I.J. Image, Graphics and Signal Processing, **2026, 4, 139-164**
Published Online on August 8, 2026 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2026.04.08

Kidney Stone Detection Using Medical Imaging and Computational Intelligence - A Review

Karthick P.
Vellore Institute of Technology, Vellore, India
E-mail: karthick.p2024@vitstudent.ac.in
ORCID iD: https://orcid.org/0009-0004-5365-7337

Chiranji Lal Chowdhary*
Vellore Institute of Technology, Vellore, India
Email: chiranji.lal@vit.ac.in
ORCID iD: https://orcid.org/0000-0002-5476-1468
*Corresponding author

Received: August 12, 2025; Revised: December 22, 2025; Accepted: April 20, 2026; Published: August 8, 2026

Abstract: Nephrolithiasis (kidney stone disease) is a common urological disease that has a high clinical and economic impact. The early diagnosis is needed to avoid complications like obstruction of the ureter, infection, impaired kidney functioning. Traditional imaging modalities, such as ultrasonography, kidney-ureter-bladder radiography, and non-contrast computed tomography, are common but have a number of limitations, specifically their operator dependence, radiation, and low sensitivity to small or radiolucent stones. This review follows a PRISMA-based methodology to conduct a systematic review of studies published between 2015 and 2025 on the topic of computational intelligence methods such as artificial intelligence, machine learning, and deep learning to detect kidney stones based on medical images. Major scientific databases were considered in studies according to imaging modality, preprocessing method, model architecture and performance measures. Deep learning models, especially, Convolutional Neural Networks and U-Net-based frameworks, are highly effective in detection and segmentation tasks and have been reported to have accuracy of 86 to 99.9 percent, Dice coefficients over 0.85 and AUC of up to 0.99 in controlled data. Hybridization to combine ML classifiers, including Support Vector Machines, further improves the performance of classification. Yet, these outcomes are commonly limited through small datasets, class imbalance, external validation, and overfitting, which have an impact on real-life generalization. The use of computational intelligence has greatly improved the detection of kidney stones by enhancing automation, precision, and reproducibility. However, there are still major issues, such as the standardization of the dataset, interpretability of the models, and limitations to the clinical implementation. Explainable AI, federated learning, and 3D volumetric analysis should be prioritized in future research to create diagnostic systems.

Index Terms: Kidney Stones Detection, Nephrolithiasis, Medical Imaging, Deep Learning, Convolutional Neural Networks, U-Net, Computational Intelligence, Segmentation, Artificial Intelligence.

1. Introduction

1.1. Importance of Early and Accurate Kidney Stone Detection

Kidneys are very important in ensuring metabolic balance through filtering of waste products, fluid and electrolyte homeostasis. Kidney stone disease (nephrolithiasis or urolithiasis) is a clinically important disease, which is associated with the development of crystalline deposits in the urinary tract [1,2]. These deposits are of varying size and composition that may present with symptoms, including hematuria, excruciating flank pain, urinary obstruction, and more severely, renal failure, and urosepsis. Therefore, the clinical burden of the disease requires that it is diagnosed promptly and accurately to avoid unnecessary health risks [3, 4]. The impact of a delayed or misdiagnosis is huge. Unidentified stones may result in remaining infection and progressive decline of the renal function and hydronephrosis [5]. Patients may experience repeated visits to the emergency room, have to endure chronic pain, or be treated unnecessarily because of misdiagnosed or not well localized stones. Consequently, there is a strong clinical justification for the advancement and incorporation of sophisticated diagnostic models which provide high sensitivity and specificity. Speeding up, standardizing and optimizing the reliability of diagnostic protocols represents the primary goal for nephrological and urological imaging [6, 7].

1.2. Prevalence and Burden of Kidney Stone Disease

Over the past three decades, kidney stone disease has increased globally, with 10-15% of the population affected. Recurrence rates are high, with half of depressive episodes requiring a second within five to 10 years. Risk factors include dietary, genetic, occupational, and environmental factors [11]. Recent clinical observations show a high prevalence of 7%-13% in India, particularly in the northern states due to climatic and groundwater variations [12]. It is also increasing in urban areas due to sedentary life, low water consumption, high intake of animal protein, processed food or salt ones. In addition, more younger men are being afflicted, in contrast to previous patterns in which attacks were most common in middle-aged men [7]. Countries like India with the status of developing countries have an added barrier in terms of disparity in healthcare infrastructure, access to diagnostics and affordability of advanced imaging modalities. Therefore, novel and cost-effective diagnostic methodologies are desperately needed to stem the increasing burden of this disease worldwide [15].

1.3. Limitations of Conventional Diagnostic Methods

Ultrasound (US), non-contrast computed tomography (NCCT), and plain radiography (KUB X-rays) are all common ways to check for kidney stones. But these approaches have some problems that make them less sensitive and specific. KUB X-rays are widely available and not too costly, but they don’t pick up small or radiolucent stones very well, and they don’t see stones that are on top of bone or in places that are difficult to access very well [17]. Ultrasound, which is commonly used for primary assessment as it is safe and non-invasive, can detect larger calculus in the kidneys or bladder. But it performs poorly in detecting ureteral stones, particularly < 5 mm ones, and is operator dependent. It is less accurate in obese patients, in the presence of bowel gas, and for deep anatomic structures [14, 15]. The gold standard for urinary tract stones detection is non-contrast CT (NCCT), which has a specificity and sensitivity of nearly 100%. However, its potential is limited, particularly for young patients and routine follow-up, by its high cost, limited availability within rural areas, and associated radiation burden. The radiation dose is additionally raised by repeated exposure to stone a future occurrence [15, 17]. Computational intelligence in the detection of kidney stones is achieved through a number of steps, which determine the process properly.

1.4. Rise of Computational Intelligence and Imaging in Healthcare

Computational Intelligence (CI) refers to a collection of advanced computational paradigms inspired by biological and social systems such as: artificial neural networks (ANNs), fuzzy logic, support vector machines (SVMs), genetic algorithms and, more recently, deep learning (DL) [6]. In the field of health care, CI applications are being widely introduced in medical diagnosis, image analysis, disease forecast and decision support. It’s capability to handle high-dimensional, complex data and to identify non-linear patterns, makes it suitable for medical purposes [15]. The applications of CI are not merely limited to detection. Higher-end models are able to stratify stones according to composition (such as calcium oxalate, uric acid, cystine) or infer likely recurrence based on historical imaging findings and patient metadata, recommending personally tailored treatment courses [7]. It doesn’t learn like the static algorithms because the more data it gets, the better it becomes, it is capable of learning and updating [8].Another important

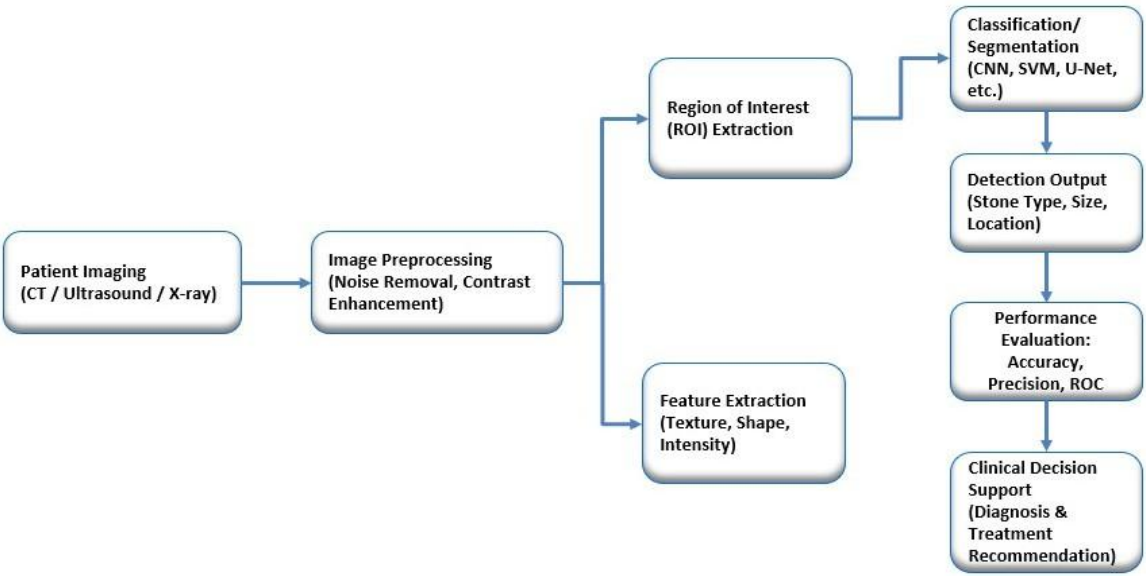

Fig. 1. General workflow of kidney stone detection using medical imaging and computational intelligence.

advantage is the possibility of real time image analysis. In an emergency department environment, automated tools can quickly flag obstructive stones or propose an urgency level that can result in immediate actions [21]. In addition, these techniques can be integrated to PACS (Picture Archiving and Communication Systems), providing an unobtrusive approach to radiological practices. Indeed, the democratization of AI implies AI systems can be deployed in constrained environments, such as from the cloud or on the edge [19, 24]. The use of AI-assisted ultrasound, for example, has the potential to overcome operator dependence automatically detecting and annotating suspected stones and may ultimately facilitate more accessible and equitable diagnoses. Fig. 1 shows the overall process behind kidney stone detection, such as acquiring images, preprocessing images, feature extraction, model training, and integrating the models with clinical applications, with the computational intelligence being applicable in all these areas.

This figure emphasizes the chronological interaction among preprocessing, feature extraction, and model development, in which mistakes in initial stages can have a profound impact on general detection performance.

1.5. Objectives and Scope of the Review

This review thoroughly presents recent developments of computational intelligence in kidney stone detection, emphasizes the main challenges, and suggests feasible future directions for both research and clinical applications. The key aims of this review are:

1. The paper provides a comprehensive overview of kidney stone disease and includes the clinical background, types, symptoms, risk factors and limitations in current diagnostic testing.
2. This review examines the medical imaging modalities of CT, ultrasound, X-ray and MRI, as well as preprocessing methods that can improve the quality of the image and its performance in detection.
3. The paper offers a detailed analysis of machine learning, deep learning, and hybrid methods of automated kidney stone detection and classification.
4. This review includes feature extraction, selection and segmentation techniques (existing along with advanced and deep learning-based approaches) for accurate localizing stones.
5. Finally, we address the challenges in datasets, evaluation metrics and upcoming research directions such as explainable AI, federated learning and embedding intelligent systems into clinical practice.

Despite these objectives, several critical research gaps remain, which are discussed in the following section.

1.6. Research Gap and Contributions

Although remarkable progress has been made in the computational intelligence performed kidney stone detection, several critical challenges remain still unresolved. First, the existing models learnt are not general due to the absence of large-scale annotated datasets and standardized protocols. Second, numerous studies report strong performance without external validation which raises the concern of overfitting and clinical validity. Third, existing attention paid to model interpretability remains limited despite being critical for clinical adoption. Moreover, literature so far concentrates only on single-item models and lacks a coherent comparative short-hand across imaging modalities as well as algorithmic approaches. This review makes a contribution to address these deficits in the following ways.

1. This work presents a systematic comparative analysis of machine learning and deep learning models on different imaging modalities.
2. It provides an in-depth assessment of published performance indicators with respect to dataset size, bias and generalizability.
3. The issue highlights major obstacles and future avenues of research, including explainable AI, federated learning, and 3D volumetric analysis.

2. Review Methodology

2.1. Study Design and Search Strategy

This review is undertaken based on Preferred Reporting Items of Systematic Reviews and Meta-Analyses (PRISMA) in order to ensure methodological rigor, transparency, and reproducibility. Using PRISMA, it is feasible to find, filter, and choose the pertinent research in the field of kidney stone detection using computational intelligence. This structure will ensure that the review makes its selection process objective and well documented thus, enhancing the credibility and reliability of the results. A comprehensive and systematic literature search was performed in a number of well-known scientific databases, including IEEE Xplore, PubMed, Scopus, and ScienceDirect. These databases have been selected in such a way that there is a broad coverage of engineering and medical research areas. The search included only the publications published in 2015-2025 as it was the latest progress in the field of machine learning and deep learning techniques applied to medical imaging. A combination of keywords and Boolean operators was used to retrieve relevant studies. The search terms were variation of kidney stones detection, nephrolithiasis imaging, deep learning, machine learning, artificial intelligence, CT imaging, ultrasound and medical image segmentation. This was done to observe extensive clinical and computer coverage of the problem.

2.2. Inclusion and Exclusion Criteria

To guarantee that the quality and relevance of the selected works were maintained, the inclusion and exclusion criteria were formulated. The inclusion criteria were that they must be dealing with machine learning or deep learning algorithms used to detect kidney stones and must be imaging modalities, such as CT, ultrasound, or X-ray. Only peer-reviewed journal articles and conference papers in English were taken into account. Only non-imaging studies, insufficient methodological information or experimental validation, and review articles without original technical contribution were filtered out of the studies. Duplicates and incomplete records were also done away with during the screening process. These criteria ensured that only good and relevant studies were involved in the analysis.

