FMDB *Transactions on Sustainable Health Science Letters*

Enhanced Deep Learning Models for Early Detection of Renal Failure from CT Imaging

S. Rubin Bose[1,*], J. Angelin Jeba[2], R. Regin[3], S. Tarunraj[4], Edwin Shalom Soji[5], S. Suman Rajest[6]

[1]Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[2]Department of Electronics and Communication Engineering, S.A. Engineering College, Chennai, Tamil Nadu, India.
[3,4]Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[5]Department of Computer Science, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India.
[6]Department of Research and Development & International Student Affairs, Dhaanish Ahmed College of Engineering, Chennai, Tamil Nadu, India.
rubinbos@srmist.edu.in[1], angelinjeba@saec.ac.in[2], reginr@srmist.edu.in[3], stharunraj2002@gmail.com[4], edwinshalomsoji.cbcs.cs@bharathuniv.ac.in[5], sumanrajest414@gmail.com[6]

Abstract: Chronic kidney disease (CKD) is a progressive condition affecting over 12% of the global population and can lead to end-stage renal disease (ESRD) if not managed effectively. Early prediction of renal failure in CKD patients is crucial for timely intervention, which can slow or prevent disease progression and reduce the need for costly treatments like dialysis or transplantation. However, traditional diagnostic methods, such as serum creatinine tests and nuclear imaging, often lack precision and are invasive. This paper examines the application of optimized deep learning models, particularly those utilizing computed tomography (CT) imaging, for the early detection of renal failure. Key approaches include AI-assisted segmentation of renal-enhanced CT images and advanced models like YOLOv8, which have shown promise in accurately identifying kidney abnormalities and assessing risks. By leveraging these cutting-edge technologies, the goal is to improve early detection, enhance patient outcomes, and reduce healthcare costs, addressing the global burden of CKD. The proposed YOLOv8 model obtained a precision up to 0.986, recall up to 0.969, mAP50 up to 0.989, and mAP50-95 up to 0.972 across BOX and MASK predictions.

Keywords: Chronic Kidney Disease; Kidney Diagnostics; Machine Learning; Renal Segmentation; Non-Invasive Diagnosis; Healthcare Systems; Artificial Intelligence; Medical Imaging; Kidney Health.

Received on: 28/03/2024, **Revised on:** 02/06/2024, **Accepted on:** 27/07/2024, **Published on:** 01/09/2024

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSHSL

DOI: https://doi.org/10.69888/FTSHSL.2024.000251

Cite as: S. R. Bose, J. A. Jeba, R. Regin, S. Tarunraj, E. S. Soji, and S. S. Rajest, "Enhanced Deep Learning Models for Early Detection of Renal Failure from CT Imaging," *FMDB Transactions on Sustainable Health Science Letters*., vol.2, no.3, pp. 152–163, 2024.

1. Introduction

*Corresponding author.

Chronic kidney disease (CKD) is a progressive condition characterized by the gradual loss of kidney function over time, affecting millions of individuals worldwide. The kidneys play a crucial role in maintaining homeostasis by filtering waste products from the blood, regulating fluid and electrolyte balance, and producing hormones vital for various bodily functions. When kidney function declines, it can lead to a myriad of health complications, ultimately culminating in end-stage renal disease (ESRD), which necessitates costly interventions such as dialysis or kidney transplantation. The significance of early prediction and detection of renal failure in CKD patients cannot be overstated, as timely interventions can significantly improve patient outcomes and mitigate the escalating burden on healthcare systems [11]. The World Health Organization (WHO) has recognized CKD as a major global health issue, with prevalence rates estimated to be over 12% in the general population. The condition is particularly concerning as it often remains asymptomatic in its early stages, leading to delayed diagnosis and treatment [12]. Significant kidney damage may have already occurred when symptoms such as fatigue, swelling, and changes in urine output become evident [13]. This progression emphasizes the urgent need for effective early detection strategies that can identify individuals at risk of deteriorating kidney function before severe complications arise [14].

One of the primary challenges in managing CKD is its heterogeneous nature; the disease can stem from various etiologies, including diabetes mellitus, hypertension, glomerulonephritis, and polycystic kidney disease, among others [15]. This variability complicates the diagnostic process, as each underlying cause may require a distinct therapeutic approach. Traditional methods for assessing kidney function primarily involve serum creatinine measurements, which, while widely used, have limitations [16]. Creatinine levels can be influenced by several factors, including muscle mass and diet, leading to potential misinterpretations of renal function. Moreover, these tests are not sensitive enough to detect early declines in kidney function, often resulting in late-stage diagnosis [17]. To address these challenges, recent advancements in medical imaging and artificial intelligence (AI) have emerged as promising avenues for enhancing the early prediction and detection of renal failure in CKD patients. The advent of non-invasive imaging techniques, such as computed tomography (CT) and magnetic resonance imaging (MRI), allows for detailed kidney structure and function visualization [18]. AI-driven analysis of these images can significantly improve diagnostic accuracy by identifying subtle abnormalities that may not be discernible to the human eye [19]. For instance, deep learning algorithms have been developed to automate the segmentation of renal structures in CT images, facilitating the measurement of renal volume and other critical parameters that can provide insights into kidney health [20].

The integration of AI into medical imaging is particularly advantageous due to its capacity to process vast amounts of data quickly and accurately [21]. Machine learning models can be trained on large datasets to recognize patterns associated with renal abnormalities, ultimately leading to enhanced predictive capabilities [22]. For example, algorithms such as YOLOv8 (You Only Look Once) have demonstrated effectiveness in real-time object detection and classification tasks, making them suitable for analyzing medical images [23]. By applying these advanced models to renal imaging, healthcare providers can obtain crucial information regarding the presence of tumours, cysts, or other anomalies that may indicate a decline in kidney function. In addition to improving diagnostic accuracy, AI technologies can facilitate estimating the glomerular filtration rate (GFR), a key indicator of kidney function [24]. GFR quantifies the volume of blood filtered by the kidneys per minute and is traditionally measured using invasive techniques or calculated from serum creatinine levels [25]. Recent studies have shown that AI models can estimate GFR more accurately by analyzing renal volume and other imaging characteristics, thus providing a non-invasive alternative that enhances patient comfort and reduces healthcare costs [26].

The early prediction of renal failure in patients with CKD is vital, as it enables healthcare providers to implement timely interventions that can slow disease progression [27]. For example, lifestyle modifications, such as dietary changes and increased physical activity, can be prescribed based on individual risk profiles identified through advanced imaging and AI analysis [28]. Furthermore, pharmacological treatments can be tailored to target specific underlying causes of CKD, optimizing therapeutic outcomes [29]. By identifying high-risk patients early, healthcare systems can reduce the incidence of ESRD, thereby lessening the financial burden associated with dialysis and transplantation, which are often necessitated when CKD reaches its final stages [30]. Moreover, the emphasis on early detection aligns with the broader goals of public health initiatives to improve population health and reduce healthcare disparities [31]. CKD disproportionately affects certain populations, including older adults and individuals with comorbid conditions such as diabetes and hypertension [32]. Healthcare providers can improve access to timely interventions for vulnerable populations by implementing targeted screening programs and leveraging AI-driven technologies, ultimately enhancing health equity [33].

