# A Dual Stage Computational Framework for Automated Classification and Precision Segmentation of Kidney Stones in CT Scans

Zuhier Humady Hussien [1, 2], Rabab Saadoon Abdoon[1] and Nihad Abdulameer Salih[1]
[1]*Department of Physics, Faculty of Science, University of Babylon, Al-Najaf Road, 51002 Hillah, Babylon, Iraq*
[2] *Al-Furat Al-Awsat Teaching Hospital, Najaf Health Directorate, Iraqi Ministry of Health, 54001 Najaf, Iraq*
*{sci753.zuhair.hamadi, sci.rabab.saadoon, sci.nihad.abdul}@student.uobabylon.edu.iq*

Keywords: Kidney Stones, CT Scans, Deep Learning, Resnet-50, Image Segmentation, MSER, Region Growing.

Abstract: The incidence of renal calculi is a serious global health concern, and accurate diagnostic procedures play a key role in treating this disease. The current study proposes a dual-stage framework to automate the classification and accurate segmentation of kidney stones utilizing a computed tomography scan image. In the first part, a CNN model using a transfer-learning ResNet-50 architecture was formulated to classify renal stone images into normal, fragmentable (F-type), and non-fragmentable (NF-type) types. The proposed CNN architecture was trained to yield a high diagnostic accuracy, with the final network achieving 99.39% accuracy. In the second stage, the image segmentation is performed using the following three methods: maximally stable extremal regions (MSER), Region Growing, and a proposed new Hybrid image segmentation technique. The hybrid proposed image segmentation technique leverages a key strategy that combines region growing and MSER segmentation to extract the periphery of the kidney stone and yield accurate morphological descriptors. The results indicate that the three methods used in this research are anatomically consistent. Also, the segmentation of pre-filtered images minimizes errors and demonstrates efficacy and feasibility to optimize ESWL procedures to increase efficiency through accurate kidney stone analysis.

## 1 INTRODUCTION

In this study, the perception of the kidney stone classification, F-type (fragmentable stones), NF-type (non-fragmentable stones), and images that do not contain a stone (Normal) was employed by employing the multi-step framework in accordance with the CNN algorithm. This has led to the application of the study of classification and segmentation as an integral solution to optimize ESWL efficiency [1], [2]. In this work, the study has employed three methods in the form of segmentation to attain appropriate accuracy in the description of kidney stone boundaries. They are Maximally Stable Extremal Regions (MSER), Region Growing, and a hybrid of the aforementioned. Theoretically, the MSER method is characterized by its resistance to linear transformations of varying image intensity, while the growth of the region ensures spatial connectivity; combining the two methods aims to take advantage of intensity stability and regional homogeneity. To ensure clinical accuracy, the performance of the three segmentation methods was verified using accuracy (%) and percentage relative difference (PRD) to attain optimal precision [3], [4]. The primary contribution of this work lies in the synergy between deep transfer learning for type identification and morphological analysis for stone extraction, providing a precise tool for urological diagnostics despite high-density interference from adjacent bone structures.

## 2 MATERIALS AND METHODS

### 2.1 Experimental Dataset and Data Acquisition

Throughout this research, MATLAB (version 2023a) was used to perform all the required computations

related to image enhancement, classification, and segmentation.

The system was validated using two datasets: a large repository of 12,000 computerized tomographic (CT) images obtained from the Kaggle database [5], which were used for initial model training. Advanced segmentation was validated on 24 clinical helical CT scans. Eight images were used in this research, four of which were of fragmentable stones and four of which were of non-fragmentable kidney stones. These studies were acquired on a Siemens scanner, with a slice thickness of 1 mm, for detailed evaluation of the results within the framework of different segmentation techniques. The provided clinical data had been verified by a specialist urologist in relation to. To maintain diagnostic accuracy and to serve as a manual reference for performance verification.

## 2.2 Theoretical Foundation of Automated Classification

The first stage of the provided system relies on deep transfer learning for the automatic identification of stone types. The theoretical advantage of this technique is based on the utilization of resnet50, which is a deep residual learning approach for directly extracting feature representations from raw data [6]. By training the architecture on a purified set, it can bypass the vanishing gradient problem via the corresponding shortcut links, allowing for high-precision differentiation of normal, fragmentable, and non-fragmentable kidney anatomy [7]. The training process is optimized through the utilization of the Adam optimizer, which is known for stable weight convergence in complex medical image analysis [7], [8].

## 2.3 Theoretical Basis for Advanced Segmentation

The second stage adopts a multi-algorithmic method to detect stone boundary, supported by the following theoretical bases:

1) Maximally stable extremal regions (MSER). This approach uses the idea of maximally stable components that remain stable under varying levels of thresholding. Due to its robust nature, localizing high-density calcified structures in challenging radiographic backgrounds can be achieved with this method [10].
2) Region Growing. This approach is based on spatial continuity and similarity of local intensity characteristics. The method uses an automated initial seed point. Through a series of iterations, neighboring pixels are accumulated, provided they satisfy given variance requirements to ensure a coherent stone body extraction in the CT images [9], [11].
3) Hybrid approach (Region Growing+MSER). In this new integration, a synergy is created between region growing connectivity and MSER structural stability. Ideally, theoretically, this integration is expected to overcome limitations in various algorithms, including the reduction of interferences from neighboring bone structures due to similarities in radiodensity [12].

## 2.4 Proposed Methodological Workflow

Figure 1 shows a systematic summary of the integrated operational sequence for the developed framework, presenting a logical sequence from data acquisition to final clinical diagnosis.

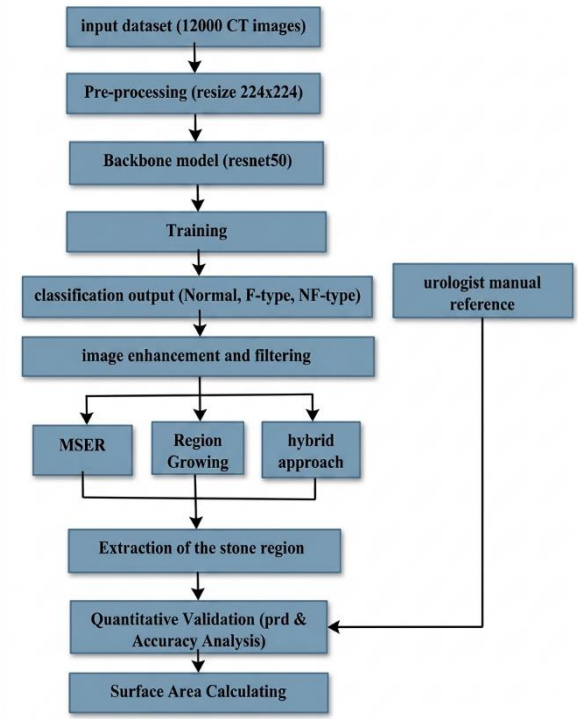



Figure 1: Comprehensive methodological workflow of the proposed dual-stage system for renal stone classification and precision segmentation.

