공학석사 학위논문

딥러닝 기반의 ^{99m}Tc-diethylenetriamine pentaacetic acid 신장 스캔을 이용한 신사구체 여과율 측정

Deep learning-based prediction of split glomerular filtration rate with ^{99m}Tc-diethylenetriamine pentaacetic acid renal scan

울산대학교대학원

의 과 학 과

박 병 수

Deep learning-based prediction of split glomerular filtration rate with ^{99m}Tc-diethylenetriamine pentaacetic acid renal scan

지도교수 주 세 경 오 정 수

이 논문을 공학석사학위 논문으로 제출함

2024 년 8 월

울 산 대 학 교 대 학 원

의 과 학 과

박 병 수

박병수의 공학석사학위 논문을 인준함

심사위원 김 용 일 

심사위원 한 상 원 

심사위원 오 정 수 

울 산 대 학 교 대 학 원

2024 년 8 월

I. Abstract

Purpose: To automate glomerular filtration rate (GFR) measurement by developing deep learning (DL) models for generating automated regions of interest (ROIs) on ^{99m}Tc-diethylenetriamine pentaacetic acid (^{99m}Tc-DTPA) renal scans and/or for directly regressing the GFR from the ^{99m}Tc-DTPA renal scans.

Methods: Manually-drawn ROIs as well as the corresponding GFR values were retrieved from a Picture Archiving and Communications System were used as ground-truth (GT) (or silver standard) labels and target values, respectively. To this end, we developed two models: one using a two-dimensional U-Net convolutional neural network (CNN) architecture (ROI generator network) with multichannel input to automatically generate kidney and background ROIs, from which GFR was calculated using the Gates formula, and another model using a two-dimensional encoder CNN architecture (GFR regressor network) with multichannel input to directly predict GFR values without the need for ROIs, respectively. The agreement between GFR values from GT and DL ROIs was evaluated using Lin's concordance correlation coefficient (CCC) and slope coefficients for linear regressor analyses. Bias and 95% limits of agreement (LOA) were assessed using Bland-Altman plots.

Results: A total of 24364 scans (12821 patients) were included. Regarding the ROI generator network, we found excellent concordance between GT and DL GFR for left (CCC 0.982, 95% confidence interval [CI] 0.981–0.982; slope 1.004, 95% CI 1.003–1.004), for right (CCC 0.969, 95% CI 0.968–0.969; slope 0.954, 95% CI 0.953–0.955), and for both kidneys (CCC 0.978, 95% CI 0.978–0.979; slope 0.979, 95% CI 0.978–0.979). Bland-Altman analysis revealed minimal bias between GT and DL GFR, with mean differences of −0.2 (95% LOA −4.4–4.0), 1.4 (95% LOA −3.5–6.3) and 1.2 (95% LOA −6.5–8.8) mL/min/1.73 m² for left, right and both kidneys, respectively. Regarding the regressor network, GT and DL estimated GFR values is as follows for the left kidney (CCC 0.969, 95% CI 0.969–0.970), right kidney (CCC 0.969, 95% CI 0.968–0.969), and for both kidneys (CCC 0.972, 95% CI 0.971–0.972). Bland-Altman analysis revealed with mean differences of −0.06 (95% LOA −5.5–5.4), 0.15 (95% LOA −5.4–5.7) and 0.08 (95% LOA −8.9–9.1) mL/min/1.73 m² for left, right and both kidneys. Notably, 19960 scans (81.9%) showed an absolute difference in GFR of less than 5 mL/min/1.73 m² by the ROI generator network. Similarly, the regressor network showed an absolute difference in GFR of less than 5 mL/min/1.73 m² in 18,770 scans (77.0%).

Conclusion: Our ROI generator network and GFR regressor network exhibited excellent performance in the generation of ROIs and estimate GFR on ^{99m}Tc-DTPA renal scans. This automated approach could potentially reduce manual effort and enhance the precision of GFR measurement in clinical practice.

Contents

II. List of Table, Supplementary Tables

III. List of Figures, Supplementary Figures

threshold from the raw image to identify the kidney regions. (d) Heatmap extracted from the final layer of the network. (e) Heatmap from (d) zoomed to a size of 256x256 pixels. (f) Final heatmap obtained by identifying the region overlapping the most with the clustered kidney region in (c) and applying a threshold of 0.1.

Figure 8. Representative images showing a low concordance between GT ROIs and DL ROIs. Each row displays a 2–3-minute summed image (left), the same image with superimposed ROIs (middle) and a corresponding CT image or intravenous pyelogram (right). Panel (a) shows an example where the GT ROI includes a tumor in the upper pole of the left kidney (yellow arrow), whereas the DL ROI does not. Panel (b) shows a case where the GT ROI excludes the tumor in the lower pole of left kidney (yellow arrow), but the DL ROI includes it. Panel (c) presents a left ectopic kidney in the pelvic cavity that was not delineated by the DL ROI (yellow arrow).

Supplementary Figure 1. Scatter plots showing the agreement between GT GFR and DL GFR for left kidney, right kidney and both kidneys across four cross-validation folds. Dashed lines represent lines of equality.

Supplementary Figure 2. Scatter plots showing the agreement between GT GFR and regressor GFR for left kidney, right kidney and both kidneys across four cross-validation folds. Dashed lines represent lines of equality.

Supplementary Figure 3. Bland-Altman plots illustrating the agreement between GT GFR and DL GFR for left kidney, right kidney and both kidneys across four cross-validation folds. Solid lines represent mean differences. Dashed lines represent 95% limits of agreement.

Supplementary Figure 4. Bland-Altman plots illustrating the agreement between GT GFR and regressor GFR for left kidney, right kidney and both kidneys across four cross-validation folds. Solid lines represent mean differences. Dashed lines represent 95% limits of agreement.

1. Introduction

Accurate measurement of glomerular filtration rate (GFR) is crucial for the diagnosis and management of renal diseases, as well as for adjusting medication dosages and monitoring kidney dysfunction across various clinical settings.

^{99m}Tc-diethylenetriamine pentaacetic acid (^{99m}Tc-DTPA) is an extensively utilized radiopharmaceutical for measuring GFR. This compound exhibits minimal protein binding properties and is excreted solely through glomerular filtration without tubular secretion [1]. Split renal GFR is the most important quantitative marker derived from radionuclide renography with ^{99m}Tc-DTPA and is crucial for clinical decision making in many clinical situations, including hydronephrosis, vesicoureteral reflux, renal artery stenosis, tumourous conditions, and renal transplant donor evaluation [2].

Calculation of GFR using a gamma camera often involves the Gates formula, necessitating the establishment of regions of interest (ROIs) for both kidneys and their background area [3, 4]. Because only 20% of renal blood flow is generally filtered by the glomeruli [5], ^{99m}Tc-DTPA images show relatively low signal-to-noise ratio with low renal uptake and high background activity. Therefore, GFR measurement by the Gates method might be sensitive to the ROI settings [6]. Commercially available software provides manual or automated tools for ROI drawing for GFR calculation on ^{99m}Tc-DTPA scans. In cases where the kidneys have irregular shapes or overlap with adjacent organs such as the liver or spleen, automated approaches can face challenges. These issues can limit the routine clinical use of automated methods. On the other hand, manual ROI drawing is considered the most accurate method, though it is often time-consuming and dependent on the operator. Automated approaches, although potentially enhancing convenience and reproducibility, may face limitations in routine clinical use due to these challenges. Manual ROI drawing remains the most accurate method but is often time-consuming and operator-dependent [7].

Artificial intelligence, particularly convolutional neural network (CNN) algorithms, has gained traction in nuclear medicine imaging analysis [8-14]. U-Net-based deep learning (DL) algorithms, in particular, have shown promise in tasks such as tumor delineation on PET/CT scans [15]. With access to large datasets with well-controlled labels, deep learning-based image segmentation holds potential for improving the accuracy and reproducibility of GFR measurement on ^{99m}Tc-DTPA scans, thereby reducing manual effort and time.

The U-Net model (ROI generator network) is a type of CNN specifically designed for biomedical image segmentation [16]. Its architecture consists of a contracting path to capture contexts or features and a symmetric expanding path for precise localization. The contracting path follows the typical architecture of a CNN, involving repeated application of convolutions, followed by max pooling for downsampling. The expanding path consists of upsampling and concatenation with the corresponding feature maps from the contracting path, allowing the model to accurately segment structures within the image.

In the context of ^{99m}Tc-DTPA scans, the ROI generator network is trained to automatically generate ROIs for the kidneys. This involves using large datasets of labeled scans to teach the model how to identify and delineate kidney boundaries accurately. By automating the ROI generation process, the ROI generator network reduces the need for manual intervention, which can be time-consuming and prone to variability.

The GFR regressor network [18], on the other hand, is used to predict the GFR directly from the input ^{99m}Tc-DTPA scans without the need for precise ROI delineation. This model is typically a type of regression-based CNN that learns to map the input image data to a continuous output representing the GFR value. The training process involves using labeled data where the GFR values are known, allowing the model to learn the relationship between the image features and the GFR. This approach can be beneficial in clinical settings as it simplifies the workflow and potentially offers quick and reliable GFR measurements. In addition to this, our team has explored two methodologies for building the GFR regressor network. One approach uses a scratch basis by taking only the encoder structure from the ROI generator network, which allows the model to learn directly from the raw data without pre-learned weights. The other approach utilizes a pretrained encoder from the ROI generator network, incorporating this pretrained component into the GFR regressor network's fully connected layers for training. This method leverages previously acquired knowledge, potentially accelerating the learning process and improving performance.

By conducting both approaches, i.e., the ROI generator network for ROI segmentation and the GFR regressor network for direct GFR estimation, the study aims to enhance the accuracy and efficiency of GFR measurement in clinical practice. The ROI generator network ensures precise and accurate ROI delineation, which is crucial for accurate GFR calculation using traditional methods, while the GFR regressor network offers a more streamlined approach by directly estimating GFR from the scans. These models, trained with extensive datasets

of ^{99m}Tc-DTPA scans with labelled data retrieved from our Picture Archiving and Communication System (PACS), aims to streamline and enhance the GFR measurement process in clinical practice.

Hence, we propose the utilization of ROI generator network for the automatic generation of kidney ROIs. This model, trained with extensive datasets of ^{99m}Tc-DTPA scans with labelled ROI data retrieved from our Picture Archiving and Communication System (PACS), aims to streamline and enhance the GFR measurement process in clinical practice. In addition to the ROI generator network, we developed a GFR regressor network that predicts GFR directly. To calculate kidney depth, which is essential for accurate GFR estimation, this regressor network concatenates age, height, and weight within the network. This dual approach leverages the strengths of both networks: the U-Net for its explainable ROI generation and the regressor network for its direct GFR prediction, offering a comprehensive solution for improved kidney function assessment.

