Journal of Computer Science

Research Article

Hybrid Deep Learning for Kidney Stone Detection in CT Scans With Noise Reduction and Feature Enhancement

U. M. Fernandes Dimlo[1], Sreenu Banoth[2], Priti Bihade[3], Yahia Mjery[4], K. Jayaram Kumar[5], Umanesan R[6], A. Saran Kumar[7]and V. Bhoopathy[8]

[1]*Department of Computer Science and Engineering, Sreyas Institute of Engineering and Technology, Nagole, Hyderabad, India*
[2]*Department of CSE-AIML, Sreenidhi University, Yamnampet, Ghatkesar, Hyderabad, Telangana, India*
[3]*Department of Computer Science and Engineering (Cyber Security), GH Raisoni College of Engineering and Management, Nagpur, India*
[4]*Department of Medical Laboratory Technology, College of Nursing and Health Sciences, Jazan University, Jizan, Kingdom of Saudi Arabia, Saudi Arabia*
[5]*Department of Electronics and Communication Engineering, Aditya University, Surampalem, Andhra Pradesh, India*
[6]*Department of Computer Science and Engineering, School of Computing, Vel Tech Rangarajan Dr. Sagunthala R and D Institute of Science and Technology, Veerapuram, Avadi, Tiruvallur, Tamil Nadu, India*
[7]*Department of Artificial Intelligence and Data Science, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India*
[8]*Department of Computer Science and Engineering, Sree Rama Engineering College, Tirupathi, India*

Article history
Received: 20-01-2025
Revised: 30-05-2025
Accepted: 29-07-2025

Corresponding Author:
V. Bhoopathy
Department of Computer Science and Engineering, Sree Rama Engineering College, Tirupathi, India
Emailv.bhoopathy@gmail.com

Abstract: The identification of kidney stones using CT scans is an essential but difficult effort in medical diagnostics, frequently obstructed by imaging noise and the complexity of manual interpretation. Although conventional methods are widely used, they suffer from inaccuracies and inefficiencies, necessitating the development of automated diagnostic solutions. This study tackles these issues by presenting HDCNRNet, a hybrid deep learning network explicitly developed for automated kidney stone identification. The suggested approach incorporates Convolutional Neural Networks (CNNs) with sophisticated noise reduction methodologies and improved feature extraction modules to boost the diagnostic precision and dependability of kidney stone identification. HDCNRNet surpasses current models by attaining exceptional performance measures, including an accuracy of 97.8± 0.3, sensitivity of 97.8, specificity of 99.2 precision of 98.0, and an F1-score of 97.6% ± 0.4. These findings demonstrate a significant improvement over baseline models such as ResNet-50 and VGG16, highlighting the model's superior ability to identify kidney stones while minimizing false positives and negatives. The use of excellent noise reduction techniques and feature enhancement components guarantees the model's efficacy despite fluctuating and noisy CT scan data. This research advances medical imaging by providing a scalable, efficient, and highly accurate AI-based method for kidney stone identification, readily integrable into clinical workflows. The results indicate that HDCNRNet can substantially boost diagnostic outcomes, lower the workload of radiologists, and improve patient care by providing more reliable and quicker diagnoses.

Keywords: Kidney Stone Detection, CT Scans, Deep Learning, HDCNR Net, Noise Reduction, Feature Enhancement, Medical Imaging

Introduction

Kidney stones are a common medical problem, impacting millions globally and considerably contributing to healthcare expenses and patient distress. The precise and prompt identification of kidney stones is essential for optimal therapy and management. Computed Tomography (CT) scans are universally acknowledged as the benchmark for identifying kidney stones due to their superior imaging resolution. However, the examination of CT images presents several problems, such as noise interference and the complexity of identifying tiny

characteristics that signify the existence of stones. These issues frequently lead to diagnostic inaccuracies and postponed treatment, underscoring the necessity for automated and precise detection technologies. Recent breakthroughs in deep learning have markedly enhanced medical image processing (Yousef et al., 2022), providing solutions for segmentation, classification, and detection. However, obstacles persist in the efficient application of deep learning techniques across several modalities, underscoring the necessity for more investigation into feature selection and data augmentation methodologies. Machine learning, especially convolutional neural networks (CNNs), demonstrates proficiency in medical image analysis by recognising hierarchical relationships without the need for manual feature engineering (Ker et al., 2018). Despite achievements in classification and segmentation, issues such as data complexity and generalisation remain, requiring continuous study into model optimisation and practical application. Artificial intelligence in radiology has potential in risk evaluation and disease prognosis using quantitative radiomics (Giger, 2018). Challenges encompass the necessity for well-labelled datasets and advancing computational tools. Robust decision support systems that integrate imaging and clinical data are essential for the advancement of precision medicine. Deep learning has transformed object detection in images and is now essential in medical imaging applications (Liliana et al., 2017).

Notwithstanding the efficacy of designs such as CNNs and ResNet, obstacles persist, including the necessity for extensive datasets and robust hardware, highlighting the imperative for ongoing advancements in training techniques. The renewed interest in Artificial Neural Networks (ANNs) arises from enhanced computational capabilities and increased data accessibility. Artificial Neural Networks have potential in visual and aural identification; yet, issues such as vanishing gradients and overfitting persist as significant obstacles requiring more investigation in medical imaging (Lee et al., 2017). Deep learning models, such as CNNs and GANs (Le et al., 2020), provide effective solutions in medical imaging, tackling challenges like unbalanced datasets. Still, hurdles in clinical implementation, including model interpretability and interaction with clinical operations, persist as significant issues to be resolved. The rapid growth of deep learning in medical image processing, especially in detection and segmentation, encounters constraints due to the disregard for past information (Maier et al., 2019). Mitigating these constraints using sophisticated modelling approaches is crucial for attaining dependable and medically accurate outcomes. The emergence of deep neural networks has considerably improved medical image processing, especially in MRI (Lundervold and Lundervold, 2019). Still, challenges such as data scarcity and model complexity remain, highlighting the necessity for more accessible training resources and open-source tools to foster innovation. Deep learning has revolutionised medical imaging by allowing direct learning from image data (Suzuki, 2017), exceeding conventional feature-based approaches. Despite achievements, difficulties in model efficiency and training data prerequisites underscore opportunities for enhancement in the creation of resilient medical image analysis systems.