As shown in Figure 1, the system has been divided into four phases of clinical reliability. The first stage uses a convolutional neural network (CNN) to identify normal kidneys and classify the type of stones in the affected kidney. The second phase includes edge improvement using filters to increase the edge gradient. Next, a special segmentation process by MSER, Region growing, and the Hybrid approach is carried out by the system in its third phase. Finally, the surface area, extracting, and morphological features are extracted (equivalent diameter, perimeter, eccentricity, center of mass) for a quantitative validation of (PRD) and an accuracy analysis compared to the manual drawing by a urologist. This takes place in the fourth stage of the workflow.

## 2.5 Stage I: Deep Transfer Learning for Automated Classification

### 2.5.1 Data Acquisition and Purification

In the initial step, the classification of renal CT images is performed into three distinct categories: images of normal kidneys (Normal), images of kidneys with fragmentable-type stones (F-type), and images of kidneys with non-fragmentable-type stones (NF-type). The main dataset used in the study entails 12,000 images from the Kaggle dataset [5], They were distributed equally across the three files, with 4000 images per file.

The dataset was subjected to an auditing process to ensure that the highest level of diagnostic reliability was achieved. During this process, inconsistent image sites were then corrected. The verification was both crucial and useful to remove ambiguity regarding the classification of different types of stones based on whether they were fragmentable or non-fragmentable for the convolutional neural network (CNN) model to learn from anatomical details at high resolution, as opposed to inconsistent classification across the three files.

### 2.5.2 Network Architecture and Fine-Tuning

The classification framework leverages the resnet50 backbone [13], which is a deep architecture known for its efficiency in medical feature extraction [2]. A transfer learning strategy was adopted-a pre-trained model, full fine-tuning on the purified renal dataset. For the new sub-network, the original classification layers were removed and replaced with a fully connected layer designed for three classes, a softmax layer, and a final classification output layer [14]. The fully connected layer behaves as a classifier, mapping the extracted features to the three target categories. It is followed by a softmax layer that converts the numerical output scores into a probability distribution so that the final prediction will be delivered as a confidence level for every class. To accelerate the adaptation, a WeightLearnRateFactor of 20 was introduced into the new layers, forcing the network to give high importance to the unique textural patterns of kidney stones.

Table 1: Comprehensive summary of training configurations and parameters for the proposed CNN model.

| Category | Parameter | Value |
|---|---|---|
| Dataset Statistics | Total images | 12,000 images |
| | Number of classes | 3 (Normal, F-type, NF-type) |
| | Training samples | 8,400 images |
| | Validation samples | 1,800 images |
| | Testing samples | 1,800 images |
| Model Architecture | Backbone architecture | ResNet50 |
| | Input image size | 224 x 224 (RGB) |
| | Data augmentation | Rotation, Translation, Reflection |
| Hyperparameters | Optimizer | Adam |
| | Initial learning rate | 0.0001 (1e-4) |
| | Learning rate schedule | Piecewise (Drop factor 0.1 every 5 epochs) |
| Training Performance | Mini-batch size | 8 |
| | Max epochs | 15 |
| | Total iterations | 15,750 |
| | Iterations per epoch | 1,050 |
| | Hardware resource | Single GPU |
| | Elapsed training time | 766 minutes (approx. 12.8 hours) |

### 2.5.3 Hyperparameter Optimization and Experimental Tuning

The model is optimized using an (Adam) optimizer [8]. To find out the optimal training configuration, several hyperparameters were tested. In this model, it was trained for 15 epochs with a mini-batch size of 8, which is the number of images as a single batch that are processed in each iteration to

update the model weights. A piecewise learning rate schedule was used, where the model is reduced by a factor of 0.1 at regular intervals of 5 epochs, starting with an original rate of 0.0001. represents the initial learning rate that contributes to controlling the step size during the optimization process. This intensive training took about 12.8 hours to reach a plateau [3]. Finally, the training configurations are briefly described in Table 1.

The training procedure was instrumental in ensuring that the model was able to capture the minute textural characteristics of kidney stones while ensuring maximum generalization and minimal chances of overfitting on CT scan images. The optimized arrangement, as outlined in Table 1, was instrumental in ensuring precise weight convergence for the resnet50 model.

### 2.5.4 Baseline Comparison and Selection

In justifying the choice of the model, the model was tested against two preliminary experimental configurations, one using a simple CNN model, while the other employed the optimized model with the SGDM optimizer [15]. As indicated in Figure 2, the shift from the basic models to the optimized model showed considerable improvement in the accuracy of the model on the test data, thus validating the implementation of the deep residual learning model in solving the diagnostic problem.

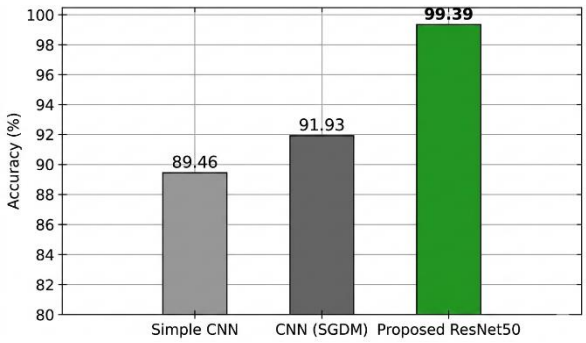



Figure 2: Performance evolution across different experimental model configurations.

Explanation of the data in Figure 2 outlines the fact that the optimized resnet50 model is significantly better than the standard CNN models. Such high accuracy is indicative of the robustness of the model in recognizing the minute patterns in kidney stones with near-perfect precision.

### 2.5.5 Performance Optimization and Training

To ensure maximum precision in the diagnostic capabilities of the resnet50 model, the model was put through a systematic hyperparameter optimization procedure. The procedure was Important role in ensuring that the model was able to converge optimally under varying parameters, including the resolution of the images, the initial learning rate, and the training duration. The experiment was able to establish that the standard resolution of 224x224 pixels was critical in ensuring that the model was able to capture the subtle characteristics of kidney stones. Additionally, the initial learning rate of 0.0001, regulated under a piecewise schedule, was critical in ensuring maximum precision.

The training process of the optimized model is depicted in Figure 3. As indicated by the accuracy and loss curves, the system was able to converge at a stable accuracy of 98.83% after 15 epochs or 15,750 iterations. The total training time taken by the model, which was 12.8 hours, ensured that the model was able to learn the textural patterns rather than memorizing the samples, thus ensuring convergence where the loss function was reduced to zero. This large training period is scientifically necessary for deep networks to achieve optimal minimum weight stability, a process that has been confirmed in recent sources to directly aid in higher diagnostic accuracy in complex medical datasets [16].

This Figure 3 further emphasizes the stability of the convergence process in the resnet50 model during the entire training phase. The low variance between the training and validation curves indicates high levels of generalization, thus confirming the successful training process in learning the characteristics of kidney stones without overfitting (over-personalizing) [10].

### 2.5.6 Classification Evaluation and Benchmarking Protocol

To validate the diagnostic capabilities of the optimized resnet50 model, the diagnostic efficacy was tested using an independent test set. The classification evaluation was conducted in accordance with the following key protocols:

1) Generating a confusion matrix to analyze the classification errors.
2) Generating comprehensive metrics to calculate precision, sensitivity, specificity, f1-score, etc.

