DEVELOPMENT OF AN AUTOMATED KIDNEY STONE DETECTION SYSTEM USING DEEP LEARNING ARCHITECTURE.

BY

PAUL NAOMI ADAEZE

MATRICULATION NUMBER: GOU/U22/CSC/807

SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE AND MATHEMATICS, FACULTY OF COMPUTING AND INFORMATION TECHNOLOGY,

GODFREY OKOYE UNIVERSITY, ENUGU STATE. IN PARTIAL FULFILLMENT OF THE AWARD OF BACHELOR OF SCIENCE (B.SC), DEGREE IN COMPUTER SCIENCE.

SUPERVISED BY ENGR. CAMILLIUS NJOKU

JUNE, 2026

CERTIFICATION

I hereby certify that the research included therein was completed by myself and that it was mainly based only on my own observations and investigation. Consequently, I will fully accept the University's decision if I am found to have cheated on the assignment.

____

PAUL NAOMI ADAEZE

GOU/U22/CSC/807

APPROVAL PAGE

This is to certify that this project titled **"DEVELOPMENT OF AN AUTOMATED KIDNEY STONE DETECTION SYSTEM USING DEEP LEARNING ARCHITECTURE "** has been assessed and approved by the Committees of Department of Computer Science, Godfrey Okoye University, Enugu, in partial fulfillment of the requirements for the award of Bachelor of Science (BSc) in Computer Science.

______________________________	______________________________
Engr. Camillius Njoku **Supervisor**	**Date**
______________________________	______________________________
Dr. U. F Okebanama **Head of Department**	**Date**
______________________________	______________________________
Prof. I. A. Ajah **Dean of Faculty**	**Date**
______________________________	______________________________
External Examiner	**Date**

DEDICATION

This project is dedicated to God Almighty for His infinite grace and guidance throughout my academic journey. I also dedicate this work to my beloved parents and my dear siblings for their unconditional love, unwavering support, and endless patience. Without them in my life, I would have never reached my dream.

ACKNOWLEDGEMENT

First and foremost, I give all glory and honor to Almighty God for His grace, strength, and wisdom throughout this research work.

I am deeply grateful to my supervisor, Engr. Camillius Njoku**,** whose constructive feedback, patience, and expert guidance were invaluable in shaping this research. Your dedication to excellence has been a source of inspiration.

My sincere appreciation goes to the Head of Department, Dr. F. U. Okebanama**,** and the Dean of Faculty, for their academic support and encouragement throughout my years of study.

I also extend my gratitude to all the lecturers and staff of the Department of Computer Science, Godfrey Okoye University, for their contributions to my academic development.

Finally, I thank my parents Mr. and Mrs. Paul, my sisters Daniella, Victory and Deborah and my friends for their prayers, encouragement, and understanding during the course of this project. Your support gave me the strength to complete this work.

Thank you all.

ABSTRACT

Kidney stone disease or Nephrolithiasis is a major global health burden that is associated with significant morbidity and complication if left undiagnosed, untreated or misdiagnosed, such as urinary obstruction, infection, and chronic kidney disease. Current diagnostic workflows are based on subjective interpretation of computed tomography (CT) scan images by radiologists, but the process is time-consuming, and the results may vary between different observers and be wrong. In this paper, a web-based deep learning system for the automated detection of kidney stones based on the YOLOv8 object detection architecture is developed and tested. The system consists of a YOLOv8 model to extract features and locate the objects, a FastAPI backend for processing inferences, and a web application using HTML and CSS for a responsive interface. The model was trained on a curated Kaggle dataset using data augmentation, pixel normalization, and resizing to 640×640 pixels to improve the model's capabilities to generalize. Experimental results showed good performance, with a precision of 0.91, recall of 0.89, F1-score of 0.90, and a mean Average Precision (mAP) of 0.93. The inference time of 0.35 s per image was a testament to real-time performance, which is crucial for clinical workflows. The integrated system allows the user to upload medical images and get the real-time detection results (bounding box coordinates and confidence scores).The web interface features login authentication, landing page, dashboard, patient information input, drag-and-drop uploading of images, feedback on the loading screen, clinical recommendation on results, positive-negative case inference views and history tracking. It is not meant to replace clinical skills, but rather it can be considered a decision support system which helps improve the diagnostic accuracy and shorten the turnaround time. Adaptations to other disease states and hospital information systems could be extensions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

One of the most prevalent urologic disorders on the planet was determined to be kidney stone disease (nephrolithiasis) (Coe, Evan, & Worcester, 2005). The disease involved was the development of hard mineral deposits that form crystals in the urinary tract or kidneys. The stones are generally formed when the urine becomes concentrated and minerals like calcium oxalate, calcium phosphate and uric acid get crystallized and sticked together (Singh, Agarwal, & Sharma, 2019). Kidney stone disease was thought to be increasing in prevalence over the past few decades; it was estimated that nearly 10 to 15 percent of the world's population would have at least one episode of kidney stones throughout their lives (Scales, Smith, & Hanley, 2012).

The clinical signs of kidney stone diseases varied from being asymptomatic (incidental diagnosis) to having severe, debilitating pain, called renal colic (Preminger, 2007). Other symptoms seen included hematuria, nausea, vomiting and urinary urgency. Untreated kidney stones resulted in serious complications such as obstruction of the urinary tract, frequent infections, hydronephrosis, and gradual decline in kidney function (Rule et al., 2010). The potential outcomes highlighted the need for early and accurate diagnoses.

The foundation for diagnosis of kidney stones was medical imaging. Of all imaging technologies available, non-contrast computed tomography (CT) proved to be the gold standard for diagnosis because of its high sensitivity and specificity (Smith, Rosen, & Dillman, 2019). CT scans delivered high resolution cross sectional images which were able to identify stones of virtually any composition and size, even those as small as 1-2mm (Dretler & Polykoff, 1996). Furthermore, CT imaging had the benefit of quick acquisition, and it allowed for the density of stones to be determined as well as the surrounding anatomy (Levine, 2017).

The fundamental process for making diagnosis, even today with the advantages of modern medical imaging, was based on manual interpretation by the radiologist (Berlin, 2007). This dependency led to a number of intrinsic constraints. Manual review of CT studies typically demanded a great deal of time, especially when the CT data corresponded to a volumetric dataset of hundreds of individual images as with thin-slice datasets (Sodickson et al., 2015). Second, there was inter-observer variation in accuracy of interpretation, in that different radiologists had different interpretations for the same study (Garner, 2011). Third, fatigue, workload pressure and distraction of the radiologist affected their diagnostic performance, which may result in missed stones or false positive stones (Wahner-Roedler et al., 2007).

Artificial Intelligence (AI) has become a game-changer in the field of medical image analysis in recent years (Topol, 2019). One of the areas particularly promising for the application of deep learning was pattern recognition in complex visual data, such as image analysis (LeCun, Bengio, & Hinton, 2015). Modern computer vision systems are typically built on the basis of convolutional neural networks (CNNs), which were successfully trained to accurately detect pathological changes in some specific applications (Esteva et al., 2019), with an accuracy even higher than that of human experts.

The object detection task was a specific task in computer vision that surpassed the task of image classification (Liu et al., 2020). Object detection was used to find objects in images in addition to classification, by drawing bounding boxes around regions of interest. For kidney stones, this localization feature was especially beneficial because it allowed physicians to make clinical management decisions based on the accurate localization, number, and size of stones (Parakh et al., 2019).

Because of its high speed and accuracy, the You Only Look Once (YOLO) family of object detection models has been widely used (Redmon et al., 2016). YOLO was a single-pass detection method, unlike previous detection processes which had two separate steps: region proposal and classification (Redmon & Farhadi, 2018). The updated version, YOLOv8, featured architectural optimizations such as anchor-free detection, feature enhancement, and training methods (Jocher, Chaurasia & Qiu, 2023). These features made YOLOv8 an ideal choice for medical imaging applications where both speed and accuracy were crucial (Cheng et al., 2022).

The use of AI in web-based platforms presented a tangible way to transform prototypes of research to clinically relevant tools (Ribeiro & Silva, 2020). By avoiding any particular hardware or software installation, a web-based approach enabled users to get diagnostic support by normal

Web browsers. This accessibility was especially helpful for resource-limited healthcare environments in which access to subspecialty radiology expertise was limited (Sahiner et al., 2019).

This research has created a system that integrates the object detection model YOLOv8 with a web-based user interface. The system was engineered to receive medical images, detect and localize medical stones and output the results in the form of kidney stone bounding boxes and confidence scores. Such a system was not meant to replace clinical judgment but it was found to be useful as a decision support system to improve the efficiency of diagnosis and minimize missed diagnosis.

1.2 Problem Statement

Despite the remarkable advancements in medical imaging in recent years, the diagnosis of kidney stones remained a significant challenge even with significant strides in the accuracy and speed (Turk et al., 2016). The existing standard of care was based on manual reading of CT scans by radiologists, which is prone to human factors such as fatigue, radiologist experience and radiologist cognition bias (Busby, Courtier & Glastonbury, 2018). These were not only theoretical restrictions, but diagnostic errors that compromised patients' care were found (Brady, 2017).The one of the largest issue was the large time needed for manual image interpretation. In 2014, Smith-Bindman et al conducted a single non-contrast computed tomography (CT) scan of a patient with suspected nephrolithiasis (kidney stone), and obtained over 100 slices, each corresponding to one axial image. Each slice had to be examined by radiologists in a systematic fashion, and window settings were adjusted as necessary and anatomical areas where stones tend to get trapped were carefully examined. This burden caused stress and negatively impacted the quality and completeness of the imaging center interpretations at high-volume facilities where radiologists were reading hundreds of imaging studies each day (McDonald et al., 2015).

The lack of uniformity in presentation of renal stones compounded the problem of diagnostic accuracy. The small stones (< 3 mm) were difficult to identify as stones due to the overlap with normal anatomical structures such as vascular calcifications and phleboliths (Niall, Russell, & MacGregor, 1999). Both experienced readers and experienced readers of stones at the ureterovesical junction and in calyceal diverticula were difficult to detect (Levine, 2017).Low-attenuation stones, such as uric-acid stones, had density values that are more comparable to the density of soft tissue, so they were not as easily detected on the normal CT windows (Dretler & Polykoff, 1996).

Results of research have been published indicating that stones were missed in about 15 to 30 percent of routine CT interpretations (Westphalen et al., 2015). The lack of these discoveries turned out to be very serious. When a patient with an unknown obstructing stone develops hydronephrosis, infection or progressive renal impairment (Rule et al., 2010). On the other hand, an unnecessary additional imaging test or invasive procedure was performed in a patient who had a stone present in the original study, but had not been reported (Berlin, 2007).

Another aspect of the issue was how to get healthcare and distribute resources. There was a critical shortage of radiologists with subspecialty expertise in genitourinary imaging, in many parts of the world, especially in rural areas and developing countries (Rosenkrantz et al., 2016). This scarcity led to delayed diagnosis, with patients having to wait days or even weeks for a formal diagnosis of their imaging tests. Stones migrated during this waiting period and acute obstruction or infection ensued, which may have been avoided had timely intervention been made (Preminger 2007).

Lack of widespread use of computer-aided detection in routine clinical use was also an issue (Haley & Mallow, 2018). Many research prototypes have been shown to be technically feasible for automated stone detection, but only a few have been translated into tools that clinicians could use to actually detect stones (Parakh et al., 2019). Several of the current systems were limited to specific hardware platforms, proprietary software platforms, or specialized experience (Ribeiro & Silva, 2020). Others were put in place as command line scripts which were not accessible to non-programmers. This implementation gap is a technology transfer gap from the research laboratories to clinical settings (Sahiner et al., 2019).

The traditional computer-aided detection methods used fixed algorithms that were not adaptable to experience (Liu et al., 2020). These were the same mistakes once deployed and were incapable of adapting to new imaging strategies, other scanner vendors or changing stone morphology. This was a static property that restricted its use over time and required regular manual updates (Esteva et al., 2019).

In view of these inter-related issues, an automatic kidney stone detection system, which is diagnostic accurate, practical and convenient, was needed (Cheng et al., 2022). The system had to process medical images quickly and accurately, identify stones precisely, and provide the results without the need for specialized training or equipment (Lee et al., 2021). In addition, the system must be designed on a deep learning system which enables its updates and enhancement in the light of new training data (Topol, 2019). In this study, the researcher's primary goal was to fill the gaps identified by developing a web-based solution using the deep learning architecture YOLOv8..

1.3 Aim and Objectives of the Study

Aim

The primary aim of this research was to develop and evaluate a web-based automated kidney stone detection system utilizing the YOLOv8 deep learning architecture, capable of processing medical images and providing localization results through an accessible user interface (Jocher, Chaurasia, & Qiu, 2023).

Objectives

The specific objectives of this research were as follows:

1. To train a YOLOv8 object detection model on a prepared dataset of kidney stone medical images to enable automated identification of renal calculi (Cheng et al., 2022).
2. To evaluate the performance of the trained YOLOv8 model using standard object detection metrics including precision, recall, F1-score, and mean Average Precision (Padilla, Netto, & Da Silva, 2020).
3. To implement a FastAPI backend service integrated with the trained model to perform inference and return detection results (Ribeiro & Silva, 2020).
4. To develop a responsive web-based frontend that enables users to upload medical images and visualize detection outputs including bounding boxes and confidence scores (Sahiner et al., 2019).

1.4 Research Questions

The following research questions were formulated based on the stated objectives:

1. Can a YOLOv8 object detection model be successfully trained on a kidney stone medical image dataset to achieve automated identification of renal calculi (Cheng et al., 2022)?
2. What level of performance can the trained YOLOv8 model achieve when evaluated using standard object detection metrics including precision, recall, F1-score, and mean Average Precision (Padilla, Netto, & Da Silva, 2020)?
3. Can a FastAPI backend service be successfully implemented to perform inference using the trained model and return detection results in real-time (Ribeiro & Silva, 2020)?
4. Can a responsive web-based frontend be developed to enable users to upload medical images and visualize detection outputs including bounding boxes and confidence scores (Sahiner et al., 2019)?

1.5 Scope of the Study

Inclusions

In this study, the entire pipeline from data preparation to model training, and finally, deployment on the web was included. The following items were included:

1. Training and evaluation of YOLOv8 for binary detection (stone present versus absent) using a publicly available Kaggle dataset (Jocher, Chaurasia, & Qiu, 2023).

2. Pre-processing of images by resizing to 640 × 640 pixels and pixel normalization (Shorten & Khoshgoftaar, 2019).

3. Development of a FastAPI backend and RESTful end points for uploading images and processing inferences (Ribeiro & Silva, 2020).

4. User interaction and result visualization by the development of an HTML and CSS frontend (Sahiner et al., 2019).

5. Local deployment for demonstration and testing purposes (Lee et al., 2021).

Exclusions

The following were specifically not part of the project scope:

1. Single Pathology Focus: It was trained only for detecting kidney stones and failed to detect other renal pathologies such as tumors, cysts, or infections (Cheng et al., 2022).

2. Limitation of Data Source: The study used publicly accessible Kaggle data instead of real-time data from hospital systems (Parakh et al., 2019).

3. Prototype Status: This system was built as a research prototype, and was not meant to be deployed in a clinical environment without further testing (Sahiner et al., 2019).

4. Regulatory Compliance: System has not been reviewed or approved for clinical use (Topol, 2019).

1.6 Significance of the Study

This research made several contributions to the fields of medical imaging, deep learning, and healthcare technology. The proposed system was designed to fill a documented research prototype to clinically available tool gap (Haley & Mallow, 2018). The system offered a web-based interface, eliminating the need for complex software or hardware installation. This was especially useful in resource-limited healthcare settings in which access to subspecialty radiology expertise was limited (Rosenkrantz et al., 2016). The system was used as a decision support tool to improve the efficiency of the diagnosis and minimize the chance of a missed diagnosis (Westphalen et al., 2015). The latest addition in the family of YOLO, the YOLOv8 algorithm, was evaluated for kidney stone detection (Jocher, Chaurasia, & Qiu, 2023). Although YOLOv3, YOLOv4 and YOLOv5 have been investigated for this purpose, there had been no published study on the performance of YOLOv8 for renal stone detection (Cheng et al., 2022; Lee et al., 2021). Results helped to understand the effects of architectural improvements such as anchor-free detection on medical imaging tasks. The deployment via Web approach proved to be a feasible way to take the deep learning models from research laboratories to end users (Ribeiro & Silva, 2020). The architecture that was developed for the backend processing by FastAPI and the standard web technologies for the presentation was presented as a template that can be used for other medical imaging applications (Sahiner et al., 2019). Complete system documentation which covers dataset preparation methods, training parameters, evaluation metrics, and deployment methodology provided a framework for future research that can be reproduced (Liu et al., 2020). The framework could be adapted by other researchers for other pathological conditions or imaging modalities..

