OPEN ACCESS
https://doi.org/10.70905/bmcj.06.02.0547
Review Article

A REVIEW ON KIDNEY STONE DETECTION AND DIAGNOSIS USING CT KUB IMAGING: BRIDGING EMERGING TECHNOLOGIES AND CLINICAL PRACTICE

Khurshid Alam[1], Fakhre Alam[2], Asad Ullah[3]

[1]Department of Urology, Miangul Abdul Haq Jahanzeb Kidney Hospital, Swat, Pakistan.
[2]Department of Computer Science and IT, University of Malakand, Pakistan.
[3]Department of Computer Science and IT, University of Malakand, Pakistan.

Correspondence:
Dr. Fakhre Alam
Email: akhre.uom@gmail.com

Abstract

Background: Kidney stone disease is a common condition impacting a significant global population, with increasing incidence rates associated with lifestyle changes and dietary factors. Traditional diagnostic methods, including ultrasound and X-rays, have limitations in detecting specific stone types and sizes. CT KUB imaging, however, remains the gold standard for diagnosis.
Objective: This paper aims to evaluate the advancements in artificial intelligence (AI) methods in automating the detection, segmentation, and classification of kidney stones using CT KUB images. Additionally, it investigates the integration of multi-modal data, including patient clinical data, to enhance diagnostic accuracy and treatment personalization.
Material and Methods: The search of the structured literature in PubMed, Scopus, IEEE Xplore, and Google Scholar was conducted in terms of the keywords that included kidney stone, urolithiasis, AI, deep learning, and CT and were published in the period between January 2015 and March 2025. The inclusion criteria included peer-reviewed articles that used AI to detect or classify stones on CT KUB. The total number of records that were screened was 482, 58 full texts evaluated and 32 studies were included according to PRISMA guidelines.
Results: The reported diagnostic performances among AI models showed a range of sensitivities of 88 to 98, specificities of 85 to 97 and AUC of 0.85 to 0.96. Improved segmentation accuracy and diagnostic speed was observed to be constant in Convolution Neural Networks (CNNs), hybrid models and radiomics-based frameworks.
Conclusion: Despite the promising potential of AI in clinical settings, challenges such as data scarcity, model interpretability, and generalization remain. The paper proposes future research directions, particularly in combining explainable AI with federated learning technologies, aimed at seamless real-time clinical integration. Emerging technologies have the potential to revolutionize kidney stone management and urological treatment protocols.
Keywords: Kidney Stones, Artificial Intelligence, CT KUB, Deep Learning, Convolution Neural Networks, Detection, Diagnosis.
Cite this article: Alam K, Alam F, Ullah A, A Review on Kidney Stone Detection and Diagnosis Using CT KUB Imaging: Bridging Emerging Technologies and Clinical Practice . BMC J Med Sci. 2025; 6(2):94-103 https://doi.org/10.70905/bmcj.06.02.0547

Introduction:

The medical condition known as kidney stone disease or nephrolithiasis appears frequently in millions of global patients. Urine stone development in kidneys or bladder or ureters causes severe pain which leads to urinary tract infections alongside increasing healthcare expenses[1]. The global burden of disease study indicates that kidney stones affect 12% of the total global population as their numbers continually increase in nations of both developed and developing status[2]. Lifestyle elements including high salt and oxalate consumption, obesity, inadequate hydration together with genetic predispositions drive this upward uric acid trend. Kidney stones that recur and their associated complications create substantial medical care expenses for healthcare systems across the world.

The classification system for kidney stones includes calcium oxalate, uric acid, cystine and struvite components. Detecting and identifying these stones represents an essential step for doctors to create proper treatment

Authorship Contribution: [1,3]Substantial contributions to the conception or design of the work; or the acquisition, Data analysis, Literature review, [1]Drafting the work or revising it critically for important intellectual content, [3,4,6]Final approval of the version to be published, Topic Selection & Supervision

Funding Source: none
Conflict of Interest: none

Received:May,03, 2025
Accepted:Sept ,26, 2025
Published: Dec,30,12,2025

strategies. Ultrasound and plain X-ray methods which doctors commonly use for diagnosis present several limitations like poor sensitivity and incapability to identify specific stones while also requiring operator expertise[4].

Kidney stones receive their diagnosis through non-contrast CT KUB (Kidneys, Ureters, and Bladder) imaging because it shows excellent accuracy through both high sensitivity and specificity5. Manual CT image assessment demands prolonged attention with substantial workload and contains human mistakes when used in high-throughput clinical environments.

Figure 1 shows the process of stone nucleation, growth, and aggregation within renal tubules.

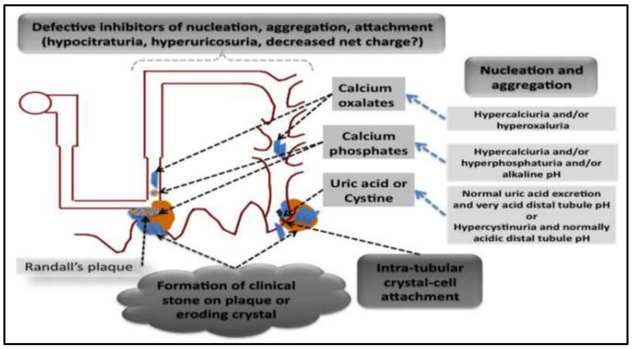

Figure 1. The process of stone nucleation, growth, and aggregation within renal tubules[3].

