



Comprehensive 3D Analysis of the Renal System and Stones: Segmenting and Registering Non-Contrast and Contrast Computed Tomography Images

Zhuo Chen^{1,2} · Chuda Xiao² · Yang Liu³ · Haseeb Hassan¹ · Dan Li² · Jun Liu² · Haoyu Li¹ · Weiguo Xie² · Wen Zhong³ · Bingding Huang¹

Accepted: 9 March 2024
© The Author(s) 2024

Abstract

Detecting and accurately locating kidney stones, which are common urological conditions, can be challenging when using imaging examinations. Therefore, the primary objective of this research is to develop an ensemble model that integrates segmentation and registration techniques. This model aims to visualize the inner structure of the kidney and accurately identify any underlying kidney stones. To achieve this, three separate datasets, namely non-contrast computed tomography (CT) scans, corticomedullary CT scans, and CT excretory scans, are annotated to enhance the three-dimensional (3D) reconstruction of the kidney's complex anatomy. Initially, the research focuses on utilizing segmentation models to identify and annotate specific classes within the annotated datasets. Subsequently, a registration algorithm is employed to align and combine the segmented results, resulting in a comprehensive 3D representation of the kidney's anatomical structure. Three cutting-edge segmentation algorithms are employed and evaluated during the segmentation phase, with the most accurate segments being selected for the subsequent registration process. Ultimately, the registration process successfully aligns the kidneys across all three phases and combines the segmented labels, producing a detailed 3D visualization of the complete kidney structure. For kidney segmentation, Swin UNETR exhibited the highest Dice score of 95.21%; for stone segmentation, ResU-Net achieved the highest Dice score of 87.69%. Regarding Artery, Cortex, and Medulla segmentation, ResU-Net and 3D U-Net show comparable performance with similar Dice scores. Considering the Collecting System and Parenchyma, ResU-Net and 3D U-Net demonstrate similar performance in Dice scores. In conclusion, the proposed ensemble model shows potential in accurately visualizing the internal structure of the kidney and precisely localizing kidney stones. This advancement improves the diagnosis process and preoperative planning in percutaneous nephrolithotomy.

Keywords 3D kidney reconstruction · Abdominal CT scans · Renal diseases · Segmentation and registration · Percutaneous Nephrolithotomy (PCNL)

Zhuo Chen, Chuda Xiao and Yang Liu These authors contributed equally to this work.

✉ Wen Zhong
gzgyzhongwen@163.com

✉ Bingding Huang
huangbingding@sztu.edu.cn

¹ College of Big Data and Internet, Shenzhen Technology University, Shenzhen, China

² Wuerzburg Dynamics Inc., Shenzhen, China

³ Department of Urology, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

1 Introduction

The kidney plays a crucial role in the human body by purifying harmful substances from the bloodstream. However, it can also give rise to different irregularities, such as congenital kidney disease (Schedl, 2007), kidney infection (Flores-Mireles et al., 2015), diabetic nephropathy (Lim, 2014), polycystic kidney disease, urinary tract infection (Murray et al., 2021), and kidney stones (Piao et al., 2015). Among these, kidney stones, also known as renal calculi, nephrolithiasis, or urolithiasis, are a common urological condition that affects approximately 1 in 11 people and is associated with multiple complications (Zisman et al., 2015). Healthcare professionals use various imaging techniques such as

ultrasound (US), X-ray, magnetic resonance imaging (MRI), and computed tomography (CT) to diagnose and detect kidney stones. While these imaging methods can identify kidney stones within the urinary tract, the intricate anatomical structure of the kidney makes it difficult to definitively exclude the presence of underlying stones.

CT scans utilize a combination of X-rays and computer technology to generate detailed images of the inner structure of the kidney. These scans can be performed with or without contrast, depending on the requirements. Advanced techniques like high-speed or dual-energy CT can detect tiny kidney stones (Cuingnet et al., 2012). However, detecting and localizing kidney stones in CT images can be challenging. The similarity in shape and intensity between stones and non-stone structures can pose difficulties for computer-assisted detection (CAD) of stones and how to manage large volumes of high-resolution CT data. Furthermore, the varying sizes, diverse appearances, deformable and small shapes from slice to slice, and the prevalence of negative samples in an abdominal CT scan present severe challenges for automatic CAD systems.

To address the challenges, artificial intelligence (AI) and CAD systems have increasingly turned to deep-learning algorithms (Yang et al., 2020). These technologies have been applied in the medical field, particularly image segmentation (Kakhandaki & Kulkarni, 2023). Furthermore, AI has been utilized in radiomics to analyze CT, MRI, and US images for various purposes, such as identifying stone dimensions, detecting stones and stone composition, predicting spontaneous stone passage, and forecasting outcomes of endourological operations (Hameed et al., 2021). For instance, our previous research (Li et al., 2022) aimed to develop an automated system for detecting kidneys and kidney stones using non-contrast CT (NCT) scans. However, normal CT images alone do not provide a complete representation of the intricate inner structure of the kidney, including the size of kidney stones, medulla, and vessels. Consequently, more detailed information is necessary for accurate diagnosis and effective preoperative planning in percutaneous nephrolithotomy (PCNL) (Ansari et al., 2022).

The main objective of this research is to reconstruct a three-dimensional (3D) model of the kidney that includes all its anatomical components, including renal stones, renal arteries, renal pelvis, and so on. As mentioned above, NCT scans only allow visualization of kidney stones and the contours of the kidneys, while other anatomical structures cannot be seen. To address this limitation, three annotated datasets were created with classes for stones and cortex NCT scans, enhanced corticomedullary TC (CTC) with classes for artery, renal cortex, and renal medulla, and enhanced CT of excretory (CTE) with classes for renal cortex and collection system. By combining these three-phase CT scans (NCT, CTC, and CTE), the reconstruction of the complex structure

of the kidney could be enhanced. Initially, the phase-specific segmentation models using the three-phase abdominal CT data were employed to achieve this. Since the three-phase CT scans are captured at different time intervals, the position of the kidneys may shift. To tackle this issue, our proposed experiments aim to align the segmentation results spatially and merge them into a single image that accurately represents the complete 3D anatomical structure of the kidneys.

In this work, the main aim is to create an ensemble model combining segmentation and registration tasks in abdominal CT scans. This model will leverage non-contrast and contrast CT scans to visualize the inner structure of the kidney and accurately locate any potential underlying stones. The main contributions of this work are outlined below.

- (1) **Dataset Annotation:** Annotated three distinct datasets consisting of non-contrast and contrast CT scans, with each dataset assigned its specific labeled classes.
- (2) **Segmentation Models:** Three segmentation models were adopted and trained using the annotated three-phase CT datasets. These models were then tested and validated to generate accurate segmented outputs.
- (3) **Rigid Registration:** This algorithm spatially aligns and merges the segmented results of different classes, resulting in a comprehensive and accurate 3D representation of the kidney's anatomical structure.

It should be noted that prior to this research, no existing system had been developed to simultaneously utilize both contrast CT and NCT scans and incorporate both segmentation and registration processes for reconstructing the complete 3D kidney and localization of renal stones. Therefore, to the best of our knowledge, we are the first to do so. The rest of this paper is organized as follows. Section 2 provides an overview of the related work in the field, followed by a description of the materials and methods in Section 3. Section 4 presents the results and key findings. Lastly, in Section 5, a detailed conclusion that summarizes the overall research is provided.

2 Related Work

This research aims to develop an ensemble system considering contrast and non-contrast CT scans. The ultimate goal is to integrate image registration and segmentation techniques to predict the presence of renal stones and the renal collection system. Unlike previous methods that predominantly relied on either segmentation or registration alone, our approach combines both processes. We have extensively reviewed related studies encompassing various aspects such as kidney segmentation, kidney stone segmentation, renal cortex, medulla, pelvicalyceal system, and registration.

Notably, renal segmentation on contrast-enhanced CT scans offers valuable spatial context and morphology information, which contributes to the distinctiveness of the proposed approach.

An example of a practical model-based approach for computer-aided kidney segmentation in abdominal CT images is presented (Lin et al., 2006). This method specifically took into account the anatomical structure of the kidney and utilized contrast-enhanced CT images. The results of this approach demonstrated an average correlation coefficient of up to 88% compared to manual segmentation. However, it is important to note that this method is considered conventional and not recent. In (Thong et al., 2018), a convolutional neural network (CNN) was trained to predict the class membership of the central voxel in two-dimensional (2D) patches. By densely applying the CNN over each slice of a contrast-enhanced CT scan, kidney segmentation is achieved. Efficient predictions were achieved by transforming fully connected layers into convolutional operations. Additionally, a semi-automated segmentation method for kidneys using a graph-cuts technique was introduced (Shim et al., 2009). However, it is worth noting that this particular method solely relied on contrast-based modalities and did not incorporate the combination of both NCT and contrast CT scans.