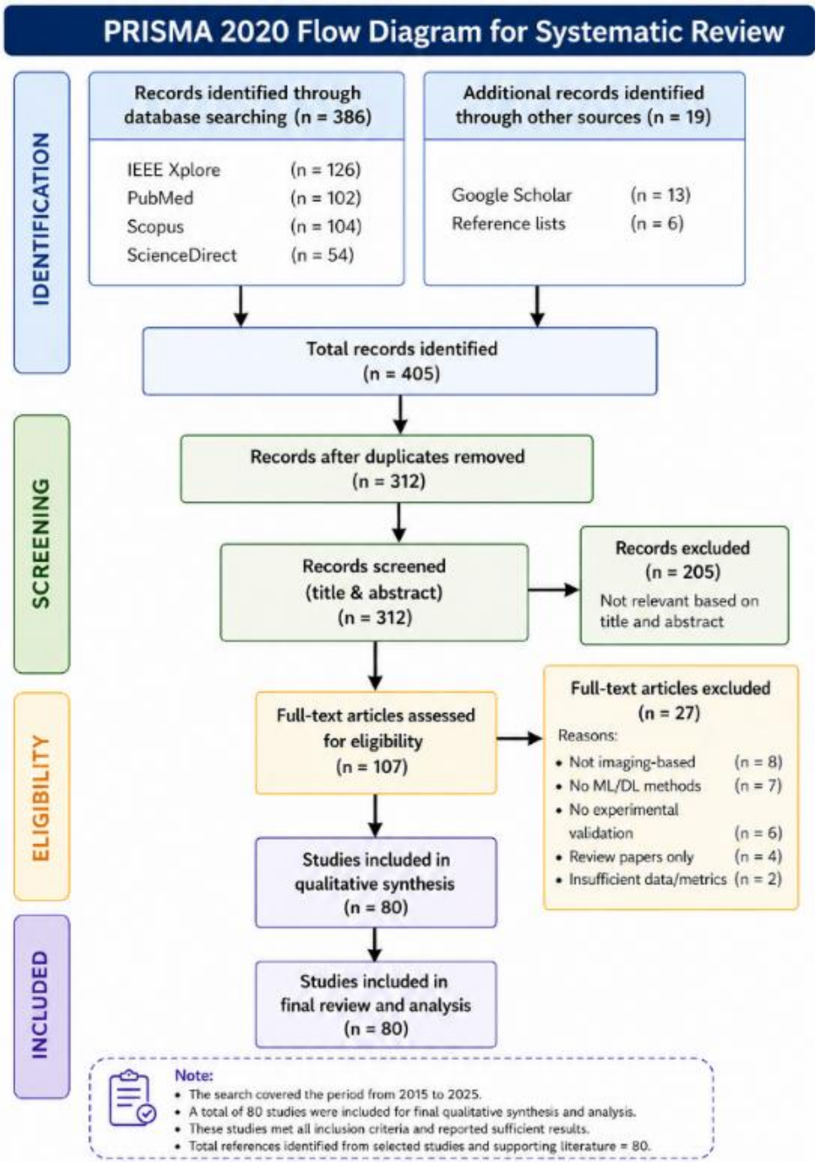

Fig. 2. Prisma flow diagram.

2.3. Study Selection Process

The study selection process was grounded on the PRISMA workflow that involves four steps of study selection including identification, screening, eligibility, and inclusion. To begin with, a considerable number of articles was retrieved via the selected databases. Duplicates were filtered and the remaining studies were filtered using their titles and abstracts. After, full-text articles were considered, in terms of eligibility, based on the set inclusion and exclusion criteria. Finally, a sample of research was selected to be studied in a qualitative manner. Finally, a set of studies was selected for qualitative analysis. The data that is relevant to the study was systematically extracted in both selected studies to enable them be compared. It was comprised of dataset properties (size and type), imaging modality, model architecture (e.g., CNN, U-Net, SVM), evaluation measures (accuracy, Dice coefficient, IoU, and sensitivity) and the most valuable findings. In addition, any constraints and potential biases of each study were documented.

This type of systematic data mining enabled a comprehensive comparison of different approaches and the detection of trends and gaps in the studies. A total of 405 records were identified with the help of database search and other sources. Following the elimination of 93 duplicates, 312 studies have been left to be screened in terms of titles and abstracts. According to the screening process, the number of full-text articles to be considered with regard to their eligibility was 107, 27 of these articles were excluded in accordance with the predefined criteria. Finally, 80 articles were involved in the qualitative analysis. Fig. 2 shows the study selection process.

2.4. Risk of Bias Analysis

Each of the studies included in critically evaluated on the basis of possible bias. The size of the dataset, its diversity and representativeness, and the validation methodology applied (e.g., cross-validation or external validation) were considered in the analysis. Of specific concern was the risk of overfitting, especially in studies which assert very high accuracy based on small datasets. They observed that many studies cannot be externalized and use small datasets, which are single-centres, which may lead to bias estimates of performance and decrease the possibility of generalization. This disadvantage highlights the need of standard datasets and multi-Centre validation in the future. This is a systematic methodology that offers transparency and reproducibility of the review procedure. To simplify the process of selecting the pipeline, a PRISMA flow diagram is given, which helps in visualizing the process.

3. Proposed Taxonomy of AI-Assisted Detection of Kidney Stones

To close this gap of lack of a single framework in the extant literature [71], this review provides an in-depth taxonomy of kidney stones detection on the basis of application of computational intelligence. This taxonomy systematizes the entire diagnostic pipeline into stages enabling a systematic comparison and analysis of alternative approaches. It also provides a systematic perspective that takes into consideration technical, clinical and implementation aspects of AI-based systems. This section gives the overview of the proposed framework. The proposed taxonomy divides the kidney stone detection systems into six big steps: image acquisition, pre-processing, feature extraction, model development, output generation, and clinical integration. All the stages are rather significant components of the whole pipeline and it is one of the aspects of the efficiency and effectiveness of the diagnostic system. Fig. 3 depicts the proposed taxonomy which offers an all-inclusive end-to-end architecture that combines the technical and clinical sides as compared to the current methods which deal with isolated components only.

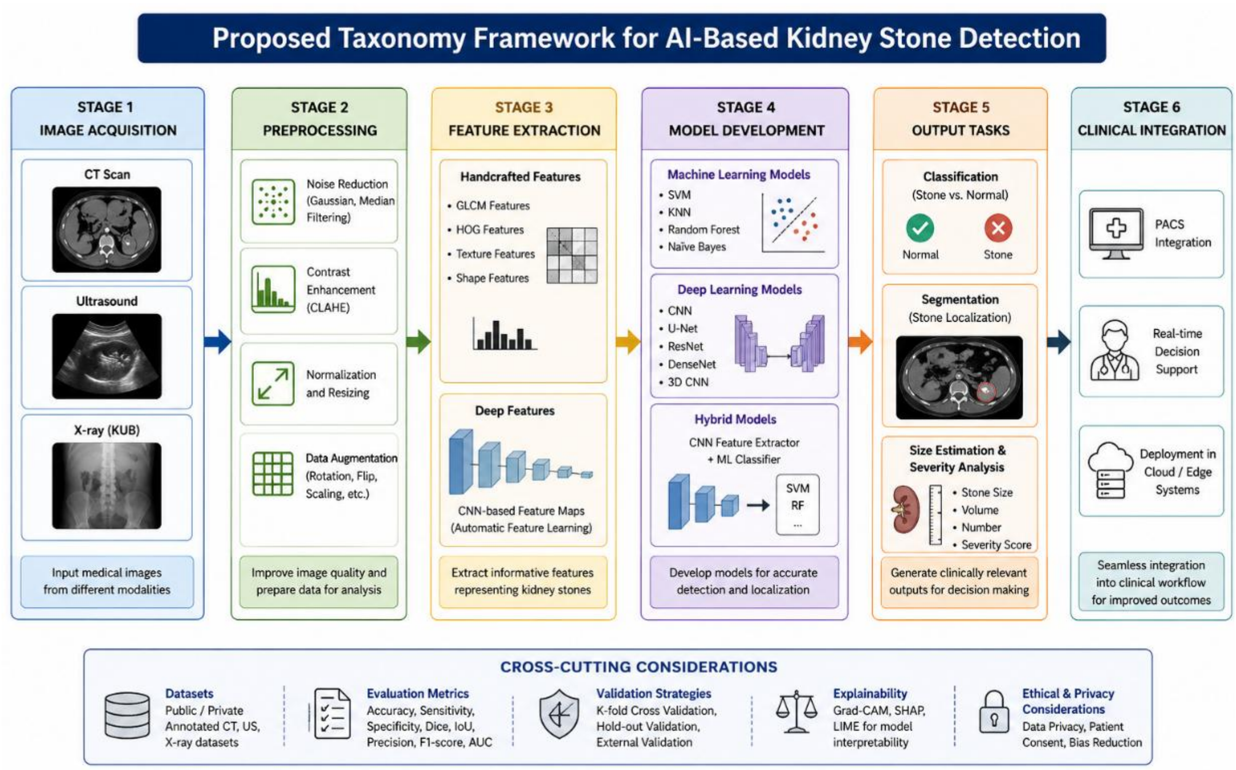

Fig. 3. Proposed taxonomy for AI-based kidney stone detection.

3.1. Stage-wise Description

The structured pipeline underlines the significance of the fact that most of the existing studies focus on the development of models yet minimal attention is given to clinical integration and deployment, which implies that the knowledge gap in the study is acute. The medical imaging is obtained during the image acquisition phase through the assistance of CT scans, ultrasound and X-ray (KUB) [73]. Modality influence detection of accuracy and clinical relevance to a great extent. The preprocessing stage includes noise reduction, contrast enhancement (e.g. CLAHE) and normalization of the image, to improve image quality and prepare data to be analyzed. This plays a very important role in order to enhance the visibility of features and reduce artifacts. In the feature extraction stage, handcrafted features (GLCM, HOG) and deep learning-based features are utilized. Deep learning models are able to automatically learn hierarchical features on raw data in comparison to the traditional methods, which rely on manually constructed descriptors. Development phase of the model includes machine learning models (e.g., SVM, KNN), deep learning architectures (e.g., CNN, U-Net, ResNet), and hybrid approaches [74], which are based on the integration of both approaches. Deep learning model tends to perform better than the traditional methods and requires more data [72]. Output is interested in classification (stone vs normal), segmentation (localization of stones) and size estimation. These are products where clinical decisions are highly relied upon. Finally, there is the stage of clinical integration, which is an AI model implementation into healthcare systems including integration into PACS, real-time decision support, and implementation through cloud or edge computing platform. The Taxonomy has made significant contributions that are important.

The suggested taxonomy provides:

- A combined platform in which to compare machine learning and deep learning.
- A systematic process of image capture to clinical practice.
- A systematic architecture of identifying the research and optimization opportunities gaps.

This article presents a new concept of a review framework that takes a unified approach of integrating clinical workflow, AI-based computational pipelines, into one taxonomy of kidney stones detection. However, unlike the current reviews, which emphasize the separate elements of the model performance or modalities considered, the proposed taxonomy can allow the systematic comparison of the modalities, preprocessing strategies, model architectures, and clinical deployment considerations. Moreover, it tackles key gaps like diversity of datasets, absence of external validation and clinical integration. This wholesome outlook offers more in-depth opinions on the advantages and constraints of present strategies and lays the groundwork of subsequent studies to strong, clinically viable AI systems.

4. Kidney Stones: Clinical Background

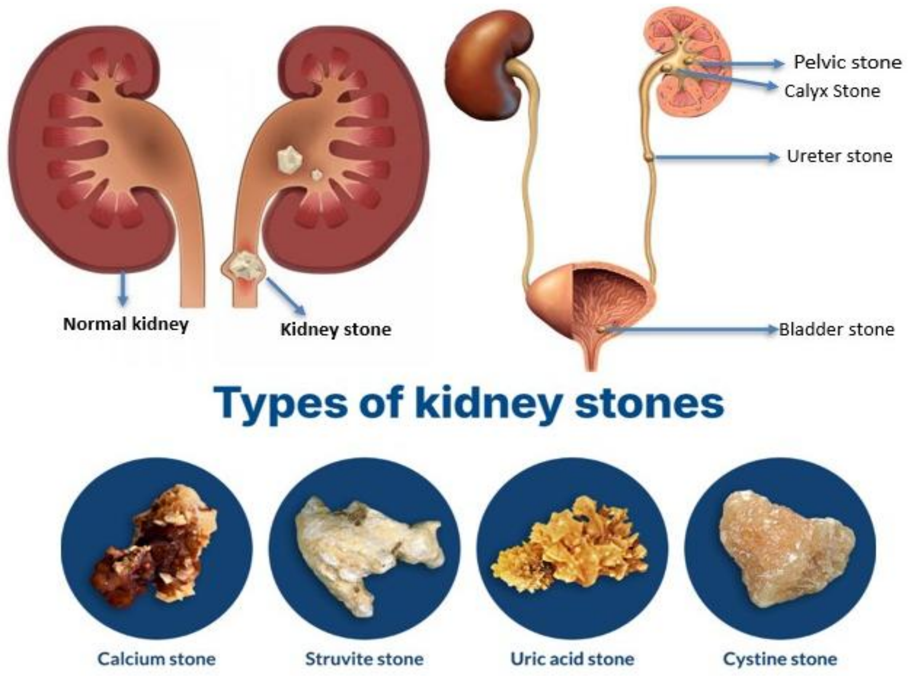

Fig. 4. Kidney stone forming places and types.

Kidney stones (renal calculi) are formed by the accumulation of salt and mineral salt deposits within the renal system [9]. These stones form when urine becomes supersaturated with calcium, oxalate, uric acid, or cystine to the extent that there are not enough inhibitors to crystallize. The process of stone formation is a sequence of events, including nucleation, crystal growth, aggregation, and retention within the nephron or urinary tract [15]. Nephrolithiasis is common and widespread condition and is the disease of people of world. Roughly 10-15% of the population are expected to develop kidney stones at some point in their life; within 10 years of the first episode about 50% will suffer a recurrence [1, 15]. The global prevalence of kidney stones is increasing worldwide, primarily due to changes in dietary habits, lifestyle and exposure to environmental and metabolic risk factors [9]. Fig. 4 describes the anatomy of the areas where most kidney stones are located. The spatial awareness is essential in the recognition of regions of interest (ROI) in AI systems since it directs the model awareness to clinical areas of interest.

4.1. Types of Kidneys Stones

Kidney stones are primarily classified based on their chemical composition. The four most common types include calcium oxalate, uric acid, struvite, and cystine stones. Each type is associated with distinct etiologist, risk profiles, and clinical management approaches. Table.1. illustrates the main types of kidney stones, their clinical variability can impose considerable dilemmas on automated detection systems. The composition and density cause a difference in imaging properties, which causes some type of stones to be more challenging to identify especially in the low-resolution or noisy imaging circumstances.

Table.1. Types of Kidneys Stones [8].

S No.	Type of Stone	Composition	Frequency (%)	Causative Factor(s)
1	Calcium	Calcium oxalate or calcium phosphate or both	60–85	Hypercalciuria, primary hyperparathyroidism, low urine citrate level
2	Struvite(Triple Phosphate)	Mixture of magnesium, ammonium, and phosphate	10–15	Urinary tract infection
3	Uric Acid	Uric acid (anhydrous or dihydrate)	5–10	Hyperuricosuria, acidic urine pH, or both
4	Cystine	Sulphur-containing amino acid	1–2.5	Cystinuria

4.1.1. Calcium Oxalate Stones

Calcium oxalate is the most common type of stone, occurring in 70 − 80% of cases. It is developed in acidic urine due to hypercalciuria, hyperoxaluria, hypocitraturia, and low fluid intake. Oxalate is found in foods like spinach, tea, chocolate, and nuts. Excessive oxalate intake or endogenous synthesis can cause supersaturation in urine. People with gastrointestinal diseases or gastric bypass may be more at risk. Calcium oxalate stones are radiopaque and rigid [18].