Due to the current emerging technologies such as AI and DL, medical imaging has transformed into an automated system for identification, categorization and segmentation of different pathological conditions6. The accuracy level of AI and DL models exceeds that of human experts when detecting kidney stones through image interpretation.AI systems shorten diagnosis time and decrease human mistakes while ensuring reliable and repeatable outcomes. Automated kidney stone segmentation and classification from CT KUB images has improved using Convolution Neural Networks (CNNs) and hybrid models thus improving diagnostic workflows[7].

The integration of AI and DL models between radiological data and patient history together with metabolic disorders and lifestyle factors enhances stone detection accuracy along with treatment outcome predictions. These technological breakthroughs have the capability to improve clinical choices and deliver targeted treatments that decrease kidney stone recurrence[8]. The implementation of artificial intelligence systems for detecting and categorizing stones should become significant for urological medical practice in the years ahead.

This review serves three essential purposes: (1) it consolidates contemporary AI and DL research about kidney stone detection alongside deep learning approaches and their clinical implementation effects in stone detection; and (2) To identify and evaluate the limitations and challenges within the field that affect data quality and model validation and clinical integration; and (3) to presents recommendations about future directions which combine multi-modal data sources while implementing explainable AI and federated learning to improve diagnosis of kidney stones.

Applications:

The practical applications of artificial intelligence in kidney stone assessment and diagnosis surpass conventional medical procedures to enhance both patient health outcomes and treatment processes. The main applications are:

Automated Detection and Segmentation: AI and DL models can automatically detect and segment kidney stones in CT KUB images while eliminating the need for manual labor which prevents human error in analysis[11]. Such models have the ability to find stones throughout their developmental process by identifying even tiny calcifications which traditional diagnostic methods cannot detect.

Stone Classification: AI and DL systems examine kidney stones to determine their dimensions as well as form and density levels and chemical properties to help doctors choose optimal treatment plans. The treatment of kidney stones depends on their type since calcium oxalate needs extracorporeal shock wave lithotripsy (ESWL) but uric acid requires medical dissolution therapy[12].

Predictive Modeling for Recurrence: AI and DL systems forecast the likelihood of stone reoccurrence through examination of patient information including personal data and clinical records and imaging scans[13]. Clinicians can use patient-specific information to deliver personalized preventive care that decreases stone formation possibilities.

Clinical Decision Support Systems (CDSS): Computer systems integrated within clinical workflows assist health professionals to reach faster accurate choices[14]. The system uses patient-specific data in combination with CT imaging information to suggest individualized treatment paths which enhances kidney stone disease management[15].

Real-Time Assistance in Surgery: AI and DL technologies helps surgeons perform more effective and shorter percutaneous nephrolithotomy (PCNL) procedures by providing instant stone detection assistance during the procedure[16].

Radiomics and Multi-Modal Data Fusion: The fusion of AI and DL systems with radiomics alongside multi-modal imaging technologies such as CT scan and ultrasound or MRI produces enhanced insights about kidney stone pathology which enhances diagnostic precision and promotes improved treatment decisions[17].

The medical field of kidney stone diagnosis stands to benefit from AI and DL technologies since they unlock opportunities to enhance diagnostic precision together with speedier testing and individualized therapeutic approaches. AI and medical imaging research development maintains a strong potential to form the future direction of urological clinical practice.

Material and Method:

The Preferred Reporting Items of Systematic Reviews and Meta-Analyses (PRISMA) guidelines have been used to conduct this review where necessary. Though it was based on the narrative review, structured search and selection procedures were underused to cover all possible areas and to be transparent.

Search Strategy:

Four electronic databases, Pub Med, Scopus, IEEE Xplore, and Google Scholar, were searched systematically in literature. The search included the studies published in the years 2015-17. The final search was conducted on March 15, 2025. The search terms were as follows used in different Boolean combinations:

("kidney stone" OR "renal calculus" OR "urolithiasis") AND ("artificial intelligence" OR "machine learning" OR "deep learning") AND ("CT" OR "computed tomography" OR "KUB").

Inclusion criteria:(a) English-language articles wherein peer-reviewed articles; (b) studies that use AI or DL to detect, segment, or identify kidney stones using CT or CT KUB imaging; (c) studies that provide quantitative diagnostic performance measures (e.g. sensitivity, specificity, AUC).

Exclusion criteria :(a) Non-AI research; (b) reviews, case and editorial studies; (c) research studies that do not utilize CT-based imaging; (d) non-English articles.

Study Selection Process:

The first search found records totaling 482, out of which 146 were removed due to duplication and 58 full-text articles were sifted through to determine their eligibility of inclusion into this review, 32 articles were included.

Data Extraction and Synthesis:

Each study was analyzed to extract relevant information, such as the AI technique, the dataset used, the sample size, the evaluation metrics (sensitivity, specificity, AUC), and important findings. The summarization of the results was done through an organized evidence table to aid in methodological and performance outcome comparisons.

