Original Article - Endourology/Urolithiasis

Investig Clin Urol 2022;63:441-447.
https://doi.org/10.4111/icu.20220062
pISSN 2466-0493 • eISSN 2466-054X

INVESTIGATIVE AND CLINICAL UROLOGY
ICUROLOGY

Prediction of the composition of urinary stones using deep learning

Ui Seok Kim[1], Hyo Sang Kwon[1], Wonjong Yang[1], Wonchul Lee[1], Changil Choi[1], Jong Keun Kim[1], Seong Ho Lee[1], Dohyoung Rim[2], Jun Hyun Han[1]

[1]Department of Urology, Hallym University Dongtan Sacred Heart Hospital, Hwaseong, [2]Department of Cognitive Science, Yonsei University, Seoul, Korea

Purpose: This study aimed to predict the composition of urolithiasis using deep learning from urinary stone images.

Materials and Methods: We classified 1,332 stones into 31 classes according to the stone composition. The top 4 classes with a frequency of 110 or more (class 1: calcium oxalate monohydrate [COM] 100%, class 2: COM 80%+struvite 20%, class 3: COM 60%+calcium oxalate dihydrate [COD] 40%, class 4: uric acid 100%) were selected. With the 965 stone images of the top 4 classes, we used the seven convolutional neural networks (CNN) to classify urinary stones and compared their classification performances.

Results: Among the seven models, Xception_lr0.001 showed the highest accuracy, precision, and recall and was selected as the CNN model to predict the stone composition. The sensitivity and specificity for the 4 classes by Xception_lr0.001 were as follows: class 1 (94.24%, 91.73%), class 2 (85.42%, 96.14%), class 3 (86.86%, 99.59%), and class 4 (94.96%, 98.82%). The sensitivity and specificity of the individual components of the stones were as follows. COM (98.82%, 94.96%), COD (86.86%, 99.64%), struvite (85.42%, 95.59%), and uric acid (94.96%, 98.82%). The area under the curves for class 1, 2, 3, and 4 were 0.98, 0.97, 1.00, and 1.00, respectively.

Conclusions: This study showed the feasibility of deep learning for the diagnostic ability to assess urinary stone composition from images. It can be an alternative tool for conventional stone analysis and provide decision support to urologists, improving the effectiveness of diagnosis and treatment.

Keywords: Artificial intelligence; Deep learning; Endoscopy; Machine learning; Urolithiasis

INTRODUCTION

Artificial intelligence (AI) has made significant strides in interpreting perceptual information, allowing machines to analyze complex data better. Among the various technologies of AI, the most popular is deep learning (DL), part of a broader family of machine learning methods based on a neural network structure inspired by the human brain. In recent years, the application of such DL technology to medical and biomedical research fields has increased exponentially [1].

When a specific problem is solved using DL, the developed algorithm would have derived the result. Algorithms can build models based on training data, apply experiences, and make predictions and decisions on new, unknown data obtained through training. To understand A more clearly

Received: 13 February, 2022 • **Revised:** 31 March, 2022 • **Accepted:** 17 April, 2022 • **Published online:** 25 May, 2022
Corresponding Author: Jun Hyun Han https://orcid.org/0000-0002-8452-1916
Department of Urology, Hallym University Dongtan Sacred Heart Hospital, 7 Keunjaebong-gil, Hwaseong 18450, Korea
TEL: +82-31-8086-2730, FAX: +82-31-8086-2728, E-mail: junuro@naver.com

and use it, it is essential to select a topic suitable for DL and collect data, rather than simply input data and adjust variables to produce results. After that, it is necessary to go through several stages of verifying the already created algorithm and applying it to actual clinical practice.

Urolithiasis is one of the most common diseases in Korea and Western countries, and it can be a serious medical problem for the working-age population [2-4]. With the development of optical technology and endoscopic instruments, transurethral endoscopic surgery through a natural orifice in the field of urology has been increasing rapidly in recent years [5,6]. Currently, it is possible to acquire high-quality images in real-time for urolithiasis or lesions during most urological endoscopic surgeries. Predicting the stone's composition and strength using real-time images during surgery has several advantages in treating urolithiasis. It can enable the surgeon to select a more appropriate tool during surgery and help make decisions about treatment, such as administering antibiotics before and after surgery.