Despite the promising potential of AI and advanced imaging techniques in the early detection of renal failure, several challenges remain. One significant hurdle is the robust validation of these technologies in diverse clinical settings [34]. While initial studies demonstrate encouraging results, further research is necessary to establish the generalizability and reliability of AI algorithms across different populations and imaging modalities [35]. Additionally, integrating these technologies into clinical workflows poses logistical challenges, requiring collaboration among radiologists, nephrologists, and data scientists to ensure seamless implementation [36]. The early prediction of renal failure and the management of chronic kidney disease represent critical areas of focus in contemporary healthcare. As CKD continues to impact a significant portion of the global population, innovative solutions leveraging medical imaging and artificial intelligence advancements are essential for improving patient outcomes

[37]. By enhancing diagnostic accuracy, enabling timely interventions, and addressing healthcare disparities, these technologies can potentially transform the landscape of kidney disease management [38]. Continued research and collaboration among healthcare professionals, technologists, and policymakers will be vital in realizing the full benefits of these innovations, ultimately paving the way for a future where the burden of chronic kidney disease is significantly reduced.

2. Literature Review

Luo et al. [1] emphasize the urgent need for effective early detection methods for chronic kidney disease (CKD), which affects over 12% of the global population. Despite advancements in renal replacement therapy since the 1960s, which have prolonged the lives of patients with end-stage renal disease (ESRD), the treatment remains prohibitively expensive, with the number of recipients expected to reach 5.4 million by 2030. CKD's inconspicuous symptoms and infrequent physical ex-aminations often lead to late diagnoses. While the estimated glomerular filtration rate (eGFR) is a critical measure of kidney function, its reliance on invasive serum creatinine (SCR) testing and lack of precision hinder early detection. However, advances in medical imaging technology provide promising alternatives, with AI and radiomics significantly enhancing lesion recognition and segmentation. Renal fibrosis, a key factor in CKD progression, can be detected through texture analysis, revealing structural changes in renal tissue. Recent studies demonstrate the potential of MRI-based texture analysis in identifying early-stage renal injury and correlating with Scr levels. This study focuses on developing an AI-based automatic segmentation model utilizing renal-enhanced computed tomography (CT) images for quantitative CKD assessment, facilitating clinicians' efforts in early detection, diagnosis, and treatment of CKD through accurate renal function evaluation.

Herts et al. [2] present a study aiming to develop a model for estimating glomerular filtration rate (GFR) in healthy individuals, such as renal transplant donors, using renal volume measurements derived from multidetector computed tomographic (CT) scans along with serum creatinine levels, height, weight, race, and age. The study compares this model's performance with the Modification of Diet in Renal Disease (MDRD) equation. Conducted as a retrospective study, it included data from 244 individuals who underwent renal donor evaluation over two years. An automated segmentation algorithm was utilized to measure renal parenchymal volume from CT images, while GFR was measured using the urinary clearance of iodine 125 (125I) iothalamate. The model was developed using analysis of covariance with significant variables such as renal volume, age, serum creatinine level, and weight, which strongly correlated with GFR measured using 125I-iothalamate clearance. In contrast, sex, race, and height did not correlate significantly. The resulting model positively correlated with the GFR measured by 125I-iothalamate clearance and outperformed the MDRD equation in all six measurements. Consequently, this renal volume-based model provides a reliable alternative for estimating donor GFR from CT scans, reducing the need for invasive 125I- iothalamate clearance methods.

Sasikaladevi and Revathi [3] approach the early and automatic detection of chronic kidney diseases (CKD) using deep learning techniques applied to CT scan images. The study focuses on diagnosing prevalent kidney conditions such as stones, cysts, and tumours. It utilizes a dataset comprising 12,446 unique CT images, categorized into cysts (3,709 images), normal kidneys (5,077 images), stones (1,377 images), and tumours (2,283 images). The methodology involves extracting deep features from these images and constructing hypergraphs, which are then used in a Hypergraph Convolutional Neural Network (HCNN) for representational learning. The proposed model achieved a superior validation accuracy of 99.71%, outperforming other state-of-the-art algorithms. This robust digital-twin framework for kidney disease diagnosis aids nephrologists in better prognosis of renal abnormalities and demonstrates the potential of deep learning in healthcare diagnostics.

Correa-Medero et al. [4] investigate whether differential kidney function, traditionally assessed through nuclear medical imaging, can be effectively evaluated using contrast-enhanced CT scans combined with deep learning and radiomic features. The study analyzed data from patients who underwent kidney nuclear scanning at Mayo Clinic sites between 2018 and 2022, using CT scans per-formed within three months of the nuclear scans, excluding those who had urological or radio-logical interventions during this period. A segmentation model was employed to analyze both kidneys, extracting 2D and 3D radiomic features to predict differential kidney function. The study utilized 1,159 cases from Arizona and Rochester as the internal dataset and 39 cases from Florida as the external test set. A random forest model using 3D delta radiomics features demonstrated promising results, achieving an area under the curve (AUC) of 0.85 for the internal and 0.81 for the external datasets. The model showed specificity and sensitivity of 0.84 and 0.68 for the internal and 0.70 and 0.65 for the external set. The study concludes that this automated approach can derive important differential kidney function information from CT scans, potentially reducing the need for expensive and radioactive nuclear medicine scans in early-stage assessments.

Pallab et al. [5] emphasize the growing role of machine learning in the healthcare industry, particularly for detecting and predicting kidney abnormalities (KA), which are becoming increasingly common in Bangladesh. This rise in kidney-related health issues is exacerbated by inadequate information and poor lifestyle choices, highlighting an urgent need for effective methods to monitor kidney health. To address this public health concern, the Kidney Abnormality, Monitoring, and Analytics (KAMA) project aims to develop an advanced machine learning system capable of quickly and accurately identifying kidney

conditions and distinguishing between normal and abnormal states. This study divided the dataset into training and testing subsets to ensure precise identification of kidney abnormalities. The models employed, Google Net and a custom-designed Convolutional Neural Network (CNN), delivered the most promising results for this task, demonstrating the potential of machine learning in enhancing kidney health diagnostics.