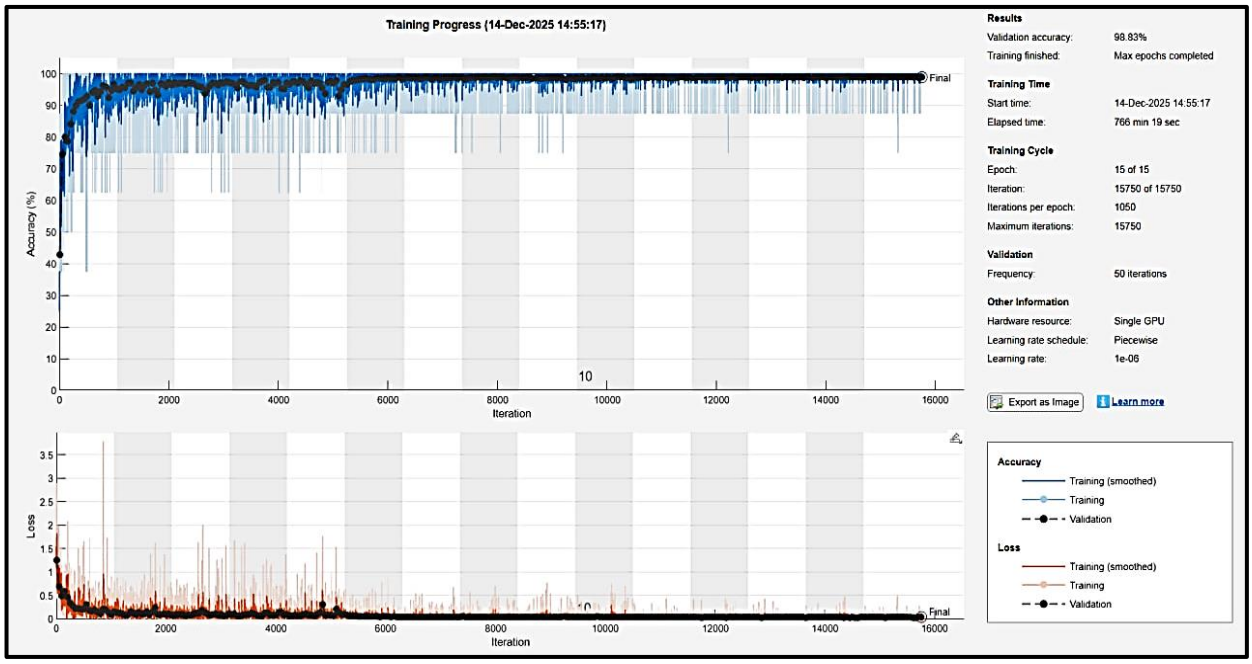
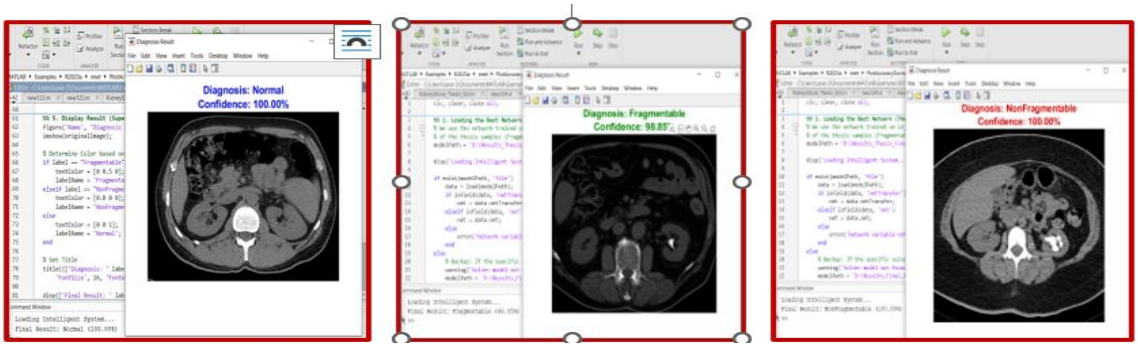



Figure 3: Training and validation progress of the optimized CNN model.




(a) (b) (c)

Figure 4: Sample GUI diagnostic outputs for a) normal, b) fragmentable, and c) non-fragmentable stones.

To validate the clinical importance of the study, the classification evaluation was benchmarked against the state-of-the-art research in the domain. The numerical outcomes of the benchmarking process, along with the subsequent statistical analysis, have been included in the results section.

Upon validation, this model succeeded in achieving a High accuracy percentage in testing accuracy, namely 99.39%, as shown in Figure 3, Consequently, the framework was used in the creation of an intelligent diagnostic interface (GUI) with the intent of ascertaining the practical applicability of the proposed system in diagnosing kidney stones, as shown in Figure 4.

The notes are clearly shown in Figure 4 emphasizing the capabilities of the system in providing instantaneous and accurate diagnostic predictions. Each of the sample cases provided emphasizes the GUI capabilities in interpreting the complex resnet50 classification systems and simplifying classifications into an easy to apply formula. Such successful integration provides a connection between the complex

level of classification systems and practical applications of urological surgery. Table 2 shows the classification results for the images used in the research, ensuring a level of transparency and reliability in diagnostic predictions. Having established a level of high-performance classification, the subsequent section provides an overview of the implementation of advanced segmentation systems for the accurate extraction of stone region boundaries.

Table 2: Automated diagnosis results for the test images through the GUI.

| Images | Predicted Class (GUI) | Confidence Score (%) |
|---|---|---|
| Stone_1_F | Fragmentable | 99.92 |
| Stone_2_F | Fragmentable | 98.93 |
| Stone_3_F | Fragmentable | 99.75 |
| Stone_4_F | Fragmentable | 99.98 |
| Stone_1_NF | Non-Fragmentable | 100 |
| Stone_2_NF | Non-Fragmentable | 100 |
| Stone_3_NF | Non-Fragmentable | 100 |
| Stone_4_NF | Non-Fragmentable | 99.98 |

## 2.6 Stage II: Kidney Stone Segmentation

The second stage of the proposed system provides an overview of extracting the stone region from the classified CT images in the first stage. This includes an initial enhancement phase, followed by the application of segmentation systems for accurate quantification.

### 2.6.1 Image Enhancement

To improve the clarity level in CT images and enhance the boundaries of kidney stones, a sequential filtering system is applied to the images, incorporating the Prewitt, Laplacian, and unsharp masking (USM) filters. The progressive effect of the enhancement filters is provided in Figures 5 and 6 for fragmentable and non-fragmentable kidney stones, respectively.