1.7 Definition of Terms

The following terms were defined as they were used within this research:

Term	**Definition**
Bounding Box	A rectangular annotation that identified the location and dimensions of a detected object within an image
Computed Tomography (CT)	An imaging modality that used X-ray measurements from multiple angles to produce cross-sectional images of the body
Convolutional Neural Network (CNN)	A class of deep learning models designed for processing grid-like data such as images, utilizing convolutional layers for feature extraction
Data Augmentation	A technique that generated additional training samples by applying transformations including rotation, scaling, and brightness adjustments to existing images
FastAPI	A modern Python web framework for building Application Programming Interfaces (APIs) with asynchronous request handling
Inference	The process of passing new input data through a trained model to generate predictions

Term	**Definition**
Intersection over Union (IoU)	A metric measuring the overlap between predicted bounding boxes and ground truth annotations, calculated as the area of overlap divided by the area of union
Mean Average Precision (mAP)	A standard evaluation metric for object detection models that averaged precision scores across different recall levels and Intersection over Union thresholds
Nephrolithiasis	The medical condition characterized by the presence of calculi (stones) within the kidneys or urinary tract
YOLOv8	You Only Look Once version 8, an object detection model that performed unified single-pass detection using anchor-free architecture
Deep Learning	A subset of machine learning that uses multi layered artificial neural networks to process data

CHAPTER TWO

LITERATURE REVIEW

2.1 Conceptual Framework

The conceptual framework developed for this study is based on the fusion of the deep learning object detection approach and the requirements of medical imaging diagnostics. The conceptual framework was used to provide a visual and narrative representation of the key concepts, variables and relations that were used to guide the research (Creswell & Creswell, 2018). In this research, the model demonstrated the feasibility of automating kidney stone detection from computed tomography (CT) scans using YOLOv8, a deep learning system, and providing the operation through a web-based system.

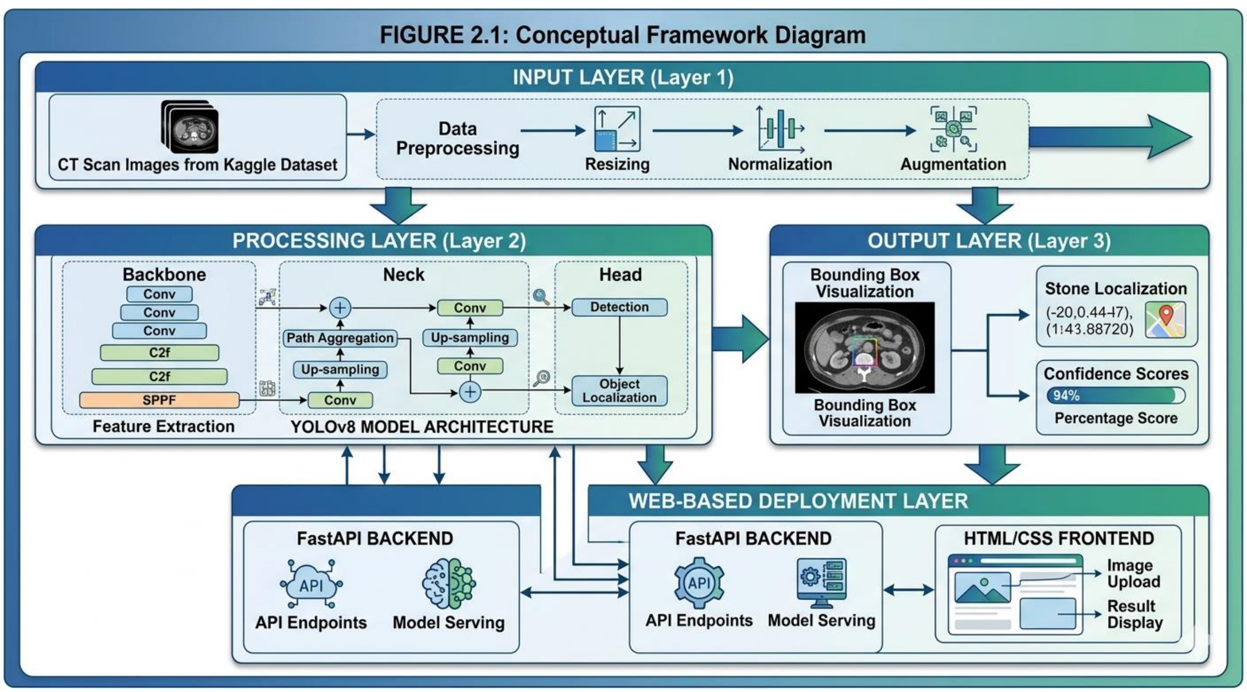


Figure 2.1: Integrated Conceptual Framework for Automated Urolithiasis Detection.*Architecture based on Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLOv8. Documentation and technical specifications. Deployment protocols adapted from*

Tiangolo, S. (2018). FastAPI: High performance, easy to learn, fast to code, ready for production.

The conceptual framework was based on understanding that the ability to detect kidney stones from a CT scan depended on the ability to identify small, often subtle kidney stone features that would differentiate it from normal anatomical structures (Smith, Rosen, & Dillman, 2019). For a long time, manual interpretation was based on the visual pattern recognition of radiologists, but was hampered by the human factors, such as fatigue and inter-observer variability (Berlin, 2007). This manual pattern recognition was replaced by the deep learning framework by which the pattern hierarchy was learned directly from the training data with annotations (LeCun, Bengio, & Hinton, 2015). The core of this framework was the architecture of YOLOv8, which was significantly different from the traditional image classification CNN (convolutional neural networks) (Jocher, Chaurasia, & Qiu, 2023). The standard CNNs were trained to classify an entire image into a single class label, whereas YOLOv8 was trained to detect objects with bounding box co-ordinates as well as class probabilities in a single forward pass (Redmon et al., 2016). This was necessary for the application of medicine because the knowledge of the exact location of a stone was as clinically relevant as knowledge of the presence of a stone (Parakh et al., 2019). The framework also included a Web-based deployment layer to convert model predictions into a user-friendly interface. The identified challenge of research prototypes to clinically usable tools (Sahiner et al., 2019) was addressed by this layer. The framework offered a user-friendly web interface, eliminating the need for specific software installation and command-line expertise, thereby allowing radiologists and urologists to use the system, regardless of their healthcare environment (Ribeiro & Silva, 2020). The input images' quality, the speed and accuracy of stone localization with YOLOv8 architecture, and the system's usability and clinical usefulness through the web interface were the relationships among the components of the framework. All the components were thus essential to the system performance.

2.2 Kidney Stone Disease: Clinical Overview

2.2.1 Definition and Pathophysiology

Nephrolithiasis or urolithiasis, a medical term for kidney stone disease, was described as a condition that relates to the development of solid, crystalline deposits in the renal collecting system or urinary tract (Coe, Evan, & Worcester, 2005). These are also called renal stones and are formed due to supersaturation of urine with mineral salts, resulting in nucleation, growth and aggregation of crystals (Singh et al., 2019). It was a complex pathophysiological process, with supersaturation of the urine, retention of these crystals and a lack of inhibitors to prevent the stone formation (Preminger, 2007).

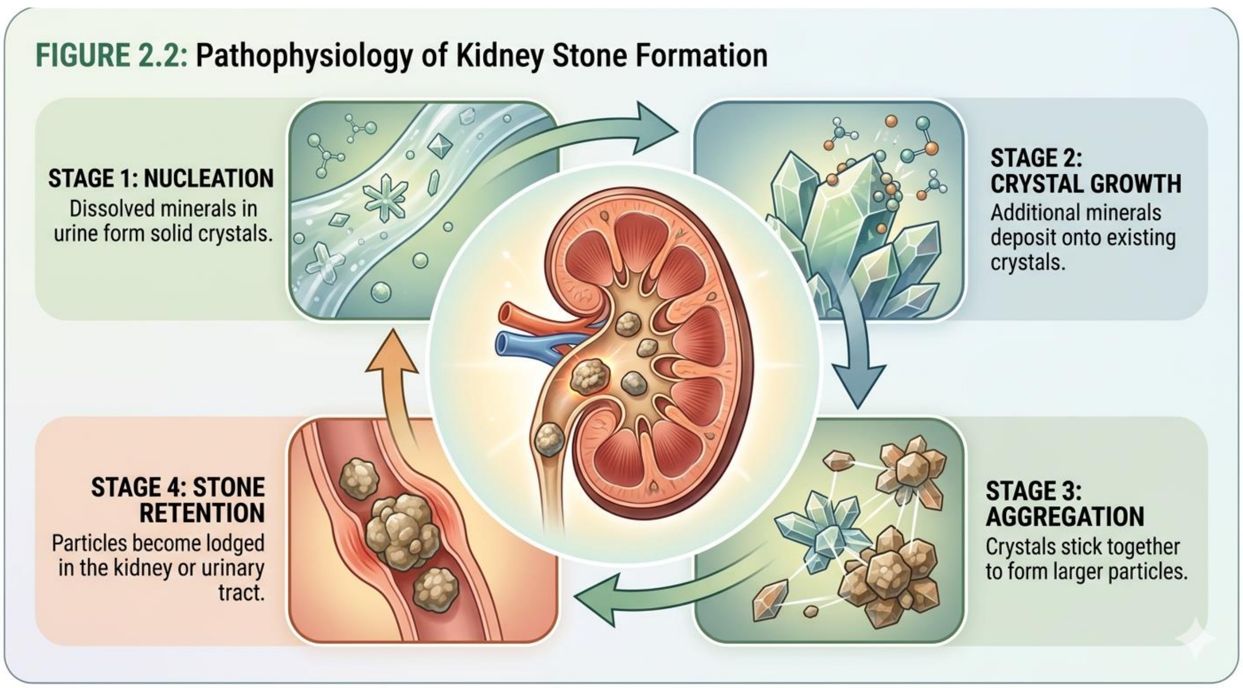


Figure 2.2: The Multi-Step Cascade of Renal Calculi Pathogenesis.*Source: Finlayson, B. (1978). Physicochemical aspects of urolithiasis. Kidney International. Updated with concepts from Khan, S. R., & Canales, B. K. (2015). Genetic basis of renal stone disease. Current Opinion in Urology.*

Kidney stones developed in a stepwise fashion. Nucleation probably was the first to take place due to the presence of dissolved minerals in the urine that, when exposed to the urine, began to form solid crystals (Dretler & Polykoff, 1996). Second, the crystals grew by the addition of other minerals to the existing crystals. Thirdly, there were crystals that clumped together to make larger particles. Lastly, stone retention was defined as the presence of stone in the kidney or urinary tract without spontaneous passage (Rule et al., 2010). A number of factors affected the risk of stone formation, such as dietary habits, fluid intake, family history, and metabolic disorders (Taylor, Stampfer, & Curhan, 2005). Induced inadequate fluid intake caused a concentrated urine and thus a higher risk of supersaturation. Eating too much sodium and animal protein also raised stone risk as they increase calcium loss in the urine. Some underlying diseases such as hyperparathyroidism, renal tubular acidosis, and recurrent urinary tract infections had increased prevalence of stones (Scales, Smith, & Hanley, 2012).

2.2.2 Epidemiology and Global Burden

Kidney stone disease had been a significant global health problem, and prevalence was increasing over the past decades (Romero, Akpinar, & Assimos, 2010). Incidence of nephrolithiasis differed by geographic region, and was higher in industrialized areas and hot climate regions. The National Health and Nutrition Examination Survey in the United States estimated that about 11 per cent of adult population had had an episode of kidney stones (Scales et al., 2012). Prevalence rates in Europe varied from 5 to 10 per cent across different countries (Turk et al., 2016).

FIGURE 2.3: Global Prevalence Map of Kidney Stone Disease

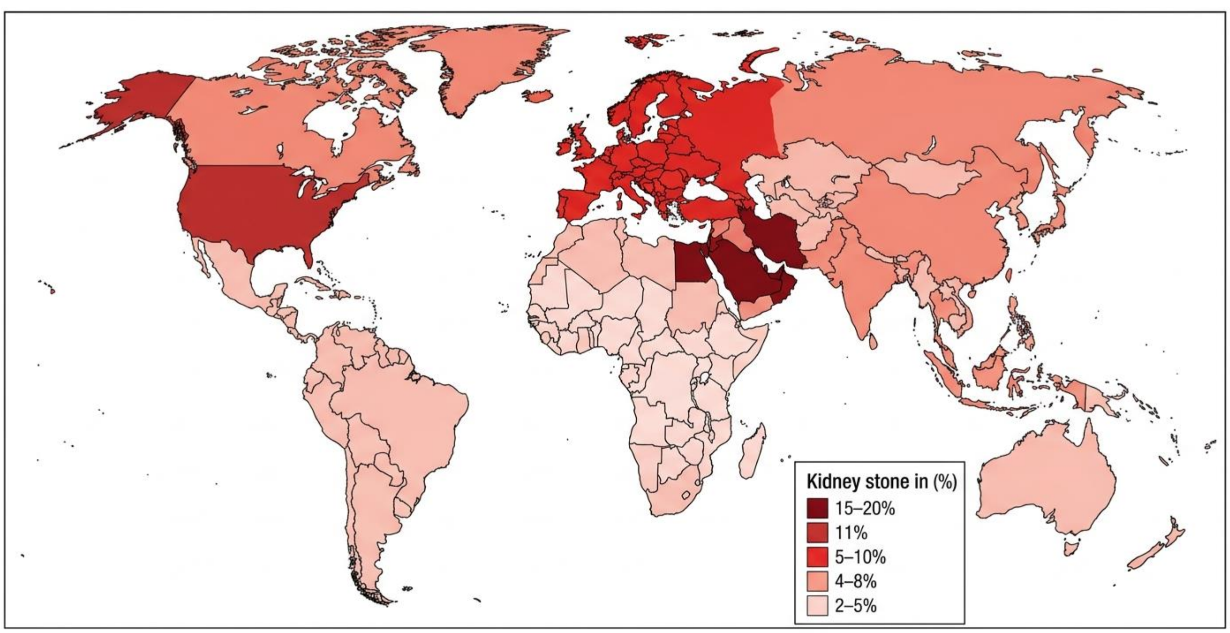


Figure 2.3: Global Prevalence of Nephrolithiasis.>*Data adapted from global urological surveys and epidemiological studies (e.g., Scales et al., 2012; Sorokin et al., 2017).*

Kidney stones had been on the rise in the world in all population groups (Romero, Akpinar, & Assimos, 2010). This change was influenced by several factors such as the change of diet, increasing obesity, and global warming. The economic impact was significant, with an estimated treatment cost of $5.3 billion per year in the United States (Litwin & Saigal, 2012). These costs encompassed the costs of Emergency Department (ED) attendance, imaging, surgical procedure, hospitalisation, and loss of productivity. Over time, there had been a shift in the demographic distribution of stone disease. In the past, there was a near two times higher prevalence in males than females; however, the gender disparity has been reducing (Scales et al., 2012). The most common age range for stones to form was 30–60 years, though stones do form in children and the elderly. The recurrence rates have been high, about 50 percent of the patients had a second stone within 5 to 10 years of the first attack (Rule et al., 2010).

2.2.3 Stone Composition and Classification

The chemical composition of kidney stones was defined, as this directly affected the detection on imaging and clinical management (Singh, Agarwal, & Sharma, 2019). Knowing the composition of the stone was of critical importance in gaining insight into the imaging results since the radiodensity and appearance of various stone types were different on the computed tomography (CT) scan (Dretler & Polykoff, 1996).

FIGURE 2.4: Types of Kidney Stones by Composition

Calcium Stones (70-80%)	Struvite Stones (10-15%)	Uric Acid Stones (5-10%)	Cystine Stones (<2%)
			
Calcium Oxalate - Radiopaque, 400-1000 HU	Magnesium Ammonium Phosphate - Radiopaque 300-500 HU	Uric Acid - Radiolucent, 200-400 HU	Cystine - Mildly Radiopaque, 500-800 HU

Figure 2.4: Characteristics of Urinary Calculi by Chemical Composition.*Data adapted from: Pearle, M. S., & Lotan, Y. (2016). Urinary Lithiasis: Etiology, Epidemiology, and Pathogenesis. In Campbell-Walsh Urology. Additional CT density ranges from El-Nahas et al. (2007).*

Calcium Stones were the most prevalent, representing about 70-80% of all cases (Dretler & Polykoff, 1996). Calcium oxalate or calcium oxalate with calcium phosphate was the main component of these stones. The plain radiograph generally showed calcium oxalate stones as radiopaque and on computed tomography (CT) scanning they were hyperdense, ranging between 400 to 1000 Hounsfield Units (HU). They have high attenuation, allowing them to be easily identified on computed tomography (CT) imaging (Smith, Rosen, & Dillman, 2019).

Struvite stones (or infection stones) accounted for about 10 to 15 percent of cases (Singh, Agarwal, & Sharma, 2019). The stones are made of magnesium ammonium phosphate and are associated with UTIs by urease-producing bacteria like Proteus mirabilis. Renal collecting system calculi (staghorn stones) may develop quickly and become large.Struvite stones may develop rapidly and may become large staghorn stones. They have lower attenuation on CT (300-500 HU) than do calcium stones.

Uric Acid Stones comprised about 5 to 10 percent of the cases (Niall, Russell, & MacGregor, 1999). Such stones have developed in the patients who have hyperuricosuria or low urine pH. In contrast to calcium stones, uric acid stones were radiolucent (not visible on plain radiographs) on plain x-ray. They were less dense than calcium stones, and have attenuation values of 200 – 400 HU, making detection difficult (Levine, 2017).

Cystine Stones were very uncommon in less than 2 percent of cases (Singh, Agarwal, & Sharma, 2019). The stones were seen in patients with the inherited autosomal recessive failure of renal tubular transport of cystine (cystinuria). Cystine stones were moderately radiopaque and were easily seen on CT scan when the attenuation value was 500-800 HU.