Wang et al. [6] identify preoperative predictors of aggressive pathology in cT1 solid renal cell carcinoma (RCC) by integrating clinical features with qualitative and quantitative CT parameters and developing a nomogram model for prediction. In a retrospective study of 776 cT1 RCC patients who underwent partial or radical nephrectomy between 2018 and 2022, significant predictors of aggressive pathology included neutrophil-to-lymphocyte ratio, distance to the collecting system, CT necro-sis, tumour margin irregularity, peritumoral neovascularity, and RER-NP. The study utilized four-phase contrast-enhanced CT scans and logistic regression to build a nomogram, demonstrating an area under the curve (AUC) of 0.854 in ROC analysis. This nomogram model effectively predicts aggressive pathology and aids in treatment and follow-up decisions.

Pande and Agarwal [7] emphasize that impaired renal function poses a significant health challenge, especially given the global shortage of nephrologists. This shortage has increased demand for AI-driven systems that can autonomously detect kidney abnormalities. The study introduces a deep learning model, YOLOv8, designed to identify kidney diseases such as stones, cysts, and tumours from a dataset of 12,446 CT images. The dataset, sourced from hospitals in Dhaka, is categorized into four groups: cyst, tumour, stone, and normal. The YOLOv8 model achieved an accuracy of 82.52%, a precision of 85.76%, a recall of 75.28%, an F1 score of 75.72%, and a specificity of 93.12%, outperforming traditional diagnostic methods. The study highlights the model's potential in enhancing the early detection and management of chronic kidney disease, providing valuable support to healthcare providers, and improving patient outcomes.

Zheng et al. [8] present UroAngel, a deep learning-based system specifically developed for the non-invasive prediction of single-kidney function levels in patients with obstructive nephropathy (ON). The system employs a 3D U-Net model to segment renal parenchyma from CT urography images and uses logistic regression for classifying renal function. In a retrospective analysis involving 520 ON patients, UroAngel demonstrated a high accuracy of 0.918 in predicting renal function stages. This performance surpassed the Modification of Diet in Renal Disease (MDRD) and Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equations and two expert radiologists within a validation cohort of 40 ON patients. With a Dice similarity coefficient of 0.861 for renal cortex segmentation, UroAngel offers a reliable and non-invasive approach to assessing single-kidney function, showcasing its potential to improve clinical decision-making in the management of ON.

Zhang et al. [9] outline a deep learning pipeline for kidney disease diagnosis, starting with acquiring large, high-quality medical image datasets from diverse sources to ensure accuracy and reduce bias. The next step is data preprocessing, which involves techniques like noise reduction, normalization, and augmentation to improve model performance and prevent overfitting. Accurate data labelling is crucial for supervised learning tasks. The model development phase includes selecting appropriate architectures, such as U-Net for segmentation or ResNet for classification and optimizing the model through weight adjustments and hyperparameter tuning. During the model evaluation, metrics like accuracy, precision, F1-score, and Dice coefficient are used to assess the model's effectiveness. Deep learning applications in kidney disease range from automated segmentation of renal tumours and structures to differential diagnosis and grading of renal masses, highlighting its potential to improve diagnostic accuracy and treatment planning.

He et al. [10] developed an interactive, non-invasive AI system to predict the malignancy risk in cystic renal lesions (CRLs). Through a multicenter retrospective study involving 715 patients, they introduced a geodesic-based 3D segmentation model for CRL segmentation and a classification model based on the Spatial Encoder Temporal Decoder (SETD) architecture. The classification model utilized a 3D-ResNet50 network to extract spatial features and a gated recurrent unit (GRU) network to decode temporal features from multi-phase CT images. The segmentation model's performance was measured by sensitivity, specificity, intersection over union (IOU), and Dice similarity. In contrast, the classification model was assessed using the area under the receiver operating characteristic curve (AUC), accuracy, and decision curve analysis (DCA). The models demonstrated strong performance in both validation (AUC = 0.973, ACC = 0.916, Dice = 0.847, IOU = 0.743, SEN = 0.840, SPE = 1.000) and testing datasets (AUC = 0.998, ACC = 0.988, Dice = 0.861, IOU = 0.762, SEN = 0.876, SPE = 1.000). This study emphasizes the AI system's capability to accurately distinguish between benign and malignant CRLs, thereby improving clinical decision-making and potentially reducing unnecessary follow-ups and overtreatment.

3. Methodology

The methodology for addressing kidney disease and renal failure leverages advanced medical imaging analysis through the YOLOv8 model, a state-of-the-art real-time object detection system. This model is designed to detect, classify, and accurately localize kidney anomalies within medical images, providing detailed insights into their size, shape, and location. We utilize an extensive dataset of over 2000 meticulously labelled kidney scans, ensuring rigorous training and validation of the model.

Throughout the development process, we rigorously validate the model's accuracy and performance using this dataset, including testing real-world patient data to ensure consistent and precise results [39]. One of the key advantages of the proposed approach is its real-time detection and localization capabilities, which empower healthcare professionals with timely, actionable information. This enables swift and informed decision-making, particularly crucial in critical healthcare scenarios. Figure 1 shows the Functional block diagram of YOLOv8 for Renal Failure Detection.

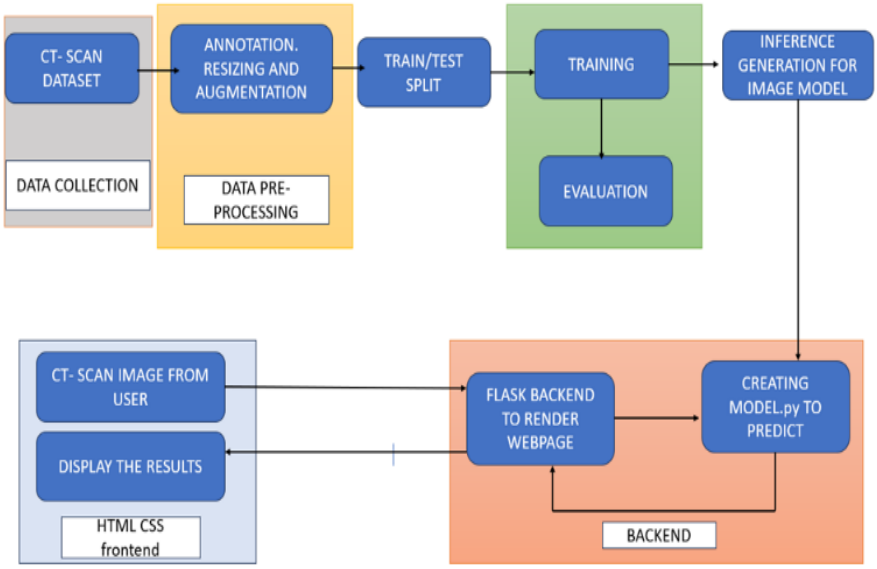


Figure 1: Functional block diagram of YOLOv8 for Renal Failure Detection

The YOLOv8 model is a real-time object detection system for identifying, classifying, and pinpointing kidney anomalies or lesions in medical images. YOLOv8 is known for its ability to accurately detect objects within images, making it well-suited for detailed analysis of kidney scans, including determining the size, shape, and location of anomalies [40]. A comprehensive dataset of over 2382 kidney scan images is utilized to train and validate the proposed model. Many of these samples are sourced from Indian healthcare institutions, ensuring that the model is attuned to the specific characteristics of the local population. Each image in this dataset is meticulously labelled by experts, providing a robust foundation for the model's learning process.

