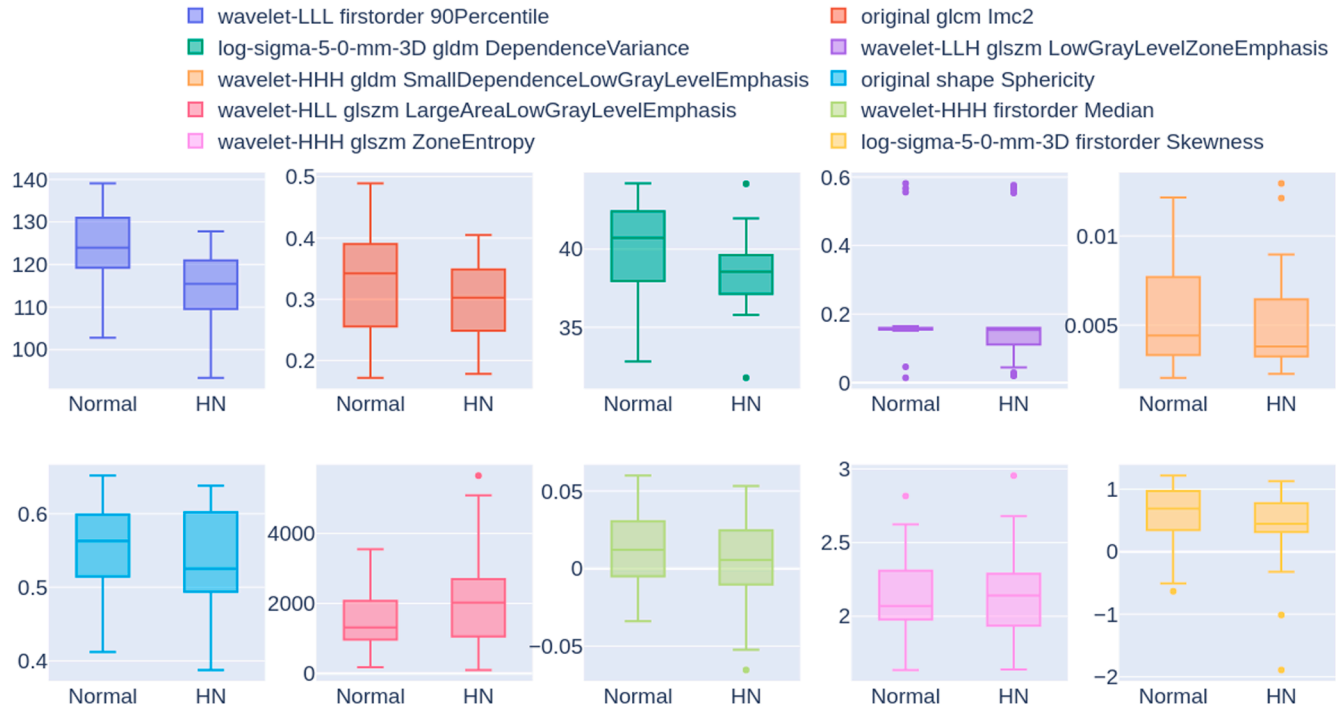

Fig. 6. Boxplots of 10 most important radiomics features selected in the radiomics model using automated segmentations in the test set. Data are shown for cases with hydronephrosis (*HN – hydronephrosis)* and cases without hydronephrosis (=normal). Different radiomics features are shown in different colors.

might support diagnosis or highlight findings prior to the radiological report. However, it can not be used as a stand-alone option and critical review by a radiologist is crucial. Regarding also the limitations that are discussed below, further studies are necessary to critically address the potential clinical application of AI-based algorithms for hydronephrosis detection in clinical routine.

The main limitation of this study is its retrospective character. The used radiomics python-based radiology is an open-source product and therefore easy to implement. However, the application of this algorithm requires prior image extraction and image analysis combined with technical know-how and standardization of the analysis which limits the use in clinical routine at the moment and might be addressed by further technical advances and staff training. Main limitations of radiomics are further the lack of generalizability and the high susceptibility to image artifacts which might be addressed in external validation studies.

Another limitation is that the patient collective was comparatively low, and the developed deep learning algorithm has to be investigated in a larger patient collective. Since artificial intelligence and deep learning algorithms are increasing especially in the field of radiology it is important to focus on one algorithm after developing and testing different algorithms, which artificial intelligence has achieved the best results and can be applied in clinical routine.