If elements such as struvite exist, the surgeon will think more seriously about the countermeasures for postoperative infection and implement appropriate medical treatment accordingly. In addition, information on recurrent stones and metabolic stones can be provided before the conventional stone component test results are released.

In recent years, computer vision and DL have been used to detect many different diseases and lesions in the body automatically [7]. Object detection is a method that is used to recognize and detect other objects present in an image or video and label them to classify these objects, and the technology has been significantly advanced. In addition, DL-based medical image analysis can be applied to computer-aided diagnostics to provide decision support to clinicians and improve the accuracy and efficiency of various diagnostic and treatment processes [7].

Most urinary stones are composed of different chemical compositions in varying proportions. Therefore, pure single composition-containing stones only account for a small proportion of urolithiasis. Black et al. [8] reported that deep convolutional neural networks (CNN) on 63 human kidney stones of various components in a previous pilot study could be usefully used to predict kidney stone composition with good calls in digital photographs. This study aimed to investigate the prediction of the components of urolithiasis by DL algorithm using images of urinary stones obtained by endoscopic surgery.

MATERIALS AND METHODS

1. Datasets

The protocol of this study was conducted in accordance with good clinical practice guidelines and the Declaration of Helsinki and was approved by the Institutional Review Board Committee for Human Subjects at Hallym University Dongtan Sacred Heart Hospital (IRB no. HDT 2022-01-004).

From January 2018 to March 2021, a total of 1,332 stones were obtained through endoscopic surgery. The components of each stone fragment were verified with the conventional stone analysis, and images of the stones were captured on a digital camera.

For conventional stone analysis, Fourier-transform infrared spectroscopy (FT-IR, Green cross labs, Yongin, Korea) method was used. This is a method using the physical properties of molecules forming stones to absorb infrared rays in a specific wavelength range that coincides with their intrinsic oscillation period. Objective and quantitative results can be obtained because the most suitable component and its composition ratio are found after comparing with the standard spectrum for the intrinsic spectrum of a substance. In addition, the sample processing process is quick and simple, and it is known as the gold standard method for stone analysis so far because it can accurately distinguish various components even with a very small amount of less than 100 g [9].

These 1,332 stones were analyzed for components through stone analysis, and as a result of classification according to the proportion of stone components, a total of 31 classes were classified (Table 1). The top 4 classes with more than 110 frequencies out of 31 categories, a total of 965 stones were used in the DL algorithm. The composition of the top four classes were as follows. Class 1 (n=469): calcium oxalate monohydrate (COM) 100%, Class 2 (n=240): COM 80% and struvite 20%, Class 3 (n=137): COM 60% and calcium oxalate dihydrate (COD) 40%, Class 4 (n=119): uric acid 100% (Table 1).

2. Preprocessing of image and architecture of the deep convolutional neural network

We used the 7 CNN models, DenseNet201 [10], ResNet152 [11], ResNet152_FC3 [11], Xception [12], Xception dropout0.8 [12], Xception_Ir0.001 [12] and Xception_Ir0.001_FC3 [12] to classify the urinary stone images and compared their classification performances.

The original image data was cropped with 160×160 size and no more preprocessing as resize and contrast adjusting were not used. For the training data, the shortage of samples was solved by expanding the data set through data augmen-

