Deep Learning Approaches for Kidney Stone Detection: A Systematic Review

Abstract:

The introduction of artificial intelligence (AI) in medical diagnostics has been gaining speed over the last few years, especially in the field of nephrology. Kidney stone disease (KSD) is a widespread and recurrent condition and needs proper and prompt diagnosis. This systematic review investigates the current trends in deep learning (DL) to the identification and classification of kidney stones in the different imaging modalities. The identification of the relevant articles was conducted through a thorough search in PubMed, IEEE Xplore and Google Scholar, including the articles published in 2020-2025 with the total of 700 publications, of which 49 articles were included in the research. It was found that the number of publications sharply increased in 2022 and was caused by the growing interest in AI-based urological diagnostics. Imaging was done by CT, ultrasound, X-ray and endoscopy, CT being the most popular, because it displayed a better spatial definition. Classical DL architectures such as VGG16, ResNet, Inception, and Xception presented high accuracy, especially that related to CT-based tasks. New models like Swin Transformer, RT-DETR, AMC-AM, CNN-U-Net proved effective in dealing with noisy data of ultrasounds, recruitment inequalities, and multi-scale detection. Nevertheless, the constraints still exist as the majority of the studies have been based on publicly available data (e.g., KiTS19, KiTS21, Kaggle) that is demographically and clinically biased. The metrics used to evaluate were not reported consistently and little emphasis on external validation and explainability were placed on any of the metrics in use. Although old machine learning algorithms were sometimes used, deep neural networks were popular because of their exceptional functions with image data. In general, DL has great potential in the improvement of the KSD diagnosis and treatment. Future research should focus on explainable artificial intelligence (XAI), multimodal dataset, and real-time applications enabled by the edge computing technology to facilitate clinical implementation. With the development of AI and, especially, large language models and federated learning, nephrology solutions with more clinical use cases are to be anticipated.

Keywords: artificial intelligence, machine learning, deep learning, stone, disease, detection, prediction, healthcare

1. Introduction:

Artificial Intelligence is now very useful in healthcare as it can handle medical data with remarkable efficiency. Within AI, machine learning (ML) allows systems to learn from data and make accurate predictions for cases that have not been seen before. As deep learning has advanced quickly within ML, AI has delivered highly promising results in analyzing medical images. Deep learning has been recognized in the field of kidney stone detection for automatically identifying and classifying stones in different imaging tests which aids in accurate diagnosis, preparing for treatment and monitoring the condition.

The human excretory system is responsible for maintaining homeostasis by removing metabolic wastes from the body. The main organs of this excretory system include the kidneys, ureters, bladder, and urethra. Of all these organs, the kidneys are the main filtering organs that remove toxins, specifically nitrogen wastes such as urea, from the blood. These wastes are eliminated from the body by excretion through urine with the help of the coordinated effort of the urine's passage through the ureters, bladder, and urethra [1].

One of the most common excretory system disorders is nephrolithiasis, or kidney stone disease. Kidney stones are hard crystal formations created by the concentration and solidification of dissolved minerals and salts, mainly within the renal pelvis [2]. Stones usually form because of imbalances of urine pH, decreased urine output, food factors, or underlying metabolic and genetic diseases. Most of the stones are made of calcium oxalate or calcium phosphate (about 80%), followed by struvite (10%), and uric acid stones (9%). Few purine-containing stones and drug-induced ones are described, usually related to certain genetic diseases, but are harder to identify by using traditional methods [3].

The passage of kidney stones along the urinary tract is usually associated with severe and severe pain, usually localized at the lower back, flanks, or lower abdomen. The pain is often accompanied by scalding or painful urination, hematuria (presence of blood in urine), nausea, vomiting, and fever [4]. If left untreated or if it is a repeated occurrence, it leads to obstruction of the urinary tract, renal impairment, or even renal failure, which considerably lowers quality of life and is life-threatening in severe cases [5,6]. Since the passage of stones hinders the process of filtration and excretion, it directly impairs kidney function and leads to potential damage over a longer period.

Epidemiological studies indicate an increasing worldwide burden of kidney stone disease. Lifetime prevalence is calculated to be 12% among males and 7% among females in developed nations. Incidence is disparate globally, at 1–19% for Asia, 4% for South America, and 5–10% for Europe. In a multicentric survey from seven nations, incidence varied from 114 to 720 cases per 100,000 population, with a prevalence ranging from 1.7% to 14.8%, with a trend of an ongoing increase [3,7]. These disparities are explained by variations in food, climate, genetic susceptibility, and accessibility of medical care.

In spite of developments in biochemical, imaging, and spectroscopy methods for analysis of stones, there is no single method available that gives a complete picture of the composition and internal architecture of the calculi. Traditional chemical and physical tests are often unable to correctly diagnose rare or compound stones, particularly those with metabolic or drug-induced etiologies [8]. This diagnostic void emphasizes the need for accurate, timely, and complete detection methods, particularly with an increase in incidence and recurrence of nephrolithiasis globally.

In response to these challenges, Deep learning models are able to address these challenges by increasing the accuracy of medical diagnoses, analyzing images as they are produced and handling the difficulties of traditional approaches.This analysis examines the present state of deep learning in detecting kidney stones, showing current issues, future research paths and opportunities for clinical implementation.

2. Methodology

Each study in the systematic literature review revealed the latest developments in deep learning as applied to kidney stones detection between 2020 and 2025. The review also offers a description of the current trends in the automated detection based on deep learning regarding different imaging modalities and main trends and datasets that are utilized, as well as evaluation strategies. As illustrated in Fig. 1, objective definition and designing research questions marked the beginning of the process. Thereafter, a thorough search plan was formulated that entailed the choice of databases and determination of keywords. Relevance and quality of the chosen studies were guaranteed with the help of inclusion and exclusion. The important information extracted by data extraction included the type of datasets, and architecture of the models, and evaluation statistics. Lastly, the studies chosen were classified by the imaging modality, the deep learning method, and the performance results providing insight into the present-day and the future of this field of research.

Fig 1. Process of study analysis

2.2 Research Objectives

The main purpose of this review is to analyze in depth the current methods of deep learning and machine learning that identify and classify kidney stones through CT, X-ray, or ultrasound images. The purpose of this review is to spotlight common research paths, notable architectures, used datasets, and the main technical issues in 3D segmentation. The analysis sorts studies according to imaging methods, types of models, and ways the data is applied so as to list any gaps and highlight the latest methods being used in AI-based diagnostics for kidney stones.

2.3 Research Queries

The following are some research questions that are stated to begin the systematic review:

RQ1: What major type of research is being done for detecting kidney stones?
RQ2: What imaging modalities are commonly used for kidney stone diagnosis?
RQ3: What are the technical backgrounds of deep-learning architectures used in kidney stone detection?
RQ4: What are the latest trends in kidney stone recognition using AI?
RQ5: What are the publicly available datasets used for kidney stone detection research?
RQ6: Which measures were used to report the performance of automatic kidney stone detection using deep learning, and how well did these approaches actually perform?

RQ1 looks at studies published between 2020 and 2025 to get a clear picture of how research in deep learning for kidney stone detection has evolved in recent years. RQ2 explores which imaging techniques are most commonly used and considered most effective in this area. RQ3 takes a closer look at the deep learning methods being applied, aiming to understand their strengths and limitations. RQ4 dives into the latest techniques and approaches being explored for detecting kidney stones. RQ5 focuses on the datasets used to train and test these models, helping to highlight what data is available and how it's being used in current research. Lastly, RQ6 looks at the performance measures that are used to evaluate various deep learning methods and determines how well they detect kidney stones.

2.4 Search Strategy

In order to achieve a full scope of research within the medical and technical communities, such respectable digital libraries and scientific databases were used in this review as ScienceDirect, Elsevier, IEEE Xplore, Google Scholar, Springer, MDPI, Nature, and Wiley Online Library. The choice of these platforms was made to cover a large scope of literature about the usage of machine learning (ML) and deep learning (DL) algorithms in the detection, classification, and diagnosis of kidney stones with the use of medical imaging modalities, CT, ultrasound, and X-ray.