These limitations might restrict the findings of the study for clinical application at the moment. However, this approach combining deep learning-based segmentation and radiomics for hydronephrosis detection adds a further piece to the rapidly growing AI-based algorithms in medical image analysis which becomes more and more important also in clinical routine. Further (multi-center) studies addressing the applicability and implementation of automatic segmentation and radiomics extraction, including also external validation of the algorithm, in clinical practice are necessary.

5. Conclusion

An approach combining deep-learning based automated segmentation of the kidney's parenchyma and radiomics analysis allows an accurate and fast automatic detection of hydronephrosis in unenhanced low-dose CT scans of the abdomen and is comparable to manual segmentation of board-certified radiologists. This algorithm is not dependent on changes in the renal pelvis and the presence of calculi and might therefore also be applied in patients with concurrent parapelvic cysts. In the rapidly growing era of AI-based analysis of medical images, inclusion of this algorithm might help radiologists in clinical routine especially in times of high throughput, might act as a second reader and increase time efficiency. Further studies, including external validation, are necessary to further adapt this method for use in clinical routine.

6. Statements and declarations

This study was partially funded by the Medical Faculty of the University of Augsburg. Florian Schwarz reports a relationship with Siemens Healthineers that includes: speaking and lecture fees. Thomas Kroencke reports a relationship with Siemens Healthineers that includes funding grants.

CRediT authorship contribution statement

Judith Becker: Writing – original draft, Visualization, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Piotr Woźnicki:** Writing – review & editing, Visualization, Validation, Software, Methodology, Formal analysis, Conceptualization. **Josua A. Decker:** Writing – review & editing, Supervision, Methodology, Formal analysis. **Franka Risch:** Writing – review & editing, Software, Formal analysis, Data curation. **Ramona Wudy:** Writing – review & editing, Resources, Investigation. **David Kaufmann:** Writing – review & editing, Resources, Investigation. **Luca Canalini:** Writing – review & editing, Software, Methodology, Investigation, Formal analysis. **Claudia Wollny:** Writing – review & editing, Resources, Project administration, Funding acquisition. **Christian Scheurig-Muenkler:** Writing – review & editing, Supervision, Resources. **Thomas Kroencke:** Writing – review & editing, Supervision, Resources, Project administration. **Stefanie Bette:** Writing – review & editing, Validation, Supervision, Formal analysis, Conceptualization. **Florian Schwarz:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: [Thomas Kroencke reports financial support was provided by Siemens Healthineers. Florian Schwarz reports financial support was provided by Siemens Healthineers. Judith Becker reports financial support was provided by University Augsburg. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.].

Data availability

Data will be made available on request.

Acknowledgements

This study received financial support from the University Augsburg. The open access publication of this article was supported by the DFG sponsored Open Access Fund of the University of Augsburg.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejrad.2024.111677.

References

[1] W. Wang, J. Fan, G. Huang, J. Li, X. Zhu, Y. Tian, L. Su, Prevalence of kidney stones in mainland China: a systematic review, Sci. Rep. 7 (2017) 41630, https://doi.org/10.1038/srep41630.
[2] T. Yasui, M. Iguchi, S. Suzuki, K. Kohri, Prevalence and epidemiological characteristics of urolithiasis in japan: national trends between 1965 and 2005, Urology 71 (2008) 209–213, https://doi.org/10.1016/j.urology.2007.09.034.
[3] A. Hesse, E. Brändle, D. Wilbert, K.-U. Köhrmann, P. Alken, Study on the prevalence and incidence of urolithiasis in Germany comparing the years 1979 vs 2000, Eur. Urol. 44 (2003) (2003) 709–713, https://doi.org/10.1016/S0302-2838(03)00415-9.
[4] M. Gottlieb, B. Long, A. Koyfman, The evaluation and management of urolithiasis in the ED: A review of the literature, Am. J. Emerg. Med. 36 (2018) 699–706, https://doi.org/10.1016/j.ajem.2018.01.003.
[5] N.J. Rukin, Z.A. Siddiqui, E.C.P. Chedgy, B.K. Somani, Trends in upper tract stone disease in england: evidence from the hospital episodes statistics database, Urol. Int. 98 (2017) 391–396, https://doi.org/10.1159/000449510.
[6] Deutsche Gesellschaft für Urologie e.V. (DGU), S2k-Leitlinie Diagnostik, Therapie und Metaphylaxe der Urolithiasis, Https://Www.Awmf.Org/Leitlinien/Detail/Ll/043-025.Html (2019).
[7] R.J. Gaspari, K. Horst, Emergency ultrasound and urinalysis in the evaluation of flank pain, Acad. Emerg. Med. 12 (2005) 1180–1184, https://doi.org/10.1197/j.aem.2005.06.023.
[8] M. Okumus, Correlation of volume, position of stone, and hydronephrosis with microhematuria in patients with solitary urolithiasis, Med. Sci. Monit. 19 (2013) 295–299, https://doi.org/10.12659/MSM.889077.
[9] J.M. Weinrich, P. Bannas, M. Regier, S. Keller, L. Kluth, G. Adam, F.O. Henes, Low-dose CT for evaluation of suspected urolithiasis: diagnostic yield for assessment of alternative diagnoses, Am. J. Roentgenol. 210 (2018) 557–563, https://doi.org/10.2214/AJR.17.18552.
[10] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures, they are data, Radiology 278 (2016) 563–577, https://doi.org/10.1148/radiol.2015151169.
[11] V. Kumar, Y. Gu, S. Basu, A. Berglund, S.A. Eschrich, M.B. Schabath, K. Forster, H. J.W.L. Aerts, A. Dekker, D. Fenstermacher, D.B. Goldgof, L.O. Hall, P. Lambin,

