



Deep-learning-aided forward optical coherence tomography endoscope for percutaneous nephrostomy guidance

CHEN WANG,^{1,7} PAUL CALLE,^{1,7} NU BAO TRAN TON,¹ ZUYUAN ZHANG,² FENG YAN,¹  ANTHONY M. DONALDSON,¹ NATHAN A. BRADLEY,³ ZHONGXIN YU,⁴ KAR-MING FUNG,^{5,6} CHONGLE PAN,^{2,8} AND QINGGONG TANG^{1,9}

¹Stephenson School of Biomedical Engineering, University of Oklahoma, Norman, OK 73072, USA

²School of Computer Science, University of Oklahoma, Norman, OK 73072, USA

³Department of Urology, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA

⁴Children's Hospital, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA

⁵Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA

⁶Stephenson Cancer Center, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA

⁷These authors contributed equally to this work

⁸cpan@ou.edu

⁹qtang@ou.edu

Abstract: Percutaneous renal access is the critical initial step in many medical settings. In order to obtain the best surgical outcome with minimum patient morbidity, an improved method for access to the renal calyx is needed. In our study, we built a forward-view optical coherence tomography (OCT) endoscopic system for percutaneous nephrostomy (PCN) guidance. Porcine kidneys were imaged in our experiment to demonstrate the feasibility of the imaging system. Three tissue types of porcine kidneys (renal cortex, medulla, and calyx) can be clearly distinguished due to the morphological and tissue differences from the OCT endoscopic images. To further improve the guidance efficacy and reduce the learning burden of the clinical doctors, a deep-learning-based computer aided diagnosis platform was developed to automatically classify the OCT images by the renal tissue types. Convolutional neural networks (CNN) were developed with labeled OCT images based on the ResNet34, MobileNetv2 and ResNet50 architectures. Nested cross-validation and testing was used to benchmark the classification performance with uncertainty quantification over 10 kidneys, which demonstrated robust performance over substantial biological variability among kidneys. ResNet50-based CNN models achieved an average classification accuracy of $82.6\% \pm 3.0\%$. The classification precisions were $79\% \pm 4\%$ for cortex, $85\% \pm 6\%$ for medulla, and $91\% \pm 5\%$ for calyx and the classification recalls were $68\% \pm 11\%$ for cortex, $91\% \pm 4\%$ for medulla, and $89\% \pm 3\%$ for calyx. Interpretation of the CNN predictions showed the discriminative characteristics in the OCT images of the three renal tissue types. The results validated the technical feasibility of using this novel imaging platform to automatically recognize the images of renal tissue structures ahead of the PCN needle in PCN surgery.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Percutaneous nephrostomy (PCN) was first described in 1955 as a minimally invasive, x-ray guided procedure in patients with hydronephrosis [1]. PCN needle placement has since become a valuable medical resource for minimally invasive access to the renal collecting system for drainage, urine diversion, the first step for percutaneous nephrolithotomy (PCNL) surgery and

other therapeutic intervention, especially when the transurethral access of surgical tools into the urological system is difficult or impossible [1–7]. Despite being a common urological procedure, it remains technically challenging to insert the PCN needle correctly in the right place. During PCN, a needle penetrates the cortex and medulla of the kidney to reach the renal pelvis. Conventional imaging modalities have been used in PCN puncture. Ultrasound technique, as a commonly used medical diagnostic imaging method, has been utilized in PCN surgery for decades [8–11]. Additionally, fluoroscopy and computed tomography (CT) are also employed in PCN guidance and sometimes they are used with ultrasonography simultaneously [10–14]. However, due to the limited spatial resolution, these standard imaging modalities have proven to be inadequate for accurately locating the needle tip position [15,16]. The failure rate of PCN needle placement is up to 18%, especially in nondilated systems or for complex stone diseases [17,18]. Failure of inserting the needle into the targeted location in the kidney through a suitable route might result in severe complications [19–21]. Moreover, fluoroscopy has no soft tissue contrast and, therefore, cannot differentiate critical tissues, such as blood vessels, which are important to avoid during the needle insertion. Rupture of renal blood vessels by needle penetrations can cause bleeding. Temporary bleeding after PCN placement occurs in ~95% of cases [2,17]. Retroperitoneal hematomas have been found in 13% [17]. When PCNL is followed, hemorrhage requiring transfusion has increased to 12–14% of the patients [22]. Additionally, needle punctures during PCN can lead to infectious complications such as fever or sepsis, thoracic complications like pneumothorax and hydrothorax, and other complications like urine leak or rupture of pelvicalyceal system [23,24].

Therefore, the selection of position and route of the puncture is important in PCN needle placement. It is recommended to insert the needle into the renal calyx through calyx papilla because fewer blood vessels are distributed on this route, leading to lower possibility of vascular injury [23,24]. Nevertheless, it is always difficult to precisely identify this preferred inserting route in complicated clinical setting even for experienced urologists [25]. If PCN puncture was executed for multiple times, the likelihood of renal injury will increase and the operational time will lengthen, resulting in higher risks of complications [26].

To better guide PCN needle placement, substantive research work has been done to improve the current guidance practice [9,17,27–30]. Ultrasound with technical improvements in many aspects has been utilized. For instance, contrast-enhanced ultrasound has been proved to be a potential modality in the guidance of PCN puncture [31,32]. Tracked ultrasonography snapshot is also a promising method to improve the needle guidance [32]. In order to resolve the bleeding during needle puncture, combined B-mode and color Doppler ultrasonography has been applied in PCN surgeries and it provides promising efficiency in decreasing the major hemorrhage incidence [33]. Moreover, developments in other techniques such as cone-beam CT [34], retrograde ureteroscopy [35] and magnetic field-based navigation device [36] have been utilized to improve the guidance of PCN needle access. On the other hand, endoscope is an effective device that can be assembled within PCN needle [37,38]. It can effectively improve the precision of PCN needle punctures, resulting in lower risks of complications and fewer times of insertions [38]. However, most of the conventional endoscopic techniques involving charge coupled device (CCD) cameras can only provide two-dimensional information and cannot detect subsurface tissue (in other words, visualize the tissue before the needle tip damage it) [39]. Thus, there is a critical need to develop new guiding techniques which have depth-resolved capability for PCN. Optical coherence tomography (OCT) is a well-established non-invasive biomedical imaging modality which can image subsurface tissue with the penetration depth of several millimeters [40–44]. By obtaining and processing the coherent infrared lights backscattered from reference arm and sample arm, OCT can provide two-dimensional (2D) cross-sectional images with high axial resolution (~10 μm), which is 10–100 times higher than conventional medical imaging modalities (*e.g.*, CT and MRI) [45–47]. Owing to the high speed of laser scanning and data processing,

three-dimensional (3D) images of the detected sample formed by numerous cross-sectional images can be obtained in real time [46,48]. Because of the differences in tissue structures among renal cortex, medulla and calyx, OCT has the potential to distinguish different renal tissue types. Due to the penetration limitation in biological tissues (1-2 mm), studies in kidney using OCT have mainly focused on renal cortex [44,45,49–51]. OCT can be integrated with fiber-optic catheters and endoscopes for internal imaging applications [52–54]. For example, endoscopic OCT imaging has been demonstrated in the human GI tract [55–58] to detect Barrett's esophagus (BE) [59,60], dysplasia [61] and colon cancer [62–64]. In the previous study, our lab has developed a portable hand-held forward-imaging endoscopic OCT needle device for real-time epidural anesthesia surgery guidance [65]. This endoscopic OCT setup holds the promise in PCN guidance.

Given enormous accumulation of images and inter- and intra-observer variation from subjective interpretation, computer-aided automatic methods have been utilized to accurately and efficiently classify these data [66,67]. In automated OCT image analysis, convolutional neural networks (CNN) [66,68,69] has been demonstrated to be promising in various applications, such as hemorrhage detection of retina versus cerebrum and tumor tissue segmentation [67,68,70–73].