The research strategy was structured on formulated research questions that were used to guide the selection of key keywords and phrases. Particular words like kidney stone, renal stone, nephrolithiasis, sonography, ultrasound, deep learning, machine learning, explainable AI, CNN, and diagnosis. The refinements and the broadening of results with the application of Boolean operators (AND, OR) were used where needed. The keyword strings in each platform were customized. In the case of PubMed, the search terms included the combination of the words, such as the following: kidney stone AND deep learning and classification, whereas in Springer, MDPI, they involved broader terms, such as the following: urolithiasis, image recognition, and AI-based classification.

The number of articles retrieved by the first search, which was carried out in all databases and in approximately the same time period (2020 to 2025), was rather large. But the research published before 2020 or not related to the detection of kidney stones with DL/ML algorithms was filtered out at an initial stage. All records before 2000 were also excluded from the search because it was not match the current trend of AI development. 49 research articles were identified that fulfilled the preset selection criteria, and therefore used in this systematic review.

Inclusion and exclusion were done very well. Duplicated articles that were found in various databases were eliminated. Inclusion criteria included the requirement that a study be written in English, published in Q1 or Q2 journals, and must involve the use of deep learning or other artificial intelligence methods (e.g., transfer learning, transformer-based models) to detect or predict abnormalities in the kidney stones, tumors, or cysts. Medical studies included in the eligibility criteria should have employed applicable modalities of the images, such as X-ray, CT, or ultrasound. The inclusion and exclusion criteria excluded the studies on different types of stones, similar theoretical or discussion-based investigations without actual implementation, and peer-reviewed documentary or conference papers, as well as preprints. Also, articles were removed when they considered only maternal information or were using ML/DL without speaking clearly about it.

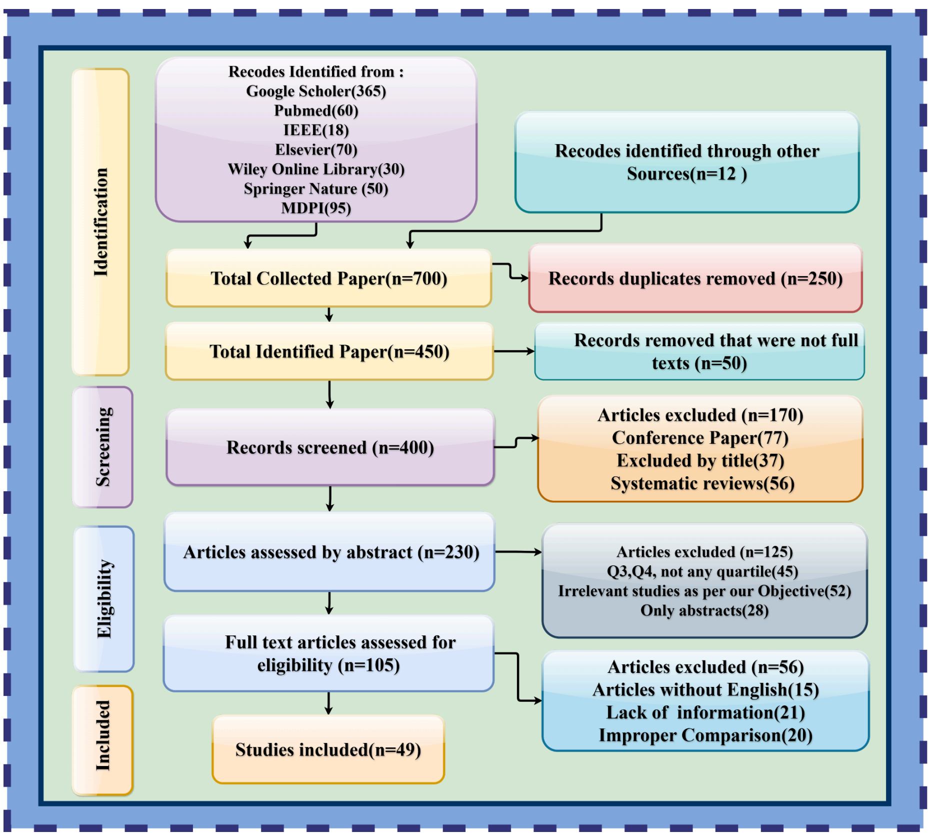

Fig 2: Diagram of Study Selection Process

The lead reviewer read the title and the abstract of all the articles. Where ambiguity prevailed the complete text was considered. During circumstances where there was an inconsistent reviewer, there was a consultation over the topic until there was an agreement. To keep transparency in the selection of studies, this review was carried out using a PRISMA-based workflow as depicted in Figure 2, where both inclusion and exclusion criteria were applied[9].

The last sample of 49 articles was identified that met all inclusion criteria and offered empirical evidence on machine learning usages regarding kidney stone detection systems. They are the center of our review as they provide helpful insights into the recent developments, the dataset usage, and the model's performance in this field. All articles that made it into the review are listed in the Supplementary Matarials presented in the supplements to this review.

2.5 Quality assessment of the study

To ensure methodologically sound and compliant inclusion, in Table.1 a research appraisal was conducted for all studies. This appraisal aimed to assess the scientific rigor, transparency, and relevance of the research on kidney stone formation and classification using deep learning techniques.

Table 1.Assessment Criteria Table

No.	Assessment Criteria	Scoring Options
Q1	Is the research problem clearly stated and relevant to kidney stone detection?	Yes (1), Somehow (0.5), No (0)
Q2	Is the deep learning methodology well described?	Yes (1), Somehow (0.5), No (0)
Q3	Is the dataset type, source, and pre-processing clearly explained?	Yes (1), Somehow (0.5), No (0)
Q4	Are performance metrics presented clearly?	Yes (1), Somehow (0.5), No (0)
Q5	Is the contribution of the study clearly stated?	Yes (1), Somehow (0.5), No (0)
Q6	Are limitations and future research directions discussed?	Yes (1), Somehow (0.5), No (0)

Each study could receive a maximum score of 6.0. The cumulative scores were then categorized into three quality tiers:

- High Quality: 5.5 – 6.0
- Moderate Quality: 4.0 – 5.4
- Low Quality: Below 4.0

Table 2.Quality assessment

Author	**Quality assessment**						**T. score**
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	
Black et al.[42]	0.5	1	0.5	1	0.5	1	4.5
Sudharson and Kokil[10]	1	1	1	1	0.5	1	5.5
Yildirim et al.[5]	0.5	1	0.5	1	1	1	5
Sudharson and Kokil[34]	1	1	1	1	0.5	0	4.5
Onal and Tekgul[43]	0.5	1	0.5	1	0.5	1	4.5
Beze et al.[40]	0.5	1	0.5	1	1	1	5
Xiang et al.[53]	0.5	1	1	1	1	1	5.5
Baygin et al[14]	1	1	1	1	1	1	6
Elton et al.[15]	1	1	1	1	0	1	5
Zhang et al.[18]	0.5	1	0.5	1	1	1	5
LOPEZ-TIRO et al.[41]	1	1	1	1	0.5	1	5.5
Patro et al.[16]	1	1	1	1	0.5	1	5.5
Kilic et al.[1]	1	1	1	1	0.5	1	5.5
Yan and Razmjooy [17]	0.5	1	1	1	1	1	5.5
Bingol et al.[19]	0.5	1	1	1	1	1	5.5
Wu et al[49]	1	1	1	1	1	1	6
Gulhane et al.[47]	1	1	0.5	1	0.5	1	5
Asif et al.[20]	1	1	1	1	0	1	5
CHAKI et al.[21]	1	1	1	1	0.5	1	5.5
Ahmed et al.[38]	1	1	1	1	1	1	6