The visual data presented in Figures 5 and 6 comprise Five columns. The first column represents the original (CT) images, the second column represents the Laplacian-filtered images, the third column represents the output of the Prewitt filter, which is used to detect edges, the fourth column represents the output of the unsharp masking filter, and the fifth column represents the enhanced image (after applying the three filters) This series of image processing enhances the contrast between the stones and the other tissues in the image, creating a high-fidelity input image that is necessary to obtain accurate extraction in the subsequent stages of segmentation. The mathematical criterion used to evaluate the performance of the image filters is as follows:

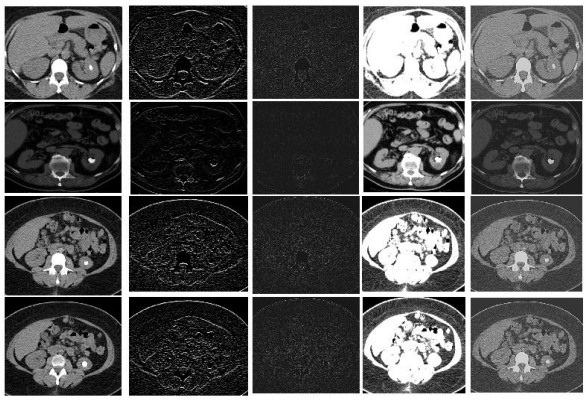


Figure 5: Enhancement results for fragmentable kidney stones.

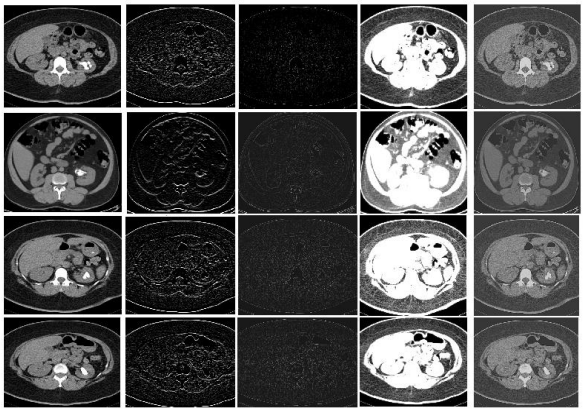


Figure 6: Enhancement results for non-fragmentable kidney stones.

### 2.6.2 Image Quality Assessment Metrics

In order to evaluate the performance of the image enhancement using the adopted filters, before the segmentation stage, three different image quality assessment metrics were used to evaluate the performance of these filters in suppressing image noise

and enhancing the structural fidelity of the image [3]. The evaluation process is based on the following mathematical criteria:

1) Mean Squared Error (MSE). MSE is an image quality metric that measures the mean square difference between the original image and the image obtained after applying the adopted filter. If the MSE is decreasing, it implies that the image has minimal distortion. On a standard scale [0, 1], values less than 0.1 represent high resolution and high similarity to the original image, as expressed in the following mathematical expression [17]:

$$MSE = 2[K(i,j) - I(i,j)] \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} \frac{1}{mn} . \quad (1)$$

where I (i, j) and K (i, j) refer to the pixel values of the original and filtered images, respectively. Lower MSE indicates high conformity and less distortion.

2) Peak Signal-to-Noise Ratio (PSNR). Measured in decibels (dB), PSNR assesses image quality relative to noise levels. The highest level above 30 dB, PSNR, leads to better image restructuring and signal preservation. Values below 20 decibels indicate high noise levels and are expressed by the following equation [3]:

$$\text{PSNR} = \left(\frac{I^2_{MAX}}{MSE}\right) 10 \log .10, \quad (2)$$

where IMAX is the maximum pixel value. Higher PSNR values reflect better image quality [17].

3) Structural Similarity Index Measure (SSIM). The SSIM evaluates perceived image quality by comparing luminance, contrast, and structural information. The values range from -1 to 1, where a value approaching 1.0 very high and almost perfect structural match [4]:

$$SSIM(\chi, \gamma) = \frac{\left(2^{2}\sigma xy + C\right)(1^{2}\mu\chi\mu\Upsilon + \mathrm{C})}{\left(2\sigma^2_{\chi} + \sigma^2_{\Upsilon} + \mathrm{C}\right)(1\mu^2_{\chi} + \mu^2_{y} + \mathrm{C})} \quad (3)$$

The quantitative performance data for these metrics across both fragmentable and non-fragmentable stones. This topic was discussed in detail in the Results section, where the optimal filtering approach is identified for the subsequent segmentation phase [18].

### 2.6.3 Segmentation Algorithms

#### 2.6.3.1 Maximally Stable Extremal Regions Method (MSER)

This method works by assessing the stability of the parts connected across a sequential range of density thresholds. A region is said to be maximally stable if its size remains relatively constant even as the threshold is varied. This makes it particularly useful for locating calcified objects [10]. In this study, the MSER detector was parameterized with a threshold delta of 17.5, like kidney stones. About overlapping radial backgrounds, the MSER detector was configured using a delta threshold of 17.5 and a region area range of 200 to 1000 pixels. This enables the system to hone in on similar to kidney stones by concentrating on regions with strong intensity stability, while filtering out distracting anatomical clutter, as shown in Figures 7 and 8.

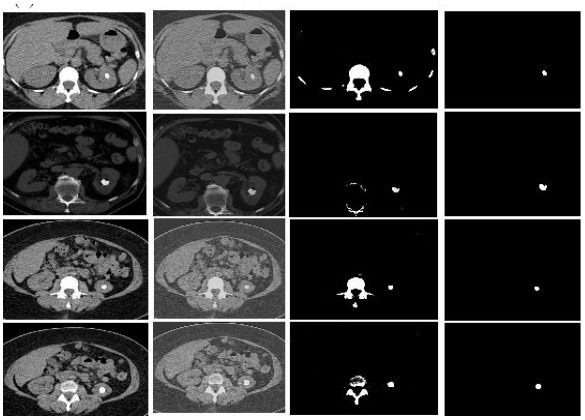


Figure 7: MSER segmentation results for fragmentable kidney stones.

Figures 7 and 8 show the results of the detection, which have been arranged in four columns. The first column represents the original input image, the second column represents the image after applying the three filters, the third column represents the stones and bones, and the fourth column represents the extracted stone. In the case of fragmentable stones, the MSER algorithm provided smooth and compact segmentation results

with clear boundaries. In the case of non-fragmentable stones, there was some partial overlap in the initial stages due to the share of the high-density intensity range with the bone structures. The MSER algorithm provided robust results, and the structural basis of the results confirms the accuracy of the subsequent steps of the extraction process.

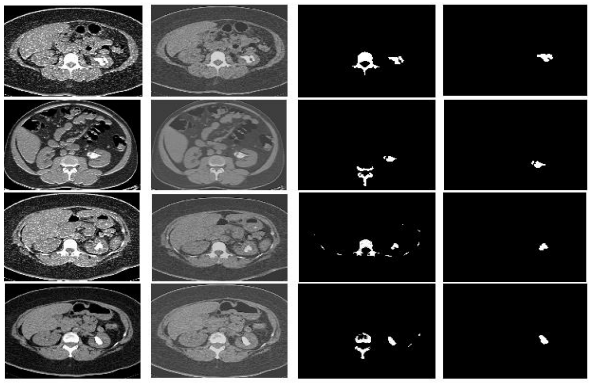


Figure 8: MSER segmentation results for non-fragmentable kidney stones.