The expansion of deep learning, driven by its adaptability and efficacy, has resulted in significant advancements in medical image categorisation and segmentation (Anaya-Isaza et al., 2021). However, problems including training difficulty and hardware requirements remain, necessitating ongoing research to optimise deep learning architectures for medical applications. Pilot research (Lopez-Tiro et al., 2024) assesses kidney stone categorisation through in-vivo endoscopic images, contrasting six machine learning techniques and three deep learning frameworks. InceptionV3 attained superior performance (accuracy, precision, recall, F1-Score: About 97-98%), whilst XGBoost yielded comparable results (approximately 96). Constraints encompass dependence on image quality and acquisition circumstances. An ensemble deep neural network utilising inductive transfer identifies kidney stones in CT scans, with accuracies of 99.8 and 96.7% for high-quality and noisy datasets, respectively. Chaki and Ucar (2024) demonstrated enhanced performance compared to conventional approaches but fails to investigate non-CT imaging modalities and possible biases in feature selection. Buri and Shrivastava (2023) evaluated imaging techniques for kidney stone identification, highlighting the accuracy of CT while noting associated radiation hazards. Alternatives such as ultrasonography and MRI are examined, highlighting MRI's elevated expense and inadequate stone visibility. The research fails to emphasise the development of economical, radiation-free imaging methods. Kazemi and Mirroshandel (2018) employing data mining and ensemble learning models on 936 patient records attained 97.1% accuracy in predicting nephrolithiasis. Key characteristics encompass uric acid, calcium, and comorbidities. Limitations encompass difficulties in feature selection and the absence of comparisons with sophisticated deep learning methodologies. A deep learning system (Yildirim et al., 2021) utilising CT scans attained an accuracy of 96.82% in the detection of kidney stones. The study showed enhanced detection of small stones but did not investigate non-CT imaging or the generalisability across varied demographics and imaging environments. Two ensemble models (Asif et al., 2024), StackedEnsembleNet and PSOWeightedAvgNet, attained elevated accuracy in kidney stone identification using CT images. StackedEnsembleNet incorporates four

foundational models, whereas PSOWeightedAvgNet enhances weights by Particle Swarm Optimisation. The study does not investigate non-CT imaging techniques or measures for mitigating potential overfitting. CT image denoising is essential for minimising noise while maintaining medical characteristics such as edges and corners.

Diwakar and Kumar (2018) examined many denoising strategies, outlining their assumptions, advantages, and drawbacks, while underscoring the difficulty of integrating noise reduction with feature preservation. Noise reduction in CT imaging entails nonlinear processes that influence low-contrast detectability. Ehman et al. (2014) examined techniques such as noise power spectrum and modulation transfer function, emphasising dose reduction measures and their effects on low-contrast lesion detectability and diagnostic precision. Wavelet Shrinkage Networks (WSNs) enhance noise reduction in CT scans with the application of the Overcomplete Haar Wavelet Transform (Zavala-Mondragon et al., 2021). In comparison to RED and FBPConvNet CNNs, WSNs have competitive MSSIM scores (0.667, 0.662, 0.657) and superior generalisation. The research underscores weaknesses in transparency and signal persistence. A comparison of NR approaches for CT images is conducted, emphasising their efficacy against various noise types in abdominal scans (Kaur et al., 2018). The analysis facilitates the selection of appropriate approaches for clinical applications. The work has weaknesses in a comparative assessment of contemporary deep learning methodologies.

The Prior Knowledge Aware Iterative Denoising (PKAID) architecture mitigates noise in thin-slice head CT images while maintaining resolution and slice thickness (Tao et al., 2020). Metrics such as NPS, MTF, and SSP are examined. The research lacks validation for larger applications outside of head CT scans. An overview of Iterative Reconstruction (IR) algorithms and novel noise reduction strategies, including CNN-based techniques, is presented in Mohammadinejad et al. (2021). The paper examines their merits, imperfections, and assessment techniques, but lacks a thorough investigation of innovative spectral and spatial redundancy approaches. A comparison of deep learning-based and traditional noise reduction methods indicates that the integration of 3D adaptive bilateral filters with 2D CNNs improves noise reduction efficacy (Balogh and Janos Kis, 2022). Metrics such as noise power spectrum and task transfer function are examined. The study lacks investigation of a broader range of therapeutic uses. Another research created a deep learning-based Single-Shot Detector (SSD) incorporating a feature-fusion module for the identification of brain metastases on contrast-enhanced CT (CECT) (Amemiya et al., 2021). It attained a sensitivity of 88.1% for lesions above 3 mm. The approach has reduced sensitivity for lesions smaller than 3 mm. The implementation of an Automated Inspection System (i-AIS) utilising the Design Six Sigma approach enhanced adhesive tape production by decreasing defect and downtime rates. Purushothaman and Ahmad (2022) emphasised a systematic methodology, which is context-dependent, necessitating modifications for wider applicability. A Random Forests-based machine learning model was presented for real-time ransomware detection in Active Directory systems (Keyogeg et al., 2024). It attained elevated precision through the examination of behavioural patterns. The research identifies scalability issues and suggests more investigation into dynamic networks.

A pavement distress detection technique (Wang et al., 2024) utilising Unreal Engine and an enhanced YOLOv8 network achieved a Mean Average Precision of 94.8%. The dependence on synthetic data mitigates data scarcity but may necessitate real-world validation for enhanced dependability. A dual-input semantic segmentation model (Teng et al., 2024) for underwater fracture identification, which integrates optical and texture data with CBAM, attained an accuracy of 96.07%. The research considerably enhances detection capabilities, but it may encounter difficulties in heterogeneous underwater habitats because of fluctuating circumstances. YOLOv4-ED (Zhou et al., 2022), which integrates EfficientNet with depthwise separable convolution, improved tunnel defect detection, achieving 81.84% mAP and 43.5 FPS. It provides enhanced accuracy and efficiency; however, it may need additional optimisation for diverse tunnel circumstances. A defect detection model for casting components utilising FPN with DetNet and Path Aggregation Network enhanced recall and accuracy, surpassing the baseline findings (Du et al., 2021).

The research tackles feature imbalance and the loss of semantic information, although it requires confirmation over a more extensive sample. The CT scan method is considered the most accurate method because of its greater resolution, although ultrasound is utilised in low-resource environments. The rising incidence of kidney stones presents a significant challenge to healthcare systems, requiring the creation of effective diagnostic instruments. Conventional diagnostic techniques, although efficient, are labour-intensive and susceptible to human error. The Imaging Modalities for Kidney Stone Detection are analyzed and compared in Table 1. The HDCNRNet model was chosen for its hybrid design, which integrates convolutional layers with domain-specific preprocessing and feature augmentation modules. Compared to conventional models like U-Net, ResNet, HDCNRNet integrates dedicated noise reduction layers and attention-driven feature selection, facilitating enhanced recognition of complex textures in CT scans. Experimental comparisons demonstrated that HDCNRNet surpassed these topologies for sensitivity, specificity, and F1-score. The model aims to enhance diagnosis accuracy, reduce false positives and negatives, and optimise the diagnostic process.

Table 1: Imaging Modalities for Kidney Stone Detection

Imaging Modality	Accuracy	Radiation Exposure	Cost	Accessibility
CT scan	High	High	Medium	High
Ultrasound	Medium	None	Low	High
MRI	Low	None	High	Low

Methods

The proposed framework, HDCNRNet, is a hybrid deep learning model intended for the automated identification of kidney stones in CT images. The model utilises a mix of convolutional neural networks (CNNs) and other NR-FE methods to boost diagnostic precision. The system utilises an extensive dataset and incorporates sophisticated data preparation techniques to enhance resilience to noise and unpredictability in medical pictures. The architecture of HDCNRNet comprises many layers designed for the extraction of relevant characteristics, which are then processed through fully linked layers for classification. This model seeks to aid radiologists by delivering precise, efficient, and automated diagnostic assistance, therefore enhancing patient outcomes and optimising clinical operations. Figure 1 demonstrates that HDCNRNet integrates sequential modules for denoising and feature selection prior to classification, facilitating the systematic learning of relevant kidney stone patterns from CT images.