Liu And Ghadimi[22]	1	1	1	1	0.5	1	5.5
Zhu et al[23].	0.5	1	0.5	0	1	1	4
Yenikekaluva et al[24].	1	1	1	1	0.5	1	5.5
Islam et al[25]	1	1	1	1	1	1	6
Villanueva et al[52].	1	0	1	1	0.5	1	4.5
Cui et al[26]	1	1	1	1	0.5	1	5.5
C Venkata Narasimhulu[35]	0.5	1	1	0	1	1	4.5
Caglayan et al[11].	1	1	1	1	0.5	1	5.5
Viswanath et al[36].	1	0.5	0.5	1	1	1	5
Li et al[27]	1	1	1	1	1	1	6
Kavoussi et al [51]	1	0	1	1	0.5	0.5	4
Babajide et al[28].	1	0	1	1	0.5	1	3.5
Steuwe et al[29l.	1	1	0.5	0	0.5	1	4
Kaviani et al[3].	1	1	1	1	0.5	1	5.5
Asif et al[7].	1	1	1	1	0.5	1	5.5
Sundaramoorthy et al[30].	1	1	1	0	0.5	1	4.5
Maqsood et al[8].	1	1	1	1	1	1	6
Karthikeyan et al[4]	1	1	1	1	1	1	6
Kim et al[33].	0.5	1	1	1	0.5	0.5	4.5
Zhu et al[31]	1	1	0.5	1	0	1	4.5
Kulandaivelu et al[46]	1	1	1	1	0	1	5
Vasanthi et al[32]	1	1	0.5	1	0	0.5	4
Pimpalkar et al[12]	1	1	1	1	1	1	6
Asaye et al[37]	1	1	1	1	0	1	5
Alghamdi and Amoudi et al[5]	1	0.5	1	1	1	1	5.5
Zapataa et al[39]	1	1	1	1	0	0.5	4.5
Kumar et al[13]	1	0.5	1	1	0	1	5.5
Bhandari et al[45]	1	1	1	1	1	1	6

Abdimurotovich and Cho[48]	1	1	1	1	1	0.5	5.5

From Table.2 the reviewed 49 studies had 9 studies with a trial score of 6.0, 17 studies had 5.5, 10 studies had 5.0, 9 trials had 4.5, 4 had 4.0 and One study had a score of 3.5, there were no studies that scored below 3.5. This led to 26 Was thought of (53%) having a high-quality rating, 22 (45 %) a moderate one, and 1 (2 %) low quality. The vast majority of the studies provided powerful methodology with most of them specifying their goals, extensive models of learning, data, imaging schemes, and performance indicators very clearly. Nonetheless, only a small number of them described data and presented reasons behind innovations and discussed limitations in detail. It lacked also the reproducibility and had insufficient information of data manipulation as well as the clinical translatability. The moderate quality studies also did not discuss overfitting and dataset biases a lot.

3. Result

In this section, the main findings and revelations from the studies reviewed in this thorough study are presented. It investigates crucial parameters such model accuracy, imaging modality types, dataset distribution, and the performance of various deep learning architectures. A comprehensive picture of the state of AI-based kidney stone detection is given by the analysis, which examines changes over time in addition to variables like country of origin, journal quality, and dataset characteristics. The goal is to draw attention to the models that are working well, the imaging techniques that are working well, and any areas in which the research is lacking or inconsistent. Ultimately, these insights are intended to guide future work, helping researchers refine their approaches and move closer to developing AI tools that are not only accurate but also ready for real-world clinical use.Our research question-specific findings are presented in distinct subsections below.

3.1 RQ1. What major type of research is being done for detecting kidney stones?

Kidney stone detection research has shifted from using traditional approaches to deep learning and new kinds of features in the years from 2020 to 2025 see figure.16. In 2020 and 2021, the main activities were focused on using straightforward CNNs and transferred architectures like VGG16 and ResNet to scan images captured by CT and ultrasounds, mostly using set data without much variation[5, 10,11]. Initially, neural networks were made for deciding between stones and normal abdominals, as they lacked a lot of available data and had little variation. Still, a major technological shift kicked in starting in 2022, as researchers started to focus more on hybrid and complicated models for spotting, identifying, categorizing and differentiating stones in images, including those of calcium oxalate, uric acid and cystine. An important innovation here was the AMC-AM method that brought together VGG16, ResNet and Inception modules to form one model that provides attention-based, multi-scale features. Advanced techniques like MSD-CMPA for automatically adjusting hyperparameters made the system much better suited to cases of imbalance, noise and low resolution[12,13].

CT imaging can clearly identify body structures and is relevant for both detection and analysis. it was the modality covered by the highest number of journals (over 30) in the research document [2-5,7,8,11,13-33]. 5 journals featured ultrasound and another two featured X-ray, mostly because they are considered low-cost and non-invasive[1,10,34-38]. Among publishers, Elsevier and Springer were the leading disseminators, with each publishing 15 and 13 articles, respectively, meaning both play key roles in multidisciplinary DL and medical imaging areas. Leading journal publications were seen in 2024 (15 journals), representing the maturation of common multimodal and combined approaches, yet published journal volumes dropped to 5 in 2025 which might reflect an effort to apply the research innovations in clinical use.

All in all, research efforts from 2020 to 2025 moved from basic diagnostic tools to modern AI systems that aid doctors in their decisions. From Fig.3 Between 2022 and 2025, there was a rapid increase in advanced DL models, fusion of different types of data, and checking their usefulness in clinical practice. These advancements prove both how advanced technology is today and how focused the field is on solutions that actually work, which makes this period key in the growth of automated kidney stone detection and classification.

