

Development and external validation of a machine learning-based model to classify uric acid stones in patients with kidney stones of Hounsfield units <800

Ben H. Chew

University of British Columbia, Stone Centre at Vancouver General Hospital

Victor KF. Wong

University of British Columbia, Stone Centre at Vancouver General Hospital

Abdulghafour Halawani

King Abdulaziz University

Sujin Lee

Infinyx, AI research team

Sangyeop Baek

Infinyx, AI research team

Hoyong Kang

Infinyx, AI research team

Kyo Chul Koo (✉ gckoo@yuhs.ac)

Yonsei University College of Medicine

Research Article

Keywords: decision support techniques, machine learning, urolithiasis, validation

Posted Date: July 6th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3133615/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Urolithiasis on September 30th, 2023. See the published version at <https://doi.org/10.1007/s00240-023-01490-y>.

Abstract

The correct diagnosis of uric acid (UA) stones has important clinical implications since patients with a high risk of perioperative morbidity may be spared surgical intervention and be offered alkalization therapy. We developed and validated a machine learning (ML)-based model to identify UA stones from non-UA stones. An international, multicenter study was performed on 202 patients who received percutaneous nephrolithotomy for kidney stones with HU < 800. Data from 156 (77.2%) patients were used for model development, while data from 46 (22.8%) patients from a multinational institution were used for external validation. A total of 21,074 kidney and stone contour-annotated computed tomography images were trained with the ResNet-18 Mask R-convolutional neural network algorithm. Finally, this model was concatenated with demographic and clinical data as a fully-connected layer for stone classification. Our model was 100% sensitive in detecting kidney stones in each patient, and the delineation of kidney and stone contours was precise within clinically acceptable ranges. The development model provided an accuracy of 99.9%, with 100.0% sensitivity and 98.9% specificity, in distinguishing UA from non-UA stones. On external validation, the model performed with an accuracy of 97.1%, with 89.4% sensitivity and 98.6% specificity. SHAP plots revealed stone density, diabetes mellitus, and urinary pH as the most important features for classification. Our ML-based model accurately identified and delineated kidney stones and classified UA stones from non-UA stones with the highest predictive accuracy reported to date. Our model can be reliably used to select candidates for an earlier-directed alkalization therapy.

1. Introduction

The selection of a treatment modality for patients with kidney stones is based on stone size, location, density, instrument availability, and patient comorbidities. This choice affects postoperative outcomes and perioperative morbidity. Treatment modalities recommended by contemporary guidelines include ureteroscopic lithotripsy, percutaneous nephrolithotripsy (PCNL), extracorporeal shockwave lithotripsy, and alkalization therapy [1–3]. For selected stones, surgical treatment options may be equally effective regarding the outcome. Therefore, a specific treatment modality can be recommended based on the physician's expertise and patient preference [1]. On the other hand, surgical intervention may compromise perioperative outcomes and increase morbidity in patients with significant comorbidities. In some patients, non-surgical modalities that are less invasive may be beneficial.

Uric acid (UA) stones comprise 5–20% of all urolithiasis and are primarily explained by a low urinary pH, with a minority of patients exhibiting high urinary excretion of UA [4, 5]. Medical dissolution therapy with urinary alkalization is the cornerstone of the medical management of UA stones, with a reported success rate of 80% and cost savings compared to surgical intervention [5]. The prediction of UA stones is not entirely accurate and is based on Hounsfield units (HU) < 450 on non-contrast-enhanced computed tomography (NCCT), radiolucency on plain radiography, and low urinary pH < 6 [6]. However, UA stones outside these ranges are often underdiagnosed due to a lack of reliable diagnostic systems predicting UA components and concerns about stones mistakenly thought to be UA [7]. As a result, alkalization therapy

is reported to be underused despite its potential benefits [6]. Correct differential diagnosis of UA stones from other stone components has important clinical implications since patients with a high risk of perioperative morbidity may be spared surgical interventions.