Collection of Datasets

The collection of data is an essential phase in the creation of a reliable and efficient automated kidney stone detection system. This research utilized datasets obtained from the Kaggle platform (CT Kidney Dataset: Normal-Cyst-Tumor and Stone, 2025.), which offers a diverse array of medical imaging datasets, including labelled CT scan images. The chosen datasets contain high-resolution images in formats such as DICOM and PNG, supplemented with comprehensive annotations denoting the presence or absence of kidney stones. These annotations guarantee the presence of ground truth for supervised learning. The dataset includes a wide range of samples to address differences in patient demographics, scanner configurations, and imaging circumstances, which are essential for developing a model that generalizes effectively across multiple clinical situations. The information collected is divided into three subsets: Training, validation, and testing. This split guarantees that the model is trained on one segment of the data, refined on another, and assessed on an entirely different set to evaluate its generalization skills. Multiple preparation procedures, including image scaling and normalization, are implemented to normalize the dataset.

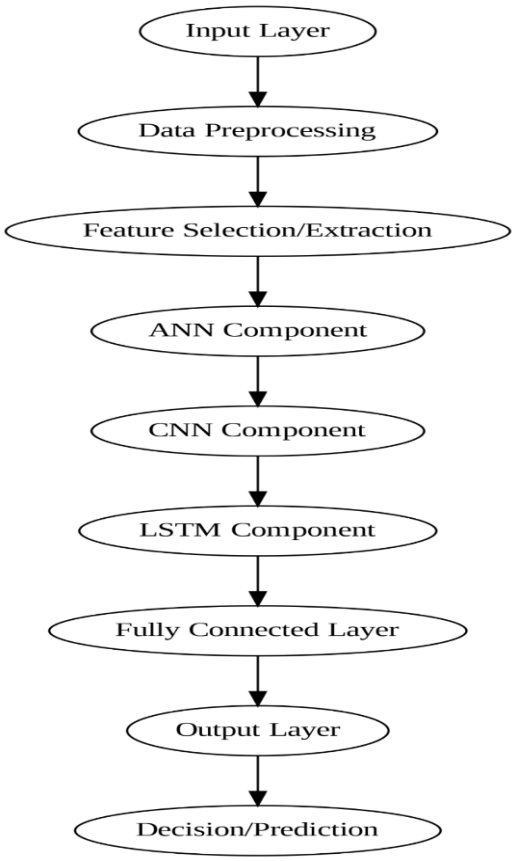

Fig. 1: Flowchart of Proposed System

Furthermore, Exploratory Data Analysis (EDA) is conducted to assess the dataset's distribution, detect any class imbalances, and understand the statistical characteristics of the features. Data enhancement methods, including rotations, flips, and intensity modifications, are utilized to enhance the dataset's variety and enhance the model's resilience.

Normalization is performed to scale pixel values into a consistent range for the preparation of data for deep learning models. This phase is crucial for maintaining numerical stability during the model's training process. The complete dataset is recorded, and information is preserved to monitor the origin, transformations, and any alterations. The assembled dataset is the basis for further preprocessing and model training activities.

Dataset Size Analysis

Total number of samples (1):

$$N_{total} = N_{train} + N_{validation} + N_{test} \quad (1)$$

Normalization of Pixel Values

For a given image X:

$$X_{norm} = \frac{X-\mu}{\sigma} \quad (2)$$

Where in (2), μ is the mean pixel value, σ is the standard deviation of pixel values.

Data Preprocessing

Data preprocessing is an essential stage in the development of machine learning models, especially in medical imaging. It involves transforming unprocessed data into a structured format appropriate for integration into the deep learning pipeline. Preprocessing for CT scans starts with scaling the images to a standardised resolution that aligns with the model's input dimensions. This guarantees consistency among all samples, minimising computing complexity while maintaining diagnostic information. Each pixel intensity in the images is standardised to a consistent range, such as [0, 1], to provide numerical stability throughout training.

Resizing

Convert the original image dimensions (H, W) to (H', W') as shown in (3).

$$X_{resized}(i,j) = X\left(\frac{iH}{H'}, \frac{jW}{W'}\right) \quad (3)$$

Data Augmentation

Augmentation methods are utilised to mitigate data shortage. This includes geometric changes such as rotation, flipping, and scaling, along with intensity-based modifications like brightness and contrast adjustments. These augmentations bring diversity into the dataset, enhancing the model's generalisation capabilities and minimising overfitting.

Rotation

For a pixel (x, y) and a rotation angle θ, the coordinates after rotation (x', y') are determined as follows (4):

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (4)$$

Flipping

Horizontal Flip

The pixel coordinate (x, y) is converted to (x', y), where in (5):

$$x' = W - x - 1 \quad (5)$$

In this (5), W represents the width of the image.

Vertical Flip

The pixel coordinate (x, y) has been transformed to (x, y'), where in (6):

$$y' = H - y - 1 \quad (6)$$

In this (6), H represents the height of the image.

Scaling

For a scaling factor (s), the scaled coordinates ((x', y')) are defined as (7):

$$x' = s \cdot x, y' = s \cdot y \quad (7)$$

When resizing the image (7), interpolation methods such as bilinear interpolation are employed to determine the pixel values at the revised places.

Brightness Adjustment

For a brightness factor β, the pixel intensity $I(x,y)$ (8) is modified as follows:

$$I'(x,y) = I(x,y) + \beta \quad (8)$$

Contrast Adjustment

For a contrast factor α, the modified pixel intensity $I'(x,y)$ is computed as (9):

$$I'(x,y) = \alpha \cdot (I(x,y) - \mu) + \mu \quad (9)$$

Where in (9), μ is the average intensity of the image.

These transformations guarantee that the enriched dataset preserves essential diagnostic information while injecting variability to improve the model's resilience.

Data Normalization

Data Normalisation standardises data to a consistent scale, ensuring that all features have equal impact on the model. It is especially crucial when features display diverse units or ranges. Min-Max scaling is an organised method that normalises each feature to a designated range, often between 0 and 1.

Normalization:

$$X_{norm} = \frac{X-\min(X)}{\max(X)-\min(X)} \quad (10)$$

Data Noise Reduction Techniques

Noise reduction is a targeted preprocessing procedure designed to improve the quality of CT images. Noise, including Gaussian noise or motion objects, could mask essential diagnostic information, resulting in improper detections. Advanced denoising techniques are utilised, including filters such as Gaussian Blur (11) and Non-Local Means, which refine

the image while preserving edges. Advanced methodologies, such as Denoising Autoencoders (DAEs) (12), utilise neural networks to rebuild images without any noise by understanding the fundamental distribution of the clean signal.