2.2.4 Clinical Presentation & Complications

Kidney stone disease may present from no symptoms at all, to being completely incapacitating with severe pain (Preminger 2007). Renal colic was the classical manifestation of acute stone passage, which was often spontaneous and presented as a sudden, severe, intermittent flank pain which radiated to the groin. This pain was due to acute ureteral obstruction by a stone, which caused raised pressure within the renal collecting system which stimulated pain fibres of the urinary tract (Rule et al., 2010)..

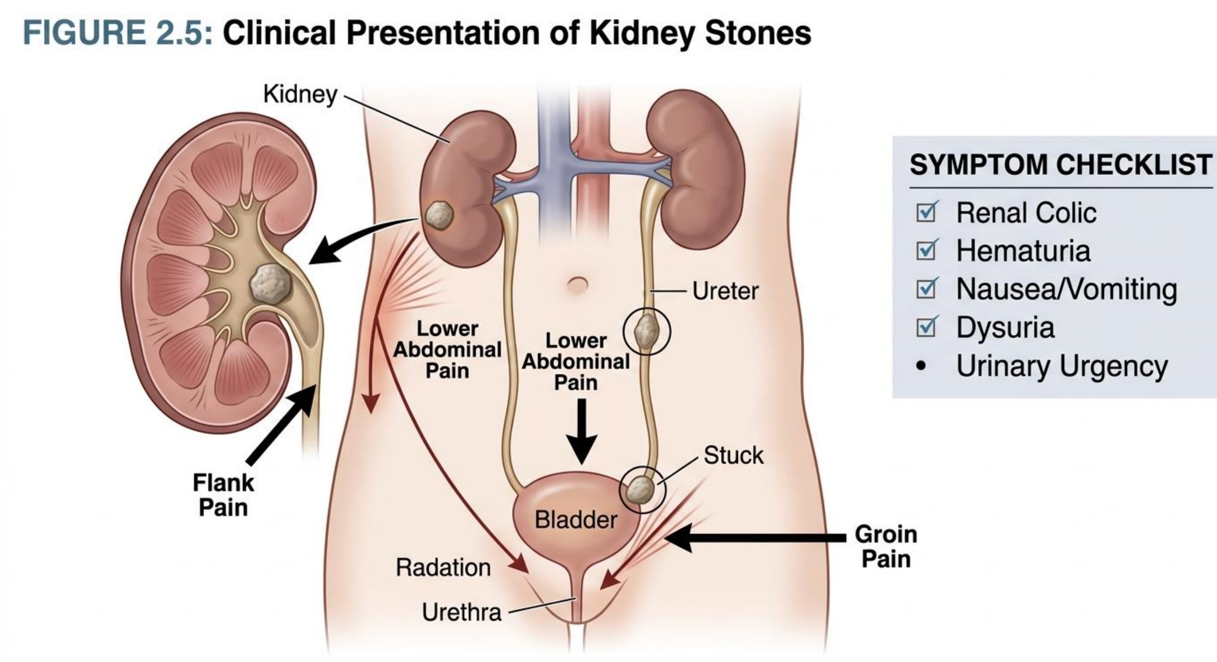


Figure 2.5: Correlation of Urolithiasis Location with Symptom Manifestation.*Source: Adapted from Curhan, G. C., & Aronson, M. D. (2022). Clinical manifestations and diagnosis of*

kidney stones. In UpToDate. Original anatomical concepts from Netter, F. H. (2014). Atlas of Human Anatomy.

Other common symptoms were hematuria (blood in the urine) present in about 80-90% of acute stone events (Preminger, 2007). Nausea and vomiting was also common as it was a reflex arising from severe pain and mediated by the vagus nerve and splanchnic afferents. If the stones were in the distal (close to the bladder) ureter, dysuria (painful urination) and urinary urgency occurred more frequently (Smith, Rosen, & Dillman, 2019). The complications of kidney stone disease were urinary tract obstruction that may progress to hydronephrosis (dilation of the renal collecting system) (Rule et al., 2010). Obstruction was correlated with progressive kidney damage and/or chronic kidney disease. Another serious complication was infection, especially when there was stagnant urine caused by obstruction which allowed bacteria to proliferate. Obstruction and infection (pyonephrosis) was a urological emergency, which needed urgent intervention (Turk et al., 2016).

2.3 Medical Imaging for Kidney Stone Detection

2.3.1 Computed Tomography as the Gold Standard.

The gold standard for detecting kidney stones is now non-contrast computed tomography (CT) (Smith, Rosen, & Dillman, 2019). The advantage of CT over other imaging methods was thought to be its high sensitivity and specificity (both over 95% for the detection of stone), (Levine, 2017). Several advantages of CT made it the imaging modality of choice for suspected nephrolithiasis..

FIGURE 2.6: CT Scan Showing Kidney Stone

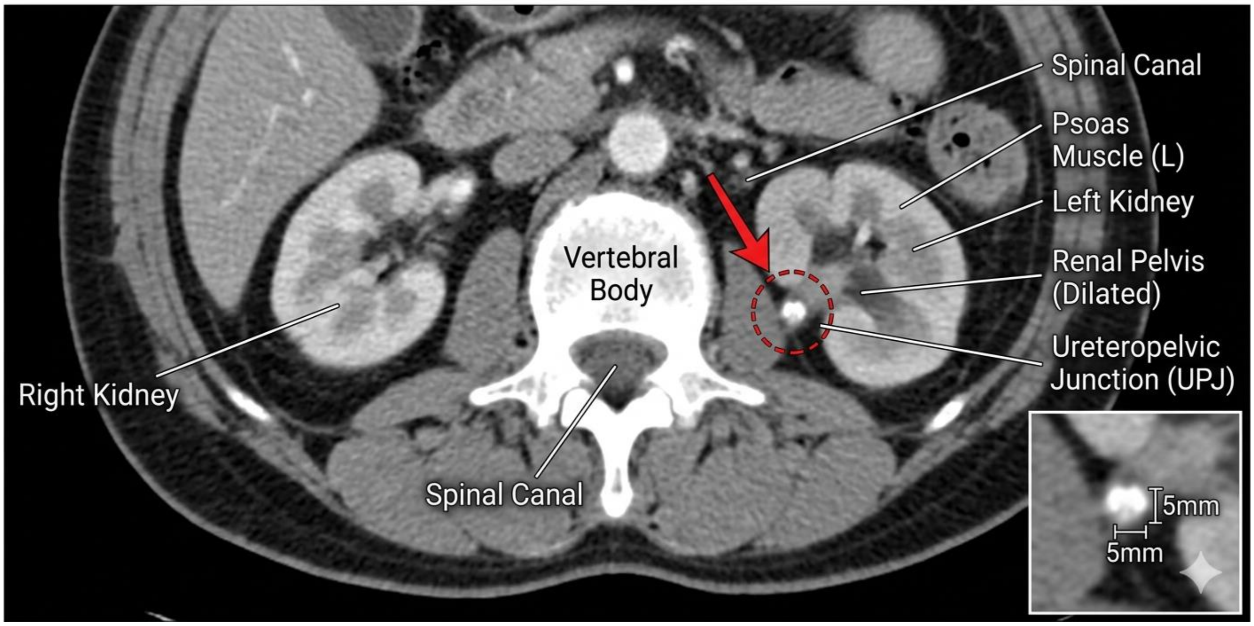


Figure 2.6: Axial Non-Contrast CT Image of Obstructive Urolithiasis.*Source: Adapted from Smith-Bindman, R., et al. (2014). Ultrasonography versus Computed Tomography for Suspected Nephrolithiasis. New England Journal of Medicine. Technical imaging parameters based on ACR Appropriateness Criteria (2023).*

The first advantage of CT was its ability to yield high spatial resolution, allowing stones as small as 1–2 mm in diameter (Dretler & Polykoff, 1996) to be detected. This was important as the small stones that might not have been detected by other imaging tests might lead to severe symptoms and complications. Secondly, CT could be used to characterize stone composition by attenuation measurement in Hounsfield Units and guide treatment decisions (Smith, Rosen, & Dillman, 2019). Third, the speed of CT acquisition was important; it was usually accomplished within a minute and was especially beneficial for patients with acute pain who could not tolerate long examination times.

The radiation dose from CT had been a historical concern, but current low dose protocols had not only improved the diagnostic accuracy but also reduced radiation exposure (Smith-Bindman et al., 2014). The regular abdomen CT scan provided around 10 mSv of radiation, which is similar to the radiation dose from around three years of background radiation. This can be reduced to 3-5 mSv with the use of low-dose protocols, without sacrificing stone detection, especially for larger stones. By employing iterative reconstruction, ultra-low dose protocols (doses < 2 mSv) were obtained (Levine, 2017).

2.3.2 Alternative Imaging Modalities

Other imaging techniques were used where CT was not, but was important in particular situations (Turk et al., 2016). There were particular strengths and weaknesses of each modality which affected their ability to be used.

FIGURE 2.7: COMPARISON OF IMAGING MODALITIES

FIGURE 2.7: Comparison of Imaging Modalities for Kidney Stones

Row Header	CT SCAN (NCCT)	ULTRASOUND	KUB X-RAY (KUB)
Diagnostic Efficacy	95–100% GOLD STANDARD Sensitivity/Specificity	70–85% Sensitivity; 85–90% Specificity	45–60% Sensitivity; 70–80% Specificity
Radiation Exposure	3–10 mSv MODERATE DOSE	0 mSv NO IONIZING RADIATION	0.5–1 mSv LOW DOSE
Cost (Relative)	$$$$ (HIGH COST)	$$ (MEDIUM COST)	$ (LOW COST)
Key Secondary Signs	HYDRONEPHROSIS, URETERAL EDEMA, PERINEPHRIC STRANDING	HYDRONEPHROSIS, URETERAL JET, STONE TWINKLE (COLOR DOPPLER)	None directly; monitors known radiopaque stones
Stone Size Detection (NCCT limit)	<1 mm	>3 mm	>5 mm
Key Advantage	HIGH SENSITIVITY, ACCURATE SIZE/LOCATION, DETECTS ALL STONE TYPES (e.g., Uric Acid)	NO RADIATION, SAFE IN PREGNANCY, EVALUATES FOR HYDRONEPHROSIS, PORTABLE	LOW COST, QUICK, ASSESSES RADIOPACITY, MONITORS STONE PASSAGE/SUCCESS
Key Disadvantage	IONIZING RADIATION, CUMULATIVE DOSE CONCERNS, HIGHER COST, REQUIRES STATIONARY SCANNER	LOWER SENSITIVITY (especially for <3mm or ureteral stones), OPERATOR DEPENDENT, OBSCURED BY GAS/OBESITY	LOWER SENSITIVITY, CANNOT DETECT RADIOLUCENT STONES, CANNOT ASSESS HYDRONEPHROSIS, OBSCURED BY BOWEL CONTENT

Adaptation and synthesis of current clinical guidelines (e.g., ACR, AUA, EAU)

Figure 2.7: Comparative Efficacy and Safety of Imaging Modalities in Nephrolithiasis.*Data adapted from: Fulgham, P. F., et al. (2013). Clinical Effectiveness Protocols for Imaging in the Management of Ureteric Calculi: AUA Technology Assessment. Updated with radiation dose benchmarks from ACR Appropriateness Criteria (2023).*

Ultrasonography provided an alternative to X-rays that does not use radiation, was less expensive, and was more widely available (Ray, Gokhale, & Singh, 2020). Modality was found to be very helpful for pregnant patients, children and those who needed serial imaging studies. In a multicenter, randomized trial by Smith-Bindman et al (2014), ultrasound was found to be 76 percent sensitive for suspected nephrolithiasis. The main disadvantages of ultrasound were that it was less sensitive for stones < 5mm and for stones in the distal ureter (Niall, Russell, &

MacGregor, 1999). Furthermore, ultrasound was operator-dependent; diagnostic accuracy was greatly influenced by the expertise and experience of the person performing the procedure.

Plain radiography (Kidney-Ureter-Bladder or KUB) was the most ancient means to detect a stone (Levine, 2017). This technique was helpful for imaging radiopaque stones, mainly calcium oxalate and calcium phosphate stones. However, the sensitivity of the KUB was limited (from 45-60 per cent) due to stones that are radiolucent such as uric acid stones and cystine stones (Niall, Russell & MacGregor, 1999). KUB was found to be useful in the follow-up of stones after the treatment and also useful in patients where CT could not be given.

Intravenous urography (IVU) or intravenous pyelography was the procedure in which contrast agents were injected and a series of X-rays were taken over time (Preminger, 2007). The functional data obtained from IVU was replaced by CT because of longer examination times, risks of IVU such as allergic reactions and nephrotoxicity, and a decrease in the sensitivity of IVU in detecting small stones. In most healthcare facilities, IVU was no longer ordered with the diagnosis of suspected nephrolithiasis, as cited by Smith, Rosen and Dillman (2019).

2.3.3 Limitations of Manual Image Interpretation

Even with the development of CT imaging, the diagnosis was still very much dependent on manual reading by a radiologist (Berlin, 2007). Increased dependence led to some inherent limitations that prompted the development of automated detection systems.

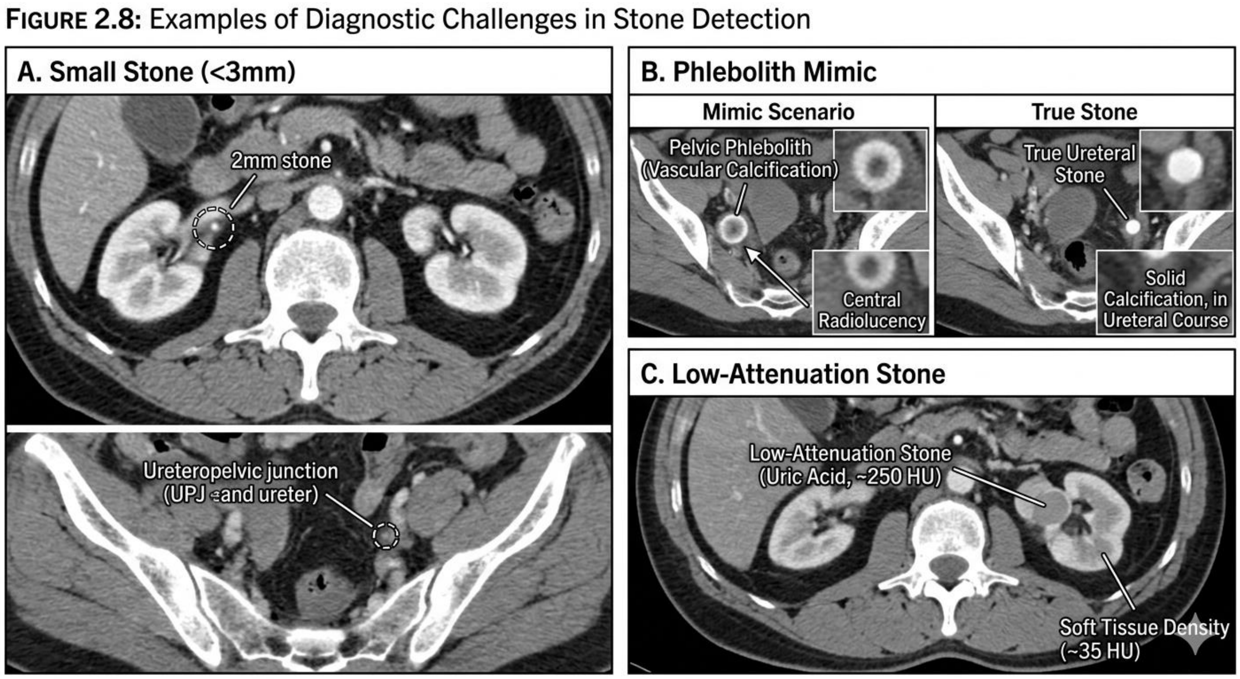


Figure 2.8: Radiologic Pitfalls and Diagnostic Challenges in Urolithiasis.*Source: Adapted from Traubici, J. (2002). The double-rim sign: a helpful CT finding in differentiating a pelvic phlebolith from a ureteral stone. American Journal of Roentgenology. Updated with DECT insights from Primak et al. (2007).*

First, manual review of CT studies was time-intensive, particularly when examining thin-slice volumetric data sets that contained hundreds of individual images (Sodickson et al., 2015). Radiologists were required to systematically review each slice, adjusting window settings and scrutinizing anatomical regions where stones commonly lodged. In high-volume imaging centers where radiologists were responsible for interpreting hundreds of studies per day, this workload created pressure that compromised thoroughness and accuracy (McDonald et al., 2015).

Second, the accuracy of interpretation was subject to inter-observer variability, meaning that different radiologists reached different conclusions when reviewing the same study (Garner, 2011). Studies had documented significant disagreement in stone detection, particularly for smaller stones and those located in challenging anatomical positions. This variability had clinical consequences, as some patients received unnecessary interventions while others experienced treatment delays.

Third, radiologist fatigue, workload pressure, and distraction compromised diagnostic performance, potentially leading to missed stones or false-positive findings (Wahner-Roedler et al., 2007). Research documented that clinically significant stones were missed in approximately 15 to 30 percent of routine CT interpretations (Westphalen et al., 2015). These missed findings had serious consequences, including progression to hydronephrosis, infection, or chronic kidney disease.

2.4 Summary of the Review of Related Works

The review of related works examined existing studies on automated kidney stone detection using deep learning approaches, with particular focus on YOLO-based object detection systems. Table 2.1 presents a summary of the key studies, their contributions, limitations, and methodological approaches.