Y. Balagurunathan, R.A. Gatenby, R.J. Gillies, Radiomics: the process and the challenges, Magn. Reson. Imaging 30 (2012) 1234–1248, https://doi.org/10.1016/j.mri.2012.06.010.
[12] D. Gu, Y. Hu, H. Ding, J. Wei, K. Chen, H. Liu, M. Zeng, J. Tian, CT radiomics may predict the grade of pancreatic neuroendocrine tumors: a multicenter study, Eur. Radiol. 29 (2019) 6880–6890, https://doi.org/10.1007/s00330-019-06176-x.
[13] P. Lambin, R.T.H. Leijenaar, T.M. Deist, J. Peerlings, E.E.C. de Jong, J. van Timmeren, S. Sanduleanu, R.T.H.M. Larue, A.J.G. Even, A. Jochems, Y. van Wijk, H. Woodruff, J. van Soest, T. Lustberg, E. Roelofs, W. van Elmpt, A. Dekker, F. M. Mottaghy, J.E. Wildberger, S. Walsh, Radiomics: the bridge between medical imaging and personalized medicine, Nat. Rev. Clin. Oncol. 14 (2017) 749–762, https://doi.org/10.1038/nrclinonc.2017.141.
[14] M. Bogowicz, D. Vuong, M.W. Huellner, M. Pavic, N. Andratschke, H.S. Gabrys, M. Guckenberger, S. Tanadini-Lang, CT radiomics and PET radiomics: ready for clinical implementation?, The Quarterly Journal of Nuclear Medicine and Molecular Imaging : Official Publication of the Italian Association of Nuclear Medicine (AIMN) [and] the International Association of Radiopharmacology (IAR), [and] Section of the Society Of... 63 (2019) 355–370. Doi: 10.23736/S1824-4785.19.03192-3.
[15] C.D. Scales, A.C. Smith, J.M. Hanley, C.S. Saigal, Prevalence of kidney stones in the United States, Eur. Urol. 62 (2012) 160–165, https://doi.org/10.1016/j.eururo.2012.03.052.
[16] D. Li, C. Xiao, Y. Liu, Z. Chen, H. Hassan, L. Su, J. Liu, H. Li, W. Xie, W. Zhong, B. Huang, Deep segmentation networks for segmenting kidneys and detecting kidney stones in unenhanced abdominal CT images, Diagnostics 12 (2022) 1788, https://doi.org/10.3390/diagnostics12081788.
[17] E.J. Lim, D. Castellani, W.Z. So, K.Y. Fong, J.Q. Li, H.Y. Tiong, N. Gadzhiev, C. T. Heng, J.-Y.-C. Teoh, N. Naik, K. Ghani, K. Sarica, J. De La Rosette, B. Somani, V. Gauhar, Radiomics in urolithiasis: systematic review of current applications, limitations, and future directions, J. Clin. Med. 11 (2022), https://doi.org/10.3390/jcm11175151.
[18] T. De Perrot, J. Hofmeister, S. Burgermeister, S.P. Martin, G. Feutry, J. Klein, X. Montet, Differentiating kidney stones from phleboliths in unenhanced low-dose computed tomography using radiomics and machine learning, Eur. Radiol. 29 (2019) 4776–4782, https://doi.org/10.1007/s00330-019-6004-7.
[19] L. Tang, W. Li, X. Zeng, R. Wang, X. Yang, G. Luo, Q. Chen, L. Wang, B. Song, Value of artificial intelligence model based on unenhanced computed tomography of urinary tract for preoperative prediction of calcium oxalate monohydrate stones in vivo, Ann Transl Med 9 (2021) 1129, https://doi.org/10.21037/atm-21-965.
[20] B.M. Zeeshan Hameed, B. Somani, N. Naik, A. Talasila, M. Shah, S. Reddy, G. Sachdev, R. Hussein Beary, P. Hegde, Application of deep learning convolutional neural network in prediction of stone location, skin to stone distance and composition in renal lithiasis: a single center pilot study, Eur. Urol. 79 (2021) S336, https://doi.org/10.1016/S0302-2838(21)00624-2.