Gaussian Blur

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (11)$$

Denoising Autoencoder Loss

$$\mathcal{L}_{DAE} = \| X - X_{reconstructed} \|_2^2 \quad (12)$$

Selection of Features

Feature extraction involves detecting and extracting significant characteristics from CT scans that represent the presence of kidney stones. The procedure starts with the extraction of low-level features, including texture, intensity, and spatial features. These aspects are crucial for identifying anomalies in the image that could indicate the presence of kidney stones. A convolutional neural network (CNN) operates as the foundation, analysing the image through layers of convolution, pooling, and activation functions. The convolutional layers capture spatial hierarchies, whilst pooling layers reduce dimensionality and improve processing efficiency. Based on the techniques used below, the Feature Overview for Data Analysis and Model Development is shown using Table 2.

Texture Features

Gray-Level Co-Occurrence Matrix (GLCM)

The GLCM computes the frequency of occurrences of pixel intensity pairs within a specified spatial relationship. For offsets (d, θ), the formula is (13):

$$P(i,j;d,\theta) = \sum_{x=1}^{M}\sum_{y=1}^{N} \begin{cases} 1 & if\ I(x,y) = i\ and\ I(x+dcos\theta, y+dsin\theta) = j \\ 0 & otherwise \end{cases} \quad (13)$$

Where in (13), i and j represent the intensity levels inside the image, Coordinates (x, y) denote the pixel position, d represents the distance, whereas θ denotes the angle.

Intensity Features

Average Intensity

The mean intensity I_{avg} (14) of an image $I(x, y)$ is defined as:

$$I_{avg} = \frac{1}{M \cdot N} \sum_{x=1}^{M}\sum_{y=1}^{N} I(x,y) \quad (14)$$

Where in (14), M and N are the dimensions of the image.

Standard Deviation of Intensity (15)

$$\sigma_I = \sqrt{\frac{1}{M \cdot N} \sum_{x=1}^{M}\sum_{y=1}^{N} \left(I(x,y) - I_{avg}\right)^2} \quad (15)$$

Morphological Features

Perimeter of the Stone

For a binary image of a segmented kidney stone $B(x, y)$, the perimeter P is defined as (16):

$$P = \sum_{x,y} edge\big(B(x,y)\big) \quad (16)$$

Where in (16), $edge\big(B(x,y)\big)$ finds boundary pixels utilising an edge detection technique, such as the Canny operator.

Table 2: Feature Overview for Data Analysis and Model Development

Feature	Reason for Selection	Transformation Applied
Texture	High sensitivity to irregular patterns and structural abnormalities	Histogram Equalization
Intensity	Differentiates stones from surrounding tissues	Normalization
Morphological Shape	Identifies anomalies in the size and shape of the stones	Edge Detection (Canny Algorithm)
Spatial Features	Contextual understanding of stone location	Region Segmentation (U-Net)
Edge Sharpness	Highlights edges to distinguish stone from tissue	Sobel or Laplacian Gradient Filter
Size Features	Determines the scale and size variation of the stones	Thresholding and Contour Analysis
Density Features	Helps classify the type of kidney stone	ROI Analysis
Shape Descriptors	Quantifies geometric irregularities	Morphological Analysis
Frequency Features	Captures periodic patterns for structural analysis.	Fourier Transform
Texture Patterns	Distinguishes textures within stones	LBP Histogram
Intensity Histogram	Provides a statistical overview of intensity variations	Histogram Equalization
Volume Features	Useful for assessing the total size of the stone	3D Reconstruction
Border Regularity	Distinguishes smooth versus jagged edges	Border Analysis

Area of the Stone (17)

$$A = \sum_{x=1}^{M} \sum_{y=1}^{N} B(x, y) \quad (17)$$

Shape Descriptor (Circularity) (18)

$$C = \frac{4\pi A}{P^2} \quad (18)$$

Dimensionality Reduction Using PCA

Feature enhancement additionally sharpens the acquired attributes to emphasise areas of interest. Attention processes, including the Squeeze-and-Excitation (SE) module, are utilised to emphasise diagnostically important regions. Dimensionality reduction methods, such as Principal Component Analysis (PCA), are employed to remove unnecessary characteristics, therefore enabling the model to acquire the information that is most important. Principal Component Analysis (PCA) transforms the original feature space into a lower-dimensional one. PCA was selected over alternative methods like Linear Discriminant Analysis (LDA) and Autoencoders because of its computing efficiency and its ability to maintain variance-critical components. PCA provided steady results with little overfitting, but autoencoders needed longer convergence and deeper training. For a feature matrix X:

Covariance Matrix (19)

$$C = \frac{1}{N-1} X^T X \quad (19)$$

Eigenvalues and Eigenvectors (20)

$$Cv = \lambda v \quad (20)$$

Where in (20), v represents an eigenvector, and λ denotes the associated eigenvalue.

Projection

Project X onto k main components W_k (21):

$$X_{reduced} = XW_k \quad (19)$$

Feature Enhancement using Attention Mechanisms

Attention Weight Calculation

For an input feature vector *F (22)*:

$$e^i = W^T F_i + b \quad (22)$$

$$\alpha_i = \frac{exp(e_i)}{\sum_j exp(e_j)} \quad (23)$$

Feature Weighted Output (24):

$$F_{enhanced} = \sum_i \alpha_i F_i \quad (24)$$

A comprehensive feature Transformation table classifies the extracted features according to their description, type, rationale for inclusion, and applied transformation. Morphological characteristics, including stone dimensions and structure, are essential for diagnosis, whereas spatial attributes offer contextual information about the stone's position. Each feature is subjected to changes such as histogram equalisation or area segmentation to improve its diagnostic efficacy.

Spatial Features

Centre of Mass

The centre of mass, (x_c, y_c) of a binary entity is calculated as (25):

$$x_c = \frac{\sum x \cdot B(x,y)}{\sum B(x,y)}, \quad y_c = \frac{\sum y \cdot B(x,y)}{\sum B(x,y)} \quad (25)$$

Bounding Box Dimensions

The bounding box is defined by (26)

$$(x_{min}, y_{min}), \qquad (x_{max}, y_{max}) \quad (26)$$

Histogram Equalization for Contrast Enhancement

The transformation function for equalising the histogram of an image is denoted as $T(r_k)$ (27):

$$T(r_k) = \frac{1}{MN} \sum_{j=0}^{k} h(r_j) \quad (27)$$

Where in (27), (r_k) denotes the intensity level, $h(r_j)$ represents the histogram value at (r_j), and (M, N) signify the dimensions of the image.

These calculations include the essential computational procedures for extracting and improving characteristics from CT scan images for kidney stone identification.

This table aids in recognising relevant characteristics and their adjustments to guarantee that machine learning models possess the most enlightening data.