https://doi.org/10.4111/icu.20220062

Table 1. The default distribution of stones used in the analysis

Class	COM (%)	COD (%)	CA (%)	ST (%)	UA (%)	AU (%)	CY (%)	BR (%)	Other (%)	N
1	+ (100)									469
2	+ (80)			+ (20)						240
3	+ (60)	+ (40)								137
4					+ (100)					119
5	+ (50)	+ (25)	+ (25)							88
6	+ (40)		+ (20)	+ (40)						57
7			+ (20)	+ (80)						44
8			+ (85)	+ (15)						34
9	+ (95)	+ (5)								30
10		+ (65)	+ (35)							24
11			+ (100)							13
12	+ (90)				+ (10)					9
13	+ (20)							+ (80)		8
14	+ (80)				+ (20)					8
15	+ (50)		+ (25)	+ (25)						5
16	+ (20)	+ (80)								5
17				+ (100)						5
18	+ (50)			+ (50)						5
19	+ (65)		+ (35)							5
20	+ (80)		+ (10)	+ (10)						5
21				+ (90)		+ (10)				3
22				+ (20)		+ (80)				3
23				+ (70)		+ (30)				3
24							+ (100)			3
25	+ (20)					+ (80)				2
26						+ (100)				2
27	+ (50)				+ (50)					2
28	+ (25)		+ (25)	+(50)						1
29	+ (50)		+ (20)	+ (30)						1
30				+ (80)		+ (20)				1
31									+ (100)	1
Sum (N)	1,077	284	277	407	138	14	2	8	1	1,332

COM, calcium oxalate monohydrate; COD, calcium oxalate dihydrate; CA, carbonate apatite; ST, struvite; UA, uric acid; AU, ammonium urate; CY, cysteine; Br, Brushite.

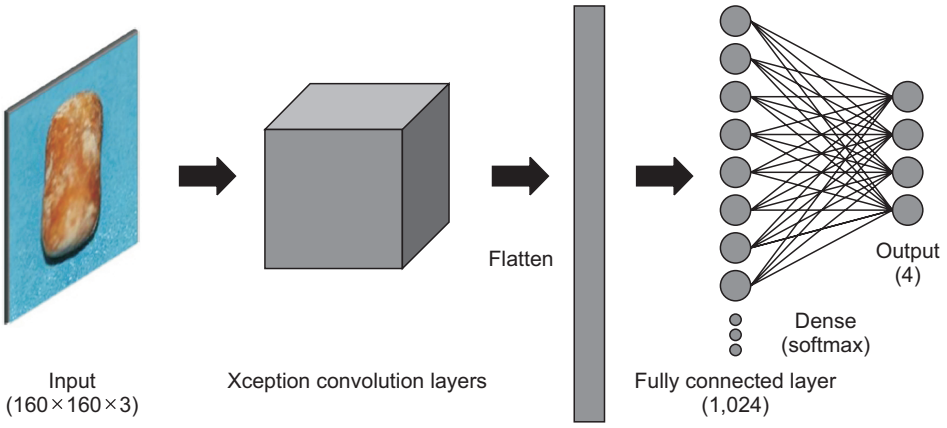

Fig. 1. Block representation showing the layer organization of Xception deep learning model.

tation. We used transfer learning based on the 7 CNN models. The overall structure is Input (160×160×3) - Convolution layer - Flatten - Fully connected layer - Dense (4, softmax) and 7 CNN models were tested by adjusting the Convolution layer and Fully connected layer. Fully connected layers of each CNN architecture were replaced with flattened layers, dense with 1,024 nodes and dropout of 0.5 drop rate. The last fully connected layer has nodes of category size followed by softmax (Fig. 1). Each hyperparameter for training is as follows; Adam optimizer with learning rate 0.001, cross-entropy loss function, 256 batch size, 1,000 epoch, early stopping after 150 patience for validation loss, and class weighting for unbalanced data distribution. 10-fold cross-validation was used, which means 90% of data was used for training, and the remaining 10% of data was used for tests. The model was tested with metrics of accuracy, precision, recall rate, F1 score, and area under the curve (AUC). In addition, the sensitivity and specificity of each class were analyzed to determine how accurately the same class of stone can be predicted, and the sensitivity and specificity of individual stone components were analyzed.

RESULTS

Among 1,332 stones, pure stones were 54% (7 types, 720 stones) and mixed stones were 46% (24 types, 612 stones) (Fig. 2). A total of 965 stones were used for the prediction of stone composition, including pure stones and mixed stones. Among the seven CNN models (DenseNet201, ResNet152, ResNet152_FC3, Xception, Xception_dropout0.8, Xception_lr0.001, Xception_lr0.001_FC3), Xception_lr0.001 showed the highest accuracy of 0.91 (0.03), precision of 0.92 (0.03), and recall of 0.90 (0.04), and was selected as the CNN model to predict the stone composition (Table 2, Fig. 3).