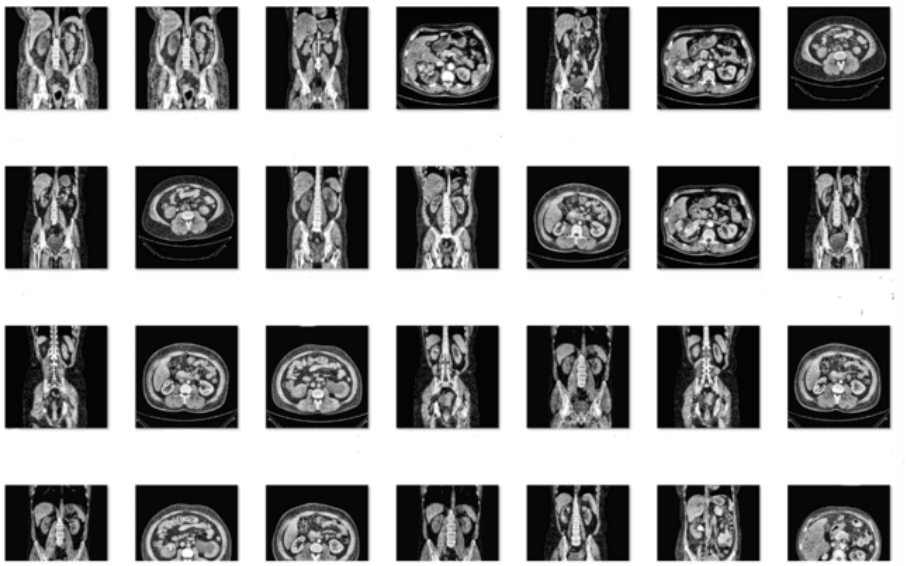


Figure 2: Sample Data from the Dataset

The YOLOv8 model is the central component of our system, processing medical images to detect and localize kidney anomalies. The YOLOv8 model undergoes thorough training and validation with this labelled dataset, enhancing its ability to accurately identify and localize kidney anomalies. The system generates detailed reports that offer healthcare professionals critical insights into the identified anomalies, supporting informed decision-making in diagnosis and treatment. The performance of the model is evaluated using the test data. The YOLOv8 model produces reliable results. A key advantage of this approach is the model's capability for real-time detection and localization of kidney anomalies, which provides healthcare professionals with timely and actionable information crucial for making quick decisions in urgent clinical situations.

A large dataset of 1743 CT scan images of the kidney is collected, annotated for cysts, stones, and tumours and preprocessed and augmented to generate 2382 images. The dataset is then split into the ratio 7:2:1 for valid training and testing. Figure 2 shows the samples from the dataset.

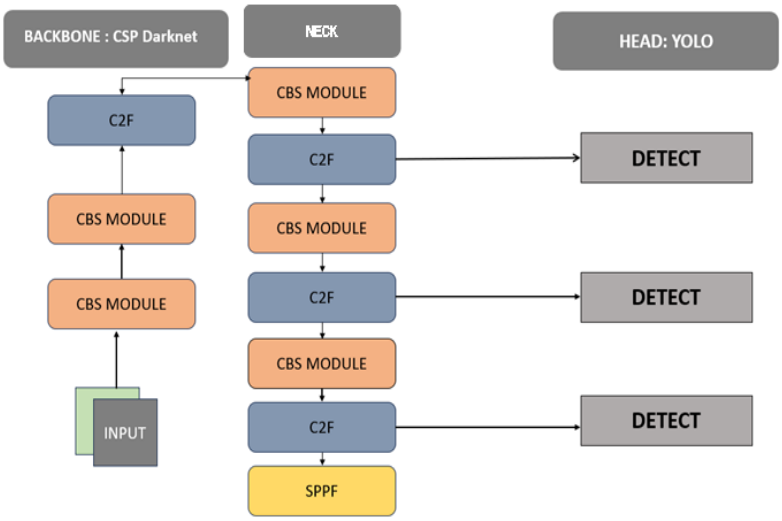


Figure 3: YOLO v8 architecture

Figure 3 shows the YOLOv8 architecture. The YOLOv8 incorporates several key components: the backbone, CSPDarknet53, provides a robust feature extraction framework to enhance the model's object identification and classification capabilities. Utilizing PANet (Path Aggregation Network), the neck improves model performance by aggregating information from different layers. Finally, the head processes the feature maps from the backbone and neck to generate predictions, including bounding boxes, objectness scores, and class probabilities for the objects within the input image. CSP Bottleneck with 2 convolutions and Feature Fusion (C2f) - The C2f module is crucial in improving gradient flow throughout the network. It integrates two parallel branches, enhancing information exchange between layers, which is essential for maintaining the model's performance and accuracy.

Spatial Pyramid Pooling Fusion (SPPF) - The SPPF module is responsible for spatially segmenting the input data into various regions and independently pooling features from each segment. This enables the model to recognize objects of different scales and sizes within the images. YOLOv8 is an object detection model that performs instance segmentation. It localizes and identifies each object instance in an image. The output of the YOLOv8 segmentation model includes several critical components that provide detailed information about the detected objects. First, the model predicts the number of masks with bounding boxes, indicating the number of detected anomalies. For each box, the model provides the coordinates of the x and y values, precisely defining the location of the detected region within the image. Additionally, the output includes the area of the object in square pixels, offering insights into the anomaly's size and the perimeter of the object in pixels, which outlines the shape. The model also generates mask weights, which contribute to the accuracy of the segmentation, along with class confidences that quantify the certainty of the classification for each detected object.

Transfer learning, a valuable technique for quickly retraining a model on fresh data without retraining the entire network, was used to train this model. This provides for quicker training times and uses fewer resources than standard training. The predicted and true values ratio is calculated during the iteration using the loss function. The Total Loss function is expressed in Equation 1 [41].

$$L_{Tot} = L_{cl} + L_{cn} + L_{bo} \quad \ldots\ldots\ldots (1)$$

Where L_T is the total loss, L_{clss} is represented as a classification loss and expressed in equation 2. L_{cnf} is denoted as confidence loss and expressed in equation 3, and L_box is the bounding box loss [41].

$$L_{cl} = \sum_{i=0}^{x^2} X_i^{ob} \sum_{j=0}^{R} [(Pi(e) - \widehat{Pi}\,(e))^2] \ldots\ldots\ldots\ldots (2)$$

$$L_{cnf} = \sum_{i=0}^{x^2} \sum_{j=0}^{R} X_i^{ob} [(Ei - \widehat{Ei})^2] + \beta_{noob} \sum_{i=0}^{x^2} \sum_{j=0}^{R} X_i^{noob} [(Ei - \widehat{Ei})^2] \ldots\ldots\ldots (3)$$

Where Pi(e) is denoted as a probability to be an object. X_i^ob and X_i^noob are represented as the indicator function. Ei is denoted as the objectness [41].