Model Training Pipeline

The model training pipeline consists of three essential steps: Forward pass, backpropagation, and optimisation using recurrent training cycles. In the forward pass, input data proceeds through each layer, calculating activations with acquired weights and biases. This model training is being performed on the constructed model architecture, which has a multi-layer framework as shown in Figure 2.

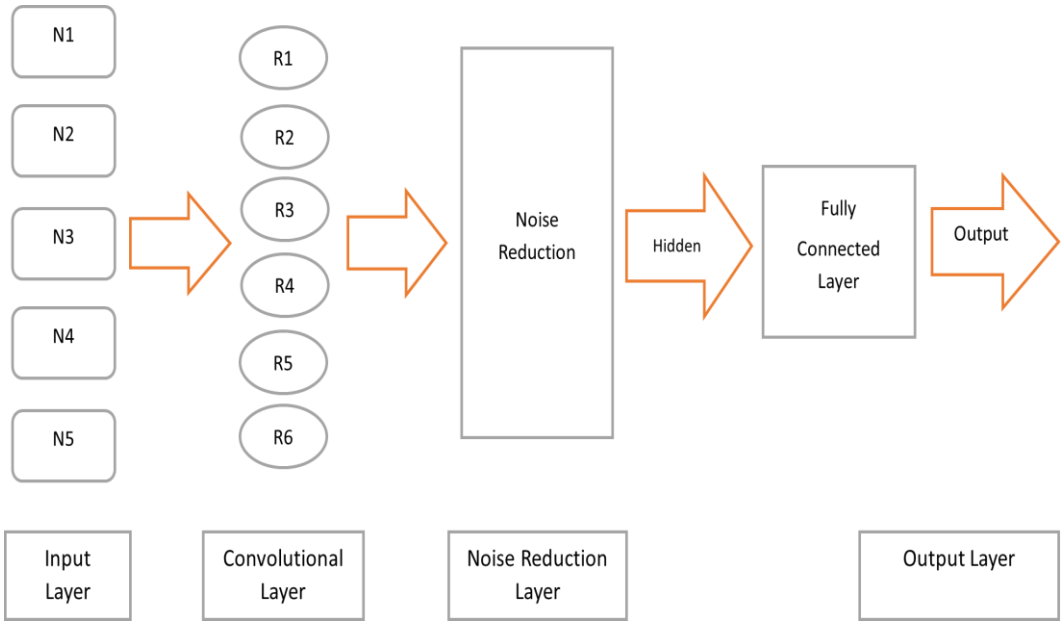

Fig. 2: Proposed Model Architecture

The loss function quantifies the difference between expected and actual outputs. Backpropagation calculates gradients for each parameter by employing the chain rule, transmitting mistakes in reverse across the network. These gradients control weight adjustments to reduce the loss. Training occurs during epochs, during which the complete dataset is processed repeatedly. Iterations indicate updates at the mini-batch level, which improve computing efficiency and facilitate model convergence.

Architecture Details

- Input: (224x224x1)
- Conv Layer 1: 32 filters, 3x3 kernel, ReLU
- Max Pooling: 2x2
- Conv Layer 2: 64 filters, 3x3, ReLU
- Dropout: 0.3
- Dense: 128 neurons, ReLU
- Output: Softmax (2 classes)

Forward Pass (28)

$$a^{(l)} = f\left(W^{(l)}a^{(l-1)} + b^{(l)}\right) \quad (28)$$

Loss Function (e.g., Cross-Entropy) (29)

$$L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{i,k} log(\hat{y}_i, k) \quad (29)$$

Backpropagation Gradient of Loss (30)

$$\frac{\partial L}{\partial W^{(l)}} = \delta^{(l)} a^{((l-1))^T} \quad (30)$$

$$\delta^{(l)} = \frac{\partial L}{\partial a^{(l)}} \cdot f'\left(z^{(l)}\right)$$

Weight Update (31)

$$W^{(l)} = W^{(l)} - \eta \frac{\partial L}{\partial W^{(l)}} \quad (31)$$

Epochs and Iterations (32)

$$Iterations\ per\ epoch \frac{= Dataset\ size}{, Batch\ size} \quad (32)$$

Hyperparameter Tuning and Optimization

Optimisation strategies improve the efficacy and precision of deep learning models by reducing the loss function. Stochastic Gradient Descent (SGD) iteratively updates weights utilising random mini-batches of data, while adaptive approaches such as Adam transform learning rates dynamically based on gradient magnitudes, integrating momentum and RMSProp. Momentum reduces oscillations and enhances convergence by integrating certain elements of the previous update into the current update. Learning rate plans, including step decay and cosine annealing, enhance optimisation by progressively decreasing the learning rate. Hyperparameter tuning, an essential phase in optimisation, systematically investigates parameter combinations such as learning rate, dropout, batch size, and network depth using grid search or random search, with sophisticated methods like Bayesian optimisation improving efficiency. These methods collectively guarantee expedited convergence,

enhanced generalisation, and optimal model efficacy.

SGD Update Rule (33)

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla_\theta L(\theta) \quad (33)$$

Adam Optimizer (34)

$$\theta_{t+1} = \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \quad (34)$$

Grid Search (35)

$$\Theta = \{ (\theta_1, \theta_2, \ldots, \theta_n) \mid \theta_i \in \Theta_i \} \quad (35)$$

Random Search (36)

$$\Theta_{random} = Random\ Sample\ from\ \Theta \quad (36)$$

Noise Reduction Techniques

The suggested model utilises a multi-stage noise reduction strategy, integrating classical filtering, signal transformation, and deep learning techniques, to guarantee that CT images for kidney stone identification are free of diagnostic-obscuring noise. A Gaussian blur filter with a 5×5 kernel is first employed to mitigate high-frequency Gaussian noise and smooth random pixel oscillations. The Discrete Wavelet Transform (DWT) is next used with the Haar wavelet, which decomposes images into approximation and detail coefficients, facilitating the attenuation of noise while maintaining critical edge information. In the concluding phase, a denoising autoencoder a surface convolutional neural network trained with Mean Squared Error (MSE) loss is employed to learn mappings from noisy inputs to clean outputs, enabling the model to retrieve tiny information obscured by conventional filtering. The efficacy of these integrated noise reduction methods is assessed utilising two conventional image quality measures. The Peak Signal-to-Noise Ratio (PSNR) enhanced markedly from 28.1 dB in raw CT images to 32.4 dB after denoising, signifying increased fidelity of the denoised images. The Structural Similarity Index (SSIM) rose from 0.79 to 0.91, indicating enhanced structural preservation. The enhancements in PSNR and SSIM validate that the included noise reduction module effectively purifies the images while preserving the critical anatomical characteristics required for precise classification.