Some studies also considered segmentation of the renal cortex, medulla, and pelvicalyceal system. For example, one such study introduced an automatic segmentation algorithm to accurately segment and label these three specific renal structures. The algorithm utilized a coarse-to-fine framework based on deep neural networks (Tang et al., 2021). Xinjian et al. (Chen et al., 2012) conducted a study where the authors developed and validated an automated method for segmenting the renal cortex on contrast-enhanced abdominal CT images obtained from kidney donors. Their objective was to track the volume change of the renal cortex after kidney donation. However, it is important to note that the modality used in their study was limited to contrast-enhanced CT scans. Similarly, in Li et al. (2011), a conventional approach was adopted to address renal cortex segmentation. This method introduced a novel graph construction scheme to tackle this specific segmentation problem. However, the proposed method only considered the segmentation of the renal cortex. Takazawa et al. (Takazawa et al., 2018) proposed a straightforward anatomical classification method for the pelvicalyceal system, specifically designed for endoscopic surgery. Their approach did not involve any registration module but allowed for sharing common intrarenal information, aiding the development of effective treatment strategies. Another method (Tang et al., 2010) successfully achieved the cortex, medulla, and collection system segmentation using clustering and inferring modules. In the final module, a two-stage recognition process was employed to label each

cluster, first distinguishing between the kidney and background and then identifying the different compartments, including the cortex, medulla, and collection system. The comparison of their method demonstrated promising results in both anatomical and functional analysis. However, it is important to note that the proposed method used magnetic resonance modality instead of CT. In a more recent study (Korfiatis et al., 2022), a deep neural network-based model was developed to segment the kidney cortex and medulla specifically in contrast-enhanced CT images.

In addition to segmenting the cortex, medulla, and collection system, several studies have also focused on the segmentation and localization of kidney and renal stones. Yildirim et al. (Yildirim et al., 2021) proposed a deep-learning model for detecting kidney stones using coronal CT images. The proposed model achieved an accuracy of 96.82% in detecting kidney stones based on CT images. In another study (Elton et al., 2022), a 3D U-Net model was employed to segment and detect kidney stones. The system incorporated gradient-based anisotropic denoising, thresholding, and region-growing techniques. It achieved a sensitivity of 0.86 at 0.5 false positives per scan on a challenging test set consisting of low-dose CT scans with numerous small stones. A different approach was proposed in Baygin et al. (2022), where the authors introduced ExDark19, a method that efficiently detects kidney stones using CT images. The authors employed iterative neighborhood component analysis to select the most informative feature vectors and then utilized a k-nearest neighbor classifier for stone detection with a ten-fold cross-validation strategy. A closely related work (Parakh et al., 2019) focused on the accuracy of a cascading CNN for urinary stone detection on unenhanced CT images. The study evaluated the performance of pre-trained models enriched with labeled CT images. However, the pre-trained weights were borrowed from ImageNet and fine-tuned on their in-house GrayNet dataset.

The aforementioned studies primarily focused on straightforward segmentation and detection tasks and typically utilized single imaging modalities such as contrast CT or NCT. However, the segmentation of renal structures can be challenging due to factors like limited field-of-view and patient variability. Therefore, there is a need for strategies that can generate accurate spatial mapping to align images acquired at different time points, and one such strategy is image registration. So far, various conventional methods that utilize registration techniques in renal structure segmentation have been developed. For instance, Martin et al. (Spiegel et al., 2009) proposed a kidney segmentation technique that specifically considered contrast-enhanced CT scans. Their approach addressed the crucial correspondence problem by employing non-rigid image registration. By doing so, the proposed research was able to generate more detailed active shape models and achieve improved segmentation results.

Another optimized and automated method (Linguraru et al., 2012) uses medical image data for segmenting multiple abdominal organs, including the kidney. This method involved graph cuts and non-linear registration, enabling simultaneous segmentation from four-dimensional CT data. However, it should be noted that this particular method did not specifically focus on the segmentation of renal stones. Khalifa et al. (Khalifa et al., 2017) proposed an automated framework for 3D kidney segmentation from dynamic contrast-enhanced CT. Their approach integrated discriminative features from current and prior CT appearances into a random forest classification approach.

A study by Yang et al. (2014) introduced an automatic method for kidney segmentation based on multi-atlas image registration. Their approach involved a two-step framework that aimed to achieve accurate segmentation results. The average Dice similarity coefficient and surface-to-surface distance between their segmentation results and the reference standard were reported as 0.952 and 0.913 mm, respectively. Another technique (Erdt & Sakas, 2010) also focused on kidney segmentation. The researchers developed a deformable model-based approach incorporating local shape constraints to prevent the model from deforming into neighboring structures while allowing global shape adaptation to the data. Several other methods (Cuingnet et al., 2012; Khalifa et al., 2011; Skalski et al., 2017) have been proposed for segmentation purposes adopting the registration procedures. However, none of these studies

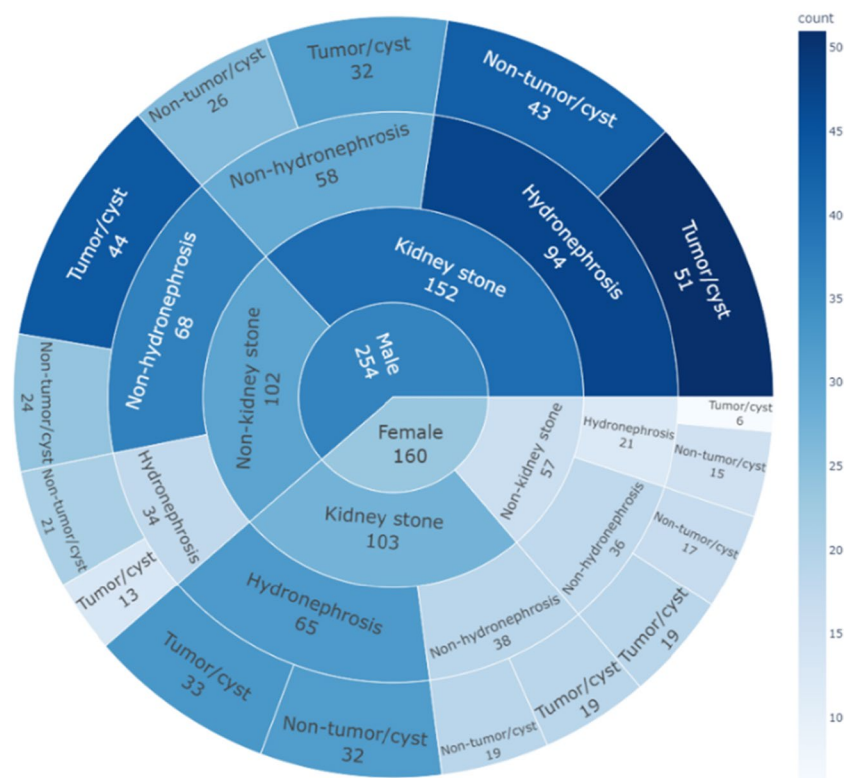
specifically addressed the combination of contrast-enhanced CT and NCT for renal system segmentation and localization. Furthermore, the commonly used method for kidney stone examination in China is CT scanning. As compared to MRI, CT scanning can clearly show the position, size, and shape of kidney stones, making it more accurate in detecting them. Additionally, CT scanning is more cost-effective and has shorter scanning times. Therefore, our intended model addresses this gap by considering both contrast-enhanced CT and NCT scans to generate 3D segmented visualizations of the renal system and renal stones.

3 Materials and Methods

3.1 Data Acquisition and Distribution

A dataset consisting of abdominal CT scans from a total of 500 patients was collected for this study. The dataset included three phases of CT scans: NCT, CTC, and CTE. The data was obtained from the First Affiliated Hospital of Guangzhou Medical University between 2014 and 2022 and adhered to the relevant data acquisition standards. Moreover, the dataset comprises data which is exclusively sourced from the GE Medical Systems Revolution CT. To ensure the quality and reliability of the imaging data, a careful curation process was performed, resulting in a final dataset of 464 patients with a total of 1392 CT scans. Figure 1 illustrates

Fig. 1 Visualization of the distribution of the curated data



the data distribution according to the patient's sex and diagnostic information.

The dataset is diverse and encompasses a wide range of medical conditions. Figure 1 provides an overview of the dataset, which includes 254 male and 160 female patients. The age of the patients varies from 9 to 90 years old. Regarding diagnoses, 61.6% of the patients in the dataset have been diagnosed with kidney stones. Among them, 59.8% are male, and 64.4% are female. Hydronephrosis, a condition characterized by the swelling of the kidney due to the accumulation of urine, affects 56.7% of the patients in the dataset, with 58.4% of males and 53.8% of females being diagnosed with it. Tumors or cysts are present in 52.4% of the patients, with 64.5% male and 35.5% female.

3.2 Three-phase Abdominal CT Annotations

The curated three-phase CT images were annotated using the 3D Slicer tool, following clinical requirements, instructions, and the need for image segmentation. The annotation process consisted of three steps. Firstly, students majoring in biomedical engineering performed the initial annotation of the curated data. Secondly, an experienced instructor thoroughly reviewed the annotated data to identify errors or formatting issues. Finally, clinical experts with domain expertise conducted a comprehensive review of the annotations. The clinical experts examined the labeled structures and made any necessary adjustments or corrections if inappropriate labels were identified. Regarding the three-phase CT labeling, specific annotations were applied to each phase. In the case of the NCT, annotations included the kidney and kidney stones. For the CTC, annotations encompassed the artery, renal cortex, and renal medulla. Similarly, the CTE phase featured annotations of the renal collection system and renal parenchyma. The detailed annotations in three-phase Abdominal CT images are depicted in Fig. 2.