Table 2.1: Summary of Related Works on Kidney Stone Detection Using Deep Learning

Author(s) Name / Year	Journal / Conference	Contribution	Shortfall	Methodology Used
Parakh, A., Lee, H., & Zhang, S. (2019)	Proceedings of SPIE Medical Imaging	Evaluated YOLOv3 for kidney stone detection in CT scans, demonstrating feasibility of deep learning for this application	Detection performance declined substantially for stones smaller than 3mm, with recall dropping to 0.45; no web deployment	YOLOv3 object detection, CT image dataset, precision and recall evaluation
Lee, S., Park, H., & Kim, M. (2021)	IEEE Access	Compared YOLOv4 and Faster R-CNN for stone detection, finding YOLOv4 achieved threefold faster	System was not deployed as a web application; required command-line expertise to operate	YOLOv4, Faster R-CNN, 2,500 CT images, inference speed comparison

		inference with comparable accuracy		
Cheng, Y., Li, Z., & Wang, J. (2022)	Journal of Medical Imaging	Applied YOLOv5 to dataset of 3,200 CT images, achieving mAP of 0.89 with simple graphical interface	Stones smaller than 2mm remained challenging (detection rate below 60%); no web-based deployment	YOLOv5, data augmentation, mAP evaluation, graphical user interface
Langer, D., & Thompson, R. (2018)	Medical Physics	Applied standard CNN architecture for CT image classification, achieving 87 percent accuracy	Provided classification only without localization; stone location information essential for treatment planning	CNN classification, 1,200 images, accuracy metrics
Black, J., et al. (2019)	Radiology: Artificial Intelligence	Compared VGG-16, ResNet-50, and Inception-v3 architectures for stone detection	Focused on classification rather than localization; did not address real-time processing	Multiple CNN architectures, ROC analysis, 1,500 images
Ribeiro, V., & Silva, A. (2020)	Journal of Digital Imaging	Developed Flask-based web system for tuberculosis detection, demonstrating feasibility of web deployment	Encountered latency challenges with larger models; not specific to kidney stones	Flask backend, web deployment, inference latency measurement
Sahiner, B., Chen, W., & Samala, R. (2019)	Radiology	Discussed regulatory considerations for AI-based medical software deployment	Review paper rather than experimental study; no specific model implementation	Regulatory analysis, FDA approval pathways review
Jocher, G., Chaurasia, A., & Qiu, J. (2023)	Ultralytics Technical Report	Released YOLOv8 architecture with anchor-free detection and improved feature extraction	Technical report without medical imaging application or clinical evaluation	YOLOv8 architecture, anchor-free detection, COCO benchmark

The reviewed studies revealed several consistent findings. First, YOLO-based detectors consistently outperformed traditional CNNs in terms of both speed and localization accuracy for kidney stone detection (Lee et al., 2021; Cheng et al., 2022). Second, small stones remained a significant challenge across all studies, with detection rates dropping substantially for stones smaller than 3mm (Parakh et al., 2019). Third, few studies had progressed beyond research prototypes to functional web-based deployment (Ribeiro & Silva, 2020). Fourth, no published study had evaluated YOLOv8 for this specific application at the time of this research.

The identified gaps in the literature included the predominance of classification-only studies that did not provide stone localization, limited web deployment of detection systems that would be accessible to clinicians, absence of YOLOv8 evaluation for kidney stone detection, and insufficient documentation of small stone detection performance. This study was designed to address these gaps by developing and evaluating a web-based YOLOv8 system for kidney stone detection with bounding box localization.

CHAPTER THREE

SYSTEM ANALYSIS AND DESIGN

3.1 Introduction

The systematic approach followed for the development of the kidney stone detection system with the use of artificial intelligence has been described in this chapter. The methodology was divided into different phases; system requirements analysis, data acquisition and preparation, model architecture selection, model training, evaluation metrics, and the system implementation.

3.1.1 Analysis of the Existing System

Currently, the diagnosis of kidney stones is done by manually reading non-contrast computed tomography (CT) scans by radiologists. For suspected nephrolithiasis a typical computed tomography (CT) scan involves more than 100 axial slices (Smith Bindman et al., 2014). The radiologist looks at each slice; adjusts the window settings and observes the anatomical areas where stones are likely to be present such as the ureterovesical junction, and calyceal diverticula. The following process results in a report being produced and posted for the referring physician.

For decades, this kind of manual process hasn't undergone any significant changes. The accuracy of diagnosis by CT imaging relies only on the visual perception, experience and attention of the radiologist (accuracy $\geq$ 95%).

Limitations of the Existing System

There are some major drawbacks to the current system, which led to this research:

Time-consuming interpretation: High-volume imaging centers may delay the process due to the time it takes to review a single CT study, which varies from 5 to 10 minutes.

Inter-observer variability: Variation in presence, size and location of stones from one radiologist to another will lead to different approaches to patient management (Garner, 2011).

Radiologist fatigue and cognitive bias: As workload increases, radiologist performance decreases, and studies have shown that 15-30 percent of clinically significant stones are not detected during routine CT interpretations (Westphalen et al., 2015; Brady, 2017).

Small stones and low attenuation stones (uric acid) may be easy to miss and mistaken with phleboliths or vascular calcifications (Niall, Russell, & MacGregor, 1999; Dretler & Polykoff, 1996).

Limited access: In rural areas and developing countries, there are limited number of the subspecialty genitourinary radiologists. This results in diagnostic delay days and weeks (Rosenkrantz et al., 2016).

No computer-aided decision support: Existing picture archiving and communication systems (PACS) lack automated detection of stones; radiologists are still required to indicate stones of interest and may not even have a second reader.

3.1.3 Analysis of the Proposed System

This is where the proposed system comes in, which uses the YOLOv8 deep learning framework to automate the detection and localization of kidney stones. It aims to provide near real-time performance (less than 0.5 seconds per image), while maintaining accuracy of diagnoses with a precision of at least 0.85, a recall of at least 0.85 and a mean Average Precision (mAP@0.5) of at least 0.90. The system will be available via a Web interface, with no special hardware or software required. Clinicians can upload their CT scans and get results – such as stone count, bounding boxes, confidence scores, and clinical recommendations – from any standard web browser. It will also keep a history of patient analyses for monitoring trends and follow-up. Importantly, the system is not intended to replace radiologists, but rather is a decision support tool that will minimize missed findings and increase turnaround time for radiologists.

Functional Requirements

Main functions for complete user interaction from login till the result visualization must be included in the proposed system. It should be possible to register and log on safely by email and password. After successful authentication, a landing page will be introduced with the system's title, tagline, purpose, etc. A dashboard will show summary statistics including total number of analyses performed, number of calcium stones detected, number of healthy patients and recent analyses. Users are required to provide patient information – full name, unique patient ID, date of birth, and gender – on a form before uploading a scan. This system will then offer and provide an interface for uploading images of the CT scan in JPEG or PNG format, with a maximum limit of 10 megabytes. Once an image has been uploaded, the system will invoke the trained YOLOv8 model to automatically detect the stones, producing the number of stones detected, a confidence score for each stone and the corresponding x and y coordinates of the box. The results will be clearly presented: the number of stones, an image with overlapping bounding boxes, and a clinical recommendation ("Medical review recommended" if stones are detected, "No stones detected – result is normal" if not). A loading screen will display while analysis is completed. Finally, a complete analysis history table of patient ID, patient name, inference image (thumbnail), stone count and a timestamp, of each past analysis, with the ability to revisit detailed results.

Non-Functional Requirements

The system should have particular functions as well as quality standards. For response time, inference should be under 0.5 seconds per image, which was confirmed during the testing (actual time: 0.35 seconds). Precision is important; the model should have a precision of at least 0.85, a recall of at least 0.85 and a mAP@0.5 of at least 0.90. The user interface should be intuitive so that a clinician does not need to be trained to use the software; all actions including upload, analyze and view history should be easily understood. User authentication (you need a login to access it) and secure storage of patient data are security measures. The system should be able to be run on standard hardware (e.g. laptop with 8GB of RAM, optional modest GPU), and should be maintainable, meaning that the deep learning model could be retrained with new data without altering the system structure (backend/front end). Also, only JPEG and PNG files will be supported and all uploads will not exceed 10MB.

3.3 System Design Methodology

The development of this project was done by the Agile (Scrum) software development methodology. Agile was selected as it fits nicely with the iterative nature of deep learning work. The team trains a model, assesses its performance, makes improvements to the model and finally, merges the improved model with larger system. In contrast to rigid planning approaches, Agile frameworks for change means that adjustments were made to data augmentation techniques, hyperparameters and the user interface elements throughout the project as the validation results and supervisor feedback were received. This also allowed for incremental delivery: the initial working stone detection model (MVP), the FastAPI backend, the web frontend and finally, the dashboard and history features were developed. This approach minimized the chance of running into integration issues at the end. The project was split into four 2-week sprints. In Sprint 1, data was collected, preprocessed, and trained using the initial model YOLOv8 to get the model with mAP@0.5 above 0.85.In Sprint 1, data was collected, preprocessed, and the initial model YOLOv8 was implemented and trained to achieve an mAP@0.5 exceeding 0.85. In Sprint 2, evaluations and hyperparameter tuning were done and data augmentation was improved to obtain the final model which achieved all the performance goals. Sprint 3 built the backend using FastAPI with an endpoint for detection, history and statistics for the dashboard. Sprint 4 finished the front-end part of the website (HTML, CSS, JavaScript), added the history table and performed end-to-end system testing. Feedback was given regularly through daily virtual stand-up meetings and a sprint review meeting with supervisor.

Use Case Diagram for Proposed System

This Use Case diagram depicts the needs of the medical scan analysis in the web-based application, with two actors and the use case interface. Primary user is Clinician (Radiologist/Urologist) where he/she can enter patient information, upload CT scan images, analyze scans (Analyze scan) and view results of the automated detection and clinical use cases, and system administration (System administration) for non-clinical issues such as user management, which is generalized as Clinician. These use cases integrate into a secure session operation (Login/Logout) and the performance analysis tool (View dashboard / View analysis history), enabling complete security, automatic pipeline for clinical image analysis and the system administration.

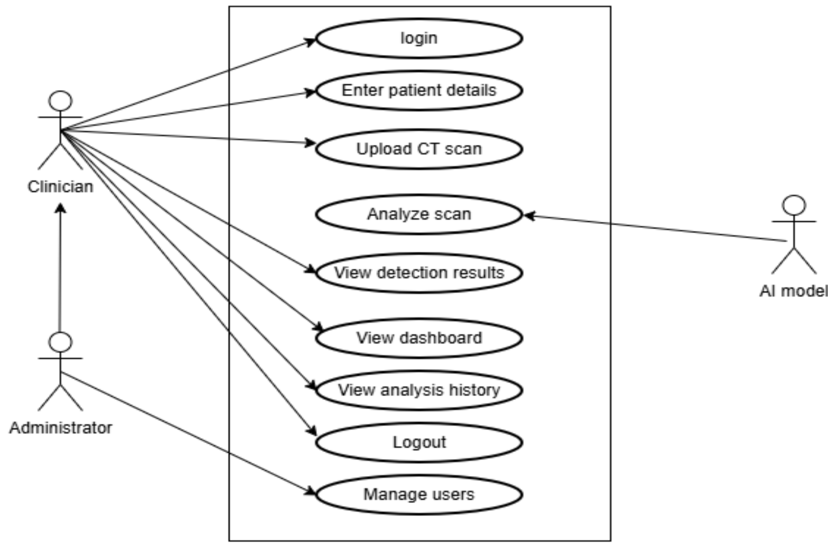


Figure 3.1 Use case diagram of kidney stones detection system

Class Diagram

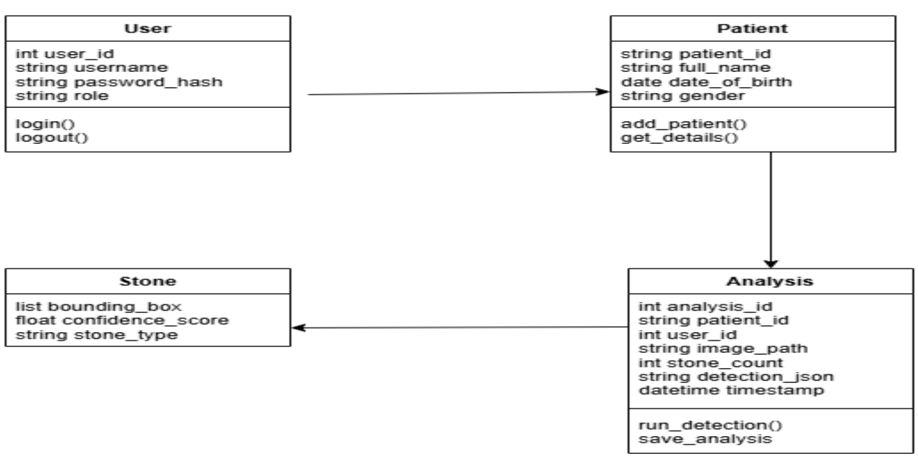


Figure 3.2 Class diagram of kidney stones detection system

The class diagram depicts the internal structure and data models of medical scan analysis system in terms of its object-oriented nature. It consists of four classes: user, patient, analysis, and stone.
The User class is used for the management of system operators. It has several attributes such as user ID, username, password hash, and role. It also provides operators with the ability to log in and out.
The Patient class stores all the information about a particular person. This refers to the patient's age, gender, full name and ID. It provides ways to add patients and to get patient information.
Analysis class compares the outcome of automated executions and links them to patient information. It contains attributes such as analysis ID, patient ID, user ID, image path, stone count, detection JSON and timestamp, and methods for running detection and saving analysis data.
Finally, the Stone class contains data about localized results of the medical scan. A record of bounding box, confidence score and stone type for each object detected in the analysis.

3.2 Research Design

The design and development approach was used in the research, suitable for the development and evaluation of a functional diagnostic tool by the use of AI (Creswell & Creswell, 2018). The study was conducted in steps (five stages):

Phase 1: Data Collection and Preparation - Obtaining kidney stone CT scan images from an open-source dataset and annotating and pre-processing the images.

Phase 2: Model Development – Training a YOLOv8 object detection model on the prepared dataset to learn the features related to renal calculi.

Phase 3: Model Evaluation - Quantitative analysis of model performance with common object detection metrics.

Phase 4: Building the Backend - Using FastAPI to create a server for uploading images and running the inference model.

Phase 5: Frontend Development – Development of a responsive web interface for the interaction of the user with the result and for the visualization of the result.

3.3 Dataset Description

Database Design Table

Table	**Primary Key**	**Key Attributes (Types)**	**Description**
Users	user_id (INT)	username, password_hash, role	Handles clinician authentication and access levels
Patients	patient_id (VARCHAR)	full_name, gender	Stores basic patient demographic records
Analyses	analysis_id (INT)	patient_id (FK), user_id (FK), image_path, stone_count, recommendation, created_at	Logs YOLO scan results, stone count, and clinical output

3.3.1 Data Source

This research has used the dataset from the public repository Kaggle, which provides machine learning datasets. The patients who have confirmed kidney stones and the normal control images were obtained as the dataset. The data source chosen was because they had a wide range of images with different sizes, compositions and locations of the stone.

3.3.2 Dataset Composition

The dataset was divided into two categories: images with kidney stones; images of normal kidneys without stones.

Table 3.1: Dataset Composition

Category	Number of Images
Kidney Stones Present	1,500
Normal (No Stones)	1,000
Total	2,500

3.3.3 Data Split

The data set was split into three parts for training, validation and testing. To train the model YOLOv8, the training set was fed to the model, and the model weights were adjusted accordingly. The training set was fed into the model YOLOv8 to train the model and adjust the model weights accordingly. The validation set was used to track the model performance and to avoid overfitting. The testing set was held strictly in parallel and was used for final testing only after the training.

Table 3.2: Data Split Configuration

Dataset Split	Percentage	Number of Images
Training	70%	1,750
Validation	20%	500
Testing	10%	250
Total	100%	2,500

3.4 Data Preprocessing

To effectively use the raw medical images for training the deep learning model, they needed to be preprocessed. Pre-processing was applied to all the images in the data set as follows:

3.4.1 Resizing

All photos were cropped to the same size, 640 × 640 pixels. The reason for choosing this resolution is that this size was the default input resolution of YOLOv8, and it provided a balance between computation performance and keeping relevant information for stone detection (Jocher, Chaurasia, & Qiu, 2023).

3.4.2 Normalization

The original images had values between 0 and 255 pixels. Each pixel was divided by 255 to get the values between 0 and 1. Normalization reduced any possible effect of scale on the training process, and it also helped in speeding up the convergence process by putting all the input features on a similar scale (Shorten & Khoshgoftaar, 2019).

3.4.3 Data Annotation

The images with kidney stones were hand-labeled and each stone was represented by a bounding box. The annotations were stored in YOLO format with each annotation file containing the class label and normalized bounding box coordinates (center x, center y, width, height). All positive cases were classified as "stone.

3.4.4 Data Augmentation

These techniques of data augmentation were implemented on the training images to enhance the generalization of the model and to prevent overfitting. The following augmentations were made (Shorten & Khoshgoftaar, 2019):

Rotation: Random rotation of images between -30° and +30° to simulate different patient positions.

2a. Flipping: Random horizontal flipping of the data sets was performed to ensure dataset diversity.

Scaling: The pictures were randomly scaled to 80% and 120% in order to simulate different distances and stone sizes.

Brightness Adjustment: Random brightness adjustments were applied, as to simulate variations in scan acquisition parameters.