Applications of Artificial Intelligence in Kidney Stone Detection and Classification:

Modern urology faces a complete diagnostic and decision-making transformation because of Artificial Intelligence (AI) through machine learning (ML) and deep learning (DL) applications. The processing capabilities of AI systems result in better performance when detecting kidney stones together with their segmenting and classifying stages and predicting treatment outcomes[18]. Research indicates that AI systems now match or surpass the accuracy rates of radiologists in CT-based stone detection operations while operating in real-time clinical environments.

Table 1: Summary of Key Studies on AI-Based Kidney Stone Detection and Diagnosis Using CT KUB.

Author (Year)	AI Technique / Model	Dataset Type / Size	Performance Metrics	Key Findings
Park et al. (2023)	CNN-based Detection Model	500 CT KUB scans	Sensitivity 95.2%, AUC 0.93	Achieved high real-time detection accuracy for small stones.
Li et al. (2022)	3D U-Net (Segmentation)	400 multi-slice CTs	DSC 0.89	Accurate volumetric stone segmentation.
Kim et al. (2023)	Deep-Learning Segmentation	250 CT cases	DSC 0.87	Automated delineation of urinary stones.
Akkaş et al. (2022)	Logistic Regression + ML	300 patients	AUC 0.91	Predicted need for surgical intervention
Isha & Shah (2023)	Radiomics + AI	200 CT scans	Accuracy 92%	Predicted stone composition effectively.
Abdimurotovich & Cho (2024)	Optimized YOLOv5	600 CT images	Sensitivity 96%, Specificity 94%	Fast real-time detection with high accuracy.
Sánchez et al. (2024)	ML Model (Random Forest)	1,000 patient data	AUC 0.90	Predicted urolithiasis risk from EHR data.
Deol & Kavoussi (2025)	Hybrid CNN + Radiomics	350 CT scans	AUC 0.94	Combined image and clinical features improved prediction.

AI Models for Kidney Stone Detection:

The localization and detection of renal calculus are performed mostly by Convolution Neural Networks (CNNs) that automatically extract hierarchical information by analyzing CT KUB images. They are networks that can examine patterns in images and spatial features to detect kidney stones with high precision with minimum handling of manual features. The systems based on CNN have demonstrated good results in separating stones and the surrounding tissues especially when the images are complex or in noisy environments.

A number of object detection architectures have been used successfully such as YOLO (You Only Look Once), Faster R-CNN, Efficient Net and Mask R-CNN. These models can locate the position and boundaries of the stones within the CT volumes at a high rate and with high sensitivity and specificity. As an example, Park *et al.*[19]have mentioned a CNN-based detecting system with a sensitivity of 95.2, and Abdomurotovich and Cho20have optimized the YOLOv5 model with a sensitivity of 96% and specificity of 94 %, which is able to track in real time.

Latest research has enhanced further the sensitivity of the models by incorporating attention mechanisms and multi-scale feature extractions, which enable networks to capture finer grained textures of small calculi and adjust to the changes in image quality. Also, it has been demonstrated that ensemble methods that utilize several CNN configurations can be more accurate in detecting objects and can minimize false negatives.

There have also been hybrid machine learning models, in which image features found by CNN are used with Support Vector Machines (SVMs) or Random Forest classifiers to fine tune predictions. These composite methods combine the merits of deep learning with the traditional machine learning, which leads to a better reliability of diagnoses and generalization on a variety of datasets.

Altogether, AI-based detection systems have a similar performance with the expert radiologists in detecting kidney stones on CT KUB images. Nevertheless, the majority of models are still in the proof-of-concept stage and there is little prospective validation and multi-centre testing. In order to support clinical translation, large scale dataset standardization, model explains ability, and real-time integration in Picture Archiving and Communication Systems (PACS) and Electronic Health Records (EHRs) should be prioritized, in the future.

Segmentation of Kidney Stones in CT KUB Images

The deep learning-based models, especially the U-Net, and its extensions, such as Residual U-Net (ResU-Net) and nnU-Net, have demonstrated outstanding performance in outlining the location of kidney stones in CT scans. These architectures are ideal in the case of biomedical image segmentation because of encoder-decoder architecture and skip connections, which enable the preservation of fine spatial data but also the context of the information across different image scales.

Li *et al.* [21] noted that a modified 3D U-Net with residual connections was employed to segment the renal calculi on multi-slice CT volumes. The model attained the Dice similarity coefficient (DSC) of 0.89, which implies that there was a high level of overlap between the segmentation produced by the model automatically and the expert manual annotations. The DSC metric is often applied in medical imaging, and it measures the accuracy of the segmentation, or how close the predicted areas are to the actual stone borders of the area.

In general, these deep learning models have dramatically increased the accuracy and reliability of stone detection and boundary definition relative to the human-based technique. Nevertheless, the performance of segmentation may still differ based on the image resolution, type of scanner, and thickness of slice, and density of stones. Future studies are to be aimed at creating more robust and multi-institutional models that would be able to retain high levels of segmentation accuracy in various CT acquisition protocols and patients.

Classification of Stone Type and Composition

The recent progress in the field of artificial intelligence (AI) has made it possible to create predictive models that utilize CT-based radiomics features such as texture pattern, Hounsfield Unit (HU) distribution, and three-dimensional morphological features to categorize the type of kidney stones with high accuracy. Such methods have been recognized to classify with over 90 percent highlights the high ability of the CT image based radio density profile and spectral features to discriminate.