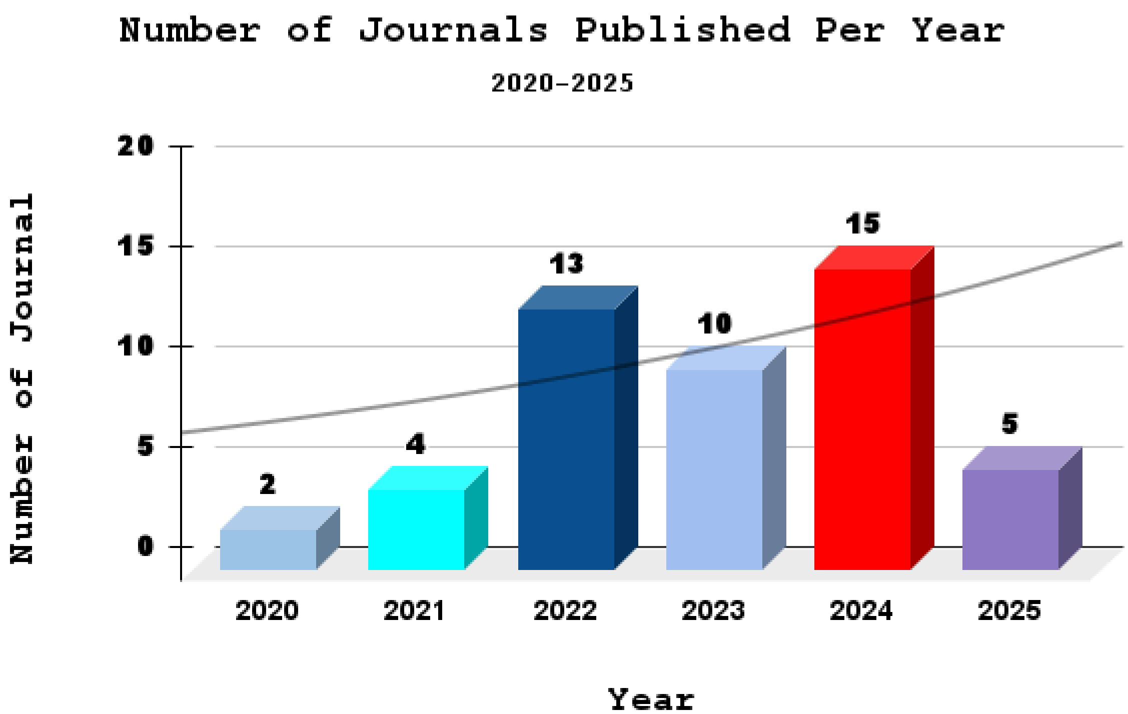

Fig 3: Number of Journals in each Year

3.2 RQ2. What imaging modalities are commonly used for kidney stone diagnosis?

Across the 49 selected studies, a total of five imaging modalities were identified across the reviewed studies, with Computed Tomography (CT) being the most frequently used and widely supported. From fig.4 CT was discussed in 30 journal articles and is considered the gold standard for kidney stone detection due to its ability to accurately identify stones of varying size, composition, and density[2-5,7,8,11,13-33]. Its high-resolution imaging and three-dimensional visualization capabilities make it indispensable for both diagnosis and treatment planning. Ultrasound ranked second in frequency, especially in cases involving children, pregnant patients, or low-resource settings, where minimizing radiation exposure is critical. Although it provides less detail than CT, ultrasound is still valued for detecting larger stones and signs of obstruction[10,34-37]. X-ray (KUB), while less sensitive and limited to radiopaque stones, remains relevant in basic screening and resource-constrained environments[1,38]. Endoscopy appeared in four studies and is primarily used during surgical procedures for stone visualization and removal [33,39-41]. Microscopic imaging, referenced in two articles, plays a crucial role post-surgery by enabling chemical and structural analysis of stones, which informs recurrence prevention strategies[42,43]. Overall, CT remains the dominant imaging modality for kidney stone detection and classification, with ultrasound, endoscopy, and microscopy offering valuable support depending on clinical context. This distribution aligns with current best practices and provides guidance for future diagnostic strategies and research directions.

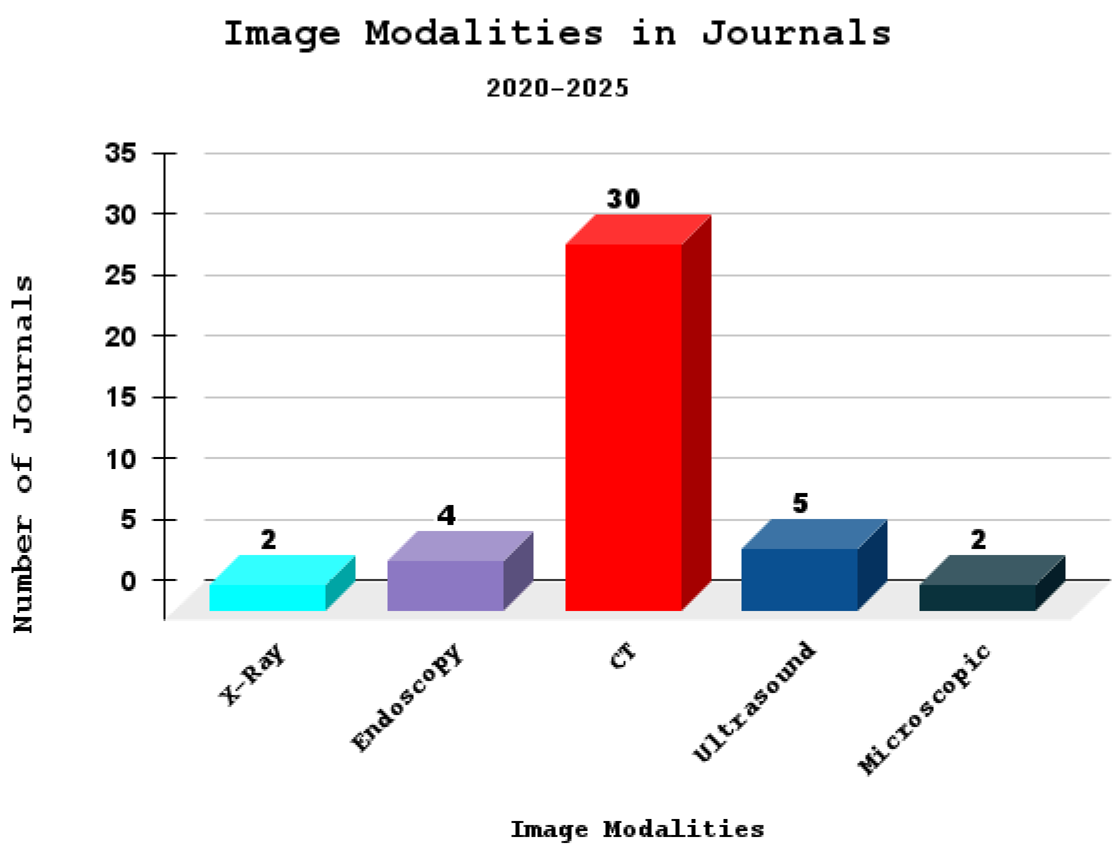

Fig 4: Number of journals across the modalities

3.3 RQ3. What are the technical backgrounds of deep-learning architectures used in kidney stone detection?

The architectures of deep learning (DL) architectures to detect kidney stones in 2020-2025 used convolutional neural networks (CNNs), residual networks, encoder-decoder, multi-scale feature extraction, and transformers as primary components. These models were created to overcome issues in medical imaging which include insufficient annotated data, modality differences, imaging noise and consistent localization accuracy. Most architectures used included CNN-based models (e.g. VGG16, AlexNet, InceptionV3 and Xception) where they extract spatial and texture-based features and identify and classify stones [13,20,42,43]. ResNet and its variations (e.g. ResNet-101, XResNet-50, etc.) were popular due to deep feature extraction and solution to vanishing gradients [5,10,34]. The U-Net and 3D U-Net were better at performing the segmentation task because they created skip connections, thus preserving both contextual and spatial representation [15,26,27]. Multi-scale feature analysis at a lower computational cost was possible using InceptionV3 and Xception [13,20]. To enhance generalization and classification, hybrid and ensemble models, e.g., VGG16, ResNet, and Inception combined; or home-built models, e.g., StoneNet, have entered the stage [2,44]. Transformer classifications of models were used such as Swin Transformers to improve understanding of the global context. YOLO was employed in real-time application aims at detecting objects in real-time. At the same time, interpretable classical ML models, such as Random Forest (RF) and XGBoost, were introduced to radiomic features regarding small or structured forms of data [33].In Fig. 5 there is a basic flow of the deep learning model architecture. In sum, CNNs played important roles in classification, ResNets in deep feature representation, U-Nets in segmentation, Inception/Xception in multi-scale features extraction, YOLO in fast detection, and transformers in contextual reasoning, which were constituted as the technological pillar of AI in kidney stone diagnosis.

