Enhanced Kidney Stone Detection through Hybrid Crow-Cuckoo Search Optimized Convolutional Deep Belief Network Model

G. Ramesh Babu [1,*], N. Pushpalatha [1], Ganesh Khekare [2], Krishnamoorthy [3], Yousef A. Baker El-Ebiary [4], and S. Anjali Devi [5]

[1] Department of Computer Science and Engineering, Marri Laxman Reddy Institute of Technology and Management, Dundigal, India
[2] School of Computer Science and Engineering, Vellore Institute of Technology, India
[3] Department of Computer Science and Engineering, Panimalar Engineering College, Chennai, India
[4] Faculty of Informatics and Computing, UniSZA University, Malaysia
[5] Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India
Email: rameshbabu0552@gmail.com (G.RB.); pushpalatha523@mlritm.ac.in (N.P.); khekare.123@gmail.com (G.K.); krishnamoorthymuniyan@gmail.com (K.); yousefelebiary@unisza.edu.my (Y.A.B.E-E); swarnaanjalidevi@gmail.com (S.A.D.)
*Corresponding author

***Abstract*—Kidney stone detection is a critical task in medical imaging due to its significant impact on patient treatment. Early detection of kidney stones is crucial, as untreated cases can cause complications. However, machine learning methods for detecting kidney stones in medical imaging face challenges such as data availability, model complexity, interpretability, generalization, and ethical considerations. This study introduces an innovative hybrid approach that combines deep learning with optimization techniques to enhance kidney stone detection. The proposed method integrates watershed segmentation to improve feature extraction and facilitate more precise delineation of object boundaries, enhancing the quality of extracted features. This method uses a Convolutional Neural Network (CNN) to extract features and a Deep Belief Network (DBN) for classification. The CNN learns spatial hierarchies and edges, while the DBN has a multilayer structure for effective classification. Additionally, a hybrid optimization scheme that integrates Crow with Cuckoo Search Optimizer is introduced to further enhance the model's performance. This approach outperforms existing methods such as Resnet-50, DA-CNN, and Deep Learning (DL) with ResNet, achieving a remarkable accuracy of 99.35% in kidney stone detection. It significantly increases the accuracy of kidney stone detection due to the creative combination of data pre-processing, feature extraction, and hybrid Marker-based segmentation with a Crow-Cuckoo Search Optimizer (CCSO). Detecting kidney stones using existing methods such as Magnetic Resonance Imaging (MRI), Ultrasounds, and X-ray imaging faces challenges such as extended scanning times, invasiveness, radiation exposure, and inaccuracies. This study emphasizes using deep learning and optimization techniques to improve kidney stone detection and diagnostic accuracy in medical imaging, leading to better patient care.**

***Keywords*—kidney stone detection, watershed segmentation, hybrid crow-cuckoo search optimizer, convolutional neural network, deep belief network**

I. Introduction

Kidney stones have become increasingly common, with their prevalence rising from 3.8% in the latter part of the 1970s to 8.8% by the end of the 2000s. This percentage continued to increase, stabilizing at 10% prevalence from 2013 to 2014. Kidney stones could happen to about 11% of men and 9% of women. Conditions including diabetes, hypertension, and obesity can make this risk worse. A kidney stone is a hard item created by urine's chemical composition. Kidney stones can be made of calcium oxide, acid in the urine, a substance called or cystine. Major lower back pain, blood in the urine, feeling sick, throwing up, a fever, chills, and murky or smelly urine are typical signs. Additionally, risk factors for kidney stones encompass diet, chronic diseases (such as diabetes), and an unhealthy lifestyle (leading to a high Body Mass Index (BMI)). Consuming excess fruits, vegetables, vitamin C, vitamin B6, or animal proteins can increase the risk of forming kidney stones [1]. Frequent kidney stone symptoms include fever, discomfort, and feeling sick, which might be confused for indications of other illnesses. Kidney stones, which can vary in dimension and form and are made of hard, pebble-like material, develop in your kidneys as a result of a high level of some mineral compounds in urine [2]. These stones can lead to kidney failure and diminish quality of life by obstructing the urinary system [3]. They can result in renal failure, hamper daily functioning, and cause discomfort. At present, the fusion of computer vision and deep learning methods has found utility in diverse realms of medical imaging. Deep learning

Manuscript received January 12, 2024; revised March 1, 2024; accepted April 18, 2024; published March 14, 2025.

models have showcased effective utilization across various sectors, notably excelling in tasks such as segmentation within the domain of medical imaging. A large number of researchers have extensively utilized medical image processing to investigate renal concerns and conditions connected to kidney stone diseases [4].

Imaging techniques, with X-rays, ultrasounds, and Computed Tomography (CT), aid in detecting kidney stones. CT is particularly favoured for its sensitivity and specificity. It allows size and location assessment, aiding in the chance of stone passing without surgery categorization [5]. CT's ease of generating 3D images makes it a preferred choice in hospitals for kidney stone screening. Recent studies highlight Convolutional Neural Networks' (CNNs) ability to handle tasks like segmentation. CNNs can autonomously train on real data without human input, extracting features as they traverse the network layers. In medic/al imaging, CNNs show promise in case triaging and workflow optimization, crucial in emergency settings. Ultrasound (US) imaging is a practical method, providing anatomical data for diagnostics. Renal stones appear as echo-generating focal areas with acoustic shadowing in ultrasound images. The use of ultrasound technology assists doctors in locating and detecting stones. Detecting stones in abdominal X-ray images poses a complex challenge for medical professionals and radiologists due to a multitude of obstacles. Variability in the size and shape of the stones, coupled with their diverse locations along the urinary tract, adds to the complexity. Additionally, identifying stones in X-ray images can be troublesome at times. The development of an automated system for segmenting urinary stones holds significant value for the initial detection of stones and providing diagnostic aid to medical practitioners.

Numerous researchers have put forth various methodologies aimed at detecting or segmenting urinary stones, particularly in ultrasound or CT-scan images [6]. The effectiveness of deep learning often hinges on the quantity of training data available. Unlike some other domains, medical image datasets are constrained by factors such as the expensive nature of data collection, privacy constraints, and challenges associated with accurate image annotation that demands specialized expertise. Furthermore, an issue frequently encountered in medical fields is the imbalance between different classes, where the number of normal samples far exceeds the instances containing lesions [7]. Furthermore, conventional machine learning techniques might not encompass all the pertinent insights embedded within intricate patterns of neural signals. The efficacy of standard classifiers is largely contingent on the choice of features employed in model training. The process of extracting and pinpointing the most suitable features can pose difficulties when dealing by neural signals. The intricate and multidimensional nature of neuro imaging data renders it better suited for the application of deep learning methodologies [8]. Medical image segmentation is vital for diagnostics and interventions. Computer Tomography (CT) is commonly used, and recent studies have shown progress in isolating renal volumes in CT scans for analysis [9]. The proposed approach will utilize marker-controlled based watershed segmentation, a more refined version of watershed segmentation. Markers indicating the likely regions of kidney stones will be identified and used to guide the segmentation process, leading to more precise and controlled segmentation outcomes. The proposed research focuses on advancing kidney stone detection through the utilization of a novel Hybrid Crow-Cuckoo Search Optimizer (CCSO) and Convolutional Deep Belief Network model is used as a classifier. This work aims to increase the precision and effectiveness of kidney stone detection by integrating the strengths of both CCSO and convolutional Deep Belief Network (DBN) classifier. By synergistically combining these techniques, the study seeks to develop a more robust and effective approach for identifying kidney stones in medical images.

In the realm of medical imaging, the focus has recently shifted towards leveraging Magnetic Resonance Imaging (MRI), Ultrasounds, and X-ray modalities for the detection and diagnosis of kidney stones. While these techniques offer valuable insights into the presence and characteristics of kidney stones, they present distinct challenges that necessitate further research. MRI and Ultrasounds, although non-invasive, often require extended scanning times, potentially impeding timely diagnosis and treatment decisions [10]. Additionally, the invasiveness of MRI and Ultrasound procedures can introduce sample errors and discomfort to patients. On the other hand, X-ray imaging provides quicker results but exposes patients to ionizing radiation, raising concerns about cumulative exposure and health risks. Furthermore, both MRI and Computerized Tomography (CT) scans involve complex calculations that, if incorrect, could compromise the accuracy of diagnosis, potentially affecting patient safety. To address these limitations, researchers have explored the use of extensive datasets and predictive models to generate constant predictions. This approach promises non-invasiveness, cost-effectiveness, and ease of implementation. However, the challenge lies in refining and optimizing these image processing techniques while navigating the advantages of advanced deep learning methods [11]. Achieving a comprehensive solution that maximizes accuracy, minimizes invasiveness, and ensures patient safety is the focal point of current research in kidney stone detection using MRI, Ultrasounds, and X-ray imaging modalities.

The key contribution of this study are as follows:

- Frost and Anisotropic Diffusion Filter (FADF) is a method of filtering pictures that keeps edges and textures intact while removing noise. It is especially helpful in fields like medical imaging, where it is crucial to reduce noise without sacrificing crucial features.
- The kidney segmentation procedure makes use of Marker-Controlled Watershed Transform (MCWT). It is an approach to segmenting images that seeks to minimize over segmentation. It makes

use of the idea of flooding a geographic relief to divide a picture into several parts.

- CCSO: Crow-Cuckoo Search Optimization In order to solve the problem of high-dimensional feature spaces in classification problems, CCSO is utilized for feature selection. To effectively choose important characteristics, it combines the Crow Search Optimizer (CSO) with the Cuckoo Optimization Algorithm (COA).
- When working with a large number of characteristics, feature selection is essential for enhancing classification model performance. The most pertinent characteristics for the kidney stone detection job are chosen using CCSO.

The rest of this investigation is as follows: The existing research on detection issues utilizing different optimization techniques is examined in Section II. Problem statements were covered in Section III. The proposed approach is explained in Section IV. Section V talks about the performance review and system models that were built theoretically are included. In Section VI, the study conclusion is finished.