4.1.2. Uric Acid Stones

Uric acid stones, which cause 5–10% of kidney stones, can be detected in urine with a pH of less than 5.5. This condition, residing overweight or obese, metabolic syndromes, diets high in purines, and drinking too much alcohol can all raise uric acid levels. Because uric acid stones fail to show up on X-rays, a non-contrast CT scan or ultrasound may be necessary to make a diagnosis. It is important to find and classify them early since they can be dissolved in medicine by treating the urine with potassium citrate or sodium carbonate [6].

4.1.3. Struvite stones

Struvite, also referred to as infection stones, are connected with constant UTIs brought on by bacteria which generate uric acid. They lead to struvite crystals to form by reducing urine to ammonia. Staghorn calculi are frequently caused by these stones, which usually grow quickly and fill the renal pelvis. Due to their size, risk of kidney failure and collaboration with infection, they are radiopaque and need surgical management. Because UTIs are frequent in women, the impact is disproportionately greater in this population [13].

4.1.4. Cystine stones

Cystine stones are uncommon, accounting for only 1 to 2 percent of all kidney stones. They are caused by the genetic defect cystinuria, an autosomal recessive disorder of impaired renal tubular reabsorption of the cystine amino acid. Because cystine is relatively insoluble in urine, overexcretion of this substance causes the production of smooth, waxlike crystals, which sometimes are shaped like hexagonal tablets [17]. They commonly emerge during the early years of life and early adolescence and tend to recur frequently. Their detection may be difficult as their radiopacity is very low and they may go unnoticed by routine radiographs. It is frequently necessary to perform CT scan for better visualizations. Because of their genetic cause and low spontaneous remission rate, cystine stones frequently require long-term medical and dietary therapy [13]. Fig. 5 depicts the variation of the kidney stones and their occurrence percentage, which directly affects the imaging features and complexity of detection in the AI-based system.

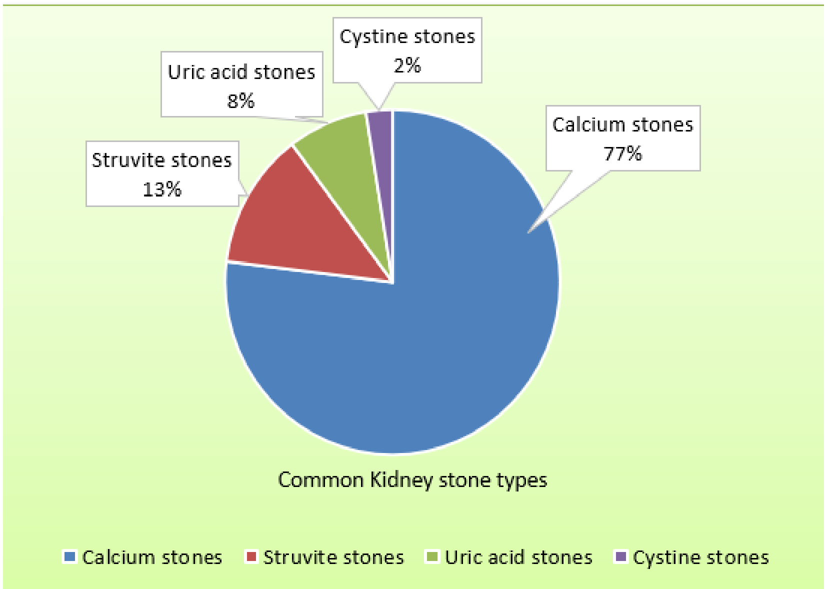

Fig. 5. Common kidney stone types.

4.1.5. Symptoms and Risk Factors

Renal stones look varied based on their size, where they are, and how much they block the flow of urine. Small stones that don't block anything may not cause any symptoms, but stones that do cause symptoms usually cause renal colic, which is a sharp, crampy pain in the flank that spreads to the lower abdomen, groin, or genitals. The symptoms include urination, feeling sick, vomiting, diarrhea, pressure, fever, chills, and symptoms of systemic poisoning. Kidney stones can lead to problems like disabled uropathy, hydronephrosis, progressive renal failure, acute kidney damage, and even life-threatening disorders like pronephroses and urosepsis. Kidney stones need to be treated right away since they might cause problems like pronephroses and urosepsis [16]. As such, timely diagnosis, the use of effective interventions in the acute setting and a preventive approach are integral parts of the comprehensive care [5].Various risk factors may lead to kidney stone formation, such as genetic susceptibility, metabolic abnormalities, dietary factors, and the environment. Genetic susceptibility represents an important factor with subjects presenting a genetic familiar predisposition for the disease, being at significantly higher risk. Inherited metabolic disorders, such as primary hyperoxaluria or cystinuria, affect the actual urinary excretions of stone-forming substances [12]. Fig. 6. illustrates kidney stone symptoms and risk factors. Combining this information of clinical with imaging data can improve prediction modeling and individual diagnosis.

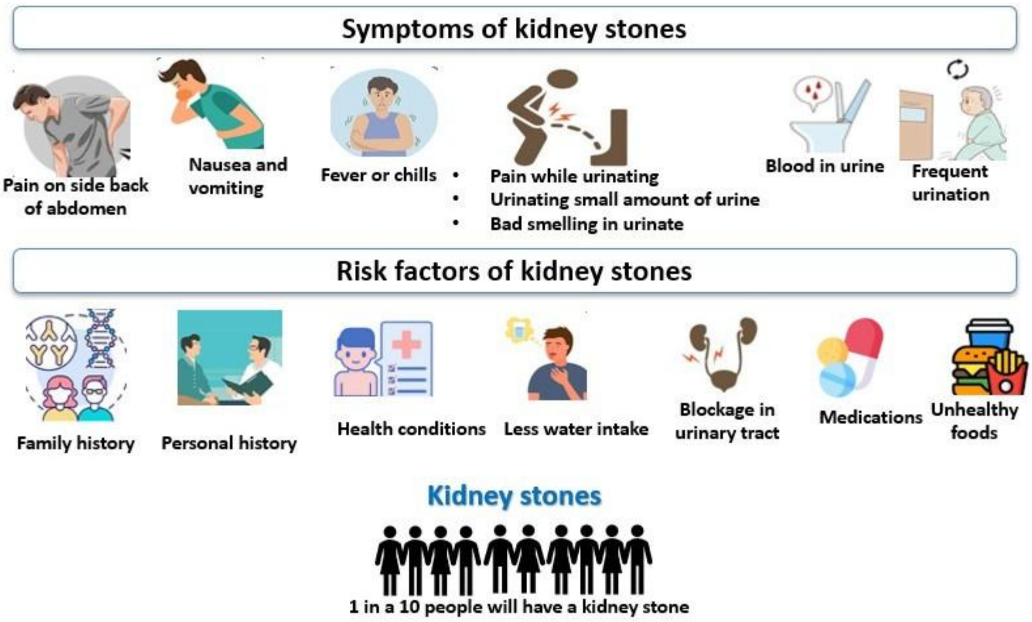

Fig. 6. Symptoms and Risk Factors of kidney stones.

High levels of animal protein, sodium, and refined sugar raise the excretion of calcium and uric acid in the urine, while oxalate in plant foods can lead to oxalate absorption and the development of calcium oxalate stones [19]. One of the most relevant factors is the modification of fluid intake. Inadequate hydration causes urine to be more concentrated, favoring precipitation of solutes. Populations in hot climates, or who are performing heavy labour with no replacement of fluid, are especially prone [19]. Obesity and metabolic syndrome have been reported as emerging factors in the etiology of stones. These all attract risk rather than an explanation for these diseases such as insulin resistance, acid urine pH and high level of urinary uric acid. Gastrointestinal disorders, such as Crohn's disease, ulcerative colitis, or previous intestinal bypass surgery, modify calcium absorption and oxalate absorption in the gut and predispose patients to hyperoxaluria. Some medications, such as diuretics, antacids, corticosteroids, as well as too much vitamin C or D, can also promote stone formation by different pathways [15, 16].

4.2. Diagnosis Process in Clinical Settings

Diagnosis and Management of Ureteral Stones The diagnosis of ureteral stones in the clinical practice proceeds through a thorough analysis of the clinical picture, laboratory exams and imaging studies [10, 13]. The diagnostic workup starts with thorough history taking (type, localization, radiation of pain; previous stones; diet; hydration; family history; systemic symptoms). On examination, a costovertebral angle tenderness, abdominal guarding, or evidence of systemic infection can be noted. Imaging has a central role in documenting the existence, size, location, and number of stones. NCCT is considered as the gold standard because of its high sensitivity and specificity in detecting stones of all compositions and determining the degree of obstruction. Less sensitive than CT, ultrasonography is the preferred initial modality for pregnant women and children because it is safe and can identify hydronephrosis [10, 13]. Simple abdominal radio- graphs (KUB) are helpful to follow up on known radiopaque stones and may not be useful to detect uric acid and small stones. In certain patients, intravenous urography or dual-energy CT may be used to evaluate function and distinguish stone composition [11, 12, 17]. These (clinical and imaging) considerations are the basis of the design of computational intelligence methods which are presented in the subsequent sections.

5. Medical Imaging Modalities in Kidney Stone Detection

Imaging is key for the identification, characterization, and management of kidney stones. It allows the clinician to evaluate the size, number, anatomic location and characteristics of the stones as well as any associated complications, including hydronephrosis or urinary obstruction. Several imaging techniques, with their own advantages and drawbacks, are used according to the clinical setting, characteristics of the patient, and availability of healthcare facilities [10]. The four most common imaging modalities employed in modern nephrolithiasis work-up are ultrasound, KUB (X-ray), computed tomography (CT) and magnetic resonance imaging (MRI). Here we provide a critical assessment of these modalities in the context of clinical relevance to kidney stone detection [22]. Fig.7. compares the imaging modalities in the detection of kidney stones, where CT imaging has high resolution and high accuracy, but has the risk of radiation exposure whereas ultrasound is safe and cost effective, but less sensitive with regard to smaller sized kidney stones. The considerations highlight the importance of paying attention to the choice of modality depending on the clinical scenario and justify the application of computational intelligence to increase detection results.

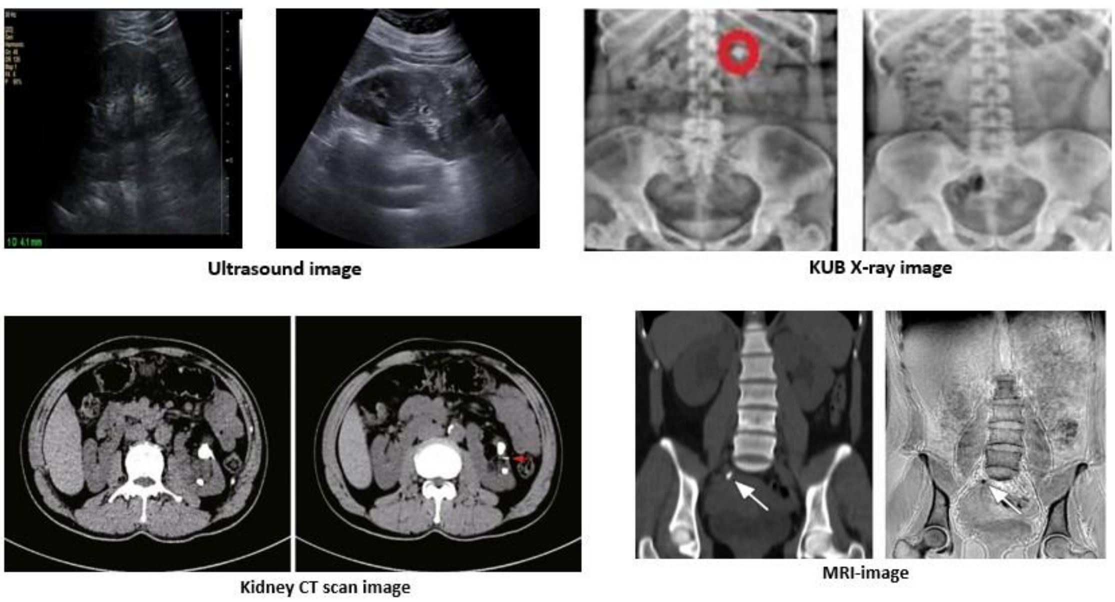

Fig. 7. Visual comparison of imaging modalities and their outputs for kidney stone detection [2, 6, 23].

5.1. Ultrasound: Safe, Low Resolution

Ultrasonography is known for being a safe, non-invasive, and readily available technique. As an imaging technique, it utilizes high-frequency sound waves to produce live images of the kidneys, ureters, and bladder. Ultrasonography is usually a first-line diagnostic option, especially in individuals where radiation should be avoided, such as in pregnant women or pediatric patients. It's a repeatable and patient friendly tool, free of ionizing radiation [13]. In clinical practice, US is excellent for detecting stones within the kidney itself whilst indirectly indicating hydronephrosis caused by urinary outflow obstruction. Nonetheless, the diagnostic sensitivity is much lower compared to CT, especially for the diagnosis of small stones ($< 5mm$), ureter stones and stones in anatomical difficult regions. Optician dependence also leads to variability in image quality and interpretation [14]. The visibility may also be restricted by body habitus and bowel gas. However, even with these limitations, ultrasound continues to play a significant role as an expedient initial examination in the setting of suspected nephrolithiasis, in follow-up situations, and in patients who should not be exposed to radiation [13, 15, 44].

5.2. X-ray (KUB): Low Cost, Limited Accuracy

Another conventional modality often used in the evaluation of nephrolithiasis is the kidney-ureter-bladder (KUB) X- ray. It is cheap, readily available, and simple, and thus, is a common method in outpatient and emergency departments [16]. KUB demonstrates radiopaque stones (most frequently calcium-based stones such as calcium oxalate and struvite) that appear white relative to the grayscale of soft tissue structures. But the diagnostic accuracy of KUB is relatively low. A large proportion, e.g. uric acid and cystine stones, are radiolucent and cannot be visualized by X-ray. Small stones and those situated over bony structures or bowel can be overlooked, resulting in false-negative findings. Furthermore, KUB only offers limited anatomical information and is unable to evaluate the obstruction degree or stone density [9, 16, 19]. Although it has limitations, it is frequently combined with ultrasound or CT to document the migration or dissolution of known radiopaque stones during follow-up. However, in low resource settings, it may still be useful as an inexpensive first pass test [19].