2. Materials and methods

2.1 Patients

This study was conducted at a single center (Asan Medical Center) and included data from patients who underwent a ^{99m}Tc-DTPA renal scan for GFR measurement at our institution from October 2009 to November 2021. Scans were excluded if they met any of the following criteria: 1) presence of a single kidney (such as post-nephrectomy or transplant kidney); 2) severely decreased split renal function (relative function < 20%) [19]; 3) inconsistency in the number of frames (not 80); and 4) patient age ≤ 15 years [20]. The study adhered to the Declaration of Helsinki and institutional guidelines, with approval from the local institutional review board (IRB No. 2022-0333). Informed consent was waived due to the retrospective nature of the study.

Electronic medical records and ^{99m}Tc-DTPA renal scans were reviewed retrospectively by experienced nuclear medicine physician (S.H., S.H.), and serum creatinine, blood urea nitrogen (BUN) level, and GFR estimated by the CKD-EPI equation [21] within 1 month of the scan were collected.

2.2 ^{99m}Tc-DTPA scan acquisition and report

As outlined in our previous study, quantitative ^{99m}Tc-DTPA renal scans were conducted following specific protocols [22]. Patients were instructed to consume 300–500 mL of water 30–60 minutes before the scan, with intravenous hydration provided if necessary. NSAIDs were discontinued 2–3 days prior to the scan, and patients were instructed to void just before the procedure. Intravenous administration of 185 MBq of ^{99m}Tc-DTPA was performed, with imaging conducted using dual-head gamma cameras equipped with either low-energy high-resolution or all-purpose collimators. Dynamic imaging was performed over 20 minutes, resulting in 80 frames, with the injection site included in the scan field for monitoring. Radioactivity of the syringe pre- and post-injection was measured to determine the exact injection dose, factoring in decay correction.

Images were analyzed prospectively by expert nuclear medicine radiologic technologists using Syngo workstation (Siemens Healthineers, Forchheim, Germany). Manual delineation of ROIs for each kidney was conducted, excluding any present tumors. The perirenal background ROI was placed inferolaterally to the kidney center, ensuring adequate distance from neighboring organs to minimize scatter activity.

GFR was calculated based on renal uptake during a 2–3-minute interval (frames 9–12) using the gamma camera-based Gates method [3, 4]. Kidney depth calculation employed the Tonnesen formula [23] until July 2014, transitioning to the Taylor formula [24] thereafter, with manual calculation for patients with renal anomalies [25]. For patients with renal anomalies or variations in kidney position, kidney depth was manually calculated on CT images. Final structured reports were reviewed and approved by experienced nuclear medicine physicians (S.H., S.Y.C., and D.H.M.) before uploading to the in-house PACS system known as PetaVision.

2.3 Parsing PACS data

Data sets comprising structured reports (1132 × 860 matrix) and 80 raw frame images were obtained in the Digital Imaging and Communications in Medicine (DICOM) format. Internal software was employed for automatic anonymization of the images. Various parameters such as kidney and background ROIs, time-activity curves, patient demographics (age, height, weight), split kidney functions, kidney depth, and GFR values were extracted from the structured reports using character recognition techniques [26]. To elaborate on the method for extracting digits, we extracted each digit (0–9) from the values presented in the report format. We then applied a matched filter approach, utilizing patches from the areas where the digits were located. The correlation between these patches and the extracted digits (0–9) was assessed to verify the validity of the digit extraction process. Each kidney depth, depth corrected kidney counts, GFR is calculated using the following formula:

Tonnesen kidney depth [cm]

$$\text{Left kidney depth} = 13.2 * \left(\frac{Weight}{Height}\right) + 0.7$$

$$\text{Right kidney depth} = 13.3 * \left(\frac{Weight}{Height}\right) + 0.7$$

Taylor kidney depth [cm]

$$\text{Left kidney depth} = 15.31 * \left(\frac{Weight}{Height}\right) + 0.22 * Age + 0.077$$

$$\text{Right kidney depth} = 16.17 * \left(\frac{Weight}{Height}\right) + 0.027 * Age - 0.94$$

Left/Right kidney counts [cpm]

$$\text{Depth– corrected Left kidney counts} = \frac{Left\ kidney\ count - Left\ background\ count}{exp(-0.153 * Left\ kidney\ depth)}$$

$$\text{Depth– corrected Right kidney counts} = \frac{Right\ kidney\ count - Right\ background\ count}{exp(-0.153 * Right\ kidney\ depth)}$$

Total kidney uptake [%]

$$\text{Total kidney uptake} = \frac{Depth\ corrected\ kidney\ count}{pre-injection\ count - post-injection\ count} * 100$$

Total/Left/Right GFR [ml/min]

$$\text{Total GFR} = 9.8127[\text{ml/min} * \%]\ \text{Total Kidney uptake} - 6.82519$$

$$\text{Left GFR} = Total\ GFR * \frac{Left\ kidney\ counts}{Total\ Kidney\ counts}$$

$$\text{Right GFR} = Total\ GFR * \frac{Right\ kidney\ counts}{Total\ Kidney\ counts}$$

Normalized GFR [ml/min/1.73m^2]

$$\text{Normalized GFR} = \frac{Total\ GFR}{1.73m^2}$$

Extraction of manual ROIs, which were routinely traced by nuclear medicine technologists, for each kidney and background was performed using color-based edge detection methods. This served as the ground-truth (GT) label for the following DL processing, where manual ROIs defined by nuclear medicine technologists may suffer from considerable inter-rater variability, in turn, we can consider this as “silver standard”. These ROIs were then superimposed onto the 2–3-minute summed images to estimate GFR (Figure 1). Validation of the extraction process was conducted by comparing kidney counts from the structured reports with those obtained

through extraction. The report format depicted ROIs as contours (red/green for kidneys, yellow/blue for backgrounds) rather than filled areas. Typically, these contours formed a closed circuit, providing contiguous pixel connections. In cases of open contours, we applied morphological operations in Python, including binary closing and hole filling, to ensure the integrity of the ROIs. Scaling of ROI dimensions from the report format image space (192 x 192 or 166 x 166) to the raw data space (64 x 64) was performed before the overlay onto raw data.

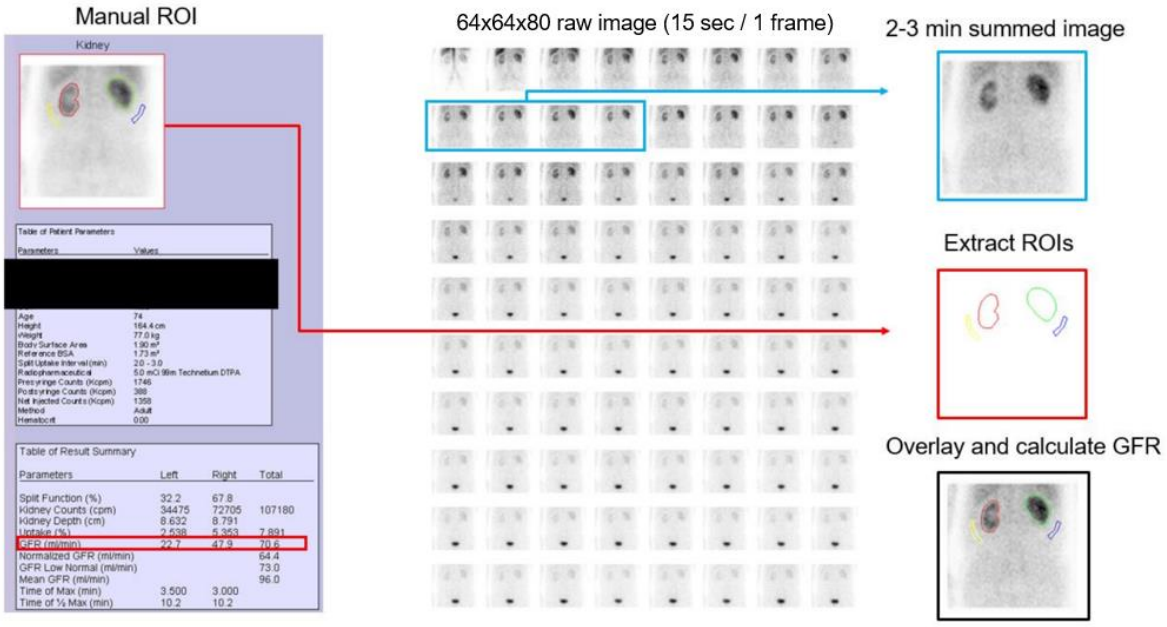


Figure 1

2.4 Deep CNN-based ROI generator network and GRF regressor network: Network Architecture and Training/Test

Schematic diagrams illustrating the dataset setup and the architecture of both DL models are available in Figure 2. GT labels were established using the ROIs extracted from each kidney and the background. A dataset comprising images summed over 2–3 minutes with a resolution of 256 × 256 pixels, along with their corresponding GT ROIs, was randomly split into training (75%) and testing (25%) sets. To ensure model effectiveness and prevent overfitting, a four-fold cross-validation approach was implemented, creating four distinct test groups, each containing 25% of the data. [11, 27]