II. Related Works

Utilizing coronal Computed Tomography (CT) images, the researcher devised an automated kidney stone detection mechanism employing Deep Learning (DL) techniques. The automated detection method involved extracting multiple cross-sectional CT images from each individual. To prevent the model from memorizing images from the same patient during training, the raw photos were enhanced using techniques like zooming and mild rotation (10). The set of parameters of the XResNet-50 model were adjusted using the Adaptive Moment Estimation (ADAM) optimization technique in conjunction using the XResNet-50 neural network framework. This approach yielded results both in the output layer and in the Region of Interest (ROI), emphasizing accuracy. However, the model's limitation lies in its inability to classify kidney stones based on size, and it should be noted that the dataset was obtained solely from a single source. Patro *et al.* [12] utilized gradient-based anisotropy noise reduction, thresholding, and area expanding approaches after kidney segmentation using a 3D U-Net model. A 13-layer Convolutional Neural Network (CNN) classifier was then used to distinguish kidney stones from locations that had been mistakenly labelled as positives. The 3D U-Net segmentation primarily aimed to separate the two most prominent objects, effectively eliminating erroneous segmentations beyond the confines of the kidneys. In terms of stone measurements, manual acquisition of stone volume is a time-consuming process. Research suggests that volume measurements are more consistent than metrics such as stone diameter. However, the limitation of this study is its inability to automatically track changes in kidney stone volume over time.

In the context of kidney stone detection, the Support Vector Machine (SVM) algorithm is utilized within the domain of machine learning. This approach involved employing pre-processing techniques to extract dataset features, coupled with data augmentation to expand the dataset size for more effective model training. The versatility of the SVM algorithm enabled its application in both image classification and regression tasks, leading to the successful recognition of kidney stone types. However, a limitation of this methodology is its inability to effectively manage datasets with multiple layers or dimensions. By applying, both the segmentation method as well as artificial neural network developed by the multi-kernel k-means clustering algorithm, the clustering algorithm expand the feature set to higher dimensions and learn non-linear boundaries. The XGBoost classifier offers a training approach that is extremely scalable and prevents overfitting. The system is composed of four modules: (i) preprocessing, (ii) feature extraction, (iii) classification, and (iv) segmentation. Though segmenting and detecting are the findings of this research, the deep-learning network is complicated and has a long inference time [13].

Given the limited resolution of CT images, manual kidney stone detection can be challenging, potentially resulting in false alarms. Research tackled this concern by introducing a deep learning framework encompassing a Kronecker product-based convolution approach. This innovation not only discerns kidney stones but also streamlines feature extraction by circumventing convolution overlap. To enhance kidney stone detection, the study employs pre-processing techniques to extract the Region of Interest (ROI) from the image. A spectrum of image enhancement methods, including embossing, region identification, and cropping, has been developed. The embossing technique transforms 2D images into 3D representations. However, the model may encounter difficulties in accurately classifying images that are blurry or exhibit low contrast. Turk *et al.* [14] introduced models based on the V-Net architecture for dataset segmentation. The network design entails an encoder on the left side and a decoder on the right side. This architecture employs a convolutional structure for both feature extraction and resolution reduction, following a well-defined pathway. The encoder facet of the V-Net architecture comprises sections dedicated to calculating feature sets, while the decoder portion strives to achieve two-channel volumetric segmentation. The imaging data and corresponding ground-truth labels were anonymized and offered in the Neuroimaging Informatics Technology Initiative (NIFTI) format. However, the study did highlight an extended training period, noting that the model's complexity is relatively high.

In another study of kidney stone detection, the author conducted the imaging system and machine learning research on kidney illnesses, with a focus on automated segmentation for the kidneys in CT images. The suggested training plan has two levels of segmentation. The original image is first enhanced and trimmed to include both kidneys in one region. The cropped area is further subdivided into cubes for the first segmented network to process. The findings of the first stage in addition the original image serve as the input for the second stage, which enables the secondary segmentation network to carry out accurate kidney segmentation and kidney stone identification. This research demonstrates high precision in kidney as well as stones in the kidney classification in three dimensions. Still encounters difficulties while using to annotate and split the anatomy of the core kidney. In this research, Islam *et al.* [15] used the K-Nearest Neighbours (KNN) and Random Forest (RF) algorithms for detecting kidney stones. To choose the best settings, researchers used Random and Grid Search methods from the scikit-learning toolkit. To provide precise findings with no overfitting, which a Random Forest made up of 160 Decision Trees (DTs) was built for the outermost imaging sub-dataset, and 40 DTs were used for the sectional sub-dataset. The 40-component descriptor vectors for an alternate classifier (KNN) were divided by four 10-component parameter vectors. One of these held Local Binary Pattern (LBP) data, while the other three included data from the Herpes Simplex Virus (HSV) colour channels. These descriptor vectors were all used as the input for different KNN classifiers. This tactic produced outstanding accuracy. The study did not, however, go into great detail in examining colours and textures.

The literature review on kidney stone detection encompasses various approaches, primarily utilizing deep learning and machine learning algorithms to analyze medical images like CT scans. While some studies focus on automated detection mechanisms using deep learning techniques, others combine segmentation and classification algorithms for more accurate results. The SVM algorithm has been employed for image classification and regression tasks, with some limitations in managing multi-layer datasets. Clustering algorithms have also been used for feature expansion, along with classifiers like XGBoost, yet these methods can be complex and time-consuming. Innovations like a Kronecker product-based convolution and the V-Net architecture aim to streamline feature extraction and segmentation, but they may face challenges with blurry or low-contrast images. Overall, the existing literature highlights a demand for faster and more accurate detection methods, especially in handling large and complex datasets, which motivated the current research to propose a Hybrid Crow-Cuckoo Search Optimized Convolutional Deep Belief Network Model for enhanced kidney stone detection.

III. Materials and Methods

A. Data Collection

The Picture Archiving and Communication System (PACS) of multiple healthcare facilities in Dhaka, Bangladesh, provided the raw data for this research. This dataset encompassed patients who had previously received diagnoses of kidney tumours, cysts, normal conditions, or kidney stones. Considering the guidelines for entire abdominal and urogram tests, both Coronal and Axial slices were retrieved compared to contrast along with non-contrast images. The meticulous selection process involved isolating Dicom studies, one diagnosis at a time, and curating batches of Dicom images that captured the region of interest (ROI) corresponding to each radiological finding. To ensure privacy and data integrity, personal patient information and metadata were excluded from the Dicom images. Additionally, the Dicom images were transformed into lossless jpg image format. To maintain accuracy, each image finding was subsequently reviewed and validated by both a radiologist and a medical technologist. The resulting dataset is comprised of 12,446 distinct data instances. These instances are distributed as follows: 3,709 for cysts, 5,077 for normal cases, 1,377 for kidney stones, and 2,283 for tumours [15]. Fig. 1 shows the proposed mechanism.

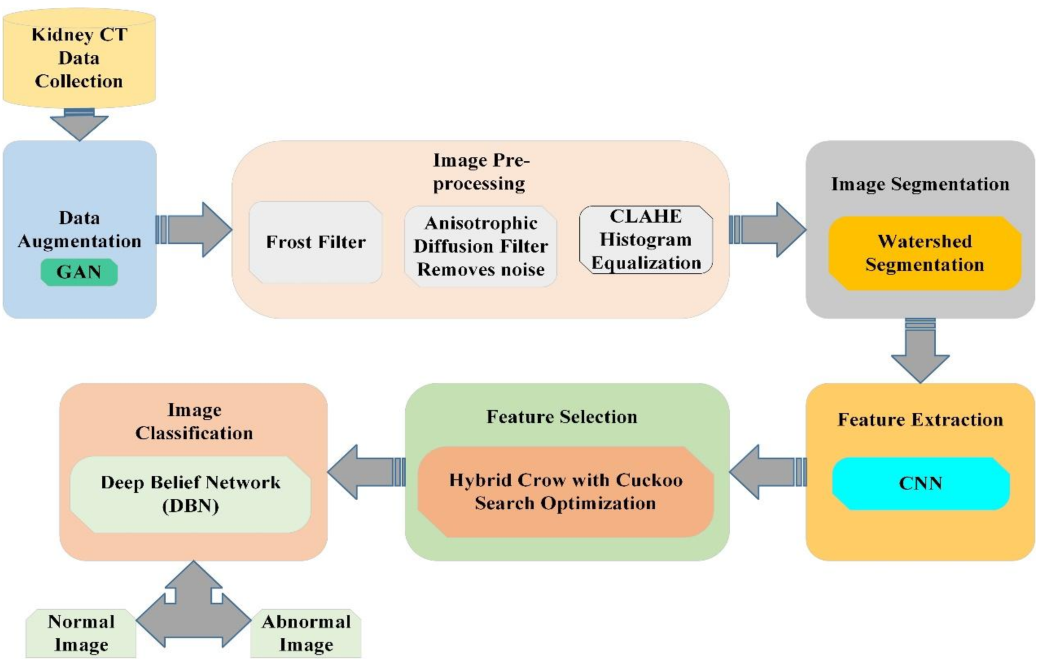

Fig. 1. Flow diagram of kidney stone detection using CNN.

B. Data Augmentation Using Generative Adversarial Network (GAN)

Data augmentation using GANs for kidney stone detection involves training a Generative Adversarial Network (GAN) to generate synthetic kidney stone images that resemble real kidney stone images. These synthetic images can be combined with the original dataset of kidney stone images to create a larger and more diverse dataset for training the kidney stone detection model. Below .is a step-by-step explanation of the process.

1) Prepare the original dataset

Collect a dataset of kidney stone images along with appropriate labels (positive for images containing kidney stones and negative for those without kidney stones). Preprocess the images as necessary for training.

2) Architecture of GAN

Design the structure of the Generator and Discriminator networks.