Common machine learning approaches, including Support Vector Machines (SVMs) and Random Forests, have been extensively applied to discriminate between stone types based on the learning of quantitative imaging characteristics of the CT scan. But more recent models make use of deep learning models, especially the Convolution Neural Network (CNNs) which learn high-level imaging attributes that provide a more effective representation of stone composition and structure than handcrafted attributes.

Modern research has also combined the dual-energy CT scan to radiomics-based deep learning to improve the compositional analysis. These multi-based systems integrate both image derived attributes and pertinent clinical data of the patient like geographic, biochemical details and stone history into a composite diagnostic profile. Such combination of imaging and clinical data can be used to have a more complete assessment, which enhances the ability to classify and even personalized therapy planning.

As an example, Isha and Shah[22]showed that radiomics, with AI-based classifiers, yielded an overall accuracy of 92% in predicting stone composition, whereas Deol and Kavoussi23also were able to reach an AUC of 0.94 with hybrid CNN -radiomics models. The findings highlight the increased capabilities of multimodal systems based on AI to inform treatment choices- including the choice between medical dissolution therapy of uric acid stones and shock wave lithotripsy (ESWL) of calcium oxalate stones.

Although these advancements have been made, the majority of the existing models are constrained by dataset size, imaging heterogeneity, and unstandardized ground truth on the basis of chemical analysis. Future research must focus on the development of multi-centered datasets using confirmed compositional data to allow strong, generalizable AI models to be used in clinical deployment.

AI-Based Treatment Planning and Outcome Prediction

AI models are being actively developed to support personalized treatment planning. Predictive models can estimate the likelihood of spontaneous stone passage, ESWL success, or recurrence risk based on stone size, location, density, and patient factors24. For instance, Haifler, M. *et al*.[25] developed a machine learning model using gradient boosting and logistic regression to predict the need for surgical intervention in ureteral stone cases, with an AUC of 0.91. Other studies use recurrent neural networks (RNNs) and ensemble classifiers for outcome prediction post-URS or PCNL.

Integration of Multimodal Data and Future Directions

Emerging research combines imaging data with electronic health records (EHRs), laboratory data, and genetic profiles for more holistic risk modeling. These multimodal AI systems enhance decision-making, offering insights beyond what radiological imaging alone can provide. Additionally, explainable AI (XAI) frameworks are being introduced to increase clinician trust and regulatory compliance[26].AI-based mobile and web-based clinical decision support tools are also in development, potentially democratizing access to expertise in low-resource settings.

Artificial Intelligence (AI), particularly machine learning (ML) and deep learning (DL), is revolutionizing the landscape of urological imaging by enabling automated, accurate, and rapid detection, segmentation, and classification of kidney stones. With the rise of high-resolution CT imaging and large annotated datasets, state-of-the-art deep learning architectures are achieving near-radiologist level performance in detecting renal calculi, estimating composition, and supporting treatment planning. Table 2summarizes the comparative overview of AI Techniques for Kidney Stone Detection and Segmentation.

Table 2. A Comparative Overview of AI Techniques for Kidney Stone Detection and Segmentation.

Technique	Key Principles	Strengths	Weaknesses		Applications	Limitations
U-Net (2D/3D)	Encoder-decoder architecture for semantic segmentation	Accurate boundary delineation; good for small datasets	Sensitive to class imbalance; requires extensive annotation		2D/3D kidney and stone segmentation in CT KUB slices	Performance drops in low-contrast or noisy CTs
ResU-Net	U-Net with residual blocks	Better gradient flow; improved segmentation performance	Slightly higher computational cost		High-resolution segmentation of renal parenchyma and stones	Requires tuning for different scanners
Swin UNETR	Transformer-based hierarchical representation	Captures long-range dependencies; state-of-the-art 3D segmentation	High memory requirements; complex training		Multi-organ segmentation including stones in CT urography	Limited pre-trained models for urology datasets
YOLOv5 / YOLOv8	Real-time object detection via regression and classification	Fast inference; good localization accuracy	Less effective on very small or low-contrast stones		Real-time triage of CT or KUB images for stone detection	Needs large annotated bounding boxes; less precise for segmentation
Bi-LSTM	Bidirectional sequence learning to enhance spatial dependencies	Excellent classification accuracy; outperforms CNN/LSTM in temporal dependencies	High model complexity; less interpretable		Detection of presence/absence of stones from CT features (accuracy ~99%)	Does not localize stones; used only after pre-segmentation
CNN + Radiomics Hybrid	Combines DL feature maps with handcrafted radiomic features	High specificity; useful for stone type prediction	Complex pipeline; needs radiologist input for feature calibration		Stone composition classification; predicting treatment response	Hard to deploy in real-time systems
XGBoost + BOA (Butterfly	Ensemble classifier with metaheuristic-	Strong performance in structured/tabular data;	Not suitable for raw image processing		Feature-based classification from	Not end-to-end; depends on