The performance of the YOLOv8 model is determined using various parameters like Precision, Recall, F1-score, and Prediction time. The various parameters are expressed in equations 4, 5, and 6 [41].

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad \ldots. (4)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad \ldots. (5)$$

$$F1 - Score = 2.\frac{Precision\ .Recall}{Precision + Recall} \quad \ldots. (6)$$

4. Results and Discussion

The CT scan dataset is used to train and test the YOLOv8 model. Using AdamW optimizer, a six-teen-batch deep learning-based Renal Cyst, Stone, and Tumor detection model is trained. AdamW has a learning rate of 0.001429 momentum of 0.9 with parameter groups 66 weight (decay=0.0), 77 weights (decay=0.0005), and 76 biases (decay=0.0). The YOLOv58 proposed model is trained using a transfer learning technique from the COCO128 dataset with pre-trained weights. The proposed YOLOv8 model is a lightweight and time-efficient detecting model with high precision. Table 1 lists the performance metrics of the YOLOv8 Model [41].

At Epoch 10, the model demonstrated promising results with high precision, recall, and mAP scores, indicating its early potential for accurate object detection. As training progressed, performance significantly improved at Epoch 50, further refining the model's ability to detect and localize objects accurately. At Epoch 100, it consistently maintained high precision, recall, and mAP scores across various IOU thresholds, showcasing its reliability and robustness. Finally, at Epoch 200, the model exhibited remarkable performance with consistently high precision, recall, and mAP scores, underscoring its accuracy and effectiveness in detecting and localizing objects.

Table 1: Performance metrics of the YOLOv8 Model

Epochs	BOX				MASK			
	Precision	Recall	mAP_{50}	mAP_{50-95}	Precision	Recall	mAP_{50}	mAP_{50-95}
10	0.935	0.886	0.944	0.872	0.935	0.886	0.945	0.845
50	0.981	0.958	0.982	0.955	0.977	0.954	0.981	0.922
100	0.981	0.969	0.989	0.972	0.978	0.967	0.987	0.94
200	0.986	0.962	0.985	0.97	0.986	0.962	0.986	0.945

YOLOv8 offers a comprehensive suite of segmentation models (nano, small, medium, large, and extra-large), catering to various performance requirements. The pre-trained weights from the COCO dataset are used to train the YOLOv8 models. The training and testing phases employ the customized dataset. The AdamW optimization approach is used to improve the training model. With a batch size of 16, the model has been trained for different models of YOLOv8 for 200 epochs. The YOLOv8 model shows high performance across BOX and MASK predictions, with precision reaching up to 0.986, recall up to 0.969, mAP50 up to 0.989, and mAP50-95 up to 0.972 after 200 epochs. Table 2 lists the performance metrics across various YOLOv8 Models.

Table 2: Average prediction speed across YOLOv8 models

Yolo model	BOX				MASK			
	Precision	Recall	mAP_{50}	mAP_{50-95}	Precision	Recall	mAP_{50}	mAP_{50-95}
Nano	0.935	0.886	0.944	0.872	0.935	0.886	0.945	0.845
Small	0.948	0.917	0.961	0.891	0.948	0.917	0.962	0.872
Medium	0.944	0.912	0.955	0.883	0.948	0.915	0.961	0.869
Large	0.932	0.881	0.951	0.876	0.932	0.881	0.951	0.858
Extra large	0.919	00.897	0.949	0.889	0.915	0.899	0.955	0.864

In comparing the five versions of YOLOv8 architecture models, notable variations are observed in their performance metrics and speed. The metrics, including precision, box recall, mAP50, and mAP50-95, provide insights into the model's object detection accuracy, while speed indicates its computational efficiency. Among the configurations, the "Nano" model demonstrates a strong balance between speed and accuracy, making it a compelling choice. It achieves high precision and recall values, indicating an ability to detect objects accurately. Additionally, the mAP50 and mAP50-95 scores are competitive, reflecting robust object recognition capability across a range of IoU thresh-olds. Table 3 Lists the average prediction speed across YOLOv8 Models.

Table 3: Average prediction speed across YOLOv8 models

Yolo model	Preprocess (ms)	Inference (ms)	Post-process (ms)	Total (ms)
Nano	1.8	2.5	2.7	8.0
Small	0.5	6.9	4.6	12.0
Medium	0.4	13.4	6.0	19.8
Large	0.5	20.6	3.3	24.4
Extra large	0.3	33.2	3.4	36.9

Furthermore, the "Nano" model offers a remarkable speed advantage, balancing computational efficiency and accuracy. This model is well-suited for real-time or resource-constrained applications, where speed and accuracy are crucial factors. As a result, the "Nano" model is the preferred choice as it provides a favourable trade-off between speed and accuracy, making it an excellent candidate for a wide range of practical applications. In this study, we evaluated the performance of an object detection model on a dataset of 270 images. The model was trained for 10, 50, 100 and 200 epochs and achieved the following results.