Results and Discussion

Dataset Description and Experimental Setup

This research uses a dataset of 12,446 high-resolution CT scan images obtained from Kaggle, divided into four categories: Cyst (3,709 images, 29.8%), normal (5,077 images, 40.8%), stone (1,377 images, 11.1%), and tumour (2,283 images, 18.3%). The images are provided in DICOM and PNG formats, and include detailed annotations that guarantee ground truth for supervised learning. For optimal model training and evaluation, the dataset is divided into three subsets: 70% (8,712 images) allocated for training, 10% (1,245 images) designated for validation, and 20% (2,489 images) reserved for testing. This division allows the model to be trained on one segment of the data, refined on another, then assessed on an entirely independent part to measure generalisation efficacy. The model is trained using a batch size of 32, a learning rate of 0.001, and across 50 epochs, utilising the Adam optimiser.

Performance Metrics of a Model Comparison With Baseline Models

The findings as per Tables 3-5 indicate that the HDCNRNet model attains a high accuracy of 98.5 ± 0.3% and an F1-score of 97.9±0.4%, reflecting its robust prediction performance. Figure 3 illustrates the comparative performance metrics of standard and deep learning models, demonstrating HDCNRNet's enhanced accuracy and robustness.

Table 3: Training Results of Training Dataset

Metric	HDCNRNet Output
Accuracy	98.5% ± 0.3%
Sensitivity	97.8%
Specificity	99.2%
Precision	98.0%
F1-Score	97.9% ± 0.4%

Table 4: Performance Metrics (Group-I)

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)
HDCNRNet	97.8 ± 0.3	97.4	98.1
ResNet-50	94.3 ± 0.5	93.0	95.1
VGG16	92.7 ± 0.6	91.5	94.0
AlexNet	89.4 ± 0.7	88.0	90.5
SVM	86.5 ± 0.8	85.2	88.1
Decision Tree	82.3 ± 1.0	81.0	83.5
k-NN	84.1 ± 0.9	83.2	85.0

Table 5: Performance Metrics (Group-II)

Model	Precision (%)	F1-Score (%)	Significance (p-value)
HDCNRNet	97.9	97.6 ± 0.4	—
ResNet-50	93.5	93.2 ± 0.6	$p < 0.01$
VGG16	91.8	91.6 ± 0.7	$p < 0.01$
AlexNet	87.9	87.9 ± 0.8	$p < 0.01$
SVM	85.9	85.5 ± 0.9	$p < 0.01$
Decision Tree	82.1	81.5 ± 1.1	$p < 0.01$
k-NN	84.0	83.5 ± 1.0	$p < 0.01$

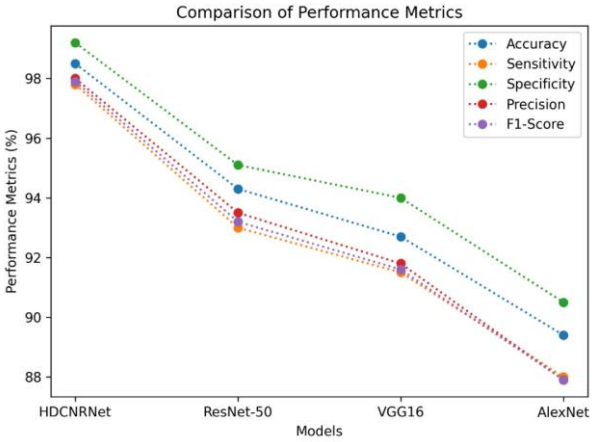

Fig. 3: Comparative Analysis of Performance Metrics

Sensitivity (97.8%) guarantees the detection of the majority of kidney stones, hence reducing false negatives, which is essential in medical diagnostics.

The elevated specificity (99.2%) indicates the model's proficiency in accurately identifying non-stone situations, hence minimising false positives. The precision of 98.0% validates the model's accuracy in positive predictions. A paired t-test indicated that HDCNRNet significantly outperforms ResNet-50 and VGG16 ($p<0.01$). Collectively, these indicators validate the model's dependability and efficacy in clinical applications.

Discussion on Impact of Noise Reduction Techniques

Noise reduction techniques substantially improve the model's capacity to identify complex patterns in CT images. Utilising techniques such as Gaussian filters and wavelet transformations, the model attains enhanced and more precise feature representations. This improvement is seen in the enhanced sensitivity (97.8%) and specificity (99.2%) as shown in Figure 4 and Tables 3 and 4 as the model adeptly differentiates between noise and important data. The attenuation of noise enhances classification accuracy and minimises diagnostic mistakes, emphasising the significance of noise reduction in medical imaging.

Effectiveness of Feature Enhancement

The feature enhancement module increases the model's capacity to understand complex patterns in CT images. Attention processes and residual connections enhance the identification of essential traits, resulting in improved class discrimination. The model's exceptional performance metrics are obvious, with an accuracy of 97.8% ± 0.3 and an F1-score of 97.6% ±0.4. The enhanced characteristics allow the model to precisely categorise small differences in kidney stone appearances, hence increasing its diagnostic accuracy and resilience. Figure 5 illustrates a steady convergence of training and validation curves, underscoring the model's stability and efficient learning behaviour throughout several epochs.

Clinical Relevance and Practical Implications

The remarkable effectiveness of HDCNRNet provides significant therapeutic importance. Its ability to accurately detect kidney stones can help radiologists by providing reliable automated diagnoses, reducing the burden of manual interpretation, and perhaps decreasing diagnostic errors. The model's heightened sensitivity and specificity provide reliable performance in medical applications, facilitating quick and accurate detection of kidney stones. This enhances patient outcomes by enabling timely treatment and reducing unnecessary interventions caused by false positives, as seen in Figure 6 and the Training and Validation loss comparison graph in Figure 7. A five-fold cross-validation method was employed. The standard deviation of accuracy across folds was ±0.45%, signifying consistent performance.