Herein we demonstrated a forward OCT endoscope system to image the kidney tissues lying ahead of PCN needle during PCN surgery. Images of renal cortex, medulla and calyx were obtained from ten porcine kidneys using our system. A tissue type classifier was developed using the ResNet34, ResNet50, and MobileNetv2 CNN architectures. Nested cross-validation and testing [74–76] was used for model selection and performance benchmarking to account for the large biological variability among kidney through uncertainty quantification. The predictions from the CNN models were interpreted to identify the important regions in the representative OCT images used by CNN for the classification.

2. Materials and methods

2.1. Experimental setup

In our project, the OCT endoscope was built based on the swept-source OCT (SS-OCT) system. Figure 1 shows the schematic of our forward endoscopic OCT system. The light source of the SS-OCT system has the center wavelength of 1300 nm and bandwidth of 100 nm [39]. The wavelength-swept frequency (A-scan) rate is 200 kHz with ~25 mW output power. As demonstrated in Fig. 1, laser provided by the light source was first split by a fiber coupler (FC) into two paths with 3% and 97% of the whole laser power, respectively. The 3% of the laser power was delivered into the Mach-Zehnder interferometer (MZI), which generated a frequency-clock signal for triggering the OCT sampling procedure and provided it to data acquisition (DAQ) board. The remaining 97% laser transmitted through a circulator in which the light runs only in one direction. Therefore, the light entering port 1 only emitted out from port 2, and then it was evenly split to the reference arm and sample arm of a fiber-based Michelson interferometer. Backscattered light from both arms formed interference fringes at the FC and transmitted to the balanced detector (BD). The interference fringes from different depths received by the BD were encoded with different frequencies. The output signal from BD was further transmitted to the DAQ board and computer for processing. Cross-sectional information can be obtained through Fourier transform of interference fringes [65].

A gradient-index (GRIN) rod lens with a diameter of 1.3 mm was stabilized in front of the galvanometer scanning mirror (GSM). The endoscope we used has the diameter of 1.3 mm, the length of 138.0 mm and the view angle of 11.0°. The total outer diameter (O.D.) including the GRIN rod lens and the protective steel tubing is around 1.65 mm. The proximal GRIN lens entrance of the endoscope was placed close to the focal plane of the objective lens. The GRIN lens can preserve the spatial relationship between the entrance and the output (distal end) and further to the sample. Therefore, one or two directional scanning can be readily performed on

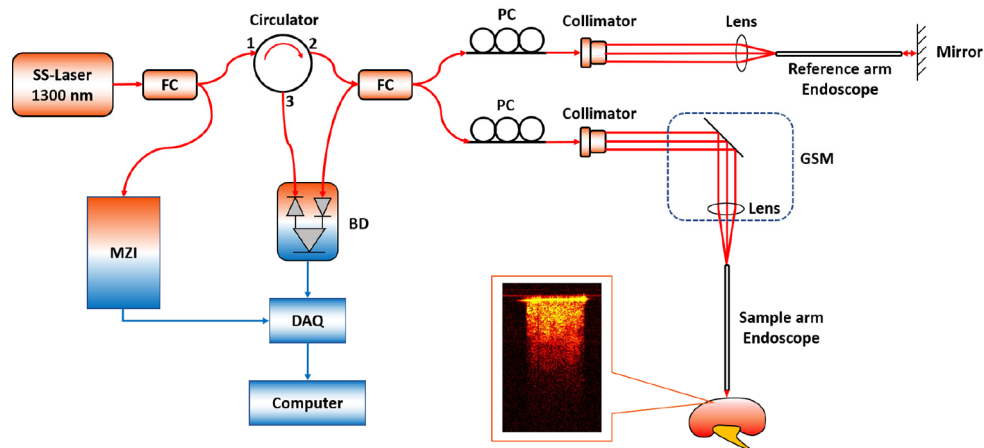


Fig. 1. Schematic of endoscopic OCT system. MZI: Mach-Zehnder interferometer; FC: fiber coupler; PC: polarization controller; GSM: galvanometer scanning mirror; BD: balanced detector; DAQ: data acquisition board.

the proximal GRIN lenses surface to create 2D or 3D images. In addition, the same GRIN rod lens was put in the light path of reference arm for the purpose of compensating light dispersion and expanding the length of reference arm. Polarization controllers (PCs) were put in both arms to decrease the background noise. The system had the axial resolution of $\sim 11 \mu\text{m}$ and lateral resolution of $\sim 20 \mu\text{m}$ in tissue. The lateral imaging field-of-view (FOV) was around 1.25 mm. The sensitivity of system was optimized to 92 dB, calculated using a silver mirror with a calibrated attenuator.

2.2. Data acquisition

Ten fresh porcine kidneys were obtained from a local slaughterhouse. The cortex, medulla and calyx of the porcine kidneys were exposed and imaged in the experiment. Renal tissue types can be identified from the anatomic appearance. The OCT endoscope was placed against different renal tissues for image acquisition. To mimic the clinical situation, we applied some force during imaging the *ex-vivo* kidney tissues to generate tissue compression. The 3D images of $320 \times 320 \times 480$ pixels on X, Y and Z axes (Z presents the depth direction) were obtained with the pixel size of $6.25 \mu\text{m}$ on all three axes. Therefore, the size of the original 3D images is $2.00 \text{ mm} \times 2.00 \text{ mm} \times 3.00 \text{ mm}$. For every kidney sample, we obtained at least 30 original 3D OCT images for each tissue type and each 3D tissue scanning took no more than 2 seconds. Afterwards, the original 3D images were separated to 2D cross-sectional images as shown in Fig. 2. Since the GRIN lens is cylindrical, the 3D OCT images obtained were also in the cylindrical shape. Therefore, not all the 2D cross-sectional images contained the same structural signal of the kidney. Only the 2D images with sufficient tissue structural information (cross-sectional images that close to the center of the 3D cylindrical structures) were subsequently selected and utilized for the image preprocessing. At the end of imaging, tissues of cortex, medulla and calyx of the porcine kidneys were excised and processed for histology to compare with corresponding OCT results. The tissues were fixed with 10% formalin, embedded in paraffin, sectioned ($4 \mu\text{m}$ thick) and stained with hematoxylin and eosin (H & E) for histological analysis. Images were taken by Keyence Microscope BZ-X800. Sectioning and H & E staining was carried out by the Tissue Pathology Shared Resource, Stephenson Cancer Center (SCC), University of Oklahoma Health Sciences Center. The Hematoxylin (cat#3801571) and Eosin (cat# 3801616) were purchased from

Leica biosystems and the staining was performed utilizing Leica ST5020 Automated Multistainer following the HE staining protocol at the SCC Tissue Pathology core.

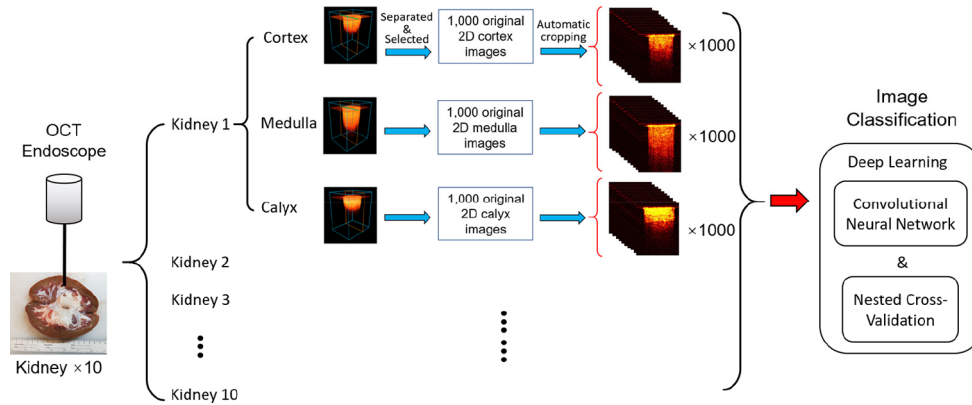


Fig. 2. Illustration of data acquisition and processing steps

Although the three tissue types showed different imaging features for visual recognition, it will take time and expertise for doctors to differentiate them during surgeries. In order to improve the efficiency, we developed deep learning methods for automatic tissue classification based on the imaging data. Figure 2 shows the overall process of the data acquisition and processing. In total, ten porcine kidneys were imaged in this study. For each kidney, 1,000 2D cross-sectional images were obtained for cortex, medulla, and calyx, respectively. For the purpose of convenient analysis and increasing the speed of deep-learning process of the OCT images, a custom MATLAB algorithm was designed to recognize the surface of the kidney tissue on the 2D cross-sectional images. It automatically cropped the images from the size of 320×480 to 235×301 . Therefore, all the 2D cross-sectional images have the same dimensions and cover the same FOV before deep-learning processing.