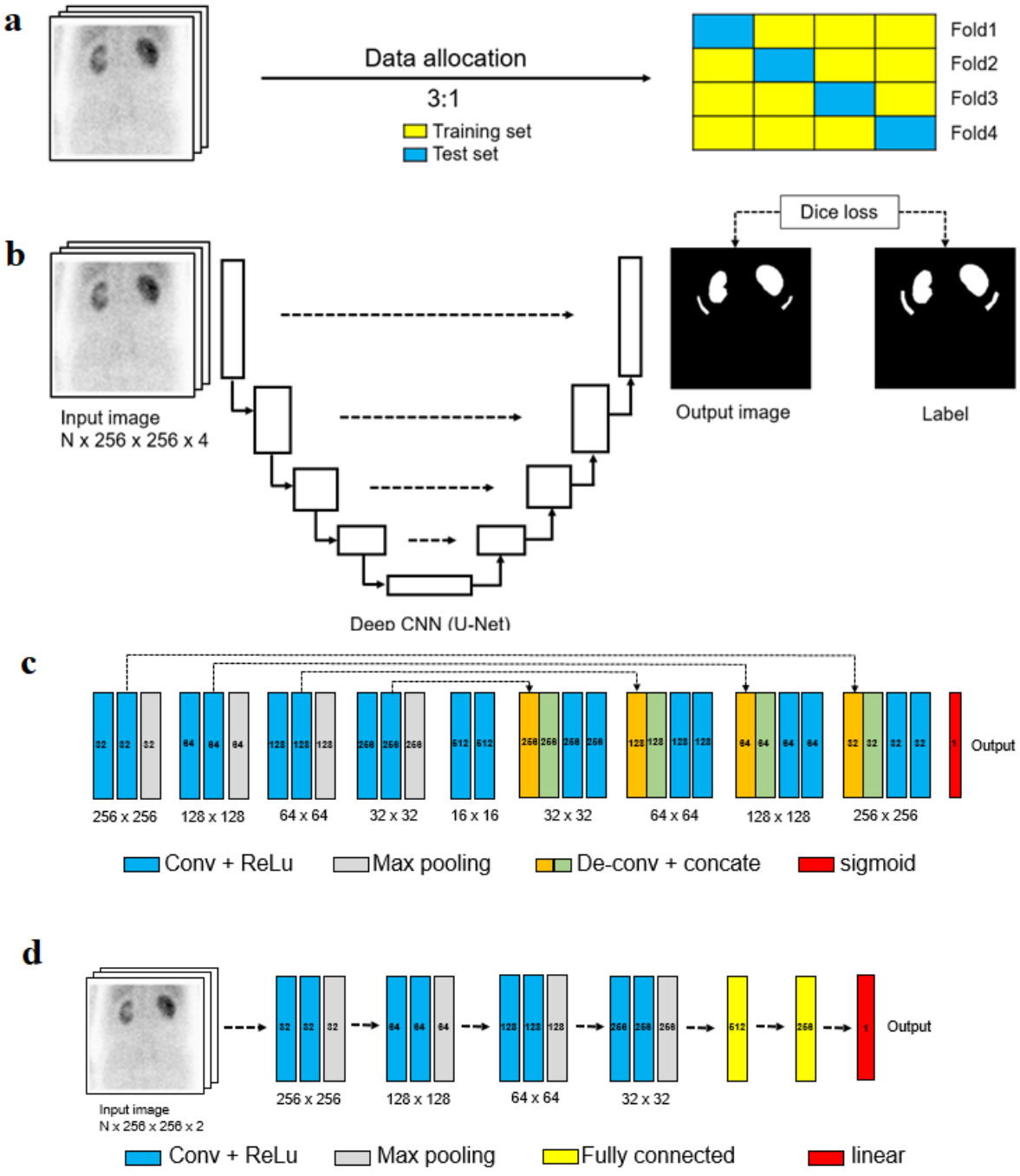


Figure 2

ROI generator network

For the ROI generator network, a 2D ROI generator network architecture was developed to accurately replicate the GT ROIs. This model employed a continuous CNN structure alternating between contracting and expanding paths, integrating residual learning through skip connections to generate segmented images. To enhance model generalization, data augmentation techniques including random translations, rotations, scaling, and shearing were applied to both input data and labels. The Dice similarity coefficient (DSC) served as the loss function, optimized using the adaptive moment estimation algorithm. Rectified linear unit activation functions were utilized in convolutional layers, while a sigmoid function was employed as the final activation function to produce binary ROI outputs, and training began with an initial learning rate of 1e-5. Following deep-learning training, a post-processing step was implemented to refine the generated masks, employing a 2D-connected component analysis-based approach to eliminate islands and fill holes.

$$Dice\ Similarity\ Coefficient = \frac{2|P \cap G|}{|P| + |G|}$$

$$Loss\ function = 1 - Dice\ Similarity\ Coefficient$$

Following ROI generation by DL, a post-processing step was implemented to refine the generated masks, employing a 2D-connected component analysis-based approach to eliminate islands and fill holes. Since variability was observed in the spatial and morphological characteristics of the GT background ROIs, we developed an automated algorithm to generate background ROIs in cases where DL-based background ROIs were not generated. This algorithm performed morphological operations such as dilation, subtraction, and angle selection on each 2D kidney mask to create pie-shaped background ROIs positioned at the 5 and 7 o'clock directions for the right and left kidneys, respectively.

GFR regressor network

Our team has developed a deep CNN architecture to accurately estimate the GFR from ^{99m}Tc-DTPA renal scans. In our approach, we utilized different regressor networks to enhance the accuracy and interpretability of GFR estimation. The first model, a scratch-based GFR regressor network, leverages an encoder structure derived from the ROI generator network.

The second approach involves using a pretrained encoder from the ROI generator network to construct the GFR regressor network. This pretrained model has already learned from a broader dataset, capturing generalized features that are useful for interpreting the renal images. By integrating this pretrained encoder into our GFR regressor network, we benefit from enhanced feature extraction capabilities, which can lead to more accurate and reliable GFR estimations.

Specifically, by combining input data such as age, height, and weight with the features extracted by the encoder, the model can more accurately calculate kidney depth and enhance GFR predictions. We use Mean Squared Error (MSE) as the loss function to minimize the difference between predicted and actual GFR values, ensuring that the model's predictions are as accurate as possible. The final layer of our GFR regressor network employs a linear activation function, providing a continuous range of GFR outputs, which is essential for precise clinical evaluations.

$$Mean\ Squared\ Error = \frac{1}{\mathrm{n}}\sum_{i=0}^{n}(y_i - \hat{y}_i)^2$$

Utilizing various regressor networks, including those with pretrained weighted encoders and models trained from scratch, represents a significant advancement in applying deep learning techniques to medical imaging. This approach allows us to leverage the strengths of different model architectures, thereby providing a more robust and effective tool for clinical assessments. By comparing these different regressor networks, we can enhance GFR estimation accuracy and explore a wider range of potential applications for our models in the medical field.

To determine which parts of the input the GFR regressor network focuses on to produce results, we employed grad-CAM to generate heatmaps. Given that the GFR regressor network lacks a parameter to differentiate between the left and right kidneys, we used masked inputs for training and evaluation. Heatmaps were generated from the final layer of the network, and to assess their quality, we resized them to 256x256 pixels to match the size of the labeled images. Next, we applied K-means clustering to the masked input images and used a threshold of 60% of the maximum value to approximately identify the kidney regions. We calculated the overlap between each cluster's area and the identified kidney region, selecting the best cluster with the greatest overlap to define the final heatmap. Finally, we evaluated the heatmap against the label ROI using the Dice coefficient.

All DL models were implemented by tensorflow-based Python scripts (version 2.6.0) on a system equipped with a GeForce NVIDIA RTX 3090 GPU and an AMD Ryzen Threadrippers 3960X CPU.

Since our study has weaker notion of GT (or silver standard), the performance of all DL models were evaluated using the Concordance Correlation Coefficient (CCC) of GFR values generated by DL models versus that of GT.

2.5 Statistical analysis

Continuous data were reported as mean ± standard deviation or median with interquartile range (IQR). CCC and slope coefficients from linear regressor analyses were utilized to assess agreements between: 1) kidney counts on the standard PACS report form and those extracted from the GT ROIs (to validate ROI extraction); 2) GFR values derived from GT ROIs versus those from DL ROIs (to evaluate DL model performance); and 3) GFR values derived from GT ROIs versus those directly predicted from the DL regressor. Bland-Altman plots were employed for visual assessment of potential bias and 95% limits of agreement (LOA) between these variables. Statistical analyses were conducted using R (version 4.1.1; R Foundation for Statistical Computing, Vienna, Austria) and Python (version 3.8; Python Software Foundation, Wilmington, DE).

3. Results

3.1 Baseline characteristics

Out of 29,550 scans conducted during the study period, 3,334 were excluded due to single kidney presence, 1,653 due to severely decreased split renal function, 178 due to a different number of slices (not 80), and 21 due to patient age ≤ 15 years. Consequently, 24,364 scans from 12,822 patients were included in the DL process and analysis. The number of scans performed using each gamma camera is summarized in Supplementary Table 1. ^{99m}Tc-DTPA scan sessions per patient are summarized in Supplementary Table 2.

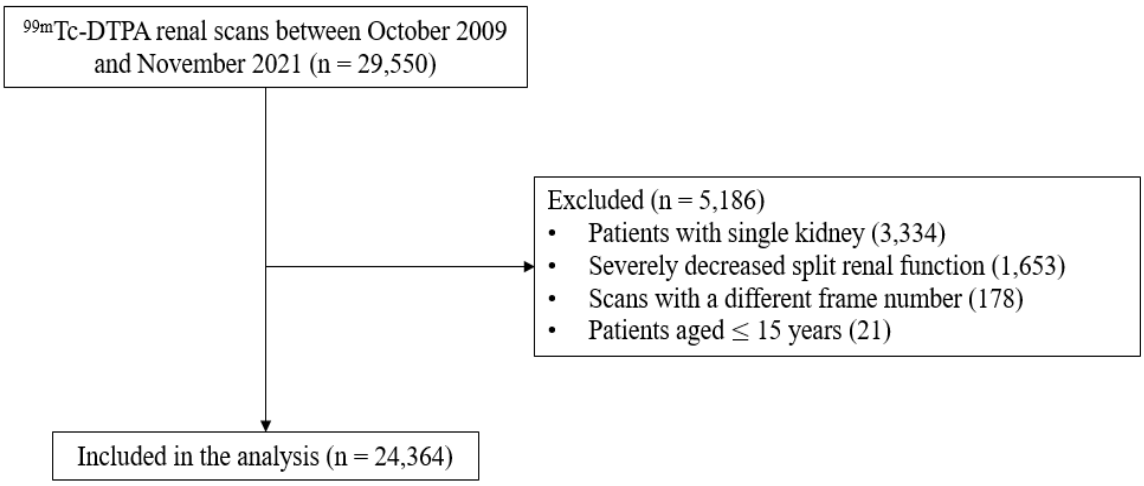


Figure 3

The median patient age was 57.0 years (IQR, 47.0–65.0), median creatinine level was 0.87 (IQR, 0.73–1.01), median BUN level was 14.0 (IQR, 12.0–18.0), and median GFR estimated by the CKD-EPI equation was 91.0 (IQR, 77.0–101.0). A total of 9,835 ^{99m}Tc-DTPA scans were performed on patients following partial nephrectomy. Serum creatinine, BUN, and estimated GFR (by the CKD-EPI equation) levels within one month from the date of ^{99m}Tc-DTPA scan were available in 23,260, 12,777, and 17,284 scans, with median time intervals of 0 (IQR 0–2), 0 (IQR 0–1), and 0 (IQR 0–5) days, respectively. These baseline characteristics of the patients are summarized in Table 1.

3.2 ROI extraction

The kidney and background ROIs extracted from the PACS standard report forms demonstrated excellent agreement with left kidney counts (CCC 0.992, 95% CI 0.992–0.993; slope 0.991, 95% CI 0.990–0.991) and right kidney counts (CCC 0.991, 95% CI 0.99–0.991; slope 1.015, 95% CI 1.014–1.016) calculated from the extracted GT ROIs (Figure 4).