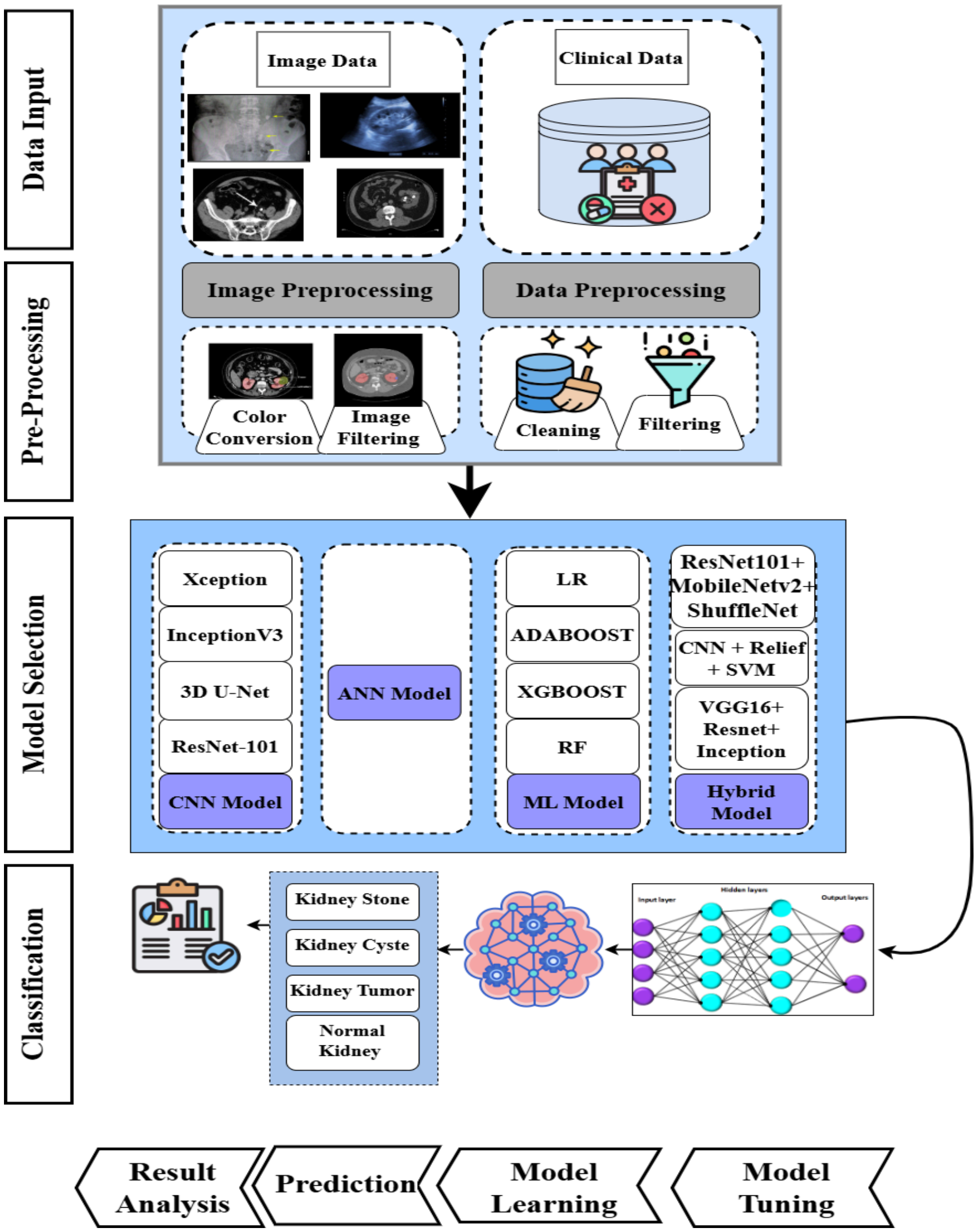

Fig 5. Overall diagram of all deep learning methods to the detection of kidney stone

3.4 RQ4 . What are the latest trends in kidney stone recognition using AI?

Recent advances in AI-based kidney stone recognition involve the predominance of CNNs, namely VGG16, ResNet, Inception, and Xception, that have been shown to achieve high accuracies regarding CT-based classification and detection assignments, usually over 98% [7,14,16]. State-of-the-art networks such as Xception and InceptionV3 even achieve almost perfect performance on standard benchmarks. The use of hybrid models like AMC-AM with several backbones used in CNN and optimized with MSD-CMPA solves the issue of class imbalance and generalization [12,13]. Models based on transformers such as Swin Transformer and RT-DETR are being developed because they have global context sensitivity and are real-time-efficient in detecting objects even in the presence of noisy information, such as ultrasound information, despite their inability to give precise stone features [32]. Multimodal data integration and ensemble learning are also some of the shifting directions evidenced through the integration of models such as VGG16, ResNet, and Inception to enhance segmentation [2]. Nevertheless, problems continue to ensue; numerous models are taught using small, local data, which impedes generalization and X-ray-only methods lack non-calcified stones [5,22,38]. In general, the area is leaning toward the models which are accurate, efficient, and interpretable. The future of health care is in standardizing the datasets, the potential utilization of various sources of data and creating clinically adaptable artificial intelligence systems.

3.5 RQ5. What are the publicly available datasets used for kidney stone detection research?

A number of publicly available datasets can be used to identify kidney stones based on ultrasound, CT and X-ray data. The datasets of ultrasound-images.com and ultrasoundcases. info (2020-2021) contain balanced images distributed across four categories -stone, cyst, tumor, and normal- which makes it appropriate in developing models that address the diverse quality of images [10,34]. Another publicly available resource related to the stone vs. normal prediction is the Kidney Stone Detection dataset available on GitHub because of its high-resolution structure [34]. The large-scale Kaggle CT dataset (Dataset-06) has more than 12,000 images in the four categories: cyst, stone, tumor, and normal within CT imaging. It has been widely applied whereby it is diverse and voluminous and helps in classification, as well as decision-support modeling [8,17,19,21,22,25,45]. Larger data such as the ACR Data Science Institute dataset, and Li et al. data set of Zenodo CT scans provide well-annotated images and metadata, which can alleviate the problem of class imbalance and can lead to more generally applicable models[15,27]. KITS19 and KITS21 also are broadly applied in the segmentation literature, as they are highly annotated and used in surgical planning[46]. Also, the Kaggle dataset provided by Vasanthi et al. contains 1054 training images of ultrasound CT Braun and labeled specifically to detect the presence of a stone [32]. In X-ray, DUSX dataset can be used as a cheaper replacement to CT but is problematic since anatomical features are overlapped [1,38]. From Table.3 datasets alone allow consistent benchmarking, algorithm comparison, and creation of clinical models of kidney stone detection in various imaging modalities.

Table 3. Dataset Source

Authors	Type	Amount	Class	Link
Sudharson and Kokil(2020)[10]	Ultrasound	4940	Cyst 1235, Normal 1235,Stone 1235,Tumor 1235	**Dataset-01:**https://www.ultrasound-images.com/#google_vignette **Dataset-02:**https://www.ultrasoundcases.info/
Yildirim et al.(2021)[5]	CT	433	278 Kidney Stone 165 normal.	**Dataset-03:**https://github.com/yildirimozal/Kidney_stone_detection
Sudharson and Kokil (2021)[34]	Ultrasound	4940	Cyst 1235, Normal 1235,Stone 1235,Tumor 1235	**Dataset-01:**https://www.ultrasound-images.com/#google_vignette **Dataset-02:**https://www.ultrasoundcases.info/
Baygin et al(2022)[14]	CT	433	278 Kidney Stone 165 normal.	**Dataset-03:**https://github.com/yildirimozal/Kidney_stone_detection
Elton et al(2022)[15]	CT	180	91 kidney stone, 89 normal	**Dataset-04:**https://www.acr.org/data-science-and-informatics/ACR-Data-Science-Institute
Li et al(2022)[27]	CT	500	119 scan thickness, 209 stone,50 normal	**Dataset-10:**https://zenodo.org/records/6042410
Patro et al(2023)[16]	CT	433	278 Kidney Stone 165 normal.	**Dataset-03:**https://github.com/yildirimozal/Kidney_stone_detection
Kilic et al.(2023)[1]	X-Ray	630	558 kidney stone 72 normal	**Dataset-05:**https://github.com/ugrkilc/DUSX-Dataset
Yan and Razmjooy (2023)[17]	CT	12,446	Cyst 3709, Normal 5077, Stone 1377, Tumor 2283	**Dataset-06:**https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone?resource=download
Bingol et al(2023)[19]	CT	12,446	Cyst 3709, Normal 5077, Stone 1377, Tumor 2283	**Dataset-06:**https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone?resource=download
Bhandari et al.(2023)[45]	CT	12,446	Cyst 3709, Normal 5077, Stone 1377, Tumor 2283	**Dataset-06:**https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone?resource=download
Gulhane et al(2024)[47]	Clinical Data	79	56 % kidney stones 44 % is Normal Kidney	**Dataset-07**:https://link.springer.com/chapter/10.1007/978-1-4612-5098-2_45
Asif et al(2024)[20]	CT	433	278 Kidney Stone 165 normal.	**Dataset-03:**https://github.com/yildirimozal/Kidney_stone_detection
CHAKI et al(2024)[21]	CT	12,446	Cyst 3709, Normal 5077, Stone 1377, Tumor 2283	**Dataset-06:**https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone?resource=download
Liu And Ghadimi(2024)[22]	CT	12,446	Cyst 3709, Normal 5077, Stone 1377, Tumor 2283	**Dataset-06:**https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone?resource=download
Islam et al(2022)[39]	CT	12,446	Cyst 3709, Normal 5077, Stone 1377, Tumor 2283	**Dataset-06:**https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone?resource=download
Maqsood et al.(2024)[8]	CT	12,446	Cyst 3709, Normal 5077, Stone 1377, Tumor 2283	**Dataset-06:**https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone?resource=download

