Towards 3D-Dense Ultrasound Image Simulation from 2D CT Scans for Ultrasound-Guided Percutaneous Nephrolithotomy: A Progressive Training Approach from Basic to Advanced Simulator Complexity

Sathiyamoorthy Selladurai[a,*], James Watterson[b], Rebecca Hibbert[c] and Carlos Rossa.[a]

[a]Carleton University,, Department of Systems and Computer Engineering, Ottawa, ON, Canada
[b]University of Ottawa, , Ottawa, ON, Canada
[c]Mayo Clinic, Department of Radiology, Rochester, MN, USA

ARTICLE INFO

Keywords:
3D Ultrasound Imaging
Simualtion
PCNL
CT
Simulator

ABSTRACT

Ultrasound-guided renal access for percutaneous nephrolithotomy (usPCNL) is a common technique used to remove large kidney stones through an incision in the patient's back. The procedure requires a high level of dexterity and is associated with a steep learning curve. Advanced simulation tools can enhance clinical training and provide interventionists with a platform to rehearse the operation with pre-operative patient data. This paper presents the first step towards the development of such a simulator. We propose a new framework and algorithms to generate volumetric ultrasound images from preoperative 2D computed tomography (CT). First, successive CT scans are interpolated to augment the dataset and increase spatial resolution. Each scan is then converted into an ultrasound image based on principles of linear acoustics and spatial impulse response. These ultrasound images are then combined to form two volumetric images, one derived from the original sparse CT scans, and one with the denser data. New images can then be formed along arbitrary imaging planes not captured in the original CT data. The obtained images are compared with real images acquired experimentally, and further evaluated quantitatively. We demonstrate that the peak signal-to-noise ratio (PSNR) in the simulated images shows a significant enhancement for the denser CT scan datasets, with an improvement of 34.8% for Dataset 1 and 46.5% for Dataset 2. The second part of the paper proposes a multi-level approach where kidney stones and kidney contours are segmented from CT scans and fused on top of ultrasound images. Four levels of assistance are proposed, ranging from full anatomical information, to no assistance with increasing levels of attenuation, increasing complexity and realism to build the trainee's expertise progressively. This work lays the foundation for developing the first usPCNL simulator, which could enhance training and procedural outcomes in this complex medical procedure.

1. Introduction

Percutaneous nephrolithotomy (PCNL) is a minimally invasive procedure to remove large kidney stones that cannot be treated with other methods such as extracorporeal shock wave lithotripsy or ureteroscopy. In this procedure, the interventionist accesses the kidney through a small incision on the patient's back and uses a nephroscope to remove the stones under image guidance Türk, Petrik, Sarica, Seitz, Skolarikos, Straub and Knoll (2016); Metzler, Holt and Harper (2021); Labate, Modi, Timoney, Cormio, Zhang, Louie, Grabe and de la Rosette (2011). Fluoroscopy has been the preferred imaging modality as it provides real-time guidance and clear visualization of instruments and the urinary tract. However, it exposes patients and staff to ionizing radiation Yang, Wen, Chen and Chen (2019); Kumari, Kumar, Wadhwa, Aron, Gupta and Dogra (2006). Ultrasound-guided PCNL (usPCNL) has emerged as an alternative modality as it is radiation-free and allows for better visualization of soft tissue structures Pu, Wang, Tang, Yuan, Li, Bai, Wang, Wei and Han (2015). In addition, usPCNL can be performed with patients in the supine position, which is more comfortable for obese patients and those with cardiopulmonary disorders or skeletal deformities El-Shaer, Abdel-Lateef, Torky, Elshaer et al. (2019).

Despite these advantages, the adoption of usPCNL has been limited by a steep learning curve associated with navigating surgical instruments while interpreting real-time ultrasound(US) images. This learning curve is often attributed to the level of dexterity required to coordinate the US probe and the tool, the limited spatial resolution of ultrasound images, and the presence of image artefacts Sahan, Sarilar, Savun, Caglar, Erbin and Ozgor (2020); Beiko, Razvi, Bhojani, Bjazevic, Bayne, Tzou, Stoller and Chi (2019). To smoothen the learning curve, traditional training methods often involve practising the procedure on phantoms or cadavers, which may not fully replicate real-world scenarios. As a result, there is a critical need for advanced simulation tools that offer realistic and accessible training opportunities. Canalichio, Berrondo and Lendvay (2020); Ferraguti, Farsoni and Bonfè (2022); Farcas, Reynolds and Lee (2021); Doizi and Koskas (2022)

High-fidelity simulators replicate anatomical and procedural complexities, providing trainees with a safe and controlled environment to practice and hone their skills. These simulators can also offer real-time feedback and performance metrics, which are essential for effective skill development Ajemba, Ikwe and Iroanya (2024); Hamacher,

*Sathiyamoorthy Selladurai
SelladuraiSathiyamoo@cunet.carleton.ca (S. Selladurai)
ORCID(s): 0000-0001-9475-5997 (S. Selladurai)

Whangbo, Kim and Chung (2018); Gomaa, Grafton-Clarke, Saratzis and Davies (2023); Ritchie, Pacilli and Nataraja (2023). An example is Marion Surgical's K181 fluoroscopy-based PCNL simulator - a virtual reality platform that emulates PCNL procedures using a virtual reality headset and a haptic device. The simulator can import real, pre-operative Computed Tomography (CT) images of the patient to simulate real-time fluoroscopy images, providing opportunities not only for skills development but also surgical rehearsal. To date, no commercially available PCNL simulator uses ultrasound image guidance Sainsbury, Lacki, Shahait, Goldenberg, Baghdadi, Cavuoto, Ren, Green, Lee, Averch et al. (2020); Rassweiler and Teber (2016).

The first step towards implementing a usPCNL simulator would be to recreate realistic volumetric ultrasound images of the patient, either through pre-operative ultrasound scans or by simulating such images from routinely used CT scans. A collection of 2D images can be converted into a volumetric image that can be sliced and displayed in real time as the trainee controls the position and orientation of a virtual ultrasound probe in the simulator. This paper presents the first step towards the development of a usPCNL simulator, that is, a method to simulate real-time ultrasound images along arbitrary imaging direction from pre-operative parallel 2D CT scans. The image simulator employs a multi-level approach with increasing levels of complexity and assistance, tailored to progressively build the trainee's expertise. It integrates segmented images of the kidney contour and the location of any kidney stones obtained from CT scans. By overlaying these segmented images on top of real-time US images, the simulator offers a comprehensive visual representation that allows the trainee to focus on one aspect of their training at a time. The proposed progressive training approach with varying levels of assistance is structured as follows:

- **Level 1**: Kidney stones and kidney contours are segmented and overlaid on top of real-time ultrasound images, allowing trainees to focus on tool steering;
- **Level 2**: Segmented kidney stone images are removed, emphasizing anatomical recognition and navigation without the presence of stones;
- **Level 3**: Segmented kidney stones are overlaid on real-time US images, but segmented kidney contours are not provided, challenging trainees to identify and navigate towards the stones without the aid of kidney contours;
- **Level 4**: All segmented images are removed, and varying levels of US attenuation are introduced to simulate different tissue densities and complexities, mimicking real-world challenges.

Such a progressive training framework can reduce the learning curve associated with ultrasound-guided PCNL.

To implement the proposed framework we follow the overall workflow shown in Fig. 1. Firstly, preoperative CT images of a subject with kidney stones are acquired. A region of interest (ROI) is chosen in each CT image to limit the boundaries to what a conventional ultrasound transducer can reach. The selected ROI is resampled using bilinear interpolation to augment the number of pixels in each image and make denser ultrasound scattering later on. Given the limited spatial resolution of CT slices, a second interpolation algorithm is employed to generate additional slices between the original CT scans. This is elaborated in detail in Section II. From each slice in the augmented CT dataset, a scattering map with normally distributed signal strength is created. Based on the principles of linear acoustics and computation of the spatial impulse response, speckles are simulated by randomly placed scatterers with strength randomly chosen from a normal distribution. The Tupholme-Stepanishen method is then used for calculating pulsed ultrasound fields from which a corresponding ultrasound image is created. The process of converting CT scans into 2D ultrasound images is explained in Section III. In Section IV, these simulated 2D ultrasound images are arranged in a 3D voxel, thereby extending data formation to slices taking along directions not captured in the original dataset. The final step, presented in Section V, is the thresholding of original CT images to segment the kidney contour and kidney stones and then fuse. The remainder of the paper defines performance metrics to evaluate and compare the generated US images against the real-images obtained with the same phantom, presents the obtained experiential results followed by a discussion and conclusions.

Towards implementing the framework above, this paper offers the following contributions:

1. A physics-based method for generating volumetric ultrasound images from 2D CT scans using interpolation and acoustic modeling with tissue attenuation.
2. A multilevel system that incrementally builds expertise in ultrasound-guided PCNL, transitioning from guided overlays to realistic challenges.
3. Fusion of segmented kidney and stone contours for guided learning and gradual removal of assistance to challenge anatomical recognition and procedural navigation forms the foundation for the PCNL simulators.