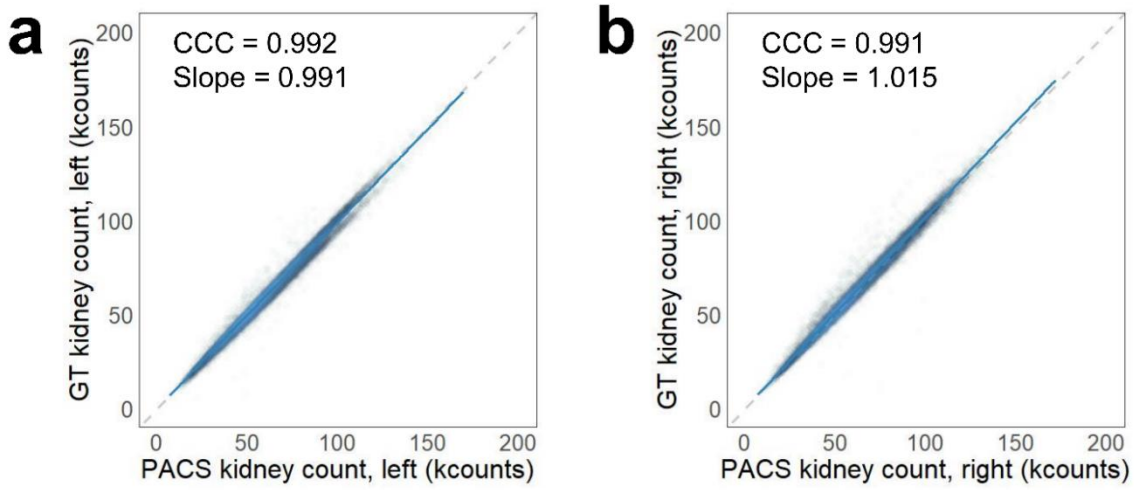


Figure 4

3.3 ROI generation-based GFR calculation and direct GFR prediction

ROI generator Network

The mean DSCs across all test sets were 0.928 (95% CI: 0.888–0.968), 0.927 (95% CI: 0.884–0.970), and 0.565 (95% CI: 0.563–0.567) for the left kidney, right kidney, and background ROIs, respectively. The DL model failed to generate a background ROI in 428 scans, and an automatic algorithm-based background ROIs were applied in these cases. The study demonstrates a remarkable concordance between GFRs derived from GT and DL ROIs when utilizing the ROI generator network architecture. Specifically, the CCC for the left kidney was 0.982 (95% CI 0.981–0.982) with a slope of 1.004 (95% CI 1.003–1.004). For the right kidney, the CCC was 0.969 (95% CI 0.968–0.969) with a slope of 0.954 (95% CI 0.953–0.955), and for both kidneys combined, the CCC was 0.978 (95% CI 0.978–0.979) with a slope of 0.979 (95% CI 0.978–0.979). (Figure 5) The Bland-

Altman analysis revealed mean differences between GFR from GT and DL ROIs of -0.2 (95% LOA: -4.6 to 4.2), 1.4 (95% LOA: -2.1 to 4.9), and 1.2 (95% LOA: -5.2 to 7.6) mL/min/1.73 m^2 for the left, right, and both kidneys. For 19,960 scans (81.9%) (Figure 6), the absolute difference between GFR from GT and DL ROIs in both kidneys was less than 5 mL/min/1.73 m^2. The total GFR values showed a percent difference of less than 10% between the DL ROI and GT ROI methods in 22,023 (90.4%) cases and exhibited a percent difference of less than 20% in 24,133 (99.1%) cases.

GFR regressor Network

The results from the scratch-based GFR regressor network exhibited a high level of concordance, similar to the pretrained encoder-based model. Specifically, the CCC for the left kidney was 0.969 (95% CI 0.969–0.970), for the right kidney it was 0.969 (95% CI 0.968–0.969), and for both kidneys combined, it was 0.972 (95% CI 0.971–0.972). The Bland-Altman analysis for the GFR regressor network revealed mean differences of -0.06 (95% LOA −5.5–5.4), -0.15 (95% LOA −5.4–5.7), and -0.08 (95% LOA −8.9–9.1) mL/min/1.73 m² for the left, right, and both kidneys, respectively. 18,770 (77.0%) scans showed a difference of less than 5 mL/min/1.73 m^2, 20,580 (84.4%) cases had a percent difference within 10%, and 22,424 (98.0%) cases had a percent difference within 20%. These results are comparable to those obtained from the pretrained encoder-based model, indicating that both approaches are effective in estimating GFR values accurately. This similarity underscores the robustness of our models in handling the complexities of renal function assessment through imaging.

Among the various methods attempted to accurately define the heatmap, applying K-Means clustering and a 60% threshold of the maximum value from the raw image to identify the kidney region yielded the best results in terms of alignment with the GT kidney ROI. This method achieved a Dice coefficient of 0.796 for the left kidney and 0.744 for the right kidney. These results indicate that this approach effectively highlights the regions of interest that the model focuses on, significantly improving interpretability. Figure 7 illustrates this method, showcasing its superior performance compared to previous attempts and demonstrating its utility in enhancing clinical relevance by accurately identifying anatomical regions of interest.

The heatmap generated by the ROI generator network effectively highlights the ROI, providing a clear understanding of the areas the model focuses on. This enhanced visualization is crucial for clinical

interpretation, assisting medical professionals in identifying key areas of concern. In comparison, while the GFR regressor network also generates heatmaps that generally align with the GT kidney ROI, they are significantly less clear and less intuitive than those produced by the ROI generator network. This suggests that the ROI generator network provides more discernible and actionable insights, which are essential for clinical decision-making. Despite achieving accurate GFR estimates, the heatmaps from the GFR regressor network did not consistently match the expected anatomical regions of interest, highlighting the superior interpretability of the ROI generator network approach.

Representative cases where ROI generator network failed to draw an ROI or showed low concordance compared with GT are presented in Figure 8. Upon visual review, poor renal uptake, severe hydronephrosis, large renal tumor, or an ectopic kidney was often accompanied in the cases where kidney ROI was not drawn or had a low concordance.