- Generator: The Generator creates fake kidney stone images using input of random noise. Numerous convolutional layers frequently followed through up-sampling layers in the design. (e.g., transposed convolutions) to generate higher-resolution images that resemble kidney stone patterns.
- Discriminator: The Discriminator is a binary classifier that aims to distinguish between real kidney stone images from the original dataset and synthetic kidney stone images produced by the Generator. The architecture usually involves of multiple convolutional layers followed by down-sampling layers (e.g., max-pooling) to lower the image dimensions.

3) GAN training

The GAN training process involves the following steps:

Step 1: Training the Discriminator

- Randomly select a batch of real kidney stone images from the original dataset.
- Generate a batch of synthetic kidney stone images using the Generator and random noise as input.
- Label the real kidney stone images as "real" (assigned a label of 1) and the synthetic kidney stone images as "fake" (assigned a label of 0).
- Train the Discriminator on the combined batch of real and synthetic kidney stone images to correctly classify them as real or fake.
- Update the Discriminator's weights based on the loss calculated from its predictions.

Step 2: Training the Generator

- Generate a new batch of synthetic kidney stone images using the Generator and random noise as input.
- Label these synthetic kidney stone images as "real" (flipped labels) since the Generator seeks to produce convincing visuals that will deceive the Discriminator. Train the Generator to produce synthetic kidney stone images that the Discriminator is more likely to classify as real.
- Update the Generator's weights based on the loss calculated from the Discriminator's predictions on the synthetic kidney stone images.

Fig. 2 illustrates the architecture for data augmentation using GANs. The generator tries to minimize the difference between the distribution of generated data and the real data. The generator loss is typically defined using the binary cross-entropy loss. Let $D(x)$ be the discriminator's output when given the real data $'x'$, and $D\big(G(z)\big)$ be the discriminator's output when given the generated data $G(z)$ (where z is the random noise input to the generator). The generator loss L_G is defined as in Eq. (1).

$$LG{=}{-}logDGz \tag{1}$$

The discriminator aims to correctly classify real data as real and generated data as fake. It seeks to maximize the difference between the output of the discriminator for created data and real data. The binary cross-entropy loss is also commonly used to define the discriminator loss. The discriminator loss L_D is defined as in Eq. (2).

$$LD{=}{-}logDx{-}log1{-}DGz \tag{2}$$

The total loss for the GAN is a combination of the generator and discriminator losses. Typically, the generator and discriminator are updated alternately during training. The total loss L_{Total} is defined as Eq. (3).

$$LTotal{=}LG{+}LD \tag{3}$$

During training, the goal is to find the optimal parameters for the generator and discriminator that minimize L_{Total} or L_T. This adversarial training process iterates multiple times to improve the generator's ability to produce realistic data and the discriminator's ability to distinguish between real and fake data.

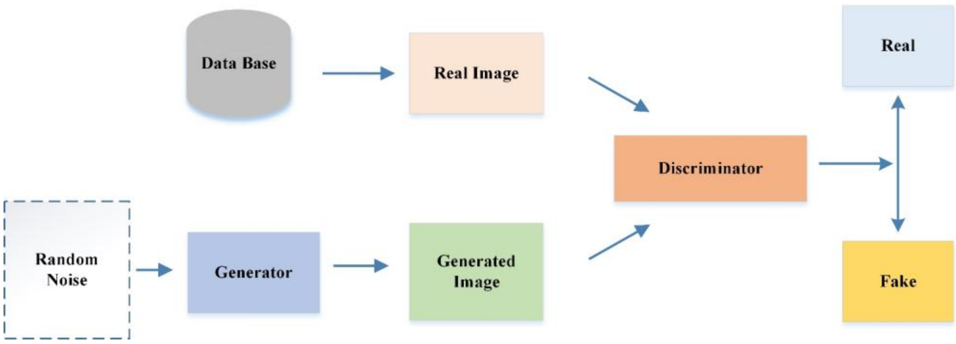

Fig. 2. Architecture of data augmentation using GAN.

C. Pre-Processing Using Frost Anisotrophic Diffused Filter (FADF)

The Frost and Anisotropic Diffusion Filter (FADF) are image filtering techniques used in pre-processing to enhance or denoise images before feeding them into machine learning models. They are commonly used for improving the quality and clarity of images, which can be especially beneficial for tasks like object detection or classification.

1) Frost Anisotrophic Diffused Filter (FADF)

The Frost Anisotropic Diffusion Filter, a variation of the Frost filter, represents a non-linear diffusion technique designed to eliminate noise from images while retaining their edges and textures [16]. This filter excels at mitigating speckle noise and other common forms of image noise, making it particularly valuable in domains such as medical imaging and remote sensing. By diffusing pixel values in response to the local gradient and neighbouring pixel data, the filter effectively smooths areas with uniform intensity, while simultaneously preserving intricate details and sharp boundaries. The mechanism of the Diffused Frost Filter involves iteratively solving a Partial Differential Equation (PDE) across the image. This equation orchestrates the diffusion of pixel values over time, leveraging both local gradient information and adjacent pixel values. The outcome is a smoothing effect that maintains the integrity of edges and fine details. The mathematical representation of the Diffused Frost Filter can be articulated as follows. Given an input image is $L_T(x, y)$ at coordinates (x, y), the Diffused Frost Filter can be defined using the following PDE as in Eq. (4).

$$\partial LT \partial t = div\ Dx,\ y,\ t \nabla LT \tag{4}$$

where, $\frac{\partial L_T}{\partial t}$ represents the time derivative of the image intensity, i.e., the rate of change of the pixel values over time. $div()$is the divergence operator, which measures the rate of change of the vector field ∇I at each point. ∇L_T is the gradient of the image intensity, which represents the direction and magnitude of the change in intensity at each pixel. $D(x, y, t)$ is the diffusion coefficient, which controls the rate of diffusion at each pixel location and time. The diffusion coefficient $D(x, y, t)$ is calculated based on the local gradient information at each pixel in Eq. (5).

$$D(x, y, t) = exp(-|\nabla L_T|^2/k^2) \tag{5}$$

where, k is a diffusion parameter that controls the strength of diffusion. Higher values of k result in stronger diffusion and smoother images. To solve the PDE, numerical methods such as finite differences or partial differential equation solvers are used. The Diffused Frost Filter is applied iteratively until convergence, with the process of diffusion smoothing out the noise in the image while preserving important features and edges. It's important to note that the implementation and parameter tuning of the Diffused Frost Filter may vary depending on the specific application and the characteristics of the input images. The choice of κ and the number of iterations can affect the denoising performance and computational efficiency. Anisotropic diffusion stands out as a widely-used image processing approach, offering noise reduction and edge-conserving smoothing. Introduced by Perona and Malik in the 1990s, anisotropic diffusion emerged as an alternative to conventional isotropic diffusion methods that tend to blur edges and fine features during noise reduction. This mathematical technique selectively smooths images, safeguarding crucial features such as edges and boundaries. Achieving this balance involves adapting the diffusion process according to local image gradients, allowing diffusion along smooth regions while impeding diffusion across edges. In essence, the Frost Anisotropic Diffusion Filter encapsulates these principles, offering a robust approach to noise reduction that respects image structures. Mathematically, anisotropic diffusion is typically described using Partial Differential Equations (PDEs). The basic form of the anisotropic diffusion PDE is as follows in Eq. (6) [17].

$$\begin{cases} \frac{\partial L_T(x,y,t)}{\partial t} = div\big(G(\|\nabla L_T(x, y, t)\|) \cdot \nabla L_T(x, y, t)\big) \\ L_T(x, y, 0) = L_{T_0} \end{cases} \tag{6}$$

where, t represents time; $L_T(x, y, t)$ is an image at time t; div denotes the divergence operator; $\nabla L_T(x, y, t)$ is the gradient image; $g(\cdot)$ I the diffusion function and L_{T_0} is the real image $(t = 0)$. In anisotropic diffusion, the amount of diffusion is higher in regions where a small gradient (low intensity variation) is detected. Conversely, the diffusion is reduced in regions where a higher gradient magnitude (high intensity variation) is detected. This behaviour allows anisotropic diffusion to smooth the image in areas with low contrast or smooth regions while preserving edges and boundaries in areas with high contrast or sharp intensity transitions. These requirements are satisfied by the nonnegative decreasing conductance function that is chosen as the diffusivity function as follows in Eq. (7).

$$\begin{cases} \lim_{x \to 0} g(x) = 1 \\ \lim_{x \to \infty} g(x) = 0 \end{cases} \tag{7}$$

The Alliance Defending Freedom (ADF) suggests the use of two distinct diffusion coefficients in this paper described by Eqs. (8) and (9).

$$g_1(\|\nabla L_T\|) = \frac{1}{1+\left(\frac{\|\nabla L_T\|}{k}\right)^2} \tag{8}$$

$$g_2(\|\nabla L_T\|) = exp\left[-\left(\frac{\|\nabla L_T\|}{k}\right)^2\right] \tag{9}$$

The diffusion process is modified using the threshold parameter, k. The gradient $\|\nabla L_T\|$ acts as the edge detector. Consequently, when the gradient magnitude is higher than the threshold value $k(g = 0)$, the diffusion process is halted across edges. On the other hand, in uniform regions where the gradient magnitude is smaller than k, diffusion is encouraged to be maximal $(g = 1)$.