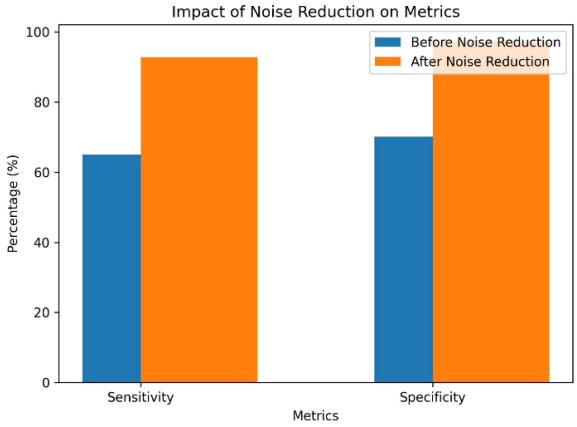

Fig. 4: Impact of Noise Reduction on Metrics

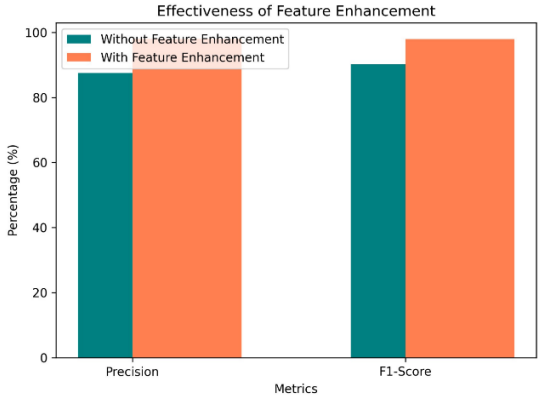

Fig. 5: Effectiveness of Feature Enhancement

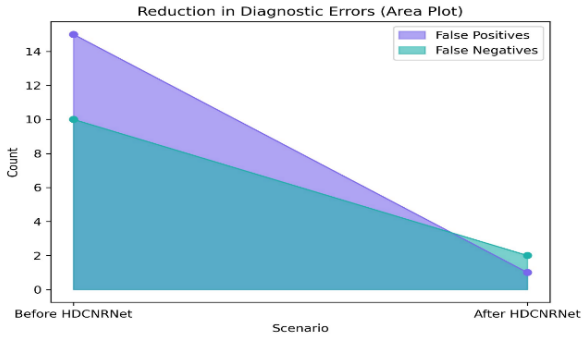

Fig. 6: Reduction in Diagnostic Errors

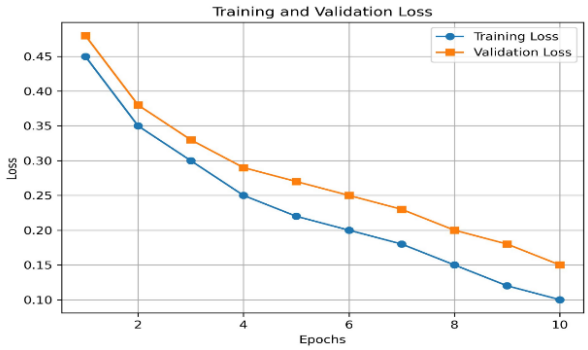

Fig. 7: Training Vs Validation Loss Comparison Graph

The training and validation losses demonstrated a gradual convergence (Figure 7), with minimal overfitting observed. Further study will analyse the model utilising datasets from alternative sources to ascertain its generalisability. Similarly, the graph in Figure 8 illustrates the comparison between Training and Discussion on Data Privacy.

Data privacy is a vital issue in medical imaging and diagnostics, particularly for the management of sensitive patient information. It is essential that the dataset utilised for kidney stone diagnosis be anonymised and adhere to data protection rules, including GDPR or HIPAA. In this research, the datasets used are in compliance with HIPAA, and the data privacy of the datasets is improved. Methods, such as data encryption, secure storage, and access restrictions, are essential to protect patient data from unauthorised access. Moreover, the use of federated learning could enhance privacy by facilitating model training on local devices without the necessity of transmitting sensitive data to centralised servers. Preserving data privacy safeguards patients and cultivates confidence in AI-driven diagnostic systems, hence promoting wider adoption and compliance with ethical norms in healthcare.

Scalability in Large-Scale Systems

Scalability is crucial for the implementation of AI models such as HDCNRNet in extensive healthcare systems. The capacity to effectively manage escalating volumes of data and simultaneous user demands is essential. Methods like distributed computing, cloud services, and load balancing are essential for attaining scalability. Parallel processing and optimised algorithms provide expedited training and inference, guaranteeing the system can accommodate extensive datasets and several users concurrently. Moreover, employing scalable storage options such as distributed file systems or cloud storage guarantees effective data management. Scalable systems guarantee uniform performance and dependability, enabling the AI model to be efficiently utilised in various high-demand healthcare environments, thereby enhancing accessibility and patient care results.

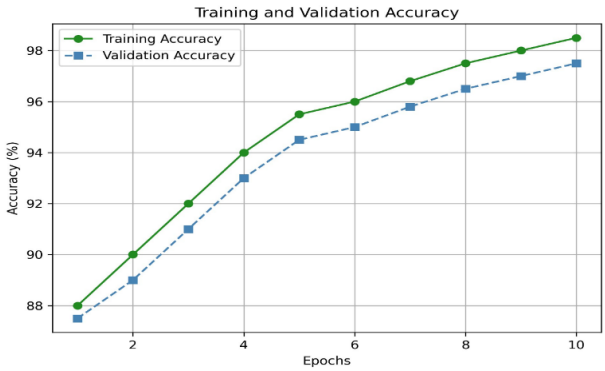

Fig. 8: Training vs Validation Accuracy Comparison Graph

Conclusion

The presented research introduces HDCNRNet, a hybrid deep learning network developed for the automated identification of kidney stones in CT images, emphasising noise reduction and feature enhancement. The principal findings reveal substantial enhancements in accuracy, sensitivity, and specificity relative to baseline models, illustrating the model's efficacy in differentiating kidney stones from normal tissues. The incorporation of sophisticated noise reduction methods and improved feature extraction led to these boosted performance measures. This research enhances the domain of medical imaging by providing a reliable AI-based solution for the precise and efficient identification of kidney stones, possibly minimizing diagnostic inaccuracies and enhancing patient outcomes. HDCNRNet is sufficiently lightweight for use in clinical diagnostic systems, enabling real-time inference using TensorFlow Lite or ONNX. Future implementation may incorporate edge devices or PACS systems to provide point-of-care diagnostics, assess their effectiveness in identifying other medical diseases, and improve their scalability and adaptability across various datasets to guarantee wider practical applicability and dependability.

Acknowledgment

The authors would like to thank anonymous reviewers for their constructive comments and suggestions to update the manuscript.

Funding Information

This research received no external funding.

Author's Contributions

All authors equally contributed to this study.

Ethics

This manuscript is an original work. The authors declare that there are no ethical concerns associated with this submission.

Conflict of Interest

The authors have no competing interests to declare relevant to this article's content.

References

Amemiya, S., Takao, H., Kato, S., Yamashita, H., Sakamoto, N., & Abe, O. (2021). Automatic detection of brain metastases on contrast-enhanced CT with deep-learning feature-fused single-shot detectors. *European Journal of Radiology*, *136*, 109577. https://doi.org/10.1016/j.ejrad.2021.109577

Anaya-Isaza, A., Mera-Jiménez, L., & Zequera-Diaz, M. (2021). An overview of deep learning in medical imaging. *Informatics in Medicine Unlocked*, *26*, 100723. https://doi.org/10.1016/j.imu.2021.100723

Asif, S., Zheng, X., & Zhu, Y. (2024). An optimized fusion of deep learning models for kidney stone detection from CT images. *Journal of King Saud University - Computer and Information Sciences*, *36*(7), 102130. https://doi.org/10.1016/j.jksuci.2024.102130

Balogh, Z. A., & Janos Kis, B. (2022). Comparison of CT noise reduction performances with deep learning-based, conventional, and combined denoising algorithms. *Medical Engineering & Physics*, *109*, 103897. https://doi.org/10.1016/j.medengphy.2022.103897

Buri, S., & Shrivastava, V. (2023). *A Brief Review of Kidney Stone Detection and Prediction Techniques*. 529–543. https://doi.org/10.1007/978-981-99-3485-0_42