[21] J. Zheng, H. Yu, J. Batur, Z. Shi, A. Tuerxun, A. Abulajiang, S. Lu, J. Kong, L. Huang, S. Wu, Z. Wu, Y. Qiu, T. Lin, X. Zou, A multicenter study to develop a non-invasive radiomic model to identify urinary infection stone in vivo using machine-learning, Kidney Int. 100 (2021) 870–880, https://doi.org/10.1016/j.kint.2021.05.031.
[22] A. Parakh, H. Lee, J.H. Lee, B.H. Eisner, D.V. Sahani, S. Do, Urinary stone detection on CT images using deep convolutional neural networks: evaluation of model performance and generalization, Radiol Artif Intell 1 (2019) e180066.
[23] F. Homayounieh, R. Doda Khera, B.C. Bizzo, S. Ebrahimian, A. Primak, B. Schmidt, S. Saini, M.K. Kalra, Prediction of burden and management of renal calculi from whole kidney radiomics: a multicenter study, Abdom. Radiol. 46 (2021) 2097–2106, https://doi.org/10.1007/s00261-020-02865-0.
[24] M.A. Mazurowski, M. Buda, A. Saha, M.R. Bashir, Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI, J. Magn. Reson. Imaging 49 (2019) 939–954, https://doi.org/10.1002/jmri.26534.
[25] F. Isensee, P.F. Jaeger, S.A.A. Kohl, J. Petersen, K.H. Maier-Hein, nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation, Nat. Methods 18 (2021) 203–211, https://doi.org/10.1038/s41592-020-01008-z.
[26] B. Kocak, B. Baessler, S. Bakas, R. Cuocolo, A. Fedorov, L. Maier-Hein, N. Mercaldo, H. Müller, F. Orlhac, D. Pinto Dos Santos, A. Stanzione, L. Ugga, A. Zwanenburg, CheckList for evaluation of radiomics research (CLEAR): a step-by-step reporting guideline for authors and reviewers endorsed by ESR and EuSoMII. Insights Imaging 14 (2023) 75, https://doi.org/10.1186/s13244-023-01415-8.
[27] J.J.M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R.G.H. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, H.J.W.L. Aerts, Computational radiomics system to decode the radiographic phenotype, Cancer Res. 77 (2017) e104–e107, https://doi.org/10.1158/0008-5472.CAN-17-0339.
[28] J. Ranstam, J.A. Cook, LASSO regression, Br. J. Surg. 105 (2018) 1348, https://doi.org/10.1002/bjs.10895.
[29] R. Tibshirani, Regression shrinkage and selection via the Lasso, J. R Stat. Soc. Ser. B Stat. Methodol. 58 (1996) 267–288.
[30] P. Woznicki, F. Laqua, T. Bley, B. Baeßler, AutoRadiomics: a framework for reproducible radiomics research, Front. Radiol. 2 (2022) 919133, https://doi.org/10.3389/fradi.2022.919133.
[31] L. Mills, E.J. Morley, Z. Soucy, G.M. Vilke, S.H.F. Lam, Ultrasound for the Diagnosis and Management of Suspected Urolithiasis in the Emergency Department, J. Emerg. Med. 54 (2018) 215–220, https://doi.org/10.1016/j.jemermed.2017.09.020.
[32] R. Alexa, J. Kranz, R. Kramann, C. Kuppe, R. Sanyal, S. Hayat, L.F. Casas Murillo, T. Hajili, M. Hoffmann, M. Saar, Harnessing artificial intelligence for enhanced renal analysis: automated detection of hydronephrosis and precise kidney segmentation, Eur. Urol. Open Sci. 62 (2024) 19–25, https://doi.org/10.1016/j.euros.2024.01.017.
[33] O.F. Miller, S.K. Rineer, S.R. Reichard, R.G. Buckley, M.S. Donovan, I.R. Graham, W.B. Goff, C.J. Kane, Prospective comparison of unenhanced spiral computed tomography and intravenous urogram in the evaluation of acute flank pain, Urology 52 (1998) 982–987, https://doi.org/10.1016/S0090-4295(98)00368-9.