The anisotropic diffusion equation is derived as follows in Eqs. (10) and (11).

$$L_{T_{t+1}}(s) = L_{T_{t+1}}(s) + \frac{\lambda}{|\eta_s|}\sum_{p\in\eta_s} g_k\left(\left|\nabla L_{T_{s,p}}\right|\right)\nabla L_{T_{s,p}} \quad (10)$$

where L_T represents a separately extracted picture, s indicates a discrete 2-D grid's pixel position, t represents the repetition stage, g is the conductance function and k is the gradient threshold parameter.

$$\nabla L_{T_{s,p}} = L_{T_t}(p) - L_{T_t}(s),$$

$$p \in \eta_s = \{N, S, E, W, NE, SE, SW, NW\} \quad (11)$$

Edges are regions in an image where there are significant local changes in intensity. Frequently, the local maxima or minima of the first derivative or the zero-crossings of the second derivative, respectively, are used to identify sites that are on an edge. While the second derivative displays a zero crossing, the first derivative shows a point of maximum at the edge. In accordance with a finite difference, first derivative is calculated in the following ways as Eqs. (12) and (13).

$$f^{'}(x,y) = \lim_{l\to 0}\frac{f(x+l,y)-f(x,y)}{h} \quad (12)$$

$$f^{''}(x,y) = \lim_{l\to 0}\frac{f^{'}(x+l,y)-f^{'}(x,y)}{h} \quad (13)$$

2) Histogram Equalisation—Contrast Limited Adaptive Histogram Equalization (CLAHE)

Histogram Equalization (HE) is a fundamental approach for improving the contrasting qualities of an image. Histogram equalization aims at equally distributing an image's intensity levels across the whole range of possibilities, producing a more harmonious and aesthetically beautiful image with greater visibility. It is easily capable of overexposing the picture. By using adaptive histogram equalization in specific locations, the Contrast Limited Adaptive Histogram Equalization (CLAHE) approach is employed to improve the contrast of an image. The final result is divided into fewer pixels or regions, and every region's histogram is separately equalized. For Knoxville Utilities Board (KUB) imaging, CLAHE has successfully increased contrast. This method makes it possible to distinguish stones clearly with the backdrop**.** In CLAHE, each pixel's intensity value is adjusted based on the histogram of its local neighborhood, ensuring that the contrast enhancement is limited and does not lead to unnatural artifacts. This method is particularly useful in improving the visibility of details in both dark and bright areas of the image simultaneously. The results after adjusting only the luminance channel of an HSV image are far better than those after adjusting each channels for a Bureau of Governmental Research (BGR) image when CLAHE is being used to color images, since it frequently relies on the lightness component [18].

The Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm can be summarized in the following steps.

- Divide the input image into smaller non-overlapping tiles or regions.
- Compute the histogram of each tile, representing the distribution of pixel intensities within that region.
- Apply standard histogram equalization to each individual tile, which stretches the intensity range to make the histogram more uniform.
- To avoid over-amplification of noise, clip the histogram by limiting the maximum pixel frequency within each tile. If a pixel frequency exceeds the defined clip limit, redistribute the excess pixels uniformly among the intensity levels.
- Reassemble the processed tiles to form the final enhanced image.

The algorithm's key feature is its adaptiveness to the local characteristics of the image, as it equalizes each tile independently, taking into account the varying contrast levels in different regions. This makes CLAHE effective in improving the visibility of details while avoiding undesirable artifacts caused by global histogram equalization methods.

D. Image Segmentation Using Marker Based Watershed

Image segmentation using watershed segmentation is a common method for separating an image into various areas or parts based on certain image properties, such as intensity, texture, or gradient information. The watershed algorithm is primarily used for over segmentation, meaning it tends to divide regions even when they should be part of the same object. To obtain meaningful segments, watershed segmentation often requires the use of markers to guide the process. Two structures are used in the processing stages that follow. The kidney's intensity, which is normally around 30 HU in most abdominal CT scans, is the initial characteristic. However, due to the contrast-enhanced nature of the analysed CT series, a broader value range (0–200 HU) is considered. Throughout the further processing steps, the image is denoted as $I(x, y, z)$ by Wieclawek [19]. The second feature focuses on a geometric constraint related to the anatomical location of the kidneys. The limit enclosing the region of interest is established for further research by analyzing the body and skeleton views. This enclosing box's core line is recognized and highlighted across the spine position. The area within it is divided into two smaller compartments, each of which holds either the left or right kidney. The kidney' primary placements are determined by the point at which they intersect between their diagonals. Additionally, all of the non-zero voxels from the image were used for further evaluation $I(x, y, z)$ that are indicated by the mask are taken into consideration.

In the initial mask, two distinct areas are defined. One area corresponds to the kidney, while the foreground is represented by the other one. These regions are referred to as the "object marker" and the "background marker", correspondingly, in the next stage known as "Preliminary kidney segmentation". The binary mask provided indicates a significant number of voxels comprising the kidneys. To derive the h value in the HMAX transform, the sum of the mean value and standard deviation of nonzero voxels along the diagonal is employed was expressed in Eq. (14).

$$HMAX_h\big(L_T(x,y,z)\big) = R^{\delta}_{L_T}\big(L_T(x,y,z) - l\big) \quad (14)$$

The process reduces the visibility of voxels in the kidney in $L_T(x,y,z)$ by reducing the intensity maxima below the threshold level h. The function $R^{\delta}_{L_T}(\bullet)$ performs morphological reconstruction through dilation. In other words, it reconstructs the image $L_T(x,y,z)$ by iteratively dilating the image while preserving certain structural elements, as in Eq. (15).

$$R^{\delta}_{m}\big(L_T(x,y,z)\big) = \delta^{i}_{m}\big(L_T(x,y,z)\big) \quad (15)$$

The resulting image contains voxels that are identified by the mask, forming a subset of pixels denoted as D. This subset plays a crucial role in determining the threshold value, which is defined as Eq. (16).

$$th_k = \bar{D} + \sigma_D \quad (16)$$

where, mean value of D is $\bar{D}$. In the 3D implementation, the binary image, which meets the condition $L_T(x,y,z) > th_k$, undergoes an opening operation. This is tailed by morphological reconstruction to eliminate objects that are touching the borders of the rectangular shape. All these operations are performed in the three-dimensional space of the image to ensure comprehensive processing and accurate results. The procedure of removing objects is repeated until there is more than one remaining binary object within each bounding box. When the volume of the objects is similar, both of them are retained. However, in cases where the volumes differ significantly, the smaller object is removed. This process is commonly referred to as nephrectomy. The volume of the binary items, which serve as kidney seeds, is often a little less than that of the kidneys themselves. While parts of the processing, a basis go through a stage called marker creation.

1) Marker generation for kidney segmentation

In the marker-controlled watershed transform, researcher anticipate the presence of two markers. The initial marker, also known as an "object marker", is associated with the image region that was obtained in the previous stage. Each slice containing the kidney is given a rectangular solid shape in order to create the backdrop marker. A medium-sized structural element is also used to dilate the renal area morphologically. This element's size is 10% of the smallest measured size in the (x,y,z) directions and is dependent on the size of the kidney. The final step is to flip the photograph to produce a mask that emphasizes the kidney background.

2) Marker generation for kidney segmentation

The Marker-Controlled Watershed Transform (MCWT) is used in the kidney segmentation process's final stage. A dependable and flexible method for picture segmentation is the marker-controlled watershed algorithm given in this research [20]. The MCWT is an enhanced watershed transform that aims to lessen over-segmentation. It is simple to illustrate the idea behind this technique using a 1D image quality or even a 2D-gray level image. The image is treated by the watershed transform as a topographic relief, similar to a landscape that has been wet. All local minima in the image are where flooding begins, and as the water level rises, all basins are filled. These structures are built to stop water from various basins from mixing when they would otherwise meet. If the water level hits the highest point in the landscape, the procedure comes to an end. The ensuing arrangement of structures symbolizes the segmentation of the image into various sections, which is a fundamental idea in image processing. Its gradient magnitude is further examined due to watershed segments cross the brightest pixels. The background marker effectively masks out irrelevant areas, while the object marker identifies regions that should remain connected and not be split during the segmentation process. As a result, a limited area is created where the edges of the kidneys are located. To prepare the gradient magnitude image, The watershed structures and the borders of the kidneys cross. Two 3D techniques—average spacing filtering and morphological opening—are used to lessen the effect of noise on the gradient computation. The following equation is used to determine the gradient magnitude in 3D space as expressed in Eq. (17).

$$\|\nabla L_T\| = \sqrt{\left(\frac{\partial}{\partial x}L_T\right)^2 + \left(\frac{\partial}{\partial y}L_T\right)^2 + \left(\frac{\partial}{\partial z}L_T\right)^2} \quad (17)$$

Even with a relatively modest gradient window size, the gradient magnitude reveals additional features in addition to the kidney edge, particularly when there are nearby objects with strong edges. Markers are used to guarantee that just the desired edges are produced. Effectively revealing a corridor with the kidney edge is the marker-defined area. As a result, the watershed dams precisely line up with the limits of the kidneys, producing a highly precise separation of the kidneys.

3) Algorithm for Marker-Controlled Watershed Transform (MCWT) kidney segmentation

a) Step 1: Pre-processing

- Convert the input 3D abdominal CT scan image $L_T(x,y,z)$ into a grayscale image.
- Apply any necessary preprocessing steps, such as noise reduction and intensity normalization, to enhance the image quality.

b) Step 2: Marker generation

- Generate the object marker by extracting the kidney region obtained from a preliminary segmentation step. This serves as the seed for the kidney.
- Generate the background marker by performing a rectangular convex hull for each slice containing the kidney. Additionally, apply a moderately sized structural element and structural dilatation (size based on kidney size) to increase the kidney region

c) Step 3: Gradient magnitude image

- Compute the gradient magnitude of the image, which highlights edges and gradients of intensity variations.

d) Step 4: Marker-controlled watershed transform

- Combine the object marker and background marker to create the marker image.
- Apply the watershed transform to the gradient magnitude image using the marker image as a constraint.
- This process segments the image into regions, and the watershed dams align with the kidney boundaries

e) Step 5: Refinement

- Perform further processing on the segmented image, such as noise reduction, morphological operations, and region merging, to refine the kidney segmentation.

f) Step 6: Post-processing

- Postprocess the final segmentation results to remove any remaining artifacts or unwanted regions.
- Obtain the segmented kidney regions as the final output of the algorithm.