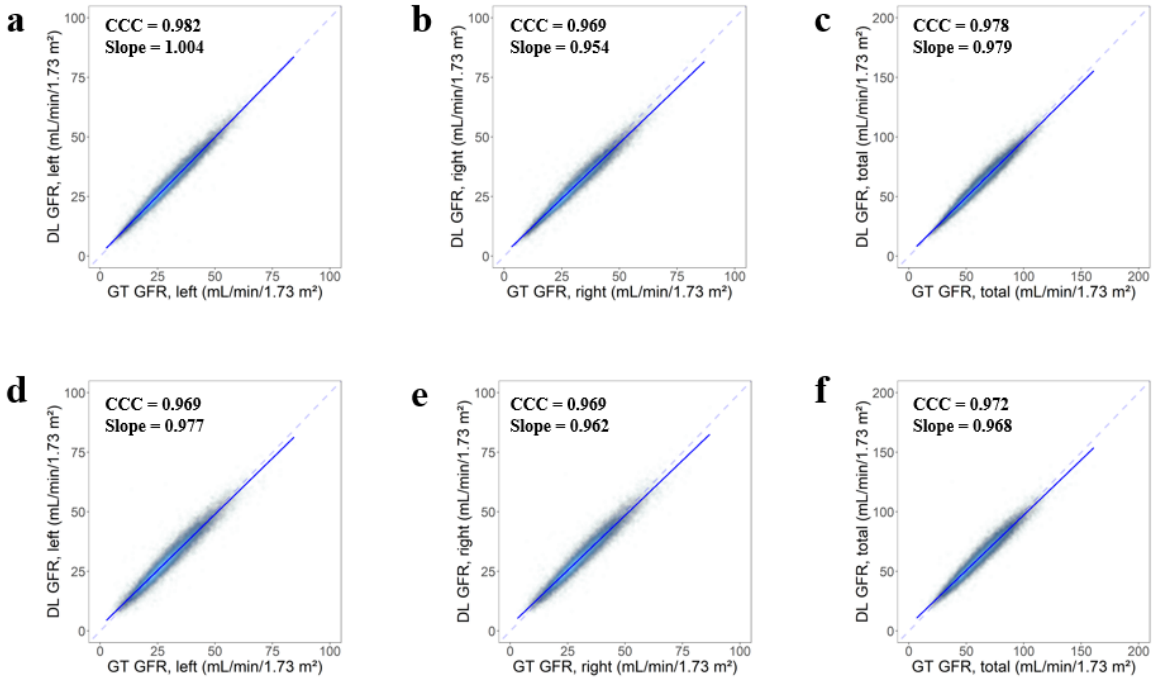


Figure 5

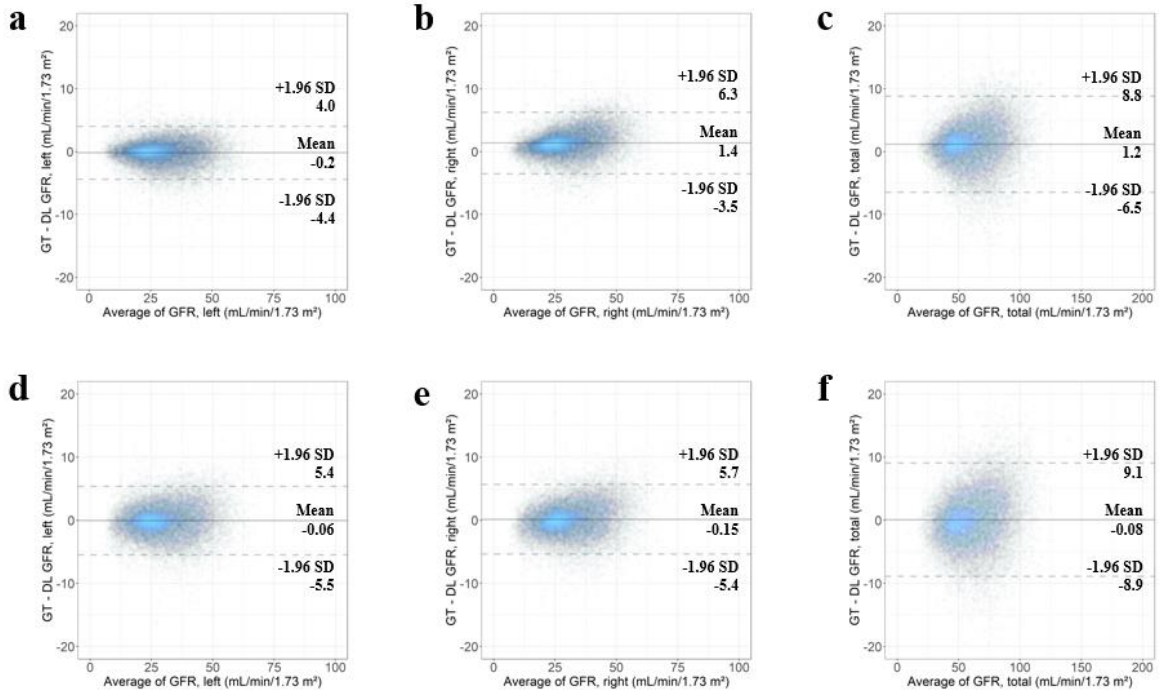


Figure 6

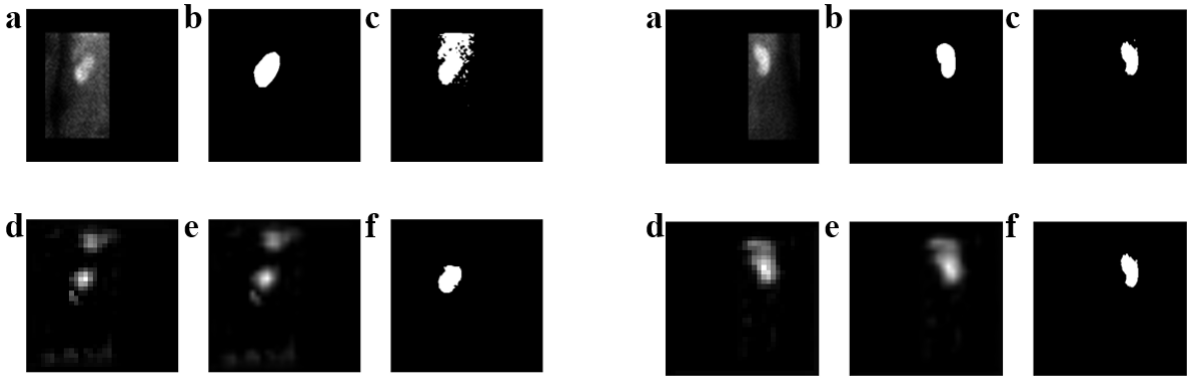


Figure 7

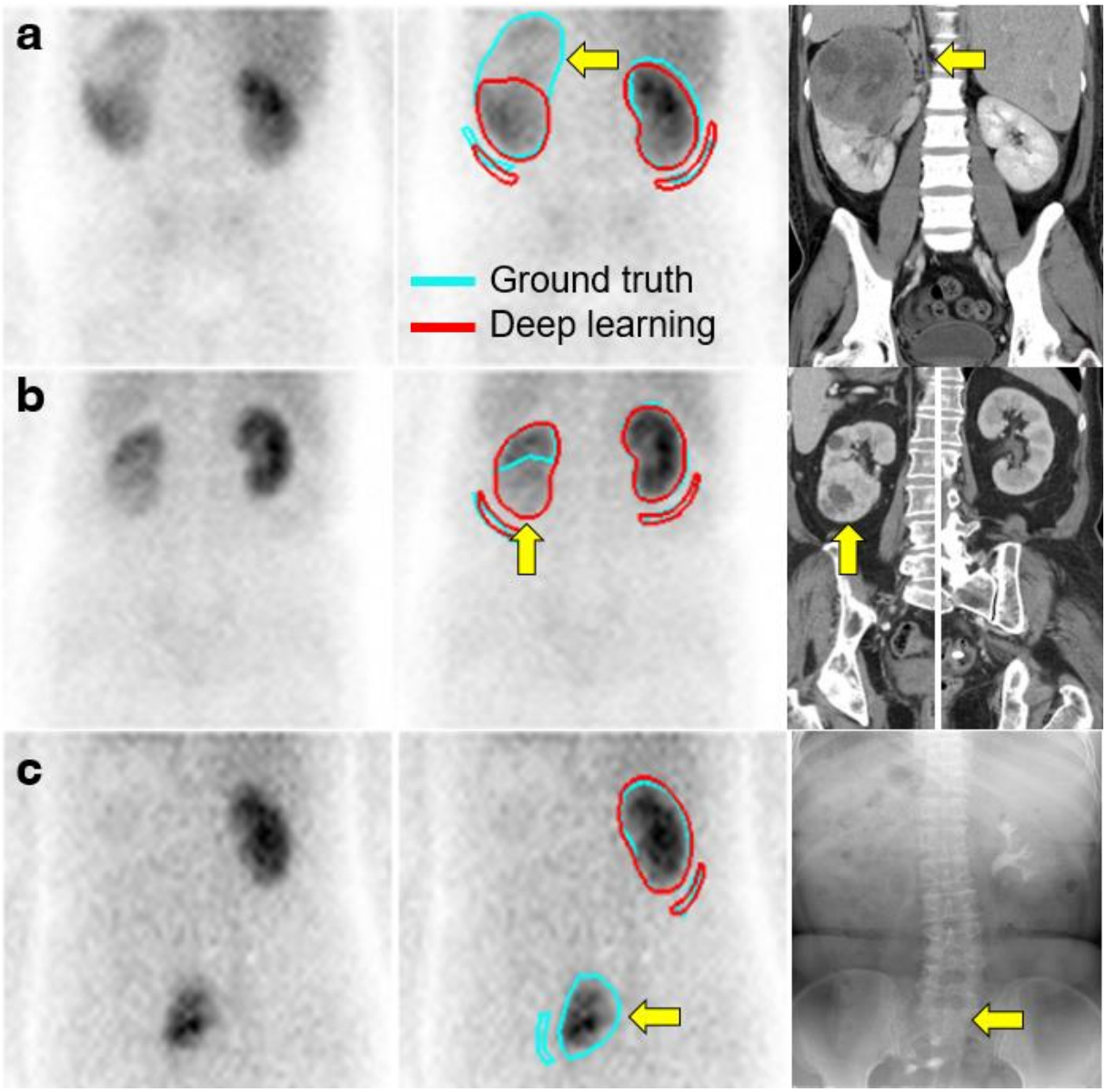


Figure 8

4. Discussion

ROI generator Network

Our study showcased the efficacy of a 2D-CNN with a U-Net structure in accurately delineating kidney ROIs and measuring GFR on ^{99m}Tc-DTPA renal scans. The GFR values obtained from DL-based ROIs closely matched those from manually-drawn ROIs, with an absolute difference of less than 5 mL/min/1.73 m² in 81.9% of cases and a relative difference within 10% in 90.4% of cases. This suggests that using our DL model, 80–90% of patients may undergo GFR measurement without requiring additional manual correction effort. Even in the remaining 10–20% of cases, editing the DL-generated ROI could significantly reduce the overall effort needed for scan analysis. Remarkably, our DL model processed 24,364 ^{99m}Tc-DTPA scans in less than a second each, demonstrating remarkable efficiency. While ^{99m}Tc-DTPA scans are widely used for GFR measurement, they suffer from limited precision and variability, which our DL model can help mitigate by minimizing human error in the image analysis process. [28, 29]

Regarding unique aspect of our ROI generator network research is leveraging existing ROI labels from standard reports stored in our PACS, eliminating the need to generate extensive labels, a common challenge in DL training. Our fully automated approach, from downloading DICOM files to processing images, sets a precedent for large-scale DL studies involving ROIs in related fields. Comparing our approach with Pi et al.'s study, where CNN regressor was applied to derive kidney counts or GFR values directly from dynamic renal images, we speculate that generating ROIs may enable more accurate GFR measurement than direct prediction. Additionally, our DL model's explainable GFR values can be corrected by modifying the ROI if needed, offering flexibility and accuracy in measurement. Moreover, our model for setting kidney and background ROIs can be applied not only for GFR measurement but also for generating renograms to assess renal conditions.

GFR regressor Network

An exceptional milestone in our methodology was the successful extraction of GFR values without the reliance on creating ROIs, representing a significant leap forward in renal scan analysis. By harnessing the capabilities of deep learning, our networks autonomously discerned and quantified pertinent features directly from the input scans, eliminating the need for the conventional ROI generation step. This innovative approach

not only streamlined the analysis process but also underscored the potential of deep learning in revolutionizing medical image analysis.

To estimate the GFR, our GFR regressor network used features such as age, height, and weight, incorporating these parameters into a single input to predict GFR values directly from the scans. This model employed mean squared error as its loss function, with a linear activation function for the output layer to produce continuous GFR values. The results indicated that the model effectively predicted GFR values with an absolute difference of less than 5 mL/min/1.73 m² in 77.0% of cases and within 10% in 84.4% of cases. This shows the potential utility of our GFR regressor network for direct GFR estimation, offering a streamlined workflow without the need for ROI delineation.

Comparing our approach with Pi et al.'s study, where a CNN regressor was applied to derive kidney counts or GFR values directly from dynamic renal images, we speculate that generating ROIs may enable the model to measure GFR with results similar to those obtained from direct prediction. Additionally, our DL model's explainable GFR values can be corrected by modifying the ROI if needed, offering flexibility and accuracy in measurement.

Limitations and Future Studies

Despite these strengths, our study has limitations. Firstly, due to its single-institutional nature and the exclusion of cases with severely reduced renal function, the generalizability of our findings to broader patient populations may be limited. Additionally, the presence of interrater variability in the delineation of GT ROIs by different radiologists introduces uncertainty into the accuracy of the GT. This variability leads to inconsistencies in the inclusion or exclusion of tumors within the ROIs, which in turn affects the predictions of our DL model. As a result, the lack of uniformity in the GT may contribute to an underestimation of the CCC values. Furthermore, while our DL model demonstrates superior reproducibility over manual methods in many cases, direct evidence supporting this claim is lacking, warranting further research to validate its performance and reliability.

We addressed the concern of potential overestimation due to multiple scans per patient through two key safeguards. We ensured a sufficient time interval between scans to capture distinct physiological states, minimizing the likelihood of similar results due to unchanged conditions. Additionally, we used different

imaging devices and interpreted by various radiologists to minimize systematic biases. This enhanced the generalizability and reliability of our findings.

The grad-CAM visualizations demonstrated a significant improvement in the explainability of the GFR regressor network. By utilizing left and right masked inputs, the network produced more accurate and relevant heatmaps that closely aligned with the labeled ground truth. This approach allowed the GFR regressor network to focus its feature extraction processes on the specific anatomical regions of interest for each kidney, thereby enhancing the interpretability of the model's predictions. In previous analyses, grad-CAM heatmaps exhibited discrepancies, often showing activity near both kidneys when predicting the GFR for one side. This issue likely arose because the input image used for the renal scan included information from both kidneys, whereas the labels were designated for the GFR of each kidney separately. By using masked inputs, the network can now target the relevant kidney, resulting in heatmaps that better align with the clinically labeled ROIs.