5.3. CT scan: Gold Standard, High Radiation

Computed tomography (CT), especially non-contrast helical CT (NCCT), has become the current gold standard for the detection of kidney stones. This method is highly sensitive and specific, being able to identify stones up to 1 to 2 mm diameter independently of their chemical nature. NCCT offers valuable anatomical information and precise localization of stones in renal pelvis, ureters, or bladder [8]. Examination is helpful in assessing stone size, number, shape, and density as measured in units of Hounsfield (HU) which are essential data for clinical decision-making especially with regard to the possible use of ESWL) [37]. CT scans are particularly helpful in emergency situations, where quick diagnosis is crucial, or for complicated cases in which other conditions (such as appendicitis, diverticulitis or an abdominal aortic aneurysm) need to be excluded [17]. Additionally, CT will demonstrate complications such as perinephric stranding, hydronephrosis, or obstruction, and visualize anatomical variations relevant to surgical planning [13, 17, 20]. The significant disadvantage of CT, however, is radiation. Standard-dose CT imaging entails exposure to substantially larger amounts of ionizing radiation than other modalities, which is a major concern, especially in young subjects or in patients with multiple refractory stone disease, necessitating frequent imaging [20]. In parallel, low-dose and ultra-low-dose CT techniques have been introduced, that considerably diminish the radiation exposure without decreasing the diagnostic power in the majority of patients. Despite the developments of these superior protocols, CT remains too expensive and less accessible in resource-challenged situations. Despite this, its higher diagnostic yield renders it the first-line modality in the majority of adults presenting with renal stone or 'stone disease' [19, 20].

5.4. MRI: Rare for Stones, Valuable in Soft Tissue

Magnetic resonance imaging (MRI) is an advanced non-ionizing imaging modality that exhibits superior definition of soft tissues. It uses magnetic fields and radio waves to create a sharp image of an anatomical structure [22]. But MRI can't be applied for kidney stone because it has limited ability in detecting the calcified objects. Most urinary stones typically appear hypointense on routine MRI sequences and are difficult to visualize as seen on CT or to a lesser extent as in ultrasound. Although MRI is of little value for the direct visualization of stones, MRI can play a supplementary role in specific clinical situations [19]. It is most valuable for evaluating soft tissue urinary tract structures, congenital anomalies, and renal masses. Among pregnant patients, in whom there is aversion to radiation and inconclusive ultrasonography, MRI may offer useful insight into the urinary tract and potential causes of obstruction [11]. Advanced MR techniques such as MR urography (MRU) and T2-weighted imaging may be able to demonstrate indirect signs of obstruction, such as hydronephrosis or dilatation of the renal pelvis, even when the stone is not visualized. MRI has additional value for those with contraindications to iodinated contrast agents used in CT, such as those with contrast sensitivities or severe renal dysfunction. But the price and long scanning time, limited accessibility and the challenge to display urinary system disease are limited it is application in the regular diagnosis of nephrolithiasis [11, 19]. Table.2. provides a comparison of different imaging modalities indicating that they are remarkably punishing in terms of trade-off.

Table 2. Comparison of Imaging Modalities Used in Kidney Stone Detection.

Modality	Pros	Cons	Resolution/ Accuracy	Cost	Clinical Consideration
Non-contrast CT (NCCT)	Gold-standard: highest sensitivity/specify; stone composition estimation (HU)	Higher radiation; overestimates stone size by 30–50%	91–100% sensitivity, 95–100% specificity	Moderate–high (CT scan)	Preferred for accurate diagnosis; limited use in repeated scans
Low-dose Ultra-low NCCT	Maintains diagnostic accuracy; reduced radiation	Reduced-HU accuracy; resolution loss for composition assessment	Similar sensitivity/specify; HU measurements compromised	Slightly lower than standard NCCT	Suitable for follow-up imaging with lower radiation risk
Dual-energy CT (DECT)	Excellent for identifying uric acid stones; differentiates compositions	Higher radiation; specialized equipment needed	Near 100% for UA vs non-UA stones	Higher (multi-energy CT setup)	Useful for treatment planning (stone composition analysis)
Digital Tomosynthesis	Lower dose than CT; near-MDCT detection for stones >5 mm	Poor for small (<5 mm) or ureteral stones; 3D recon limits	Sensitivity 94% overall; <60% for stones <5 mm	4,000 RMB ($560); cost-effective vs CT	Limited clinical use; not reliable for small stones
Ultrasound (US)	No radiation; low cost; bedside use	Lower sensitivity; inaccurate size estimation; operator dependent	Sensitivity lower than CT, <80% for small stones	Low	First-line screening, especially in pregnancy
MRI / MR-urography	No radiation; good for pregnant/pediatric patients; functional info	Poor for stone detection; signal voids can be ambiguous	Not standardized for stones; limited sensitivity	High	Used when CT is contraindicated

CT is the most accurate in its diagnosis but it exposes radiation and is also costlier. Ultrasound is less sensitive, operator-dependent, and safer and more available. KUB X-ray is affordable and has the weakness of not being able to spot radiolucent stones. These disadvantages indicate that there is no optimal modality, and hence, AI-based solutions could enhance the detection accuracy and assist human clinicians in their decision-making processes.

6. Image Preprocessing and Enhancement in Kidney Stone Detection

The combination of computational intelligence and medical imaging in kidney stone detection brings in new dimension of diagnostic accuracy and effectiveness. Nevertheless, the performance or robustness of an automatic or semi-automatic diagnostic system relies essentially on the quality of the input images [23]. Raw medical images, such as ultrasound, X-ray, and CT images commonly suffer from the distortion of noise, low-contrast, lighting discrepancy,

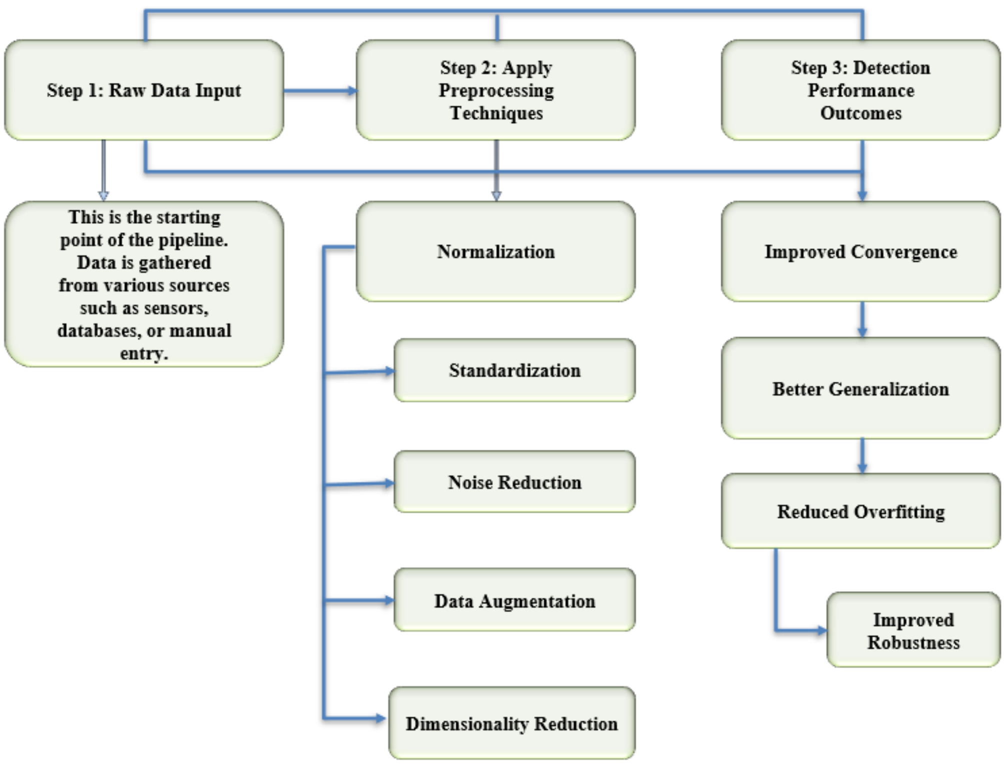

Fig. 8. Comparison of preprocessing techniques and their impact on detection performance.

and resolution variabilities. Consequently, an effective image preprocessing chain is crucial to achieve enhancement in diagnostically relevant features, while suppressing irrelevant or potentially misleading artifacts. It usually contains noise reduction, contrast enhancement, normalization, filtering and region of interest (ROI) extraction, and these steps are essential to enhance the accuracy and efficiency of the following computational tasks, e.g., segmentation, classification, and detection [23]. Fig. 8. shows the effects of the various preprocessing methods on the performance of detection. It is evident that more sophisticated methods like CLAHE and anisotropic diffusion filtering will make feature more visible than the simpler approaches to filtering. But too much preprocessing can leave artifacts or misrepresent significant structural information and thereby impacts negatively on models. Consequently, it is essential to design a preprocessing pipeline that will lead to maximum detection accuracy.

6.1. Noise Removal and Contrast Enhancement

Reduction of noise is one of the most important aims of pre-processing in the case of kidney stone imaging [25]. A significant number of medical images, and particularly ultrasound images, are corrupted by speckle noise, which is represented by a noise interference with a grained pattern that adds noise to image and to some extent degrades the clarity of the image. For CT images, the random noise from the imaging hardware or the patient movement can degrade the edge definition considerably. Poor processing of such noise can make unclear the edges of kidney stones or produce false positives, thus diminishing diagnosis efficiency [26]. To overcome this, many tools are used to denoise the images. Linear filtering techniques, like the averaging or mean filter, are simple techniques that efficiently reduce uniform noise, but also degrade edges and fine structures [10]. More sophisticated approaches like the median filter provide an edge-preserving alternative, where the value of each pixel is replaced by the median value in its neighborhood, hence removing salt- and-pepper noise without smearing the boundaries of stones [31]. After denoising, approaches enhancing the contrast of kidney stones against anatomical background have been employed to enhance the stones detection. Tumor or stone grayscale value is very likely to be very close to the intensity of the background tissue resulting in metal shadowing due to the poor contrast [28]. Histogram equalization is one of the most popular methods for redistributing the intensity levels and enhancing the global contrast. However, when the regions of interest are very small compared to the background (e.g. small calculi), global histogram equalization can seriously over-enhance irrelevant regions [28]. In this case, adaptive histogram equalization (AHE) or enhanced adaptive histogram equalization (CLAHE) is employed. These methods perform local contrast enhancement which results in the better appearance of small or low-contrast features (e.g., small kidney stones) while reducing background noise [30]. Though noise filtering and contrast filtering enhance the quality of the image, wrong filtering can result in disappearance of fine structural features especially in the small kidney stones. This means reliability of detection requires a trade-off between the noise reduction and edge preservation.

6.2. Normalization and Filtering

In practice, image acquisition conditions are often inconsistent since diversity of devices, protocols, patient position etc., exists. This variability has a potentially strong impact to the use of machine learning as well as deep learning since these algorithms assumes a random input data. Image normalization is thus a crucial preprocessing step to bring a quantitative consistency to the range of pixel intensity values of the images. This normalization guarantees that the computed model retrieves you the difference of the structure, and not on irrelevant differences of brightness or contrast [29]. The most common types of normalization are min-max, which normalizes the pixel values to a certain range (e.g., 0 to 1), and z-score normalization, which subtracts each pixel by its mean and divides by its standard deviation, to produce zero-mean, unit-variance images [31]. Gaussian filtering, as the one previously mentioned for noise removal, can similarly be modified to blur slight intensity changes and preserve essential edge detail. Bilateral filtering is much more complex and puts the weight on spatial closeness and intensity similarity so that the smooth is aware of the edge and it is particularly useful in terms of retaining small stones boundaries [30]. The median filter is still widely used in ultrasound image preprocessing for its ability to deal with impulse noise without blurring edges of anatomical structures greatly. In complex workflows, anisotropic diffusion filtering [16] can also be performed which iteratively smoothies the image while it also selectively preserves edges (very useful when working with CT imaging in which the balance between high frequency noise and edge details should be maintained) [33]. Although normalization helps to better converge and enhance the robustness of the models, inconsistent preprocessing among datasets may bring in bias and decrease overall generalizability. This points to the importance of standardized medical imaging pipeline preprocessing protocols.

6.3. ROI Segmentation for Stone Detection

After the denoising, enhancement, and normalization, the region of interest (ROI) extraction is other important pre-processing technique [34]. For ROI extraction, the specific anatomical area suspected to contain the kidney stone was selected, i.e. the renal pelvis, together with the calyces including the ureters depending on the suspicion. This operation reduces the computational complexity, filters out irrelevant background information and improves the efficiency and accuracy of further algorithms. Conventional approaches for ROI selection are thresholding (pixels above or below a given intensity threshold) and edge detection (based on Sobel, Prewitt or Canny edge detection methods, which identify rapid changes in intensities, i.e., anatomical edges) techniques [35]. For example, in computed tomography (CT) images, the stones are often observed as high-intensity surrounded by low-intensity soft tissue, and these are well-suited to the threshold-based segmentation algorithm [39]. More elaborate methods are region growing methods, in which a seed point is selected, either interactively or automatically, followed by the detection of neighboring pixels sharing similar features, until all such pixels are linked together to form a connected region [36]. Morphological operators, i.e., dilation, erosion, opening, and closing can then be used to refine the ROI mask, suppress noise pixels and close small holes in the segmented object. The ROI extraction is especially important in CT and ultrasound-based work where noise and artifacts may resemble stone structures. Focusing the approach on the areas of interest minimizes the chance of false positives and makes more effective use of computing resources [38, 39]. Furthermore, proper ROI delineation helps in enhancing the reproducibility of quantitative measurements of stone size, shape and location, imperative for a clinical decision-making purpose [36]. Although it has its benefits, the extraction of ROI is hard in low-contrast, or complex anatomy, resulting in partial segmentation or false positives. These limitations are being addressed by increasingly exploring the use of deep learning-based approaches. Preprocessing is also a key aspect of boosting machine learning and deep learning model performance. The input images should be of high quality to allow a better extraction of features and high accuracy of segmentation and classification. Poor preprocessing or inconsistency in preprocessing can seriously impair the model performance, especially in heterogeneous data. Preprocessing thus is suggested as the well-thought-out part presented in AI-founded systems of kidney stone detectors.