The Marker-Controlled Watershed Transform (MCWT) kidney segmentation algorithm presented in this study demonstrates a robust and flexible approach for accurate kidney segmentation in 3D abdominal CT scans. By employing object and background markers, the algorithm effectively localizes the kidney region while minimizing over-segmentation and undesired splitting. The utilization of gradient magnitude and marker-based constraints enables the algorithm to precisely identify the kidney boundaries, even in the occurrence of strong edge objects in the image. The incorporation of 3D operations further enhances the accuracy of the segmentation process. The experimental results on a diverse set of abdominal CT scans showcase the algorithm's capability to provide reliable and consistent kidney segmentation across different cases. Moreover, the algorithm's ability to handle varied kidney sizes and complex anatomical structures contributes to its practicality in clinical applications. The MCWT kidney segmentation algorithm offers a promising solution for automating and improving the efficiency of kidney segmentation in medical imaging, paving the way for advanced analysis, diagnosis, and treatment planning in renal-related medical scenarios.

E. Feature Extraction Using CNN

Convolutional Layers, Rectified Linear Unit (ReLU) Layers, Pooling Layers, and Fully-Connected Layers are the four types of layers that commonly make up a Convolutional Neural Network (CNN) structure. User provides a set of secure interactive protocols connecting the two edge servers, S_1 and S_2 in order to achieve privacy-preserving feature extraction within CNNs.

1) Convolutional layer

The convolution operation fundamentally involves computing dot products between filters and localized sections of the input. This operation follows the principles of associativity and distributivity in relation to addition. Each of the pixels in a convolution's receptive area are combined into just one value.

2) ReLU layer

The activation layer serves the purpose of infusing non-linearity into the network. In contemporary neural networks, it is generally recommended to employ Rectified Linear Units (ReLU) as the activation function. Consequently, focus initially lies on describing the confidentiality-preserving ReLU layer.

3) Pooling layer

Down sampling is carried out by the pooling layer throughout the spatial dimensions, which include width and height. As mentioned above, popular pooling techniques include things like average pooling and maximum pooling. It is simple to prove that determining the average maintains qualities of correlation and distributiveness through additions in the situation of average pooling, which returns the mean of values within a rectangular neighbourhood. Both of the edge servers can still separately carry out the pooling activities if we include normal pooling into network. A little more complicated procedure is introduced by the Max pooling. It locates and displays the amount with the highest output inside a specified rectangle zone.

4) Fully connected layer

Every neuron within this layer establishes connections with each neuron in the preceding layer. To link the neurons between two layers, the Fully Connected (FC) layer, which also includes weights and biases, is utilized. The Fully Connected (FC) layer, which links the neurons to the other layers, is made up of both biases and weights as well as the neurons.

F. Feature Selection Using Hybrid Crow-Cuckoo Search Optimizer

Following feature extraction, hybrid Crow—Cuckoo Search Optimisation (CCSO) is used to choose the key features. Due to the performance of the classification process being affected by a high number of characteristics and the complexity of the calculation. The CSO (Crow Search/ Optimizer) makes an effort to mimic the social intelligence and food-gathering techniques of crow flocks [21]. The lifestyle of a group of birds known as cuckoos served as an inspiration for COA (Cuckoo Optimization Algorithm). The invention of this optimization method was inspired by the lifestyle, egg-laying characteristics, and breeding of these birds [22]. The hybrid Crow-Cuckoo Search Optimizer (CCSO) describes the suggested feature selection procedure step by step. The movements of each crow are caused by two basic characteristics: (1) locating the other flock members' hiding places; and (2) defending one's own hiding places. In a typical CSA, a flock of crows would spread out and explore the whole decision space in search of the ideal hiding places (global optima). In order to ensure the versatility of efficient optimization algorithms across diverse space characterized by d dimensions. To start a scenario where N crow individuals (indicative of the flock size) are randomly positioned within a space having d dimensions [23]. At the t^{th} iteration, the location of the i^{th} crow individual within the search space is denoted as $x(i, t)$, constituting a viable array of decision variables. Furthermore, each individual crow has the capability to remember the location of the most favourable hiding spot discovered. During the t^{th} iteration, $m(i, t)$ denotes the hideout position for the i^{th} crow individual, representing the best spot identified by that individual up to that point.

Following this, each crow individual will engage in motion guided by the fundamental principles of the CSA: (1) securing its individual hiding place, and (2) detecting the concealed locations (hidden spots) of other group members. Let's consider that during the t^{th} iteration, the j^{th} crow individual endeavours to retrieve food from its concealed location $[m(j, t)]$. At the same time, the i^{th} crow chooses to tail the j^{th} individual with the intention of pilfering the stored food. Under such conditions, two potential outcomes may emerge: (a) the j^{th} crow individual remains unaware of being trailed, resulting in the exposure of its hideout location to the i^{th} crow individual; or (b) the j^{th} crow individual detects the existence of a potential thief, prompting a strategic makeover by the j^{th} crow aimed at deception. Under the first scenario, the inattentiveness of the j^{th} crow individual would permit the i^{th} crow to identify and loot the concealment of the j^{th} crow. In this particular situation, the realignment of the i^{th} crow's position can be achieved through the following approach as described in Eq. (18),

$$x_{(i,t+1)} = x_{(i,t)} + r_i \times fl_{(i,t)} \times \left[n_{(j,t)} - x_{(i,t)}\right] \quad (18)$$

Here, r_i represents a randomly chosen number following a uniform distribution within the interval of $[0, 1]$; while $fl(i, t)$ signifies the Flight distance of an individual crow during the t^{th} iteration. Among the algorithm's parameters, $fl(i, t)$ holds the ability to influence the algorithm's search effectiveness. Let's consider that when fl takes on lower values, it guides the algorithm to perform a targeted local search in the proximity of $x(i, t)$. In contrast, higher fl values cause the algorithm's search to encompass a broader range. Speaking in terms of optimization, diminishing fl values contribute to heightened result concentration, whereas increasing fl values encourage result variation. Effective optimization algorithms demonstrate qualities of both robust intensification and diversification. There might also be situations where the j^{th} crow individual perceives being trailed by a fellow flock member (such as the i^{th} crow). Consequently, to safeguard its food supply from potential theft, the j^{th} crow would slyly glide over an area without hiding spots. To replicate this behaviour in the context of the CSA, a random location in the i^{th} crow would be allotted d-dimensional decision space. In both of the aforementioned circumstances, it is possible to create a formula for detecting individual crow activity are as follows in Eq. (19).

$$x_{(i,t+1)} = \begin{cases} x_{(i,t)} + r_i \times fl_{(i,t)} \times \left[n_{(j,t)} - x_{(j,t)}\right] & r_j \geq AP_{(j,t)} \\ a \quad random \quad position & otherwise \end{cases} \quad (19)$$

where, r_j represents a selected at random from a similar distribution within the range of $[0,1]$, and $AP_{(j,t)}$ signifies the AP of the j^{th} crow during the t^{th} iteration. The flight length (fl) and the Awareness Probability (AP) are the two fundamental elements that govern intensification and diversification in the CSA. It is likely that the crow flock members will discover the hiding places would rise by lowering the awareness probability. Therefore, the CSA often concentrates its search efforts in the area around potential hiding places. On the other side, by raising the AP, the crows have a greater tendency to search the decision space randomly, which would actually reduce the likelihood that the plunderers would find the actual hiding places. The j^{th} crow would slyly fly over a location where one can't hide, protecting its food supply provided by the pillager. The i^{th} crow would be considered to be somewhere random in the d-dimensional decision space in order to mimic such an action in the CSA. In conclusion, the two mentioned situations of the individuals of the crow species tailing can be defined as: The flight length and awareness probability are the two key variables that may be changed in the population-based optimization technique known as Cloud Security Alliance (CSA). Because of these qualities, Cloud Security Alliance (CSA) is a promising solution for challenging engineering optimization issues. The Cuckoo Optimization Algorithm (COA) is used to carry out continuous nonlinear optimization. COA was inspired by the way of life of the cuckoo family of birds. The development of this optimization method was based on the breeding, lifestyle, the egg-laying habits of these birds. The algorithm's foundation is the struggle for survival. Some of them die while battling it out to survive. The cuckoos who make it through relocate to better regions, where they begin breeding and laying eggs.

Finally, the cuckoos that have survived have merged into a community with the same profit rate. These cuckoos have some eggs that they will lay in the nests of other species. The likelihood of some of these eggs being reared and developed into cuckoos is higher if they resemble the host's eggs. The host discovers more eggs, which are demised. The rate of raised eggs indicates how suitable the location is. A place will earn more if there are more eggs to be produced there. Therefore, one criterion for the cuckoos to be optimized will be the circumstance in which more eggs are survived. Cuckoos look for the optimum location to extend the lifespan of their eggs. Cuckoos develop into adult birds that establish colonies and communities after hatching. Every town has a unique environment in which to live. An algorithm called COA maximizes the profit function. The cost function has to be multiplied by a minus sign in order to apply COA to address the issue. A habitat matrix with a size of $N_{pop} \times N_{var}$ is produced to begin optimization. Then, a certain number of randomly chosen eggs is determined for each habitat matrix. In the wild, each cuckoo lays 5–20 eggs. In various iterations, each cuckoo's egg specification uses these values as the upper and lower bounds. Every genuine cuckoo lay eggs within a certain range. In an optimization challenge with both upper and lower limits, each cuckoo has an ELR that is inversely proportional to the overall number of eggs, the number of eggs currently deposited, and the upper and lower limits that comprise the difficulties components for var_{hi} and var_{low} . Thus, ELR is described as Eq. (20):

$$ELR = \alpha \times \frac{No.of\ current\ cukoos\ eggs}{Total\ number\ of\ eggs} \times (var_{hi} - var_{low}) \quad (20)$$

where, α is a variable, by which the maximum ELR is set. As the Crow-Cuckoo Search Optimizer (CCSO) is a hybrid of the Crow Search Algorithm (CSA) and Cuckoo Optimization Algorithm (COA), it utilizes their fitness functions. For the kidney stone detection problem, the fitness function typically revolves around the evaluation of a candidate solution's quality, which can vary depending on the specific problem at hand. This evaluation can be based on various metrics such as accuracy, precision, recall, or the F1–Score, depending on the study's objectives and constraints. For example, accuracy measures the percentage of correctly identified instances, while precision measures the ratio of true positives to the sum of true and false positives. Similarly, recall measures the ratio of true positives to the sum of true positives and false negatives, and the F1–Score is the harmonic mean of precision and recall, providing a balance between the two. The specific fitness function used in the optimization process would depend on the desired balance between these metrics and the problem's requirements.