Several non-invasive techniques have been developed to classify UA from non-UA stones, including 24-hour urinalysis, nomograms using clinical and HU parameters, and dual-energy CT (DECT) [8–10]. However, these classification systems were limited by their model developments being based on a specific stone component, the necessity of access to particular CT scanner types, and, most importantly, the lack of external validation, which precludes their general applicability to clinical use. A growing body of evidence indicates that machine learning (ML) models may improve the accuracy of disease diagnosis and treatment outcomes compared to conventional discriminant analyses [11]. Given the multifactorial nature of the formation of specific stone components, we sought to develop an ML-based model that incorporates a comprehensive set of demographic, clinical, and CT data to better classify UA stones.

The aims of our study were (1) to train and develop an ML-based urinary stone recognition algorithm that can automatically identify the stone location and delineate its contour relative to the kidney and (2) to utilize this algorithm to develop a prediction model that incorporates demographic, clinical, and NCCT data to classify UA stones from other stone components.

2. Materials and Methods

2.1. Study cohort

An international, multicenter, cross-sectional study was conducted on 202 patients who underwent PCNL for kidney stones with HU < 800 between March 2005 and November 2018. The development cohort consisted of 156 (77.2%) patients from the Stone Centre at Vancouver General Hospital in Canada, while the external validation cohort comprised 46 (22.8%) patients from Gangnam Severance Hospital in Seoul, South Korea, a multinational institution with patients of a distinct ethnic background. Patients with incomplete data, including unknown stone composition, were excluded from the analysis.

This study was approved by the institutional ethics committees of the University of British Columbia's Clinical Research Ethics Board (H14-00475) and Gangnam Severance Hospital (2019-0838-001) after reviewing the protocols employed. All study procedures complied with the principles of the 1946 Declaration of Helsinki and its 2008 update.

2.2. Data acquisition

The demographic and clinical data that potentially affect stone components and NCCT images were retrospectively collected from the patients' medical records. Twenty-six preoperative demographic and clinical data included: patient age, body mass index, gender, American Society of Anesthesiologists score, stone history, number, bilaterality, multiplicity, HU, urinary pH and nitrite positivity, urine culture, serum levels of sodium, potassium, calcium, glomerular filtration rate, UA, phosphate, and the presence of

comorbidities including cerebrovascular disease, chronic kidney disease, chronic obstructive pulmonary disease, diabetes mellitus, hypertension, gout, and neurological disease prior to PCNL (Table 1). Stone component data for both the development and external validation cohorts were based on analyses performed using Fourier-transform infrared spectrometry, carried out at Lifelabs, Burnaby, BC, Canada, and Green Cross Laboratories, Yongin, South Korea, respectively.

Table 1
Demographic and clinical characteristics of the development and external validation cohorts.

	Development cohort	External validation cohort	<i>P</i>
No.	156 (77.2%)	46 (22.8%)	<i>NS</i>
Age (year)	59.0 (47.0–68.3)	54.0 (37.0–64.5)	<i>0.030</i>
Body mass index (kg/m ²)	29.1 (24.3–33.3)	24.8 (21.8–27.6)	<i>< 0.001</i>
Gender			<i>0.467</i>
Male	101 (64.7%)	26 (56.5%)	
Female	55 (35.3%)	20 (43.5%)	
Hounsfield units	499.4 (413.6–629.3)	545.5 (487.3–699.5)	<i>< 0.001</i>
Stone, number	3.0 (2.0–4.5)	4.0 (3.0–6.5)	<i>0.065</i>
Stone component			
UA	41 (26.3%)	13 (28.3%)	<i>0.143</i>
Struvite	56 (35.9%)	14 (30.4%)	<i>0.508</i>
Calcium oxalate	54 (34.6%)	17 (37.0%)	<i>0.236</i>
Cystine	5 (3.2%)	2 (4.3%)	<i>0.987</i>
Urinary pH	5.5 (5.0–6.5)	6.5 (5.5–7.3)	<i>0.003</i>
Nitrite positivity	32 (20.5%)	9 (19.6%)	<i>0.566</i>
Urine culture positivity	31 (19.9%)	8 (17.4%)	<i>0.648</i>
Serum			
Sodium	140.0 (139.0–142.0)	140.0 (140.0–141.0)	<i>0.587</i>
Potassium	4.1 (3.8–4.4)	4.20 (4.10–4.38)	<i>0.864</i>
Calcium	2.33 (2.26–2.42)	2.31 (2.24–2.47)	<i>0.873</i>
Glomerular filtration rate	69.0 (50.0–94.0)	90.0 (74.0–121.5)	<i>< 0.001</i>
UA	339.5 (273.5–431.5)	298.0 (206.5–350.5)	<i>0.018</i>

Values are presented as number (%) or median (interquartile range).