2.3. Training of the convolutional neural networks

CNN was used to classify the images of renal cortex, medulla, and calyx. Three CNN architectures were tested, including Residual Network 34 (ResNet34) [77], Residual Network 50 (ResNet50) [77] and MobileNet2 [78], using Tensorflow 2.3 [79] in open-ce version 0.1.

Pre-trained ResNet50 and MobileNet2 models on the ImageNet dataset [80] were imported from the Keras library [81]. The output layer of the models was changed to one containing 3 softmax output neurons for cortex, medulla, and calyx. The input images were preprocessed by resizing to the 224×224 resolution, replicating the input channel to 3 channels, and scaling the pixel intensities to $[-1, 1]$. Model fine-tuning was conducted in two stages as described in [82]. First, the output layer was trained with all the other layers frozen. The optimizer, stochastic gradient descent (SGD), was used with a learning rate of 0.2, a momentum of 0.3, and a decay of 0.01. Then, the entire model was unfrozen and trained. The SGD with Nesterov momentum optimizer was used with a learning rate of 0.01, a momentum of 0.9, and a decay of 0.001. Early stopping with a patience of 10 and a maximum number of epochs 50 was used for the Pre-trained ResNet50. Early stopping with a patience of 20 and a maximum number of epochs 100 was used for MobileNet2.

The ResNet34 and ResNet50 architectures were also trained using randomly initialized weights. ResNet34 [77] was obtained from [82]. The mean pixel in the training dataset was used to center the training, validation, and test datasets. The input layer was modified to accept only one input channel in the OCT images and the output layer was changed for the classification of the three

tissue types. For ResNet50, the optimizer SGD with Nesterov momentum with learning rate 0.01, momentum 0.9 and decay 0.01 was used. ResNet50 was trained with a maximum of 50 epochs, early stopping with a patience of 10, and a batch size of 32. For ResNet34, the Adam optimizer was used with learning rate 0.001, beta1 0.9, beta2 0.9999 and epsilon 1E-7. ResNet34 was trained with a maximum of 200 epochs, early stopping with a patience of 10, and a batch size of 512.

2.4. Nested cross-validation and testing

A nested cross-validation and testing procedure [74,76,83] was used to estimate the validation performance and the test performance of the models across the 10 kidneys with uncertainty quantification. The pseudo-code of the nested cross-validation and testing is shown below.

```
# 10-fold cross-testing loop
for kidney i in the 10 kidneys do
    Hold out kidney i in the test set
    # model optimization loop
    for each model configuration do
        # 9-fold cross-validation loop
        for kidney j in the remaining 9 kidneys do
            Use kidney j as the validation set
            Train a model using the remaining 8 kidneys as the training set
            Benchmark the validation performance using kidney j
        end for
        Estimate the mean validation accuracy and its std. error
    end for
    Select the best model configuration based on the validation performance
    Train a model with the selected configuration using the 9 kidneys
    Benchmark the test performance of this model using kidney i
end for
Summarize the test performance of this procedure
```

In the 10-fold cross-testing, one kidney was selected in turn as the test set. In the 9-fold cross-validation, the remaining nine kidneys were partitioned 8:1 between the training set and the validation set. Each kidney contained a total of 3000 images, including 1000 images for each tissue type. The validation performance of a model was tracked based on its classification accuracy on the validation kidney. The classification accuracy is the percentage of correctly labeled images out of all 3000 images of a kidney.

The 9-fold cross-validation loop was used to compare the performance of ResNet34, ResNet50, and MobileNetv2, and optimize the key hyperparameters of these models, such as pre-trained versus randomly initialized weights, learning rates, and number of epochs. The model configuration with the highest average validation accuracy was selected for the cross-testing loop. The cross-testing loop enabled iterative benchmarking of the selected model across all 10 kidneys, giving a better estimation of generalization error with uncertainty quantification.

Gradient-weighted Class Activation Mapping (Grad-CAM) [84] was used to explain the predictions of a selected CNN model by highlighting the important regions in the image for the prediction outcome. The interpretation implementation on ResNet50 was based on [85].

All the model development was performed on the Schooner supercomputer at the University of Oklahoma and the Summit supercomputer at Oak Ridge National Laboratory. The computation on Schooner used five computational nodes, each of which had 40 CPU cores (Intel Xeon Cascade Lake) and 200 GB of RAM. The computation on Summit used up to 15 nodes, each of which had 2 IBM POWER9 processors and 6 NVIDIA Volta Graphic cards. The source code of the model training is available at https://github.com/thePanlab/FOCT_kidney.

3. Results

3.1. Forward OCT imaging of different renal tissues

The imaging setup our OCT endoscope was demonstrated in Fig. 3(A). An adapter was used to stabilize the endoscope in front of the OCT scan lens kit. From the kidney sample shown in

Fig. 3(A), we can visually distinguish different tissue types. Renal cortex was the brown tissue on the edge of the whole kidney; Medulla can be recognized by its red renal pyramid structures distributed on the inner side of the cortex; Calyx was featured by its obvious white structure in the central kidney. Three tissue types were imaged respectively following the procedure described in Section 2.2.

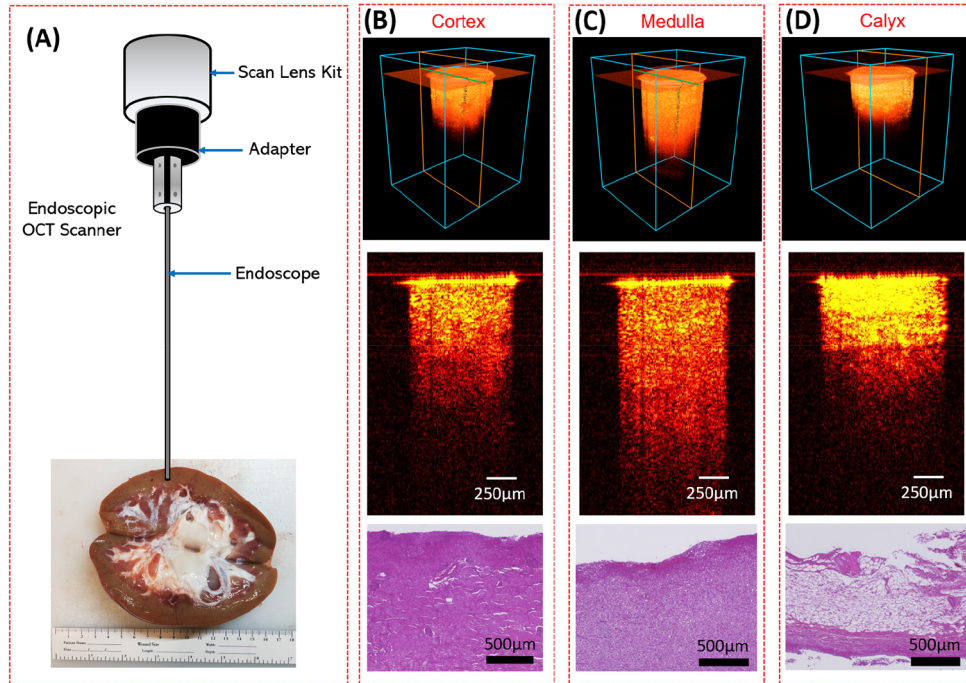


Fig. 3. (A) Sample kidney used for endoscopic OCT imaging; (B-D) Representative 3D OCT images, cross-sectional OCT images and the histology results of renal cortex (B), medulla (C) and calyx (D).