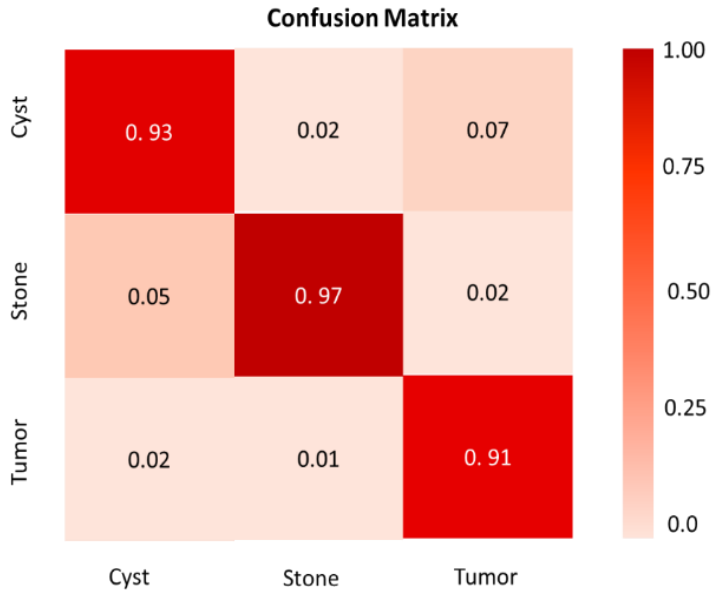
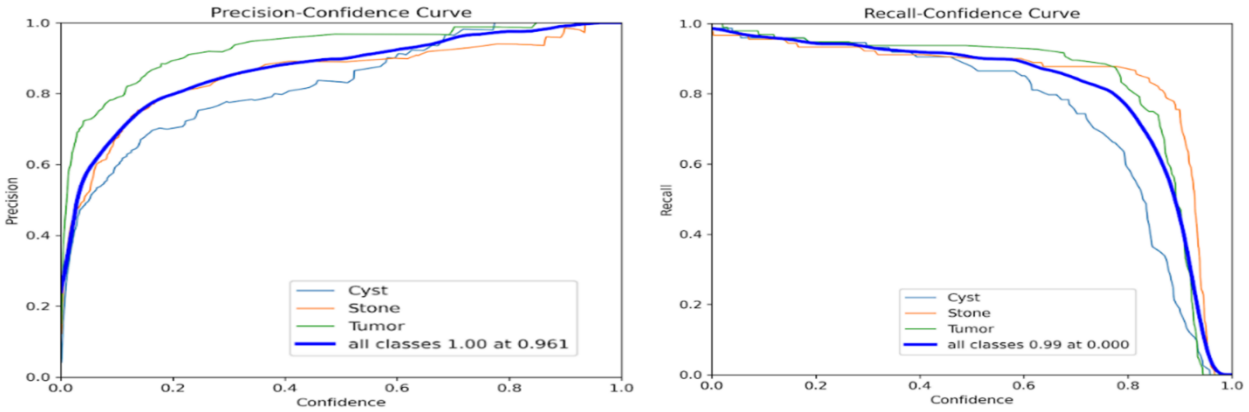


Figure 4: Confusion Matrix of Image Model.

Figure 4 shows the confusion matrix of the Yolov8 Model. The confusion matrix indicates that the model performs well in detecting kidney anomalies, with a high true positive rate across cysts, stones, and tumours. Specifically, it correctly identifies stones with a 0.97 accuracy, cysts with a 0.93 accuracy, and tumours with a 0.91 accuracy. However, there are some misclassifications: cysts are occasionally misclassified as tumours (0.07), and there are minor confusions between cysts and stones (0.02) and between stones and tumours (0.02). Overall, the model shows robust performance, though it could benefit from further refinement to reduce the low but notable misclassification rates, especially between cysts and tumours.



(a) Precision-confidence curve (b) Recall--confidence curve

Figure 5: Performance curve of YOLOv8

Figure 5 Performance curve (a) Precision-confidence curve, (b) Recall--confidence curve. The precision confidence curve for renal failure detection plots precision (the proportion of true positive pre-dictions among all positive predictions) against various thresholds, along with confidence intervals that reflect the reliability of these precision estimates. This curve helps assess how accurately the model makes positive predictions across different thresholds. By incorporating confidence intervals, you gain insight into precision values' consistency and statistical reliability, which is crucial for evaluating how well the model avoids false positives. A curve with narrow confidence intervals around high precision values indicates a robust model performance with reliable detection of true positives.

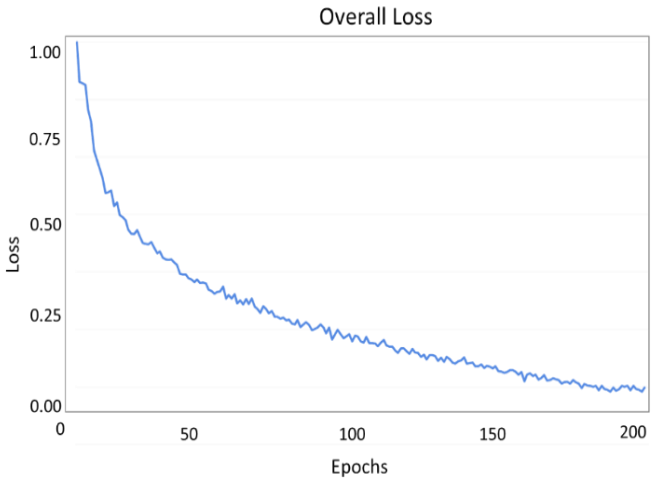
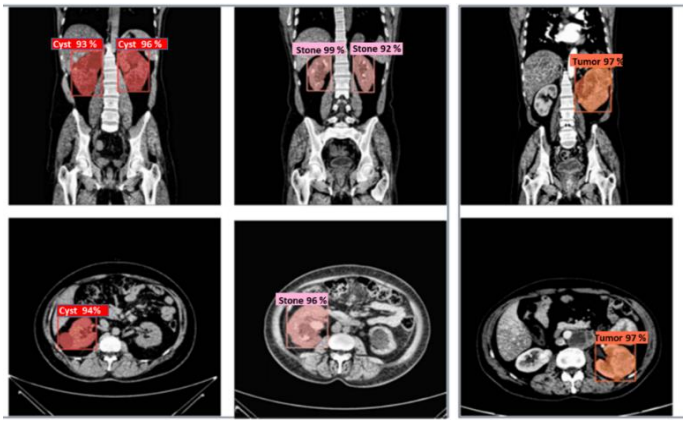


Figure 6: Loss curve of YOLOv8

On the other hand, the recall confidence curve plots recall (the proportion of actual positives correctly identified) against varying thresholds, accompanied by confidence intervals to show the reliability of recall estimates. This curve helps you understand the model's effectiveness in detecting all relevant positive instances across different thresholds. Narrow confidence intervals around high recall values suggest consistent and reliable performance in identifying true positives. Together, these curves allow you to balance precision and recall according to clinical needs, ensuring effective and dependable renal anomaly detection

while guiding optimal threshold selection based on performance reliability. Figure 6 illustrates the loss curve of YOLOv8. The overall loss curve for renal failure detection tracks the model's loss function over training epochs, with the loss value plotted on the y-axis and epochs on the x-axis. A well-behaved curve should show a downward trend, indicating that the model is effectively learning and improving. Consistent decreases in loss suggest good convergence, while erratic fluctuations or constant loss may signal issues such as inadequate training, improper learning rates, or overfitting. Comparing training and validation loss helps assess the model's performance and generalization ability. These results indicate that the model can accurately detect boxes and masks. The model's performance is also relatively consistent across all classes, with no major differences in accuracy.



(a) (b) (c)

Figure 7: Predicted Sample of YOLOv8 Model (a) Cyst, (b) Stone, (c) Tumour

Figure 7 shows the predicted sample of the YOLOv8 Model. The model performs better on mask detection than on box detection. This may be because the model can learn more discriminative features for masks than for boxes. Masks are typically more fine-grained than boxes, and they can provide more information about the shape and appearance of the object being detected. The results suggest that early stopping can be an effective way to prevent overfitting in object detection models. In this study, the model achieved its best performance at 200 epochs. Training the model for longer did not improve the model's performance, leading to a decrease in performance on the validation dataset. This suggests that the model started to overfit the training data after 200 epochs.