G. Deep Belief Network for Image Classification

A DBN [24], a kind of profound neural network, is a pictorial structure that includes multiple layers of "hidden units", where connections are made between layer rather than inside individual layers. While trained supervision DBN is used for classification, trained Un-supervision DBN computationally combines its data inputs to detects features. Unsupervised DBN describing an unstructured, constructive energy-based system with a hidden layer that serves as a visible component over the next one is called Restricted Boltzmann Machine (RBM)/autoencoders. Layer-wise biased training of DBN results in an effective neural network method. Multiple layers are present in DBN [25] as shown in Fig. 3, including an input layer, visible neurons, an output layer, and buried neurons.

The GAN-CCSO algorithm is designed for the discovery of stone in kidney images. It begins by taking an input image containing a kidney stone and applies image augmentation through a Generative Adversarial Network (GAN) for data enhancement. Subsequently, the image undergoes preprocessing, involving contrast enhancement using the CLAHE method and noise removal through a Frost Anisotropic Diffused Filter (FADF). Marker based image segmentation is performed by computing a threshold value. To retrieve relevant features, a Convolutional Neural Network (CNN) is employed to capture color, border, diameter, and shape information. CCSO is utilized to select pertinent features, followed by the computation of fitness for vultures. If the number of selected features (|s|) is greater than or equal to 1, the algorithm upgrades crow-cuckoo locations using Eq. (33); otherwise, it utilizes Eq. (34). Finally, the algorithm employs a Deep Belief Network (DBN) for classification, culminating in stone detection. This comprehensive approach integrates GAN-based augmentation, preprocessing, segmentation, feature extraction, CCSO feature selection, and DBN classification to effectively identify stone in kidney images. Fig. 4 shows flowchart of this research.

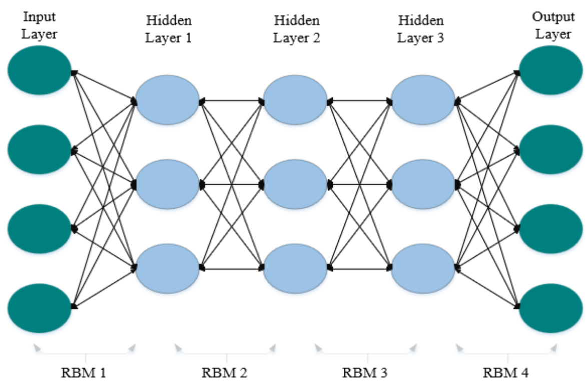

Fig. 3. DBN for classification.

Algorithm 1: GAN-CCSO

Input: Image containing kidney stone
Output: Stone detection in kidney image
Load the input image
Image augmentation using GAN
Perform preprocessing operation //contrast enhancement using CLAHE Noise removal using FADF is given in Eq. (13)
Image segmentation //compute threshold value using Eq. (24)
Feature extraction //color, border, diameter, shapes were extracted using CNN
Select feature using CCSO
Calculate the fitness of crow- cuckoo
If(|s|≥1) then
Upgrade the location crow with cuckoo using Eq. (19)
Else
Upgrade the location crow with cuckoo using Eq. (20)
Endif
Classification using DBN
End

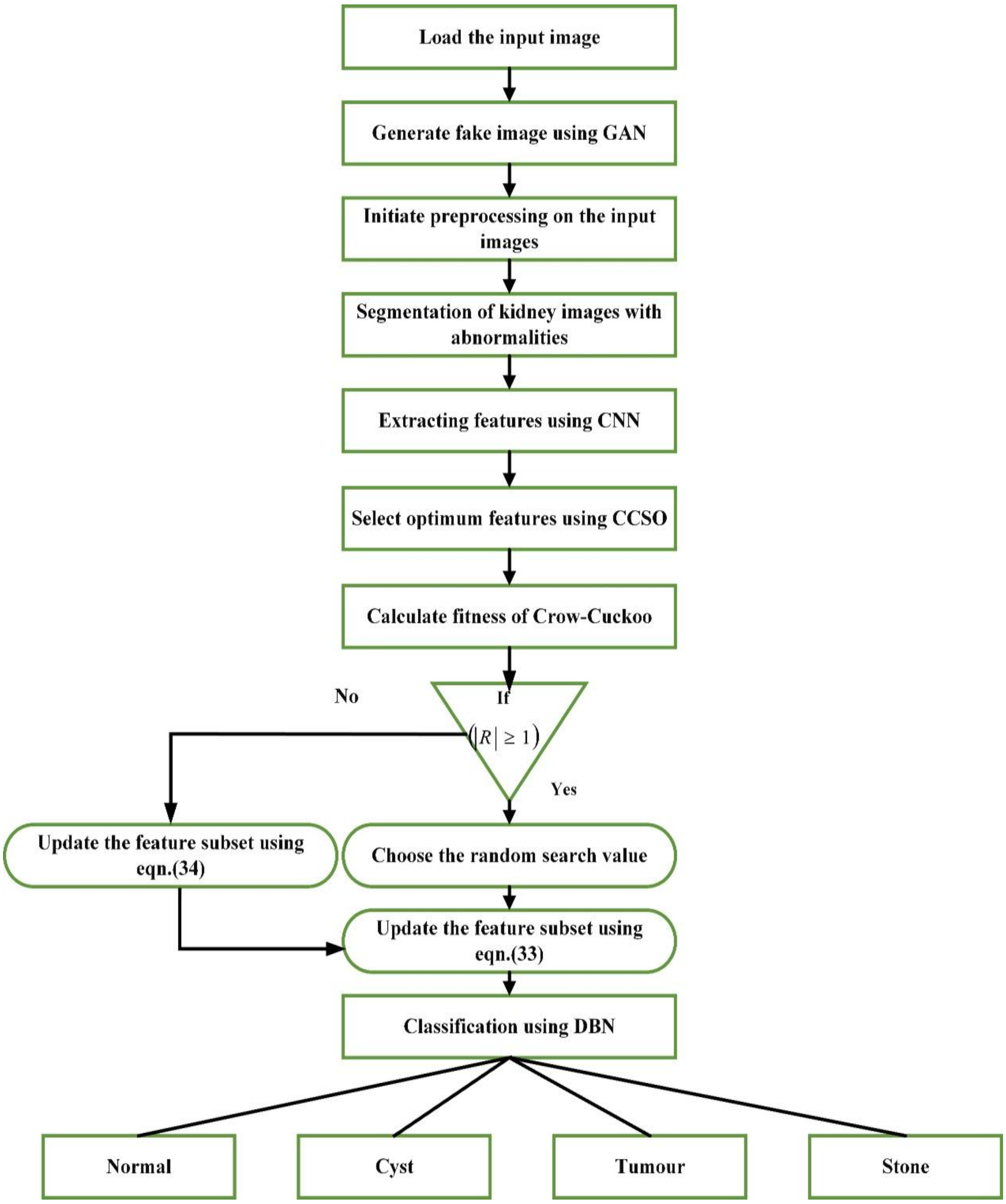

Fig. 4. Flowchart of proposed mechanism.

IV. Results and Discussion

This research is implemented on MATLAB tool has been successfully achieved. The proposed study aimed to enhance the accuracy of stone detection in kidney by employing a novel hybrid Convolutional Crow with Cuckoo Search Optimization (CCSO) with Convolutional Deep Belief Network (CDBN) framework for kidney stone detection and classification. The novel method used GAN to create realistic and variety fake images, which were then mixed with the original dataset to supplement training images and improve model generalization. In this study, Diffused Anisotrophic Frost Filter with CLAHE (DAFF-CLAHE) is utilized, and hence image quality (global contrast) is enhanced. Specific regions are identified within the image, with Marker based watershed segmentation. Hybrid CCSO was used to optimize the network design and hyperparameters, improving the model's overall performance. Extensive tests on a large-scale image were carried out, and the findings showed a considerable improvement in stone detection, classification and localization accuracy. The following metric was used to assess the model efficiency of the strategy. The main innovation of this paper is the novel hybrid approach that combines deep learning and optimization techniques for efficient and effective kidney stone detection in medical imaging. This approach leverages watershed segmentation to improve feature extraction, employing a Convolutional Neural Network (CNN) for hierarchical feature learning and a Deep Belief Network (DBN) for effective classification. Furthermore, a hybrid optimization scheme that integrates Crow with Cuckoo Search Optimizer (CCSO) is introduced to enhance the model's performance. Classified kidney images with normal, cyst, tumor, and stone are described in Fig. 5.