Figures 3(B)–3(D) show representative 3D OCT images, 2D cross-sectional images and the histology results of three renal tissues. They were featured with different imaging depth and brightness. The renal calyx had the shallowest imaging depth, but the tissue close to the surface showed the highest brightness and density. Cortex and medulla both presented relatively homogeneous tissue structures in OCT images, and the imaging depth of medulla was larger than cortex. Furthermore, compared to cortex and medulla, calyx was featured with horizontal stripes and layered structure. The transitional epithelium and fibrous tissue in the calyx may explain the strip-like structures and significantly higher brightness in comparison to the other two renal tissues. This is the significant part for PCN insertion since the goal of PCN is to reach the calyx precisely. These imaging results demonstrated the feasibility of distinguishing renal cortex, medulla, and calyx with the endoscopic OCT system.

3.2. CNN development and benchmarking results

Table 1 shows the average validation accuracies and their standard errors for the pre-trained (PT) or randomly initialized (RI) model architectures after hyperparameter optimization. RI MobileNetv2 frequently failed to learn, so only the PT MobileNetv2 model was used here. The PT ResNet50 models outperformed the RI ResNet50 models in 6 of the 10 testing folds, which indicated only a small boost by the pre-training on ImageNet. For all the 10 testing folds,

the validation accuracies of the ResNet50 models were significantly higher than those of the MobileNetv2 and ResNet34 models. Thus, the characteristic patterns of the three kidney tissues may require a deep CNN architecture to recognize. The detailed results from the 9-fold cross validation of RI ResNet34, PT MobileNetv2, RI ResNet50, and PT ResNet50 can be found in Supplementary Table 1-Table 4, respectively.

Table 1. The averages and standard errors of the validation accuracies from the nested cross-validation and testing procedure. The architecture with the highest validation performance in each testing fold is indicated in bold.

Testing folds	PT MobileNetv2	RI ResNet34	RI ResNet50	PT ResNet50
K1	84.1%±3.5%	83.8%±3.7%	87.1%±2.9%	89.2%±2.4%
K2	82.1%±3.9%	84.6%±2.8%	87.0%±3.0%	89.8%±2.4%
K3	83.9%±4.2%	84.5%±3.5%	88.5%±2.4%	86.6%±2.4%
K4	78.0%±6.3%	83.5%±2.6%	85.0%±2.4%	87.8%±2.3%
K5	83.3%±2.7%	82.8%±4.2%	87.8%±1.9%	87.7%±2.3%
K6	86.2%±2.7%	84.5%±3.6%	88.4%±2.1%	88.1%±2.4%
K7	86.0%±3.8%	82.8%±3.0%	88.6%±1.7%	86.6%±2.4%
K8	76.1%±6.0%	82.1%±3.9%	87.7%±2.6%	88.0%±2.5%
K9	77.7%±4.8%	79.4%±4.3%	82.1%±3.6%	85.1%±2.3%
K10	81.8%±4.6%	86.4%±2.7%	87.5%±3.2%	87.7%±2.1%

Table 2 shows the test accuracy of the best-performing model in each of the 10 testing folds. The output layer of the CNN models estimated three softmax scores that summed up to 1.0 for the three tissue types. When the category with the highest softmax score was selected for an image (*i.e.*, a softmax score threshold of 0.333 to make a prediction), the CNN model made a prediction for every image (100% coverage) at a mean test accuracy of 82.6%. This was substantially lower than the mean validation accuracy of 87.3%, which suggested the overfitting to the validation set by the hyperparameter tuning and early stopping. The classification accuracy can be increased at the expense of lower coverage by increasing softmax score threshold, which allowed the CNN model to make only confident classifications. When the softmax score threshold was raised to 0.5, 89.9% of the images on average were classified to a tissue type and the mean classification accuracy increased to 85.6%±3.0%. For the uncovered images, the doctors can make a prediction with the help of other imaging modality and their clinical experience.

There was substantial variability in the test accuracy among different kidneys. While three kidneys had test accuracies higher than 92% (softmax score threshold of 0.333), the kidney in the sixth fold had the lowest test accuracy of 67.7%. Therefore, the current challenge in the image classification mainly comes from the anatomic differences among the samples. For instance, Figs. 4(A) and 4(B) shows the receiver operating characteristic (ROC) curves of the prediction results from kidney No. 5 and No. 10 (ROC curves of all the 10 kidneys in the 10-fold cross-testing can be found in the Supplementary data). It is clear that the prediction of kidney 5 is much more accurate than that of kidney 10. Our nested cross-validation and testing procedure was designed to simulate the real clinical setting in which the CNN models trained on one set of kidneys need to perform well on a new kidney unseen by the CNN models until the operation. When a CNN model was trained on a subset of images from all kidneys and validated on a separate subset of images from all kidneys in cross-validation as opposed to partitioning by kidneys, it achieved accuracies over 99%. This suggested that the challenge of image classification mainly stemmed from the biological differences between different kidneys. The generalization of the CNN models across kidneys can be improved by expand our dataset with kidneys at different ages or physical conditions to represent different structural and morphological features.

Table 2. Results from the 10-fold cross-testing. Average and standard error of accuracy are shown for two thresholds.

Testing folds	Threshold = 0.333		Threshold = 0.5	
	Accuracy	Coverage	Accuracy	Coverage
K1	75.4%	100%	79.0%	89.3%
K2	87.9%	100%	90.0%	93.3%
K3	71.4%	100%	76.8%	79.3%
K4	92.1%	100%	95.9%	88.4%
K5	93.2%	100%	95.6%	93.7%
K6	67.7%	100%	68.9%	94.2%
K7	83.5%	100%	86.4%	91.8%
K8	88.2%	100%	92.3%	85.0%
K9	92.2%	100%	94.8%	94.2%
K10	74.1%	100%	76.6%	90.2%
Average	82.6%±3.0%	100%	85.6%±3.0%	89.9%±1.5%

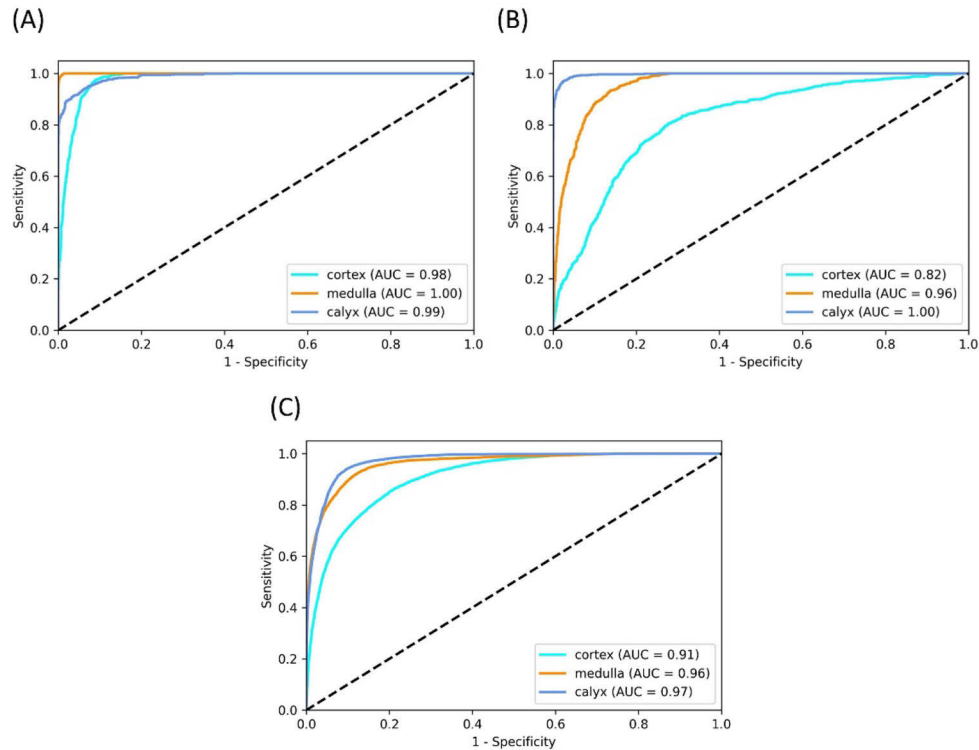
**Fig. 4.** ROC multi-class testing curves of (A) kidney 5, (B) kidney 10 and (C) the average of all kidneys.