5. Conclusion

The YOLOv8 model, trained on a CT scan dataset with the AdamW optimizer, demonstrates robust performance in detecting renal cysts, stones, and tumours. Using a learning rate of 0.001429 and a batch size 16, the model was trained for up to 200 epochs, achieving high precision and recall values, with precision up to 0.986 and recall up to 0.969. The model showed significant improvement across epochs, with the best performance metrics, including mAP50 reaching 0.989 and mAP50-95 up to 0.972, observed at 200 epochs. Among the different YOLOv8 versions, the "Nano" model offered an optimal balance between speed and accuracy, making it suitable for real-time applications. The confusion matrix reveals strong detection accuracy for cysts, stones, and tumours, though minor misclassifications suggest areas for further refinement. Performance curves indicate the model's reliability and effectiveness in detection, while the loss curve demonstrates effective learning and convergence. Overall, YOLOv8 provides a highly effective and efficient solution for renal anomaly detection, with the "Nano" model emerging as a particularly advantageous option for practical use.

Acknowledgement: We sincerely appreciate the invaluable guidance and unwavering support of our mentors, whose contributions have been instrumental in shaping this research.

Data Availability Statement: The datasets utilized in this study can be accessed upon reasonable request from the corresponding author.

Funding Statement: This research and manuscript were conducted and prepared without external financial assistance or funding.

Conflicts of Interest Statement: The authors affirm that no conflicts of interest are associated with this study.

Ethics and Consent Statement: This study adhered to ethical guidelines and received approval from the appropriate institutional review board. Before participation, informed consent was obtained from all individuals involved.

References

1. H. Luo, J. Li, H. Huang, L. Jiao, S. Zheng, Y. Ying, and Q. Li, "AI-based segmentation of renal enhanced CT images for quantitative evaluation of chronic kidney disease," Sci. Rep., vol. 14, no. 7, pp. 1–11, 2024.
2. B. R. Herts, N. Sharma, M. Lieber, M. Freire, D. A. Goldfarb, and E. D. Poggio, "Estimating glomerular filtration rate in kidney donors: A model constructed with renal volume measurements from donor CT scans," Radiology, vol. 252, no. 1, pp. 109–116, 2009.
3. N. Sasikaladevi and A. Revathi, "Digital twin of renal system with CT-radiography for the early diagnosis of chronic kidney diseases," Biomed. Signal Process. Control, vol. 88, no. 2, p. 105632, 2024.
4. R. L. Correa-Medero, J. Jeong, B. Patel, I. Banerjee, and H. Abdul-Muhsin, "Automated analysis of split kidney function from CT scans using deep learning and delta radiomics," J. Endourol., vol. 38, no. 8, pp. 817–823, 2024.
5. N. M. Pallab, M. U. Mojumdar, N. R. Chakraborty, L. Vetrivendan, and S. Akteri, "Machine Learning based Diagnosis of Kidney AbnormalityRecognition on CT Scan Images," in 2024 11th International Conference onComputing for Sustainable Global Development (INDIACom), pp. 1283–1289, IEEE, New Delhi, India, 2024.
6. K. Wang, L. Dong, S. Li, Y. Liu, Y. Niu, and G. Li, "CT features-based preoperative predictors of aggressive pathology for clinical T1 solid renal cell carcinoma and the development of a nomogram model," BMC Cancer, vol. 24, no. 1, pp. 1–13, 2024.
7. S. D. Pande and R. Agarwal, "Multi-class kidney abnormalities detecting novel system through computed tomography," IEEE Access, vol. 12, no. 1, pp. 21147–21155, 2024.
8. Q. Zheng, X. Ni, R. Yang, P. Jiao, J. Wu, S. Yang, Z. Chen, and X. Liu, "UroAngel: A single-kidney function prediction system based on computed tomography urography using deep learning," World J. Urol., vol. 42, no. 4, pp. 1–7, 2024.
9. M. Zhang, Z. Ye, E. Yuan, X. Lv, Y. Zhang, Y. Tan, C. Xia, J. Tang, J. Huang, and Z. Li, "Imaging-based deep learning in kidney diseases: Recent progress and future prospects," Insights Imaging, vol. 15, no. 2, pp. 1–13, 2024.
10. Q.-H. He, J.-J. Feng, L.-C. Wu, Y. Wang, X. Zhang, Q. Jiang, Q.-Y. Zeng, S.-W. Yin, W.-Y. He, F.-J. Lv, and M.-Z. Xiao, "Deep learning system for malignancy risk prediction in cystic renal lesions: A multicenter study," Insights Imaging, vol. 15, no. 5, pp. 1–13, 2024.
11. H. R. Abdulshaheed, H. R. Penubadi, R. Sekhar, J. F. Tawfeq, A. S. Abdulbaq, A. D. Radhi, P. Shah, H. M. Gheni, R. Khatwani, N. Nanda, P. K. Mitra, S. Aanand, and Y. NIU, "Sustainable optimizing WMN performance through meta-heuristic TDMA link scheduling and routing," Herit. Sustain. Dev., vol. 6, no. 1, pp. 111–126, 2024.
12. B. Leena and A. N. Jayanthi, "Brain Tumor Segmentation and Classification-A Review," Annals of the Romanian Society for Cell Biology, vol. 25, no. 4, pp. 11559-11570, 2021.
13. B. Leena, "Deep Learning-Based Convolutional Neural Network with Random Forest Approach for MRI Brain Tumour Segmentation," in System Design for Epidemics Using Machine Learning and Deep Learning, pp. 83-97, 2023.
14. B. Leena, et al., "Effective Calculation of Power, Direction and Angle of Lightning using Wiedemann-Franz Law," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 4, no. 3, p.11, 2014.
15. C. S. Kumar, B. S. Kumar, G. Gnanaguru, V. Jayalakshmi, S. S. Rajest, and B. Senapati, "Augmenting chronic kidney disease diagnosis with support vector machines for improved classifier accuracy," in Advances in Medical Technologies and Clinical Practice, IGI Global, USA, pp. 336–352, 2024.
16. E. Shalom Soji, S. Gnanamalar, N. Arumugam, S. S. Priscila, N. Selvam, and S. S. Rajest, "AI-Driven Computer Vision for Intelligent Home Automation and Surveillance Systems," in Explainable AI Applications for Human Behavior Analysis, IGI Global, USA, pp. 242–257, 2024.
17. G. Gnanaguru, S. S. Priscila, M. Sakthivanitha, S. Radhakrishnan, S. S. Rajest, and S. Singh, "Thorough analysis of deep learning methods for diagnosis of COVID-19 CT images," in Advances in Medical Technologies and Clinical Practice, IGI Global, USA, pp. 46–65, 2024.
18. H. R. Penubadi et al., "Sustainable electronic document security: a comprehensive framework integrating encryption, digital signature, and watermarking algorithms," Herit. Sustain. Dev., vol. 5, no. 2, pp. 391–404, 2023.
19. I. Al Barazanchi, W. Hashim, R. Thabit, R. Sekhar, P. Shah, and H. R. Penubadi, "Secure and Efficient Classification of Trusted and Untrusted Nodes in Wireless Body Area Networks: A Survey of Techniques and Applications", Forthcoming Networks and Sustainability in the AIoT Era, Springer Nature, Switzerland AG, pp. 254-264, 2024.