3.3 Registering and Fusing 3D Kidney Annotations for Registration Evaluation

To assess the performance of the registration framework, another subset of 3D annotated data is generated by registering and fusing the kidney annotations. This subset was created using data from 50 patients, each of whom had undergone three scans: NCT, CTC, and CTE, resulting in a total of 150 CT scans. Initially, the kidney annotations from the three-phase CT data were not spatially aligned, as depicted in Fig. 3, and were in their raw form. Therefore, the annotations were fused without performing any registration. To obtain comprehensive 3D annotations, manual registration of the three-phase CT data was conducted, followed by their fusion. The 3D Slicer Tool was employed for the manual registration process. Linear

transforms and rigid transformations were applied to align the spatial positions of the three-phase CT scans.

During the fusion process, specific labels are selected from the annotated data of the three phases: NCT, CTC, and CTE. The chosen labels included the stone label from NCT, the artery label from CTC, and the renal collection system and renal parenchyma labels from CTE. In overlap between labels, we applied label priorities to determine which label should take precedence. The priority order was as follows: artery > kidney stone > renal collective system > renal medulla > renal parenchyma. The results of the fusion process, incorporating the selected labels according to their priority, are depicted in Fig. 3.

3.4 Data Split and Usage

The three-phase abdominal data is divided into training and testing sets, with approximately a 9:1 ratio. This results in 414 samples for training and 50 samples for testing. The partitioning of the data is carefully performed, taking into account the distribution of kidney stone size and kidney lesions such as cysts and hydronephrosis. This ensures that both the training and testing sets have similar distributions. To evaluate the segmentation accuracy, the 150 CT scans from the testing set were annotated using the method described in Section 3.2, which focused on three-phase abdominal CT annotations. Regarding the registration evaluation, the testing samples were selected from Section 3.3, specifically the registered and fused kidney annotations.

3.5 Methodology

The proposed framework consists of two key steps: segmentation and registration, as illustrated in Fig. 4. In the first step, the phase-specific models were applied to perform the segmentation of different structures within the abdominal CT scans of the kidney. For instance, a segmentation model was utilized to segment the kidney and kidney stone labels on the NCT scans. Similarly, the segmentation model was used to segment cortex, medulla, and artery labels on CTC scans and parenchyma, collective system labels on CTE scans.

In the second step, registration is applied by aligning the segmentation results obtained from NCT and CTE scans with the segmentation outcomes from CTC scans. This allowed us to generate a comprehensive 3D representation of the inner structures of the kidney. It is crucial to highlight that CTC serves as the fixed image, while the NCT and CTE scans act as the moving images. The following sub-sections provide a detailed explanation of each step.

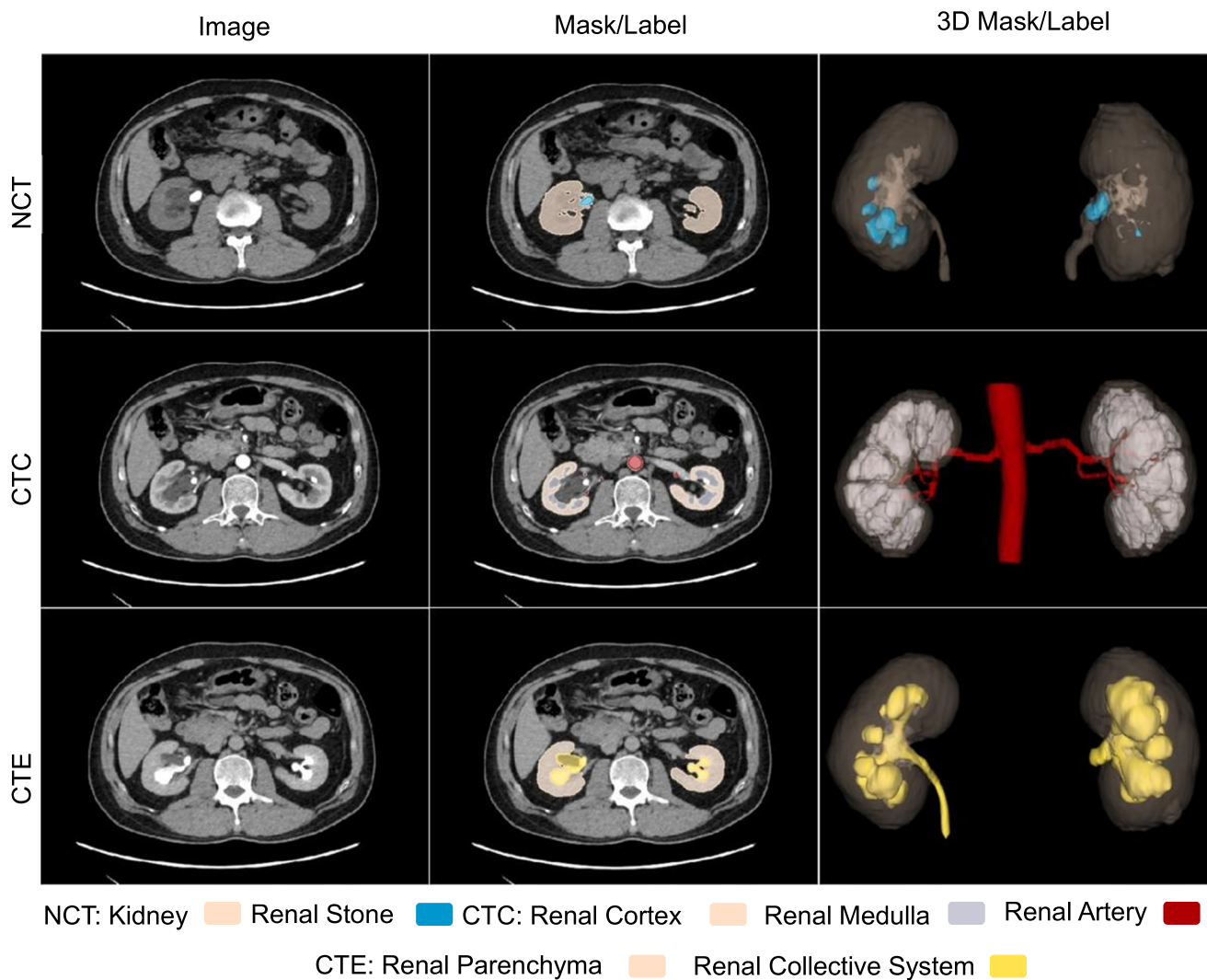


Fig. 2 Different annotations in three phases of CT images

3.5.1 Adopted Segmentation Networks

In the initial stage of the proposed framework, three state-of-the-art segmentation algorithms, 3D U-Net (Çiçek et al., 2016), Res U-Net (Zhang et al., 2018), and Swin UNETR (Hatamizadeh et al., 2021), were assessed and examined. Here is a brief overview of each segmentation network:

3D U-Net:

- The 3D U-Net architecture is designed to be trained end-to-end for segmentation.
- It consists of an encoder part that extracts relevant information from the input image and a decoder part that generates boundaries of the segmented objects.
- The network utilizes 3D convolutions, 3D max pooling, and 3D up-convolutional layers to capture spatial information.

- Batch normalization and extreme representational compression techniques are employed to accelerate convergence during training.
- It is proven to be effective for 3D medical image analysis, such as segmentation.

ResU-Net:

- Res U-Net combines the U-Net architecture with residual connections, atrous convolutions, and pyramid scene parsing pooling.
- Residual units, known for their ability to simplify training and reduce the number of parameters, are integrated into the architecture.
- Atrous convolution and pyramid scene parsing pooling mechanisms help capture contextual information.