7. Smart Applications of Artificial Intelligence for Kidney Stone Detection

The rise of computational intelligence represents a radical change in the diagnostic scenario in the current era of medicine and opens up much new hope for precision, rapidity, and automation [34]. In the domain of kidney stones (i.e., nephrolithiasis), the conventional dependence on visual image inspection has been increasingly complemented and sometimes replaced by algorithmic models capable of analyzing large collections of imaging data, detecting complex patterns, and aiding clinicians in data-enabled decisions. Computational intelligence is a giant field of methodologies consisting of machine learning (ML), deep learning (DL) and hybrid methodologies combining rule-based logic with artificial neural networks. More and more such methods are now being used on kidney stone detection, classification, size prediction and recurrence prediction from imaging and clinical data [37].

7.1. Machine Learning Approaches

Machine learning techniques are the initial method to apply hand-crafted features from medical images and supervised learning algorithms to get shape, texture, and intensity values for kidney stone detections have demonstrated success in binary and multi-class classification tasks for GS detection, separating ”stone” and ”non-stone” regions in medical images. They have achieved high sensitivity and specificity in CT and ultrasound studies, using features like gray-level co-occurrence matrix (GLCM) and histogram of oriented gradients (HOG) [6, 7]. Initial stone classification is performed with a non-parametric method called K nearest neighbors (KNN). It makes use of the k nearest points in the feature space, but it may not perform effectively on big datasets or imbalanced data, such as clinical imaging [17].

7.2. Deep Learning Approaches

Deep learning is a subset of machine learning that utilizes artificial neural networks with more than one layer in order to learn intricate features of non-linear relationships from a large amount of data. These models do not need manual feature engineering, but extract significant features automatically during training, thus are particularly suitable for image-based diagnosis [78]. CNNs are a prevalent deep learning architecture in medical imaging. They are used for the identification and classification of renal stones in CT and ultrasound data. They use convolutional layers for spatial filtering, pooling, and fully connected layers for classification to distinguish stones from other biological noise [19, 22]. Localizing kidney stones is often implemented by U-Net, which is a particularly tailored CNN model for biomedical image segmentation. An explicit encoder-decoder architecture with skip connections helps the model to clearly define the stone boundaries in the presence of noise or low resolution. CT slices have been used to segment kidney stones using trained U-Net models, facilitating the precise quantification of volume and planning of therapeutic strategies [28, 32]. Residual Networks (ResNet) use skip connections to train deep networks without disappearing gradients. They perform much better than human experts at instructing various types of stone and their compositions in addition in multi-class classification circumstances [33].Another CNN-based variant named VGGNet [40, 42], which possesses a consistent layer configuration with an extremely deep structure [29, 41, 42], has also been used in the detection of kidney stones analysis. Even though it is computationally expensive, VGGNet has shown promising results in classification, especially when applied to medical images and pre-trained through transfer learning. Being pretrained on large scale datasets, such as ImageNet, they need less domain specific data to reach high accuracy, particularly beneficial for the medical domain where labelled datasets are scarce [48].

7.3. Hybrid Approaches

A popular hybrid architecture uses a CNN as a feature extractor and a Support Vector Machine (SVM) as a classifier, encoding spatial locality features in CNN convolution layers for learning patterns from medical images, resulting in decent accuracy and robustness [36]. A different mixed approach is to combine rule-based systems and deep learning models [44]. For example, specific domain knowledge about ubiquitous locations of stones or strength of intensity can be leveraged to guide the preprocessing or region of interest (ROI) selection priors to inputting images into a deep learning pipeline is used. Such rule-based components may contribute to a decrease of false positives, a situation that may be encountered in noisy imaging environments where CNNs can misinterpret overlapping structures [6, 7]. Although machine learning techniques are based on manually crafted features and are computationally cheap, their prediction tasks are very sensitive to the quality of the features and do not always scale to more demanding tasks in imaging. On the contrary, deep learning methods learn hierarchical feature automatically and perform better in reference to detection and segmentation [79]. They however need huge datasets of annotations, and big computational capabilities. Hybrid methods aim to merge the advantages of the two strategies and make use of the deep feature significance by utilizing the classical classifiers to enhance robustness when limited data sets are involved. Although they have these benefits, issues like overfitting, interpretability, and inability to predict across datasets are crucial drawbacks in modern AI-based kidney stone detectors. Table 3. is the summary of the recent researches on kidney stone detection, but a few observations can be made. High accuracy (usually over 95%), generally, is reported in most studies although many of them use small or private datasets, limiting the ability to generalize. Very high accuracy (e.g., 99-100%) is likely achieved in these studies due to limited number of samples, lack of independent test set and overfitting. Also, there is seldom external validation, which should be questioned because of the possibility of overfitting. The recent research has been taken over by deep learning models, especially U-Net and CNN-based ones, as the models have shown better performance and the traditional machine learning methods are being gradually phased out. These results indicate that standardized assessment procedures and big data are required.

Table 3. Summary of studies on kidney stone detection using different models and datasets.

Author	Year	Model / Method	Dataset Size	Findings	Limitations
Abdalla et al. [73]	2025	CNN	3,364 CT images	98.8% accuracy	No external validation
Sharma et al. [15]	2025	ResNet101 + CNN	12,446 CT images	Very high accuracy 100%	High computational cost
Rathva et al. [53]	2025	VGG16 Ensemble	12,446 CT images	96% accuracy	Generalization issues
Obaid et al. [9]	2025	Ensemble DL	US images	High accuracy 99%	Noise sensitivity
Asaye et al. [4]	2025	KNN	US images	98.4% accuracy	Limited modality
Ahmed et al. [2]	2024	VGG16 + XAI	500 X-ray images	97.1% accuracy	Small dataset
Sharen et al. [13]	2024	Transformer	CT dataset	F1: 0.98	High computational cost
Ravisankar et al. [54]	2024	DenseNet	CT images	86% accuracy	Moderate performance
Wang et al. [63]	2023	CNN-Fusion	1,799 CT images	F1: 96%	No external validation
Rabby et al. [10]	2023	YOLOv7	CT images	High accuracy 99%	No clinical integration
Tahir et al. [66]	2023	CNN + DWT	CT data	accuracy 99%	Model complexity
Alkurdy et al. [3]	2023	CNN + VGG16	US images	High accuracy 99%	Ultrasound artifacts
Eun et al. [22]	2022	Fast R-CNN	CT images	Sensitivity: 0.93	Model variability

8. Feature Extraction and Selection in Kidney Stone Detection

In the computer aided diagnostics for the detection of kidney stones, performance of any machine learning or deep learning model depends significantly on the quality and relevance of the features extracted from the imaging modality. Feature extraction refers to the analysis step where a set of image features is derived from the raw data to be used for further processing. These features are then used as input for classifier and segmentation models, which predict whether a stone exists, its type, and its position [41]. Due to the high-dimensionality and complexity of medical images, a proper feature selection and dimensionality reduction are required for the computational efficiency improvement and over-fitting reduction.

8.1. Features: texture, shape, edge, and intensity

In the conventional image processing algorithms, handcrafted features are generated through mathematical and statistical approaches, which reflect the visual appearance and structure of objects. Textures, including those encoded by the GLCM, measure the spatial dependence among pixel intensities. Features such as contrast, correlation, entropy, homogeneity, and energy are particularly relevant for distinguishing stone and soft tissue because of their characteristic granularity and structural homogeneity [43]. These textural characteristics are particularly applicable in ultrasound and CT images, where the stones manifest as bright, high-density spotty regions with non-smooth surfaces [44]. Shape features characterize the size and shape of the detected regions. For kidney stone detection, shape descriptors (compactness, circularity, elongation, aspect) of stones can help to differentiate stones from cysts, calcifications, or anatomical variants [45]. The shape of the calculi is particularly useful in the segmentation process, since the purpose of the segmentation is to extract and measure the size and the shape of the calculi [41, 46]. With respect to the edge feature, edge information can show the area of intensity change, which is usually considered as boundaries of stones. Edge detection methods, such as the Sobel, Canny and Laplacian operators are oftentimes used to detect the edges of kidney stones in CT and X-ray images. The resulting edge maps may be employed to calculate additional data, including perimeter, boundary irregularity, and gradient strength measures that may be informative in comparing stone complexity and response to therapy [47]. Intensity features, on the other hand, describe the histogram of gray values within the ROI. These features are inexpensive to compute, yet informative, especially in CT imaging where stone density (in Hounsfield units) is an important diagnostic criterion. Statistical knowledge of the stone's composition and size is available in the average intensity, standard deviation, skewness, and kurtosis of pixel values [48, 43].

8.2. Deep Features on CNN Layers

Although handmade features allow for interpretability and certain domain-specific insights, they have low representation power for modelling complex spatial hierarchies [43]. This limitation is remedied by deep features, learned automatically from data via Convolutional Neural Networks (CNNs). In a CNN, each convolutional layer generates a corresponding set of feature maps which encode the abstractness of the input image [51]. Whereas in kidney stone detection, shallow layers of CNNs represent low-level features (e.g., edges and textures) and deeper layers correspond to high-level representations (e.g., shape and spatial arrangement). These deep features are not hand-engineered exploits but rather, features extracted from learning the disappointments of the network itself. And so, they are less likely to be the specific model that applies to this data [47].

8.3. Techniques for Dimensionality Reduction: PCA and LDA

PCA is an unsupervised way for minimizing the number of dimensions. This is accomplished by showing the initial feature space onto variance-sorted parallel parts. It prevents overfitting and speeds model training by hurrying up feature space while preserving significant categorization patterns. By making the variance among classes larger than the variance within classes, the supervised learning technique known as LDA, or linear discriminant analysis, lowers the number of dimensions. Because LDA reduces the generalization error, it is excellent for integrating classes, such as kidney stone varieties, based on features that have been retrieved [53]. PCA and LDA are the indispensable preprocessing blocks for all conventional as well as hybrid kidney stone detection methods. They can reduce feature vectors without loss of important information, enabling the development of highly efficient, scalable, and clinically practical diagnostic systems [51]. Table 4 shows the variety of the types of features used to detect kidney stones. Although the traditional features are computationally efficient, they have a limitation in dealing with complex variations of images. Deep features are more powerful, but need large datasets and computational resources and focus on the trade-off between performance and efficiency.

Table 4. Summary of feature types and their extraction techniques.

Feature Type	Extraction Techniques
Shape-based features	Compactness, elongation, sphericity, eccentricity, circularity, volume, surface area
Intensity-based	Histogram statistics (mean, variance, skewness, kurtosis, percentile)
Texture-based	Gray-level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM), gray-level size zone matrix (GLSZM), neighborhood gray tone difference matrix (NGTDM)
Wavelet features	Decomposition of image into frequency sub-bands to extract spatial and frequency features
Deep features	Automatically extracted using convolutional neural networks (CNNs) from raw image inputs
Radiomic features	Combination of shape, intensity, and texture features using tools like Radiomics

Despite the important role of feature engineering in conventional machine learning methods, handcrafted features are likely to miss the complex spatial patterns of medical images. CNNs have the benefit of deep feature extraction at the expense of interpretability. As such, proper feature representation is a major problem when trying to come up with viable kidney stone detection systems.

9. Segmentation Techniques for Stone Localization

Precise identification of the anatomic position of kidney stones in medical images poses an important step in automated diagnostic pipelines. Segmentation, dividing the image into anatomically meaningful regions, can help in accurate location and demarcation of kidney stones from other anatomic structures. Accurate segmentation will help target serves for subsequent tasks such as classification, volume quantification, and radiation planning [39, 42]. Time-consuming and effective segmentation tools are as simple as thresholding and region growing to our goal, and expensive morphological operations and semantic segmentation-based deep learning models. Both methods have their strengths and weaknesses and the decision on the method used is influenced by the imaging modality, the quality of the images and the clinical environment [45].

9.1. Techniques of thresholding and region growing

Thresholding is one of the most significant and popular segmentation methods. It works by transforming gray-scale images into binary format by using a specified intensity threshold. In the case of kidney stone detection, this process can be easily used to extract high intensity regions, which corresponds to dense calculi particularly for the CT data since the stones are generally brighter than their surrounding tissue indicating higher radio density of the stones [49]. However, simple global thresholding is sensitive to image noise and contrast variations, which result in over- or under-segmentation. Adaptive thresholding approaches are an attempt to overcome these shortcomings of a fixed threshold value by using locally computed threshold values determined from the intensity distribution in a local region containing the pixel, and hence taking into account more contextual information [51, 54]. It is spatial localization technique, factors in analysis use to come up with segmentation results that is dependent on spatial positions, defined on some initial seed point (in this case it will be a seed point indicated in the location of a stone which can be presumed manually or automatically) and expands the selected region according to the following rule: 1. Start from the initial seed point representing the stone. 2. Add the neighbouring pixels to the initially selected stone region if they have similar texture or intensity. 3. Continue until no more pixels added (or an arbitrary threshold is met) [41]. This approach is especially advantageous if stones afford uniform properties that differ from neighbouring tissues. Nevertheless, the region growing is extremely susceptible to the seed points and the similarity criteria. Incorrect selection may lead to the addition of irrelevant structures or missing stone pieces. To overcome this, region growing is frequently used in conjunction with edge detection or statistical region merging in order to enhance precision [54].

9.2. Morphological Operations

Morphological operations are fundamental ones in image preprocessing and segmentation post-processing, especially for binary images generated by thresholding or region growing. These operations are essentially set operations which use a structuring element in order to alter the topology of objects in an image [53, 55]. The basic morphological operations are dilation which increases the size of objects; erosion which reduces it; and combinations of these operations - opening and closing. In kidney stone imaging, morphological operations are often used to eliminate small artifacts and fill small gaps in the segmented regions, or separate a connecting structure from the overlapping ones [49]. For example, the small speckles as well as noise artifacts generating false positives can be subtracted with the operation of opening, and the fragmented stone parts can be linked with the closing operation [57]. Further advanced morphological operations, e.g. the watershed segmentation, use the intensity gradient to separate touching or overlapping stones applying topographic principles. These are mostly useful for thick clumps of stones or for stones that lie next to an anatomically dense boundary, like bone or vessel [54].