Unlike prior work that primarily focused on physics-based ultrasound simulation Selladurai et al. (2024); Satheesh and Thittai (2019), our contribution lies in combining CT-derived volumetric densification with a pedagogically structured, multi-level training framework for usPCNL. The progressive removal of visual cues, guided overlays, and attenuation effects represents a novel integration aimed at reducing learning curves. Furthermore, we use multiple datasets enabling both controlled validation and clinically relevant testing. To rigorously evaluate reconstruction fidelity, we employ a comprehensive set of validation metrics including Peak signal-to-noise ratio (PSNR), Dice Similarity Coefficient (DSC), and Structural Similarity Index Measure(SSIM), offering both intensity and structure-based assessments. In addition, arbitrary slices extracted from the

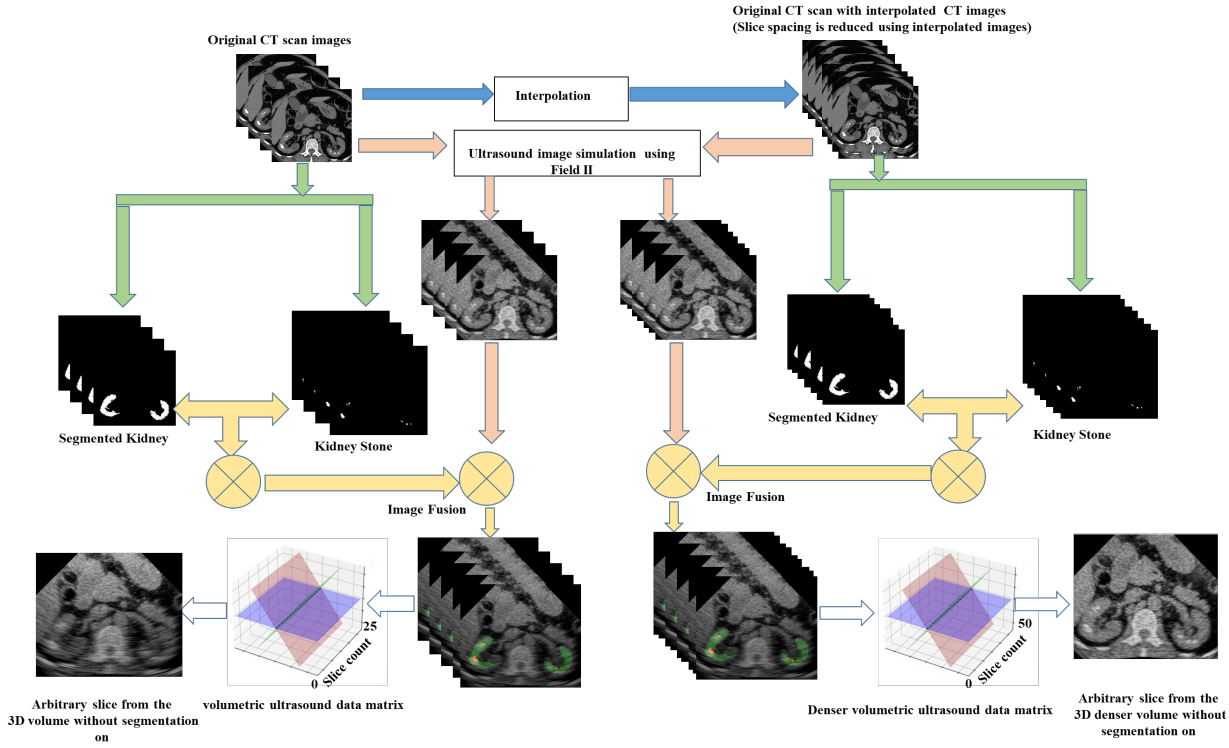

Figure 1: Workflow to generate volumetric ultrasound images from 2D CT scans. CT images are interpolated and converted into ultrasound images to form a 3D volume. New images are then created along imaging planes not captured in the original CT data.

reconstructed 3D volume are analyzed to validate robustness across varying anatomical planes. By incorporating higher-order interpolation, our framework improves spatial continuity and anatomical fidelity. Collectively, these contributions advance beyond existing work by providing a multi-dataset, multi-metric, and pedagogically adaptive simulation platform that is better aligned with both clinical realities and training needs.

2. CT IMAGE PREPROCESSING

The proposed framework begins with CT image preprocessing, we used two different datasets:

1. **Dataset 1**: A set of CT scans and the corresponding ultrasound images are acquired from a CIRS Model 057A Computerized Imaging Reference Systems (2013) triple-modality abdominal phantom, measuring 26 × 12.5 × 19 cm, was used in this study. Designed to mimic the anatomy of an adult abdomen, it is compatible with ultrasound, and MRI imaging. The phantom incorporates internal anatomical features such as the liver, a section of lung, portal vein, segments of both kidneys, abdominal aorta, vena cava, simulated spine, and six ribs. Additionally, it contains eight embedded lesions.
2. **Dataset 2:** The second dataset was collected from the Kaggle database, PACS (Picture Archiving and Communication System) archives across multiple hospitals in Dhaka, Bangladesh. Patients included in this dataset had pre-established diagnoses of kidney-related findings (tumour, cyst, stone, or normal). Both coronal and axial CT slices were selected from whole-abdomen and urogram protocols, including both contrast-enhanced and non-contrast studies, to ensure variability across acquisition conditions. Islam, Hasan, Hossain, Alam, Uddin and Soylu (2022b); Islam, Hasan, Hossain, Alam, Rabiul, Uddin and Soylu (2022a).

Both datasets undergo pre-processing as follows. Since CT slices are rectangular and have a much larger area than an ultrasound image, we first specify a region of interest (ROI) in each CT slice that corresponds to the desired imaging plane of an ultrasound transducer. For usPCNL, we consider a curvilinear ultrasound transducer with a width w and a fixed penetration depth d. ROI is defined as a sector-shaped area centered within the slice, extending to a radius of w. To standardize the ROI for further processing, the sector image—initially in polar coordinates—is transformed into a rectangular format through a polar-to-Cartesian conversion, as illustrated in Fig. 2. This transformation serves two main purposes: (1) it enables the use of a consistent simulation approach for both linear and curvilinear array transducers, and (2) it addresses the disparity between the typical pixel spacing in CT images, which is generally finer than the wavelength of the ultrasound signal.

To transform the trapezoidal ROI image into a rectangular image using polar-to-Cartesian conversion, we need to map each pixel from the polar coordinate system of the

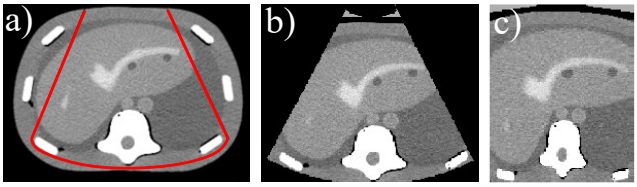

Figure 2: Images showing the input CT slice. The region of interest is marked with a black solid line in (a). Assuming the transducer is centered within the slice, the extracted CT Region of Interest (ROI) is shown in (b), before it is converted into a rectangular image in (c).

trapezoidal ROI to the Cartesian coordinate system of the rectangular image. This transformation involves converting the radial distance r and angular position θ of each pixel to Cartesian coordinates (x, y) as follows:

$$x = r\cos(\theta), \qquad y = r\sin(\theta) \tag{1}$$

where x and y are the horizontal and vertical coordinates of the pixel in the rectangular image, r is the radial distance from the origin to the pixel in the trapezoidal ROI, θ is the angle from a reference axis to the pixel in the trapezoidal ROI. To perform the reverse process, converting rectangular coordinates (x, y) back to polar coordinates (r, θ), we use:

$$r = \sqrt{x^2 + y^2}, \qquad \theta = \arctan\left(\frac{y}{x}\right) \tag{2}$$

where $\sqrt{x^2 + y^2}$ is the Euclidean distance from the origin to the point, giving the radial distance r, and $\text{atan2}\left(\frac{y}{x}\right)$ calculates the angle θ using the *atan2* arctangent function, taking into account the signs of x and y to determine the correct quadrant.

Following this step, to increase spatial resolution the ROI of two consecutive parallel images is interpolated to create new images between the two samples. One common method, bilinear interpolation Yun, Lee, Hong and Shim (2023), calculates pixel values for the new slices based on those of the neighbouring pixels in the original slices. Given a point (x, y) in the new slice, the pixel value $f(x, y)$ can be interpolated from the surrounding two pixels (x_1, y_1), (x_1, y_2), (x_2, y_1), and (x_2, y_2) in the original slices:

$$\begin{aligned} f(x, y) =& (1-\alpha)(1-\beta)f(x_1, y_1) + \alpha(1-\beta)f(x_2, y_1) \\ &+ (1-\alpha)\beta f(x_1, y_2) + \alpha\beta f(x_2, y_2) \end{aligned} \tag{3}$$

where

$$\alpha = \frac{x - x_1}{x_2 - x_1}, \quad \text{and} \quad \beta = \frac{y - y_1}{y_2 - y_1} \tag{4}$$

Bilinear interpolation was chosen for its computational efficiency and balance between smoothness and edge preservation.Preliminary comparisons with nearest-neighbour and cubic spline methods showed that bilinear interpolation minimized discontinuities between slices without introducing over-smoothing or artificial artifacts in the reconstructed 3D volume.