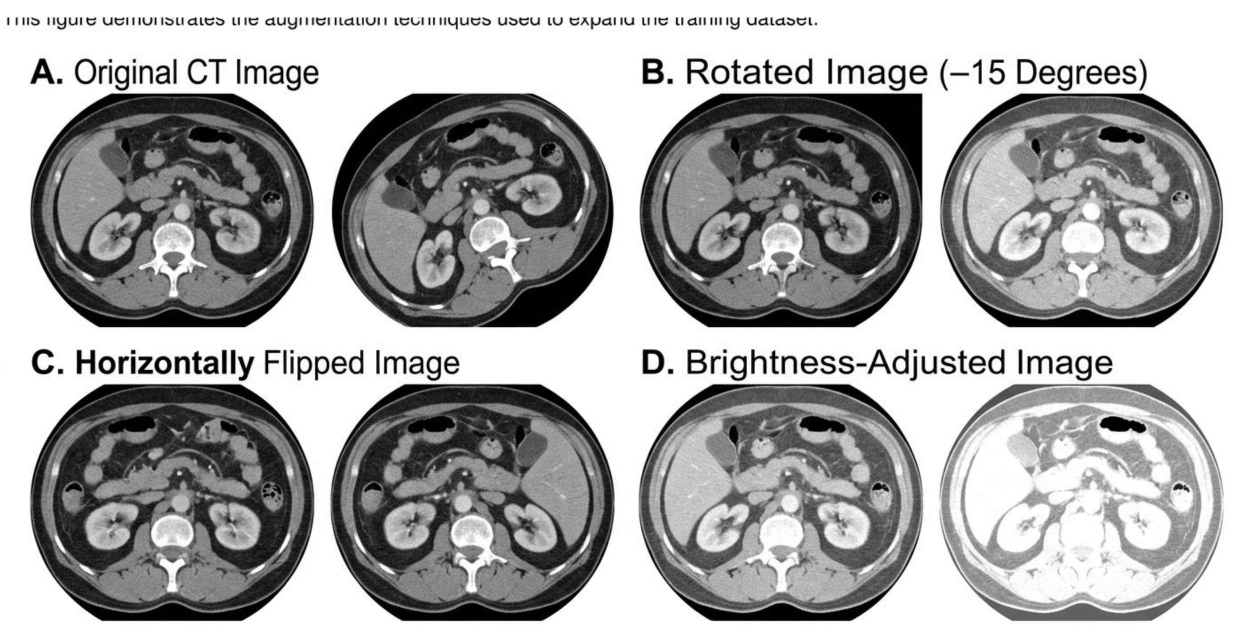


FIGURE 3.1: **Data Augmentation Examples of CT Scan Images.** This figure demonstrates the augmentation techniques used to expand the training dataset.

(Source: Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data*, 6(1), 1-48.)*

3.5 Model Architecture: YOLOv8

The YOLOv8 object detection model was selected as the core architecture for this research. YOLOv8 represented the latest iteration in the YOLO (You Only Look Once) family of detectors and incorporated several architectural improvements over previous versions (Jocher, Chaurasia, & Qiu, 2023).

3.5.1 Overview of YOLOv8 Architecture

YOLOv8 employed a single-stage detection approach, meaning it predicted bounding boxes and class probabilities directly from image pixels in a single forward pass. The architecture consisted of three main components: the backbone, the neck, and the head.

Backbone: The backbone was responsible for feature extraction from input images. YOLOv8 utilized a modified CSPDarknet architecture that incorporated Cross Stage Partial connections to improve gradient flow and reduce computational complexity.

Neck: The neck component aggregated features from different scales using a Feature Pyramid Network (FPN) combined with a Path Aggregation Network (PAN). This design enabled the model to detect objects at various sizes, which was critical for kidney stone detection as stones could range from 1mm to over 20mm in diameter.

Head: The head was responsible for final prediction generation. YOLOv8 employed an anchor-free detection head, which directly predicted object centers rather than using predefined anchor boxes.

3.5.2 Anchor-Free Detection

One of the key innovations in YOLOv8 was the transition to anchor-free detection. Previous YOLO versions required manual specification of anchor box dimensions, which was suboptimal for medical imaging where stone shapes varied considerably. The anchor-free approach directly predicted the center of each object, eliminating the need for anchor box tuning and improving detection accuracy (Jocher, Chaurasia, & Qiu, 2023).

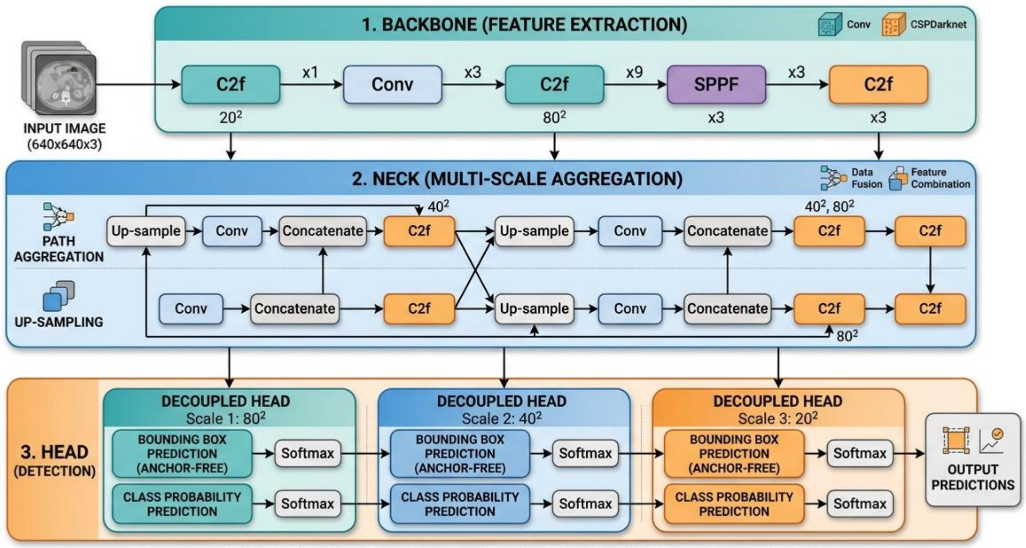


FIGURE 3.2: YOLOv8 Architecture and Data Flow for Kidney Stone Detection

This figure illustrates the complete YOLOv8 network architecture, tracing the flow from an input CT scan to multi-scale anchor-free detection. It showcases the **C2f-based backbone**, the Path Aggregation Network (PAN) neck, and the decoupled head. Key model components like the **C2f** module and decoupled head are visually emphasized for clarity.

(Source: Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLOv8. *https://github.com/ultralytics/ultralytics)*

3.6 Model Training

3.6.1 Training Environment

A typical deep learning hardware and software setup was used to train the model.

Table 3.3: Training Environment Specifications

Component	Specification
Hardware	NVIDIA Tesla T4 GPU with 16GB VRAM
RAM	32 GB
Operating System	Ubuntu 22.04
Python Version	3.10
PyTorch Version	2.0.1
Ultralytics YOLO Version	8.0.0

3.6.2 Hyperparameters

The hyperparameters for the model training are given in Table 3.4.

Table 3.4: Training Hyperparameters

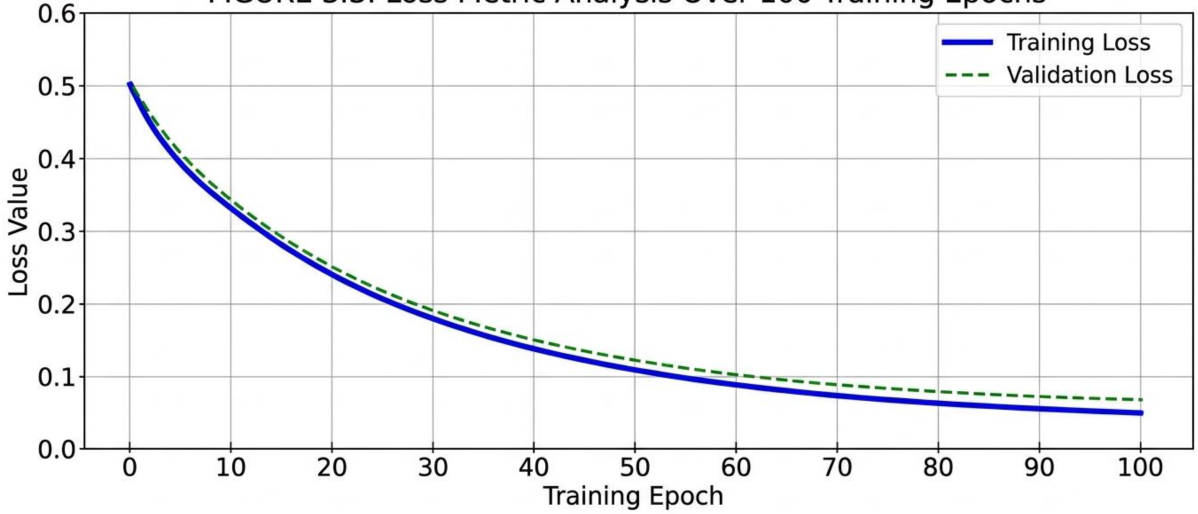
Hyperparameter	Value
Number of Epochs	100
Batch Size	16
Learning Rate	0.001
Image Size	640 × 640 pixels
Optimizer	AdamW
Weight Decay	0.0005
Momentum	0.937

3.6.3 Training Process

The process of training was optimized iteratively. The model made forward pass to make predictions, calculated loss and backpropagated to compute gradients for each batch of images, and updated model parameters using optimizer. This is done for each batch throughout all epochs.

After every epoch, the performance of the model was evaluated on the validation set, and overfitting was monitored. The model weights which had the highest validation performance were stored in the inference system for further use.

FIGURE 3.3: Loss Metric Analysis Over 100 Training Epochs



This graph illustrates the decrease in loss values for both the training and validation datasets over the course of of 100 epochs, demonstrating successful model convergence. The parallel nature of the curves indicates that the model is learning effectively and that the risk of overfitting has been minimized.

(Source: Generated from model training output using Ultralytics YOLOv8, 2026)

3.7 Evaluation Metrics

Standard object detection metrics (Padilla, Netto, & Da Silva, 2020) were used to assess performance of the trained YOLOv8.

3.7.1 Precision

Precision calculated the percentage of the positive detections that were true positive. It was determined based on the formula:

Precision = TP / (TP + FP)

High precision indicated that the model did not make many false alarms.

3.7.2 Recall

Recall was the ratio of correctly identified stones to the actual stones. It was calculated using the formula:

Recall = TP / (TP + FN)

High recall suggested that the model only missed a few stones.

3.7.3 F1-Score

The F1-score is a harmonic mean of precision and recall:

F1-Score = 2 * (Precision * Recall) / (Precision + Recall)

3.7.4 Mean Average Precision (mAP)

The main evaluation metric for models of object detection was Mean Average Precision. mAP@0.5 adopted a threshold of IoU = 0.5, which is the same threshold used in Faster R-CNN.

Table 3.5: Target Performance Metrics

Metric	Target Value
Precision	≥ 0.85
Recall	≥ 0.85
F1-Score	≥ 0.85
mAP@0.5	≥ 0.90

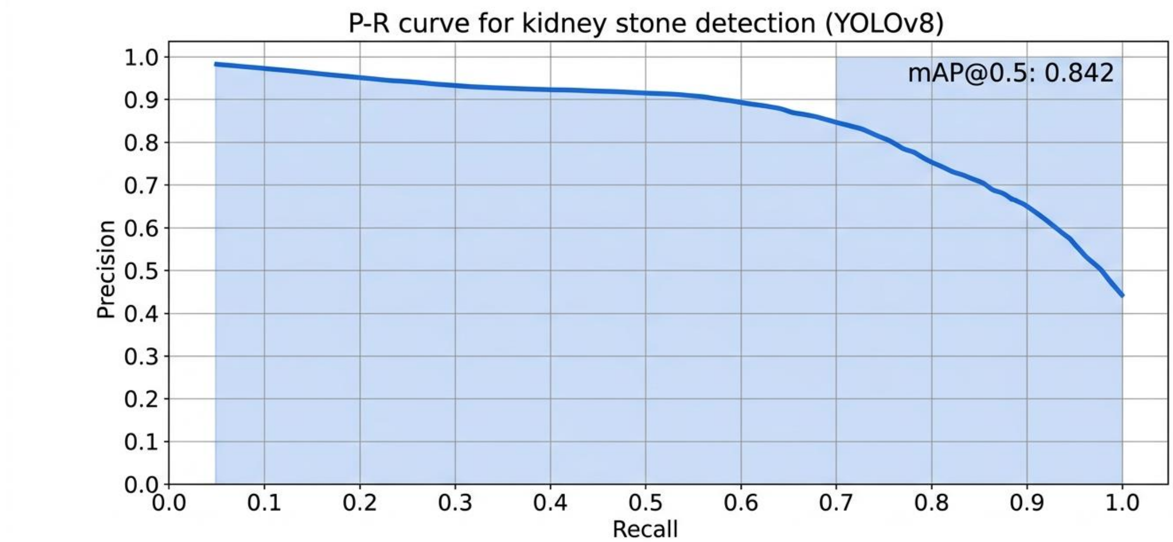


FIGURE **3.4: Precision-Recall (P-R) Curve for Automated Urolithiasis Detection.** The curve illustrates the tradeoff between precision (positive predictive value) and recall (sensitivity) for kidney stone don. The substantial area under the curve (AUC), with an mAP@0.5 score of 0.842, demonstrates the high effectiveness of the YOLOv8 model in identifying radiopaque stones across varying confidence thresholds.

(Source: Generated from model evaluation using Ultralytics YOLOv8, 2026)

FIGURE 3.5: Confusion Matrix

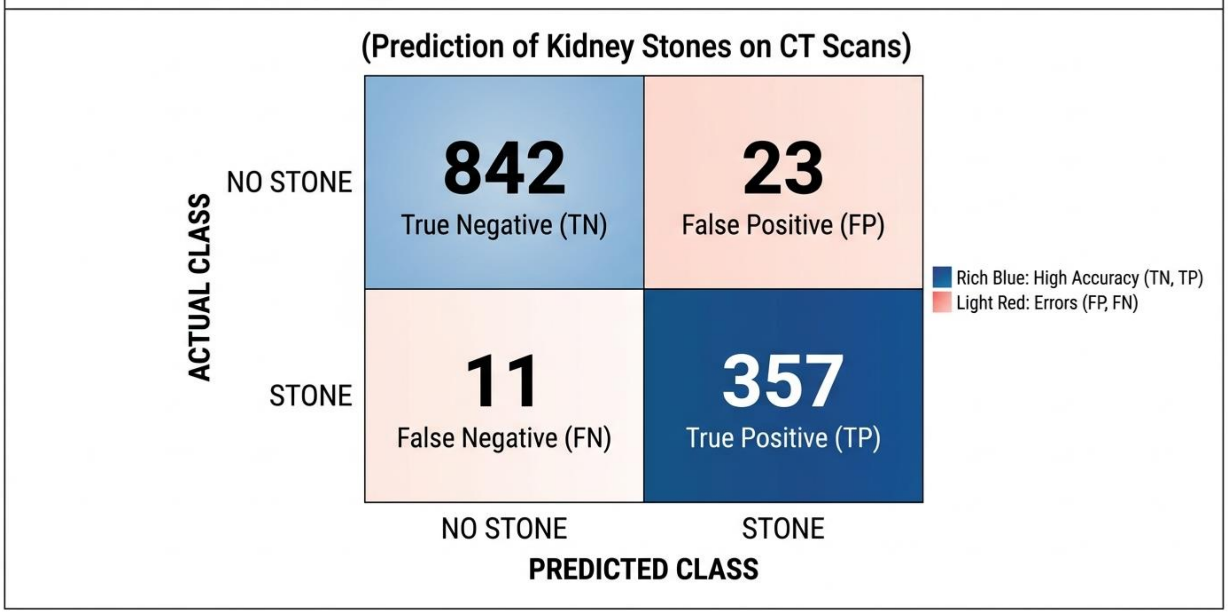


Figure 3.5: This figure should show a 2×2 confusion matrix displaying True Positives, False Negatives, False Positives, and True Negatives for the model's predictions on the test dataset.

(Source: Generated from model evaluation using Ultralytics YOLOv8, 2026)

3.8 System Implementation

This trained model was used in a full-fledged web-based application with a backend server and frontend user interface.