### 2.6.3.2 Region Growing Method

The region growing method is employed to detect the kidney stones by utilizing the local intensity uniformity and spatial continuity. The basic idea behind this technology is the automatic selection of the starting point, which is the centroid of the region of the highest intensity [10]. Then, the algorithm continues to add the neighboring points that meet the criteria of the similarity measure, which is usually set to a threshold of 0.35, serving as a tolerance parameter that defines the maximum allowable intensity deviation for pixel inclusion of the intensity difference of the region's mean intensity. The principle of iterative expansion ensures that the extracted object retains a coherent structure, which is crucial and highly effective in identifying calcifications in helical CT scan images [19]. The outcomes of this spatial expansion process are presented in Figures 9 and 10 for both fragmentable and non-fragmentable stones.

Figures 9 and 10 show that the first column represents the input images, the second column shows the growth maps of the primary area, the third column shows the areas of bone and aggregated stone, and the fourth column depicts the final isolated stone. soft outer borders and the low level of background noise for fragmentable kidney stones. While non-fragmentable stones, being of higher density, overlap adjacent bone structures, post-processing techniques such as filling of holes and morphological area filtering can alleviate these problems to a certain extent. Thus, this method facilitates easier identification of stones and their boundaries in the clinical data set.

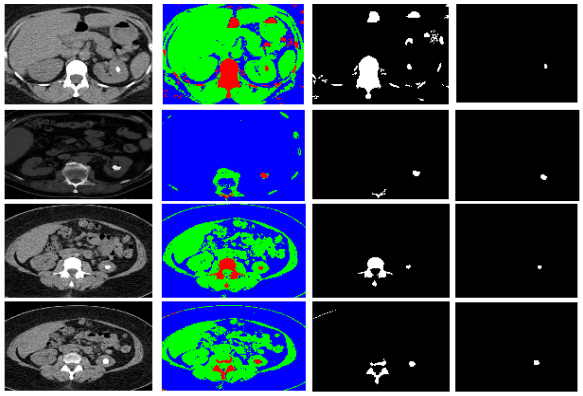


Figure 9: Segmentation of fragmentable kidney stones using the region growing approach.

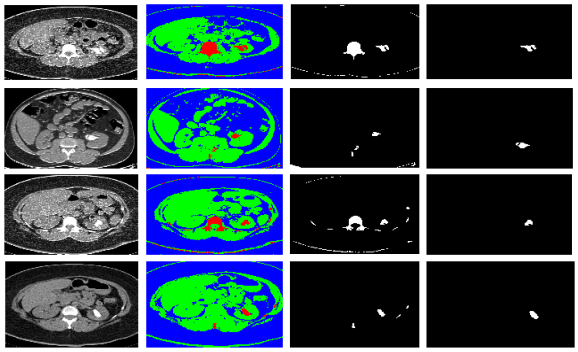


Figure 10: Segmentation of non-fragmentable kidney stones using the region growing approach.

### 2.6.3.3 Hybrid Segmentation Method

The real contribution of this research is the hybrid segmentation method, which combines the advantages of region growing and the maximally stable extremal region (MSER) technique. Region growing maintains good connectivity in homogeneous regions but fails to stop at high-density bone boundaries. On the other hand, although the MSER method is efficient in identifying stable and calcified structures, it may result in fragmented results in heterogeneous regions. By

integrating these two techniques, this method uses region growing for global homogeneity mapping and MSER for local extremal stability, creating a hybrid effect to improve the boundaries of structures while minimizing segmentation artifacts. The effectiveness of this hybrid method over the individual techniques is evident from Figures 11 and 12.

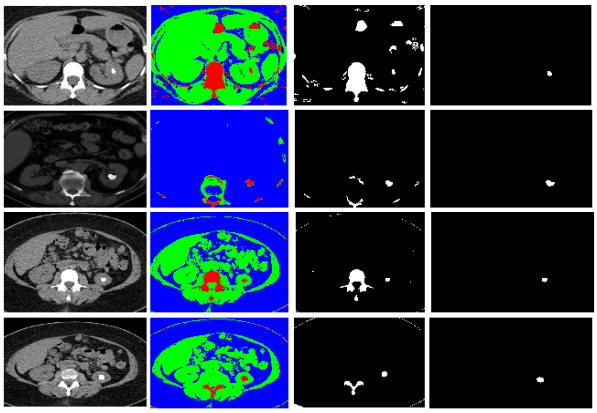


Figure 11: Results of the proposed hybrid segmentation for fragmentable kidney stones.

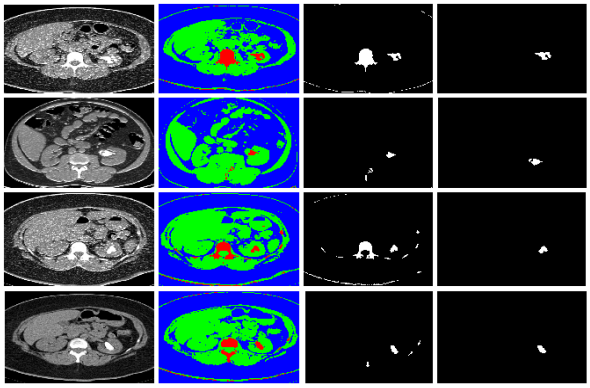


Figure 12: Results of the proposed hybrid segmentation for non-fragmentable kidney stones.

The results obtained using the Hybrid method, as presented in Figures 11 and 12, are arranged in four columns. The first column represents, includes the input scan, the second column represents the initial Region growing expansion, the third column represents the stabilization of the structure using the MSER algorithm, and the final isolates of the stone region. The fine results, depicted in the final column, have been obtained using a suite of post-processing refinement tools. First, the hybrid fusion, a multi-algorithmic technique, underwent a set of morphological opening and closing operations to smooth the boundaries of the stones. Secondly, area-based filtering and hole-filling techniques have been useful for ensuring the integrity of the stone extraction from the medical image. The results obtained proved that the Hybrid method is able to combine the results of the various algorithms in a way that is able to successfully separate the stone from the dense, homogeneous bone tissue surrounding it.

## 2.7 Reference Measurement by Urologist

In order to develop a scientifically sound benchmark to assess the accuracy of the proposed segmentation algorithms, a manual delineation process was performed by a specialist urologist. This process provides a ground truth and offers the required reference to assess the accuracy of automated extraction results and to ensure that the results obtained by computer algorithms match To ensure results meet official surgical standards [2], [7]. The process followed to develop the required reference dataset includes the following steps:

1) Image Selection and Annotation. Images were printed at high resolution. A transparent layer was placed on top of it with a thickness of 0.125 mm was employed to accurately delineate the boundaries of kidney stones using a fine-tip marker.
2) Digitization and Binarization. The images were scanned using a high-resolution scanner, and an image binarization technique was employed.
3) Quantitative processing. A quantitative processing of the annotated images was the same computational steps applied to the outputs of automated hashing were used in the three previous methods to calculate the surface area and geometric properties. As shown in the two Figures 13 and 14.

Figures 13 and 14 show the certified and high-precision reference system, which is divided into three descriptive columns: the original CT scan image with expert annotations, the scanned manual drawings, and the final binary negative image. The level of accuracy in the contours produced by the expert guarantees that the subsequent quantitative assessment, such as the calculation of the percentage relative difference (PRD) and accuracy, is performed objectively.