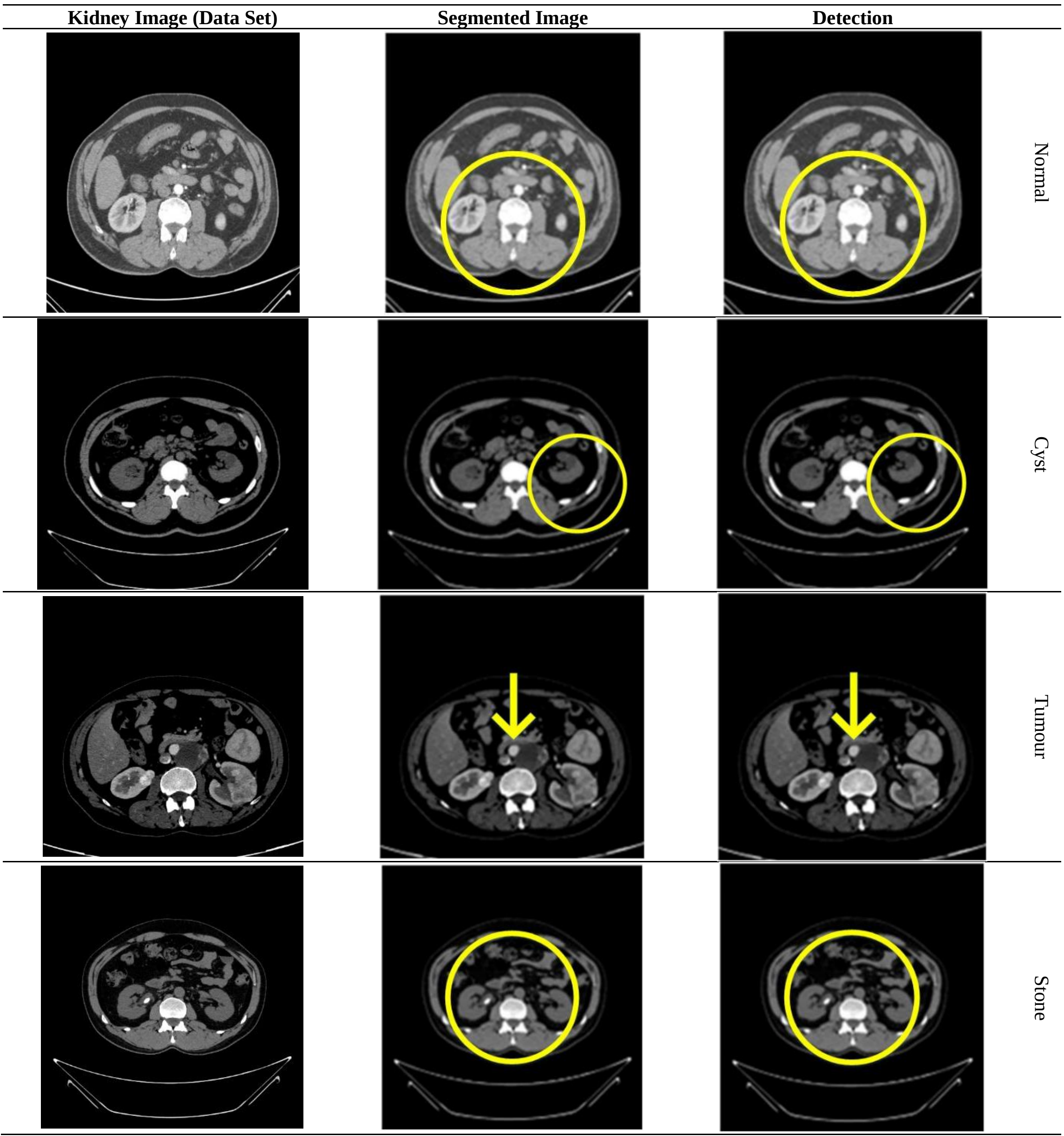

Fig. 5. Classified outcome.

For comparison, the following segmentation kidney stone detection evaluation criteria were used: precision, recall, F1–Score, and accuracy. These parameters were used to assess the model.

Accuracy: It calculates the fraction of real outcomes including true positives and true negatives across all cases investigated. It is expressed in Eq. (21).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{21}$$

Precision: The ratio of exactly anticipated positive outcomes to overall predicted positive occurrences is defined as precision. The precision is calculated using Eq. (22).

$$Precision = \frac{TP}{TP+FP} \tag{22}$$

Recall: The recall measures the proportion of genuine positive samples that were projected to be positive. Using Eq. (23), calculate the value recall.

$$Recall = \frac{TP}{TP+FN} \tag{23}$$

F1–Score: In the categorization task, recall and accuracy relate to one another. Although a high value for both is ideal, the reality is generally great accuracy with low recall, or high recall with low accuracy. To account for both recollection and accuracy, the F1–Score, which is a mean of recall and accuracy, can be employed. Eq. (24) shows the definition of F1–Score.

$$F1\text{–}Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{24}$$

Using Eq. (25), the Area Under the Curve (AUC) has been calculated to estimate overall performance.

$$AUC = \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right) \tag{25}$$

The accuracy level fluctuation graphs for the entire Hybrid Marker based segmentation-with CCSO model procedure are displayed in Fig. 6.

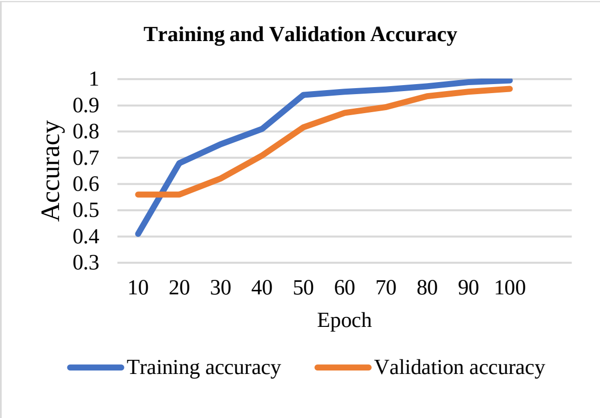

Fig. 6. Graphical depiction for training and validation accuracy of proposed method.

The loss rates fluctuation graphs for the entire Marker based segmentation-with CCSO model procedure are displayed in Fig. 7.

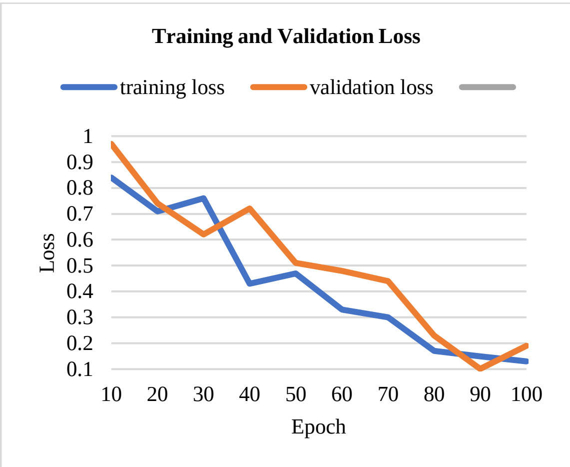

Fig. 7. Graphical representation of loss in proposed marker based segmentation-with CCSO.

Before optimization the value of accuracy for proposed Marker based segmentation is 99.35%. The accuracy achieved after optimization using CCSO is 99.40%. The fitness of CCSO is depicted in Fig. 8. Fig. 9 shows ROC curve for the Marker based segmentation-with CCSO model, and it can be seen that the ROC area is nearly close to 1, confirming the model's good stability and potential for usage as classification model for kidney stone detection.

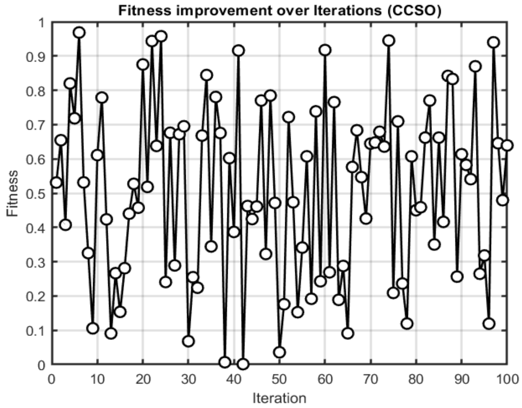

Fig. 8. Fitness improvement over iterations (CCSO).

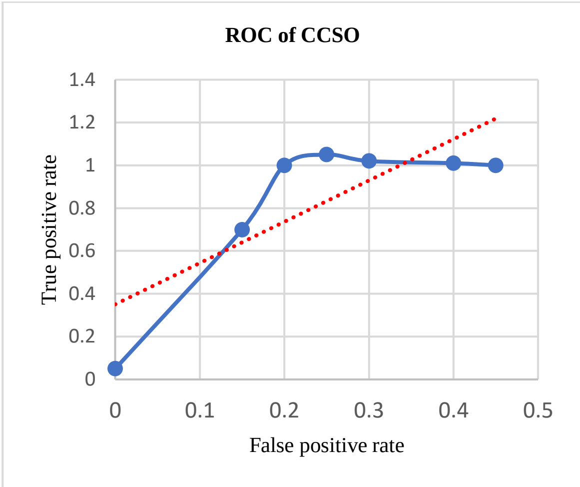

Fig. 9. ROC of CCSO.

The assessment results of the created kidney stone detection system employing the combined strategy are shown in Table I. At 99.35%, the accuracy is exceptionally high. Precision, which measures the percentage of accurate positive predictions compared to all positive forecasts, is an outstanding 99.3%. The system's capacity to accurately detect real positive cases is demonstrated by the recall measure, also known as sensitivity which is remarkably high at 98.4%. At 98.67%, the F1–Score, which balances recall and accuracy, is very impressive.

TABLE I. PERFORMANCE METRICS OF PROPOSED METHOD

Performance Metrics	Values (%)
Accuracy	99.35
Precision	99.3
Recall	98.4
F1–Score	98.67

With high values across key performance parameters, these findings show precise and reliable kidney stone detection. Fig. 10 illustrates the performance assessment of proposed Marker based segmentation-with CCSO. The comparison of the proposed method is presented in Table II.

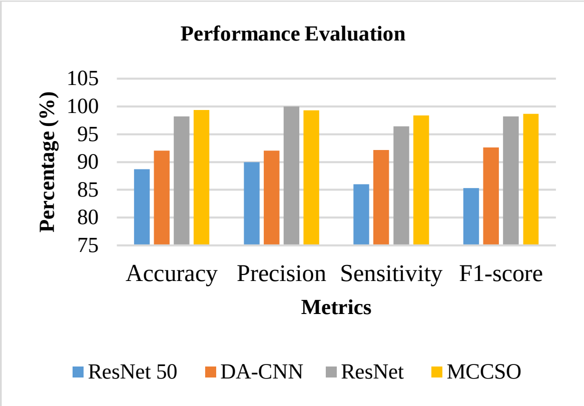

Fig. 10. Performance assessment of proposed marker based segmentation-with CCSO.