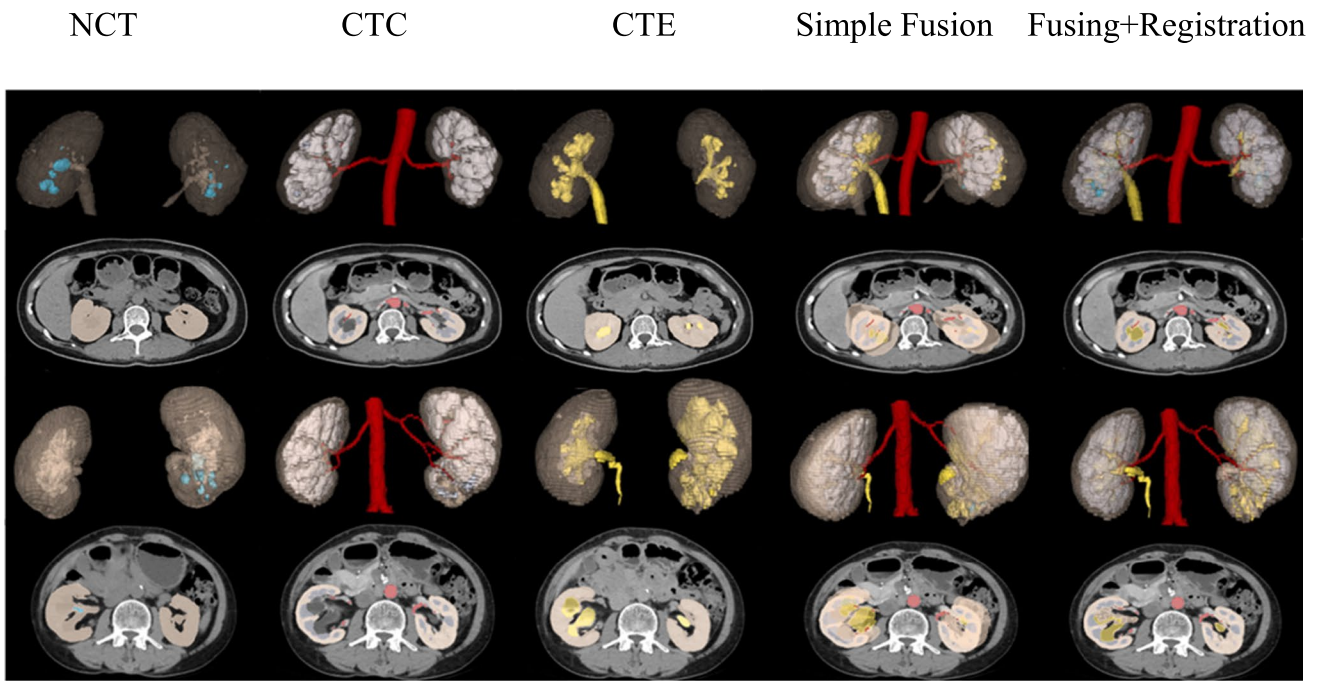


Fig. 3 Visual analysis of the comparison between simple label fusion and registered-fused annotations, where the latter procedure provided better data quality for evaluating the effectiveness of the registration technique

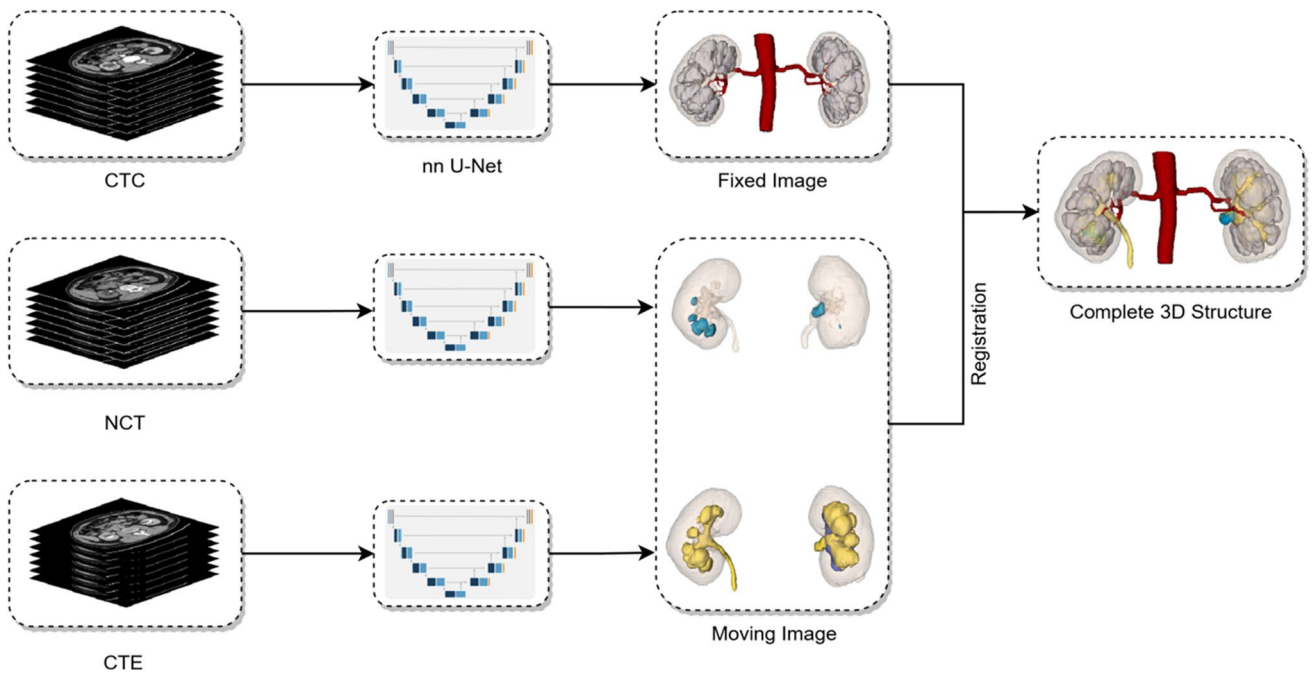


Fig. 4 Overview of our 3D framework for segmenting and registering three-phase kidney CT images

Swin UNETR:

- Swin UNETR is an extension of the U-Net architecture that incorporates a transformer-based pretraining framework designed specifically for 3D medical image tasks.
- It employs a hierarchical vision transformer as the encoder, allowing the model to learn representations of input volumes.
- Segmentation is formulated as a sequence-to-sequence prediction task.
- It employs shifted windows to extract features at multiple resolutions in the encoder.
- It utilizes skip connections to connect the decoder with the encoder at each resolution.
- It benefits from the vision transformer's ability to capture global and local representations, resulting in competitive results and accurate segmentation.

The above segmentation networks were chosen based on their performance and effectiveness in various benchmarks and medical image analysis tasks.

3.5.2 Adopted Registration Method

In this phase, a rigid registration method is implemented to tackle the challenge of kidney position variation in three-phase CT scans acquired at different time intervals. The registration approach is adopted from the ITK library (Ibanez et al., 2005). This registration method comprises six essential components: fixed image, moving image, metric (called loss function as in Section 4.2), interpolator, optimizer, and transform. The registration method inputs the segmentation results, measuring the distance between the fixed and moving images. The nearest-neighbor interpolation is used for interpolation, and a regular step gradient descent optimizer with Euler 3D transform is used for transformation.

Figure 5 illustrates each phase's initial cropping of the left and right kidneys using the segmentation outcomes. The cropped kidneys are labeled "L" and "R," respectively. The CTC phase is chosen as the fixed image, while the other two phases are treated as the moving images for registration. Registration is performed by aligning NCT.L and CTE.L with CTC.L, and NCT.R and CTE.R with CTC.R using the rigid Euler 3D transformation from the ITK library. This process achieves spatial alignment of the kidneys across all three phases. Additionally, the stone label from NCT, the artery label from CTC, and the renal collecting system and renal parenchyma from CTE are combined, resulting in comprehensive 3D segmentations of the entire kidney structure. In cases of label overlap, the predefined priority order from Section 3.3 is used to determine label priorities.

3.5.3 Training Protocol and Implementation Details

Further, the nnU-Net v1 framework (Isensee et al., 2021) is considered for training the proposed framework. The training and configuration parameters are outlined in Table 1. The "Spacing of images" parameter defines the voxel spacing or distance between slices in millimeters, with values of 0.75 mm for the first two dimensions and 1.25 mm for the third dimension. The "Hounsfield units (HU) range" specifies the range of HU values used, encompassing the 0.05th percentile to the 99.5th percentile, ensuring the inclusion of relevant image intensities. HU intensities are normalized using Z-score normalization to standardize the data. The "Data augmentation parameters" refer to the nnU-Net configuration, with default settings applied for data augmentation. The model is trained for a total of 1000 epochs. The "training patch size" indicates the dimensions of the input patches during training, with values of 96, 160, and 160 for the three dimensions, respectively. A batch size of 2 is used during training, determining the number of samples processed in each iteration. The model weights are initialized using He initialization. The "Optimizer" employed is stochastic gradient descent (SDG), with an initial learning rate of 0.01. The learning rate decay strategy implemented is ReduceLROnPlateau, which adjusts the learning rate during training based on specified conditions.

Before commencing the training, each data phase is analyzed by using the nnU-Net v1 framework. This analysis aimed to determine the optimal image preprocessing parameters for the training process. Since CT images can exhibit variations in image spacing, intensity range, and size, it is crucial to normalize each data phase under consistent conditions. This ensures improved training convergence and segmentation accuracy. Table 2 presents the data preprocessing parameters identified for NCT, CTC, and CTE data.

First, the intensity values of the CT images were adjusted to fall within the range of 0.5 and 99.5 percentile. Next, the CT images were normalized using their intensity values' mean and standard deviation. This normalization process helped to standardize the range of intensity values across the images. Finally, the normalized CT images were resampled to match the desired spacing of the target images. This preprocessing step ensured consistency in the spacing between pixels, which is important for both training and inference purposes. The training sessions were conducted on a high-performance GPU server with impressive specifications, including 1.8 terabytes of RAM, two AMD EPYC 7742 CPUs, and eight Nvidia Tesla A100 GPUs, each with 40 GB of VRAM. The operating system is Ubuntu 20.04, utilizing nnU-Net v1 and PyTorch 1.12 to ensure optimal performance and compatibility.