Figure 4(C) shows the average ROC curves for the three tissue types. The Area Under the ROC Curve (AUC) were 0.91 for cortex, 0.96 for medulla, and 0.97 for calyx. Table 3 shows the average confusion matrix for the 10 kidneys in the 10-fold cross-testing with a score threshold of 0.333 and the average recall and precision for each type of tissue. Confusion matrix from the 10-fold cross-testing for each of the 10 kidneys is shown in the Supplementary data. Cortex

was most challenging tissue type to be classified correctly and often mixed up with medulla. From the original images we found that the penetration depths of medulla were much larger than cortex in seven of the ten imaged kidneys. While in other three samples, these differences were insignificant. This may explain the challenging classification between cortex and medulla.

Table 3. Confusion matrix for classification of the three tissue types.

		Predicted			Recall
		Cortex	Medulla	Calyx	
True	Cortex	677 ± 109	204 ± 93	118 ± 93	68% ± 11%
	Medulla	75 ± 33	908 ± 41	17 ± 15	91% ± 4%
	Calyx	95 ± 30	14 ± 9	892 ± 32	89% ± 3%
	Precision	79% ± 4%	85% ± 6%	91% ± 5%	

To better understand how a CNN model classified different renal types, the class activation maps were generated to visualize heatmaps of class activation over input images. Figure 5 shows the class activation heatmaps for 3 representative images of each tissue type from the RI ResNet50 model and the PT ResNet50 model. The models and the representative images were selected from the fifth testing fold. The RI ResNet50 model performed the classification by paying more attention to the lower part of the images of cortex, to both the lower part and near the upper part of the medulla images, and to the area of the calyx images with high intensity near the needle tip. The PT ResNet50 model focused on both the upper part and lower part of the cortex images, on the middle part and/or lower part of the medulla images, and on the region close to the needle tip of the calyx images. Compared to the RI ResNet50 model, the PT ResNet50 model shifted its attention closer to the signal regions. The class classification heatmaps provided an intuitive explanation of the classification basis for the two CNN models.

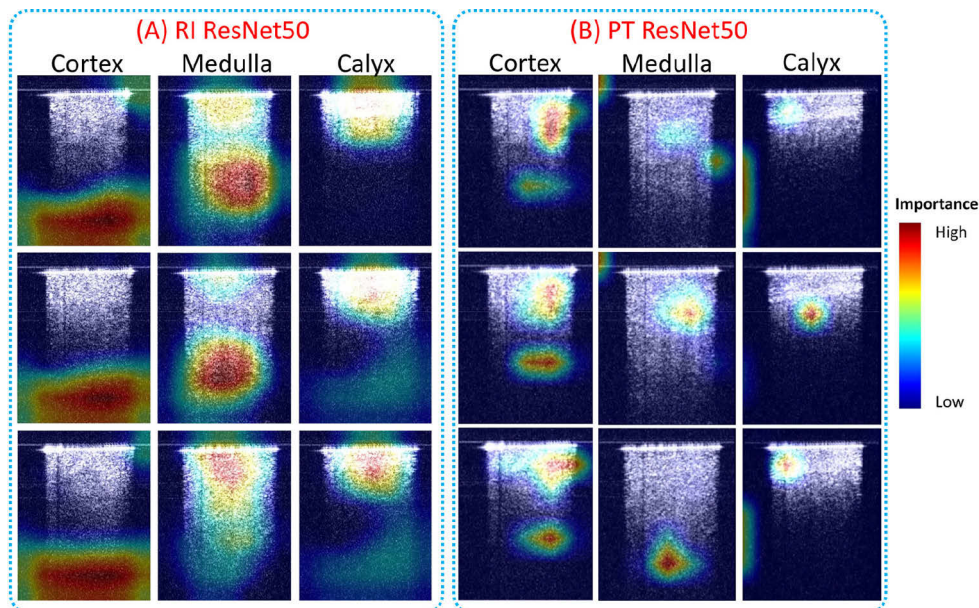


Fig. 5. Class activation heatmaps of (A) RI ResNet50, and (B) PT ResNet50 for three representative images in each tissue type. The class activation heatmaps used jet colormap and were superimposed on the input images.

4. Discussion

We investigated the feasibility of OCT endoscopic system for PCN surgery guidance. Three porcine kidney tissue types: cortex, medulla and calyx were imaged. These three kidney tissues show different structural features which can be further used for tissue type recognition. To increase the image recognition efficiency and reduce the learning burden of the clinical doctors, we developed and evaluated CNN methods for image classification and recognition. ResNet50 had the best performance compared to ResNet34 and PT MobileNetv2 and achieved an average classification accuracy of $82.6\% \pm 3.0\%$.

In the current study, the porcine kidneys samples were obtained from a local slaughterhouse without controlling the sample preservation and time after death. Biological changes may have occurred in the *ex-vivo* kidneys, including collapse of some structures of nephrons such as the renal tubules. This may make the tissue recognition more difficult, especially the classification between cortex and medulla. Characteristic renal structures in the cortex can be clearly imaged by OCT in both well-preserved *ex-vivo* human kidneys and living kidneys as previously reported [44,50,86] and verified in an ongoing study in our lab using well-preserved human kidneys. Additionally, nephron structures distributed in renal cortex and medulla are different [87]. These additional features in renal cortex and medulla will improve the recognition of these two tissue types and increase the classification accuracy of our future CNN models when imaging *in-vivo* samples or well-preserved *ex-vivo* samples. The current study established the feasibility of automatic tissue recognition using CNN and provided information for the model selection and hyper-parameter optimization in future CNN model development using *in-vivo* pig kidneys and well-preserved *ex-vivo* human kidneys.

For translating the proposed OCT probe into clinics, we will assemble the endoscope with appropriate diameter and length into the clinical used PCN needle. In current PCN puncture, a trocar needle is inserted into the kidney. Since the trocar has a hollow structure, we can fix the endoscope within the trocar needle. Then our OCT endoscope can be inserted into the kidney together with the trocar needle. After the trocar needle tip arrives at the destination (such as the kidney pelvis), we will withdraw the OCT endoscope from the trocar needle and other surgical processes can be continued. During the whole puncture, no extra invasiveness will be caused. Since the needle will keep moving during the puncture, there will be a tight contact between the needle tip and the tissue. Therefore, the blood (if any) will not accumulate in front of the needle tip. From our previous experience in the *in-vivo* pig experiment guiding the epidural anesthesia using our OCT endoscope, the presence of blood is not a big issue [65]. The diameter of the GRIN rod lens we used in the study is 1.3 mm. In the future study, we will further improve the current setup with smaller GRIN rod lens that can be fit inside the 18-gauge PCN needle which is clinically used in PCN puncture [88]. Furthermore, we will miniaturize the GSM device based on microelectromechanical system (MEMS) technology, which will enable ease of operation and is important for translating the OCT endoscope to clinical applications. The current employed OCT system has a scanning speed up to 200 kHz, the 2D tissue images in front of the PCN needle can be provided to surgeons in real time. Using ultra-high speed of laser scanning and data processing system, 3D images of the detected sample can be obtained in real time [46,48]. In the next step, we will acquire 3D images that may further improve our classification accuracy, because of the added information content in 3D images. For example, Kruthika *et al.* detected Alzheimer's disease from MRI images and showed improved performance of 3D Capsule Network (CapsNet) and 3D CNN over previous 2D approaches [89].