This interpolation doubles the number of CT slices in the dataset and leads to a denser volumetric image downstream, which, as will be shown, improves image rendering. All CT volumes had a native in-plane resolution of 0.74 × 0.74 mm (Dataset 1) and 0.70 × 0.70 mm (Dataset 2), with an inter-slice spacing of 1.0 mm and slice thickness of 1.25 mm. After ROI extraction and polar-to-Cartesian conversion, each slice was resampled to 512 × 512 pixels. Volumetric densification was performed by inserting one interpolated slice between every two adjacent CT slices using bilinear interpolation, reducing the effective voxel spacing in the z-direction to 0.5 mm. Once these CT images are generated, the next step is to convert them into ultrasound images based on linear acoustic principles.

3. 2D ULTRASOUND IMAGE SIMULATION FROM CT SCANS

Simulation of ultrasound images from CT data is carried out across all interpolated and original axial CT slices, with the transducer assumed to be located at the center of each slice. Various techniques have been proposed for generating ultrasound images from CT datasets. For instance, a low-cost ultrasound simulator has been developed for radiology training, where simulated images are produced by combining an ultrasound echo map, an absorption image, and a texture map derived from CT data; however, speckle information is not incorporated in this approach. Burger et al. Bürger, Abkai and Hesser (2008) proposed a method for simulating CT-based ultrasound images by reconstructing a three-dimensional model from CT data. They then generate ultrasound images by simulating ray propagation, beamforming, and backscattering. Reichl et al. Reichl, Passenger, Acosta and Salvado (2009) proposed the use of a computationally efficient parallel programming platform, such as CUDA, for simulating ultrasound from CT images. In another approach, Karamalis et al. Karamalis, Wein and Navab (2010) employed the Westervelt equation to model ultrasound propagation through tissues and solved it using a finite difference scheme implemented on a graphics processing unit (GPU), enabling real-time ultrasound simulation. Two-phase method, where in the first phase, they obtain a map of ultrasound echo reflections for the selected region of interest (ROI), and in the second phase, they generate a scattering image of the ROI using Field II Jensen and Nikolov (2000); Shams, Hartley and Navab (2008), based on linear acoustic principles. The final B-mode image is constructed by merging the reflection map from phase one with the scatter image derived in phase two. Kutter et al. Kutter, Shams, Wein and Navab (2009) extended this work by generating ultrasound transmission texture during phase one, which is then combined with the reflection and scatter images. Dillenseger et al. Dillenseger, Laguitton and Delabrousse (2009) introduced another approach based on the convolution model suggested by Bamber and Dickinson Bamber and Dickinson (1980). A recent technique by Szostek and

Piorkowski Szostek and Piórkowski (2016) uses ray tracing-based wavefront reconstruction. Salehi et al. Salehi, Ahmadi, Prevost, Navab and Wein (2015) proposed a method for simulating ultrasound images, utilizing a convolutional ray tracing technique and a MRI scan segmentation portion as input. Finally, Gao et al. Gao, Karimaghaloo, Sengupta and Osman (2012) compared various methods for simulating ultrasound from other imaging modalities.

Using CT images directly for scatter image generation is suboptimal, as soft tissue delineation in CT scans is limited. To improve contrast, a contrast stretching technique is applied, modifying the dynamic range of the image values to the 5th and 95th percentiles, thereby minimizing the impact of outliers. Selladurai et al. (2024); Satheesh and Thittai (2019)

In this paper, US images are simulated based on the physical principles of ultrasound imaging, including the interaction of sound waves with different tissue densities and the resulting acoustic impedance variations. The simulator generates ultrasound images by calculating the reflection and transmission of sound waves at tissue interfaces, producing realistic ultrasound images Selladurai et al. (2024); Satheesh and Thittai (2019). This is described in the next subsection.

3.1. Ultrasound Reflection and Transmission Image Generation

To generate ultrasound reflection images of the ROI, edge information is essential. In this study, Canny edge detection is applied to extract the necessary edge information and generate both the ultrasound reflection and transmission images for a curvilinear array transducer. Specular reflection of ultrasound signals occurs at the interface between two media when the dimensions of the interface exceed the wavelength of the incoming ultrasound signal. The position of these interfaces is determined from the edge map. The strength of the reflected signal is influenced by the acoustic impedance mismatch between the media and the angle of insonification. The intensity reflection coefficient (α_R) quantifies the relative amount of reflected signal Satheesh and Thittai (2019), as expressed by:

$$\alpha_R = \left(\frac{R_2 \cos\theta_i - R_1 \cos\theta_t}{R_2 \cos\theta_i + R_1 \cos\theta_t} \right)^2 \tag{5}$$

where θ_i is the incidence angle, θ_t is the transmitted angle, and R_1 and R_2 are the acoustic impedance of medium 1 and medium 2, respectively. The incidence angle is calculated as the angle between the insonification beam (emanating radially from the curvilinear array) and the local surface normal of the tissue boundary, estimated from the edge map gradient. The transmitted angle was then obtained using Snell's law. The Hounsfield units (HU) of tissues in CT images are reported to be approximately proportional to their acoustic impedance and can therefore be utilized in the calculation of the reflection coefficient, can be seen in equation 6.

Table 1
Acoustic parameters used in ultrasound simulation.

Tissue Type	Acoustic Impedance (MRayl)	Attenuation Coefficient (dB/cm·MHz)
Fat	1.38	0.6
Muscle	1.70	1.1
Kidneys	1.63	0.9
Stone	7.8	12.0
Water	1.48	0.002
Bone	7.8	20.0

However, air and bone do not follow this proportionality; therefore, the intensity reflection coefficient α_R is fixed at 0.99 for air–tissue interfaces and 0.4 for bone–tissue interfaces. The transmission coefficient is related to the reflection coefficient by $\alpha_T = 1 - \alpha_R$. Consequently, the ultrasound reflected intensity at any point (x, y) can be expressed as:

$$I_R(x, y) = I_a(x, y) \cdot \alpha_R(x, y) \cdot \cos\theta \tag{6}$$

where $I_a(x, y)$ represents the incident intensity at point (x, y), and θ is the angle between the surface normal and the direction of ultrasound wave propagation. The transmitted intensity at that point is calculated as the difference between the incident and reflected intensities. The incident intensity I_a at (x, y) can be expressed as:

$$I_a(x, y) = I_0 10^{-\alpha d f/20} \tag{7}$$

where I_0 denotes the initial incident intensity, α is the attenuation coefficient $dB/\text{cm}/\text{MHz}$, d is the depth in centimeters, and f is the frequency in megahertz.

Acoustic Parameter Definition and Mapping from Hounsfield Units

Acoustic impedance (R) values were derived from CT Hounsfield Units (HU) using a linear mapping function within the soft-tissue range (−100 to +300 HU):

$$R = R_{\text{water}} + k \times (HU - HU_{\text{water}}) \tag{8}$$

where $R_{\text{water}} = 1.48$ MRayl, $HU_{\text{water}} = 0$, and $k = 0.0006$ MRayl/HU. This relationship holds for soft tissues; for air and bone, fixed impedance values were applied since their behavior deviates from linear proportionality. Table 1 summarizes the acoustic impedance and attenuation coefficients used in Equations (5) and (7), based on literature values Bamber and Dickinson (1980); Satheesh and Thittai (2019).

3.2. Ultrasound Image Formation

The process of generating an ultrasound image involves creating a scattering strength map of the ROI and calculating the accumulated acoustic pressure field from a collection of scatterers, from which a single radio frequency (RF) line in

an image can be calculated. The final image combines 128 RF lines interpolated to form the B-mode ultrasound image using principles of linear acoustics as follows:

Scattering Map Generation with Boundaries: First, an image representing the scattering strength of the ROI is created. Simulated boundaries are introduced by defining lines in the scatterer map along which strong scatterers are placed. This step enhances the realism of the simulated ultrasound images.

Random Scatterer Distribution and Field Simulation: Next, a large number N of scatterers (e.g., 1,000,000) are randomly distributed within the ROI. The scatter amplitude and standard deviation depend on the scattering map, ensuring variability in scatterer properties. Then, we calculate and sum the acoustic pressure field $p(x, y, t)$ resulting from this collection of scatterers at a point (x, y) at time t as

$$p(x, y, t) = \sum_{i=1}^{N} A_i e^{-j2\pi \frac{f_0}{c} d_0} \tag{9}$$

where:

$$d_0 = \sqrt{(x - x_i)^2 + (y - y_i)^2} \tag{10}$$

and A_i is the amplitude of the i^{th} scatterer, f is the centre frequency of the transducer, c is the speed of sound in the medium, and (x_i, y_i) are the coordinates of the i^{th} scatterer. The resulting acoustic pressure field mimics the behavior of ultrasound waves propagating through tissue. The sum represents the contribution of each scatterer to the overall acoustic pressure at a given point.