The sensitivity and specificity for the 4 classes by Xception_lr0.001 were as follows: class 1 (94.24%, 91.73%), class 2 (85.42%, 96.14%), class 3 (86.86%, 99.59%), and class 4 (94.96%, 98.82%) (Table 3). The sensitivity and specificity of the in-

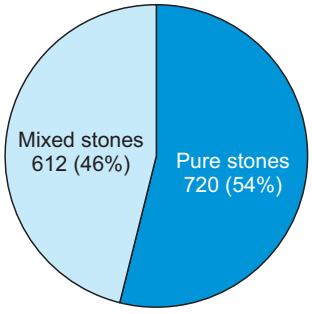

Fig. 2. Ratio of pure stones and mixed stones.

Table 2. Accuracy, precision, and recall rate of 7 convolutional neural networks models

Model	Accuracy	Precision	Recall
DenseNet201	0.82 (0.03)	0.84 (0.03)	0.81 (0.03)
ResNet 152	0.77 (0.03)	0.78 (0.03)	0.75 (0.03)
ResNet 152_FC3	0.70 (0.04)	0.64 (0.08)	0.66 (0.07)
Xception	0.89 (0.03)	0.90 (0.04)	0.88 (0.04)
Xception dropout0.8	0.89 (0.03)	0.90 (0.03)	0.88 (0.03)
Xception_lr0.001	0.91 (0.03)	0.92 (0.03)	0.90 (0.04)
Xception_lr0.001_FC3	0.87 (0.03)	0.87 (0.03)	0.87 (0.04)

Values are presented as average (error).

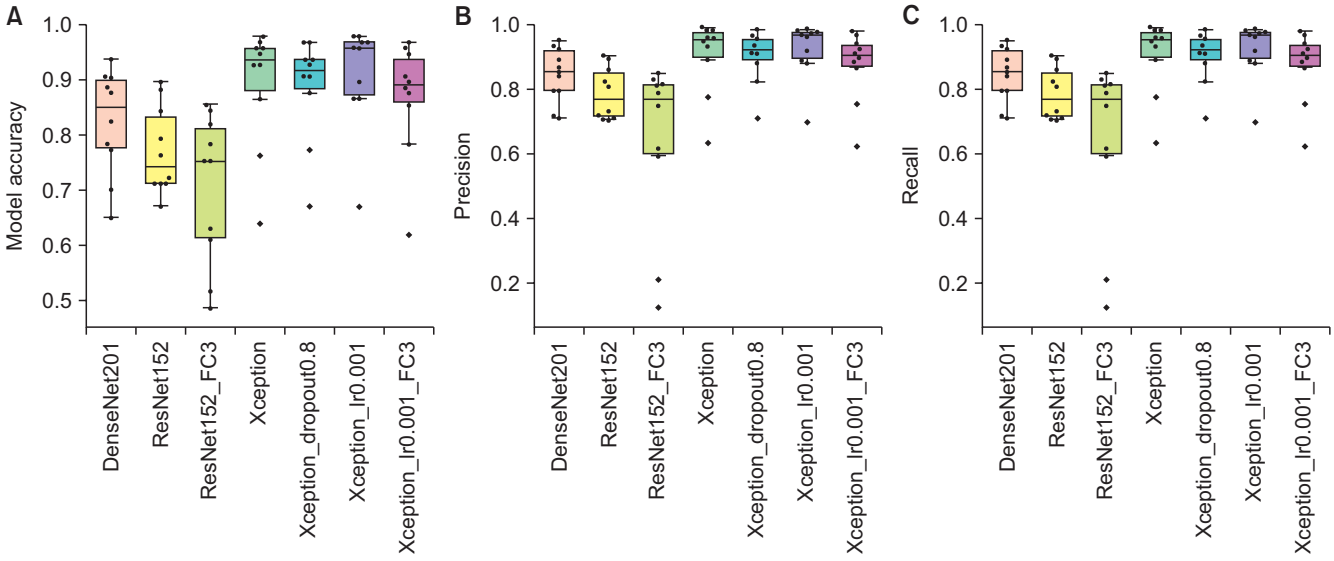

Fig. 3. Model accuracy (A), precision (B), and recall (C) rate of 7 convolutional neural networks models.