The CNN model training in this study used significant computational power. Each fold of the cross-validation took ~25 minutes for RI ResNet50 and ~45 minutes for PT ResNet50 using one NVIDIA Volta GPU. The 90 folds of the nested cross-validation for each model configuration were performed in parallel across multiple compute nodes on the Summit supercomputer. The inference was computationally efficient. It took ~50 seconds using one NVIDIA Volta GPU to

perform inference on 1000 images (i.e., ~0.05 seconds per image on average), including model loading, image preprocessing, and the ResNet50 classification. The ResNet50 classification used ~16 seconds for 1000 images or ~0.016 seconds per image. In future, the inference can be further accelerated through algorithm optimization and parallelization to make it more practical for surgical applications.

Funding. National Cancer Institute (P30CA225520); National Institute of General Medical Sciences (P20GM103639); Office of Science (DE-AC05-00OR22725); University of Oklahoma (2020 Junior Faculty Fellowship, 2020 Faculty Investment Program); Oklahoma Center for the Advancement of Science and Technology (HR19-062).

Acknowledgments. Qinggong Tang would like to acknowledge the support from Faculty Investment Program and the Junior Faculty Fellowship of University of Oklahoma; Oklahoma Health Research Program (HR19-062) from Oklahoma Center for the Advancement of Science and Technology. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. Histology service provided by the Tissue Pathology Shared Resource was supported in part by the National Institute of General Medical Sciences Grant P20GM103639 and National Cancer Institute Grant P30CA225520 of the National Institutes of Health.

Disclosures. The authors declare no conflicts of interest related to this article.

Supplemental document. See [Supplement 1](#) for supporting content.

References

1. W. E. Goodwin, W. C. Casey, and W. Woolf, "Percutaneous trocar (needle) nephrostomy in hydronephrosis," *J. Am. Med. Assoc.* **157**(11), 891–894 (1955).
2. M. Young and S. W. Leslie, "Percutaneous nephrostomy," in *StatPearls* (Treasure Island (FL), 2019).
3. L. R. Bigongiari, K. R. Lee, J. Weigel, and J. Foret, "Percutaneous nephrostomy. A non-operative temporary or long-term urinary diversion procedure," *J. Kans. Med. Soc.* **79**, 104–106 (1978).
4. S. Hellsten, J. Hildell, D. Link, and U. Ulmsten, "Percutaneous nephrostomy. Aspects on applications and technique," *Eur. Urol.* **4**(4), 282–287 (1978).
5. E. Perinetti, W. J. Catalona, C. B. Manley, G. Geise, and W. R. Fair, "Percutaneous nephrostomy: indications, complications and clinical usefulness," *J. Urol.* **120**(2), 156–158 (1978).
6. D. P. Stables, N. J. Ginsberg, and M. L. Johnson, "Percutaneous nephrostomy: a series and review of the literature," *Am. J. Roentgenol.* **130**(1), 75–82 (1978).
7. W. J. Lee, U. Patel, S. Patel, and G. P. Pillari, "Emergency percutaneous nephrostomy: results and complications," *J. Vasc. Interv. Radiol.* **5**(1), 135–139 (1994).
8. J. F. Pedersen, "Percutaneous nephrostomy guided by ultrasound," *J. Urol.* **112**(2), 157–159 (1974).
9. O. Efesoy, B. Saylam, M. Bozlu, S. Cayan, and E. Akbay, "The results of ultrasound-guided percutaneous nephrostomy tube placement for obstructive uropathy: A single-centre 10-year experience," *Turk. J. Urol.* **44**, 329–334 (2018).
10. L. LeMaitre, P. Mestdagh, J. Marecaux-Delomez, P. Valtille, F. Dubrulle, and J. Biserte, "Percutaneous nephrostomy: placement under laser guidance and real-time CT fluoroscopy," *Eur. Radiol.* **10**(6), 892–895 (2000).
11. H. G. Zegel, H. M. Pollack, M. C. Banner, B. B. Goldberg, P. H. Arger, C. Mulhern, A. Kurtz, P. Dubbins, B. Coleman, and H. Koolpe, "Percutaneous nephrostomy: comparison of sonographic and fluoroscopic guidance," *Am. J. Roentgenol.* **137**(5), 925–927 (1981).
12. N. L. Miller, B. R. Matlaga, and J. E. Lingeman, "Techniques for fluoroscopic percutaneous renal access," *J. Urol.* **178**(1), 15–23 (2007).
13. Z. L. Barbaric, T. Hall, S. T. Cochran, D. R. Heitz, R. A. Schwartz, R. M. Krasny, and M. W. Deseran, "Percutaneous nephrostomy: placement under CT and fluoroscopy guidance," *Am. J. Roentgenol.* **169**(1), 151–155 (1997).
14. A. Hacker, G. Wendt-Nordahl, P. Honeck, M. S. Michel, P. Alken, and T. Knoll, "A biological model to teach percutaneous nephrolithotomy technique with ultrasound- and fluoroscopy-guided access," *J. Endourol.* **21**(5), 545–550 (2007).
15. N. de Sousa Morais, J. P. Pereira, P. Mota, E. Carvalho-Dias, J. N. Torres, and E. Lima, "Percutaneous nephrostomy vs ureteral stent for hydronephrosis secondary to ureteric calculi: impact on spontaneous stone passage and health-related quality of life—a prospective study," *Urolithiasis* **47**(6), 567–573 (2019).
16. S. Gupta, M. Gulati, and S. Suri, "Ultrasound-guided percutaneous nephrostomy in non-dilated pelvicaliceal system," *J. Clin. Ultrasound* **26**(3), 177–179 (1998).
17. P. Ramchandani, J. F. Cardella, C. J. Grassi, A. C. Roberts, D. Sacks, M. S. Schwartzberg, and C. A. Lewis, and S. o. I. R. S. o. P. Committee, "Quality improvement guidelines for percutaneous nephrostomy," *J. Vasc. Interv. Radiol.* **14**, S277–S281 (2003).
18. P. Montvilas, J. Solvig, and T. E. B. Johansen, "Single-centre review of radiologically guided percutaneous nephrostomy using "mixed" technique: Success and complication rates," *Eur. J. Radiol.* **80**(2), 553–558 (2011).
19. O. Ozbek, H. E. Kaya, A. Nayman, T. B. Saritas, I. Guler, O. Koc, and H. Karakus, "Rapid percutaneous nephrostomy catheter placement in neonates with the trocar technique," *Diagn. Interv. Imaging* **98**(4), 315–319 (2017).