Optimization Algorithm)	based feature selection	interpretable models	without feature engineering		preprocessed CT images	handcrafted or selected features
Multi-modal CNN (CT + clinical data)	Fuses image data with patient metadata (age, labs, history)	Enables personalized diagnostics; improves generalizability	Data fusion complexity; missing clinical records affect accuracy		Risk stratification; predicting stone recurrence or complications	Requires well-integrated hospital information systems
3D V-Net	Volumetric segmentation network using 3D convolutions	Effective for volumetric kidney and stone segmentation	Requires large GPU memory; sensitive to slice thickness variations		3D CT segmentation for surgical planning and volume estimation	Lower performance on highly anisotropic CT datasets
Federated CNN	Distributed learning without centralizing patient data	Enhances privacy and cross-institutional learning	Requires compatible infrastructure; communication overhead		Multi-hospital kidney stone detection and segmentation models	Difficult to implement without unified data standards

The extensive majority of published artificial intelligence models are still at the proof-of-concept phase although artificial intelligence has demonstrated impressive potential in automating kidney stone detection and classification. The majorities of the research papers are retrospective and performed on small and single-center samples and are not externally validated on different patient populations. Few studies have even gone to prospective clinical analysis and none of them have so far proven to be having great effect on patient outcomes, workflow effectiveness, or reduction of cost in normal urological clinical practice. Moreover, there has been no formal regulatory approval of an AI kidney stone detection system like FDA clearance or CE certification to use the system in clinical settings.

Such restrictions support the importance of multi-institutional studies, imaging standardization, and prospective trials aimed at not only assessing the quality of diagnosis but also at analyzing clinical usefulness, safety, and cost-effectiveness. Unless these measures are taken, kidney stone AI-based management will remain an instrument of research to-be instead of a clinical implementation tool.

Discussion:

Technical variation in Imaging parameters

The detection of kidney stones using AI models is very sensitive to scanner consistency, slice thickness, and CT acquisition parameters. Variations in scanning maker, reconstruction algorithms, and image resolution cause domain shift that may greatly contribute to model generalization across institutions. Low dose CT scans, although reducing radiation dose, tend to produce images that are noisy and therefore difficult to detect small stones. Equally contrast enhancement may blur the edges of stones and decrease the accuracy of segmentation. More standard imaging parameters should be used or cross-scanner/cross-protocol variations should be addressed by the use of domain adaptation methods to guarantee reproducibility in future AI studies.

Stone Composition

Even though many studies on AI claim a high classification accuracy, the chemical analysis of extracted stones is the most reliable method of determining the composition. An even smaller number of studies confirmed AI predictions with dual-energy or spectroscopic CT data, and still fewer related the results with laboratory analyzed stone composition. This absence of standard data of reference creates ambiguity in model assessment. It will be important to create large, annotated databases containing imaging and compositional analysis both to promote precision AI diagnostics in urology.

Clinical Impact: Diagnostic Efficiency and Patient Outcomes

Studies on AI-assisted detection systems have shown diagnostic accuracies of between 88-98 and are commonly comparable or outperform the performance of radiologists. The systems are able to significantly shorten the interpretation time especially in the emergency or high-volume environment and can be used to help minimise unnecessary interventions like repeat imaging or exploratory procedures. The AI tools may also be used in treatment planning, such as distinguishing between uric acid and calcium oxalate stones to prescribe either medical or procedural treatment. Nonetheless, their future clinical utility or quantifiable patient outcomes remain unproven and longitudinal validation studies are required.

Operational and Implementation Requirement

Although it is claimed to be an accurate tool, the practical implementation of AI models in the radiology department will have to overcome the operational and logistical obstacles. It must be integrated with Picture Archiving and Communication Systems (PACS) and Electronic Health Records (EHRs) so that the data flow may be seamless and real-time decision support is possible. Also, feasibility in clinical practice depends on the model inference time, hardware requirements and maintenance. Light and cloud-based architectures can counter the limitations of hardware but give a cause to think about data privacy and cyber security. Economic viability of implementing AI systems to have them as routine kidney stone evaluation is also required through cost-benefit analyses. Multidisciplinary working between clinicians, engineers and healthcare administrators will be essential in achieving success in translating to enable safety, efficiency and sustainability.

Regulatory and medico-legal expectations

Clinical translations of AI systems to kidney stone detection and diagnosis have to step in line with the current regulatory and medico-legal systems to promote patient safety, accountability, and transparency. The regulatory bodies of the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA) mandate that AI-based diagnostic systems must be properly validated, risk assessed and performance evaluated prior to being sanctioned to be used in clinical practice. Currently, there are no kidney stone-specific AI models that have been officially cleared in the United States (FDA) and Europe (CE) and this information highlights the fact that they are at the early stages of development.

To be approved by the regulators, AI tools should be shown not only to be highly diagnostic but also to have clinical benefit, reproducibility, and strength across a range of imaging centers and study groups. There is a growing trend among developers to deliver explainable AI (XAI) outputs such as visualizations or explanatory decision paths to improve clinician trust and audit needs. Ongoing performance checks and post-implementation surveillance will also prevail according to the changing Software as a Medical Device (SaMD) standard.