9.3. CNN-based Semantic Segmentation (e.g., U-Net)

Although traditional methods can offer efficient and interpretable segmentation for simple cases, they are limited by their ability to address complex images involving contrast variations, noises or overlapped anatomical components. To cope with these difficulties, deep learning-based segmentation models, and especially Convolutional Neural Networks (CNNs), emerge as the state-of-the-art methodology in medical image analysis. In the family of CNNs, U-Net architecture has become a state-of-the-art model for semantic segmentation tasks especially in bio-medical imaging. In U-Net architecture, there is a shallow contracting path (encoder) to capture context, as well as a symmetric expanding path (decoder) to enable accurate localization. Skip connections between the corresponding levels of encoder and decoder paths can hold spatial information that was lost during down-sampling, and it is beneficial for the discrimination of small and irregularly shaped structures (for example kidney stones). On kidney stone identification, U-Net models are trained from CTS and US data with segmenting stones automatically, and provide high Dice/IoU scores. The ability of the model to learn suitable features i.e., shape, texture and intensity distribution, directly from the training data enables it to easily cope with the variation in stone size, composition and anatomical context. U-Net performs extremely well for small or low-contrast stones, which are commonly overlooked by conventional techniques [50]. For better performance, U-Net, but also its variants such as Attention U-Net, ResUNet and 3D UNet have been introduced [53]. These models use attention modules, residual learning and volumetric data handling, respectively, to take into consideration segmentation accuracy for challenging or multi-slice datasets. For example, 3D U-Net operates on the whole volumes instead of 2D slices and can achieve more consistent stone segmentation in 3D space a key aspect to calculate the size and volume of stones with precision [55]. However, despite their impressive performance, CNN-based segmentation models are task-specific and need large, annotated datasets to train on, which can be an impediment in clinical contexts. To mitigate this limitation, some transfer learning and data augmentation methods, which take a data-driven optimization approach to leverage models pre-trained on large public datasets and would subsequently be fine-tuned for specific cohorts (clinical populations) or applications (imaging protocols) need to be integrated [57]. The conventional segmentation methods which include thresholding and region growing are very easy and computationally friendly but very sensitive to noise and intensity changes. Conversely, deep learning-based segmentation algorithms like U-Net are more accurate and robust, however, large annotated datasets and costly to compute are required. This trade-off emphasizes the need to make the right choices regarding segmentation strategies, depending on clinical needs and the availability of data.

10. Datasets and Annotation Challenges for Kidney Stone Detection

The performance of AI-based models for kidney stone detection largely relies on the existence of annotated imaging datasets. Such datasets are the basic input for the training, validation, and test of machine learning (ML) and deep learning (DL) algorithms. Nonetheless, large scale, high-quality datasets are not easy to obtain in medical imaging [60]. Challenges including difficulty in access to the data, problems of imbalance in small scale datasets, and annotation inconsistencies prevent the model generalization and robustness. To mitigate these limitations, data augmentation methods are commonly used to increase the effectiveness of training and the robustness of the model [27].

10.1. Public vs. Private Datasets

As public data set constitutes a useful benchmark for comparison and competition among researchers to test the performance of various algorithms under the same condition. However, the publicly available datasets for detecting kidney stone are quite limited. Some databases (DICOM/ files), such as those available on platforms, such as Kaggle, or the open access ones provided by radiological societies, contain samples of CT- or X-ray images labelled with urinary tract-related diseases such as Nephrolithiasis [63]. Despite their values, the existing datasets are large but not that diverse in images and demographics. In contrast, private datasets, which are usually collected by hospitals, diagnostic labs or academic institutions, are richer in modality, image resolution and clinical metadata. Frequently, they contain CT images, ultrasound images or, occasionally, MRI sequences, along with corresponding radiologist readings. But they are difficult to access because of ethical, legal and privacy restrictions pertaining to data such as those imposed by data protection laws

including HIPAA and GDPR. In addition, variances in acquisition protocol, scanner parameters, and annotation criteria from different institutions also bring heterogeneity that makes the model training and validation challenging [59, 65]. The authors also point out that the lack of standardization between public and private data sets is a barrier to the creation of diagnostic tools that can be used universally. In reality, many successful models published in the literature are trained also on private datasets and might not suit for generalization across external datasets or various clinical scenarios [66].

10.2. Imbalanced Class, Small Size and Annotation Error

Data imbalance is one of the most lasting problems for kidney stone imaging datasets. Stone size, location, composition, and appearance vary widely, with some stones (e.g., calcium oxalate) being orders of magnitude more frequently encountered than others (e.g., cystine stones). Therefore, data frequently suffers from a sampling bias towards some classes of stones and learning becomes biased. Common models trained on such datasets may have high overall accuracy but lose their performance on less frequent cases, and are very dangerous in the clinic [67]. The second challenge is limited dataset size, which is especially troublesome given that, for deep learning, tens of thousands of labelled examples are commonly required to train reliable models. In practice, often in rural and specialized clinical facilities, only several hundreds of imaging case can be collected, and the trained model may be shallow and less generalizable. This is even more critical in pediatric or rare stone entities, in which it is inherently difficult to gather significant numbers of cases [68]. A third important limitation is annotation errors. Therefore, in medical imaging, the ground-truth labels are generally provided by radiologist's anatomical boundaries and classified stones. Nonetheless, inconsistent or wrong annotations may be produced due to inter-observer variation and subjectivity in annotation reading. Small stones may be overlooked e.g. benign calcifications may be confused for stones. Moreover, edge cases, like intersecting anatomical structures or artifacts, can lead to uncertain observations, thereby misleading the training, and deteriorating the reliability of the trained model [61]. To overcome these problems, we suggest the consensus labelling among multiple radiologists, annotation audits and a standardized labelling protocol. Semi-automated annotation which injects human oversight to guide a (initial) recommended set of annotations has been another tool that has been gaining popularity over the past few years as a way to easily annotate while maintaining control over annotation quality [63, 65]. These issues of datasets have a profound effect on the model performance and generalizability. Smaller or biased datasets may lead to models that cannot be used well in actual clinical conditions, hence the importance of standardized, diverse, and well-annotated datasets.

10.3. Data Augmentation Strategies

Due to the restricted data set access and diversity, the data augmentation has been a key part in training pipeline in kidney stone detection systems. Augmentation is the process of creating new training samples by applying stochastic operations to original images. These transformations maintain the diagnostic traits and inject variations that make models generalize more. Geometric image transformations including rotation, scaling, translation, and flipping are the most basic and widely used augmentation techniques. And they mirror patients who have natural configurations for them, to make sure models isn't hyper-sensitive to any kind of orientation and size you give it [65]. To enable the model to be robust to variations in imaging factors and scanners, such as change of brightness, contrast, and noise, we adopt photometric transformations. This is especially salient for ultrasound, where image quality varies widely [66]. Global elastic and affine transforms are more sophisticated methods that account for soft tissue deformations and interindividual variability. Such augmentations should be particularly beneficial for training segmentation models that can reliably identify stones in highly variegated anatomical presentations. Recently, synthetic data generation based on Generative Adversarial Networks (GANs) is becoming a new augmentation strategy [67, 68]. The GANs can generate realistic medical images which mimic rare pathology or imaging artefacts, thus enriching the pool of training data and contributes to the robustness of the model. Synthetic images, if validated by clinical experts, may be of great value as complementary sources of data, especially for underrepresented stone types or imaging modalities [64]. Table 5 shows that the majority of publicly available datasets are small and similar while larger datasets are private and unavailable. This incompatibility limits the reproducibility and equitable comparison of models between studies, which is a significant challenge of developing AI-based kidney stone detection.

Table 5. Summary of publicly and privately available kidney stone-related datasets.

Dataset Name	Author (Year)	Imaging Modality	Dataset Size	Availability
Abdalla CT Kidney Dataset	Abdalla et al. (2025)	CT (Axial)	3,364 original / 35,457 augmented	Public (Kaggle)
DeepStone Dataset	Patel et al. (2024)	CT, Ultrasound	1,100 patients	Private (hospital trial)
UCI Kidney Disease Dataset (Updated)	Dua & Graff (2024 up- date)	CT (Axial)	400 records	Public (UCI Repository)
KidStoneNet Dataset	Sharma et al. (2025)	CT (multi-angle)	2,050 scans	Public (ArXiv-linked)
Open Kidney Dataset	Singla et al. (2022)	Ultrasound (B-mode)	500 annotated images	Public (CC BY NC SA)
Yildirim NCCT Kidney Dataset	Yildirim et al. (2021)	CT (Coronal NCCT)	1,450 train + 346 tests (1,796 total)	Private / research use (via GitHub)

To overcome these issues, concerted actions are needed to create large-scale and standardized datasets, better annotation procedures, and effective validation plans. New methods like federated learning and semi-supervised learning can provide a chance to solve the problem of data scarcity and privacy limitations.

11. Evaluation Metrics for Kidney Stone Detection and Segmentation

In the domain of kidney stone detection using computational intelligence, it is essential to assess the model's performance using standardized evaluation metrics. These metrics provide a quantitative basis for understanding how well a model performs on classification, detection, and segmentation tasks. This document describes each metric with theoretical context and mathematical formulation.

11.1. Accuracy

Accuracy is a fundamental metric used in classification tasks. It measures the proportion of true results (both true positives and true negatives) among the total number of cases examined. However, in cases of imbalanced datasets, like kidney stone detection, where the number of negative cases may outnumber the positives, accuracy can be misleading [52]. Accuracy can be determined using Equation (1).

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{1}$$

Notation: TP is true positive (stones correctly identified), TN is true negative (non-stones correctly identified), FP is false positive (non-stones marked as stones), and FN is false negative (stones marked as non-stones). Thus, accuracy cannot be trusted as a metric for kidney stone detection, particularly when the datasets are imbalanced, and should be supplemented with sensitivity, specificity, and Dice-based measures.

11.2. Sensitivity (Recall or True Positive Rate (TPR))

Sensitivity, also known as recall or true positive rate, refers to the ability of a model to correctly identify all relevant instances of a condition. In kidney stone detection, a high sensitivity ensures that most actual stone cases are identified, reducing the risk of missed diagnoses [52]. Equation (2) provides the formula used to calculate Sensitivity (Recall or True Positive Rate (TPR)).

$$Recall\ (or)\ Sensitivity = \frac{TP}{(TP+TN)} \tag{2}$$

The issue of sensitivity is especially significant in clinical practice since the failure to detect a kidney stone (false negative) may result in delayed diagnosis and severe complications.

11.3. Specificity (True Negative Rate (TNR))

Specificity measures the proportion of actual negatives that are correctly identified. It is critical in medical diagnosis to avoid unnecessary treatments caused by false positives [53]. The formula used to calculate Specificity (3).

$$specificity = \frac{TN}{(TN + FP)} \tag{3}$$

The sensitivity and specificity need to be balanced to guarantee proper diagnosis without exposing the patient to unwarranted treatments.

11.4. F1-Score

The F1-score is the harmonic mean of precision and recall. It provides a balance between precision and recall, and it is especially useful in evaluating models on imbalanced datasets [53]. Equation (4) provides the formula used to calculate F1-Score.

$$\text{F1 Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{4}$$

F1-score is also very beneficial in imbalanced datasets since it gives a balanced value of false positive and false negative.

11.5. Receiver Operating Characteristic - Area Under Curve (ROC-AUC)

The ROC curve is a plot of the true positive rate against the false positive rate at various thresholds. The Area Under the Curve (AUC) represents the degree of separability. A higher AUC means the model is better at predicting 0s as 0s and 1s as 1s [52]. ROC-AUC is a common metric to compare the performance of different models at varying thresholds and is particularly applicable in assessing classification models in medical diagnosis.

11.6. Dice Similarity Coefficient (DSC)

The Dice Score, or Sorensen-Dice coefficient, is used to measure the overlap between two samples. It is widely used in medical image segmentation to evaluate how well the predicted region (P) matches the ground truth (G) [53]. Equation (5) provides the formula used to estimate Dice Similarity Coefficient (DSC).

$$\text{Dice Score} = \frac{2 \times |P \cap G|}{(|P| + |G|)} \tag{5}$$

Dice Similarity Coefficient and Intersection over Union are both common in segmentation tasks but Dice is more sensitive to small objects and is more commonly used in medical imaging.

11.7. Intersection over Union (IoU)

IoU, also known as the Jaccard Index, is another common metric for evaluating segmentation tasks. It measures the overlap between the predicted and actual segmentation masks [69]. Equation (6) is used to calculate Intersection over Union (IoU).

$$IoU = \frac{|P \cap G|}{|P \cup G|} \tag{6}$$

IoU, however, is a more stringent measure of overlap and can be used to assess the overall segmentation performance. Dice is commonly used in the detection of kidney stones because of their small size and shape. Overall, measurement scales are significant in determining the dependability and practical use of AI-based kidney stone detection systems. Where classification measures like accuracy, sensitivity and specificity measure the performance of detection, segmentation measures like Dice and IoU measure spatial accuracy. To have strong and clinically reliable models, a combination of these metrics is required.

12. Challenges and Limitations

Although computational intelligence has enhanced kidney stone detection to a considerable extent, there are some important challenges that inhibit the application and scalability of the technique in clinical settings.

12.1. Lack of Data and Standardization

One of the main challenges in the AI-based kidney stone detection is the lack of annotated imaging datasets and heterogeneity between them. It is also a lot of work to construct a large-scale, publicly accessible dataset for training computational or deep learning (DL) model [77]. Yet in the field of nephrology and urology, datasets can be much smaller and rarely diverse, because of data privacy, patient confidentiality and institutional separation [62]. There is also no uniformity in data acquisition protocols for different types of imaging. Variabilities in CT scan parameters, ultrasound machine brands, and image resolution, as well as patient demographics contribute to differences in the quality and the content of the images. These mismatches cause challenges for models trained on one set of data but applied on another, often leading to degradation in generalization. In addition, the lack of benchmark datasets available for kidney stone detection makes it difficult to compare performance across research [65].

12.2. Explainability and Trust in AI Systems

Another limitation that hinders clinical acceptance is the lack of interpretability of complex AI models, in particular deep neural networks. Although modern models such as CNNs can perform image classification and segmentation with very high accuracy levels, their decision-making process is seen as a "black box" by clinicians. For high-stakes medical decisions lack of transparency can be detrimental to trust, acceptance and use in clinical workflows [67]. Additionally, models can be based on unnoticeable image artifacts or irrelevant confounding features to the pathology that might compromise the model and mislead both the model and the clinician. Failure to rationalize predictions of stones in specific locations undermines the clinical applicability of the model in decision support systems, and raises questions about ethical responsibility and accountability in AI-augmented diagnostics [76].

12.3. Overfitting and Representation Learning

Overfitting is a well-known challenge in machine learning, especially when training models on small or unbalanced datasets. By modelling-to-excess, we mean that the model overfits training data, getting it very right on early datapoints, but being unable to generalize well to new test cases [80]. In an application such as for detection of kidney stones, where high anatomical and demographic variability exists, overfitting can have a profound negative impact on diagnostic accuracy [69]. Many deep learning-based models are sensitive to small differences in imaging conditions; patients' pose or the overall device calibration. Without effective regularization, data augmentation and cross-validation, there is a significantly higher risk for deriving inappropriate models that perform poorly in a real-world setting. Generalizability across populations and scanner technologies is a major limitation yet to be resolved at the clinical scale [70].