Fig. 5 The whole process of three-phase image registration utilizing the segmented outcome

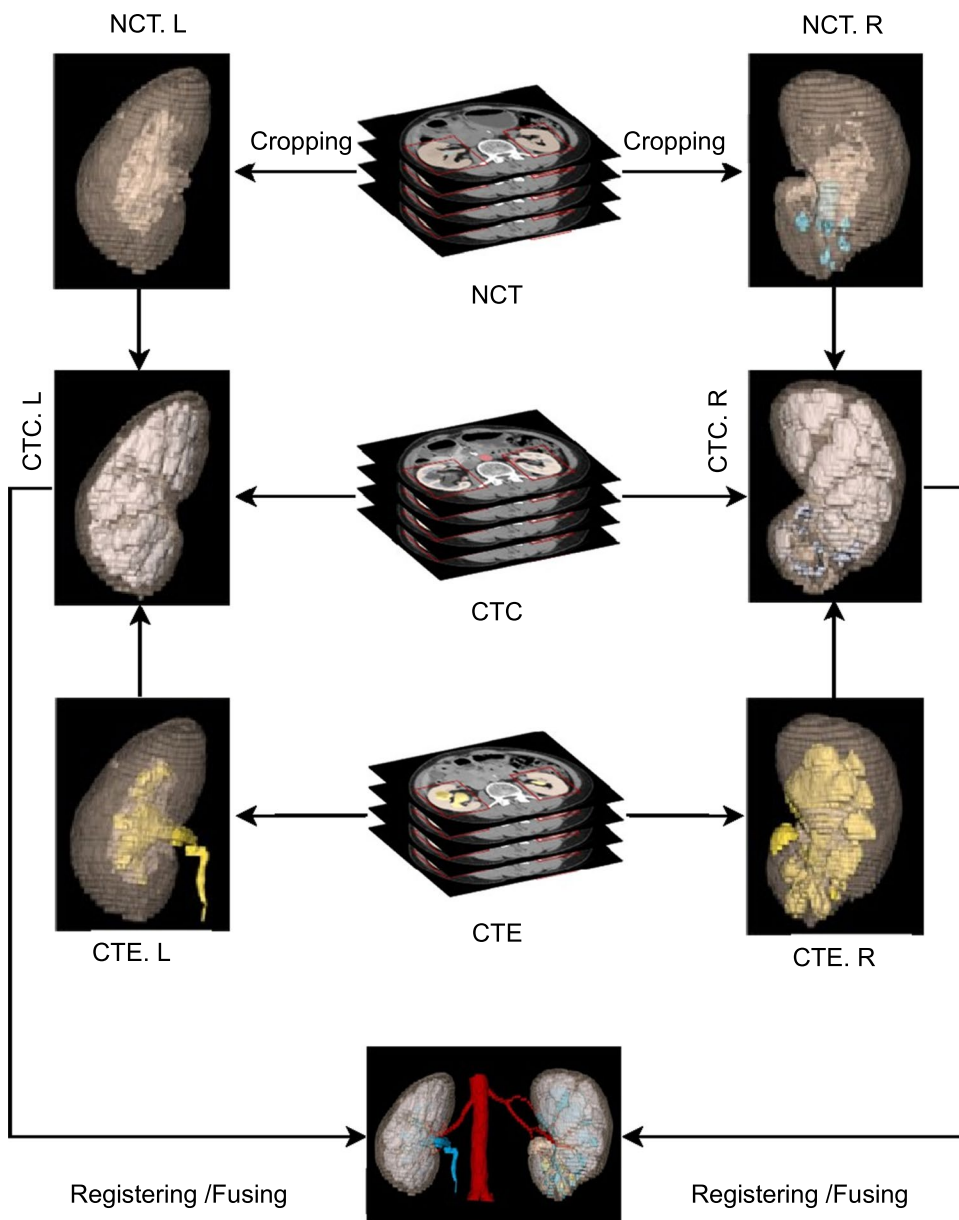


Table 1 Training parameters and configurations adopted for our training module

Hounsfield units (HU) range	From percentile 0.05 to percentile 99.5
HU intensities normalization	Z-score normalization
Data augmentation parameters	nnU-Net default configuration
Epoch	1000
Training patch size	[96, 160, 128]
Batch size	2
Weight initialization	He Initialization (He et al., 2015)
Optimizer	SGD
Learning rate	0.01
Learning rate decay	ReduceLROnPlateau

4 Experimental Results and Discussions

In this section, a comprehensive analysis is presented to demonstrate the effectiveness of proposed framework. In Section 4.1, complete experiments are conducted to compare the performance of three state-of-the-art segmentation algorithms. Subsequently, in Section 4.2, the adopted registration algorithms are evaluated using the results obtained in Section 4.1. This evaluation allowed us to measure the overall errors of the proposed framework. Two metrics are used to assess the segmentation and registration performance: the Dice Similarity Coefficient (DSC) and the Hausdorff Distance (HD). The DSC measures the spatial overlap between two samples, as demonstrated in Eq. (1), where TP denotes true positive, FP denotes false positive, and TN denotes true

Table 2 Data preprocessing parameters for three-phase data during training

Parameters	NCT Data Values	CTC Data Values	CTE Data Values
Intensity mean	32.83	110.25	133.22
Intensity standard deviation	76.50	48.12	164.09
Intensity 0.5 percentile	-37.00	11.00	-12.00
Intensity 99.5 percentile	586.00	275.00	1220.00
Target spacing	2.5, 0.754, 0.754	2.5, 0.754, 0.754	2.5, 0.748, 0.748

negative. On the other hand, the HD quantifies the distances between two sets of points, allowing us to assess the displacements between the predicted and ground truth data, as shown in Eq. (2). By considering both the DSC and HD metrics, we obtained a comprehensive understanding of the accuracy of the algorithms employed.

$$DSC = \frac{2TP}{2TP+FP+TN} \quad (1)$$

$$HD(X, Y) = \max \left\{ \max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(x, y) \right\} \quad (2)$$

4.1 Segmentation Outcomes

In this section, three state-of-the-art segmentation algorithms, namely ResU-Net, 3D U-Net, and Swin UNETR are compared using the annotated three-phase CT data mentioned above. The purpose of this comparison was to evaluate the performance of these algorithms based on various metrics and visualize the segmentation results obtained by each algorithm.

According to the results presented in Table 3, all three algorithms demonstrated competitive performance for the kidney label during the NCT phase, achieving a dice value of approximately 95%. However, when it came to the stone label, the ResU-Net algorithm outperformed both the 3D U-Net and Swin UNETR algorithms by a significant margin. Specifically, the ResU-Net achieved a dice value that was

approximately 7% higher than the 3D U-Net and 10% higher than the Swin UNETR for the stone label. In the CTC phase, both the ResU-Net and 3D U-Net algorithms outperformed the Swin UNETR algorithm. Specifically, for the artery and medulla labels, the ResU-Net and 3D U-Net achieved dice values that were approximately 2% higher than the Swin UNETR algorithm.

When analyzing the CTE phase, the results indicate that both the ResU-Net and 3D U-Net algorithms outperformed the Swin UNETR algorithm. Specifically, the ResU-Net and 3D U-Net achieved approximately 3% higher dice values for the collecting system and 1% higher dice values for the parenchyma compared to the Swin UNETR. This indicates that the ResU-Net and 3D U-Net algorithms exhibited superior segmentation performance in these areas compared to the Swin UNETR algorithm. Figure 6 provides visual representations of the segmented results produced by all three algorithms. It can be observed that all three methods yielded very similar segmentation outcomes for the kidney label but showed noticeable variations in the segmentation of kidney stones.

This can be attributed to the challenge of accurately segmenting small kidney stones, which may result in similar overall segmentation outcomes despite significant differences in dice values. Further observations by comparing the predictions of the three algorithms to the 3D ground truth revealed that the Swin UNETR algorithm performed poorly in predicting the artery compared to both the ResU-Net and 3D U-Net algorithms.