Arguing in terms of medico-legal issues, the questions of data privacy, bias in algorithms, and liability are at the heart of it. Patient imaging data has to be used in accordance with privacy laws like GDPR in Europe and HIPAA in the United States to train models. Furthermore, the issue of diagnostic error either caused by the AI system or a clinician or the institution is a complicated ethical and legal problem. The implementation of AI tools in institutions should have elaborate accountability structures, protocols on informed consent and ongoing training of users to ensure risk is minimized.

Limitations of the Review

Although this review is a thorough synthesis of the new breakthroughs in artificial intelligence to identify kidney stones and diagnose them by using CT KUB images, it is important to consider a number of shortcomings. To begin with, it is not a systematic review or a meta-analysis, it is a narrative review; hence quantitative comparisons and pooled performance estimates were not conducted. The selection of the study was, despite its organization and the use of PRISMA principles, the possible source of the selection bias. Second, it used only English-language publications and, thus, this may have filtered out any related studies that were published in other languages. Third, the reviewed studies were heterogeneous in the terms of AI model structures, dataset sizes, measures of evaluations, and validation procedures, and did not allow straightforward comparison of results. Moreover, the reviews of the studies did not provide a comprehensive report on the training data composition, external validation, and clinical deployment, which limited the capacity to evaluate the generalizability and practical applicability.

Finally, there is no standardized benchmark of kidney stone imaging datasets, which is one of the obstacles to reproducibility. The work in the future should focus on the solutions to these problems by conducting systematic meta-analyses, creating multi-center open datasets, and defining standard evaluation criteria that would offer more powerful, evidence-based clinical translation advice.

Challenges and Future Directions

Despite significant advances, several persistent challenges hinder the widespread clinical integration of AI in kidney stone detection and diagnosis.

Data Scarcity and Annotation Burden

Most AI models require large-scale, annotated datasets for robust performance. However, publicly available CT KUB datasets annotated for kidney stones are scarce, and manual annotation by expert radiologists is time-consuming and expensive27. Synthetic data generation using generative adversarial networks (GANs), data augmentation, and semi-supervised learning approaches have been proposed to mitigate this limitation.

Lack of Standardization and Protocol Variability

Differences in CT acquisition parameters (slice thickness, scanner model, contrast settings) and annotation protocols introduce domain shifts, reducing model generalizability across institutions. Domain adaptation, normalization pipelines, and harmonization of imaging protocols are needed to overcome these discrepancies.

Model Interpretability and Clinical Trust

Deep learning models often function as "black boxes," making it difficult for clinicians to understand how predictions are derived. Explainable AI (XAI) tools like Grad-CAM, SHAP, and LIME have been introduced to visualize model decisions and improve clinician trust[28].

Ethical, Legal, and Regulatory Barriers

Ethical issues such as data privacy, bias, and the opaque nature of algorithmic decisions present barriers to clinical adoption. In addition, regulatory frameworks (e.g., FDA, CE marking) require rigorous validation through multicenter trials to ensure safety, effectiveness, and transparency[29].

Integration with Clinical Workflow

Most AI systems remain standalone prototypes. For clinical impact, seamless integration with PACS systems, EHRs, and real-time decision support is essential. Interdisciplinary collaboration between radiologists, computer scientists, and software engineers is needed to ensure practical deployment and usability[30].

Future Research Directions

Future research should focus on developing federated learning systems, exploring transformer-based architectures, designing lightweight models, combining multimodal data and establishing benchmark datasets. Developing federated learning systems for training models across institutions without sharing data. Exploring transformer-based architectures for better spatial-contextual understanding. Designing lightweight models for deployment on mobile devices or point-of-care systems. Combining multimodal data (e.g., lab reports, clinical symptoms) with imaging for holistic diagnosis. Establishing benchmark datasets and shared challenges (e.g., MICCAI Kidney Stone Challenge) to accelerate reproducibility and innovation. Table 4 shows summary of technical, clinical, and regulatory challenges in AI-based kidney stone detection and segmentation, along with strategic approaches tailored for clinical deployment and research advancement.

Table 3.Comprehensive summary of technical, clinical, and regulatory challenges in AI-based kidney stone detection and segmentation, along with strategic approaches tailored for clinical deployment and research advancement.