However, despite these improvements, the explainability of the GFR regressor network still falls short compared to the ROI generator network. While the GFR regressor network's heatmaps are more accurate than before, they are not as intuitively clear or as well-aligned with the labeled ROIs as those produced by the ROI generator network. This suggests that the ROI generator network remains superior in providing clear and actionable insights for clinical decision-making.

5. Conclusions

Our 2D-CNN model, incorporating the ROI generator network architecture, showcased remarkable proficiency in segmenting kidney ROIs on ^{99m}Tc-DTPA scans. The GFR values derived from automatically generated ROIs exhibited strong agreement with those manually delineated by experts. In a significant majority of cases (80-90% of patients), the DL approach enabled GFR measurements without the need for additional manual calibration, showing negligible absolute differences or relative differences within 10%.

Additionally, the GFR regressor network demonstrated similar concordance in terms of the CCC. Although it does not provide the same level of explainability, the GFR regressor network produced comparable CCC results, suggesting its potential utility in GFR estimation from renal scans. This indicates that the GFR regressor network can serve as an efficient alternative for direct GFR estimation, offering a streamlined workflow without the need for ROI delineation.

As a result, our study paves the way for straightforward, operator-independent analysis of ^{99m}Tc-DTPA scans for GFR assessment in a consistently reproducible manner. However, variability in GT labeling due to different radiologists' interpretations may have led to an underestimation of the CCC values. Future research should continue to explore these findings across multiple institutions and in patients with a broader range of renal functions to further validate and refine our model.

6. References

1. Blaufox MD, De Palma D, Taylor A, Szabo Z, Prigent A, Samal M, et al. The SNMMI and EANM practice guideline for renal scintigraphy in adults. Eur J Nucl Med Mol Imaging. 2018;45:2218-28. https://doi.org/10.1007/s00259-018-4129-6.

2. Geist BK, Dobrozemsky G, Samal M, Schaffarich MP, Sinzinger H, Staudenherz A. WWSSF - a worldwide study on radioisotopic renal split function: reproducibility of renal split function assessment in children. Nucl Med Commun. 2015;36:1233-8. https://doi:10.1097/mnm.0000000000000380.

3. Gates GF. Glomerular filtration rate: estimation from fractional renal accumulation of 99mTc-DTPA (stannous). AJR Am J Roentgenol. 1982;138:565-70. https://doi.org/10.2214/ajr.138.3.565.

4. Gates GF. Computation of glomerular filtration rate with Tc-99m DTPA: an in-house computer program. J Nucl Med. 1984;25:613-8.

5. BYFIELD GV, TELSER SE, KEETON RW. RENAL BLOOD FLOW AND GLOMERULAR FILTRATION: AS INFLUENCED BY ENVIRONMENTAL TEMPERATURE CHANGES. Journal of the American Medical Association. 1943;121:118-23. https://doi.org/10.1001/jama.1943.02840020026006.

6. Piepsz A, Dobbeleir A, Ham HR. Effect of background correction on separate technetium-99m-DTPA renal clearance. J Nucl Med. 1990;31:430-5.

7. Prigent A, Cosgriff P, Gates GF, Granerus G, Fine EJ, Itoh K, et al. Consensus report on quality control of quantitative measurements of renal function obtained from the renogram: International Consensus Committee from the Scientific Committee of Radionuclides in Nephrourology. Semin Nucl Med. 1999;29:146-59. doi:10.1016/s0001-2998(99)80005-1.

8. Seo SY, Oh JS, Chung J, Kim SY, Kim JS. MR Template-Based Individual Brain PET Volumes-of-Interest Generation Neither Using MR nor Using Spatial Normalization. Nucl Med Mol Imaging. 2023;57:73-85. https://doi.org/10.1007/s13139-022-00772-4.

9. Gil J, Choi H, Paeng JC, Cheon GJ, Kang KW. Deep Learning-Based Feature Extraction from Whole-

Body PET/CT Employing Maximum Intensity Projection Images: Preliminary Results of Lung Cancer Data. Nucl Med Mol Imaging. 2023;57:216-22. https://doi.org/10.1007/s13139-023-00802-9.

10. Park J, Kang SK, Hwang D, Choi H, Ha S, Seo JM, et al. Automatic Lung Cancer Segmentation in [(18)F]FDG PET/CT Using a Two-Stage Deep Learning Approach. Nucl Med Mol Imaging. 2023;57:86-93. https://doi.org/10.1007/s13139-022-00745-7.

11. Han S, Oh JS, Kim YI, Seo SY, Lee GD, Park MJ, et al. Fully Automatic Quantitative Measurement of 18F-FDG PET/CT in Thymic Epithelial Tumors Using a Convolutional Neural Network. Clin Nucl Med. 2022;47:590-8. https://doi.org/10.1097/rlu.0000000000004146.

12. Jiang H, Diao Z, Yao Y-D. Deep learning techniques for tumor segmentation: a review. The Journal of Supercomputing. 2022;78:1807-51. https://doi.org/10.1007/s11227-021-03901-6.

13. Park J, Bae S, Seo S, Park S, Bang JI, Han JH, et al. Measurement of Glomerular Filtration Rate using Quantitative SPECT/CT and Deep-learning-based Kidney Segmentation. Sci Rep. 2019;9:4223. https://doi.org/10.1038/s41598-019-40710-7.

14. Sharma K, Rupprecht C, Caroli A, Aparicio MC, Remuzzi A, Baust M, et al. Automatic Segmentation of Kidneys using Deep Learning for Total Kidney Volume Quantification in Autosomal Dominant Polycystic Kidney Disease. Sci Rep. 2017;7:2049. https://doi.org/10.1038/s41598-017-01779-0.

15. Moe YM, Groendahl AR, Tomic O, Dale E, Malinen E, Futsaether CM, et al. Deep learning-based auto-delineation of gross tumour volumes and involved nodes in PET/CT images of head and neck cancer patients. Eur J Nucl Med Mol Imaging. 2021;48:2782-92. https://doi.org/10.1007/s00259-020-05125-x

16. Ronneberger O, Fischer P, Brox T, et al. U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI. 2015;18:234.241 https://doi.org/10.1007/978-3-319-24574-4_28.

17. Isensee, F., Jaeger, P.F., Kohl, S.A.A. et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 2021;18:203–211. https://doi.org/10.1038/s41592-020-01008-z

18. Pi Y, Zhao Z, Yang P, Cheng J, Jiang L, Wei J, et al. Deep regression using (99m)Tc-DTPA dynamic renal imaging for automatic calculation of the glomerular filtration rate. Eur Radiol. 2023;33:34-42. https://doi.org/10.1007/s00330-022-08970-6.

19. Nayyar R, Yadav S, Singh P, Kumar R, Seth A, Dogra PN. Outcomes of Pyeloplasty in Very Poorly Functioning Kidneys: Examining the Myths. Urology. 2016;92:132-5. https://doi:10.1016/j.urology.2016.02.045.

20. Inoue Y, Ohtake T, Homma Y, Yoshikawa K, Nishikawa J, Sasaki Y. Evaluation of glomerular filtration rate by camera-based method in both children and adults. J Nucl Med. 1998;39:1784-8.

21. Allison SJ. The CKD–EPI equation—accurately stratifying risk in CKD. Nature Reviews Nephrology. 2012;8:371. https://doi.org/10.1038/nrneph.2012.94

22. Kim HO, Chae SY, Baek S, Moon DH. Factors affecting changes in the glomerular filtration rate after unilateral nephrectomy in living kidney donors and patients with renal disease. Nucl Med Mol Imaging. 2010;44:69-74. https://doi.org/10.1007/s13139-009-0010-7.

23. Tonnesen K. Influence on the renogram of variation on skin to kidney distance and the clinical importance thereof. Radionuclides in nephrology. 1975:79-86.

24. Taylor A, Lewis C, Giacometti A, Hall EC, Barefield KP. Improved formulas for the estimation of renal depth in adults. J Nucl Med. 1993;34:1766-9.

25. Yoo IR, Kim SH, Chung YA. Development of formulas for the estimation of renal depth and application in the measurement of glomerular filtration rate in Koreans. Korean Journal of Nuclear Medicine. 2000;34:418-25.

26. Turin G. An introduction to matched filters. IRE Transactions on Information Theory. 1960;6:311-29. https://doi.org/10.1109/TIT.1960.1057571.

27. Han S, Oh JS, Lee JJ. Diagnostic performance of deep learning models for detecting bone metastasis on whole-body bone scan in prostate cancer. Eur J Nucl Med Mol Imaging. 2022;49:585-95. https://doi.org/10.1007/s00259-021-05481-2.

28. Assadi M, Eftekhari M, Hozhabrosadati M, Saghari M, Ebrahimi A, Nabipour I, et al. Comparison of methods for determination of glomerular filtration rate: low and high-dose Tc-99m-DTPA renography, predicted creatinine clearance method, and plasma sample method. Int Urol Nephrol. 2008;40:1059-65. https://doi.org/10.1007/s11255-008-9446-4.

29. De Santo NG, Anastasio P, Cirillo M, Santoro D, Spitali L, Mansi L, et al. Measurement of glomerular filtration rate by the 99mTc-DTPA renogram is less precise than measured and predicted creatinine clearance. Nephron. 1999;81:136-40. https://doi.org/10.1159/000045268.

7. Korean Abstract

목적 Glomerular filtration rate (GFR) 측정을 위한 ^{99m}Tc-diethylenetriamine penataacetic acid (DTPA) 신장 스캔에서 자동화된 관심영역(ROI)을 생성하기 위한 딥러닝 모델을 개발하였다.

방법 수동으로 그려진 ROI 와 해당 GFR 값은 Picture Archiving and Communications System (PACS) 에서 검색한 GT 레이블 및 대상 값으로 사용하였다. 이를 위해, 다차원 입력을 사용하는 2 차원 U-Net 합성곱 신경망(CNN) 아키텍처(ROI 생성기 네트워크)와 다차원 입력을 사용하는 2 차원 인코더 CNN 아키텍처(GFR 회귀 네트워크)로 구성된 두 가지 모델이 훈련되어 DL ROI 및 DL GFR 값을 생성하였다. 후자(즉, 회귀) 모델은 ROI 를 생성하지 않고 GFR 값을 생성하기 위해 사용되었다. GT 및 DL ROI 간의 GFR 값의 일치는 Lin 의 일치 상관 계수(CCC) 및 선형 회귀 분석의 기울기 계수를 사용하여 평가하였다. 바이어스 및 95% 근접한 한계(LOA)를 Bland-Altman 그림을 사용하여 평가하였다.