ASA = American Society of Anesthesiologists; ; COPD = chronic obstructive pulmonary disease; UA = uric acid

	Development cohort	External validation cohort	<i>P</i>
Phosphate	1.10 (0.92–1.29)	1.06 (0.96–1.28)	0.928
Bilaterality	42 (26.9%)	16 (34.8%)	0.040
Multiplicity	89 (57.1%)	37 (80.4%)	0.009
ASA score			0.062
0	38 (24.4%)	9 (19.6%)	
≥ 1	118 (75.6%)	37 (80.4%)	
Stone history	40 (25.6%)	10 (21.7%)	0.098
Comorbidity			
Cerebrovascular disease	18 (11.5%)	5 (10.9%)	0.438
Chronic kidney disease	12 (7.7%)	3 (6.5%)	0.768
Chronic obstructive pulmonary disease	10 (6.4%)	2 (4.3%)	0.133
Diabetes mellitus	45 (28.8%)	10 (21.7%)	0.389
Hypertension	54 (34.6%)	20 (43.5%)	0.164
Gout	3 (1.9%)	0 (0.0%)	0.418
Neurological disease	20 (12.8%)	5 (10.9%)	0.129
Values are presented as number (%) or median (interquartile range).			
ASA = American Society of Anesthesiologists; ; COPD = chronic obstructive pulmonary disease; UA = uric acid			

Standard axial NCCT images that included kidney and stone information were acquired in DICOM format. After acquisition, the images were extracted using a Python software application and saved in PNG files. A total of 14,843 and 6,231 kidney and stone images were selected to form the datasets for developing the ML model. The anatomical contours of the kidneys and stones were semi-automatically annotated using the open-source Computer Vision Annotation Tool (Intel®, CA, USA). Overall, 21,074 kidney and stone annotated images were included in the datasets for model development.

2.3. Model development

Figure 1 depicts the overall architecture of our model. Initially, the ResNet18 model framework was trained with the PyTorch feedforward deep learning library using predefined methods [12]. We used stochastic gradient descent as the optimizer with a learning rate set at 0.005 and a batch size of 32. The ResNet-18 Mask R-convolutional neural network model was chosen for this study due to its capability

and robustness in general-purpose image segmentation under limited data [13]. The kidney and stone contour-annotated image data from the development cohort was used for model training, validation, and testing in an 8:1:1 ratio, with all images in the training set being different from those in the testing set. Finally, demographic and clinical data were concatenated with the ResNet18 model in a fully-connected layer to develop the final model for stone component classification. The model was then interpreted using the SHAP algorithm to enable visual interpretation of the quantitative association between the input variables and the model's output [14].

2.4. Evaluation metrics

Prediction accuracies were analyzed according to binary (UA vs. non-UA stones) and multiclass (UA vs. calcium oxalate, struvite, and cystine stones) classifications. Prediction accuracies were measured by the precision of instance prediction, which counted the number of true positive, true negative, false positive, and false negative instances as compared with the operator's semi-automatic annotation. Prediction accuracies were compared to those of the multivariate logistic regression analyses.

2.5. Statistical analysis

Demographic and clinical characteristics between the development and external validation cohorts were compared using the two-sided Mann-Whitney U-test for the analysis of continuous variables and the chi-square test for the analysis of categorical variables. Logistic regression analysis for binary classification was performed using the same training set. Independent predictive indicators associated with UA stones in the multivariate analyses were entered into the logistic regression model. All tests were two-tailed, with statistical significance set at a $p < 0.05$. Statistical analysis was performed using IBM SPSS Statistics software ver. 21.0 (IBM Corporation, Armonk, NY) and R Statistical Package ver. 3.1.3. (Institute for Statistics and Mathematics, Vienna, Austria).