https://doi.org/10.4111/icu.20220062

Table 3. The sensitivity and specificity by class for Xception_lr0.001 (unit: %)

Class	Sensitivity	Specificity
1	94.24	91.73
2	85.42	96.14
3	86.86	99.59
4	94.96	98.82

Table 4. The sensitivity and specificity by stone composition for Xception_lr0.001 (unit: %)

Stone composition type	Sensitivity	Specificity
Calcium oxalate monohydrate	98.82	94.96
Calcium oxalate dihydrate	86.86	99.64
Struvite	85.42	95.59
Uric acid	94.96	98.82

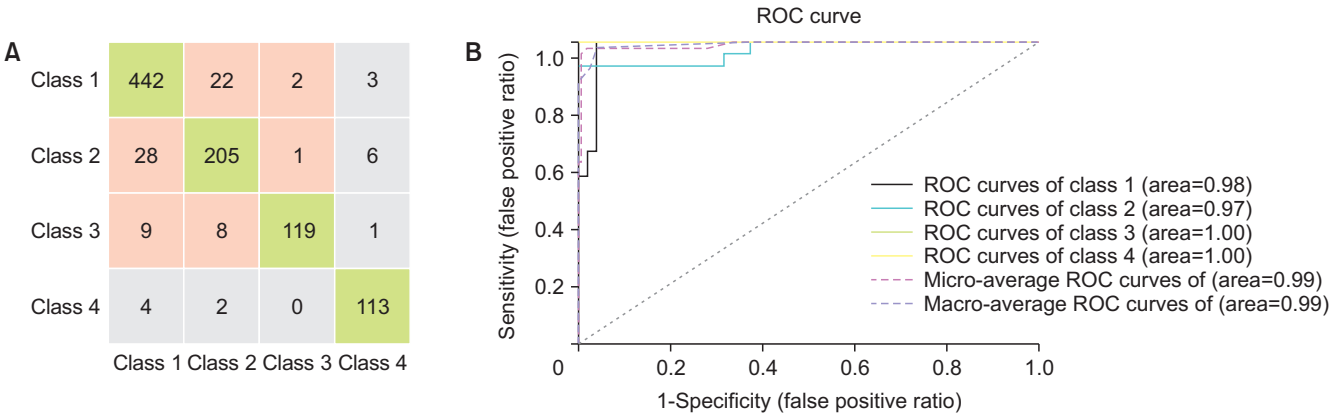

Fig. 4. Confusion matrix (A) and receiver operating characteristic (ROC) curves (B) for Xception_lr0.001.

dividual components of the stones were as follows. COM (98.82%, 94.96%), COD (86.86%, 99.64%), struvite (85.42%, 95.59%), and uric acid (94.96%, 98.82%) (Table 4). Thearea under the receiver operating characteristic curves for class 1, 2, 3, and 4 were 0.98, 0.97, 1.00, and 1.00, respectively (Fig. 4).

DISCUSSION

Machine learning is a subbranch of AI that is concerned with developing and deploying dynamic algorithms to analyze data and facilitate the identification of complex patterns [13]. DL, which trains artificial neural networks with multiple layers on large data sets, has been driving advances in AI in recent years. Many studies and applications for machine learning and DL are being conducted in urology fields such as urolithiasis, kidney cancer, bladder cancer, and prostate cancer [14]. For example, with the development of optical technology, endoscopic instruments, and lasers, transurethral endoscopic surgery for urolithiasis is rapidly increasing [5,6], and faster and more accurate decision-making is required for urologists regarding various treatment methods.