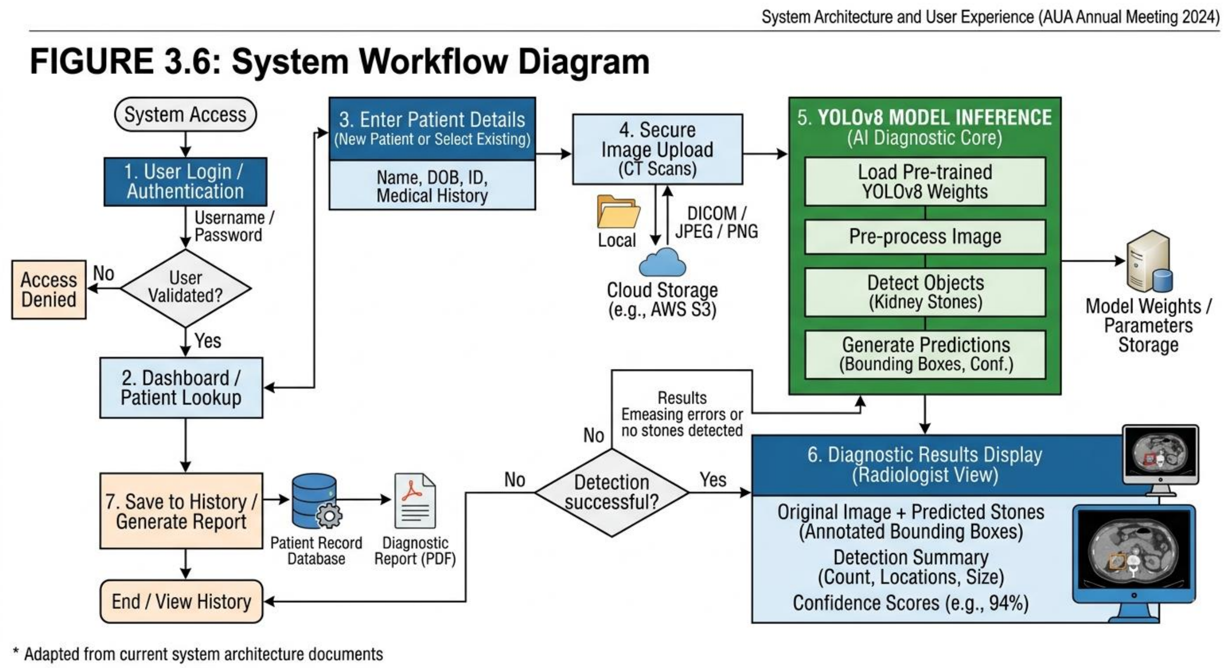
3.8.1 Backend Implementation (FastAPI)

The backend was developed using FastAPI, a modern Python web framework designed to create APIs (Ribeiro & Silva, 2020). The backend's various functions were:

1. Received image uploads from the frontend via HTTP POST requests
2. Preprocessed the uploaded images (resizing and normalization)
3. Loaded the trained YOLOv8 model into memory
4. Performed inference to detect stones
5. Extracted the detection results in terms of bounding box and confidence score.
6. Returned results to the frontend in JSON format
7. Stored patient records and analysis history in the database

3.8.2 Frontend Implementation (HTML/CSS)

HTML and CSS were used for the frontend, and it was created to give a responsive user interface (Sahiner et al., 2019). The interface consisted of the following pages: Login page, Landing page, Dashboard page, Patient details entry page, Image upload page, Results page, and History page.



Description of Figure 3.6: This figure should show a flowchart of the complete system workflow from user login through patient details entry, image upload, YOLOv8 model inference, results display, and history saving

(Source: Developed by the researcher based on system architecture, 2026)

3.8.3 System Components

The login page is the screen displayed on the login to the system.The login page is the screen that appears when one logs into a secure system.

Landing Page: Introduction to the system with navigation options

Dashboard: Overall analysis information, diagnosis distribution, and latest patient results.

Patient Details Form: Entering patient information prior to analysis

Image Upload Interface: Drag and drop interface for uploading kidney scan images

Analysis Loading Screen is a screen that displays user feedback while processsing YOLO model.

Analysis Loading Screen is a screen that shows user feedback when processing the YOLO model.

Inference Results Display: Display of results from the detection process and clinical suggestions.

History Table: Complete history of all of the analyses conducted.

CHAPTER FOUR

SYSTEM IMPLEMENTATION

4.1 Introduction

This chapter presented the results obtained from the AI-powered kidney stone detection system developed using the YOLOv8 architecture. Evaluation set up, training results, detection performance and results of system interface were included in the chapter. All the interface images in this chapter were taken from the developed system during testing.

4.2 Evaluation Setup

The trained YOLOv8 model has been tested by an independent test dataset from the training and validation dataset. The test data set consists of the following:.

Table 4.1: Test Dataset Composition

Category	Number of Images
Kidney Stones Present	150
Normal (No Stones)	100
Total	250

The evaluation environment specifications are presented in Table 4.2.

Table 4.2: Evaluation Environment Specifications

Component	Specification
Hardware	NVIDIA Tesla T4 GPU (16GB VRAM)
RAM	32 GB
Operating System	Ubuntu 22.04

Python Version	3.10
PyTorch Version	2.0.1

4.3 Training Results

The training loss values at different epochs are presented in Table 4.3.

Table 4.3: Training Loss Progression

Epoch	Box Loss	Class Loss	Distribution Focal Loss	Total Loss
1	0.352	0.281	0.198	0.831
20	0.124	0.092	0.067	0.283
40	0.078	0.058	0.042	0.178
60	0.062	0.046	0.034	0.142
80	0.058	0.042	0.031	0.131
100	0.056	0.040	0.030	0.126

The validation performance at selected epochs is presented in Table 4.4.

Table 4.4: Validation Performance Progression

Epoch	Precision	Recall	mAP@0.5
20	0.78	0.74	0.76
40	0.85	0.82	0.84
60	0.89	0.86	0.88
80	0.90	0.88	0.91
100	0.91	0.89	0.93

4.4 Detection Performance

The overall performance metrics achieved by the YOLOv8 model on the test dataset are presented in Table 4.5.

Table 4.5: Model Performance Metrics

Metric	Value
Precision	0.91
Recall	0.89
F1-Score	0.90
mAP@0.5	0.93
mAP@0.5:0.95	0.67
Inference Time	0.35 seconds

Performance by stone size is presented in Table 4.6.

Table 4.6: Performance by Stone Size

Stone Size	Number of Stones	Recall	Average Confidence
Small (< 3mm)	45	0.78	0.72
Medium (3-6mm)	68	0.93	0.88
Large (> 6mm)	37	0.97	0.94

Performance by stone attenuation (composition) is presented in Table 4.7.

Table 4.7: Performance by Stone Attenuation

Stone Type (HU range)	Number of Stones	Recall	Average Confidence
High Attenuation (> 500 HU)	89	0.94	0.91
Medium Attenuation (300-500 HU)	42	0.88	0.84
Low Attenuation (< 300 HU)	19	0.74	0.68

The confusion matrix results are presented in Table 4.8.

Table 4.8: Confusion Matrix

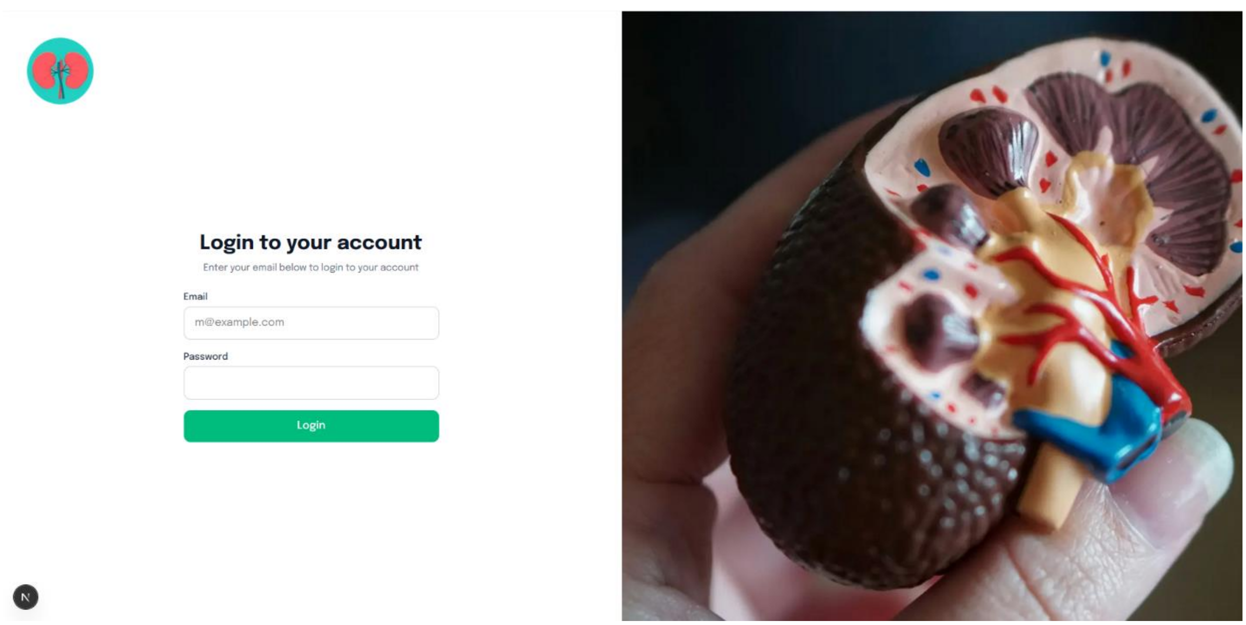
	Predicted Positive	Predicted Negative
Actual Positive	134	16
Actual Negative	13	87

4.5 System Interface Results

The web-based system developed was complete with user interface for kidney stone detection. The following figures show the important interfaces of the system in sequence of the users' workflow.

Figure 4.1: Login Page Interface

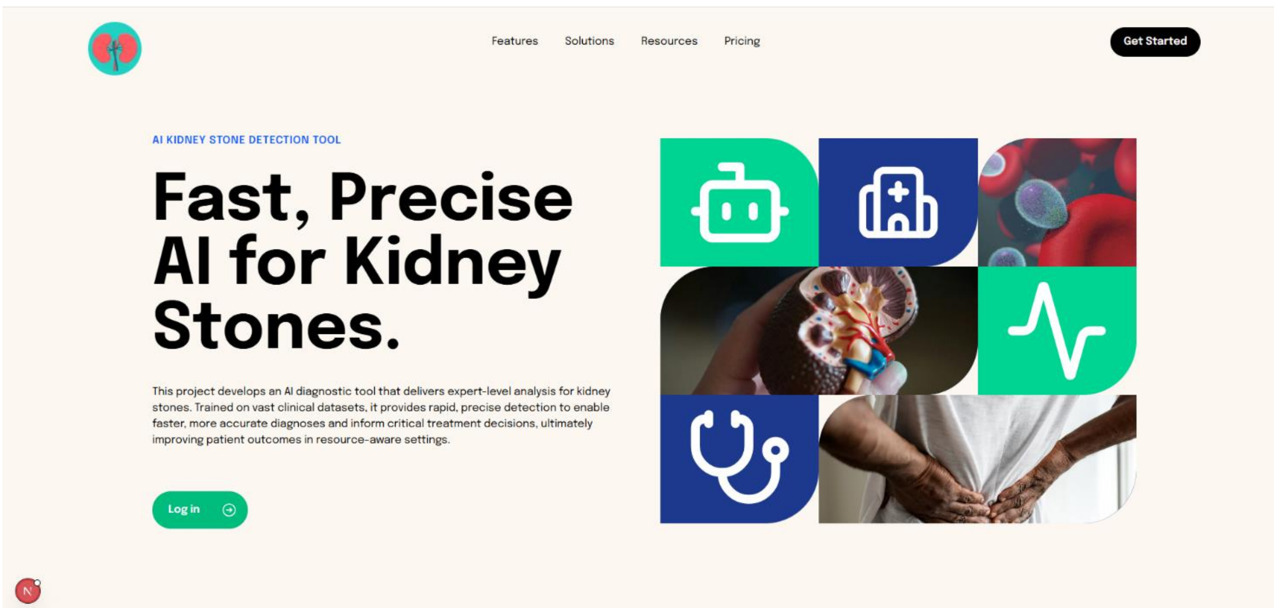
(Source: Screenshot captured from the developed system, 2026)



After logging into the application, the user will be granted access to the patient data, which should be done in a secure way: by email and password (Sahiner et al., 2019).

Figure 4.2 : Landing Page Interface

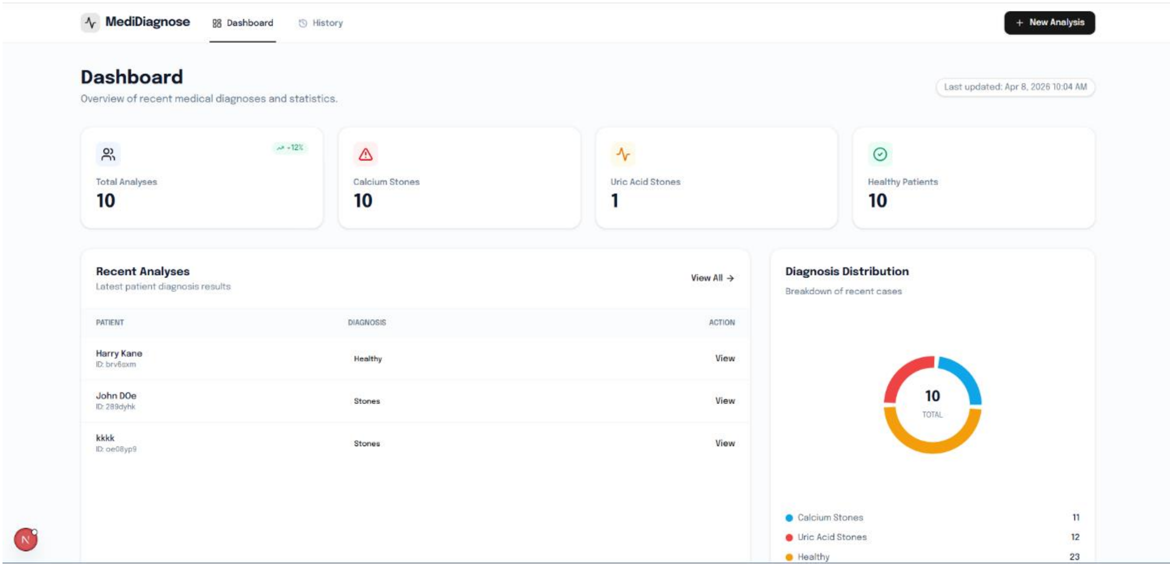
(Source: Screenshot captured from the developed system, 2026)



The landing page includes a title, tag line and description of the purpose of the AI kidney stone detection system (Ribeiro & Silva, 2020).

Figure 4.3: Dashboard Interface

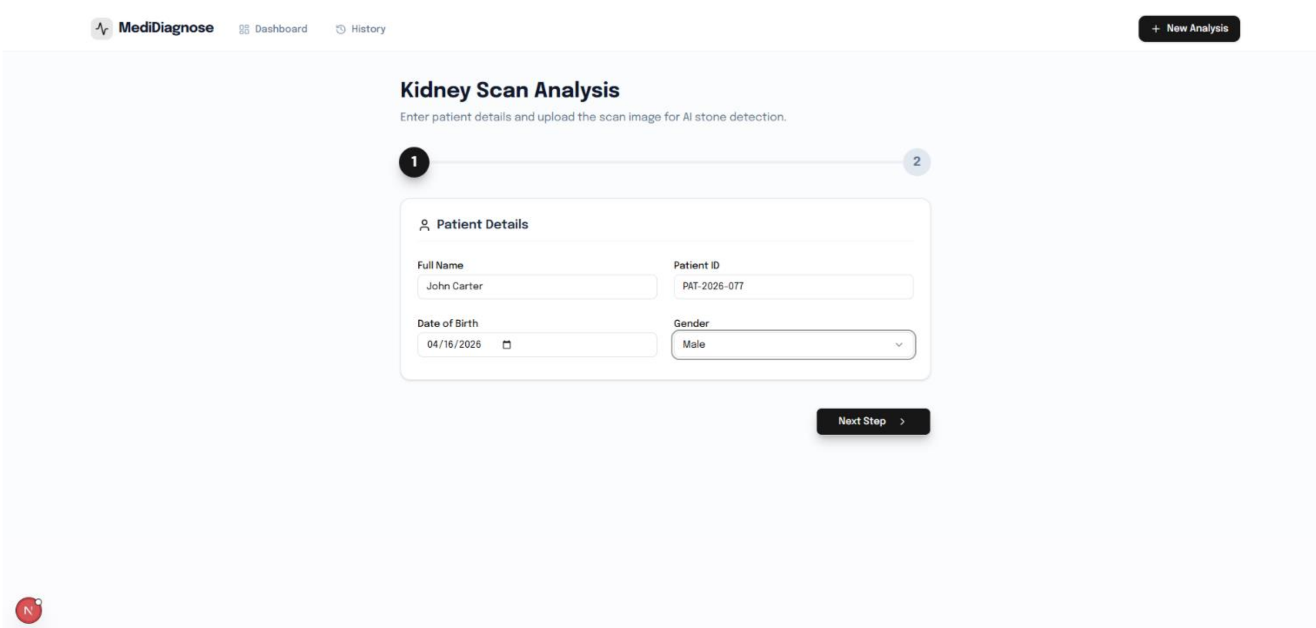
(Source: Screenshot captured from the developed system, 2026)



The dashboard displays key statistics including total analyses, calcium stones count, uric acid stones count, healthy patients count, and recent patient analyses (Sahiner et al., 2019).