20. J. Rajendran, L. Raju, and L. Bojaraj, “Analytical assessment of Schottky diodes based on CdS/Si heterostructure: current, capacitance, and conductance analysis using TCAD,” Indian Journal of Physics, vol. 98, no. 8, pp. 2775-2784, 2024.
21. K. Singh, L. M. M. Visuwasam, G. Rajasekaran, R. Regin, S. S. Rajest, and Shynu, “Innovations in skeleton-based movement recognition bridging AI and human kinetics,” in Explainable AI Applications for Human Behavior Analysis, IGI Global, USA, pp. 125–141, 2024.
22. L. Bojaraj and A. Jayanthi, “Automatic Brain Tumor Classification via Lion plus Dragon Fly Algorithm,” Journal of Digital Imaging, vol. 35, no. 5, pp. 1382–1408, 2022.
23. L. Bojaraj and A. Jayanthi, “Hybrid Feature Extraction with Ensemble Classifier for Brain Tumor Classification,” International Journal of Pattern Recognition and Artificial Intelligence, vol. 36, no. 10, p. 2250031, 2022.
24. L. Bojaraj and R. Jaikumar, “Hierarchical Clustering Fuzzy Features Subset Classifier with Ant Colony Optimization for Lung Image Classification,” in Image Processing and Intelligent Computing Systems, Taylor & Francis, United Kingdom, p. 14, 2023.
25. M. G. Hariharan, S. Saranya, P. Velavan, E. S. Soji, S. S. Rajest, and L. Thammareddi, “Utilization of artificial intelligence algorithms for advanced cancer detection in the healthcare domain,” in Advances in Medical Technologies and Clinical Practice, IGI Global, USA, pp. 287–302, 2024.
26. M. Gandhi, C. Satheesh, E. S. Soji, M. Saranya, S. S. Rajest, and S. K. Kothuru, “Image recognition and extraction on computerized vision for sign language decoding,” in Explainable AI Applications for Human Behavior Analysis, IGI Global, USA, pp. 157–173, 2024.
27. M. Kuppam, M. Godbole, T. R. Bammidi, S. S. Rajest, and R. Regin, “Exploring innovative metrics to benchmark and ensure robustness in AI systems,” in Explainable AI Applications for Human Behavior Analysis, IGI Global, USA, pp. 1–17, 2024.
28. N. T. J. Thirumurugan, T. Thirugnanam, L. Bojaraj, and L. R, “Formulation of a two-level electronic security and protection system for malls,” International Journal of Electronic Security and Digital Forensics, vol. 16, no. 1, pp. 63-72, 2024.
29. P. Shweta, L. Bojaraj, et al., “A power efficiency wireless communication networks by early detection of wrong decision probability in handover traffic,” Wireless Communications and Mobile Computing, Vol. 2022, no. 6, pp. 1-7, 2022.
30. R. Boina, “Assessing the Increasing Rate of Parkinson’s Disease in the US and its Prevention Techniques”,” International Journal of Biotechnology Research and Development, vol. 3, no. 1, pp. 1–18, 2022.
31. R. C. A. Komperla, K. S. Pokkuluri, V. K. Nomula, G. U. Gowri, S. S. Rajest, and J. Rahila, “Revolutionizing biometrics with AI-enhanced X-ray and MRI analysis,” in Advances in Medical Technologies and Clinical Practice, IGI Global, USA, pp. 1–16, 2024.
32. S. K. Sehrawat, “Empowering the Patient Journey: The Role of Generative AI in Healthcare,” International Journal of Sustainable Development Through AI, ML and IoT, vol. 2, no. 2, pp. 1-18, 2023.
33. S. K. Sehrawat, “The Role of Artificial Intelligence in ERP Automation: State-of-the-Art and Future Directions,” Transactions on Latest Trends in Artificial Intelligence, vol. 4, no. 4, p.11, 2023.
34. S. K. Sehrawat, “Transforming Clinical Trials: Harnessing the Power of Generative AI for Innovation and Efficiency,” Transactions on Recent Developments in Health Sectors, vol. 6, no. 6, pp. 1-20, 2023.
35. S. Kalimuthu, L. Bojaraj, et al., “Edge Computing and Controller Area Network for IoT data classification using Convolution Neural Network,” in IoT-enabled Convolutional Neural Networks: Techniques and Applications, River Publishers, Denmark, pp. 28, 2023.
36. P. Shah, R. Sekhar, D. Sharma, and H. R. Penubadi, “Fractional order control: A bibliometric analysis (2000–2022),” Results Control Optim., vol. 14, no.3, p. 100366, 2024.
37. A. Shawkat, B. Al-Attar, L. Abd, H. Reddy, R. Sekhar, P. Shah, S. Parihar, S. Kallam, J. Fadhil, and H. Muwafaq, “Efforts of neutrosophic logic in medical image processing and analysis,” Int. J. Neutrosophic Sci., vol. 24, no. 4, pp. 376–388, 2024.
38. T. K. Lakshmi and J. Dheeba, “Classification and Segmentation of Periodontal Cyst for Digital Dental Diagnosis Using Deep Learning,” Computer Assisted Methods in Engineering and Science, vol. 30, no. 2, pp. 131-149, 2023.
39. T. K. Lakshmi and J. Dheeba, “Digital Decision Making in Dentistry: Analysis and Prediction of Periodontitis Using Machine Learning Approach,” International Journal of Next-Generation Computing, vol. 13, no. 3, p.12, 2022.
40. Z. A. Jaaz, M. E. Rusli, N. A. Rahmat, I. Y. Khudhair, I. Al Barazanchi and H. S. Mehdy, “A Review on Energy-Efficient Smart Home Load Forecasting Techniques”, Int. Conf. Electr. Eng. Comput. Sci. Informatics, vol. 202l, no.10, pp. 233-240, 2021.
41. S. Rubin Bose, K. Varun Sharma, V. Karrthik Kishore, S. Tharunraj, G. Nikhil Srinivas and R. Regin, "Vision Based Real-Time Active Protection System Using Deep Convolutional Neural Network," 2023 International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, pp. 1-7, 2023.