결과 총 24364 개의 스캔 (12821 명의 환자)이 포함되었다. ROI 생성 네트워크에 대해서는 좌측 신장(CCC 0.982, 95% 신뢰 구간 [CI] 0.981-0.982; 기울기 1.004, 95% CI 1.003-1.004), 우측 신장(CCC 0.969, 95% CI 0.968-0.969; 기울기 0.954, 95% CI 0.953-0.955), 그리고 양쪽 신장(CCC 0.978, 95% CI 0.978-0.979; 기울기 0.979, 95% CI 0.978-0.979)에 대해 GT 및 DL GFR 사이의 일치가 우수한 것을 발견하였다. Bland-Altman 분석 결과, 좌측, 우측 및 양쪽 신장에 대한 GT 및 DL GFR 간의 평균 차이는 각각 -0.2 (95% LOA -4.4-4.0), 1.4 (95% LOA -3.5-6.3) 및 1.2 (95% LOA -6.5-8.8) mL/min/1.73 m² 임을 확인하였다. 회귀 네트워크의 경우, GT 및 DL 추정 GFR 값은 좌측 신장 (CCC 0.969, 95% CI 0.969-0.970), 우측 신장 (CCC 0.969, 95% CI 0.968-0.969) 및 양쪽 신장 (CCC 0.972, 95% CI 0.971-0.972)임을 확인하였다. 주목할 만하게, ROI 생성 네트워크에 의해 GFR 이 5 mL/min/1.73 m² 미만의 절대 차이를 보인 스캔은 19960 건(81.9%)이었다. 유사하게, 회귀 네트워크는 18770 건의 스캔(77.0%)에서 GFR 이 5 mL/min/1.73 m² 미만의 절대 차이를 보였다.

결론 우리의 ROI 생성기 네트워크와 GFR 회귀 네트워크는 ^{99m}Tc-DTPA 신장 스캔에서 ROIs 생성 및 GFR을 예측하는데 우수한 성능을 보였다. 이 자동화된 접근 방식은 임상 실무에서 수동 노력을 줄이고 GFR 측정의 정밀도를 향상시킬 수 있다.

8. Figure legends

Figure 1. Schematic representation of the manual ROI extraction and GFR calculation process utilised in this study. A structured report along with 80 raw frame images (15 second per frame) was obtained in DICOM format. The images were anonymised automatically with a black-filled square using in-house software. GFR was calculated on the basis of renal uptake of ^{99m}Tc-DTPA captured in a 2–3-minute (9–12th frame) summed image (blue arrow). The manual ROIs for each kidney and the background, identified using color-based edge detection (red arrow), were superimposed onto the summed images for GFR estimation.

Figure 2. Schematic diagrams illustrating the data allocation and deep CNN architecture (a) Data allocation strategy (training:test-set ratio of 3:1) with a four-fold cross-validation-like training scheme. (b) U-Net architecture for kidney ROI delineation using deep learning with multi-channel image input (N × 256 × 256 × 4). The output image is assessed against the ground truth using the Dice loss function. (c) The specific architecture of the U-Net for ROI delineation. (d) The specific architecture of the regressor model for estimating GFR.

Figure 3. Patient flow diagram

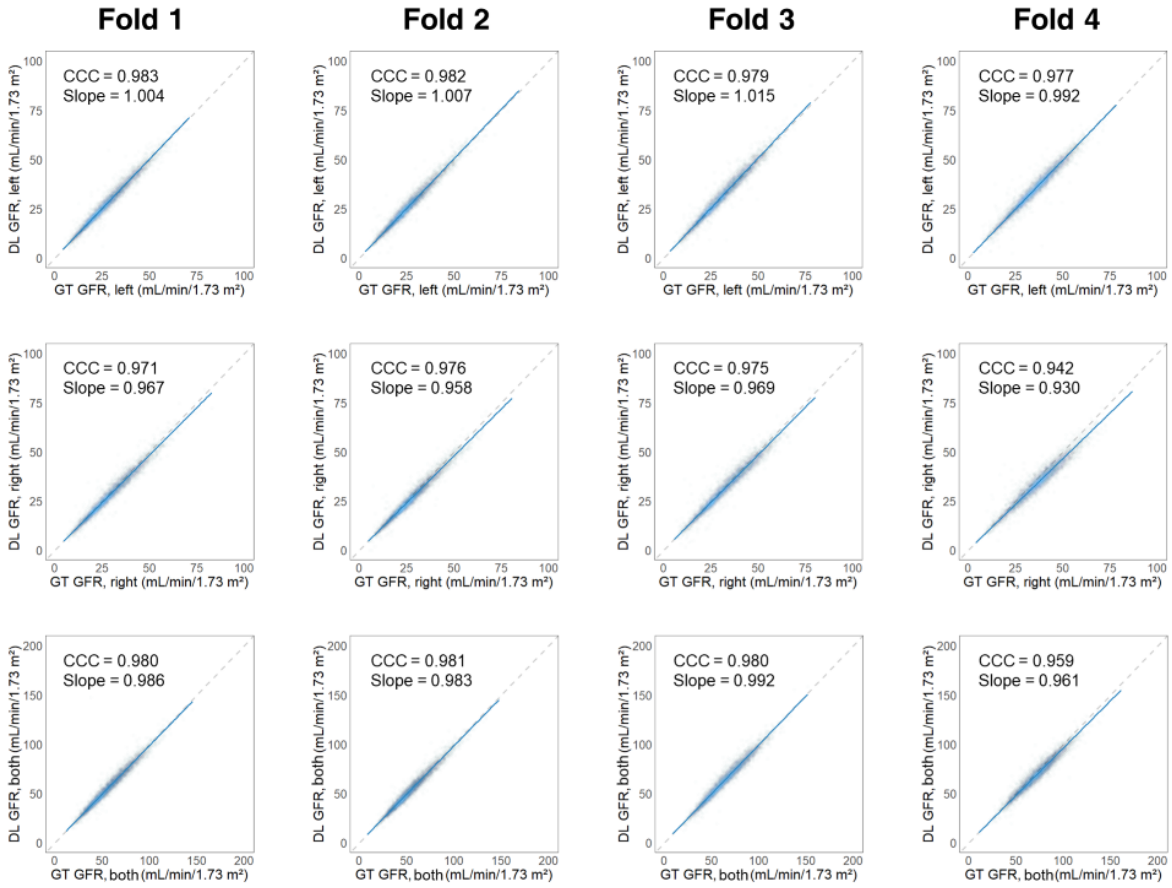
Figure 4. Scatter plots showing the agreement between the counts on the reported PACS form and the count calculated from the extracted ground-truth ROI for (a) left kidney and (b) right kidney. Dashed lines represent lines of equality.

Figure 5. Scatter plots showing the agreement between ground truth (GT) GFR and deep learning (DL) GFR for (a) left kidney, (b) right kidney and (c) both kidneys using the U-Net model, and (d) left kidney, (e) right kidney and (f) both kidneys using the regressor model. Dashed lines represent lines of equality.

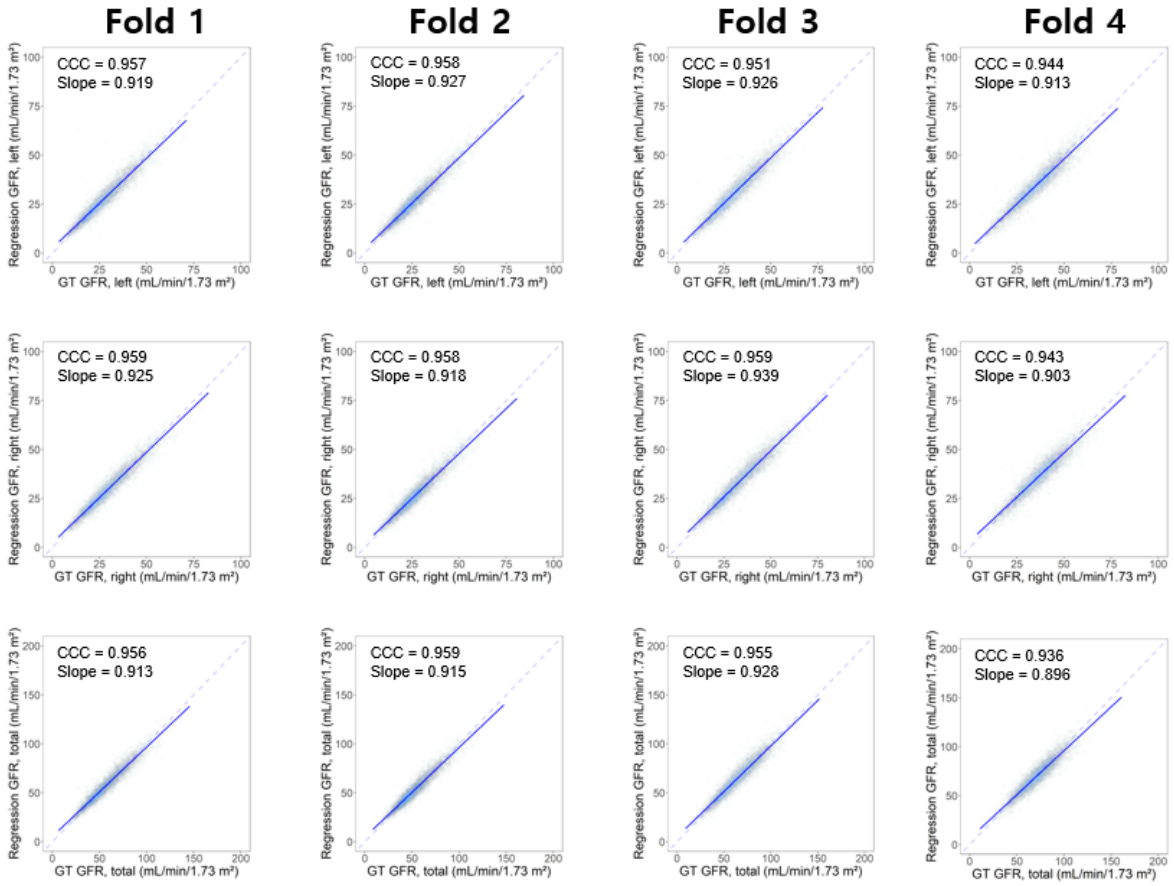
Figure 6. Bland-Altman plots illustrating the agreement between ground truth (GT) GFR and deep learning (DL) GFR for (a) left kidney, (b) right kidney and (c) both kidneys using the U-Net model, and (d) left kidney, (e) right kidney and (f) both kidneys using the regressor model. Solid lines represent mean differences. Dashed lines represent 95% limits of agreement.