Karthikeyan et al.(2024)[4]	CT	12,446	Cyst 3709, Normal 5077, Stone 1377, Tumor 2283	**Dataset-03:**https://github.com/yildirimozal/Kidney_stone_detection **Dataset-06:**https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone?resource=download **Dataset-12:**https://www.uofmhealth.org/conditions-treatments/adult-urology/kidney-stones **Dataset-13:**https://www.sciencephoto.com/search?q=kidney+stone%2C+ct+scans&license=F
Kumar et al.(2024)[13]	CT	12,446	Cyst 3709, Normal 5077, Stone 1377, Tumor 2283	**Dataset-06:**https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone?resource=download
Abdimurotovich and Cho(2024)[48]	CT	778	Not Mentioned	**Dataset-16:**https://universe.roboflow.com/tez-nwkf5/tez_roi_aug
Kulandaivelu et al.(2025)[46]	CT	Dataset-14:100 Dataset-15:300	100 Kidney Stone	**Dataset-14:**https://www.kaggle.com/datasets/user123454321/kits19-2 **Dataset-15:**https://github.com/neheller/kits21#kits21
Vasanthi er al.(2025)[32]	CT	1054	1054 Kidney stone	**Dataset-11:**https://www.kaggle.com/datasets/safurahajiheidari/kidney-stone-images/data
Pimpalkar et al.(2025)[12]	CT	12,446	Cyst 3709, Normal 5077, Stone 1377, Tumor 2283	**Dataset-06:**https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone?resource=download

3.6 RQ6: Which measures were used to report the performance of automatic kidney stone detection using deep learning, and how well did these approaches actually perform?

Most of the studies published between 2020 and 2025 that reviewed cases did not investigate regression models with the ability to generate continuous outputs [26-28,36]. Based on these studies, 14 metrics of performance were used by the researchers with accuracy, area under the receiver operating characteristic curve (AUC-ROC), sensitivity, and specificity used most frequently to evaluate classification models. It is notable that every regression came with a deep neural network (DNN) and was evaluated through specific measures that focused on Dice scores and correlation coefficients.

In the case of classification models, accuracy was reported most commonly (90 percent of articles), and highly accurate systems demonstrated outstanding performances of higher than 99 percent [8,14,21]. The sensitivity values were also rather diverse, with some studies reporting a perfect classification and the specificity values demonstrated the same range, with 90.3% accuracy and 100% accuracy [10,14,26,34]. Precision and recall balanced F1-score was especially resistant to recent analysis whereas in its predecessors, the accord was lower when dealing with more problematic subtypes of stones, such as brushite [21,42].

As to the model performance, the majority of papers have shown proof of concept with performance being significantly above chance levels [47,49]. In some other studies, (almost) perfect performance was achieved [8,14,16]. But the performance differed greatly depending on the stone type, detection of uric acid performed better in most cases than the less common compositions of cystine and struvite [42,43].

The four regression experiments indicated a good potential with alternative metrics Dice score to achieve the whole characterization of the stone[27].

The most accurate classification models were the one that had high accuracy (>99 %) and sensitivity and specificity levels were balanced across all types of stones whereas regression methods proved to be specifically useful at segmentation tasks [14,21,27]. Nonetheless, the lack of coherency in reporting on metrics lack of specificity and the presence of mostly binary classification models are the current issues facing the field [5]. Such results indicate that although the existing systems have valuable diagnostic performance, there is still a chance of better detection of rare type of stones and usage of overall regression-based strategies.

4. Discussion:

In this work, we systematically reviewed articles on deep-learning approaches for the detection of kidney stones. Although still a relatively underrepresented area of AI application compared to domains such as cancer, mental health, and chronic disease diagnostics, our systematic review reveals a clear upward trend in research activity from 2020 to 2025. As shown in Fig. 3, the number of journal publications increased from just 2 in 2020 to 15 in 2024, with a notable surge beginning in 2022 (13 publications). A slight stagnation was observed in 2021, which may be attributed to the impact of the COVID-19 pandemic, as clinical data collection and imaging access were significantly constrained during that time. This effect was particularly pronounced in non-urgent specialties such as nephrology. The decline in publication count in 2025 (5 publications) may reflect delayed project completion, shifts in research focus, or pending data publication timelines. Overall, the trend illustrates growing interest and progress in leveraging deep learning for kidney stone diagnosis.

The findings reveal a growing interest in AI-driven diagnostic support tools, particularly within nephrology and urology. Most studies used imaging-based data, with computed tomography (CT) emerging as the predominant modality due to its superior spatial resolution and ability to differentiate stone types and sizes[2-5,7,8,11,13-33]. Ultrasound, known for its safety and accessibility, was widely employed in pediatric and low-resource contexts, while X-ray, endoscopy, and microscopy played supporting roles depending on the clinical scenario[1,10,33-43].

The evolution of DL models was clear across the studies. Early research relied on basic convolutional neural networks (CNNs) like VGG16, ResNet, and Inception, achieving high accuracy, especially in binary classification tasks using CT images[11,38,40]. However, these models were limited by dataset homogeneity and narrow clinical scope. More recent studies introduced hybrid architectures such as AMC-AM, transformer-based models like Swin Transformer and RT-DETR, and advanced segmentation networks like U-Net and 3D U-Net[2,15,24-28,30,32]. These newer approaches addressed challenges like class imbalance, noise in ultrasound images, and the need for multi-class classification. YOLO-based methods were also explored for real-time object detection, reflecting a shift toward more adaptive, real-time, and clinically relevant systems[1].

Despite technical advancements, most models relied on image-only datasets—often from public sources like KITS19, KITS21, and Kaggle—which lacked demographic diversity, multi-institutional validation, and integration of patient metadata[8,12,13,17,19,21,22,39,45,46]. Unlike more mature AI applications in oncology or neurology, multimodal data fusion remains underutilized in kidney stone detection. Only a few studies incorporated patient history, clinical notes, or biochemical data, even though combining such information could significantly enhance diagnostic accuracy and support precision medicine[47,49,50,51,52,53].

Evaluation metrics across the studies were inconsistent. While accuracy, AUC-ROC, sensitivity, and specificity were commonly reported, reliance on accuracy alone can be misleading, especially in imbalanced datasets[10,14,21,26,34,42]. Notably, only 24 of the 28 studies that reported sensitivity also included specificity[1,3,7,11,16,17,20,21,26,27,35,40,47,49]. Metrics like F1-score, positive predictive value (PPV), and negative predictive value (NPV), though clinically important, were underreported[26,40,43,46]. This lack of standardized evaluation hinders reproducibility and makes performance comparisons across studies unreliable.

Moreover, few studies addressed explainability or external validation. The lack of explainable AI (XAI), real-time deployment frameworks (e.g., edge computing), and federated learning solutions raises concerns about the clinical readiness of current models[3,38,45]. Ethical issues such as dataset bias, regulatory oversight, model liability, and patient privacy were largely unexamined. While traditional machine learning methods like random forests and support vector machines remain useful for feasibility studies, sophisticated deep neural networks require larger datasets, more rigorous validation, and greater transparency to be trusted in clinical settings.

In summary, deep learning has significantly advanced kidney stone detection across diverse imaging modalities, yet challenges remain. Limited dataset diversity, minimal use of multimodal data, insufficient explainability, and non-standard evaluation practices currently impede real-world integration. Future research must adopt standardized metrics like AUC-ROC and F1-score, incorporate diverse and multisource datasets, develop interpretable and privacy-conscious AI models, and validate performance across institutions. Only through these steps can AI systems evolve from promising prototypes into reliable diagnostic tools that support clinical decision-making and improve patient outcomes in nephrology.