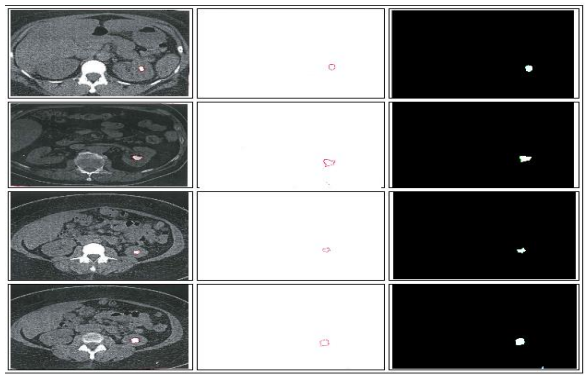


Figure 13: Urologist manual delineation generation for fragmentable kidney stones.

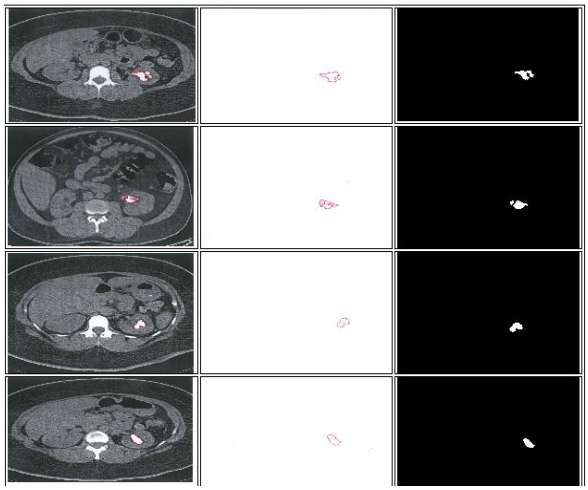


Figure 14: Urologist manual delineation generation for non-fragmentable kidney stones.

# 3 RESULTS

## 3.1 Automated Classification Performance

The classification performance of the optimized ResNet50 model was evaluated using an independent testing dataset consisting of 1,800 CT images. The proposed framework achieved an overall classification accuracy of 99.39%, indicating a high capability for distinguishing between normal kidneys, fragmentable stones, and non-fragmentable stones. The corresponding confusion matrix is presented in Figure 15.

Figure 15 demonstrates that the model achieved perfect recall for both normal kidney images and non-fragmentable stones, with all samples in these categories correctly identified. Only a small number of fragmentable stone images were misclassified, confirming the robustness and stability of the proposed classification framework.

To further evaluate the diagnostic reliability of the model, several standard performance metrics were calculated, including precision, sensitivity, specificity, and F1-score [4]. These metrics provide a comprehensive assessment of classification quality and diagnostic consistency. The final quantitative results are summarized in Table 3.

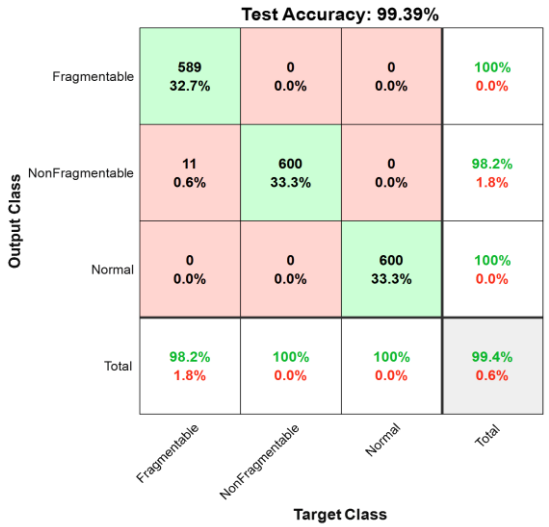



Figure 15: Confusion matrix showing the test performance of the three types of the kidney classes.

To establish the importance of these findings, the current work has been benchmarked against available literature. Table 4 provides a comparative summary of the proposed system with respect to recent literature to establish the effectiveness of the implemented system.

This study was compared with the latest studies published in international journals to demonstrate the importance of the results obtained. As indicated in Table 4, the proposed model outperformed the other applied models significantly. This proves that the optimized resnet50 framework is reliable and effective in the automated diagnostic process for kidney stones.

Table 3: Final diagnostic performance metrics for the proposed CNN model.

| Class Name | Precision (%) | Sensitivity % | Specificity (%) | F1-Score (%) |
|---|---|---|---|---|
| Fragmentable | 100.00 | 98.17 | 100.00 | 99.07 |
| Non-Fragmentable | 98.20 | 100.00 | 99.08 | 99.09 |
| Normal | 100.00 | 100.00 | 100.00 | 100.00 |
| Overall Average | 99.40 | 99.39 | 99.69 | 99.39 |

Table 4: Comparative analysis of classification accuracy with recent literature.

| Author / Reference | Year | Model Used | Precision % | Sensitivity % | Specificity% | F1-Score % | Accuracy % |
|---|---|---|---|---|---|---|---|
| Yildirim et al. [20] | 2021 | CNN ResNet-50 | 96.50 | 96.20 | 97.10 | 96.35 | 96.82 |
| Sudharson and Kokil. [21] | 2021 | CNN (ResNet+VGG) | 97.20 | 96.90 | 97.50 | 97.05 | 97.10 |
| Asif et al. [9] | 2024 | Fusion CNN | 98.80 | 98.70 | 99.10 | 98.75 | 98.81 |
| Present work | 2026 | CNN Optimized ResNet50 | 99.40 | 99.39 | 99.69 | 99.39 | 99.39 |

Table 5: Quantitative evaluation of filtering techniques for fragmentable kidney stones.

| Method | Laplacian Filtered | | | Prewitt filtered | | | Unsharp masking (USM) Filter | | |
|---|---|---|---|---|---|---|---|---|---|
| Image | MSE | SSIM | PSNR | MSE | SSIM | PSNR | MSE | SSIM | PSNR |
| Stone_1_F | 0.21 | 0.005 | 6.773 | 0.26 | 0.01 | 5.837 | 0.0005 | 0.96 | 33.16 |
| Stone_2_F | 0.03 | 0.085 | 14.71 | 0.05 | 0.07 | 13.45 | 0.0007 | 0.98 | 41.21 |
| Stone_3_F | 0.12 | 0.079 | 9.061 | 0.17 | 0.10 | 7.567 | 0.0004 | 0.96 | 33.66 |
| Stone_4_F | 0.12 | 0.081 | 9.248 | 0.17 | 0.10 | 7.669 | 0.0004 | 0.96 | 33.65 |

Table 6: Quantitative evaluation of filtering techniques for non-fragmentable kidney stones.