Table 3 The segmentation evaluation metrics of each algorithm on three-phase CT data

Phase	Annotations/ Labels	ResU-Net		3D U-Net		Swin UNETR	
		Dice (mean \pm std) %	HD (mean \pm std) mm	Dice (mean \pm std) %	HD (mean \pm std) mm	Dice (mean \pm std) %	HD (mean \pm std) mm
NCT	Kidney	94.44 \pm 6.89	19.82 \pm 8.72	94.70 \pm 6.18	17.55 \pm 35.08	95.21 \pm 4.78	17.76 \pm 5.27
	Stone	87.69 \pm 21.47	10.45 \pm 22.65	80.45 \pm 30.06	15.15 \pm 3.38	77.06 \pm 34.92	31.93 \pm 1.27
CTC	Artery	80.14 \pm 5.15	6.34 \pm 4.62	80.22 \pm 5.11	6.58 \pm 4.99	77.53% \pm 6.35	10.96 \pm 4.27
	Cortex	93.39 \pm 2.38	1.55 \pm 4.52	93.23 \pm 2.42	1.57 \pm 4.61	92.11% \pm 3.47	11.24 \pm 0.47
	Medulla	89.26 \pm 4.29	2.59 \pm 8.46	89.11 \pm 4.30	2.48 \pm 7.64	87.76% \pm 4.70	9.06 \pm 28.88
CTE	Collecting system	87.08 \pm 15.75	6.85 \pm 14.49	87.34 \pm 15.93	6.23 \pm 13.23	84.04 \pm 14.13	22.60 \pm 6.70
	Parenchyma	95.15 \pm 6.76	7.34 \pm 23.65	94.75 \pm 7.54	6.34 \pm 23.30	93.96 \pm 6.45	28.14 \pm 1.30

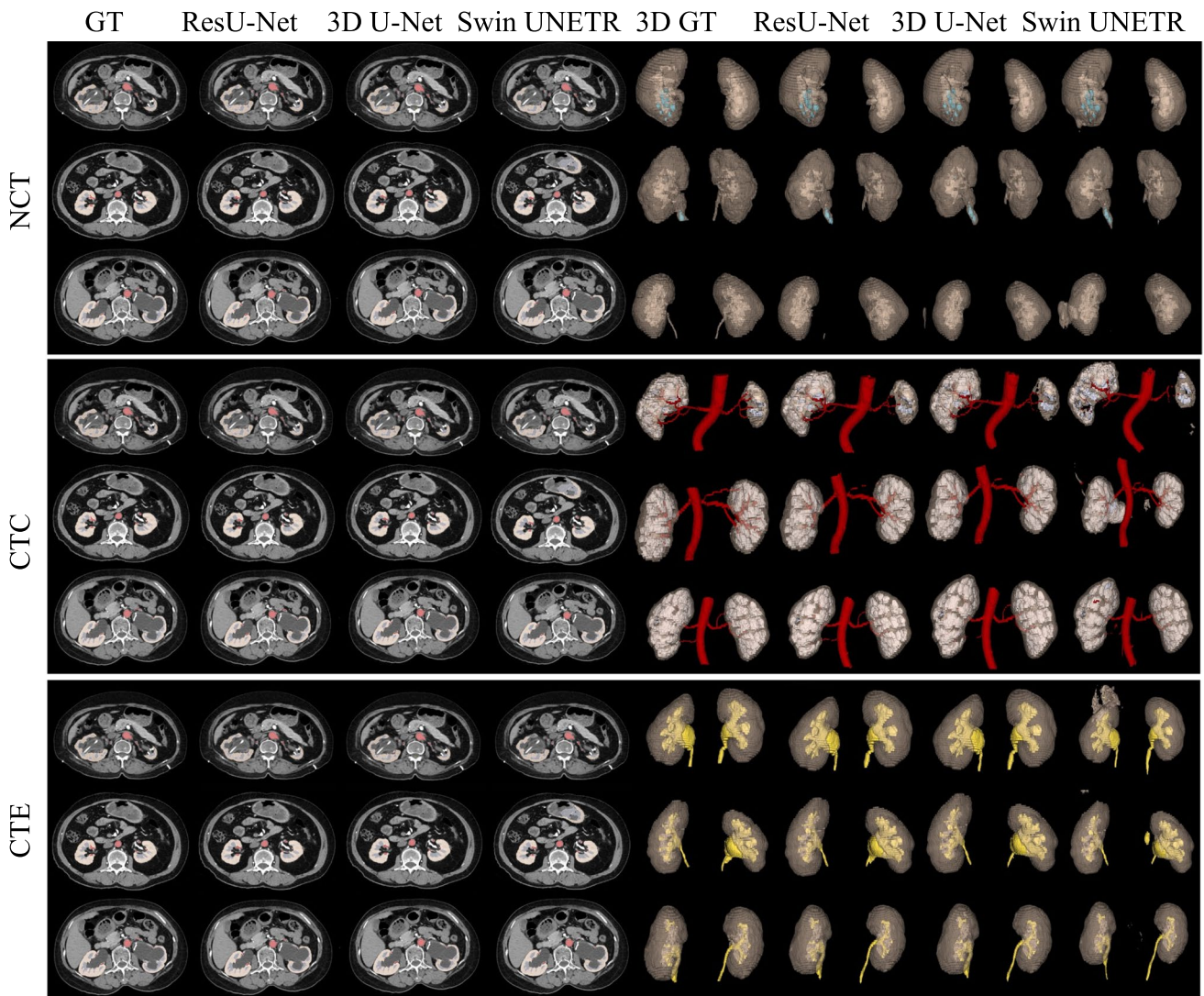


Fig. 6 Comparative analysis of ground truth (GT), segmentation results of ResU-Net, 3DU-Net, and Swin UNETR in terms of 2D and 3D views for three-phase CT images

In conclusion, when considering the segmentation of three-phase CT data, the ResU-Net algorithm outperforms both the Swin UNETR and 3D U-Net algorithms. While the Swin UNETR and 3D U-Net may perform well on certain labels, the ResU-Net algorithm consistently demonstrates competitive results across multiple labels. It achieves competitive performance for kidney, artery, and collecting system labels and stands out as the best algorithm for segmenting stones, cortex, medulla, and parenchyma. Therefore, based on the evaluation of various labels, the ResU-Net emerges as the most favorable choice for three-phase CT data segmentation.

4.2 Final 3D Registered Outcomes

This section discusses employing segmentation outcomes obtained from ResU-Net (as mentioned in Section 4.1) for the registration task, as it demonstrated the best segmentation performance compared to the other algorithms. A series of experiments were conducted to validate the effectiveness of chosen registration method. Furthermore, the adopted baseline registration method is extended by introducing three new strategies.

The first method involved directly registering the segmentation outcomes without performing any kidney cropping.

The second strategy closely resembled the baseline registration method, but the entire CT images were used instead of relying solely on segmentation outcomes. However, the segmentation outcomes were still utilized to identify and crop the relevant regions for registration, similar to the baseline approach. In contrast, the third approach relied solely on the CT images for registration without incorporating cropping steps.

Two loss functions, mutual information (MI) and normalized cross-correlation (NCC), were employed from the ITK library to assess the registration accuracy of the four methods. MI, introduced by Mayes et al. (Maes et al., 1997), calculates joint gray value histograms and is commonly utilized in medical image registration due to its versatility. Conversely, NCC measures cross-correlation at the pixel level and has a well-defined minimum. However, it is important to note that NCC can only be applied within the same imaging modality.

According to Table 4, the adopted registration method using MI achieves the best performance compared to other approaches. This success can be attributed to the following reasons. Firstly, the “Separate” method, where the left

and right kidneys are registered separately, significantly improves the registration accuracy compared to the “Whole” method. Secondly, utilizing segmentation outcomes for registration across different modalities also contributes to a significant improvement in accuracy. Figure 7 provides 3D visualizations that demonstrate the registration outcomes of each approach with different parameter settings derived from ResU-Net segmentation results. The proposed framework exhibits excellent alignment with the ground truth, surpassing other approaches in terms of accuracy. Table 5 presents the evaluation metrics for registration performance based on the segmentation results. To facilitate comprehension, the following notations are used.

MI is denoted as loss_fn1.

NCC is denoted as loss_fn2.

Registering the left and right kidneys separately is denoted as Separate.

Registering the left and right kidneys together is denoted as Whole.

Table 4 Proposed strategies according to the baseline registration approach

Methods	Input to the Registration Module	Description
Adopted Baseline Registration Method (Maes et al., 1997)	Segmented outcomes	Registering the left and right kidneys separately
Baseline ExtendedMethod 1	Segmented outcomes	Registering left and right kidneys together
Baseline ExtendedMethod 2	Entire CT image/ Segmented outcomes	Registering the left and right kidneys separately
Baseline ExtendedMethod 3	Entire CT images	Register the left and right kidneys together

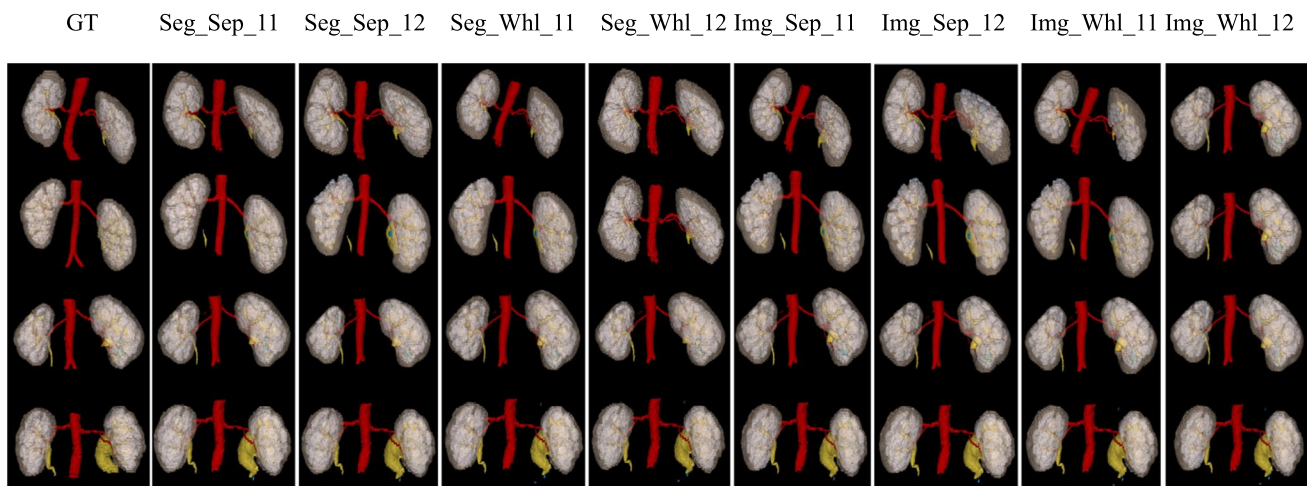


Fig. 7 Final 3D visualization of registration results based on ResU-Net segmentation