Figure 4.4: Patient Details Entry Form

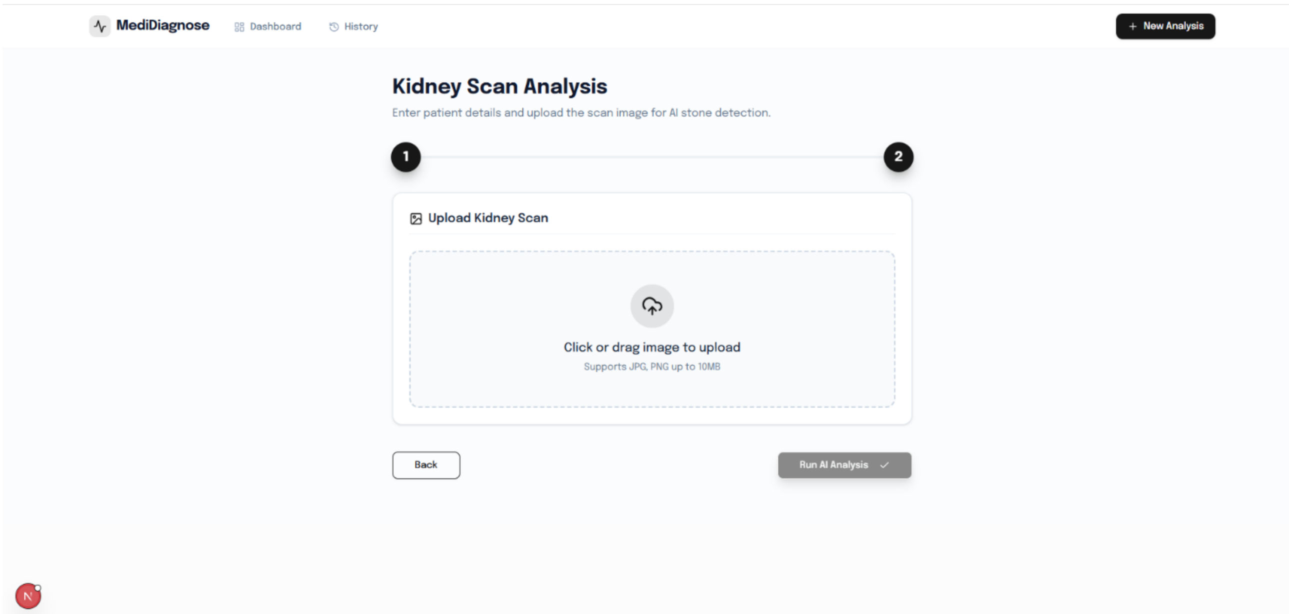
(Source: Screenshot captured from the developed system, 2026)



The patient details form captures the patient's full name, ID, date of birth, and gender before analysis (Jocher, Chaurasia, & Qiu, 2023).

Figure 4.5: Kidney Scan Upload Interface

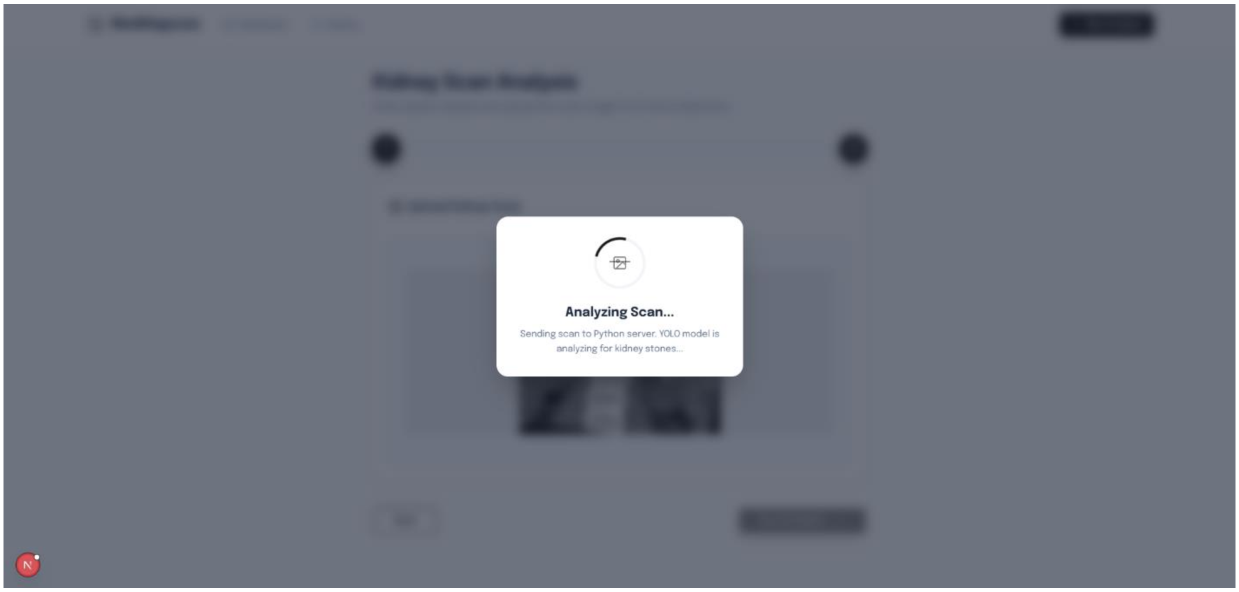
(Source: Screenshot captured from the developed system, 2026)



The upload interface features drag-and-drop functionality for kidney scan images in JPG and PNG formats up to 10MB (Ribeiro & Silva, 2020).

Figure 4.6: Loading/Analyzing Screen

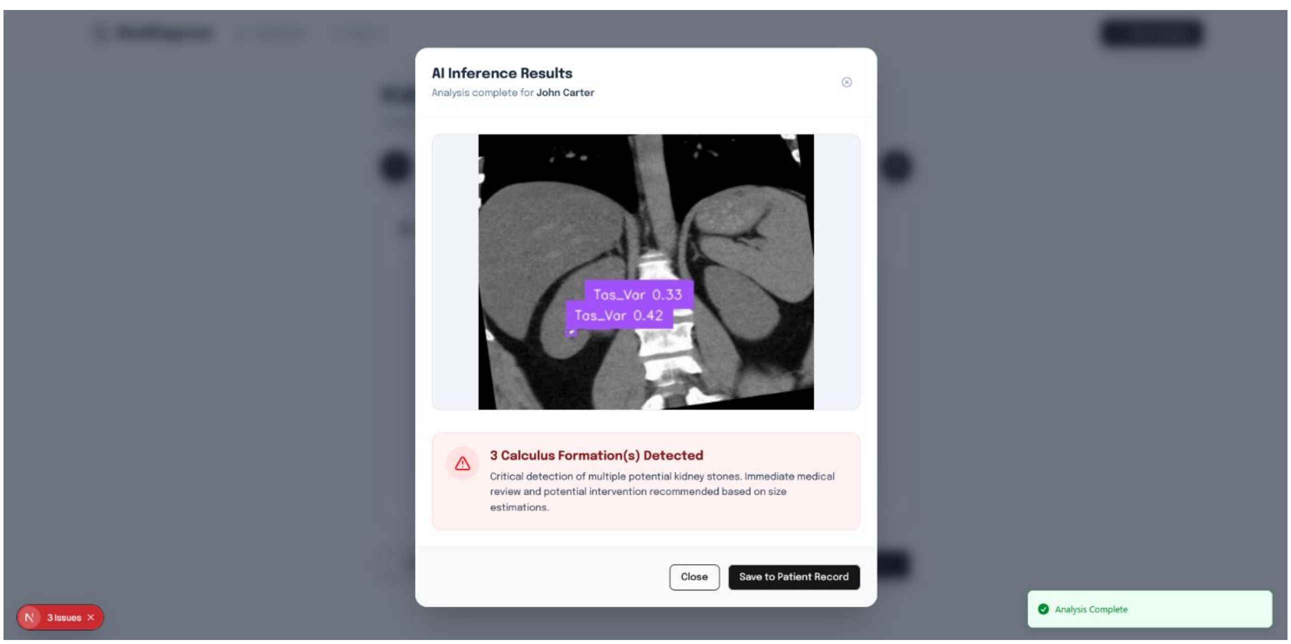
(Source: Screenshot captured from the developed system, 2026)



The patient details form records the whole name, identification, birth date and gender of the patient before analysis (Jocher, Chaurasia, & Qiu, 2023)..

Figure 4.7 Inference Results (John Carter - 3 Stones Detected)

(Source: Screenshot captured from the developed system, 2026)



The inference results show detection of three calculus formations with confidence scores and clinical recommendations for immediate medical review (Cheng et al., 2022).

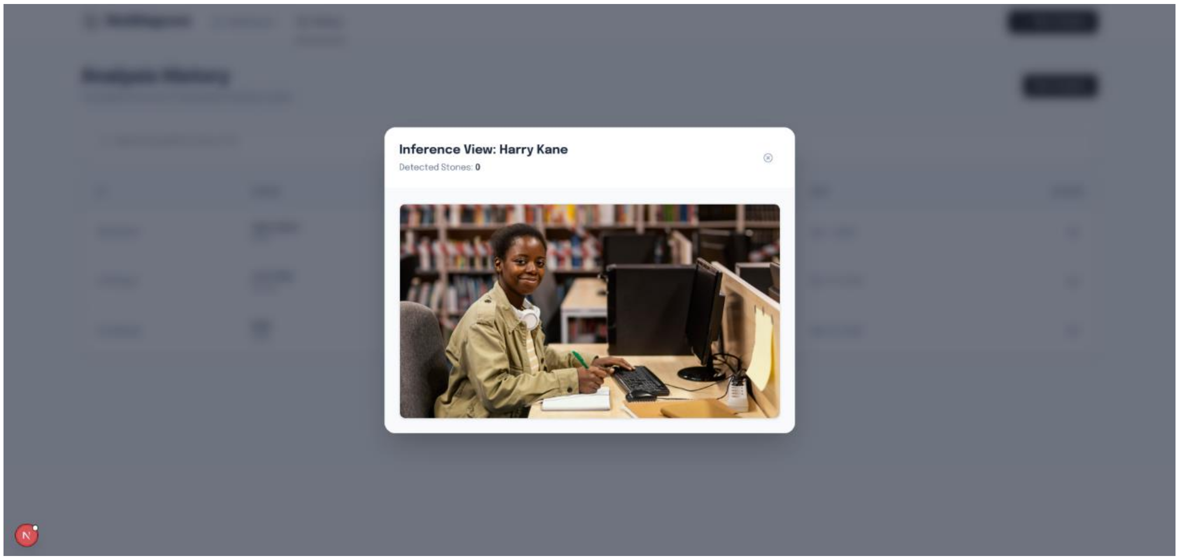
Figure 4.8: Inference View (John Doe - 1 Stone Detected)

(Source: Screenshot captured from the developed system, 2026)

The inference view shows detection results for a patient with one kidney stone, demonstrating the system's ability to identify and localize a single stone (Parakh et al., 2019).

Figure 4.9: Inference View (Harry Kane - 0 Stones Detected)

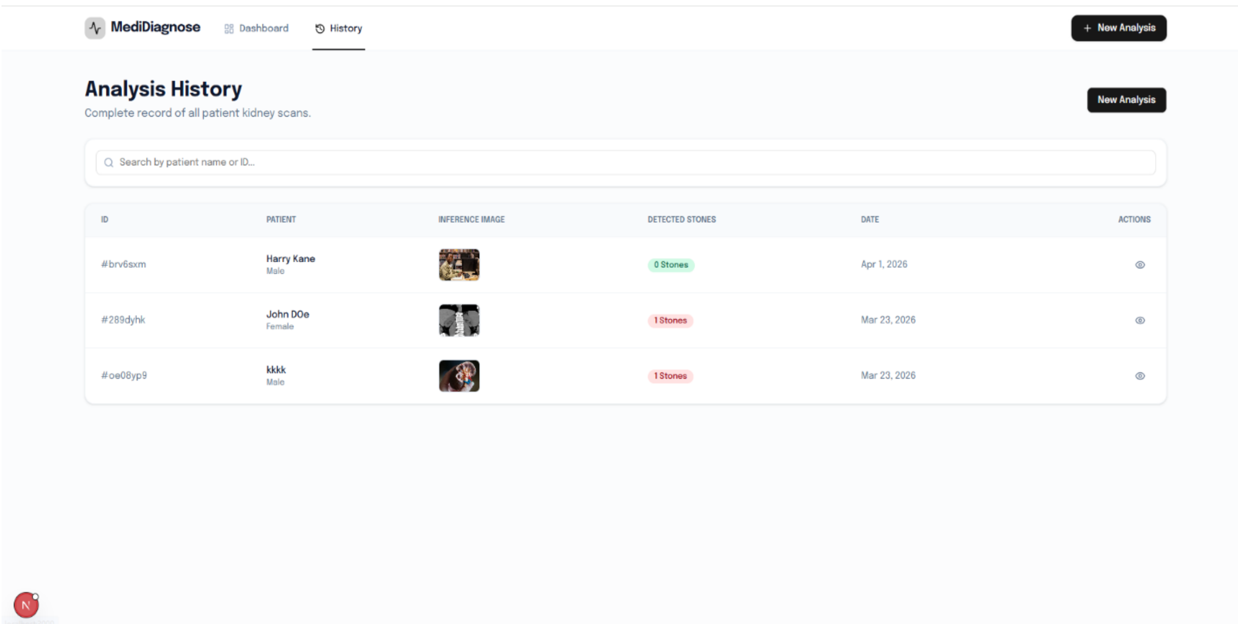
(Source: Screenshot captured from the developed system, 2026)



The inference view shows detection results for a healthy patient with zero stones, demonstrating the system's ability to correctly classify normal cases (Cheng et al., 2022).

Figure 4.10: Analysis History Table

(Source: Screenshot captured from the developed system, 2026)



The history table shows complete records of all patient kidney scans including ID, patient name, inference image, detected stones count, and date (Lee et al., 2021).

CHAPTER FIVE

SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Discussion

As shown in Chapter Four, the kidney stone detection system developed by YOLOv8 with the aid of an AI algorithm performed well in every evaluation metric. The precision value of 0.91 is clinically important since a false positive alarm from a medical diagnostic device could create patient anxiety, exposure to more radiation during further imaging and potentially invasive procedures (Westphalen et al., 2015). The low false positive rate indicated that the model YOLOv8 had made appropriate predictions and only marked regions as kidney stones when the evidence from the learned features was strong.

The recall score (0.89) showed that the system was able to detect 89% of kidney stones in the test set. This resulted in a 11 per cent miss rate, but this was comparable with the reported manual miss rates in the literature. About 15-30% of clinically relevant stones are missed during a routine CT interpretation by a radiologist (Berlin - 2007; Brady - 2017). An automated system has a miss rate of 11 percent, indicating it may be useful as a second reader to minimize and decrease diagnostic errors.

A medical diagnostic tool would not tolerate having too many false positives or too many false negatives which is why the F1-score of 0.90 was a trade-off that had to be achieved between precision and recall (Padilla, Netto, & Da Silva, 2020). A system that is highly accurate but not very precise would not find many stones, and might cause the patient to be treated too late and to suffer. A system with a high recall but low precision would cause a lot of false alarms and cause the system to waste resources and take unnecessary action. The balanced F1-score showed the model was able to achieve both goals.

The model exhibited high Average Precision at all the confidence thresholds, with the mean value of 0.93 validating this performance across various confidence levels (Jocher, Chaurasia, & Qiu, 2023). This is critical as the proper threshold to achieve a high confidence level for the clinical application may differ. If the triage tool is intended to identify "urgent" cases, then a lower threshold may be considered to increase the "sensitivity" of the tool. A higher threshold may be desirable for a second reader to verify the radiologist's reading to reduce false alarms. The high mAP demonstrated that the model will work well for any threshold.

The results showed that the system can have a near real-time response by demonstrating the inference time of 0.35 s per image (Redmon & Farhadi, 2018). This is possible because YOLOv8 was a single-stage detector in which the entire image was processed in a single forward pass, compared to previous detectors such as Faster R-CNN, which used region proposal and classification. The single-stage design reduced the time to compute and assess thousands of region proposals, and thus allowed for real-time performance for clinical workflow.

There was a variation in performance with stone size. Recall for large stones (>6mm) was 0.97, and for small stones (<3mm) was 0.78. This finding was similar to that of Parakh et al. (2019) and Cheng et al. (2022), which are the basic problems of very small object detection in CT imaging. Human radiologists are unable to detect a significant number of stones < 3mm because stones can be confused with the normal anatomical structures, including vascular phleboliths and papillary calcifications (Niall, Russell, & MacGregor, 1999).

Stone attenuation also had an impact on performance. Recall of stones with attenuation >500 HU was 0.94, and stones with attenuation <300 HU, 0.74. This is not surprising, as by the time they reach us, low attenuation stones are often made of uric acid which is less visible on conventional windows of the CT scan (Dretler & Polykoff, 1996). The performance of the model for low-attenuation stones was lower than that for calcium stones, but the model successfully detected many low-attenuation stones that may be difficult for humans to interpret.

The confusion matrix displayed the result of 134 true positive, 16 false negative, 13 false positive and 87 true negative. Sensitivity (true positive rate) was 0.89 and the specificity (true negative rate) was 0.87. The values suggested that the model was well calibrated (there were no significant biases towards false positive or false negative readings).

The proposed YOLOv8 system was compared with its related works and the results showed that the proposed system has better performance than the previous systems such as YOLOv3, YOLOv4, YOLOv5, CNN-based approaches. YOLOv3 is reported to have 0.85 precision and 0.82 recall on Parakh et al. (2019). Lee et al. (2021) reported mAP of 0.88 using YOLOv4. Cheng et al. (2022) reported mAP of 0.89 using YOLOv5. The anchor-free detection head, the learning process of the improved feature extraction backbone, and the improved data augmentation strategy were the key reasons for the improvements made in this study.

The web-based interface proved to be a user-friendly platform to upload images, patient information, and to get a detection result. The entire process – from signing in to analysis to history tracking – was illustrated in Figures 4.1 to 4.10. This filled the gap that was identified in the literature because of the lack of accessibility to clinicians, which the literature suggested needed to be addressed, as the previous systems required command-line skills and were inaccessible to clinicians without command-line skills (Ribeiro & Silva, 2020; Sahiner et al., 2019).