Challenge	Description	Technical Implications	Strategic Solutions	Domain-Specific Notes
Limited Annotated Datasets	Scarcity of high-quality, annotated CT KUB datasets	Limits model generalization; leads to over fitting	Use federated learning, semi-supervised learning, or synthetic data augmentation	Curated, multi-institutional urology image datasets are urgently needed
Data Heterogeneity	Variation in image quality, scanner types, protocols	Causes domain shift and inconsistent performance across sites	Apply domain adaptation and transfer learning; use harmonization algorithms	Vital for generalizing models across rural and urban hospitals
Class Imbalance	Under-representation of ureters/stones in segmentation tasks	Affects training stability; bias toward majority classes	Use weighted loss functions (e.g., Dice+CE), oversampling, or GAN-based data generation	Especially critical in ureter segmentation due to anatomical sparsity
Lack of Interpretability	Deep models act as "black boxes" with limited clinical trust	Clinicians hesitant to adopt models without transparent reasoning	Use Explainable AI (XAI): Grad-CAM, SHAP, or attention-based models	Needed for compliance with clinical audit and medico-legal standards
Integration with Clinical Workflow	Standalone AI systems are rarely compatible with hospital PACS/EMR systems	Limits clinical adoption despite high technical performance	Develop API-compatible and DICOM-friendly AI interfaces	Integration with radiologist workstations (e.g., GE, Philips) is critical
Model Over fitting	High training accuracy but poor generalization	Due to small datasets, excessive model complexity	Apply regularization, dropout, data augmentation, and cross-validation	Particularly common in deep 3D CNN models trained on limited CT volumes
Computational Requirements	3D segmentation and ensemble models require large memory and time	Inaccessible in low-resource or rural clinical settings	Use lightweight models (e.g., Mobile Net), quantization, or cloud-based inference	Important for deployment in developing countries or portable screening units
Regulatory and Ethical Concerns	Lack of validated benchmarks and regulatory clarity (e.g., FDA, MDR compliance)	Slows clinical deployment of AI systems	Perform multicenter clinical trials; align with FDA/CE evidence frameworks	Future tools must demonstrate safety, fairness, and accountability
Privacy and Security	Centralized training risks patient data exposure	Violates GDPR/HIPAA; impedes cross-institutional model sharing	Implement federated learning, homomorphism encryption, and secure model exchange protocols	Especially important when training across national or cross-border urology registries
Prediction and Personalization Gaps	Current models focus on detection, not on recurrence or personalized guidance	Missed opportunities for preventive and longitudinal care	Combine imaging with EHR/lab/genomic data using multi-modal AI models	Needed to transition AI from reactive diagnostics to proactive urological care

Conclusion:

Artificial Intelligence and deep learning has emerged as a transformative force in the field of medical imaging, particularly for the detection and diagnosis of kidney stones using CT KUB imaging. This review has systematically examined the recent advancements in AI methodologies, highlighting their diagnostic accuracy, automation capabilities, and clinical relevance. From classical machine learning algorithms to cutting-edge deep learning frameworks and segmentation architectures, AI and deep learning has demonstrated remarkable potential in automating stone detection, characterizing size and composition, and supporting radiologists with reliable decision support. However, despite the promise, several critical challenges persist ranging from data scarcity, standardization issues, and interpretability gaps, to regulatory and ethical concerns. The constraints placed by insufficiently annotated multi-center datasets together with inconsistent imaging protocols slow down the creation of effective generalizable models. Model deployment into real-world settings faces delays because of insufficient clarity about the models and implementation difficulties with clinical teams. Moving forward in research requires three main aspects: federated learning techniques, explainable AI systems and multimodal fusion strategies together with lightweight edge architectures and clinical workflow integration. Establishing shared benchmarking infrastructure and following regulatory guidelines through multidisciplinary cooperation will make it possible to deploy AI systems safely in routine radiological and nephrological care. The use of AI and deep learning based kidney stone diagnosis through CT KUB imaging shows potential to mature although it remains incompletely developed. HealthCare innovation enhanced by clinical testing alongside moral design elements will create future intelligent diagnostic tools specific to urology.

References:

1. Fard AM, Fard MM. Evaluation of office stones in kidney patients and how to form and treat them. Eurasian J Sci Tech. 2021;2(2):384-98.
2. Zhang L, Zhang X, Pu Y, Zhang Y, Fan J. Global, regional, and national burden of urolithiasis from 1990 to 2019: a systematic analysis for the global burden of disease study 2019. Clinical Epidemiology. 2022 Dec 31:971-83.
3. Ratkalkar VN, Kleinman JG. Mechanisms of stone formation. Clinical reviews in bone and mineral metabolism. 2011 Dec;9(3):187-97.
4. Buri S, Shrivastava V. A Brief Review of Kidney Stone Detection and Prediction Techniques. InInternational Conference on Communication and Computational Technologies 2023 Jan 28 (pp. 529-543). Singapore: Springer Nature Singapore.
5. Mahargias H, Gunawan FM. Analysis Study of Diagnostic Imaging Performance and Accuracy of Kidney Stone: A Comprehensive Systematic Review. The International Journal of Medical Science and Health Research. 2025 Feb 22;9(5):1-28.
6. Barua R. An Investigation of AI Techniques for Detecting Kidney Stones in CT Scan Images Through Advanced Image Processing. InEnhancing Medical Imaging with Emerging Technologies 2024 (pp. 133-150). IGI Global Scientific Publishing.
7. Ahmed F, Abbas S, Athar A, Shahzad T, Khan WA, Alharbi M, Khan MA, Ahmed A. Identification of kidney stones in KUB X-ray images using VGG16 empowered with explainable artificial intelligence. Scientific Reports. 2024 Mar 14;14(1):6173.
8. Papatsoris A, Alba AB, Galán Llopis JA, Musafer MA, Alameedee M, Ather H, Caballero-Romeu JP, Costa-Bauzá A, Dellis A, El Howairis M, Gambaro G. Management of urinary stones: state of the art and future perspectives by experts in stone disease. Archivio italiano di urologia, andrologia: organo ufficiale [di] Societa italiana di ecografia urologica e nefrologica. 2024 Jun 27;96(2):12703-.
9. Montatore M, Muscatella G, Eusebi L, Masino F, Gifuni R, Giannubilo W, Sortino G, Guglielmi G. Current status on new technique and protocol in urinary stone disease. Current Radiology Reports. 2023 Dec;11(12):161-76.
10. Taherdoost H. Deep learning and neural networks: Decision-making implications. Symmetry. 2023 Sep 8;15(9):1723.'
11. Chowdhury AT, Salam A, Naznine M, Abdalla DA, Erdman L, Chowdhury ME, Abbas TO. Artificial intelligence tools in pediatric urology: a comprehensive review of recent advances. Diagnostics. 2024 Sep 17;14(18):2059.
12. Isha S, Shah SZ. Use of artificial intelligence for analyzing kidney stone composition: are we there yet?. Mayo Clinic Proceedings: Digital Health. 2023 Sep 1;1(3):352-6.
13. Sánchez C, Larenas F, Arroyave JS, Connors C, Giménez B, Palese MA, Fulla J. Artificial intelligence in urology: application of a machine learning model to predict the risk of urolithiasis in a general population. Journal of Endourology. 2024 Aug 1;38(8):712-8.
14. Khalifa M, Albadawy M, Iqbal U. RETRACTED: Advancing clinical decision support: The role of artificial intelligence across six domains.
15. Deol ES, Kavoussi NL. Artificial intelligence applications in kidney stone disease. InArtificial Intelligence in Urology 2025 Jan 1 (pp. 169-191). Academic Press.
16. Gauhar V, Giulioni C, Gadzhiev N, De Stefano V, Teoh JY, Tiong HY, Taguchi K, Milanese G, Galosi AB, Somani BK, Castellani D. An update of in vivo application of artificial intelligence and robotics for percutaneous nephrolithotripsy: results from a systematic review. Current Urology Reports. 2023 Jun;24(6):271-80.
17. Zhang M, Ye Z, Yuan E, Lv X, Zhang Y, Tan Y, Xia C, Tang J, Huang J, Li Z. Imaging-based deep learning in kidney diseases: recent progress and future prospects. Insights into imaging. 2024 Feb 16;15(1):50.
18. Fuladi S, Chaturvedi H, Nallakaruppan MK, Grover V, Alshahrani H, Baza M. Efficient approach for kidney stone treatment using convolutional neural network. Traitement du Signal. 2024 Apr 1;41(2):929.
19. Kim YI, Song SH, Park J, Youn HJ, Kweon J, Park HK. Deep-learning segmentation of urinary stones in noncontrast