Table 5 Registration performance evaluation metrics based on segmentation outcomes

Modality	Methods	ResU-Net		3D U-Net		Swin UNETR	
		Dice (mean \pm std) %	HD (mean \pm std) mm	Dice (mean \pm std) %	HD (mean \pm std) mm	Dice (mean \pm std) %	HD (mean \pm std) mm
Entire CT Image	Separate + loss_f1	89.22 \pm 11.46	5.28 \pm 7.58	89.05 \pm 11.52	5.56 \pm 8.81	88.17 \pm 11.40	9.88 \pm 15.66
	Separate + loss_f2	87.20 \pm 12.04	6.18 \pm 7.95	87.48 \pm 11.97	6.24 \pm 8.97	86.42 \pm 11.96	10.61 \pm 15.59
	Whole + loss_f1	81.20 \pm 20.56	12.98 \pm 22.65	80.92 \pm 20.72	13.70 \pm 24.54	80.31 \pm 19.28	36.75 \pm 47.78
	Whole + loss_f2	81.25 \pm 18.33	11.58 \pm 19.34	80.99 \pm 18.40	12.28 \pm 21.30	79.89 \pm 17.76	35.71 \pm 47.48
Segmentation Outcomes	Separate + loss_f1	93.28 \pm 12.84	4.15 \pm 8.58	93.19 \pm 12.73	4.31 \pm 9.61	92.44 \pm 12.25	8.66 \pm 15.21
	Separate + loss_f2	92.92 \pm 12.96	4.50 \pm 8.99	93.23 \pm 12.77	4.26 \pm 9.79	92.34 \pm 12.37	8.92 \pm 15.94
	Whole + loss_f1	89.09 \pm 13.64	7.71 \pm 16.53	88.57 \pm 14.05	8.38 \pm 18.62	87.87 \pm 12.47	33.02 \pm 48.66
	Whole + loss_f2	90.52 \pm 12.54	7.19 \pm 17.37	89.99 \pm 12.76	8.02 \pm 8.02	88.95 \pm 12.09	32.50 \pm 49.00

5 Conclusion

The proposed framework integrates deep-learning algorithms with advanced segmentation and registration techniques to achieve precise visualization and localization of the intricate anatomical structures of the kidneys. By effectively addressing the challenges associated with kidney stones and the complex nature of the renal system, this framework provides valuable assistance in diagnosing and planning surgical interventions for renal diseases. The evaluation of adopted segmentation models has shown their effectiveness in accurately delineating kidney structures. The ensemble model achieved notable segmentation results, with Swin UNETR reaching a 95.21% Dice score for kidney segmentation and ResU-Net achieving 87.69% for stone segmentation. Registration successfully aligned kidneys across all phases, generating detailed 3D visualizations.

The registration process within our proposed framework was instrumental in aligning the segmented results obtained from various CT phases, allowing for the reconstruction of the complete 3D inner structure of the kidney. By fusing the information from NCT and contrast-enhanced CT scans, the framework successfully generated comprehensive 3D visualizations that encompassed stone labels, artery labels, and labels for the renal collection system and parenchyma. These final comprehensive 3D visuals offer a holistic representation of the anatomical features of the kidney, enabling improved understanding and analysis of renal diseases.

Furthermore, the utilization of a carefully curated dataset, encompassing a wide range of CT scans from diverse patients, significantly contributed to the robustness and generalizability of the framework. The inclusion of CT scans depicting various medical conditions, including kidney stones, hydronephrosis, and tumors/cysts enhanced the framework's ability to handle different pathological

scenarios. Overall, the proposed framework for 3D kidney reconstruction offers a promising solution for achieving accurate and detailed visualization of renal structures and abnormalities. Its potential to assist in diagnosis, treatment planning, and surgical interventions in renal diseases makes it a valuable preoperative planning tool in percutaneous nephrolithotomy for clinicians.

One limitation of our framework is its slow speed, mainly due to the time-consuming registration process. On our server, the registration procedure took an average of 44.5 s per testing sample, considering a total of 50 testing samples. To address this issue, we intend to replace the conventional registration algorithm with a deep learning-based approach to reduce the registration time significantly. We also plan to further improve and refine the framework by expanding the dataset and conducting extensive clinical validations. These efforts are essential to ensure the integration of the proposed framework into routine clinical practice, providing clinicians with a reliable and efficient tool for accurate kidney stone diagnosis and preoperative planning in percutaneous nephrolithotomy.

Acknowledgements We express our gratitude to our colleague and medical students who did the CT data annotations.

Authors' contributions Data curation, Yang Liu, Haseeb Hassan, Dan Li and Jun Liu; Formal analysis, Chuda Xiao, Yang Liu and Haoyu Li; Funding acquisition, Bingding Huang; Methodology, Zhuo Chen and Chuda Xiao; Project administration, Weiguo Xie and Wen Zhong; Resources, Yang Liu, Dan Li, Weiguo Xie and Wen Zhong; Software, Zhuo Chen and Chuda Xiao; Supervision, Wen Zhong and Bingding Huang; Validation, Haseeb Hassan, Jun Liu and Haoyu Li; Writing – original draft, Zhuo Chen, Haseeb Hassan and Bingding Huang; Writing – review & editing, Bingding Huang. All the authors read and approved the final manuscript in its current state.

Funding This study was supported by the Project of the Educational Commission of Guangdong Province of China (No. 2022ZDJS113)

and the School-Enterprise Cooperation Fund provided by Wuerzburg Dynamics Inc. to the Weiding Joint Laboratory of Medical Artificial Intelligence, Shenzhen Technology University.

Data Availability The CT datasets generated and analyzed during the current study are not publicly available due the fact that we assured confidentiality but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate This work was a retrospective study, which was granted a waiver for ethical approval by the Ethics Committee of the First Affiliated Hospital of Guangzhou Medical University,

Consent for publication Not applicable.

Competing interest The authors declare that they have no conflict of interest or competing interests that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ansari, M. Y., Abdalla, A., Ansari, M. Y., Ansari, M. I., Malluhi, B., Mohanty, S., Mishra, S., Singh, S. S., Abinahed, J., & Al-Ansari, A. (2022). Practical utility of liver segmentation methods in clinical surgeries and interventions. *BMC Medical Imaging*, 22(1), 1–17.
- Baygin, M., Yaman, O., Barua, P. D., Dogan, S., Tuncer, T., & Acharya, U. R. (2022). Exemplar Darknet19 feature generation technique for automated kidney stone detection with coronal CT images. *Artificial Intelligence in Medicine*, 127, 102274.
- Chen, X., Summers, R. M., Cho, M., Bagci, U., & Yao, J. (2012). An automatic method for renal cortex segmentation on CT images: Evaluation on kidney donors. *Academic Radiology*, 19(5), 562–570.
- Çiçek Ö., Abdulkadir A., Lienkamp S. S., Brox T., Ronneberger O. (2016). 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016: 19th International Conference, Athens, Greece, October 17–21, 2016, Proceedings, Part II 19: 2016*: Springer; 424–432. https://doi.org/10.1007/978-3-319-46723-8_49
- Cuingnet R., Prevost R., Lesage D., Cohen L. D., Mory B., Ardon R. (2012). Automatic detection and segmentation of kidneys in 3D CT images using random forests. In: *The 15th International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI 2012: 2012*: Springer Berlin Heidelberg; 66–74. https://doi.org/10.1007/978-3-642-33454-2_9
- Elton, D. C., Turkbey, E. B., Pickhardt, P. J., & Summers, R. M. (2022). A deep learning system for automated kidney stone detection and volumetric segmentation on noncontrast CT scans. *Medical Physics*, 49(4), 2545–2554.
- Erdt M., & Sakas G. (2010). Computer aided segmentation of kidneys using locally shape constrained deformable models on CT images. In *Medical Imaging 2010: Computer-Aided Diagnosis* (Vol. 7624, pp. 356–363). SPIE.
- Flores-Mireles, A. L., Walker, J. N., Caparon, M., & Hultgren, S. J. (2015). Urinary tract infections: Epidemiology, mechanisms of infection and treatment options. *Nature Reviews Microbiology*, 13(5), 269–284.
- Hameed, B. Z., Shah, M., Naik, N., Rai, B. P., Karimi, H., Rice, P., Kronenberg, P., & Somani, B. (2021). The ascent of artificial intelligence in endourology: A systematic review over the last 2 decades. *Current Urology Reports*, 22, 1–18.
- Hatamizadeh A., Nath V., Tang Y., Yang D., Roth H. R., & Xu D. (2022). Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In *International MICCAI Brainlesion Workshop* (pp 272–284). Springer International Publishing.
- He K., Zhang X., Ren S., & Sun J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, (pp. 1026–1034). <https://doi.org/10.1109/ICCV.2015.123>.
- Ibanez, L., Schroeder, W., Ng, L., & Cates, J. (2005). *The ITK software guide*. Kitware, Inc Clifton Park.
- Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2), 203–211.
- Kakhandaki, N., & Kulkarni, S. B. (2023). Classification of brain MR images based on bleed and calcification using ROI cropped U-Net segmentation and ensemble RNN classifier. *International Journal of Information Technology*, 15(6), 3405–3420.
- Khalifa F., Gimel'farb G., El-Ghar M. A., Sokhadze G., Manning S., McClure P., Ouseph R., & El-Baz A. (2011). A new deformable model-based segmentation approach for accurate extraction of the kidney from abdominal CT images. In *2011 18th IEEE International Conference on Image Processing* (pp. 3393–3396).
- Khalifa F., Soliman A., Elmaghraby A., Gimel'farb G., El-Baz A. (2017). 3D kidney segmentation from abdominal images using spatial-appearance models. *Computational and mathematical methods in medicine* 2017. <https://doi.org/10.1155/2017/9818506>
- Korfiatis, P., Denic, A., Edwards, M. E., Gregory, A. V., Wright, D. E., Mullan, A., Augustine, J., Rule, A. D., & Kline, T. L. (2022). Automated segmentation of kidney cortex and medulla in CT images: A multisite evaluation study. *Journal of the American Society of Nephrology*, 33(2), 420–430.
- Li, D., Xiao, C., Liu, Y., Chen, Z., Hassan, H., Su, L., Liu, J., Li, H., Xie, W., & Zhong, W. (2022). Deep segmentation networks for segmenting kidneys and detecting kidney stones in unenhanced abdominal CT images. *Diagnostics*, 12(8), 1788.
- Li X., Chen X., Yao J., Zhang X., Tian J. (2011). Renal cortex segmentation using optimal surface search with novel graph construction. In: *Medical image computing and computer-assisted intervention: MICCAI International Conference on Medical Image Computing and Computer-Assisted Intervention: 2011*: NIH Public Access; 387. https://doi.org/10.1007/978-3-642-23626-6_48
- Lim A. K. (2014). Diabetic nephropathy—complications and treatment. *International journal of nephrology and renovascular disease*:361–381. <https://doi.org/10.2147/IJNRD.S40172>
- Lin, D.-T., Lei, C.-C., & Hung, S.-W. (2006). Computer-aided kidney segmentation on abdominal CT images. *IEEE Transactions on Information Technology in Biomedicine*, 10(1), 59–65.