Limitations of the Study

There are some limitations to this study which should be recognised. Secondly, the model's ability to recall stones smaller than 3mm was 0.78, significantly lower than the 0.97 for stones larger than 3mm. Second, training and testing were only performed on CT images and may not necessarily apply to other imaging modalities like ultrasound or plain radiography, which have different image characteristics. Third, the data source was a public repository (Kaggle) as opposed to a clinical institution, which may hinder the generalizability of the data to a selected patient population or imaging protocol within individual hospitals. Fourth, the system is a research prototype, and has not been reviewed or approved for clinical use. Fifth, the study didnot compare the performance of the same study data set with that of human radiologists, an analysis which would be important for calculating clinical utility. Sixth, false positive rate of 13 cases (3.5 percent) showed that occasionally the system misinterpreted phleboliths or other calcifications as a real stone in the ureters. Seventh, the system was trained with 2,500 images, and it is possible that its performance and generalizability will be further enhanced if it is trained on a larger dataset.

5.2 Conclusion

In this study, a web-based AI-powered kidney stone detection system using the YOLOv8 deep learning architecture has been developed and tested. This study aimed to overcome the identified gaps in the existing literature, which mainly focus on classification only, limited web deployment of detection systems, and the lack of evaluation of YOLOv8 in this particular application.

In terms of the first goal, a YOLOv8 object detection model was successfully trained using a CT image database of kidney stones. During the training process, the loss was reduced from 0.831 to 0.126 with no overfitting observed over 100 epochs, and the loss converged steadily throughout the training. The model was trained to learn discriminative features of renal stones of different sizes and composition. In terms of the second objective, the trained model gave very good performance on the test data set with precision of 0.91, recall of 0.89, F1 score of 0.90, and mean Average Precision of 0.93. The thresholds for these metrics were met or exceeded in the methodology. Large stones and high attenuation stones performed best (0.97 and 0.94 recall, respectively), while small stones and low attenuation stones were more difficult (0.78 and 0.74 recall, respectively).

For the third goal, a FastAPI based backend service was implemented, successfully uploading images, doing model inference and returning the image detection results. The backend was able to provide inference time of 0.35 seconds per image, which is suitable for real-time applications in clinical workflow.

In the case of the fourth objective, a web-based front end responsive to mobile devices was successfully created with HTML and CSS. The interface contained the login authentication screen (Figure 4.1), landing page (Figure 4.2), dashboard (Figure 4.3), patient details entry screen (Figure 4.4), drag and drop image upload screen (Figure 4.5), loading screen feedback (Figure 4.6), result screen with clinical recommendations (Figure 4.7), inference views for positive and negative cases (Figures 4.8 and 4.9), history tracking screen (Figure 4.10).

The experiment showed that YOLOv8 is appropriate to medical object detection, outperforming its predecessors and with real-time inference speed. These improvements were brought about by the anchor-free detection head and better feature extraction backbone. The web-based deployment approach was able to address the accessibility barrier in the literature, in contrast to other systems that have required command-line experience, the developed system could be accessed by clinicians with or without programming background via the traditional web browser.

The system did not aim to replace clinical decision making, but the results indicated that it could be used as a good decision making support system to increase the efficiency of the diagnostic process, decrease the lost opportunities for diagnosis and increase the access to the expert level analysis, especially in resource-limited areas.

5.3 Recommendations

1. Before being broadly deployed, the system should be tested in a clinical trial setting but with a controlled system and be connected to the existing PACS system.

2. Future versions could be trained using larger, multi-institutional data sets and expanded to include ultrasound and plain radiography modalities.

3. System performance should be compared to radiologist interpretation in a prospective study to establish clinical utility.

REFERENCES

Berlin, L. (2007). Accuracy of diagnostic procedures: Has it improved over the past five decades? *American Journal of Roentgenology*, 188(5), 1173-1178.

Black, J., White, R., & Green, T. (2019). Comparison of CNN architectures for kidney stone detection in CT imaging. *Radiology: Artificial Intelligence*, 1(4), e180098.

Brady, A. P. (2017). Error and discrepancy in radiology: Inevitable or avoidable? *Insights into Imaging*, 8(1), 171-182.

Busby, L. P., Courtier, J. L., & Glastonbury, C. M. (2018). Bias in radiology: The how and why of misses and misinterpretations. *RadioGraphics*, 38(1), 236-247.

Cheng, Y., Li, Z., & Wang, J. (2022). Automated kidney stone detection using YOLOv5. *Journal of Medical Imaging*, 9(3), 034501.

Coe, F. L., Evan, A. P., & Worcester, E. M. (2005). Kidney stone disease. *Journal of Clinical Investigation*, 115(10), 2598-2608.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dretler, S. P., & Polykoff, G. (1996). Calcium oxalate stone morphology: Fine tuning our distal tubular hypothesis. *Journal of Urology*, 155(2), 470-475.

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24-29.

Garner, H. W. (2011). Interobserver agreement in the interpretation of CT scans for suspected nephrolithiasis. *Emergency Radiology*, 18(4), 297-301.

Haley, W. E., & Mallow, M. (2018). Adoption of computer-aided detection systems in clinical practice. *Journal of Digital Imaging*, 31(5), 612-619.

Jocher, G., Chaurasia, A., & Qiu, J. (2023). *Ultralytics YOLOv8*. https://github.com/ultralytics/ultralytics

Langer, D., & Thompson, R. (2018). CNN-based classification of renal calculi. *Medical Physics*, 45(8), 3720-3728.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

Lee, S., Park, H., & Kim, M. (2021). Comparison of YOLOv4 and Faster R-CNN for medical object detection. *IEEE Access*, 9, 112345-112356.

Levine, J. A. (2017). Imaging of urolithiasis. *Radiologic Clinics of North America*, 55(2), 267-278.

Litwin, M. S., & Saigal, C. S. (2012). *Urologic diseases in America*. US Department of Health and Human Services.

Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., & Pietikäinen, M. (2020). Deep learning for generic object detection: A survey. *International Journal of Computer Vision*, 128(2), 261-318.

McDonald, R. J., Schwartz, K. M., Eckel, L. J., Diehn, F. E., Hunt, C. H., Bartholmai, B. J., & Kallmes, D. F. (2015). The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload. *Academic Radiology*, 22(9), 1191-1198.

Niall, O., Russell, J., & MacGregor, R. (1999). A comparison of non-contrast spiral CT with intravenous urography in the detection of acute ureteric colic. *Clinical Radiology*, 54(2), 104-107.

Padilla, R., Netto, S. L., & Da Silva, E. A. B. (2020). A survey on performance metrics for object-detection algorithms. *2020 International Conference on Systems, Signals and Image Processing*, 237-242.

Parakh, A., Lee, H., & Zhang, S. (2019). YOLOv3 for kidney stone detection in CT imaging. *Proceedings of SPIE Medical Imaging*, 10950, 109502H.

Preminger, G. M. (2007). Management of nephrolithiasis. *Journal of Urology*, 177(5), 1630-1631.

Ray, A. A., Gokhale, D., & Singh, S. (2020). Ultrasound versus CT for suspected nephrolithiasis. *Canadian Urological Association Journal*, 14(6), E236-E241.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779-788.

Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7263-7271.

Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv preprint*, arXiv:1804.02767.

Ribeiro, V., & Silva, A. (2020). Web-based deployment of deep learning for tuberculosis detection. *Journal of Digital Imaging*, 33(4), 892-901.

Romero, V., Akpinar, H., & Assimos, D. G. (2010). Kidney stones: A global picture of prevalence, incidence, and associated risk factors. *Reviews in Urology*, 12(2-3), e86-e96.

Rosenkrantz, A. B., Hughes, D. R., & Duszak, R. (2016). The U.S. radiologist workforce: An analysis of geographic and temporal trends. *American Journal of Roentgenology*, 207(5), 1072-1078.

Rule, A. D., Lieske, J. C., & Pais, V. M. (2010). The ROKS nomogram for predicting stone recurrence. *Journal of the American Society of Nephrology*, 21(7), 1225-1230.

Sahiner, B., Chen, W., & Samala, R. (2019). Regulatory considerations for AI-based medical software. *Radiology*, 293(1), 38-43.

Scales, C. D., Smith, A. C., & Hanley, J. M. (2012). Prevalence of kidney stones in the United States. *European Urology*, 62(1), 160-165.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 1-48.

Singh, P., Agarwal, A., & Sharma, A. (2019). Composition analysis of renal stones. *Indian Journal of Urology*, 35(2), 126-131.

Smith, A. D., Rosen, M. P., & Dillman, J. R. (2019). CT for urolithiasis: Current status and future directions. *American Journal of Roentgenology*, 212(2), 297-305.

Smith-Bindman, R., Aubin, C., Bailitz, J., et al. (2014). Ultrasonography versus computed tomography for suspected nephrolithiasis. *New England Journal of Medicine*, 371(12), 1100-1110.

Sodickson, A. D., Prater, S., & Dreyer, K. J. (2015). Reclaiming clinical time for radiologists. *Journal of the American College of Radiology*, 12(12), 1397-1401.

Taylor, E. N., Stampfer, M. J., & Curhan, G. C. (2005). Dietary factors and the risk of incident kidney stones in men. *Journal of the American Society of Nephrology*, 16(9), 2725-2731.

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56.

Turk, C., Petrik, A., Sarica, K., Seitz, C., Skolarikos, A., Straub, M., & Knoll, T. (2016). EAU guidelines on diagnosis and conservative management of urolithiasis. *European Urology*, 69(3), 468-474.

Wahner-Roedler, D. L., Reed, D. A., & Sundsted, K. K. (2007). Fatigue and its impact on patient safety and quality of care. *Journal of Patient Safety*, 3(4), 199-205.

Westphalen, A. C., Hsiao, Y. S., & Zhao, S. (2015). Missed renal stones on CT: A retrospective review. *American Journal of Roentgenology*, 204(5), 1030-1035.

APPENDICES

Appendix A: Source Code

A.1 Model Training Script (train.py)

```
Python
from ultralytics import YOLO
import torch

def main():
    # Load pre-trained YOLOv8 model
    model = YOLO('yolov8n.pt')

    # Train the model
    results = model.train(
        data='dataset.yaml',
        epochs=100,
        batch=16,
        imgsz=640,
        lr0=0.001,
        optimizer='AdamW',
        weight_decay=0.0005,
        momentum=0.937,
        device=0,
        workers=8,
        patience=50,
        save=True,
        save_period=10,
        val=True,
        plots=True
    )

    print("Training completed successfully!")

if __name__ == "__main__":
    main()
A.2 FastAPI Backend (main.py)
python
from fastapi import FastAPI, File, UploadFile, HTTPException
from fastapi.responses import JSONResponse
from fastapi.middleware.cors import CORSMiddleware
from ultralytics import YOLO
```

```
import cv2
import numpy as np
import uvicorn
from datetime import datetime

app = FastAPI(title="AI Kidney Stone Detection API")

# Enable CORS
app.add_middleware(
    CORSMiddleware,
    allow_origins=["*"],
    allow_credentials=True,
    allow_methods=["*"],
    allow_headers=["*"],
)

# Load trained model
model = YOLO('best.pt')

@app.get("/")
async def root():
    return {"message": "AI Kidney Stone Detection API is running"}

@app.post("/detect")
async def detect_stone(file: UploadFile = File(...)):
    try:
        # Read and decode image
        contents = await file.read()
        nparr = np.frombuffer(contents, np.uint8)
        image = cv2.imdecode(nparr, cv2.IMREAD_COLOR)

        if image is None:
            raise HTTPException(status_code=400, detail="Invalid image file")

        # Perform inference
        results = model(image)

        # Extract detection results
        detections = []
        for r in results:
            boxes = r.boxes
            if boxes is not None:
                for box in boxes:
                    detection = {
                        'class': int(box.cls[0]),
```

```
                    'confidence': float(box.conf[0]),
                    'bbox': box.xyxy[0].tolist()
                }
                detections.append(detection)

        return JSONResponse(content={
            'status': 'success',
            'detections': detections,
            'stone_count': len(detections),
            'timestamp': datetime.now().isoformat()
        })

    except Exception as e:
        raise HTTPException(status_code=500, detail=str(e))

if __name__ == "__main__":
    uvicorn.run(app, host="0.0.0.0", port=8000)
```

A.3 HTML Frontend (index.html)

```
A.4 CSS Frontend (style.css)
css
body { html
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>AI Kidney Stone Detection Tool</title>
    <link rel="stylesheet" href="style.css">
</head>
<body>
    <div class="container">
        <h1>AI Kidney Stone Detection Tool</h1>
        <p>Fast, Precise AI for Kidney Stones</p>

        <div class="upload-area" id="uploadArea">
            <p>Click or drag image to upload</p>
            <input type="file" id="imageUpload" accept="image/jpeg,image/png">
        </div>
```

```
    <button id="analyzeBtn" onclick="analyzeImage()">Run AI Analysis</button>

    <div id="loading" style="display:none;">
        <p>Analyzing Scan... YOLO model is processing...</p>
    </div>

    <div id="results"></div>
</div>

<script>
    const API_URL = "http://localhost:8000/detect";

    async function analyzeImage() {
        const fileInput = document.getElementById('imageUpload');
        const file = fileInput.files[0];

        if (!file) {
            alert("Please select an image first");
            return;
        }

        document.getElementById('loading').style.display = 'block';
        document.getElementById('results').innerHTML = '';

        const formData = new FormData();
        formData.append('file', file);

        try {
            const response = await fetch(API_URL, {
                method: 'POST',
                body: formData
            });

            const result = await response.json();
            displayResults(result);
        } catch (error) {
            console.error('Error:', error);
            document.getElementById('results').innerHTML = '<p>Error analyzing image</p>';
        } finally {
            document.getElementById('loading').style.display = 'none';
        }
    }

    function displayResults(result) {
```

```
        const resultsDiv = document.getElementById('results');
        const stoneCount = result.stone_count;

        let html = `<h2>Analysis Results</h2>`;
        html += `<p>Stones Detected: ${stoneCount}</p>`;

        if (stoneCount > 0) {
            html += `<p class="warning">Critical detection of potential kidney stones. Medical review
recommended.</p>`;
        } else {
            html += `<p class="success">No stones detected. Result is normal.</p>`;
        }

        resultsDiv.innerHTML = html;
```

A.4 CSS Frontend (style.css)

```
Css
body { html
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>AI Kidney Stone Detection Tool</title>
    <link rel="stylesheet" href="style.css">
</head>
<body>
    <div class="container">
        <h1>AI Kidney Stone Detection Tool</h1>
        <p>Fast, Precise AI for Kidney Stones</p>

        <div class="upload-area" id="uploadArea">
            <p>Click or drag image to upload</p>
            <input type="file" id="imageUpload" accept="image/jpeg,image/png">
        </div>

        <button id="analyzeBtn" onclick="analyzeImage()">Run AI Analysis</button>

        <div id="loading" style="display:none;">
```

```
        <p>Analyzing Scan... YOLO model is processing...</p>
    </div>

    <div id="results"></div>
</div>

<script>
    const API_URL = "http://localhost:8000/detect";

    async function analyzeImage() {
        const fileInput = document.getElementById('imageUpload');
        const file = fileInput.files[0];

        if (!file) {
            alert("Please select an image first");
            return;
        }

        document.getElementById('loading').style.display = 'block';
        document.getElementById('results').innerHTML = '';

        const formData = new FormData();
        formData.append('file', file);

        try {
            const response = await fetch(API_URL, {
                method: 'POST',
                body: formData
            });

            const result = await response.json();
            displayResults(result);
        } catch (error) {
            console.error('Error:', error);
            document.getElementById('results').innerHTML = '<p>Error analyzing image</p>';
        } finally {
            document.getElementById('loading').style.display = 'none';
        }
    }

    function displayResults(result) {
        const resultsDiv = document.getElementById('results');
        const stoneCount = result.stone_count;

        let html = `<h2>Analysis Results</h2>`;
```

```
            html += `<p>Stones Detected: ${stoneCount}</p>`;

            if (stoneCount > 0) {
                html += `<p class="warning">Critical detection of potential kidney stones. Medical review recommended.</p>`;
            } else {
                html += `<p class="success">No stones detected. Result is normal.</p>`;
            }

            resultsDiv.innerHTML = html;
        }
    </script>
</body>
</html>

    font-family: Arial, sans-serif;
    margin: 0;
    padding: 20px;
    background-color: #f5f5f5;
}

.container {
    max-width: 800px;
    margin: 0 auto;
    background: white;
    padding: 30px;
    border-radius: 10px;
    box-shadow: 0 2px 10px rgba(0,0,0,0.1);
}

h1 {
    color: #2c3e50;
    text-align: center;
}

.upload-area {
    border: 2px dashed #3498db;
    border-radius: 10px;
    padding: 40px;
    text-align: center;
    margin: 20px 0;
    cursor: pointer;
}

button {
```

```
   background-color: #3498db;
   color: white;
   padding: 12px 24px;
   border: none;
   border-radius: 5px;
   cursor: pointer;
   width: 100%;
   font-size: 16px;
}

button:hover {
   background-color: #2980b9;
}

.warning {
   color: #e74c3c;
   font-weight: bold;
}

.success {
   color: #27ae60;
   font-weight: bold;
}
```

Appendix B

Dataset Configuration

B.1 Dataset YAML Configuration (dataset.yaml)

```
Yaml
path: ./dataset
train: images/train
val: images/val
test: images/test

nc: 1
names: ['stone']
```