12.4. Computational Burden and Resource Constraint

Deep learning models, such as 3D segmentation networks or multimodal architecture, there are significant computational costs. These models need strong graphic processing units (GPUs), large memory access bandwidth and storage for processing of sizable datasets features absent in many health care facilities, especially in the developing world [63].It is worth noting that the inference time of sophisticated models could be inadequate for real-time clinical applications where diagnostic timeliness is essential. Integration into hospital environments also necessitates support for hospital watchdog or industrial standard, hardware acceleration and IT support, which adds more burden for deployment in resource-limited areas. Therefore, the balance between model complexity and computational burden needs to be handled properly [64].

All these issues underscore the fact that although the performance of AI models in controlled settings has been high, the transfer of AI models into actual clinical practice has been minimal. These are the problems that should be addressed to come up with reliable, scalable, and clinically deployable kidney stone detection systems.

13. Future Directions

A number of studies have emerged with the aim of fixing these issues and facilitating clinically reliable AI-based kidney stone detector systems [75]. To address the above challenges and unlock the full potential of AI in nephrology, several exciting future directions have arisen. The above approaches, focused on technological innovation, privacy-preserving learning, explainability, and real-time in-clinical applications.

13.1. Privacy-Preserving Federated Learning

Federated learning (FL) is a decentralized machine learning scheme for learning models across different institutions, without the need to exchange patient data. Rather than maintaining a centralized database, at each hospital or clinic the model is trained using local data and only model updates (e.g., gradients and weights) are aggregated centrally [56].

This procedure provides a better and more comprehensive training set for model development, but at the same time, maintains patient anonymity. For kidney stone detection, FL can also be applied to effectively train models on multi-institution data from different imaging devices, scanner settings, and populations to achieve better generalizability and robustness [58].Furthermore, FL enables collaborative research under compliance with data protection laws like GDPR or HIPAA [64]. Furthermore, by adopting secure techniques, like differential privacy or homomorphic encryption, it can guarantee more data security in the training and the transmission [61].

13.2. Clinical AI Tools in Real Time

Real-time diagnostic aids will characterize the future of AI in nephrology by allowing AI to become a frictionless part of the clinic. These lightweight models also can be pruned, quantized, and edge-computed models that can be deployed directly on PPE or hospital servers to analyses as scan images are collected [60]. For example, an AI algorithm integrated in a CT scanner could automatically alert on suspected kidney stones as an image is taken, leading to more timely diagnosis and helping a radiologist while coping with busy periods of demand [64]. Likewise, the AI-based US apparatus might assist the technicians to concentrate on the suspected stone regions, even with lower resolution [69].

Moreover, cloud-based applications will facilitate tele-radiology, so that general practitioners in remote, poorly and underserved areas can get access to AI-based diagnostics without having to shift to specialized machines [61].

13.3. 3D Stones Detection and Measurement

Recent developments in 3D medical imaging, technology and volumetric deep learning models have created new opportunities for accurate kidney stone detection and stone sizing. This 2D analysis is limited by slice thickness, resolution and anatomical complexity. Compared to 2D-CNN, in our 3D-CNN-based and U-Net-based methods, entire CT volumes can be taken as input and thus the kidney stones can be located and segmented in 3D [66,67]. Thereby, stone volume, density and surface can be accurately assessed, which are crucial for treatment planning and outcome prediction. For instance, volume measurement can aid a urologist in making a decision on whether a stone is treatable with extracorporeal shock wave lithotripsy (ESWL) as opposed to the need for surgery [68,61,70]. Furthermore, 3D assessment allows follow-up of stones that are growing or reducing during any conservative treatment. These models could be combined with patient-specific anatomical data, and clinical data (using, for example, an electronic health record (EHR)) to begin to develop predictive models for the management of nephrolithiasis [62,63].

14. Conclusion

This review evaluated the transformation of working paradigms in kidney stone detection through convergence of medical imaging and computational intelligence, highlighting comprehensive state-of-the-art evidence along with research gaps in this interdisciplinary field. Given the increasing world-wide incidence of nephrolithiasis and the limitations of existing diagnosis methodologies, there is a strong impetus for diagnostic tools that are high-throughput, sensitive and scalable. State-of-the-art machine learning, deep learning and hybrid computational intelligence methods have shown promise of improving diagnostic accuracy, minimizing human error as well as early detection. These methods have the benefits of speed, reproducibility and ease of generalization, while being capable of managing large medical imaging datasets. Deep learning models, especially convolutional neural networks (CNNs) and their variants perform well on classification and segmentation tasks, often strengthened by hybrid approaches. Also, using preprocessing, feature extraction and dimensionality reduction approach allows the model to be robust and reliable. Nevertheless, the majority of previous models are validated only on small datasets and lack external validation, limiting their generalizability to clinical practice. In addition, issues with data availability, interpretability of models and readiness for deployment remain obstacles to clinical uptake in any meaningful way. Although computational intelligence approaches have been very helpful in improving the detection of kidney stones, follow-up studies should emphasize the development of stable and interpretable as well as clinically applicable systems. Closing that gap between research and real-life clinical deployment will be key to deliver consistent, reproducible, scalable and readily adoptable AI-based diagnostic pathways in urology.

All the Declarations and Statements

Author Contributions Statement

Karthick. P. – Conceptualization, Methodology, Writing – Drafted the initial manuscript, contributed to the literature survey, and documented the technical background of the study.
Chiranji Lal Chowdhary – Writing – Review and Editing, and Project Management: Reviewed and edited the manuscript, and Supervision: Proposed research ideas, Constructed the overall framework, and supervised project execution.

All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

The authors declare no conflicts of interest.

Funding Declaration

No funding received.

Data Availability Statement
Not Applicable.

Ethical Declarations
Not Applicable.

Acknowledgments
We sincerely thank the experts for their professional evaluation and valuable recommendations, which have contributed to improving the quality of the experiment and the reliability of its results.

Declaration of Generative AI in Scholarly Writing
The authors sincerely acknowledge the use of AI tools, which contributed to editing and improving the readability of this article. However, the content, ideas, and conclusion presented are entirely the work of the authors, reflecting their original research and insights.

Abbreviations
The following abbreviations are used in this manuscript:

CAD **-** Computer Aided Diagnosis
ML **-** Machine Learning
DL **-** Deep Learning
MRI **-** Magnetic Resonance Imaging
NCCT **-** Non-contrast Computed Tomography
CI **-** Computational Intelligence
ANN **-** Artificial Neural Network
SVM **-** Support Vector Machine
AHE **-** Adaptive Histogram Equalization
ROI **-** Region of Interest
GLCM **-** Gray-Level Co-occurrence Matrix

Appendix
Not Applicable.

References

[1] BagaRao, S., & Ghista, D. N. (2023). Chapter 9 - Renal Physiology Theory and Functional Assessment. In Biomedical Engineering of Pancreatic, Pulmonary, and Renal Systems, and Applications to Medicine (pp. 543–572). Elsevier, Inc. https://doi.org/10.1016/B978-0-323-95884-4.00009-3

[2] Ahmed, F., Abbas, S., Athar, A., Shahzad, T., Khan, W. A., Alharbi, M., Khan, M. A., & Ahmed, A. (2024). Identification of kidney stones in KUB X-ray images using VGG16 empowered with explainable artificial intelligence. Scientific Reports, 14, 6173. https://doi.org/10.1038/s41598-024-56478-4

[3] Alaskar, H., Hussain, A., Almaslukh, B., Vaiyapuri, T., Sbai, Z., & Dubey, A. K. (2022). Deep Learning Approaches for Automatic Localization in Medical Images. Computational Intelligence and Neuroscience, 2022, 1–17. https://doi.org/10.1155/2022/6347307

[4] Alelign, T., & Petros, B. (2018). Kidney Stone Disease: An Update on Current Concepts. Advances in Urology, 2018, 3068365. https://doi.org/10.1155/2018/3068365

[5] Alghamdi, H. S., & 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. https://doi.org/10.3390/diagnostics14131343

[6] 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... https://doi.org/10.1016/j.artmed.2022.102274

[7] Almuayqil, S. N., El-Ghany, S. A., Abd El-Aziz, A. A., & Elmogy, M. (2024). KidneyNet: A Novel CNN-Based Technique for the Automated Diagnosis of Chronic Kidney Diseases from CT scans. Electronics, 13(24), 4981. https://doi.org/10.3390/electronics13244981

[8] Nimavat, A., Trivedi, A., Yadav, A., & Patel, P. (2022). A review on kidney stone and its herbal treatment. Journal of Pharmacy and Pharmacology, 10, 195-209. https://doi.org/10.17265/2328-2150/2022.06.003

[9] Babajide, R., Lembrikova, K., Ziemba, J., Ding, J., Li, Y., Fermin, A. S., Fan, Y., & Tasian, G. E. (2022). Automated

machine learning segmentation and measurement of urinary stones on CT scan. Urology, 169, 41–46. https://doi.org/10.1016/j.urology.2022.07.029.
[10] Gill, K. S., Anand, V., Chauhan, R., Rawat, R. S., & Gupta, R. (2023). Kidney Disease X- Ray Image Classification Using a ResNet50V2 Model based Machine Learning Approach. 1–5. https://doi.org/10.1109/incoft60753.2023.10425099
[11] Priyadharsini, R., Balaji, V., & T, S. S. (2024). Deep Learning-Based Renal Stone Detection: A Comprehensive Study and Performance Analysis. Applied Computer Systems, 29(1), 112–116. https://doi.org/10.2478/acss-2024-0014
[12] Caglayan, A., Horsanali, M. O., Kocadurdu, K., Ismailoglu, E., & Guneyli, S. (2022). Deep learning model-assisted detection of kidney stones on computed tomography. International Brazilian Journal of Urology, 48(5), 830–839. https://doi.org/10.1590/S1677-5538.IBJU.2022.0132
[13] Cao, J., Li, F., & Huang, B. (2021). Active contour and texture fusion for CT-based kidney stone detection. Multimedia Tools and Applications, 80(3), 4423–4440.
[14] Huang, Z., Liu, Y. Y., Wu, W.-J., & Huang, K.-W. (2023). Design and Validation of a Deep Learning Model for Renal Stone Detection and Segmentation on Kidney–Ureter–Bladder Images. Bioengineering, 10. https://doi.org/10.3390/bioengineering10080970
[15] Choudhury, A., Ramesh, D., & Verma, N. (2025). KidStoneNet: Multi-angle CT dataset for robust kidney stone detection [Preprint]. arXiv.
[16] da Cruz, L. B., Araújo, J. D. L., Ferreira, J. L., Diniz, J. O. B., Silva, A. C., de Almeida, J. D. S., ... & Gattass, M. (2020). Kidney segmentation from computed tomography images using deep neural network. Computers in Biology and Medicine,123,103906. https://doi.org/10.1016/j.compbiomed.2020.103906
[17] Cui, Y., Sun, Z., Ma, S., & Yang, K. (2021). Automatic detection and S.T.O.N.E. nephrolithometry scoring of kidney stones on NCCT with deep learning. Molecular Imaging and Biology, 23(3), 436–445. http://dx.doi.org/10.1007/s11307-020-01554-0
[18] Dong, S., Wang, P., & Abbas, K. (2021). A survey on deep learning and its applications. Computer Science Review, 40, 1–22. https://doi.org/10.1016/j.cosrev.2021.100379
[19] Kethineni, L., Nithinchandra, N., Kumar, N., & Sk, S. S. (2025). Chronic Kidney Disease Prediction Based On Machine Learning Algorithms. In ITM Web of Conferences (Vol. 74, p. 01004). EDP Sciences. https://doi.org/10.1051/itmconf/20257401004
[20] Dutta, S., Chandra, S., & Paul, K. (2022). Kidney stone detection from CT using a lightweight CNN for edge devices. Journal of Medical Systems, 46(5), 53.
[21] Elsheimy, H., & Abdel-Fatah, H. (2024). Explainable AI models for kidney stone segmentation in non-contrast CT. Artificial Intelligence in Medicine, 169, 102369.
[22] 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 non-contrast CT scans. Medical Physics, 49(4), 2545–2554. https://doi.org/10.1002/mp.15518
[23] Goel, R., & Jain, A. (2020). Improved Detection of Kidney Stone in Ultrasound Images Using Segmentation Techniques (pp. 623–641). Springer, Singapore. https://doi.org/10.1007/978-981-15-0694-958
[24] Ghosh, T., & Sen, D. (2023). Comparative analysis of machine learning and deep learning models for CT-based kidney stone detection. International Journal of Medical Informatics, 176, 105090.
[25] Nithya, A., Appathurai, A., Venkatadri, N., Ramji, D. R., Palagan, C. A. (2020). Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images. Measurement, 149, 106952.
[26] He, L., Shen, G., & Zhao, D. (2023). Dual-stream CNN for detecting multiple kidney stones in single CT scans. IEEE Access, 11, 45267–45279.
[27] Islam, M., & Mehedi, M. H. (2021). CT kidney dataset: normal, cyst, tumor and stone. Data in Brief, 37, 107395.
[28] Jadhav, S., Joshi, M., & Pawar, A. (2025). AI-powered early detection of kidney stones using a hybrid CNN-LSTM model. Journal of Biomedical and Allied Sciences, 8(1), 44–51. http://doi.org/10.52783/jns.v14.2644
[29] Jalbani, A. H., Shaikh, T. A., & Memon, K. H. (2024). An improved convolutional neural network model for accurate kidney stone classification from CT images. Computerized Medical Imaging and Graphics,
[30] Kaur, G., & Singh, S. (2025). Automated detection of kidney stones and their characteristics in kidney ultrasound images. International Journal of Intelligent Systems and Applications in Engineering, 13(2), 89–95.
[31] Kaur, M., Gianey, H. K., Singh, D., & Sabharwal, M. (2019). Multi-objective differential evolution based random forest for e-health applications. Modern Physics Letters B, 33(5), 1–13. https://doi.org/10.1142/S0217984919500222
[32] Khan, A., Das, R., & Parameshwara, M. C. (2022). Detection of kidney stone using digital image processing: A holistic approach. Engineering Research Express, 4(3), 035040. https://doi.org/10.1088/2631-8695/ac8b65
[33] Alam, F., Alam, K., & Khan, N. (2025). An Automatic Approach for The Detection and Segmentation of Kidney Stone in Kub CT Images Using Mask R-CNN. Technical Journal of University of Engineering & Technology Taxila, 30(1).
[34] Billah, M. M., Al Rakib, A., Haque, M. I., Ahamed, A. S., & Hossain, M. (2024). S., Borsha KN (2024) Real- Time Object Detection in Medical Imaging Using YOLO Models for Kidney Stone Detection. European Journal of Computer Science and Information Technology, 12(7), 54-65. https://doi.org/10.37745/ejcsit.2013/vol12n75465
[35] Kim, H. S., Kim, E. J., & Kim, J. Y. (2023). Emerging trends in artificial intelligence-based urological imaging technologies and practical applications. International Neurourology Journal, 27(Suppl 2), S74–S81. https://doi: 10.5213/inj.2346286.143.
[36] Setia, S. A., Stoebner, Z. A., Floyd, C., Lu, D., Oguz, I., & Kavoussi, N. L. (2023). Computer vision enabled segmentation of kidney stones during ureteroscopy and laser lithotripsy. Journal of Endourology, 37(4), 495-501. https://doi.org/10.1089/end.2022.0511
[37] Kumar, A., & Choi, Y. I. (2025). Optimized YOLOv5 architecture for superior kidney stone detection in CT scans. Electronics, 14(4), 750. https://doi.org/10.3390/electronics13224418