- Linguraru, M. G., Pura, J. A., Pamulapati, V., & Summers, R. M. (2012). Statistical 4D graphs for multi-organ abdominal segmentation from multiphase CT. *Medical Image Analysis, 16*(4), 904–914.
- Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., & Suetens, P. (1997). Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging, 16*(2), 187–198.
- Murray, B. O., Flores, C., Williams, C., Flusberg, D. A., Marr, E. E., Kwiatkowska, K. M., Charest, J. L., Isenberg, B. C., & Rohn, J. L. (2021). Recurrent urinary tract infection: A mystery in search of better model systems. *Frontiers in Cellular and Infection Microbiology, 11*, 691210.
- Parakh, A., Lee, H., Lee, J. H., Eisner, B. H., Sahani, D. V., & Do, S. (2019). Urinary stone detection on CT images using deep convolutional neural networks: evaluation of model performance and generalization. *Radiology: Artificial Intelligence, 1*(4), e180066.
- Piao, N., Kim, J.-G., & Park, R.-H. (2015). Segmentation of cysts in kidney and 3-D volume calculation from CT images. *International Journal of Computer Graphics & Animation, 5*(1), 1.
- Schedl, A. (2007). Renal abnormalities and their developmental origin. *Nature Reviews Genetics, 8*(10), 791–802.
- Shim, H., Chang, S., Tao, C., Wang, J. H., Kaya, D., & Bae, K. T. (2009). Semiautomated segmentation of kidney from high-resolution multidetector computed tomography images using a graph-cuts technique. *Journal of Computer Assisted Tomography, 33*(6), 893–901.
- Skalski, A., Heryan, K., Jakubowski, J., & Drewniak, T. (2017). Kidney segmentation in ct data using hybrid level-set method with ellipsoidal shape constraints. *Metrology and Measurement Systems, 24*(1), 101–112.
- Spiegel, M., Hahn, D. A., Daum, V., Wasza, J., & Hornegger, J. (2009). Segmentation of kidneys using a new active shape model generation technique based on non-rigid image registration. *Computerized Medical Imaging and Graphics, 33*(1), 29–39.
- Takazawa, R., Kitayama, S., Uchida, Y., Yoshida, S., Kohno, Y., & Tsujii, T. (2018). Proposal for a simple anatomical classification of the pelviciceal system for endoscopic surgery. *Journal of Endourology, 32*(8), 753–758.
- Tang, Y., Jackson, H. A., De Filippo, R. E., Nelson, M. D., Jr., & Moats, R. A. (2010). Automatic renal segmentation applied in pediatric MR Urography. *Int J Intell Inf Process, 1*(1), 12–19.
- Tang Y., Gao R., Lee H. H., Xu Z., Savoie B. V., Bao S., Huo Y., Fogo A. B., Harris R., & de Caestecker M. P. (2021). Renal cortex, medulla and pelviciceal system segmentation on arterial phase CT images with random patch-based networks. In *Medical Imaging 2021: Image Processing* (Vol. 11596, pp. 379–386). SPIE.
- Thong, W., Kadoury, S., Piché, N., & Pal, C. J. (2018). Convolutional networks for kidney segmentation in contrast-enhanced CT scans. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 6*(3), 277–282.
- Yang, B., Veneziano, D., & Somani, B. K. (2020). Artificial intelligence in the diagnosis, treatment and prevention of urinary stones. *Current Opinion in Urology, 30*(6), 782–787.
- Yang G., Gu J., Chen Y., Liu W., Tang L., Shu H., & Toumoulin C. (2014). Automatic kidney segmentation in CT images based on multi-atlas image registration. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 5538–5541). IEEE.
- Yildirim, K., Bozdag, P. G., Talo, M., Yildirim, O., Karabatak, M., & Acharya, U. R. (2021). Deep learning model for automated kidney stone detection using coronal CT images. *Computers in Biology and Medicine, 135*, 104569.
- Zhang, Z., Liu, Q., & Wang, Y. (2018). Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters, 15*(5), 749–753.
- Zisman, A. L., Evan, A. P., Coe, F. L., & Worcester, E. M. (2015). Do kidney stone formers have a kidney disease? *Kidney International, 88*(6), 1240–1249.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Zhuo Chen received his B.Sc. degree in Information and Computational Science from Civil Aviation University of China. He is currently an Intermediate Algorithm Engineer at Wuerzburg Dynamics Inc. Shenzhen. He works on the application of artificial intelligence in the medical field, especially surgical robots.

Chuda Xiao received his MSc. degree from in computer science from Shenzhen University. Currently he is employed at Wuerzburg Dynamics Inc. Shenzhen as an algorithm engineer. His research interests include medical image analysis and computer vision.

Yang Liu Hold a PhD degree in Urology from Guangzhou Medical University. He is currently an Urologist-in-charge at the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China. His research interests include urological disease and percutaneous nephrolithotomy.

Haseeb Hassan hold a MSc. degree in Computer Science from the International Islamic University, Islamabad, Pakistan, and PhD degree in computer science from Anhui University, Hefei, China. Presently, Dr. Hassan is a Research Associate Fellow at the College of Health Science and Environmental Engineering at Shenzhen Technology University. His research focus on cutting-edge domains such as computer vision, interdisciplinary AI applications, and Medical Imaging Analysis.

Dan Li graduated from Shenzhen University with a MSc. Degree in computer science. She currently works as an algorithm engineer at Wuerzburg Dynamics Inc. Shenzhen. Her research focus on the application of deep learning algorithms in medical imaging.

Jun Liu graduated with a B.Sc. degree in Biomedical Engineering from Nanchang Hangkong University. Currently he is a medical device engineer at Wuerzburg Dynamics Inc. Shenzhen. He mainly works medical imaging and surgical robot technology.

Haoyu Li is currently a master student at Shenzhen Technology University. He mainly focuses on developing robust algorithms and models to accurately segment CT images.

Weiguo Xie received the PhD degree in Biomedical Engineering from University of Bern, Switzerland. He is the founder of Wuerzburg Dynamics Inc. Shenzhen. His research interests include bioengineering, surgery robots, medical device development.

Wen Zhong received his MSc. and PhD degrees from Guangzhou Medical University and Southern Medical University. Currently he is a Chief Physician of Urology at the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China. His research interests include the etiology and prevention of urinary stones.

Bingding Huang received the PhD degree in Computer Science from Dresden Technical University, Germany. He is currently a Distinguished Professor at the College of Big Data and Internet, Shenzhen Technology University, where he is also the head of Department of Data Science and Big Data Technology. His research interests include bioinformatics, precision medicine, artificial intelligence, medical imaging, and big data technology and applications.