3. Results

3.1. Demographic and clinical features

The demographic and clinical characteristics of patients in the development and external validation cohorts are presented in Table 1. According to stone components, UA, calcium oxalate, struvite, and cystine stones comprised 26.3%, 35.9%, 34.6%, and 3.2% in the development cohort, while 28.3%, 30.4%, 37.0%, and 4.3% in the external validation cohort, respectively. Overall, there were no significant differences between the development and external validation cohorts regarding demographic, clinical, and stone features.

3.2. Kidney and stone identification and contour delineation

Segmentation identified and delineated two anatomical elements, kidney and stone contours. Our model was 100% sensitive in detecting kidney stones in each patient. The delineation of kidney and stone contours was precise within clinically acceptable ranges, as shown in a selected sample patient (Fig. 2).

3.3. Predictive performance

The predictive accuracies of our model varied according to the stone component and classification system. Overall, the performances of the ML-based model outperformed those of the logistic regression model (Table 2).

Table 2
Predictive accuracies of the development and external validation cohorts.

Classification	Cohort	Stone type	Accuracy (%)	Sensitivity (%)	Specificity (%)
LR					
Binary	Development	UA vs. non-UA	82.4	72.4	69.6
	External validation	UA vs. non-UA	86.9	82.6	73.3
ML					
Binary	Development	UA vs. non-UA	99.9	100.0	99.9
	External validation	UA vs. non-UA	97.1	89.4	98.6
Multiclass	Development	UA	98.2	95.7	89.6
		Calcium oxalate	88.4	70.7	98.4
		Struvite	98.7	97.7	99.2
		Cystine	95.5	77.1	96
	External validation	UA	91.3	77.1	89.6
		Calcium oxalate	74.3	95.8	55.8
		Struvite	89.4	72.5	97.2
		Cystine	93.5	0.0	93.5

UA = uric acid

3.3.1. Binary classification

The development model discriminated UA and non-UA stones with an accuracy of 99.9%, with 100.0% sensitivity and 99.9% specificity (Table 2). We identified features most predictive for binary stone classification by quantifying the predictor importance of each variable. Stone density, diabetes mellitus, and urinary pH showed to be the top three contributing features for classifying UA stones. On external

validation, the model performed with a predictive accuracy of 97.1%, with 89.4% sensitivity and 98.6% specificity. The ML-based model exhibited higher performance than the multivariate logistic regression model in both development and external validation cohorts.

3.3.2. Multiclass classification

The development model discriminated UA, calcium oxalate, struvite, and cystine stones with predictive performances of 98.2%, 88.4%, 98.7%, and 95.5%, respectively (Table 2). The features most predictive for multiclass stone classification were stone density, diabetes mellitus, and urinary pH. On external validation, the model's prediction accuracies for UA, calcium oxalate, struvite, and cystine stones were 91.3%, 74.3%, 89.4%, and 93.5%, respectively. The ML-based model's performance for multiclass classification was relatively lower than that of the binary classification; however, remained within clinically acceptable ranges.

4. Discussion

The inspiration for this study arises from the unmet clinical need to accurately predict UA stones prior to selecting the optimal treatment modality. Over the last decade, technological advancements in ureteroscopes and laser lithotriptors have paved the path to an upsurge in surgical intervention regardless of stone composition [15, 16]. Although alkalization therapy for UA stones is ideal for patients with high morbidity or recurrent UA stone formers, it is commonly underused due to the lack of reliable factors predicting its outcome, concerns about the existence of heterogeneous stone composition, and patient intolerance [7–15]. Most of all, the lack of standardized protocols for predicting UA components adds complexity to making treatment decisions in real-life situations [15]. The present study is the first to develop and validate an effective predictive model incorporating NCCT images into traditional demographic and clinical data to classify UA from non-UA stones in patients with stones in the 'grey zone' HUs. External validation showed that our objective, expeditious, and non-invasive model could identify UA stones with an accuracy of 97.1%, the highest predictive performance reported to date.