The conventional methods for stone component analysis must have a stone sample and usually require more than one day. A quick prediction of the urolithiasis component helps urologists consider the patient's condition and select an appropriate treatment option. However, the surgeon's prediction of the component by looking at the stone through the endoscopic screen is limited and unreliable. Sampogna et al. [15] reported the ability of surgeons to identify the composition of stones by observing endoscopic imaging. A total of 32 clinicians from 9 countries participated in the study. The overall accuracy was 39% (250 out of 640 predictions). COD stones were accurately detected in 69.8%, COM 41.8%, uric acid 33.3%, calcium oxalate/uric acid 34.3%, and cystine 78.1%. The precision of struvite (15.6%), calcium phosphate (0%), and calcium oxalate/calcium phosphate mixture (9.3%) were significantly lower. As a result, during endoscopic procedures, the surgeon was able to identify some of the COD and cystine stones but concluded that they were unreliable in identifying the majority of the stone constituents [15].

Most stones are composed of a mixture of different chemical compositions. Pure single composition-containing stones only account for a small proportion of urolithiasis. Currently, urinary stones can be classified based on several specific chemical components, including oxalate, phosphate, apatite, struvite, uric acid, and cystine. Mixtures of these chemical compositions in a single stone are also common, resulting in a spectrum of different stone chemistries [16]. Calcium-containing stones, including COM, COD, and calcium phosphate stones, account for around 70% to 80% of stones. Struvite (magnesium ammonium phosphate) stones account for 15% of urinary calculi and are typically associated with urease-producing urinary tract infections and carry significant morbidity [17]. Uric acid urolithiasis constitutes approxi-

mately 7% to 10% of all urinary stones [18]. Cystine stones account for 1% of stones. Other stones, such as xanthine and drug-induced calculi (e.g., triamterene, indinavir), account for less than 1% of stones. There are variations in the composition of stones, which are influenced by the differences in geographical, economic, or sanitation conditions.

In this study, the pure stones were more abundant than the mixed stones. Of the total stones, 35.2% were pure stones of COM component, and pure stones of the uric acid component were second with 8.9%. When classified by stone component, COM was included in 80.9% of the total stones, struvite 30.6%, COD 21.3%, carbonate apatite 20.8%, and uric acid 10.4%. Since this study included the stones obtained through transurethral endoscopic surgery, it cannot be said that the overall trend was reflected.

Chemical composition analysis of urolithiasis using DL is expected to provide real-time information to urologists to enable quick decision-making, helping various diagnostic and treatment processes. Black et al. [8] reported that in a pilot study, a DL computer vision algorithm was run using images of kidney stones of various components, and it was useful for recognizing the components of commonly encountered kidney stones. Estrade et al. [19] also reported that the prediction of urolithiasis components using various endoscopic urinary stone images and deep CNN algorithm was useful, and reported that the prediction of mixed stones, as well as pure stones, was good. In this study, we showed that it is able to predict the urinary stone composition using DL from digital photographs. Regardless of pure stones or mixed stones, the prediction of stone components using images of urolithiasis by a DL algorithm showed overall good recall. However, pure stones of COM and uric acid components, respectively, showed higher sensitivity than mixed stones (COM and COD, COM, and struvite), but there was no significant difference in specificity.

CONCLUSIONS

This study showed the feasibility of deep learning for the diagnostic ability to assess urinary stone composition from images. It can be an alternative tool for conventional stone analysis and provide decision support to urologists, improving the effectiveness of diagnosis and treatment.

CONFLICTS OF INTEREST

The authors have nothing to disclose.

FUNDING

None.

AUTHORS' CONTRIBUTIONS

Research conception and design: Jun Hyun Han. Data acquisition: Ui Seok Kim, Hyo Sang Kwon, and Wonjong Yang. Statistical analysis: Dohyoung Rim. Data analysis and interpretation: Ui Seok Kim, Wonchul Lee, Changil Choi, Jong Keun Kim, Dohyoung Rim, and Jun Hyun Han. Drafting of the manuscript: Ui Seok Kim and Jun Hyun Han. Critical revision of the manuscript: Ui Seok Kim and Jun Hyun Han. Administrative, technical, or material support: Wonchul Lee, Changil Choi, and Jong Keun Kim. Supervision: Seong Ho Lee. Approval of the final manuscript: all authors.