Calculating Scattering Strength: The scattering strength of each scatterer can be determined based on the tissue properties at its location. If ϕ_i denotes the scattering strength of the i^{th} scatterer given the scattering map $S(x, y)$, then:

$$\phi_i = S(x_i, y_i) \tag{11}$$

where (x_i, y_i) are the coordinates of the i^{th} scatterer.

Calculating Radio Frequency (RF) Lines: To calculate a single RF line in the ultrasound image, the responses from a collection of scatterers are summed. If the RF line signal at depth z as $s(z)$, the sum of the contributions from all scatterers within the imaging plane is:

$$s(z) = \sum_{i=1}^{N} A_i e^{-j2\pi \frac{f_0}{c} d_1} \tag{12}$$

where

$$d_1 = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2 + z^2} \tag{13}$$

and (x_c, y_c) and f_0 are the coordinate of the transducer centre, and the transducer's centre frequency.

Interpolating B-Mode Ultrasound Images: Once the RF lines are calculated, they are interpolated to form the B-mode image. Given the RF lines at discrete depth intervals, new RF lines at intermediate depths are generated by bilinearly interpolating the values between adjacent RF lines.

These equations(5- 11) were implemented using the Field II MATLAB toolbox. The next step is to combine the created 2D ultrasound images to form a volumetric image, from which new images can be created along arbitrary imaging places not originally captured by the CT scans.

4. 3D ULTRASOUND IMAGE FORMATION FROM SIMULATED 2D IMAGES

Once the ultrasound images are generated, they are stacked together in parallel along the z axis to form a 3D voxel. Each voxel element contains the pixel values corresponding to a particular location in the tissue volume. The density of the voxel increases with the number of interpolated slices, resulting in a denser representation of the tissue.

Given a set of M 2D ultrasound images $I_i(x, y)$ where $i = 1, 2, ..., M$, each representing a cross-sectional view, a 3D ultrasound volume $V(x, y, z)$ can be constructed by stacking these images along the z-axis:

$$V(x, y, z) = \sum_{i=1}^{N} I_i(x, y) \cdot \delta(z - z_i) \tag{14}$$

where z_i represents the z-coordinate of the i^{th} image plane, and δ is the Dirac delta function.

To form a 2D slice from the 3D voxel along a different imaging plane, a process called slicing is performed. Slicing involves selecting a plane within the 3D volume and extracting the pixel values along that plane. Let the desired imaging plane be defined by vector $\mathbf{n}$ normal to that plane, whose tail is at (x_n, y_n, z_n). If (x, y, z) is any point on the plane, then the vector $\langle x - x_n, y - y_n, z - z_n \rangle$ lies entirely inside the plane and must be perpendicular $\mathbf{n}$, that is,

$$\mathbf{n} \cdot \langle x - x_n, y - y_n, z - z_n \rangle = 0 \tag{15}$$

If n_x, n_y, n_z are the components of $\mathbf{n}$, then

$$n_x(x - x_n) + n_y(y - y_n) + n_z(z - z_n) = 0 \tag{16}$$

or rather:

$$n_x x + n_y y + n_z z = d \tag{17}$$

where $d = n_x x_n + n_y y_n + n_z z_n$. Once the plane is defined, it can be discretized into pixels that take the value of the closest pixel in the 3D volume. After all pixels on the plane are populated, the resulting 2D image from the 3D voxel can be interpolated using the bi-linear interpolation method.

5. EXPERIMENTAL RESULTS: IMAGE GENERATION

Each simulation level is evaluated through a series of experiments. The fidelity of the simulated images is assessed

by comparing them with real ultrasound images obtained experimentally. Quantitative metrics such as image quality, contrast, and anatomical accuracy are used to evaluate the performance of the algorithm. We used two different datasets:

1. **Dataset 1**: A set of 2D CT scans and corresponding 2D ultrasound images were acquired from a CIRS triple-modality abdominal phantom. Ultrasound imaging of the phantom was performed using a SONIX TOUCH Q+® scanner (Ultrasonix, Analogic Corporation, Peabody, MA, USA). These ultrasound images are used as ground truth for comparison with the simulated ultrasound images.
2. **Dataset 2:** The second form Kaggle datasets Islam et al. (2022b,a) containing CT scans with kidney stones saved in 'png' format.

To evaluate the similarity between the original and simulated images, the performance metrics are used:

1) Normalized Root Mean Square Error (NRMSE): NRMSE is a dimensionless measure that quantifies the difference between an image and a reference image, computed as follows:

$$\text{NRMSE} = \sqrt{\frac{1}{X \cdot Y} \sum_{x=1}^{X} \sum_{y=1}^{Y} \frac{\left[I_{in}(x,y) - I_{ref}(x,y)\right]^2}{I_{max} - I_{min}}} \quad (18)$$

where x and y are the vertical and horizontal coordinates of a given pixel, and I_{in} and I_{ref} represent the input and reference images, respectively, with dimensions $X \cdot Y$, and I_{max} and I_{min} denote the maximum and minimum pixel values in the images, set to 1 and 0, respectively. Here, x and y correspond to the vertical and horizontal pixel coordinates, respectively. I_{in} and I_{ref} denote the input and reference images, each of $X \cdot Y$. The terms I_{max} and I_{min} represent the maximum and minimum pixel intensities, set to 1 and 0, respectively

2) Dice Similarity Coefficient (DSC): DSC measures similarity between two images as the ratio of twice the intersection of two sets to the sum of the cardinalities of the two sets. In the context of image segmentation, it quantifies the agreement between the segmented regions and ground truth as

$$\text{DSC} = \frac{|X| + |Y|}{2|X \cap Y|} \quad (19)$$

where X and Y represent the sets of pixels in the segmented region and ground truth, respectively, and $|X|$ and $|Y|$ denote the cardinalities of sets X and Y, respectively. The DSC is compared between the ground truth ultrasound image obtained experimentally and simulated ultrasound images.

3) Structural Similarity Index Measure (SSIM): SSIM assesses the fidelity of the simulated ultrasound images with respect to the ground-truth images. SSIM is a perception-based index that considers image degradation as perceived changes in structural information, alongside changes in luminance and contrast. Unlike traditional error summation metrics such as mean squared error, SSIM provides a better measure of perceptual similarity by evaluating corresponding local patterns of pixel intensities that have been normalized for luminance and contrast. The SSIM between two equivalent sections of images I_1 and I_2 is defined as:

$$\text{SSIM}(I_1, I_2) = \frac{(2\mu_1\mu_2 + c)(2\sigma_{12} + d)}{(\mu_1^2 + \mu_2^2 + c)(\sigma_1^2 + \sigma_2^2 + d)} \quad (20)$$

where μ_1 and μ_2 are the mean pixel intensity of I_1 and I_2, σ_1^2 and σ_2^2 are their variances, and σ_{12} the covariance. The constants c and d are used to stabilize the division with weak denominator values (typically defined as functions of the image dynamic range). SSIM is computed locally using a sliding window, and the final index is the mean of all local SSIM values, ensuring that localized distortions are accounted for across the image. The values range from -1 for no similarity to 1 for identical images.

4) Peak signal-to-noise ratio (PSNR): The PSNR between two images is used as a quality measurement between the image slice taken from volumetric data created from sparse CT data, and the corresponding image slice taken from the volumetric ultrasound image simulated with augmented, interpolated CT slices along the same random imaging plane. The higher the PSNR, the better the quality of the reconstructed image. This is done on both datasets and calculated as:

$$PSNR = 10 \log_{10} \left(\frac{R^2}{\epsilon} \right) \quad (21)$$

where R is the maximum fluctuation in the input image data, and ϵ is the mean square error:

$$\epsilon = \frac{\sum_{X,Y} [I_1(x,y) - I_2(x,y)]^2}{X \times Y}, \quad (22)$$

in which X and Y are the number of rows and columns in the input image, respectively. While both the mean-square error and PSNR are used to compare image compression quality, the mean-square error represents the cumulative squared error between the compressed and original image.

5) Speckle Quality Evaluation: To assess the visual realism of the speckle pattern, we computed the *Speckle Signal-to-Noise Ratio (S-SNR)* in homogeneous regions, defined as:

$$S\text{-}SNR = \frac{\mu}{\sigma} \quad (23)$$

where μ and σ represent the mean and standard deviation of pixel intensities, respectively. A higher S-SNR indicates smoother and more uniform speckle texture.

All simulations use a curvilinear transducer with specifications outlined in Table 2, positioned at the centre of an axial CT slice. The depth of ultrasound signal penetration in both original database images is standardized at 180 mm.

5.1. 2D image simulation

Fig. 3(a) displays an axial CT slice from Dataset 1, with the region of interest (ROI) outlined by a solid black

Table 2
Ultrasound transducer parameters.