Our model has several implications for improving the current standard of care through its implementation in clinical practice. First, the input variables, including demographic, clinical, and NCCT data, are those readily available in real-world practice, which supports the general applicability of our model. Previously reported stone component classification models utilizing imaging data generally require time-consuming manual analysis of HU parameters or additional examination using specific CT scanner types, such as DECT, which may not be available across all practice settings [8, 17–19]. In contrast, our automated model has the potential to be integrated into any electronic medical records system that utilizes coding algorithms to be utilized as a decision support system. Such a system may reduce the time required for classification and avoid additional radiation exposure and costs.

Second, we selected patients with stones of relatively low HUs for the model development since these stones pose a diagnostic dilemma in clinical decision-making for alkalization therapy [5]. We selected stones with HUs < 800 in order to include struvite and cystine stones, in addition to UA stones, that are

characterized as having a completely distinct management approach. The multiclass classification model provided a relatively lower performance than the binary classification. However, the overall performance was excellent and surpassed that of the conventional multivariate logistic regression model, providing a reliable diagnostic standard for treatment decision-making. Lastly, the architecture of our model and its working principle allow future refinements. Our model can additionally integrate intraoperative laser lithotripsy data and has the potential to provide patient-specific optimal laser settings for maximal fragmentation efficiency according to each stone feature.

Several strengths of our study are worth mentioning. First, external validation of prediction models is essential before their use in clinical practice. Since validation samples should be obtained from different but plausibly relevant cohorts, the performance of our model was validated with an external cohort comprised of patients from an international institution with distinct ethnic backgrounds. Discrimination performance is usually observed to be inferior in the external validation cohort compared to the development cohort [20]. Nevertheless, the performance of our external validation cohort was non-inferior compared to that of the development cohort, indicating the validity and feasibility of our model. Second, we incorporated a comprehensive set of demographic, clinical, and NCCT imaging data that are potentially associated with stone components for the model development. Moreover, the dataset was considered of high quality, with all input variables of the development and external validation cohorts being manually reviewed and incorporated without any missing data, which may have contributed to its high predictive performance.

This study is not without limitations. First, mixed component stones were excluded from the development and external validation cohorts. Although the distinction between pure UA stones and mixed component stones is crucial in the decision-making of alkalization therapy, only stones with pure components were included. Since the extent of the UA component beyond which the stone has to be defined as mixed is unclear, subsequent studies incorporating quantitative analysis of mixed stones, will need to be performed to screen optimal patients who would be amenable to medical therapy. Second, a population-based database with a larger number of subjects may provide better generalizability. Albeit, we utilized institutional data, which provided a comprehensive and high-quality dataset, to maximize predictive performance. Lastly, performances declined for the multiclass classification, indicating uncertainty of clinical usefulness, especially in classifying cystine stones. The likely explanation is the limited number of cystine stones in both the development and external validation cohorts. Notwithstanding these limitations, the advantages of our model over previously reported tools classifying predicting stone components indicate its feasibility and general applicability to be implemented into real-world clinical practice.

5. Conclusions

We developed and externally validated an ML-based model to identify and delineate kidney stones and classify UA stones from other stone components. With the highest predictive performance reported to date, our model can be reliably used to select candidates for an earlier-directed alkalization therapy in

patients with kidney stones within the 'grey zone' HUs. Further modification in the ML algorithm incorporating cases with mixed component stones would be warranted for more sophisticated predictions.

Declarations

6. Acknowledgments

This study was supported by a faculty research grant of Yonsei University College of Medicine (6-2022-0173).

7. Conflict of interest

No competing financial interests exist for all authors, including Ben H. Chew, Victor KF. Wong, Abdulghafour Halawani, Sujin Lee, Sangyeop Baek, Hoyong Kang, and Kyo Chul Koo.