[34] J.-H. Wang, S.-H. Shen, S.-S. Huang, C.-Y. Chang, Prospective comparison of unenhanced spiral computed tomography and intravenous urography in the evaluation of acute renal colic, J. Chin. Med. Assoc. 71 (2008) 30–36, https://doi.org/10.1016/S1726-4901(08)70069-8.
[35] O. Niall, J. Russell, R. MacGregor, H. Duncan, J. Mullins, A comparison of noncontrast computerized tomography with excretory urography in the assessment of acute flank pain, J. Urol. 161 (1999) 534–537.
[36] M. Müller, R. Heicappell, U. Steiner, E. Merkle, A.J. Aschoff, K. Miller, The average dose-area product at intravenous urography in 205 adults, Br. J. Radiol. 71 (1998) 210–212, https://doi.org/10.1259/bjr.71.842.9579185.
[37] V. Castiglione, F. Jouret, O. Bruyère, B. Dubois, A. Thomas, D. Waltregny, A.-C. Bekaert, É. Cavalier, R. Gadisseur, Épidémiologie de la lithiase urinaire en Belgique sur base d'une classification morpho-constitutionnelle, Nephrol. Ther. 11 (2015) 42–49, https://doi.org/10.1016/j.nephro.2014.08.003.
[38] C. Seitz, H. Fajkovic, Epidemiological gender-specific aspects in urolithiasis, World J. Urol. 31 (2013) 1087–1092, https://doi.org/10.1007/s00345-013-1140-1.
[39] D. Prezioso, E. Illiano, G. Piccinocchi, C. Cricelli, R. Piccinocchi, A. Saita, C. Micheli, A. Trinchieri, Urolithiasis in Italy: an epidemiological study, Archivio Italiano Di Urologia e Andrologia 86 (2014) 99, https://doi.org/10.4081/aiua.2014.2.99.
[40] M. Ben Abdallah, M. Blonski, S. Wantz-Mezieres, Y. Gaudeau, L. Taillandier, J.-M. Moureaux, Statistical evaluation of manual segmentation of a diffuse low-grade glioma MRI dataset, in: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2016: pp. 4403–4406. Doi: 10.1109/EMBC.2016.7591703.
[41] R.B. Dubey, M. Hanmandlu, S. Vasikarla, Evaluation of Three Methods for MRI Brain Tumor Segmentation, in: In: 2011 Eighth International Conference on Information Technology: New Generations, 2011, pp. 494–499, https://doi.org/10.1109/ITNG.2011.92.
[42] I. Arganda-Carreras, V. Kaynig, C. Rueden, K.W. Eliceiri, J. Schindelin, A. Cardona, H. Sebastian Seung, Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification, Bioinformatics 33 (2017) 2424–2426, https://doi.org/10.1093/bioinformatics/btx180.
[43] Guanyu Yang, Jinjin Gu, Yang Chen, Wangyan Liu, Lijun Tang, Huazhong Shu, C. Toumoulin, Automatic kidney segmentation in CT images based on multi-atlas image registration, in: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2014: pp. 5538–5541. Doi: 10.1109/EMBC.2014.6944881.
[44] E. Zhao, Y. Liang, H. Fan, Contextual information-aided kidney segmentation in CT sequences, Opt. Commun. 290 (2013) 55–62, https://doi.org/10.1016/j.optcom.2012.10.033.
[45] L.B. da Cruz, J.D.L. Araújo, J.L. Ferreira, J.O.B. Diniz, A.C. Silva, J.D.S. de Almeida, A.C. de Paiva, M. Gattass, Kidney segmentation from computed tomography images using deep neural network, Comput. Biol. Med. 123 (2020) 103906, https://doi.org/10.1016/j.compbiomed.2020.103906.