| Method | Laplacian Filtered | | | Prewitt filtered | | | Unsharp masking (USM) Filter | | |
|---|---|---|---|---|---|---|---|---|---|
| Image | MSE | SSIM | PSNR | MSE | SSIM | PSNR | MSE | SSIM | PSNR |
| Stone_1_NF | 0.1180 | 0.109 | 9.2792 | 0.173 | 0.1254 | 7.5981 | 0.0004 | 0.9602 | 33.8085 |
| Stone_2_NF | 0.107 | 0.170 | 9.6668 | 0.149 | 0.181 | 8.2424 | 0.0004 | 0.9685 | 34.58374 |
| Stone_3_NF | 0.130 | 0.123 | 8.8417 | 0.183 | 0.138 | 7.3644 | 0.0004 | 0.9616 | 33.97120 |
| Stone_4_NF | 0.122 | 0.118 | 9.1261 | 0.176 | 0.132 | 7.5442 | 0.0003 | 0.9612 | 34.01071 |

Table 7: Comparative surface area (pixels) and PRD analysis for the three ways.

| Image ID | Urologist Reference | MSER | | Region Growing | | Hybrid Method | |
|---|---|---|---|---|---|---|---|
| | Area | Area | PRD % | Area | PRD % | Area | PRD % |
| Stone_1_F | 195 | 202 | 3.59 | 192 | 1.54 | 193 | 1.03 |
| Stone_2_F | 472 | 478 | 1.27 | 473 | 0.21 | 482 | 2.12 |
| Stone_3_F | 268 | 260 | 0.57 | 269 | 0.37 | 270 | 0.75 |
| Stone_4_F | 531 | 541 | 2.99 | 522 | 1.69 | 530 | 0.19 |
| Stone_1_NF | 1271 | 1256 | 1.18 | 1248 | 1.81 | 1253 | 1.42 |
| Stone_2_NF | 472 | 468 | 0.85 | 461 | 2.33 | 462 | 2.12 |
| Stone_3_NF | 784 | 782 | 0.26 | 746 | 4.85 | 762 | 2.81 |
| Stone_4_NF | 812 | 820 | 0.99 | 818 | 0.74 | 824 | 1.48 |

Table 8: Consolidated geometrical parameter analysis for kidney stone.

| Image ID | Parameter | MSER | Region Growing | Hybrid Method | Urologist reference |
|---|---|---|---|---|---|
| Stone_2_F | equivalent diameter | 25.15 | 24.38 | 24.77 | 24.20 |
| | perimeter (pixels) | 87.89 | 81.39 | 82.97 | 80.25 |
| | eccentricity | 0.640 | 0.648 | 0.642 | 0.6576 |
| | center of mass (x, y) | 304, 360 | 304, 360 | 304, 360 | 304, 360 |
| Stone_2_NF | equivalent diameter | 22.54 | 22.31 | 22.32 | 22.76 |
| | perimeter (pixels) | 90.36 | 86.84 | 86.83 | 90.17 |
| | eccentricity | 0.6713 | 0.698 | 0.697 | 0.6727 |
| | center of mass (x, y) | 211, 265 | 211, 265 | 211, 265 | 211, 265 |

## 3.2 Evaluation of Image Filtering Performance

The effectiveness of the enhancement techniques applied to the images in the clinical dataset was evaluated using the metrics defined in Section 2.6.2. Tables 5 and 6 consolidate the quantitative results for fragmentable and non-fragmentable stones, respectively, for a comparative analysis of the mean squared error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity index measure (SSIM).

The results quantified in Tables 5 and 6 show how the unsharp masking (USM) filter invariably performed better within all categories. Compared to conventional quality metrics, an ideal MSE approaches zero, a high-quality PSNR exceeds 30 db, and an SSIM value approaches 1.0. The unsharp masking (USM) filter showed impressive fidelity: it produced an impressively low MSE (averaging ≈ 0.0004) and a superior PSNR (averaging > 33 db), which combined confirmed high signal integrity and residual noise. Moreover, the high SSIM values (> 0.96) indicate that the structural information from the original CT images was highly preserved. For contrast, the other two filters, Laplacian and Prewitt, failed to achieve the minimum threshold of 20 db with low SSIM values, hence failing to retain key anatomical details. The unsharp masking (USM) approach is thus identified as the best technique for enhancement because of its high-fidelity base for accurate stone segmentation.

## 3.3 Quantitative Validation of Stone Segmentation and Area Estimation

After the pre-processing stage, the three advanced algorithms were tested on eight representative clinical cases, including four fragmentable and four non-fragmentable stones. To determine the deviation between the automated process of stone extraction and the clinical ground truth, the percent relative difference (PRD) was used as the primary accuracy metric to perform a normalized evaluation of the differences in surface area, as indicated by the following formula [22]:

$$\mathrm{PRD} = \left|\frac{\mathrm{A_{method}} - \mathrm{A_{reference}}}{\mathrm{A_{reference}}}\right| \times 100\ \% \quad (4)$$

where ($\mathrm{A_{method}}$) is an automatically calculated stone area, and ($\mathrm{A_{reference}}$) is an area labeled by a urologist. It is possible to judge the quality of the segmentation based on this formula, where smaller PRD values indicate higher accuracy between the manual and automated stone area calculation methods. Calculating the stone area and the PRD value for fragmentable kidney stones below in Table 7.

Table 7 shows a summary of quantitative extraction results, where the surface area extracted using the MSER, Region growing, and Hybrid techniques is compared with that of the reference urologist. The results of the extracted area obtained by the three methods indicate a high degree of agreement with the findings of the reference urologist. with most of the results remaining well below the 5% threshold of acceptability. It is interesting to note that the hybrid method had the highest precision, reaching a PRD of 0.19% for stone_4_f; the highest value was 2.81 for stone_3_NF. This confirms its high ability to accurately define the boundaries of kidney stones despite density variations in CT scan images.

## 3.4 Geometrical Consistency and Spatial Accuracy Analysis

To assess the geometrical accuracy of the algorithms used in this study, geometrical analysis was conducted

on representative cases. To avoid repetition and to provide brevity and conciseness of the data, only two cases were selected, namely, stone_2_F and stone_2_NF, for fragmentable and non-fragmentable kidney stones, respectively. Quantitative evaluation of geometrical invariants is important to assess the accuracy of kidney stone extraction algorithms in medical scenarios [11]. As shown in the following results in Table 8, geometrical parameters such as equivalent diameter, perimeter, eccentricity, and center of mass are compared with the reference urologist.

The quantitative data, which were listed in Table 8 The hybrid method achieves the best formal match with the reference data. With regard to the fragmentable material (stone_2_F), the method provided values for eccentricity and coordinates of the center of mass that were similar to those provided by the expert. This ensures a high topological similarity. With regard to the non-fragmentable material (stone_2_NF), the hybrid method provided an accurate description of the perimeter and diameter. This method has thus successfully overcome the effect of high-density interference that usually affects image extraction.

# 4 DISCUSSION

The results obtained with regard to classification accuracy (99.39%) indicate that the use of the resnet50, with data purification, was the main reason for the success of the classification process. This ensured that the model learned discriminative features related to the kidney. The use of edge enhancement filters such as Laplacian, Prewitt, and Unsharp masking (USM) filters proved to be effective in enhancing the visibility of the contours of the kidney stones.