8. Author contributions

Conceptualization: Kyo Chul Koo and Ben H. Chew; Methodology: Victor KF. Wong and Abdulghafour Halawani; Formal analysis and investigation: Sujin Lee, Sangyeop Baek, and Hoyong Kang; Writing - original draft preparation: Kyo Chul Koo; Writing - review and editing: Victor KF. Wong, Abdulghafour Halawani, Sujin Lee, Sangyeop Baek, and Hoyong Kang; Funding acquisition: Kyo Chul Koo; Resources: Ben H. Chew; Supervision: Kyo Chul Koo and Ben H. Chew

9. Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

10. Data availability

The data sets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. Geraghty RM, Davis NF, Tzelves L, et al. (2022) Best Practice in Interventional Management of Urolithiasis: An Update from the European Association of Urology Guidelines Panel for Urolithiasis 2022. *Eur Urol Focus*; doi: 10.1016/j.euf.2022.06.014.
2. Quhal F, Seitz C (2021) Guideline of the guidelines: urolithiasis. *Curr Opin Urol* 31:125-129; doi: 10.1097/MOU.0000000000000855.

3. Turk C, Petrik A, Sarica K, et al. (2016) EAU Guidelines on Interventional Treatment for Urolithiasis. *Eur Urol* 69:475-482; doi: 10.1016/j.eururo.2015.07.041.
4. Trinchieri A, Montanari E (2018) Biochemical and dietary factors of uric acid stone formation. *Urolithiasis* 46:167-172; doi: 10.1007/s00240-017-0965-2.
5. Shekarriz B, Stoller ML (2002) Uric acid nephrolithiasis: current concepts and controversies. *J Urol* 168(4 Pt 1):1307-1314; doi: 10.1097/00005392-200210010-00003.
6. Tsaturyan A, Bokova E, Bosshard P, et al. (2020) Oral chemolysis is an effective, non-invasive therapy for urinary stones suspected of uric acid content. *Urolithiasis* 48:501-507; doi: 10.1007/s00240-020-01204-8.
7. Moore J, Nevo A, Salih S, et al. (2022) Outcomes and rates of dissolution therapy for uric acid stones. *J Nephrol* 35:665-669; doi: 10.1007/s40620-021-01094-y.
8. McGrath TA, Frank RA, Schieda N, et al. (2020) Diagnostic accuracy of dual-energy computed tomography (DECT) to differentiate uric acid from non-uric acid calculi: systematic review and meta-analysis. *Eur Radiol* 30:2791-2801; doi: 10.1007/s00330-019-06559-0.
9. Moreira DM, Friedlander JI, Hartman C, et al. (2013) Using 24-hour urinalysis to predict stone type. *J Urol* 190:2106-2011; doi: 10.1016/j.juro.2013.05.115.
10. Wiessmeyer JR, Ozimek T, Struck JP, et al. (2022) Comprehensive Nomogram for Prediction of the Uric Acid Composition of Ureteral Stones as a Part of Tailored Stone Therapy. *Eur Urol Focus* 8:291-296; doi: 10.1016/j.euf.2021.02.001.
11. Lim B, Lee KS, Lee YH, et al. (2021) External Validation of the Long Short-Term Memory Artificial Neural Network-Based SCaP Survival Calculator for Prediction of Prostate Cancer Survival. *Cancer Res Treat* 53:558-566; doi: 10.4143/crt.2020.637.
12. He K, Zhang X, Ren S, et al. (2015) Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 770-778.
13. Shoaib MA, Lai KW, Chuah JH, et al. (2022) Comparative studies of deep learning segmentation models for left ventricle segmentation. *Front Public Health* 10:981019; doi: 10.3389/fpubh.2022.981019.
14. Lundberg S, Lee S (2017) A unified approach to interpreting model predictions. Presented at: 31st International Conference on Neural Information Processing Systems; December 4-9, 2017; Long Beach, California, USA.
15. Abou-Elela A (2017) Epidemiology, pathophysiology, and management of uric acid urolithiasis: A narrative review. *J Adv Res* 8:513-527; doi: 10.1016/j.jare.2017.04.005.
16. Breda A, Territo A, Lopez-Martinez JM (2016) Benefits and risks of ureteral access sheaths for retrograde renal access. *Curr Opin Urol* 26:70-75; doi: 10.1097/MOU.0000000000000233.
17. Qin L, Zhou J, Hu W, et al. (2022) The combination of mean and maximum Hounsfield Unit allows more accurate prediction of uric acid stones. *Urolithiasis* 50:589-597; doi: 10.1007/s00240-022-01333-2.