20. C. Vignali, S. Lonzi, I. Bargellini, R. Cioni, P. Petruzzi, D. Caramella, and C. Bartolozzi, "Vascular injuries after percutaneous renal procedures: treatment by transcatheter embolization," *Eur Radiol.* **14**(4), 723–729 (2004).
21. M. Dagli and P. Ramchandani, "Percutaneous nephrostomy: technical aspects and indications," *Semin Intervent Radiol.* **28**(04), 424–437 (2011).
22. W. J. Lee, A. D. Smith, V. Cubelli, G. H. Badlani, B. Lewin, F. Vernace, and E. Cantos, "Complications of percutaneous nephrolithotomy," *Am. J. Roentgenol.* **148**(1), 177–180 (1987).
23. K. A. Hausegger and H. R. Portugaller, "Percutaneous nephrostomy and antegrade ureteral stenting: technique-indications-complications," *Eur Radiol.* **16**(9), 2016–2030 (2006).
24. I. Kyriazis, V. Panagopoulos, P. Kallidonis, M. Ozsoy, M. Vasilas, and E. Liatsikos, "Complications in percutaneous nephrolithotomy," *World J. Urol.* **33**(8), 1069–1077 (2015).
25. E. Radecka and A. Magnusson, "Complications associated with percutaneous nephrostomies. A retrospective study," *Acta Radiol.* **45**(2), 184–188 (2004).
26. H. Egilmez, I. Oztoprak, M. Atalar, A. Cetin, C. Gumus, Y. Gultekin, S. Bulut, M. Arslan, and O. Solak, "The place of computed tomography as a guidance modality in percutaneous nephrostomy: analysis of a 10-year single-center experience," *Acta Radiol.* **48**(7), 806–813 (2007).
27. P. Sonawane, A. Ganpule, B. Sudharsan, A. Singh, R. Sabnis, and M. Desai, "A modified Malecot catheter design to prevent complications during difficult percutaneous nephrostomy," *Arab J. Urol.* **17**(4), 330–334 (2019).
28. P. Alken, G. Hutschenreiter, R. Gunther, and M. Marberger, "Percutaneous stone manipulation," *J. Urol.* **125**(4), 463–466 (1981).
29. Q. Liu, L. Zhou, X. Cai, T. Jin, and K. Wang, "Fluoroscopy versus ultrasound for image guidance during percutaneous nephrolithotomy: a systematic review and meta-analysis," *Urolithiasis* **45**(5), 481–487 (2017).
30. B. X. Liu, G. L. Huang, X. H. Xie, B. W. Zhuang, X. Y. Xie, and M. D. Lu, "Contrast-enhanced US-assisted percutaneous nephrostomy: a technique to increase success rate for Patients with nondilated renal collecting system," *Radiology* **285**(1), 293–301 (2017).
31. X. W. Cui, A. Ignee, T. Maros, B. Straub, J. G. Wen, and C. F. Dietrich, "Feasibility and usefulness of intra-cavitary contrast-enhanced ultrasound in percutaneous nephrostomy," *Ultrasound Med. Biol.* **42**(9), 2180–2188 (2016).
32. T. Ungi, D. Beiko, M. Fuoco, F. King, M. S. Holden, G. Fichtinger, and D. R. Siemens, "Tracked ultrasonography snapshots enhance needle guidance for percutaneous renal access: a pilot study," *J. Endourol.* **28**(9), 1040–1045 (2014).
33. M. H. Lu, X. Y. Pu, X. Gao, X. F. Zhou, J. G. Qiu, and J. Si-Tu, "A comparative study of clinical value of single B-mode ultrasound guidance and B-mode combined with color doppler ultrasound guidance in mini-invasive percutaneous nephrolithotomy to decrease hemorrhagic complications," *Urology* **76**(4), 815–820 (2010).
34. C. M. Hawkins, K. Kukreja, T. Singewald, E. Minevich, N. D. Johnson, P. Reddy, and J. M. Racadio, "Use of cone-beam CT and live 3-D needle guidance to facilitate percutaneous nephrostomy and nephrolithotripsy access in children and adolescents," *Pediatr. Radiol.* **46**(4), 570–574 (2016).
35. C. A. Uribe, H. Osorio, J. Benavides, C. H. Martinez, Z. A. Valley, and K. S. Kaler, "A New Technique for Percutaneous Nephrolithotomy Using Retrograde Ureterscopy and Laser Fiber to Achieve Percutaneous Nephrostomy Access: The Initial Case Report," *J. Endourol. Case Rep.* **5**(3), 131–136 (2019).
36. G. A. Krombach, A. Mahnken, J. Tacke, G. Staatz, S. Haller, C. C. A. Nolte-Ernsting, J. Meyer, P. Haage, and R. W. Gunther, "US-guided nephrostomy with the aid of a magnetic field-based navigation device in the porcine pelvicaliceal system," *J. Vasc. Interventional Radiol.* **12**(5), 623–628 (2001).
37. R. A. Miller and J. E. Wickham, "Percutaneous nephrolithotomy: advances in equipment and endoscopic techniques," *Urology* **23**(5), 2–6 (1984).
38. W. Isac, E. Rizkala, X. Liu, M. Noble, and M. Monga, "Endoscopic-guided versus fluoroscopic-guided renal access for percutaneous nephrolithotomy: a comparative analysis," *Urology* **81**(2), 251–256 (2013).
39. Q. G. Tang, J. T. Wang, A. Frank, J. Lin, Z. F. Li, C. W. Chen, L. Jin, T. T. Wu, B. D. Greenwald, H. Mashimo, and Y. Chen, "Depth-resolved imaging of colon tumor using optical coherence tomography and fluorescence laminar optical tomography," *Biomed. Opt. Express* **7**(12), 5218–5232 (2016).
40. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography," *Science* **254**(5035), 1178–1181 (1991).
41. A. F. Fercher, W. Drexler, C. K. Hitzenberger, and T. Lasser, "Optical coherence tomography - principles and applications," *Rep. Prog. Phys.* **66**(2), 239–303 (2003).
42. P. M. Andrews, Y. Chen, M. L. Onozato, S. W. Huang, D. C. Adler, R. A. Huber, J. Jiang, S. E. Barry, A. E. Cable, and J. G. Fujimoto, "High-resolution optical coherence tomography imaging of the living kidney," *Lab Invest.* **88**(4), 441–449 (2008).
43. J. G. Fujimoto, M. E. Brezinski, G. J. Tearney, S. A. Boppart, B. Bouma, M. R. Hee, J. F. Southern, and E. A. Swanson, "Optical Biopsy and Imaging Using Optical Coherence Tomography," *Nat. Med.* **1**(9), 970–972 (1995).
44. P. M. Andrews, H. W. Wang, J. Wierwille, W. Gong, J. Verbese, M. Cooper, and Y. Chen, "Optical coherence tomography of the living human kidney," *J. Innov. Opt. Health Sci.* **07**(2), 1350064 (2014).
45. Q. Li, M. L. Onozato, P. M. Andrews, C. W. Chen, A. Paek, R. Naphas, S. A. Yuan, J. Jiang, A. Cable, and Y. Chen, "Automated quantification of microstructural dimensions of the human kidney using optical coherence tomography (OCT)," *Opt. Express* **17**(18), 16000–16016 (2009).