Parameter	Value
Operating frequency	2.5 MHz
Bandwidth	2-5 MHz
Transducer element pitch	0.47 mm
Number of elements	128
Speed of sound	1540 m/s
Sampling frequency	40 MHz

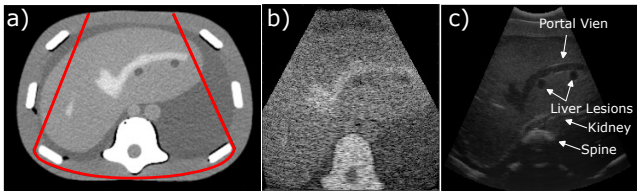

Figure 3: Comparison between (a) ground-truth CT slice with chosen sector ROI marked with a black solid line, (b) simulated ultrasound image, and (c) ground-truth ultrasound image acquired experimentally. This demonstrates the fidelity of our simulation framework relative to actual ultrasound data .

line. Within the ROI, the anatomy includes a superficial fat layer (appearing dark), an underlying muscle layer (slightly brighter than fat), the liver featuring two lesions (visible as dark circular areas) along with the hepatic vein, the vertebra (bright region), the abdominal aorta and vena cava (two circular structures positioned above the vertebra), and adjacent soft tissues (darker areas) containing part of the kidney. The simulated and actual ultrasound images are shown in Fig. 3(b) and 3(c), respectively. All of the images are displayed at the 60 dB dynamic range. All 25 simulated ultrasound images are compared quantitatively against the corresponding real images using RMSE, SSIM and S-SNR as metrics. The average RMSE, DSC, SSIM for all 25 slides are respectively 0.235 ± 0.051, 0.914 ± 0.062, and 0.79 ± 0.05. The average S-SNR for simulated images was 1.82 ± 0.24, closely matching the experimentally acquired ultrasound images (1.95 ± 0.21), confirming realistic speckle characteristics. These results show that the real and simulated ultrasound images are closely. A RMSE closer to zero represents that the generated images and ground truth images are quite similar, SSIM closer to one represents the two images are similar. Selladurai and Thittai (2018). The reported DSC value of 0.0914 ± 0.062 reflects intensity overlap between the simulated and ground-truth grayscale images rather than segmentation accuracy. Since DSC is primarily intended for binary segmentation, its numerical value appears low in this context. Therefore, PSNR and NRMSE provide more meaningful indicators of overall image similarity.

5.2. 3D image simulation

In both datasets, CT images are acquired with a spacing of 1 mm and thickness of 1.25 mm. To realize a denser volumetric ultrasound simulator, the 1 mm spacing between the slices is further reduced to 0.5 mm using the interpolation

Table 3
PSNR for volumetric data from each dataset

Dataset	Original volume	Denser volume
1	18.01 ± 2.19	24.27 ± 2.56
2	16.14 ± 2.70	23.65 ± 2.94

described earlier. To this end, 25 axial CT slices were chosen from the database and pre-processed. Bilinear interpolation was employed between each pair of adjacent slices to generate an additional slice in between, resulting in the generation of another 25 CT slices. These original and interpolated CT data are further used to simulate ultrasound data, which is then used to form a 3D volume.

Two US volumetric images were created for each dataset, one that contained only the original 25 slices and one with the extra 25 interpolated slices, resulting in a total of 50 slices, hereafter referred to as the denser volume.

Each US volumetric image was then sliced along the 10 arbitrary imaging planes, and the corresponding 2D images were generated. Figure 4 illustrates examples from three of these planes. PSNR is computed between corresponding slices from each volume. The results, presented in Table 2,show that the denser volumetric image resulted in a PSNR improvement of 25% and 27% for data sets 1 and 2, respectively, compared to the original volumetric data, suggesting that denser data improves image quality.

6. MULTI-LEVEL TRAINING WITH SEGMENTATION AND FUSION

Each CT scan from Dataset 2 is segmented to identify the kidney and surrounding anatomical structures. The segmentation process is performed using a combination of manual annotation and automated algorithms to ensure accuracy and consistency. For kidney segmentation, a soft-tissue intensity threshold (30–80 HU) was applied to extract renal parenchyma. Morphological closing filled small intraparenchymal voids, while opening removed small background fragments. The resulting binary mask captured the overall kidney contour, providing a structural reference for fusion Müller, Tibyampansha, Mildenberger, Panholzer, Jungmann and Halfmann (2023). For kidney stone segmentation, a higher HU threshold (>200) approach was used to detect the calcified kidney stone, leveraging the stone's higher radiodensity compared to surrounding tissues Akkasaligar, Biradar and Kumbar (2017). The final stone masks were extracted. After segmentation, the kidney and the kidney stone were fused with the ultrasound image. The ultrasound overlay (red for stones, green for kidneys) provides real-time soft tissue visualization, while the segmented CT data enhances anatomical clarity. This multimodal fusion aids in stone localization, procedural planning, and real-time tool guidance, see Fig. 6

The simulator is designed with a multi-level training framework to progressively build the trainee's expertise.

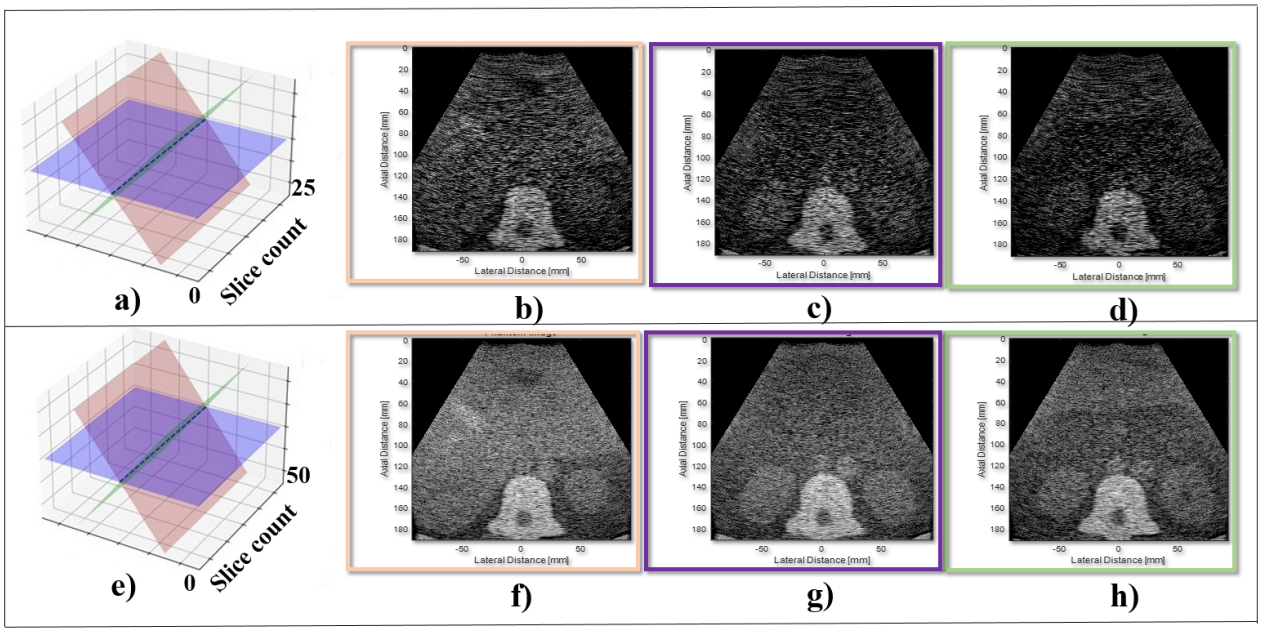

Figure 4: (a) volumetric data formed from the original 25 CT slices, (b) ultrasound image formed from the slice shown in orange in (a), (c) ultrasound image formed from the slice shown in blue in (a),(d) ultrasound image formed from the slice shown in green in (a), (e) denser volumetric data formed from ultrasound images from 50 CT slices, (f) ultrasound image from the slice shown in orange in (e), (g) ultrasound image from the slice shown in blue in (e), and (h) ultrasound image formed from the slice shown in green in (e).

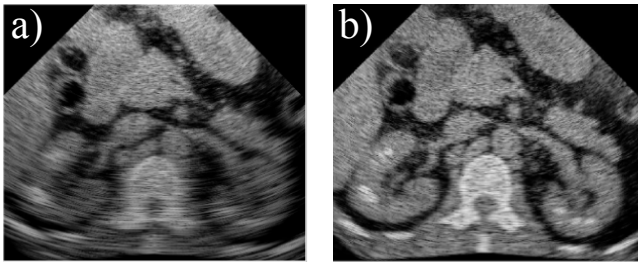

Figure 5: (a) shows arbitrary slices, formed form volumetric data of original 25 CT slices,(b) denser ultrasound image arbitrary slice formed from volumetric data corresponding 00 to 50 CT slices.

Each level introduces increasing complexity and realism, tailored to enhance specific skills required for PCNL:

Level 1 - Comprehensive image fusion: Both kidney stones and kidney contours segmented from CT scans are overlaid on ultrasound images. Fig. 6 shows a simulated slice (a), (b) shows the arbitrary slice taken from the volumetric image shown in (c). This level provides a comprehensive visual representation, allowing trainees to focus on tool steering towards the stones.

Level 2 - Anatomical recognition: Kidney stone images are removed, and trainees only see the ultrasound images fused with kidney images. Fig. 6 shows the result segmented from the CT scan, shows the arbitrary slice taken from (c). This level emphasizes anatomical recognition and navigation without the presence of stones.

Level 3 - Procedural Challenge: Only the segmented Kidney and overlaid on real-time ultrasound images and kidney contour images are not provided. This level challenges trainees to identify and navigate towards the stones without the aid of kidney contours.