18. Taily T, Larish Y, Nadeau B, et al. (2016) Combining Mean and Standard Deviation of Hounsfield Unit Measurements from Preoperative CT Allows More Accurate Prediction of Urinary Stone Composition Than Mean Hounsfield Units Alone. *J Endourol* 30:453-459; doi: 10.1089/end.2015.0209.
19. Nakhostin D, Sartoretti T, Eberhard M, et al. (2021) Low-dose dual-energy CT for stone characterization: a systematic comparison of two generations of split-filter single-source and dual-source dual-energy CT. *Abdom Radiol (NY)* 46:2079-2089; doi: 10.1007/s00261-020-02852-5.
20. Steyerberg EW, Vickers AJ, Cook NR, et al. (2010) Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 21:128-138; doi: 10.1097/EDE.0b013e3181c30fb2.

Figures

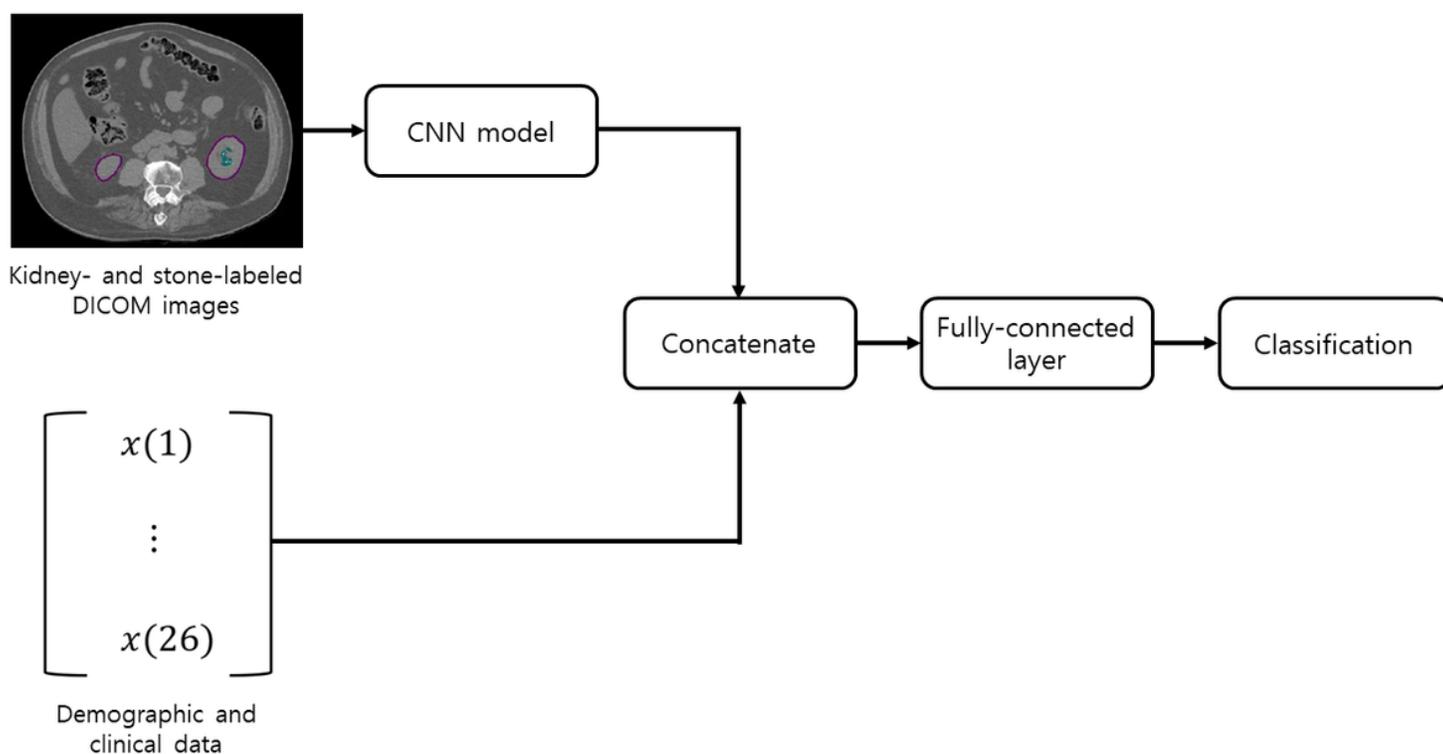


Figure 1

Overall architecture of our model: First, the ResNet18 Mask R-convolutional neural network model framework was trained using predefined methods with the PyTorch feedforward deep learning library. Next, demographic and clinical data were concatenated with the ResNet18 model in a fully-connected layer to develop the final model for stone component classification.

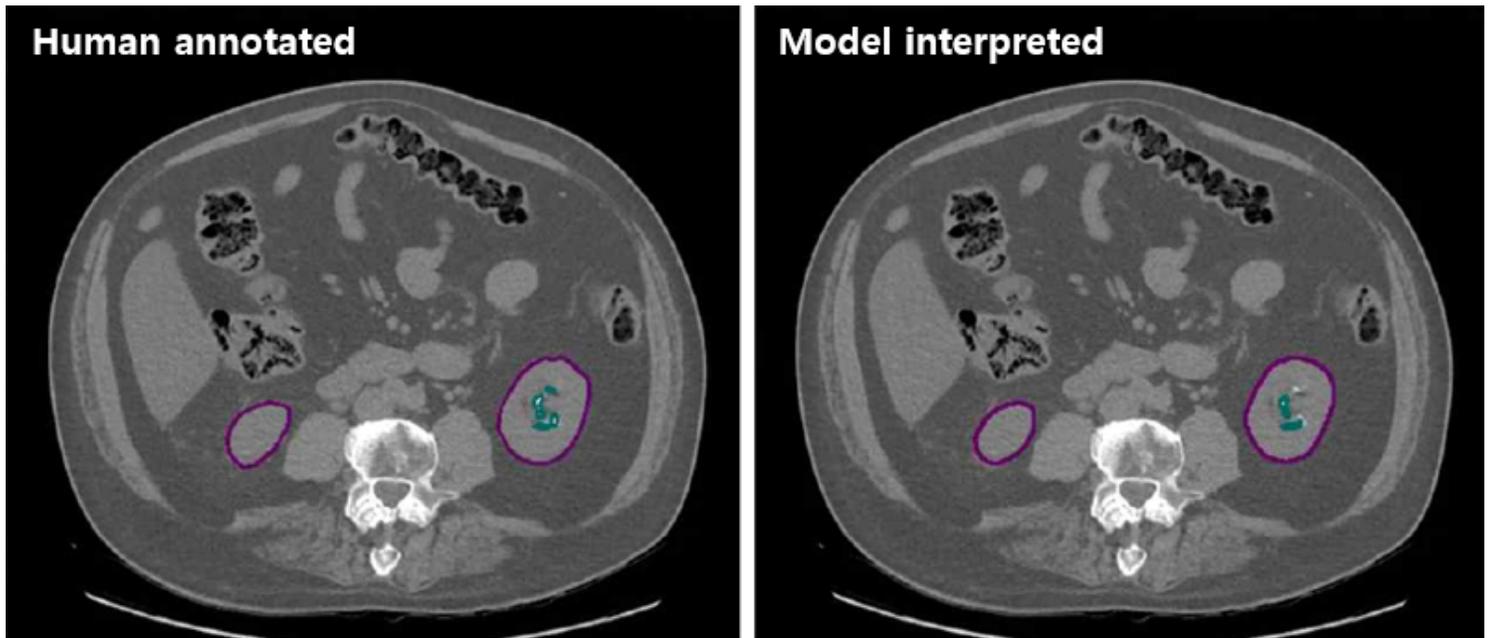


Figure 2

A patient-specific human-annotated and model-interpreted kidney and stone contour segment of a chosen sample patient (subject no. 78).