REFERENCES

1. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60-88.
2. Sorokin I, Mamoulakis C, Miyazawa K, Rodgers A, Talati J, Lotan Y. Epidemiology of stone disease across the world. World J Urol 2017;35:1301-20.
3. Tae BS, Balpukov U, Cho SY, Jeong CW. Eleven-year cumulative incidence and estimated lifetime prevalence of urolithiasis in Korea: a National Health Insurance Service-national sample cohort based study. J Korean Med Sci 2018;33:e13.
4. Jung JS, Han CH, Bae S. Study on the prevalence and incidence of urolithiasis in Korea over the last 10 years: an analysis of National Health Insurance Data. Investig Clin Urol 2018;59:383-91.
5. Kim JK, Cho YS, Park SY, Joo KJ, Min SK, Lee YG, et al. Recent surgical treatments for urinary stone disease in a Korean population: national population-based study. Int J Urol 2019;26:558-64.
6. Kim JK, Choi C, Kim US, Kwon H, Lee SH, Lee YG, et al. Recent trends in transurethral surgeries and urological outpatient procedures: a nationwide population-based cohort study. J Korean Med Sci 2020;35:e315.
7. Chan HP, Samala RK, Hadjiiski LM, Zhou C. Deep learning in medical image analysis. Adv Exp Med Biol 2020;1213:3-21.
8. Black KM, Law H, Aldoukhi A, Deng J, Ghani KR. Deep learning computer vision algorithm for detecting kidney stone composition. BJU Int 2020;125:920-4.
9. Primiano A, Persichilli S, Gambaro G, Ferraro PM, D'Addessi A, Cocci A, et al. FT-IR analysis of urinary stones: a helpful

https://doi.org/10.4111/icu.20220062

tool for clinician comparison with the chemical spot test. Dis Markers 2014;2014:176165.
10. Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. ArXiv. 1608.06993 [Preprint]. 2016 [cited 2022 Jan 7]. Available from: https://doi.org/10.48550/arXiv.1608.06993.
11. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. ArXiv. 1512.03385 [Preprint]. 2015 [cited 2022 Jan 7]. Available from: https://doi.org/10.48550/arXiv.1512.03385.
12. Chollet F. Xception: deep learning with depthwise separable convolutions. ArXiv. 1610.02357 [Preprint]. 2016 [cited 2022 Jan 7]. Available from: https://doi.org/10.48550/arXiv.1610.02357.
13. Goldenberg SL, Nir G, Salcudean SE. A new era: artificial intelligence and machine learning in prostate cancer. Nat Rev Urol 2019;16:391-403.
14. Suarez-Ibarrola R, Hein S, Reis G, Gratzke C, Miernik A. Current and future applications of machine and deep learning in urology: a review of the literature on urolithiasis, renal cell carcinoma, and bladder and prostate cancer. World J Urol 2020;38:2329-47.
15. Sampogna G, Basic D, Geavlete P, Galán Llopis JA, Reis Santos J, Saltirov I, et al. Endoscopic identification of urinary stone composition: a study of South Eastern Group for Urolithiasis Research (SEGUR 2). Actas Urol Esp (Engl Ed) 2021;45:154-9.
16. Ramaswamy K, Killilea DW, Kapahi P, Kahn AJ, Chi T, Stoller ML. The elementome of calcium-based urinary stones and its role in urolithiasis. Nat Rev Urol 2015;12:543-57.
17. Flannigan RK, Battison A, De S, Humphreys MR, Bader M, Lellig E, et al. Evaluating factors that dictate struvite stone composition: a multi-institutional clinical experience from the EDGE Research Consortium. Can Urol Assoc J 2018;12:131-6.
18. Liu CJ, Wu JS, Huang HS. Decreased associated risk of gout in diabetes patients with uric acid urolithiasis. J Clin Med 2019;8:1536.
19. Estrade V, Daudon M, Richard E, Bernhard JC, Bladou F, Robert G, et al. Towards automatic recognition of pure and mixed stones using intra-operative endoscopic digital images. BJU Int 2022;129:234-42.