Level 4 - Realistic Complexity: All images segmented from CT scans are removed, and varying levels of ultrasound attenuation α, ranging from 0 to 1.5 in dB cm^{-1} MHz^{-1}, are introduced to simulate different tissue densities and complexities. This level mimics real-world challenges and trains users to recognize the kidneys and stones in poor visibility conditions. Segmentation overlays can be re-enabled during specific scenarios to support decision-making. For example, obese patients have more acoustic attention, making it challenging to find the kidney and stones.

To support effective skill acquisition, the training framework was designed with progressive levels of difficulty. In the initial stages, kidney and stone contours are explicitly displayed, serving as visual scaffolds that reduce cognitive load and guide novice trainees in orienting themselves within the ultrasound image. These cues help learners establish a foundational understanding of the relevant anatomy and spatial relationships. As the levels advance, these visual aids are gradually removed. This stepwise reduction encourages learners to independently recognize anatomical structures and pathologies, thereby strengthening image interpretation skills without reliance on external overlays. Such a progression is consistent with competency-based training principles, in which learners transition from guided

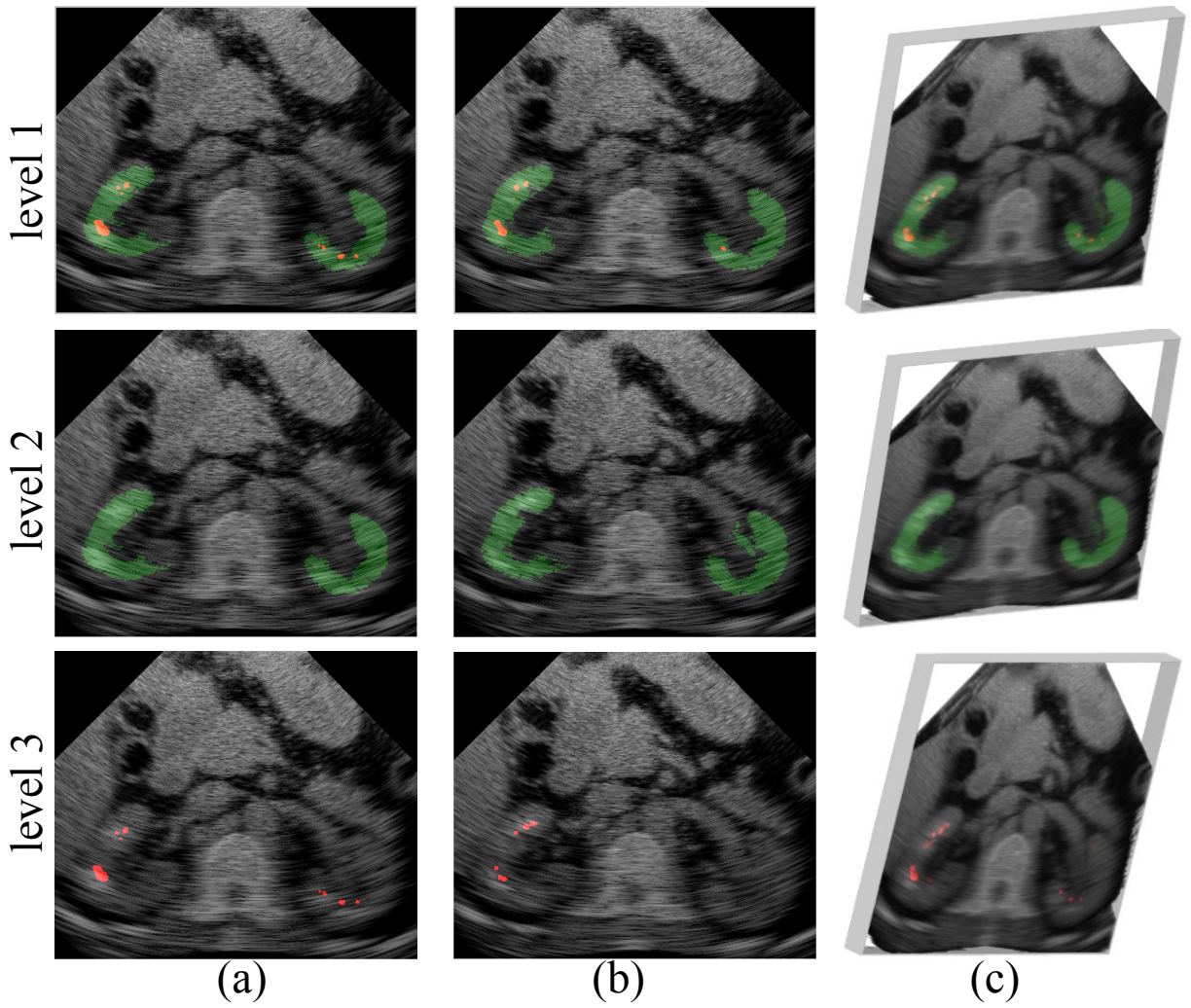

Figure 6: Assistance levels. In Level 1 both kidney stones and the kidney contours are overlaid on top of real-time ultrasound images. In Level 2, segmented images of the kidney stones are removed. In Level 3 only kidney stones from CT slices are overlaid on top of real-time ultrasound and the kidney contours are not provided. (a) is taken directly from a 2D ultrasound image corresponding to a 2D CT scan. (b) is taken from an arbitrary imaging plane from the volume shown in (c).

practice toward autonomous performance. The design ensures that skills developed in the simulated environment are transferable to real clinical scenarios, where no such visual cues are available.

To quantitatively compare the fused and unfused ultrasound images quality in Levels 1 to 3, we use contrast as a performance metric. Contrast is computed as the standard deviation (C) of pixel intensities, which represents the dispersion of intensity values around the mean. A higher standard deviation indicates better contrast, enhancing the visibility of image features. The contrast is given by:

$$C = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{24}$$

where x_i represents individual pixel intensity values, μ is the mean intensity, and N is the total number of pixels. As expected, the results summarized in Fig. 8 confirm that the addition of the segmented images improves the identification of critical anatomical structures.

Finally, DSC and SSIM were computed for the images in Level 4 by comparing simulated images with varying attenuation levels against the reference image generated at the ideal (zero) attenuation level. For each attenuation setting, 10 images were analyzed. At an attenuation level of $0.5\,\mathrm{dB}\cdot\mathrm{cm}^{-1}\cdot\mathrm{MHz}^{-1}$, the DSC and SSIM values were 0.903 ± 0.162 and 0.86 ± 0.04, respectively. These values decreased to 0.65 ± 0.92 (DSC) and 0.69 ± 0.06 (SSIM) at an attenuation level of $1.5\,\mathrm{dB}\cdot\mathrm{cm}^{-1}\cdot\mathrm{MHz}^{-1}$. This trend indicates that as attenuation increases, structural information in the images progressively diminishes. As summarized in Table 3, the denser volumetric image resulted in consistent PSNR improvement across both datasets. The average PSNR increased from 18.01 to 24.27 dB for Dataset 1 (**34.8% improvement**) and from 16.14 to 23.65 dB for Dataset 2 (**46.5% improvement**).These results confirm that denser CT

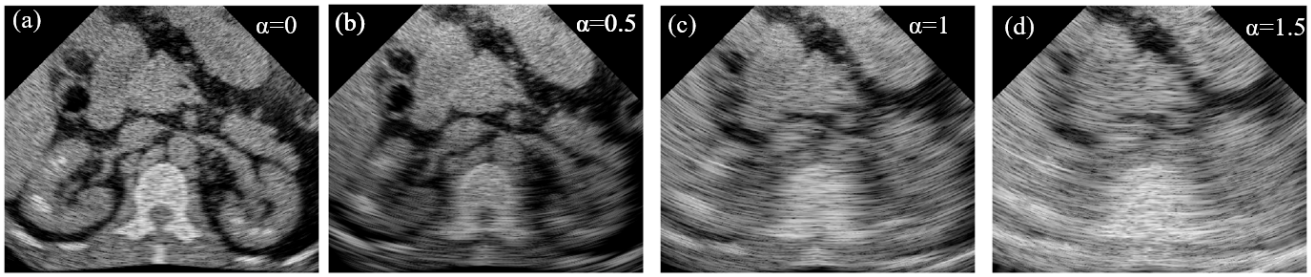

Figure 7: Level 4 removes all images segmented from CT scans and introduces increasing levels of ultrasound attenuation, simulating different tissue densities ((a) to (d) with attenuation of 0 to 1.5 dB cm^{-1} MHz^{-1}) and complexities to mimic real-world challenges. In cases where the kidney guidance is needed, the segmentation can be turned on like in Level 1. This will train the user to recognize the kidney and the stones in poor visibility conditions.