About segmentation, the hybrid method of using both region growing and MSER showed higher robustness in overcoming the issue of intensity overlap between calcified stones and bone structures. The technical advantage of this integrated approach is quantitatively substantiated by the percent relative difference (prd) and morphometric parameter analysis. While standalone methods like MSER and region growing provided acceptable results with PRD values generally below 5%, the hybrid method consistently achieved the highest precision, with PRD deviations reaching as low as 0.19%. The lowest value for the MSER method is (0.26), for the Region growing method (0.21). This accurate determination of morphological and geometrical characteristics is essential for increasing the precision of ESWL, thus reducing tissue damage and improving the outcome of the procedure.

# 5 CONCLUSIONS

This research has successfully developed a highly efficient, multi-stage system for automated diagnosis and determination of gravel area. Data purification led to improved performance of the convolutional neural network. This greatly increased the accuracy of the diagnosis. in classifying the types of stones. The MSER, Region growing, and Hybrid methods were found to be highly accurate in calculating the areas of the stones, which were extracted through the extraction phase. The results of the automated calculations were exceptionally and highly consistent with the manual assessment by the urologist, with differences remaining within a range of less than 5%. The proposed hybrid method showed a stable performance, effectively addressing the problem of density interference caused by bone structures. It can be said that the proposed system can be considered an effective tool that can be used to improve the accuracy of planning extracorporeal shock wave lithotripsy (ESWL) treatment.

# 6 ETHICAL CONSIDERATIONS

All CT images were acquired using a helical CT scanner at the Department of Urology in Al-Sadr Teaching Hospital and Al-Furat Al-Awsat Teaching Hospital, Najaf, Iraq, and also through the Kaggle data site. Ethical approval was obtained from the Najaf Health Directorate before the commencement of data collection. CT images were obtained and processed for purely scientific purposes, and all procedures performed in this study adhered to ethical and professional standards. Anonymity was ensured by removing any identifying information from the CT images of kidney stones that were specific to the patients. This was to preserve patient privacy and prevent the display of any name or address that might identify them.

## ACKNOWLEDGMENT

The authors would like to extend their thanks and appreciation to Dr. Muthanna Habeeb, a specialist in urological surgery, for His meticulous follow-up and valuable observations, which played a key role in enhancing the academic rigor and clinical validity of this work. Special thanks also go to the technicians of Al-Sadr Hospital and Al-Furat Al-Awsat Hospital in Najaf Governorate for their technical support and cooperation in collecting the Images. The authors gratefully acknowledge support from the Department of Physics, College of Science, University of Babylon, Iraq.

## REFERENCES

[1] M. N. Hossain, E. Bhuiyan, M. B. A. Miah, T. A. Sifat, Z. Muhammad, and M. F. Al Masud, "Detection and classification of kidney disease from CT images: An automated deep learning approach," Technologies, vol. 13, no. 11, p. 508, 2025.

[2] B. Oliveira, B. Teixeira, M. Magalhães, N. Vinagre, A. Fraga, and V. Cavadas, "Extracorporeal shock wave lithotripsy: Retrospective study on possible predictors of treatment success and revisiting the role of non-contrast-enhanced computer tomography in kidney and ureteral stone disease," Urolithiasis, 2024.

[3] L. Maier-Hein et al., "Metrics reloaded: Recommendations for image analysis validation," Nature Methods, vol. 21, no. 2, pp. 195–212, 2024.

[4] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool," BMC Medical Imaging, vol. 15, no. 1, p. 29, 2015.

[5] M. Islam and M. H. K. Mehedi, "CT kidney dataset: Normal-cyst-tumor and stone," Kaggle, vol. 2, 2021.

[6] Z. Chen et al., "Comprehensive 3D analysis of the renal system and stones: Segmenting and registering non-contrast and contrast computed tomography images," Information Systems Frontiers, 2025.

[7] S. N. Almuayqil, S. Abd El-Ghany, A. A. Abd El-Aziz, and M. Elmogy, "KidneyNet: A novel CNN-based technique for the automated diagnosis of chronic kidney diseases from CT scans," Electronics, vol. 13, no. 24, p. 4981, 2024.

[8] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[9] S. Asif, X. Zheng, and Y. Zhu, "An optimized fusion of deep learning models for kidney stone detection from CT images," Journal of King Saud University - Information Sciences, vol. 36, no. 7, p. 102130, 2024.

[10] S. S. G. Sahu and R. Kumar, "Automated morphometric analysis and localization of renal calculi using advanced image processing," International Journal of Biomedical Imaging, vol. 2024, 2024.

[11] D. S. N. Vasudeva and V. S. Dhaka, "Enhancing kidney stone diagnosis with AI-driven radiographic imaging: A review," Discover Artificial Intelligence, vol. 5, no. 1, p. 200, 2024, doi: 10.1007/s44163-024-00125-x.

[12] K. R. Chandra, T. H. Manoj, V. Harshitha, S. Divya, P. Swarnalatha, and S. Kareem, "Enhancing kidney stone diagnosis: A fusion approach of FCM and CNN for precise detection," in Proc. 3rd Int. Conf. Distributed Computing, Electrical Circuits and Electronics, 2024.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016.

[14] E. P. S. Win, "Nephrolithiasis detection and classification based on supervised machine learning," Journal of Data Science and Intelligent Systems, 2025.

[15] M. Zavvar and H. Ebrahimnezhad, "Accurate detection of kidney stones using attention-based graph models," Engineering Research Express, 2026.

[16] D. A. Mahmood and M. A. Muhammadali, "Hybrid U-Net architectures with ResNet50 and VGG19 for accurate CT-based kidney disease and stone segmentation," UHD Journal of Science and Technology, vol. 9, no. 2, pp. 231–250, 2025.

[17] D. C. Elton, E. B. Turkbey, P. J. Pickhardt, and R. M. Summers, "A deep learning system for automated kidney stone detection and volumetric segmentation on noncontrast CT scans," Medical Physics, vol. 49, no. 4, pp. 2545–2554, 2022.

[18] H. A. Owida, J. Al-Nabulsi, N. Al Hawamdeh, S. Abuowaida, Z. Salah, and E. A. Elsoud, "Deep learning-based kidney disease classification using ResNet50: Enhancing diagnostic accuracy," in Proc. 2nd Jordanian Int. Biomedical Engineering Conf., 2024.

[19] N. Vasudeva, V. S. Dhaka, and D. Sinwar, "Enhancing kidney stone diagnosis with AI-driven radiographic imaging: A review," Discover Artificial Intelligence, vol. 5, no. 1, p. 200, 2025.

[20] K. Yildirim, P. G. Bozdag, M. Talo, O. Yildirim, M. Karabatak, and U. R. Acharya, "Deep learning model for automated kidney stone detection using coronal CT images," Computers in Biology and Medicine, vol. 135, p. 104569, 2021.

[21] S. Sudharson and P. Kokil, "An ensemble of deep neural networks for kidney stone classification from CT images," Biomedical Signal Processing and Control, vol. 70, 2022, doi: 10.1016/j.bspc.2021.103018.

[22] Z. Zhang and U. Bagci, "Segmentation quality and volumetric accuracy in medical imaging," arXiv preprint arXiv:2404.1774, 2024.