TABLE II. COMPARISON OF PROPOSED AND EXISTING METHODS

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1–Score (%)
ResNet 50 [26]	88.7	90.0	86.0	85.3
DA-CNN (6Conv+3FCL) [27]	92.07	92.06	92.18	92.6
DL with ResNet [28]	98.2	100	96.41	98.2
Proposed Method (Marker based segmentation-with CCSO-DBN)	99.35	99.3	98.4	98.67

Table II shows the comparison between the proposed and existing methods. The accuracy of the suggested method Marker based segmentation-with CCSO (99.35%) is higher than the existing approaches ResNet 50 (88.7%), DA-CNN (92.07%), and DL with ResNet (98.2%). Fig. 10 depicts the graphic depiction of the performance metrics of proposed with existing approaches. The precision of the suggested method Marker based segmentation-with CCSO (99.1%) is higher than the existing approaches ResNet 50 (90.0%), DA-CNN (92.06%), and DL with ResNet (100%). The recall of the suggested method Marker based segmentation-with CCSO-DBN (98.4%) is higher than the existing approaches ResNet 50 (86.0%), DA-CNN (92.18%), and DL with ResNet (98.4%). The F1–Score of the suggested method Marker based segmentation-with CCSO-DBN (98.57%) is higher than the existing approaches ResNet 50 (85.3%), DA-CNN (92.6%), and DL with ResNet (98.2%).

In Table III, two distinct architectures are evaluated for their stone detection capabilities. The first architecture, which combines modified ImageNet-SB network, achieves an Area Under the Curve (AUC) of 0.936. This indicates a strong ability of the ImageNet-SB combination to victimize kidney stone. In comparison, the proposed approach, utilizing Marker based segmentation-with CCSO for feature selection in conjunction network, outperforms the former with an AUC of 0.994. The higher AUC value achieved by the proposed Marker based segmentation-with CCSO architecture underscores its superior ability to effectively distinguish stones in kidney cases from normal kidney instances, highlighting its potential as an advanced and promising model for accurate stone detection in kidney images.

TABLE III. AUC OF PROPOSED MARKER BASED SEGMENTATION-CCSO-DBN WITH EXISTING MODEL

Architecture	AUC
ImageNet-SB	0.936
Marker based segmentation-with CCSO-DBN	0.994

V. CONCLUSION

The research introduces an AI-powered framework that combines Marker based segmentation-with CCSO with Generative Adversarial Networks (GANs) and Deep Belief Networks (DBN) for kidney stone detection. This study serves as an illustrative example of CCSO' application in identifying stone detection from kidney images. GANs are harnessed to address data imbalance, generating additional data to enhance detection. Furthermore, the research devises an ensemble model that synergizes Marker based segmentation and GANs, surpassing the performance of the standalone deep learning model. Employing an iterative CCSO evolutionary method enhances a population of solutions via fitness assessment. Depending on outcomes, this approach could potentially outperform existing artificial intelligence methods. Ultimately, the proposed hybrid deep learning framework, incorporating GANs and CCS optimization, notably improves the accuracy of kidney stone detection and localization within kidney images. The integration of GANs and CCS optimization effectively tackles data limitations and optimizes deep learning models, contributing to the advancement of stone diagnosis. Future study will focus on improving the hybrid approach's scalability to bigger datasets, investigating how it may be used to a variety of medical illnesses, and dealing with practical issues including patient diversity and varying imaging quality. The actual application of this method in medical contexts will also need completing clinical trials and improving ethical issues. The deployment and clinical validation of this extremely accurate kidney stone detection model in the future might pave the path for its incorporation into standard medical practice.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

G. Ramesh Babu: conceptualization, methodology, investigation, formal analysis, writing—original draft; N. Pushpalatha: conceptualization, methodology, visualization, investigation, formal analysis; Ganesh Khekare: supervision, validation, formal analysis, writing—reviewing and editing, resources; Krishnamoorthy: methodology, investigation; Yousef A. Baker El-Ebiary: formal analysis, writing—original; S. Anjali Devi: investigation, formal analysis, supervision, validation; all authors had approved the final version.

REFERENCES

[1] K. M. Black, H. Law, A. Aldoukhi, J. Deng, and K. R. Ghani, "Deep learning computer vision algorithm for detecting kidney stone

composition: CNNs to predict kidney stone composition," *BJU Int.*, vol. 125, no. 6, pp. 920–924, Jun. 2020.

[2] G. O. Ruiz *et al.* (2022). On the in vivo recognition of kidney stones using machine learning. [Online]. Available: http://arxiv.org/abs/2201.08865

[3] 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, 104569, Aug. 2021.

[4] K. K. Patro *et al.*, "Application of Kronecker convolutions in deep learning technique for automated detection of kidney stones with coronal CT images," *Information Sciences*, vol. 640, 119005, Sep. 2023.

[5] A. Nithya, A. Appathurai, N. Venkatadri, D. R. Ramji, and C. Anna Palagan, "Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images," *Measurement*, vol. 149, 106952, Jan. 2020.

[6] W. Preedanan, I. Kumazawa, T. Kondo, and I. Junichiro, "Urinary stones segmentation in abdominal X-Ray images based on U-Net deep learning model and data augmentation techniques," in *Proc. 2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP)*, Nanjing, China, Oct. 2020, pp. 118–123.

[7] W. Preedanan *et al.*, "Improvement of urinary stone segmentation using GAN-based urinary stones inpainting augmentation," *IEEE Access*, vol. 10, pp. 115131–115142, 2022.

[8] S. Saranya, "A department of electronics and communication engineering, Sri Sairam engineering college, Chennai, India," *International Journal of Innovative Science and Research Technology*, vol. 7, no. 6, 2022.

[9] S. Saranya, "Performance evaluation of Weil codes and gold codes for application in future navigation signals," in *Proc. 2022 IEEE 2nd Mysore Sub Section International Conference*, 2022, vol. 7, no. 6.

[10] T. Borges, A. Rai, D. Raj, D. Ather, and K. Gupta, "Kidney stone detection using ultrasound images," *SSRN Journal*, vol. 1, 2022.

[11] L. B. D. Cruz *et al.*, "Kidney segmentation from computed tomography images using deep neural network," *Computers in Biology and Medicine*, vol. 123, 103906, Aug. 2020.

[12] 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, Apr. 2022.

[13] K. K. Patro *et al.*, "Application of Kronecker convolutions in deep learning technique for automated detection of kidney stones with coronal CT images," *Information Sciences*, vol. 640, 119005, Sep. 2023.

[14] F. Turk *et al*., "Kidney and renal tumor segmentation using a hybrid v-net-based model," *Mathematics*, vol. 8, no. 10, 1772, Oct. 2020.

[15] M. N. Islam, M. Hasan, M. K. Hossain, M. G. R. Alam, M. Z. Uddin, and A. Soylu, "Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from CT-radiography," *Scientific Reports*, vol. 12, no. 1, 11440, 2022.

[16] S. K. Jain and R. K. Ray, "Non-linear diffusion models for despeckling of images: Achievements and future challenges," *IETE Technical Review*, vol. 37, no. 1, pp. 66–82, Jan. 2020.

[17] S. Tebini, Z. Mbarki, H. Seddik, and E. B. Braiek, "Rapid and efficient image restoration technique based on new adaptive anisotropic diffusion function," *Digital Signal Processing*, vol. 48, pp. 201–215, Jan. 2016.

[18] M. Marsousi, K. N. Plataniotis, and S. Stergiopoulos, "Shape-based kidney detection and segmentation in three-dimensional abdominal ultrasound images," in *Proc. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Chicago, IL, Aug. 2014, pp. 2890–2894.

[19] W. Wieclawek, "3D marker-controlled watershed for kidney segmentation in clinical CT exams," *BioMed Eng OnLine*, vol. 17, no. 1, 26, Dec. 2018,

[20] B. Shah, J. Modh, and S. Shah, "A modified marker-controlled watershed algorithm with linear convolution for medical image segmentation," *International Journal of Computational Engineering Research*, vol. 2, 2012.

[21] P. Santhiya, K. Arundhathivel, P. Gomathy, and C. Loganayaki, "Cuckoo search algorithm—Artificial neural networks for kidney stone detection," *International Research Journal of Engineering and Technology*, vol. 7, no. 9, 2020.

[22] O. B. Haddad, "Advanced optimization by nature-inspired algorithms," in *Studies in Computational Intelligence*, Singapore: Springer Singapore, 2018, vol. 720.

[23] S. Krishnamoorthy, Y. Liu, and K. Liu, "A novel oppositional binary crow search algorithm with optimal machine learning based postpartum hemorrhage prediction model," *BMC Pregnancy and Childbirth*, vol. 22, no. 1, p. 560, 2022.

[24] J. Naskath, G. Sivakamasundari, and A. A. S. Begum, "A study on different deep learning algorithms used in deep neural nets: MLP SOM and DBN," *Wireless Pers Commun*, vol. 128, no. 4, pp. 2913–2936, Feb. 2023.

[25] B. Leena and A. Jayanthi, "Brain tumor segmentation and classification via adaptive CLFAHE with hybrid classification," *Int. J. Imaging Syst. Technol.*, vol. 30, no. 4, pp. 874–898, Dec. 2020.

[26] J. G. Zapata *et al.*, "A metric learning approach for endoscopic kidney stone identification," *Expert Systems with Applications*, vol. 255, 2023.

[27] L. Wang, Z. Xu, and S. Wang, "Effect of data augmentation of renal lesion image by nine-layer convolutional neural network in kidney CT," *Computer Modeling in Engineering and Sciences*, vol. 124, no. 3, pp. 1001–1005, 2020.

[28] Y.-Y. Liu, Z.-H. Huang, and K.-W. Huang, "Deep learning model for computer-aided diagnosis of urolithiasis detection from kidney–ureter–bladder images," *Bioengineering*, vol. 9, no. 12, 811, Dec. 2022.