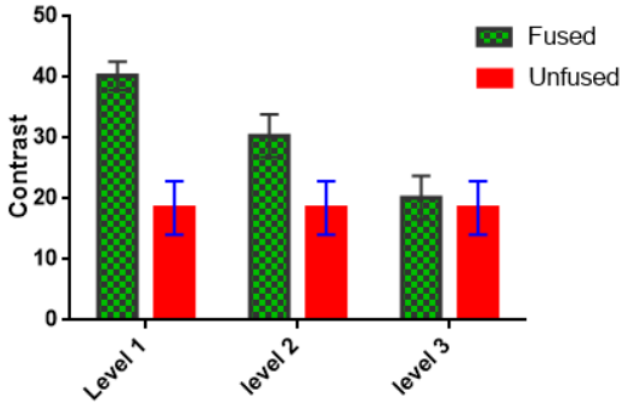

Figure 8: Contrast for fused ultrasound and segmented CT images taken from the volumetric image along an arbitrary imaging plane.

interpolation substantially enhances simulated ultrasound image quality and structural fidelity.

7. CONCLUSION AND FUTURE WORK

In this paper, we proposed a method to generate volumetric ultrasound images from 2D CT scans, laying the groundwork to create a virtual reality simulator tailored specifically for usPCNL. Our approach addresses the challenge of limited spatial resolution in CT scans by employing interpolation techniques to augment the image density, resulting in a higher-resolution CT data. These augmented CT scans were then converted into ultrasound images based on principles of linear acoustics and spatial impulse response. Performance evaluation of the image generation algorithm using normalized RMSE, SSIM, DSC, and PSNR demonstrates its effectiveness in accurately simulating ultrasound images from CT data. The comparison between simulated and ground-truth ultrasound images shows small differences, indicating acceptable simulation fidelity. The results also show that denser CT scan data led to a significant improvement in ultrasound image quality obtained along an imaging plane not captured in the original data. The ability to generate ultrasound images along arbitrary imaging planes not originally captured in the CT scans expands the versatility and effectiveness of the simulator for training scenarios.

Expanding this methodology to incorporate dynamic anatomical changes, such as tissue deformation and organ motion, would further enhance the realism of the simulation. Furthermore, the scalability of our approach should be explored to accommodate variations in patient anatomy and pathology. Customization options tailored to individual patient cases would enable personalized training scenarios, catering to diverse clinical situations encountered in real-world practice. Lastly, collaborative efforts with medical educators and practitioners would be invaluable in refining the image-processing functionality and ensuring its alignment with training objectives and clinical standards. By incorporating feedback from experts in the field, we can iteratively improve the simulation platform to better serve the needs of trainees and ultimately contribute to improved patient outcomes in usPCNL procedures. Moreover, we intended to explore machine learning techniques to personalize training scenarios based on individual trainee performance and learning needs. Ultimately, a high fidelity simulator will enable trainees to practice usPCNL procedures on patient-specific data before entering the operating room, enhancing their skills and improving patient outcomes.

In conclusion, this preliminary work represents an important step towards the development of an ultrasound 3D volumetric framework for usPCNL training, which can ultimately improve patient outcomes and safety in clinical practice. The integration of ultrasonography in PCNL brings forth several advantages, including real-time imaging, reduced radiation exposure, and improved precision in stone targeting. These factors collectively contribute to enhanced procedural outcomes and patient safety. Despite the promising findings, challenges such as operator dependence, and a steeper learning curve must be acknowledged. Further research is needed to address these limitations and refine the technique. This highlights the growing body of evidence supporting the use of usPCNL. The integration of ultrasonography offers potential advantages in stone clearance rates, safety profiles, and procedural outcomes. Future research should focus on standardized protocols, operator

training, and long-term follow-up to establish the role of ultrasonography as a routine imaging modality in PCNL procedures.

The complexity of ultrasound-guided PCNL necessitates the development of advanced simulation tools to enhance clinician training and skill development. This study presents the development of a 3D-dense ultrasound image simulator derived from 2D CT scans. The real time ultrasound image will fuse information from the CT scan, see below.

Segmented kidney images from CT scans are fused on top of the US images to enhance visualization of the kidney under ultrasound. The simulator uses a multi-level approach with an increasing level of complexity and realism, tailored to progressively build the trainee's expertise.

In Level 1 (easiest) we segment both kidney stones and the kidney contours from CT scans, and overlay them on top of real-time ultrasound images. This provides a comprehensive visual representation of the analytical structures, allowing the trainee to focus solely on tool steering towards the stones;

In Level 2 images of the kidney stones are removed. The trainee only sees the real time ultrasound images fused with kidney images segmented from the CT scan, enabling trainees to concentrate on anatomical recognition and navigation without the presence of stones. Level 3 segments kidney stones from CT slices and overlays them on top of real-time ultrasound (kidney contour images form CT scans are not provided)

Level 4 removes all images segmented from CT scans and introduces increasing levels of ultrasound attenuation, simulating different tissue densities and complexities to mimic real-world challenges. In cases where actuation is completing diring the kidney, the segmentation can be turned on like in Level 1. This will train the user to recognize the kidney and the stones in poor visibility conditions.

The simulator is designed to progress from basic to advanced levels of complexity, mirroring the increasing challenges faced in actual PCNL procedures. By incorporating these levels, the simulator aids in anatomical and procedural training while also introducing users to the subtleties of ultrasound imaging, such as varying tissue densities and attenuation effects. This progressive training framework aims to improve procedural accuracy, enhance clinical decision-making, and ultimately reduce the learning curve associated with PCNL. A limitation of the current work is the omission of surgical instrument modelling. In future iterations, we plan to represent instruments as high-scattering objects, distinct from stones, to simulate tool tracking is a crucial feature in PCNL training.

Some limitations in image quality remain. These primarily arise from the restricted spatial resolution of the preoperative CT datasets, the inherent challenges of modelling complex ultrasound physics, and the computational trade-offs necessary for efficient simulation. To objectively assess image fidelity, we have reported NRMSE, DSC, PSNR, and SSIM values, demonstrating quantitatively that the simulated images provide a sufficient basis for anatomical recognition and procedural training. Further improvements in image quality are anticipated as higher resolution datasets and future work will be directed toward further improvement in image realism and computational modelling.

The present simulation framework assumes linear acoustic propagation and a simplified mapping from Hounsfield Units to acoustic impedance, which does not fully capture tissue-dependent heterogeneity or beam-hardening effects. Speckle generation is based on randomly distributed scatterers rather than microstructural tissue properties, which may limit realism in highly heterogeneous regions. Future work will integrate nonlinear propagation models, improved impedance estimation, and biomechanical deformation to enhance realism. Also our framework is fundamentally constrained by the resolution of the CT data used for volumetric reconstruction. Anatomical details smaller than the slice spacing cannot be fully captured, which may reduce fidelity in regions containing subtle structures. To mitigate this limitation, interpolation between adjacent slices was employed. Moreover, intraoperative anatomical changes—such as respiratory motion, patient positioning, and tissue deformation—are not yet incorporated into the simulation. Modelling deformable and dynamic anatomy is essential for realistic training environments. Future efforts will therefore focus on integrating deformable models and dynamic simulations to more accurately replicate intraoperative conditions and enhance clinical applicability.

This study evaluates the technical aspects of each simulation level, the fidelity of the simulated images, and the potential impact on clinical training outcomes. The results highlight the potential of this 3D simulator to significantly enhance clinician training for usPCNL. Future directions include its integration into standard training protocols to improve procedural success rates and patient outcomes. The simulated volumetric ultrasound framework can be integrated with haptic devices to provide force-feedback during needle puncture, further improving procedural realism. Additionally, the system is compatible with augmented reality (AR) overlays, enabling AR-guided renal access planning and enhancing spatial understanding during training. These extensions constitute promising future directions for a fully immersive usPCNL simulation platform.

8. ACKNOWLEDGEMENTS

The authors express their sincere gratitude to Dr. Anila Satheesh B. and Dr. Arun K. Thittai of IIT Madras for their valuable suggestions and for kindly providing the CT data of the CIRS phantom. We also acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC).

9. Declarations

Conflict of interest The authors declare no competing interests.

References

Ajemba, M.N., Ikwe, C., Iroanya, J.C., 2024. Effectiveness of simulation-based training in medical education: assessing the impact of simulation-based training on clinical skills acquisition and retention: a systematic review. World Journal of Advanced Research and Reviews 21, 1833–1843.

Akkasaligar, P.T., Biradar, S., Kumbar, V., 2017. Kidney stone detection in computed tomography images. 2017 International Conference on Smart Technologies for Smart Nation (SmartTechCon) , 353–356doi:10.1109/SmartTechCon.2017.8358395.

Bamber, J.C., Dickinson, R.J., 1980. Ultrasonic b-scanning: a computer simulation. Physics in Medicine and Biology 25, 463–479.

Beiko, D., Razvi, H., Bhojani, N., Bjazevic, J., Bayne, D.B., Tzou, D.T., Stoller, M.L., Chi, T., 2019. Techniques–ultrasound-guided percutaneous nephrolithotomy: How we do it. Canadian Urological Association Journal 14, E104.

Bürger, B., Abkai, C., Hesser, J., 2008. Simulation of dynamic ultrasound based on ct models for medical education. Stud. Health Technol. Inform. 132, 56–61.

Canalichio, K.L., Berrondo, C., Lendvay, T.S., 2020. Simulation training in urology: state of the art and future directions. Advances in Medical Education and Practice , 391–396.