4.1 Limitation:

There are various limitations to this study that need to be mentioned. The first one is that despite our thorough search of the literature, a great number of relevant works published throughout the study period might be not captured because of time limitations or coverage of the databases we used. Also, although we used the PRISMA guidelines as much as possible, all of the requirements of a review are not met in this one, especially the parts that concern bias evaluation, synthesis methodology, and reporting of study selection [9]. We were only focusing on kidney stone detection and classification and thus the related but unrelated urological and other conditions as well as additional applications of deep learning to other regions of diagnosis were beyond our scope in this review. In addition, we did not include the names of models, e.g., ResNet, U-Net, Swin Transformer in our search strategy and thus some very relevant studies might not have been retrieved. The other key limitation is the inappropriate distribution of studies on the use of CT imaging, where the majority of included studies were carried out on the basis of CT data[2-5,7,8,11,13-33]. Relatively fewer studies which are grounded based on ultrasound or X-ray were found as a contrast [1,10,34-38]. Such an imbalance can bring a bias into the overall results, failing to give full voice to the feasibility and difficulty of implementing deep learning methods in various imaging modalities.

4.2 Conclusion

This review makes it clear that artificial intelligence (AI) and more specifically deep learning (DL) has greatly advanced towards detection and classification of kidney stones with different imaging modalities. In spite of a slightly limited presence of research on this field in comparison to the scope of application of AI in oncology or diagnosis of chronic diseases, a significant increase in studies has been observed, especially since 2022, which indicates an increase in interest in AI tools in nephrology. According to our analysis, models like convolutional neural networks (CNNs), transformer-based models, and hybrid models have been performing well over the years especially on studies that used CT imaging datasets. In the future, AI in kidney stone diagnosis has the potential of integrating large-scale, multiple, and multimodal data, which combines not only imaging data, but also prior patient histories, clinical reports, and biochemical analyses. Important to cloud-based AI adoption is the use of explainable AI (XAI) and privacy-preserving technologies, whose normalization will help promote transparency, make clinicians trust the system, and maintain regulatory compliance. In addition, new progress in federated learning and on-demand application of edge AI provide favorable opportunities to fulfill the promise of clinical translation and bedside decision support. Cumulatively, these trends indicate the revolutionary prospect of AI in enhancing the diagnostic processes as well as the provision of individualized treatment strategies and enhancing the accessibility and effectiveness of kidney stones management in contemporary urological practice.

Reference

1. Kilic, U., Karabey Aksakalli, I., Tumuklu Ozyer, G., Aksakalli, T., Ozyer, B., & Adanur, S. (2023). Exploring the Effect of Image Enhancement Techniques with Deep Neural Networks on Direct Urinary System X‑Ray (DUSX) Images for Automated Kidney Stone Detection. International Journal of Intelligent Systems, 2023(1), 3801485
2. Kulandaivelu, G., Suchitra, M., Pugalenthi, R., & Lalit, R. (2025). An Implementation of Adaptive Multi‐CNN Feature Fusion Model With Attention Mechanism With Improved Heuristic Algorithm for Kidney Stone Detection. *Computational Intelligence*, *41*(1), e70028.
3. Kaviani, P., Primak, A., Bizzo, B., Ebrahimian, S., Saini, S., Dreyer, K. J., & Kalra, M. K. (2023). Performance of threshold-based stone segmentation and radiomics for determining the composition of kidney stones from single-energy CT. *Japanese Journal of Radiology*, *41*(2), 194-200.
4. Karthikeyan, V., Kishore, M. N., & Sajin, S. (2024). End‐to‐end light‐weighted deep‐learning model for abnormality classification in kidney CT images. *International Journal of Imaging Systems and Technology*, *34*(1), e23022.
5. Yildirim, K., Bozdag, P. G., Talo, M., Yildirim, O., Karabatak, M., & Acharya, U. R. (2021). Deep learning model for automated kidney stone detection using coronal CT images. *Computers in biology and medicine*, *135*, 104569.
6. Kahani, M., Tabrizi, S. H., Kamali-Asl, A., & Hashemi, S. (2020). A novel approach to classify urinary stones using dual-energy kidney, ureter and bladder (DEKUB) X-ray imaging. *Applied Radiation and Isotopes*, *164*, 109267.

7. Asif, S., Zhao, M., Chen, X., & Zhu, Y. *StoneNet: An Efficient Lightweight Model Based on Depthwise Separable Convolutions for Kidney Stone Detection from CT Images. Interdiscip Sci. 2023 Dec; 15 (4): 633–652*.
8. Maqsood, F., Zhenfei, W., Ali, M. M., Qiu, B., Rehman, N. U., Sabah, F., ... & Sarwar, R. (2024). Artificial intelligence-based classification of CT images using a Hybrid SpinalZFNet. *Interdisciplinary Sciences: Computational Life Sciences*, *16*(4), 907-925.
9. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, *372*.
10. Sudharson, S., & Kokil, P. (2020). An ensemble of deep neural networks for kidney ultrasound image classification. Computer Methods and Programs in Biomedicine, 197, 105709.
11. Caglayan, A., Horsanali, M. O., Kocadurdu, K., Ismailoglu, E., & Guneyli, S. (2022). Deep learning model-assisted detection of kidney stones on computed tomography. International braz j urol, 48(5), 830-839
12. Pimpalkar, A., Saini, D. K. J. B., Shelke, N., Balodi, A., Rapate, G., & Tolani, M. (2025). Fine-tuned deep learning models for early detection and classification of kidney conditions in CT imaging. *Scientific Reports*, *15*(1), 10741.
13. Kumar, Y., Brar, T. P. S., Kaur, C., & Singh, C. (2024). A comprehensive study of deep learning methods for kidney tumor, cyst, and stone diagnostics and detection using CT images. Archives of Computational Methods in Engineering, 31(7), 4163-4188.
14. Baygin, M., Yaman, O., Barua, P. D., Dogan, S., Tuncer, T., & Acharya, U. R. (2022). Exemplar Darknet19 feature generation technique for automated kidney stone detection with coronal CT images. Artificial Intelligence in Medicine, 127, 102274.
15. Elton, D. C., Turkbey, E. B., Pickhardt, P. J., & Summers, R. M. (2022). A deep learning system for automated kidney stone detection and volumetric segmentation on noncontrast CT scans. *Medical Physics*, *49*(4), 2545-2554.
16. Patro, K. K., Allam, J. P., Neelapu, B. C., Tadeusiewicz, R., Acharya, U. R., Hammad, M., ... & Pławiak, P. (2023). Application of Kronecker convolutions in deep learning technique for automated detection of kidney stones with coronal CT images. Information Sciences, 640, 119005
17. Yan, C., & Razmjooy, N. (2023). Kidney stone detection using an optimized deep believe network by fractional coronavirus herd immunity optimizer. *Biomedical Signal Processing and Control*, *86*, 104951
18. Zhang, X., Zhang, G., Xu, L., Bai, X., Zhang, J., Xu, M., ... & Sun, H. (2022). Application of deep learning reconstruction of ultra-low-dose abdominal CT in the diagnosis of renal calculi. *Insights into Imaging*, *13*(1), 163.
19. Bingol, H., Yildirim, M., Yildirim, K., & Alatas, B. (2023). Automatic classification of kidney CT images with relief based novel hybrid deep model. PeerJ Computer Science, 9, e1717.
20. Asif, S., Zheng, X., & Zhu, Y. (2024). An optimized fusion of deep learning models for kidney stone detection from CT images. *Journal of King Saud University-Computer and Information Sciences*, *36*(7), 102130.
21. Chaki, J., & Uçar, A. (2024). An efficient and robust approach using inductive transfer-based ensemble deep neural networks for kidney stone detection. *IEEE Access*, *12*, 32894-32910.
22. Liu, H., & Ghadimi, N. (2024). Hybrid convolutional neural network and Flexible Dwarf Mongoose Optimization Algorithm for strong kidney stone diagnosis. *Biomedical Signal Processing and Control*, *91*, 106024.
23. Zhu, H., Kong, D., Qian, J., Shi, X., & Fan, J. (2024). The impact of deep learning image reconstruction of spectral CTU virtual non contrast images for patients with renal stones. *European Journal of Radiology Open*, *13*, 100599.