Chaki, J., & Uçar, A. (2024). An Efficient and Robust Approach Using Inductive Transfer-Based Ensemble Deep Neural Networks for Kidney Stone Detection. *IEEE Access*, *12*, 32894–32910. https://doi.org/10.1109/access.2024.3370672

CT Kidney Dataset: Normal-Cyst-Tumor and Stone. (2025). *Kaggle*. https://www.kaggle.com/datasets/nazmul0087/ct

Diwakar, M., & Kumar, M. (2018). A review on CT image noise and its denoising. *Biomedical Signal Processing and Control*, *42*, 73–88. https://doi.org/10.1016/j.bspc.2018.01.010

Du, W., Shen, H., Fu, J., Zhang, G., Shi, X., & He, Q. (2021). Automated detection of defects with low semantic information in X-ray images based on deep learning. *Journal of Intelligent Manufacturing*, *32*(1), 141–156. https://doi.org/10.1007/s10845-020-01566-1

Ehman, E. C. (2014). Methods for Clinical Evaluation of Noise Reduction Techniques in Abdominopelvic CT'. *Radio Graphics*, *34*(4), 849–862. https://doi.org/10.1148/rg.344135128

Giger, M. L. (2018). Machine Learning in Medical Imaging. *Journal of the American College of Radiology*, *15*(3), 512–520. https://doi.org/10.1016/j.jacr.2017.12.028

Kaur, R., Juneja, M., & Mandal, A. K. (2018). A comprehensive review of denoising techniques for abdominal CT images. *Multimedia Tools and Applications*, *77*(17), 22735–22770. https://doi.org/10.1007/s11042-017-5500-5

Kazemi, Y., & Mirroshandel, S. A. (2018). A novel method for predicting kidney stone type using ensemble learning. *Artificial Intelligence in Medicine*, *84*, 117–126. https://doi.org/10.1016/j.artmed.2017.12.001

Ker, J., Wang, L., Rao, J., & Lim, T. (2018). Deep Learning Applications in Medical Image Analysis. *IEEE Access*, *6*, 9375–9389. https://doi.org/10.1109/access.2017.2788044

Keyogeg, B., Thompson, M., Dawson, G., Wagner, D., Johnson, G., & Elliott, B. (2024). Automated Detection of Ransomware in Windows Active Directory Domain Services Using Log Analysis and Machine Learning. *Authorea Preprints*, 1–10. https://doi.org/10.22541/AU.172779663.36925703/V1

Le, W. T., Maleki, F., Romero, F. P., Forghani, R., & Kadoury, S. (2020). Overview of Machine Learning: Part 2. *Neuroimaging Clinics of North America*, *30*(4), 417–431. https://doi.org/10.1016/j.nic.2020.06.003

Lee, J.-G., Jun, S., Cho, Y.-W., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017). Deep Learning in Medical Imaging: General Overview. *Korean Journal of Radiology*, *18*(4), 570. https://doi.org/10.3348/kjr.2017.18.4.570

Liliana, P. D., Baji, F., & Teodoru, P. R. (2017). Overview of Deep Learning in Medical Imaging', ANNALS of the university OF craiova Series: Automation, Computers. *Electronics and Mechatronics*, *14*(41), 8–12.

Lopez-Tiro, F., Estrade, V., Hubert, J., Flores-Araiza, D., Gonzalez-Mendoza, M., Ochoa, G., & Daul, C. (2024). On the In Vivo Recognition of Kidney Stones Using Machine Learning. *IEEE Access*, *12*, 10736–10759. https://doi.org/10.1109/access.2024.3351178

Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift Für Medizinische Physik*, *29*(2), 102–127. https://doi.org/10.1016/j.zemedi.2018.11.002

Maier, A., Syben, C., Lasser, T., & Riess, C. (2019). A gentle introduction to deep learning in medical image processing. *Zeitschrift Für Medizinische Physik*, *29*(2), 86–101. https://doi.org/10.1016/j.zemedi.2018.12.003

Mohammadinejad, P., Mileto, A., Yu, L., Leng, S., Guimaraes, L. S., Missert, A. D., Jensen, C. T., Gong, H., McCollough, C. H., & Fletcher, J. G. (2021). CT Noise-Reduction Methods for Lower-Dose Scanning: Strengths and Weaknesses of Iterative Reconstruction Algorithms and New Techniques. *RadioGraphics*, *41*(5), 1493–1508. https://doi.org/10.1148/rg.2021200196

Purushothaman, K., & Ahmad, R. (2022). Integration of Six Sigma methodology of DMADV steps with QFD, DFMEA and TRIZ applications for image-based automated inspection system development: a case study. *International Journal of Lean Six Sigma*, *13*(6), 1239–1276. https://doi.org/10.1108/ijlss-05-2021-0088

Suzuki, K. (2017). Overview of deep learning in medical imaging. *Radiological Physics and Technology*, *10*(3), 257–273. https://doi.org/10.1007/s12194-017-0406-5

Tao, S., Rajendran, K., Zhou, W., Fletcher, J. G., McCollough, C. H., & Leng, S. (2020). Noise reduction in CT image using prior knowledge aware iterative denoising. *Physics in Medicine & Biology*, *65*(22), 225032. https://doi.org/10.1088/1361-6560/abc231

Teng, S., Liu, A., Wu, Z., Chen, B., Ye, X., Fu, J., Kitiporncha, S., & Yang, J. (2024). Automated detection of underwater cracks based on fusion of optical and texture information. *Engineering Structures*, *315*, 118515. https://doi.org/10.1016/j.engstruct.2024.118515

Wang, S., Cai, B., Wang, W., Li, Z., Hu, W., Yan, B., & Liu, X. (2024). Automated detection of pavement distress based on enhanced YOLOv8 and synthetic data with textured background modeling. *Transportation Geotechnics*, *48*, 101304. https://doi.org/10.1016/j.trgeo.2024.101304

Yildirim, K., Bozdag, P. G., Talo, M., Yildirim, O., Karabatak, M., & Acharya, U. R. (2021). Deep learning model for automated kidney stone detection using coronal CT images. *Computers in Biology and Medicine*, *135*, 104569. https://doi.org/10.1016/j.compbiomed.2021.104569

Yousef, R., Gupta, G., Yousef, N., & Khari, M. (2022). A holistic overview of deep learning approach in medical imaging. *Multimedia Systems*, *28*(3), 881–914. https://doi.org/10.1007/s00530-021-00884-5

Zavala-Mondragon, L. A., de With, P. H. N., & van der Sommen, F. (2021). Image Noise Reduction Based on a Fixed Wavelet Frame and CNNs Applied to CT. *IEEE Transactions on Image Processing*, *30*, 9386–9401. https://doi.org/10.1109/tip.2021.3125489

Zhou, Z., Zhang, J., & Gong, C. (2022). Automatic detection method of tunnel lining multi-defects via an enhanced You Only Look Once network. *Computer-Aided Civil and Infrastructure Engineering*, *37*(6), 762–780. https://doi.org/10.1111/mice.12836