Computerized Imaging Reference Systems, I., 2013. Triple modality 3d abdominal phantom, model 057a. http://www.cirsinc.com/products/all/65/triple-modality-3d-abdominal-phantom.

Dillenseger, J.L., Laguitton, S., Delabrousse, E., 2009. Fast simulation of ultrasound images from a ct volume. Computers in biology and medicine 39, 180–186.

Doizi, S., Koskas, L., 2022. Impact of simulation-based training in endourology: A systematic review of the literature. Progres en Urologie: Journal de L'association Francaise D'urologie et de la Societe Francaise D'urologie 32, 813–829.

El-Shaer, W., Abdel-Lateef, S., Torky, A., Elshaer, A., et al., 2019. Complete ultrasound-guided percutaneous nephrolithotomy in prone and supine positions: a randomized controlled study. Urology 128, 31–37.

Farcas, M., Reynolds, L.F., Lee, J.Y., 2021. Simulation-based percutaneous renal access training: evaluating a novel 3d immersive virtual reality platform. Journal of endourology 35, 695–699.

Ferraguti, F., Farsoni, S., Bonfè, M., 2022. Augmented reality and robotic systems for assistance in percutaneous nephrolithotomy procedures: recent advances and future perspectives. Electronics 11, 2984.

Gao, H., Karimaghaloo, Z., Sengupta, D., Osman, N.F., 2012. A fast convolution-based methodology to simulate 2-d/3-d cardiac ultrasound images. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 59, 515–524.

Gomaa, A.R., Grafton-Clarke, C., Saratzis, A., Davies, R.S., 2023. The role of high-fidelity simulation in the acquisition of endovascular surgical skills: a systematic review. Annals of Vascular Surgery 93, 405–427.

Hamacher, A., Whangbo, T.K., Kim, S.J., Chung, K.J., 2018. Virtual reality and simulation for progressive treatments in urology. International neurourology journal 22, 151.

Islam, M., Hasan, M., Hossain, M., Alam, M., Rabiul, G., Uddin, M., Soylu, A., 2022a. Ct kidney dataset: Normal-cyst-tumor and stone. kaggle.

Islam, M.N., Hasan, M., Hossain, M.K., Alam, M.G.R., Uddin, M.Z., Soylu, A., 2022b. Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from ct-radiography. Scientific Reports 12, 11440.

Jensen, J.A., Nikolov, I., 2000. Fast simulation of ultrasound images, in: 2000 IEEE Ultrasonics Symposium. Proceedings. An International Symposium (Cat. No. 00CH37121), IEEE. pp. 1721–1724.

Karamalis, A., Wein, W., Navab, N., 2010. Fast ultrasound image simulation using the westervelt equation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 243–250.

Kumari, G., Kumar, P., Wadhwa, P., Aron, M., Gupta, N.P., Dogra, P.N., 2006. Radiation exposure to the patient and operating room personnel during percutaneous nephrolithotomy. International urology and nephrology 38, 207–210.

Kutter, O., Shams, R., Wein, W., Navab, N., 2009. Visualization and gpu-accelerated simulation of medical ultrasound from ct images. Computer Methods and Programs in Biomedicine 94, 250–266.

Labate, G., Modi, P., Timoney, A., Cormio, L., Zhang, X., Louie, M., Grabe, M., de la Rosette, on behalf of the CROES PCNL Study Group, J., 2011. The percutaneous nephrolithotomy global study: classification of complications. Journal of endourology 25, 1275–1280.

Metzler, I.S., Holt, S., Harper, J.D., 2021. Surgical trends in nephrolithiasis: increasing de novo renal access by urologists for percutaneous nephrolithotomy. Journal of Endourology 35, 769–774.

Müller, L., Tibyampansha, D., Mildenberger, P., Panholzer, T., Jungmann, F., Halfmann, M.C., 2023. Convolutional neural network-based kidney volume estimation from low-dose unenhanced computed tomography scans. BMC Medical Imaging 23, 187.

Pu, C., Wang, J., Tang, Y., Yuan, H., Li, J., Bai, Y., Wang, X., Wei, Q., Han, P., 2015. The efficacy and safety of percutaneous nephrolithotomy under general versus regional anesthesia: a systematic review and meta-analysis. Urolithiasis 43, 455–466.

Rassweiler, J.J., Teber, D., 2016. Advances in laparoscopic surgery in urology. Nature Reviews Urology 13, 387–399.

Reichl, T., Passenger, J., Acosta, O., Salvado, O., 2009. Ultrasound goes gpu: real-time simulation using cuda. Proc. SPIE Med. Imaging 7261, 726116.

Ritchie, A., Pacilli, M., Nataraja, R.M., 2023. Simulation-based education in urology–an update. Therapeutic Advances in Urology 15, 17562872231189924.

Sahan, M., Sarilar, O., Savun, M., Caglar, U., Erbin, A., Ozgor, F., 2020. Adopting for supine percutaneous nephrolithotomy: analyzing the learning curve of tertiary academic center urology team. Urology 140, 22–26.

Sainsbury, B., Lacki, M., Shahait, M., Goldenberg, M., Baghdadi, A., Cavuoto, L., Ren, J., Green, M., Lee, J., Averch, T.D., et al., 2020. Evaluation of a virtual reality percutaneous nephrolithotomy (pcnl) surgical simulator. Frontiers in Robotics and AI 6, 145. doi:10.3389/frobt.2019.00145.

Salehi, M., Ahmadi, S.A., Prevost, R., Navab, N., Wein, W., 2015. Patient-specific 3d ultrasound simulation based on convolutional ray-tracing and appearance optimization, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 510–518.

Satheesh, A.B., Thittai, A.K., 2019. A fast method for simulating ultrasound image from patient-specific ct data. Biomedical Signal Processing and Control 48, 61–68. doi:10.1016/j.bspc.2018.10.004.

Selladurai, S., Thittai, A.K., 2018. Strategies to obtain subpitch precision in lateral motion estimation in ultrasound elastography. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 65, 448–456. doi:10.1109/TUFFC.2018.2793195.

Selladurai, S., et al., 2024. Towards 3d-denser ultrasound image simulation from 2d ct-scan for ultrasound-guided percutaneous nephrolithotomy training. IEEE International Conference on Systems, Man, and Cybernetics (SMC) , 3708–3714doi:10.1109/SMC54092.2024.10831528.

Shams, R., Hartley, R., Navab, N., 2008. Real-time simulation of medical ultrasound from ct images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 734–741.

Szostek, K., Piórkowski, A., 2016. Real-time simulation of ultrasound refraction phenomena using ray-trace based wavefront construction method. Computer Methods and Programs in Biomedicine 135, 187–197.

Türk, C., Petrik, A., Sarica, K., Seitz, C., Skolarikos, A., Straub, M., Knoll, T., 2016. Eau guidelines on interventional treatment for urolithiasis. European Urology 69, 475–482. doi:10.1016/j.eururo.2015.07.041.

Yang, Y.H., Wen, Y.C., Chen, K.C., Chen, C., 2019. Ultrasound-guided versus fluoroscopy-guided percutaneous nephrolithotomy: a systematic review and meta-analysis. World journal of urology 37, 777–788.

Yun, H.R., Lee, M.J., Hong, H., Shim, K.W., 2023. Inter-slice resolution improvement using convolutional neural network with orbital bone edge-aware in facial ct images. Journal of Digital Imaging 36, 240–249.

Dr. Sathiyamoorthy Selladurai is a postdoctoral fellow at Department of Systems and Computer Engineering at Carleton University. Received his BEng in Electronics and Communication in 2010, and MEng in Control and Instrumentation in 2012, both from Anna University, India. He earned his Ph.D. degree in ultrasound elastography from IIT Madras, India, in 2020. He worked as a postdoctoral fellow (2020-22) at the University of Montréal, Canada. His research interests include biomedical ultrasound, quantitative elastography, calibration of instruments, machine vision.

Dr. James Watterson is an Associate Professor, in the Faculty of Medicine, Division of Urology. Fellowship in Endourology and Shock Wave Lithotripsy, University of Western Ontario, 2002, Urology Residency, University of Ottawa, 2000. He received his MD, Queen's University, and his research interests include minimally invasive urological surgery laparoscopic surgery endourological surgery kidney stone disease.

Dr Rebecca Hibbert, M.D., F.R.C.P.C., is a board-certified radiologist who specializes in abdominal imaging and image-guided procedures. Medical school at McMaster University in Canada. She completed her diagnostic radiology residency training at the University of Ottawa in Canada and fellowship training in cross-sectional imaging and non-vascular interventions at Mayo Clinic in 2012. Dr. Hibbert's clinical practice focuses on Her research interests include interventional oncology, medical education, simulation training.

Dr. Carlos Rossa is an Associate Professor in the Department of Systems and Computer Engineering at Carleton University. He received his BEng and MSc degrees in Mechanical Engineering from the Ecole Nationale d'Ingénieurs de Metz, Metz, France, both in 2010, and earned his PhD degree in Mechatronics and Robotics from the Sorbonne Université, Paris, France, in Dec 2013 under the auspices of the Commissariat à l'Energie Atomique.