Figure 7. Visual interpretability of the GFR regressor model's heatmap. (a) Left and right masked input images. (b) Ground truth (GT) kidney ROIs. (c) Images obtained by applying KMeans clustering and a 60% maximum threshold from the raw image to identify the kidney regions. (d) Heatmap extracted from the final layer of the network. (e) Heatmap from (d) zoomed to a size of 256x256 pixels. (f) Final heatmap obtained by identifying the region overlapping the most with the clustered kidney region in (c) and applying a threshold of 0.1.

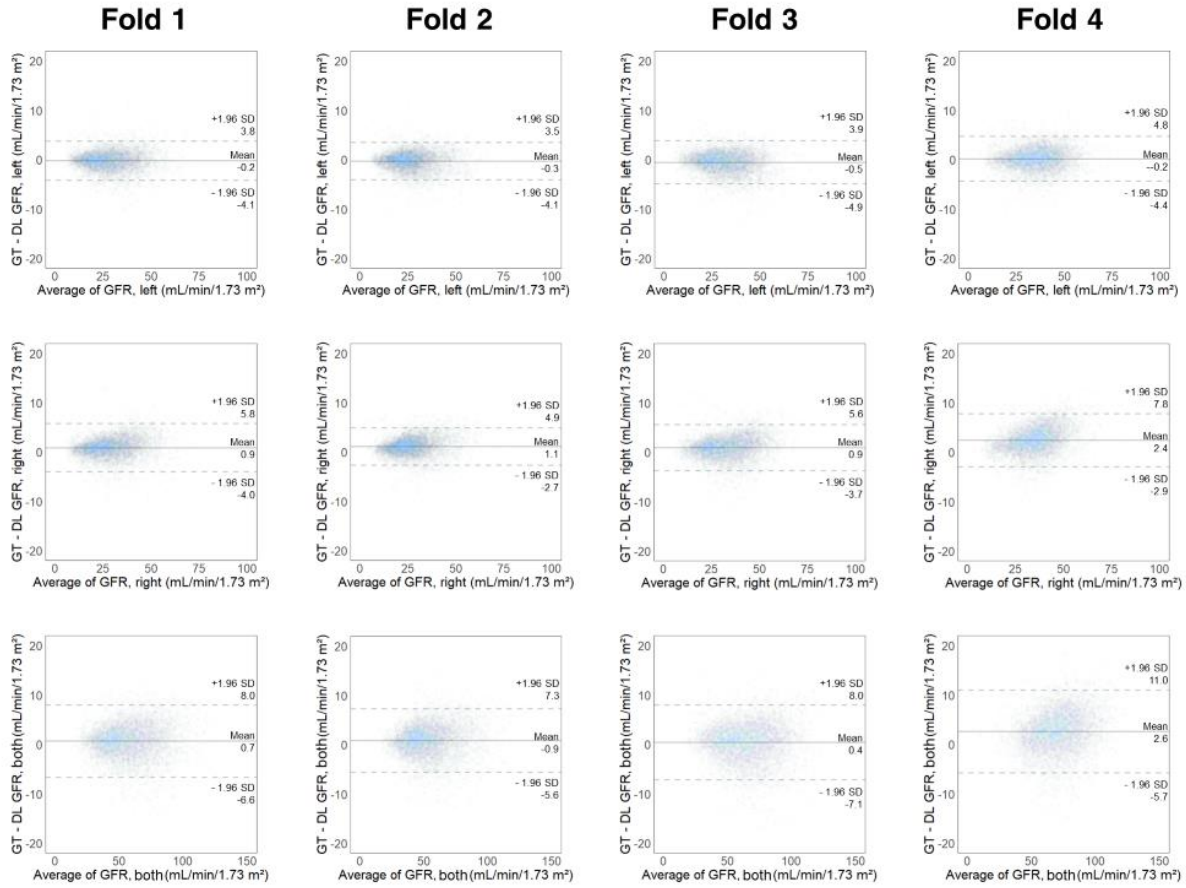
Figure 8. Representative images showing a low concordance between ground truth (GT) ROIs and deep learning (DL) ROIs. Each row displays a 2–3-minute summed image (left), the same image with superimposed ROIs (middle) and a corresponding CT image or intravenous pyelogram (right). Panel (a) shows an example where the GT ROI includes a tumour in the upper pole of the left kidney (yellow arrow), whereas the DL ROI does not. Panel (b) shows a case where the GT ROI excludes the tumour in the lower pole of left kidney (yellow arrow), but the DL ROI includes it. Panel (c) presents a left ectopic kidney in the pelvic cavity that was not delineated by the DL ROI (yellow arrow).



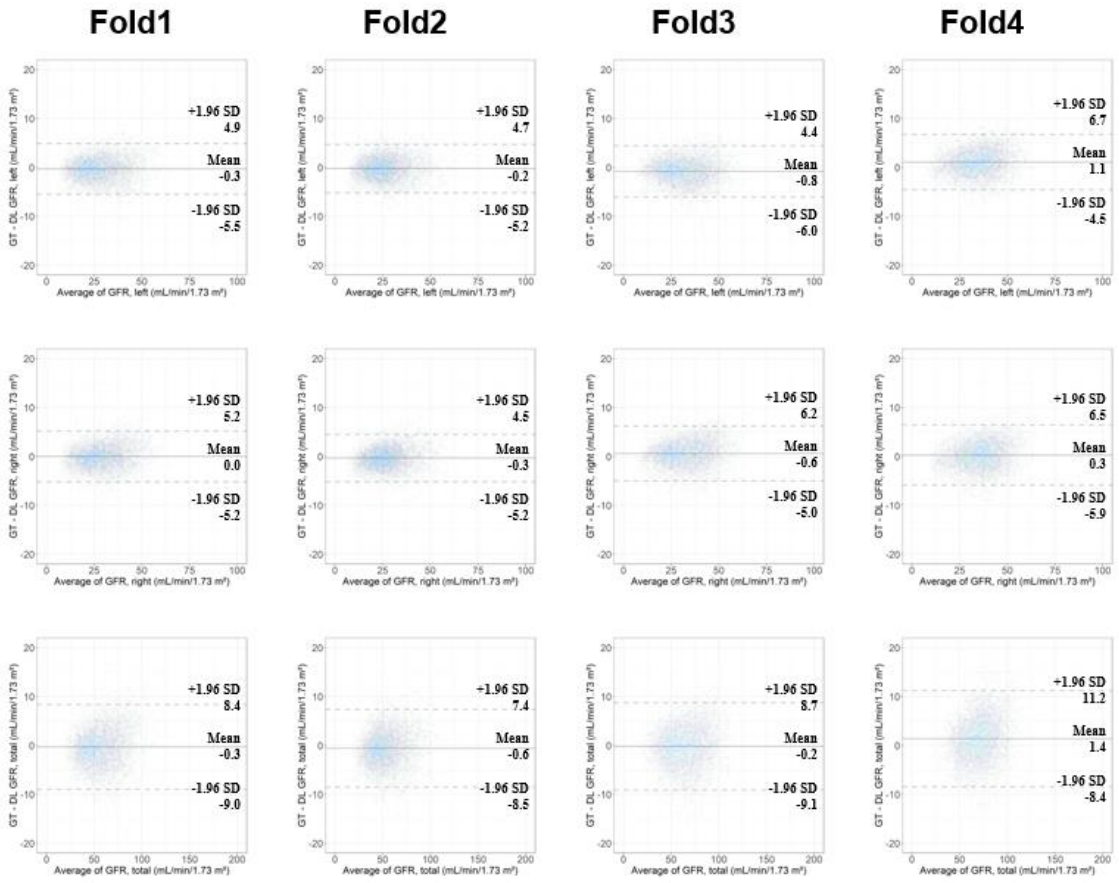
Supplementary Figure 1. Scatter plots showing the agreement between ground truth (GT) GFR and U-Net (DL) GFR for left kidney, right kidney and both kidneys across four cross-validation folds. Dashed lines represent lines of equality.



Supplementary Figure 2. Scatter plots showing the agreement between ground truth (GT) GFR and regressor model (Regression) GFR for left kidney, right kidney and both kidneys across four cross-validation folds. Dashed lines represent lines of equality.



Supplementary Figure 3. Bland-Altman plots illustrating the agreement between ground truth (GT) GFR and U-Net (DL) GFR for left kidney, right kidney and both kidneys across four cross-validation folds. Solid lines represent mean differences. Dashed lines represent 95% limits of agreement.



Supplementary Figure 4. Bland-Altman plots illustrating the agreement between ground truth (GT) GFR and regressor model GFR for left kidney, right kidney and both kidneys across four cross-validation folds. Solid lines represent mean differences. Dashed lines represent 95% limits of agreement.

9. Table

Table 1. Study characteristics

Data are expressed as number (proportion) or median.

*The estimated GFR was calculated by CKD-EPI equation

Variables	Total scans	Scans with renal mass	Scans without renal mass
Number of scans	24364	4067	20297
Number of patients	12822	2081	10741
Age at scan time (years)	56 (16–94)	56 (16–91)	56 (16–94)
Male:Female	15799:8565	2562:1505	13237:7060
History of partial nephrectomy	10525 (43%)	122 (3%)	10403 (52%)
Serum creatinine (mg/dL)	0.87 (IQR, 0.73–1.01)	0.86 (IQR, 0.71–1.00)	0.87 (IQR, 0.71–0.96)
BUN (mg/dL)	14.0 (IQR, 12.0–18.0)	14.0 (IQR, 11.0–17.0)	13.0 (IQR, 11.0–15.0)
Estimated GFR (mL/min/1.73 m^2)*			
<15	40 (0%)	4 (0%)	36 (0%)
15–30	168 (1%)	19 (0%)	149 (1%)
30–45	519 (2%)	80 (2%)	439 (2%)
45–60	1564 (6%)	271 (7%)	1293 (6%)
60–90	9161 (38%)	1553 (38%)	7608 (37%)
≥90	11451 (47%)	1941 (48%)	9510 (47%)
Not available	1461 (6%)	199 (5%)	1262 (6%)

Supplementary Table 1. The number of scans performed using each gamma camera

Cameras	Number of Scans
Intevo Bold	80
Intevo 16	39
Symbia E	11,536
Symbia E2	3,350
Symbia T2	2,083
E.cam	5,417
Evo Excel	1,859
Total	**24,364**

Supplementary Table 2. Summary of ^{99m}Tc-DTPA scan sessions per patient.

Sessions per patient	Patients (*n*)	Total Scans (*n*)	Time interval between scans (months)
1	8,743	8,743	N/A
2	1,329	2,658	5.9
3	1,056	3,168	7.5
4	542	2,168	12.1
5	289	1,445	11.9
6	263	1,578	12.1
7	300	2,100	12.1
8	221	1,768	12.1
9	56	504	12.1
10	21	210	12.0
11	2	22	11.8
Total	12,822	24,364	

Data are expressed as number or median