24. Yenikekaluva, A., Azeez, S. F., Sakegaonkar, A., Shariff, A. M., Wankhede, M., Gaikwad, S., ... & Madathiparambil Ramanathan, J. (2024). UrologiQ: AI-based accurate detection, measurement, and reporting of stones in CT-KUB scans. *Urolithiasis*, *52*(1), 170.-9
25. Islam, M. N., Hasan, M., Hossain, M. K., Alam, M. G. R., Uddin, M. Z., & Soylu, A. (2022). Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from CT-radiography. *Scientific Reports*, *12*(1), 11440.
26. Cui, Y., Sun, Z., Ma, S., Liu, W., Wang, X., Zhang, X., & Wang, X. (2021). Automatic detection and scoring of kidney stones on noncontrast CT images using STONE nephrolithometry: combined deep learning and thresholding methods. *Molecular Imaging and Biology*, *23*, 436-445
27. Li, D., Xiao, C., Liu, Y., Chen, Z., Hassan, H., Su, L., ... & Huang, B. (2022). Deep segmentation networks for segmenting kidneys and detecting kidney stones in unenhanced abdominal CT images. *Diagnostics*, *12*(8), 1788.
28. Babajide, R., Lembrikova, K., Ziemba, J., Ding, J., Li, Y., Fermin, A. S., ... & Tasian, G. E. (2022). Automated machine learning segmentation and measurement of urinary stones on CT scan. *Urology*, *169*, 41-46.
29. Steuwe, A., Valentin, B., Bethge, O. T., Ljimani, A., Niegisch, G., Antoch, G., & Aissa, J. (2022). Influence of a deep learning noise reduction on the CT values, image noise and characterization of kidney and ureter stones. *Diagnostics*, *12*(7), 1627.
30. Sundaramoorthy, S., & Jayachandru, K. (2023). Designing of Enhanced Deep Neural Network Model for Analysis and Identification of Kidney Stone, Cyst, and Tumour. *SN Computer Science*, *4*(5), 466.
31. Zhu, B., Nie, Y., Zheng, S., Lin, S., Li, Z., & Wu, W. (2024). CT-based radiomics of machine-learning to screen high-risk individuals with kidney stones. *Urolithiasis*, *52*(1), 91.
32. Vasanthi, P., Srinivasu, L. N., Teju, V., Sowmya, K. V., Stan, A., Sita, V., ... & Stan, O. (2025). Multiple kidney stones prediction with efficient RT-DETR model. *Computers in Biology and Medicine*, *190*, 110023.
33. Kim, U. S., Kwon, H. S., Yang, W., Lee, W., Choi, C., Kim, J. K., ... & Han, J. H. (2022). Prediction of the composition of urinary stones using deep learning. *Investigative and Clinical Urology*, *63*(4), 441.
34. Sudharson, S., & Kokil, P. (2021). Computer-aided diagnosis system for the classification of multi-class kidney abnormalities in the noisy ultrasound images. Computer Methods and Programs in Biomedicine, 205, 106071.
35. Narasimhulu, C. V. (2021). An automatic feature selection and classification framework for analyzing ultrasound kidney images using dragonfly algorithm and random forest classifier. IET Image Processing, 15(9), 2080-2096
36. Viswanath, K., Anilkumar, B., & Gunasundari, R. (2022). Design of deep learning reaction–diffusion level set segmentation approach for health care related to automatic kidney stone detection analysis. Multimedia Tools and Applications, 81(29), 41807-41849
37. Asaye, Y. A., Annamalai, P., & Ayalew, L. G. (2025). Detection of kidney stones from ultrasound images using machine learning algorithms. Scientific African, 28, e02618.
38. Ahmed, F., Abbas, S., Athar, A., Shahzad, T., Khan, W. A., Alharbi, M., ... & Ahmed, A. (2024). Identification of kidney stones in KUB X-ray images using VGG16 empowered with explainable artificial intelligence. Scientific Reports, 14(1), 6173.-14
39. Gonzalez-Zapata, J., Lopez-Tiro, F., Villalvazo-Avila, E., Flores-Araiza, D., Hubert, J., Ochoa-Ruiz, G., ... & Mendez-Vazquez, A. (2024). A metric learning approach for endoscopic kidney stone identification. *Expert Systems with Applications*, *255*, 124711.
40. El Beze, J., Mazeaud, C., Daul, C., Ochoa-Ruiz, G., Daudon, M., Eschwège, P., & Hubert, J. (2022). Evaluation and understanding of automated urinary stone recognition methods. *BJU international*, *130*(6), 786-798.

41. Lopez-Tiro, F., Estrade, V., Hubert, J., Flores-Araiza, D., Gonzalez-Mendoza, M., Ochoa, G., & Daul, C. (2024). On the in vivo recognition of kidney stones using machine learning. *IEEE Access*, *12*, 10736-10759.
42. Black, K. M., Law, H., Aldoukhi, A., Deng, J., & Ghani, K. R. (2020). Deep learning computer vision algorithm for detecting kidney stone composition. *BJU international*, *125*(6), 920-924.
43. Onal, E. G., & Tekgul, H. (2022). Assessing kidney stone composition using smartphone microscopy and deep neural networks. *BJUI compass*, *3*(4), 310-315.
44. Sundaramoorthy, S., & Jayachandru, K. (2023). Designing of Enhanced Deep Neural Network Model for Analysis and Identification of Kidney Stone, Cyst, and Tumour. *SN Computer Science*, *4*(5), 466.-10
45. Bhandari, M., Yogarajah, P., Kavitha, M. S., & Condell, J. (2023). Exploring the capabilities of a lightweight CNN model in accurately identifying renal abnormalities: Cysts, stones, and tumors, using LIME and SHAP. *Applied Sciences*, *13*(5), 3125.
46. Kulandaivelu, G., Suchitra, M., Pugalenthi, R., & Lalit, R. (2025). An Implementation of Adaptive Multi-CNN Feature Fusion Model With Attention Mechanism With Improved Heuristic Algorithm for Kidney Stone Detection. Computational Intelligence, 41(1), e70028.
47. Gulhane, M., Kumar, S., Choudhary, S., Rakesh, N., Zhu, Y., Kaur, M., ... & Gadekallu, T. R. (2024). Integrative approach for efficient detection of kidney stones based on improved deep neural network architecture. *SLAS technology*, *29*(4), 100159.
48. Abdimurotovich, K. A., & Cho, Y. I. (2024). Optimized YOLOv5 Architecture for Superior Kidney Stone Detection in CT Scans. *Electronics*, *13*(22), 4418.
49. Wu, Y., Mo, Q., Xie, Y., Zhang, J., Jiang, S., Guan, J., ... & Mo, C. (2023). A retrospective study using machine learning to develop predictive model to identify urinary infection stones in vivo. *Urolithiasis*, *51*(1), 84.
50. Alghamdi, H., & Amoudi, G. (2024). Using Machine Learning for Non-Invasive Detection of Kidney Stones Based on Laboratory Test Results: A Case Study from a Saudi Arabian Hospital. *Diagnostics*, *14*(13), 1343.
51. Kavoussi, N. L., Floyd, C., Abraham, A., Sui, W., Bejan, C., Capra, J. A., & Hsi, R. (2022). Machine learning models to predict 24 hour urinary abnormalities for kidney stone disease. *Urology*, *169*, 52-57.
52. Iparraguirre-Villanueva, O., Paucar-Palomino, G., & Paulino-Moreno, C. (2025). From data to diagnosis: evaluation of machine learning models in predicting kidney stones. *Neural Computing and Applications*, 1-14.
53. Xiang, L., Jin, X., Liu, Y., Ma, Y., Jian, Z., Wei, Z., ... & Wang, K. (2022). Prediction of the occurrence of calcium oxalate kidney stones based on clinical and gut microbiota characteristics. *World Journal of Urology*, *40*(1), 221-227.