computed tomography. Journal of Endourology. 2023 May 1;37(5):595-606.

20. Abdimurotovich KA, Cho YI. Optimized YOLOv5 architecture for superior kidney stone detection in CT scans. Electronics. 2024 Nov 11;13(22):4418.

21. Li D, Xiao C, Liu Y, Chen Z, Hassan H, Su L, Liu J, Li H, Xie W, Zhong W, Huang B. Deep segmentation networks for segmenting kidneys and detecting kidney stones in unenhanced abdominal CT images. Diagnostics. 2022 Jul 23;12(8):1788.

22. Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, Lu L, Yuille AL, Zhou Y. Transunet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306. 2021 Feb 8.

23. PRASUNA G, MURTHY V, REDDY VR, PATHURI SK. ENHANCING KIDNEY STONE DETECTION: INTEGRATIVE ANALYSIS OF URINE ATTRIBUTES AND MEDICAL IMAGING. Journal of Theoretical and Applied Information Technology. 2025 Feb 28;103(4).

24. Akkaş F, Culha MG, Ayten A, Danacıoğlu YO, Yildiz Ö, İnci E, Guner E, Şahin S. A novel model using computed tomography parameters to predict shock wave lithotripsy success in ureteral stones at different locations. Actas Urológicas Españolas (English Edition). 2022 Mar 1;46(2):114-21.

25. Haifler M, Kleinmann N, Haramaty R, Zilberman DE. A machine learning model for predicting surgical intervention in renal colic due to ureteral stone (s)< 5 mm. Scientific Reports. 2022 Jul 11;12(1):11788.

26. Preedanan W, Suzuki K, Kondo T, Kobayashi M, Tanaka H, Ishioka J, Matsuoka Y, Fujii Y, Kumazawa I. Improvement of urinary stone segmentation using GAN-based urinary stones inpainting augmentation. IEEE Access. 2022 Oct 31;10:115131-42.

27. Alghamdi H, Amoudi G. Using machine learning for non-invasive detection of kidney stones based on laboratory test results: a case study from a saudi arabian hospital. Diagnostics. 2024 Jun 25;14(13):1343.

28. Sadeghi Z, Alizadehsani R, Cifci MA, Kausar S, Rehman R, Mahanta P, Bora PK, Almasri A, Alkhawaldeh RS, Hussain S, Alatas B. A review of Explainable Artificial Intelligence in healthcare. Computers and Electrical Engineering. 2024 Aug 1;118:109370.

29. Pozdzik A, Grillo V, Sakhaee K. Gaps in kidney stone disease management: From clinical theory to patient reality. Urolithiasis. 2024 Apr 9;52(1):61.

30. Wada C, Laviana AA, Woodle T, Patel P, Lerman SE, Blumberg J, Nabhani J, Waterman B, Bergman J. Integrated Care in a Large Public System Improves Simple Kidney Stone Management. Urology Practice. 2025 Jan;12(1):35-42.