[38] Lai, P. C., Wang, H.C., & Chang, K.H. (2023). Design and validation of a deep learning model for renal stone detection and segmentation on kidney–ureter–bladder (KUB) images. Bioengineering, 10(8), 970. https://doi.org/10.3390/bioengineering10080970
[39] Lang, J., et al. (2022). Global trends in incidence and burden of urolithiasis from 1990 to 2019: An analysis of Global Burden of Disease study data. European Urology Open Science, 35, 37–46. https://doi.org/10.1016/j.euros.2021.10.008
[40] Li, D., Xiao, C., Liu, Y., Chen, Z., Hassan, H., Su, L., Liu, J., Li, H., Xie, W., Zhong, W., & Huang, B. (2022). Deep segmentation networks for segmenting kidneys and detecting kidney stones in unenhanced abdominal CT images. Diagnostics, 12(8), 1788. https://doi.org/10.3390/diagnostics12081788
[41] Kang, L., Zhou, Z., Huang, J., & Han, W. (2022). Renal tumors segmentation in abdomen CT Images using 3D-CNN and ConvLSTM. Biomedical Signal Processing and Control, 72, 103334. https://doi.org/10.1016/j.bspc.2021.103334
[42] Liang, X., Cheng, J., & Yin, H. (2020). Elliptical blob detection and segmentation of kidney stones using active contours. Biomedical Engineering Online, 19(1), 72.
[43] 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. doi: 10.1109/ACCESS.2024.3351178.
[44] Martin, V., Lopez-Tiro, F., Larose, C., Ochoa-Ruiz, G., & Daul, C. (2025). Assessing generalization of Segment Anything Model for ureteroscopy scene understanding. arXiv preprint arXiv:2505.17210.
[45] Moradi, E., Rahmati, M., & Rafsanjani, Z. (2022). Segmentation-based kidney stone volume assessment using U-Net on CT images. Computer Methods and Programs in Biomedicine, 209, 106374.
[46] Mua'ad, M., & Zubi, M. (2020). Analysis and implementation of kidney stones detection by applying segmentation techniques on computerized tomography scans. Italian Journal of Applied Mathematics, 43, 590–602.
[47] K. S, L. R, V. A. CP & G. V, "Enhancing Kidney Stone Recognition in Ultrasound Imaging using Hybrid Feature Extraction Techniques," 2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI), Coimbatore, India, 2024, pp. 1-5, doi: 10.1109/ICoICI62503.2024.10696452.
[48] Odenrick, A., Kartalis, N., Voulgarakis, N., Morsbach, F., & Loizou, L. (2019). The role of contrast-enhanced computed tomography to detect renal stones. Abdominal Radiology, 44(2), 652–660. https://doi.org/10.1007/s00261- 018-1778-7
[49] Omer, A. M., & Mohamed, R. A. (2024). Automated detection of kidney stones using deep CNNs from CT and ultrasound fusion. Electronics, 13(22), 4418.
[50] Patro, K. K., Allam, J. P., Neelapu, B. C., et al. (2023). Application of Kronecker convolutions in deep learning technique for automated detection of kidney stones with coronal CT images. Information Sciences, 640, https://doi.org/10.1016/j.ins.2023.119005
[51] Rahmani, A. M., et al. (2021). Machine learning in medicine: Review, applications, and challenges. Mathematics, 9(22), 1–52. https://doi.org/10.3390/math9222970
[52] Rao, T. V. N., Gaddam, A., Kurni, M., & Saritha, K. (2022). Reliance on artificial intelligence, machine learning and deep learning in the era of Industry 4.0. In Smart Healthcare Systems Design. https://doi.org/10.1002/9781119792253.ch12
[53] Rathva, H. D., & Degadwala, S. M. (2024). Kidney stone classification using ensemble fine-tuned VGG-16 on CT images. Multimedia Tools and Applications. https://doi.org/10.32628/IJSRST2512359
[54] Ravisankar, P., Balaji, V., & Hameed, S. H. T. (2024). Deep learning-based renal stone detection: A comprehensive study and performance analysis. Applied Computer Systems, 29(1), 112–116.https://doi.org/10.2478/acss-2024-0014
[55] Sassanarakkit, S., Hadpech, S., & Thongboonkerd, V. (2022). Theranostic roles of machine learning in clinical management of kidney stone disease. Computational and Structural Biotechnology Journal, 20, 6487–6494. https://doi.org/10.1016/j.csbj.2022.12.004
[56] Liu, Y. Y., Huang, Z. H., & Huang, K. W. (2022). Deep learning model for computer-aided diagnosis of urolithiasis detection from kidney–ureter–bladder images. Bioengineering, 9(12), 811. https://doi.org/10.3390/bioengineering9120811
[57] Tefik, T., Ergul, R. B., Günceler, O., Aydinoglu, A. T., Panthier, F., Ozcan, F., & Traxer, O. (2025). Mp06-11 Development and evaluation of a machine learning-based approach for real-time kidney stone detection and segmentation in robotic flexible ureterorenoscopy. The Journal of Urology, 213(5s), e146. https://doi.org/10.1097/01.ju.0001109752.81014.18.11
[58] Teng, Y., Liu, X., Pan, Z., & Wang, Y. (2020). Improved kidney stone detection on low-dose CT using U Net++ with attention gates. IEEE Journal of Biomedical and Health Informatics, 24(12), 3524–3532.
[59] Huang, Z. H., Liu, Y. Y., Wu, W. J., & Huang, K. W. (2023). Design and validation of a deep learning model for renal stone detection and segmentation on Kidney–Ureter–Bladder images. Bioengineering, 10(8), 970. https://doi.org/10.3390/bioengineering10080970
[60] Rapelang, S. L., & Obagbuwa, I. C. (2025). Hybrid Support Vector Machine-Convolutional Neural Networks Multi-Classification Models for Detection of Kidney Stones. International Journal of Imaging Systems and Technology, 35(4), e70128.
[61] Valizadeh, H., & Rezaei, S. (2022). Real-time kidney stone classification using fast Fourier feature extraction. Journal of Medical Systems, 46(3), 17. Available via PubMed Central.
[62] Viswanath, K., Anilkumar, B., & Gunasundari, R. (2022). Design of deep learning reaction–diffusion level set segmentation approach for automatic kidney stone detection – CT analysis. Multimedia Tools and Applications, 81(29), 41807–41849.
[63] Wang, F., Shen, Y., & Liu, C. (2023). Landmark-based detection of kidney stones for 3D volume reconstruction. Frontiers in Neurorobotics, 17, 115.
[64] Sheikhi, M., Sina, S., Karimipourfard, M., & Shoushtari, F. K. (2023). A study toward automatic identification of renal

stone composition in single-energy or ultra-low-dose CT scan using deep neural networks. Iran J Radiol, 20(2), e134454. https://doi.org/10.1016/j.compbiomed.2021.104569

[65] Chen, Z., Liu, Y., Hassan, H., Li, D., Liu, J., & Zhong, W. (2024). Comprehensive 3D Analysis of the Renal System and Stones: Segmenting and Registering Non-Contrast and Contrast Computed Tomography Images. Information Systems Frontiers. https://doi.org/10.1007/s10796-024-10485-y.

[66] Xin, C., Zhang, S., & Wang, Z. (2020). Multi-scale fractal analysis and deep learning for kidney stone recognition in CT. Journal of Clinical Imaging Science, 10, 88.

[67] Yazdi, R., Valizadeh, H., & Masoumi, H. (2021). Kidney stone measurement using autonomous segmentation and shape fitting on CT images. Computer Methods and Programs in Biomedicine, 202, 105933

[68] 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. https://doi.org/10.1016/j.compbiomed.2021.104569

[69] Yin, S., Peng, Q., Li, H., Zhang, Z., You, X., Fischer, K., ... & Fan, Y. (2020). Automatic kidney segmentation in ultrasound images using subsequent boundary distance regression and pixelwise classification networks. Medical image analysis, 60, 101602.https://doi.org/10.1016/j.media.2019.101602

[70] Zhang, D., López-Cruz, I., & Sánchez-González, J. (2021). U-Net-based segmentation for calcium oxalate kidney stones in CT. Journal of Imaging, 7(8), 154.

[71] Tang, Z., Tang, Z., Liu, Y., Chen, L., Tang, Z., Liu, X., & Salami, T. (2025). An optimized bidirectional recurrent neural network for kidney stone detection based on developed bald eagle search method in CT scan images. *Scientific Reports*, *15*(1), 37109. https://doi.org/10.1038/s41598-025-21103-5

[72] Degadwala, S., & Rathva, V. (2024). A review on kidney stone detection using ML and DL techniques. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, *10*(5), 101-112. https://doi.org/10.32628/CSEIT2410589

[73] Abdalla, P. A., Shakor, M. Y., Ameen, A. K., Mahmood, B. S., & Hama, N. R. (2025). Kidney stone detection via axial CT imaging: A dataset for AI and deep learning applications. *Data in Brief*, *59*, 111446. https://doi.org/10.1016/j.dib.2025.111446

[74] Rao, N. J., Girish, H., Gowrishankar, M. C., Kumar, S., & Kumar, N. (2025). A Two-Stage Deep Learning Framework for Kidney Stone Detection and Clinical Severity Grading in CT Imaging. *Informatics in Medicine Unlocked*, 101704. https://doi.org/10.1016/j.imu.2025.101704

[75] Panchal, N., Raikar, M. M., & Baligar, V. P. (2025). Kidney stone detection using deep learning model. *Procedia Computer Science*, *260*, 56-63. https://doi.org/10.1016/j.procs.2025.03.177

[76] Bala, S., Arora, K., & Doss, A. N. (2025). Improving Kidney Stone Detection with YOLOV10 and Channel Attention Mechanisms in Medical Imaging. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *7*(3), 951-963. https://doi.org/10.35882/jeeemi.v7i3.868

[77] Satsangi, A., Jain, S., & Upadhyay, S. Optimizing Kidney Stone Detection: Exploring YOLOv8 Variants for Computational Efficiency and Enhanced Accuracy. DOI: 10.5815/ijisa.2025.05.01

[78] Huang, H., Huang, Y., Du, X., Pang, J., Nong, S., Pan, X., ... & Lin, X. (2025). Contrastive Swin Transformer-Based Classification Model for Internet of Medical Things-Driven Kidney Stone Diagnosis. IEEE Internet of Things Journal. https://doi.org/10.1109/JIOT.2025.3625941

[79] Kasaraneni, C. K., Guttikonda, K., & Madamala, R. (2025). Multi-modality Medical (CT, MRI, Ultrasound Etc.) Image Fusion Using Machine Learning/Deep Learning. In Machine Learning and Deep Learning Modeling and Algorithms with Applications in Medical and Health Care (pp. 319-345). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-98728-1_16

[80] Pradhan, U., Chattopadhyay, H., Aditya, M. N., Patil, M., Baig, S. A. H., Mittal, S., & CJ, L. V. (2025, April). Explainable AI for Kidney Stone Prediction: A Machine Learning-Based Clinical Analysis. In 2025 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI) (pp. 1-6). IEEE. https://doi.org/10.1109/RAEEUCCI63961.2025.11048291

Authors' Profiles

Karthick P. is currently pursuing his Ph.D. at Vellore Institute of Technology (VIT), Vellore. He completed his M.Phil. in 2018 and M.Sc. in 2016 from Dr. MGR Chockalingam Arts College, Arni, and obtained his B.Sc. degree in 2014 from Muthurangam Government Arts College. He also holds professional teaching qualifications, including a B.Ed. (2020) from St. John's College of Education and an M.Ed. (2022) from Sri Renugambal College of Education. He has served as an Assistant Professor at DLR Arts and Science College, Arcot, contributing to teaching and academic activities. His research interests include data analysis, machine learning, and educational methodologies, and he is actively engaged in academic research and publications.

Chiranji Lal Chowdhary is a Professor in the School of Computer Science and Engineering (SCORE) at Vellore Institute of Technology (VIT) University, where he has been serving since 2010. He received his Bachelor of Engineering (B.E.) in Computer Science and Engineering from MBM Engineering College (now MBM University), Jodhpur, and his Master of Technology (MTech.) in Computer Science and Engineering from M.S. Ramaiah Institute of Technology, Bangalore. He earned his Ph.D. in Information Technology and Engineering from VIT University, Vellore. From 2006 to 2010, he worked at M.S. Ramaiah Institute of Technology, Bangalore, where he served as a Lecturer. His research interests include computer vision and image processing, with a strong focus on pattern recognition, machine learning, biometric systems, deep learning, soft computing, and computational intelligence. Dr. Chowdhary is the editor/co-editor of eight books and the author of over forty research articles in the field of computer science. He has also filed four patents based on his research contributions. He has been recognized by the Austrian Science Fund (FWF), Vienna, Austria, as an International Reviewer for long-term funded research projects. Additionally, he is involved as a Co-Principal Investigator (Co-PI) in an ongoing sponsored research project funded by the British Council under the SPARC initiative.

How to cite this paper:

Karthick P., Chiranji Lal Chowdhary, "Kidney Stone Detection Using Medical Imaging and Computational Intelligence - A Review", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.18, No.4, pp. 139-164, 2026. DOI:10.5815/ijigsp.2026.04.08