46. N. M. Israelsen, C. R. Petersen, A. Barh, D. Jain, M. Jensen, G. Hanneschlager, P. Tidemand-Lichtenberg, C. Pedersen, A. Podoleanu, and O. Bang, "Real-time high-resolution mid-infrared optical coherence tomography," *Light: Sci. Appl.* **8**(1), 11 (2019).
47. J. Fujimoto and E. Swanson, "The Development, Commercialization, and Impact of Optical Coherence Tomography," *Invest. Ophthalmol. Vis. Sci.* **57**(9), OCT1–OCT13 (2016).
48. W. Wieser, W. Draxinger, T. Klein, S. Karpf, T. Pfeiffer, and R. Huber, "High definition live 3D-OCT in vivo: design and evaluation of a 4D OCT engine with 1 GVoxel/s," *Biomed. Opt. Express* **5**(9), 2963–2977 (2014).
49. H. W. Wang and Y. Chen, "Clinical applications of optical coherence tomography in urology," *Intravital* **3**(1), e28770 (2014).
50. B. Konkel, C. Lavin, T. T. Wu, E. Anderson, A. Iwamoto, H. Rashid, B. Gaitian, J. Boone, M. Cooper, P. Abrams, A. Gilbert, Q. Tang, M. Levi, J. G. Fujimoto, P. Andrews, and Y. Chen, "Fully automated analysis of OCT imaging of human kidneys for prediction of post-transplant function," *Biomed. Opt. Express* **10**(4), 1794–1821 (2019).
51. Z. Ding, L. Jin, H.-W. Wang, Q. Tang, H. Guo, and Y. Chen, "Multi-modality Optical Imaging of Rat Kidney Dysfunction: In Vivo Response to Various Ischemia Times," in *Oxygen Transport to Tissue XXXVIII* (Springer, 2016), pp. 345–350.
52. G. J. Tearney, M. E. Brezinski, B. E. Bouma, S. A. Boppart, C. Pitris, J. F. Southern, and J. G. Fujimoto, "In vivo endoscopic optical biopsy with optical coherence tomography," *Science* **276**(5321), 2037–2039 (1997).
53. J. Xi, A. Zhang, Z. Liu, W. Liang, L. Y. Lin, S. Yu, and X. Li, "Diffractive catheter for ultrahigh-resolution spectral-domain volumetric OCT imaging," *Opt. Lett.* **39**(7), 2016–2019 (2014).
54. G. K. Sharma, G. S. Ahuja, M. Wiedmann, K. E. Osann, E. Su, A. E. Heidari, J. C. Jing, Y. Q. Qu, F. Lazarow, A. Wang, L. Chou, C. C. Uy, V. Dhar, J. P. Cleary, N. Pham, K. Huoh, Z. P. Chen, and B. J. F. Wong, "Long-Range Optical Coherence Tomography of the Neonatal Upper Airway for Early Diagnosis of Intubation-related Subglottic Injury," *Am. J. Respir. Crit. Care Med.* **192**(12), 1504–1513 (2015).
55. Y. Chen, A. D. Aguirre, P. L. Hsiung, S. Desai, P. R. Herz, M. Pedrosa, Q. Huang, M. Figueiredo, S. W. Huang, A. Koski, J. M. Schmitt, J. G. Fujimoto, and H. Mashimo, "Ultrahigh resolution optical coherence tomography of Barrett's esophagus: preliminary descriptive clinical study correlating images with histology," *Endoscopy* **39**(07), 599–605 (2007).
56. A. M. Sergeev, V. M. Gelikonov, G. V. Gelikonov, F. I. Feldchtein, R. V. Kuranov, N. D. Gladkova, N. M. Shakhova, L. B. Snopova, A. V. Shakhov, I. A. Kuznetzova, A. N. Denisenko, V. V. Pochinko, Y. P. Chumakov, and O. S. Streltzova, "In vivo endoscopic OCT imaging of precancer and cancer states of human mucosa," *Opt. Express* **1**(13), 432–440 (1997).
57. B. E. Bouma, G. J. Tearney, C. C. Compton, and N. S. Nishioka, "High-resolution imaging of the human esophagus and stomach in vivo using optical coherence tomography," *Gastrointest Endosc* **51**(4), 467–474 (2000).
58. M. J. Gora, J. S. Sauk, R. W. Carruth, K. A. Gallagher, M. J. Suter, N. S. Nishioka, L. E. Kava, M. Rosenberg, B. E. Bouma, and G. J. Tearney, "Tethered capsule endomicroscopy enables less invasive imaging of gastrointestinal tract microstructure," *Nat Med* **19**(2), 238–240 (2013).
59. J. M. Poneros, S. Brand, B. E. Bouma, G. J. Tearney, C. C. Compton, and N. S. Nishioka, "Diagnosis of specialized intestinal metaplasia by optical coherence tomography," *Gastroenterology* **120**(1), 7–12 (2001).
60. K. Liang, G. Traverso, H. C. Lee, O. O. Ahsen, Z. Wang, B. Potsaid, M. Giacomelli, V. Jayaraman, R. Barman, A. Cable, H. Mashimo, R. Langer, and J. G. Fujimoto, "Ultrahigh speed en face OCT capsule for endoscopic imaging," *Biomed. Opt. Express* **6**(4), 1146–1163 (2015).
61. G. Isenberg, M. V. Sivak, A. Chak, R. C. K. Wong, J. E. Willis, B. Wolf, D. Y. Rowland, A. Das, and A. Rollins, "Accuracy of endoscopic optical coherence tomography in the detection of dysplasia in Barrett's esophagus: a prospective, double-blinded study," *Gastrointest. Endosc.* **62**(6), 825–831 (2005).
62. A. M. Winkler, P. F. Rice, R. A. Drezek, and J. K. Barton, "Quantitative tool for rapid disease mapping using optical coherence tomography images of azoxymethane-treated mouse colon," *J. Biomed. Opt.* **15**(4), 041512 (2010).
63. S. Jäckle, N. Gladkova, F. Feldchtein, A. Terentjeva, B. Brand, G. Gelikonov, V. Gelikonov, A. Sergeev, A. Fritscher-Ravens, and J. Freund, "In vivo endoscopic optical coherence tomography of the human gastrointestinal tract-toward optical biopsy," *Endoscopy* **32**(10), 743–749 (2000).
64. P. R. Pfau, M. V. Sivak, A. Chak, M. Kinnard, R. C. K. Wong, G. A. Isenberg, J. A. Izatt, A. Rollins, and V. Westphal, "Criteria for the diagnosis of dysplasia by endoscopic optical coherence tomography," *Gastrointest. Endosc.* **58**(2), 196–202 (2003).
65. Q. G. Tang, C.-P. Liang, K. Wu, A. Sandler, and Y. Chen, "Real-time epidural anesthesia guidance using optical coherence tomography needle probe," *Quant. Imaging Med. Su.* **5**(1), 118–124 (2015).
66. C. Wang, M. Gan, M. Zhang, and D. Y. Li, "Adversarial convolutional network for esophageal tissue segmentation on OCT images," *Biomed. Opt. Express* **11**(6), 3095–3110 (2020).
67. U. Schmidt-Erfurth, A. Sadeghipour, B. S. Gerendas, S. M. Waldstein, and H. Bogunovic, "Artificial intelligence in retina," *Prog. Retin. Eye Res.* **67**, 1–29 (2018).
68. L. Y. Fang, C. Wang, S. T. Li, H. Rabbani, X. D. Chen, and Z. M. Liu, "Attention to lesion: lesion-aware convolutional neural network for retinal optical coherence tomography image classification," *IEEE Trans. Med. Imaging* **38**(8), 1959–1970 (2019).
69. L. Y. Fang, N. J. He, S. T. Li, P. Ghamisi, and J. A. Benediktsson, "Extinction profiles fusion for hyperspectral images classification," *IEEE Trans. Geosci. Remote Sensing* **56**(3), 1803–1815 (2018).

70. R. C. Date, S. J. Jesudasan, and C. Y. Weng, "Applications of deep learning and artificial intelligence in retina," *Int. Ophthalmol. Clin.* **59**(1), 39–57 (2019).
71. Q. Dou, H. Chen, L. Q. Yu, L. Zhao, J. Qin, D. F. Wang, V. C. T. Mok, L. Shi, and P. A. Heng, "Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks," *IEEE Trans. Med. Imaging* **35**(5), 1182–1195 (2016).
72. S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Trans. Med. Imaging* **35**(5), 1240–1251 (2016).
73. M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sanchez, "Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images," *IEEE Trans. Med. Imaging* **35**(5), 1273–1284 (2016).
74. S. Raschka, "Model evaluation, model selection, and algorithm selection in machine learning," arXiv preprint arXiv:1811.12808 (2018).
75. S. Varma and R. Simon, "Bias in error estimation when using cross-validation for model selection," *BMC Bioinformatics* **7**(1), 91 (2006).
76. N. Iizuka, M. Oka, H. Yamada-Okabe, M. Nishida, Y. Maeda, N. Mori, T. Takao, T. Tamesa, A. Tangoku, H. Tabuchi, K. Hamada, H. Nakayama, H. Ishitsuka, T. Miyamoto, A. Hirabayashi, S. Uchimura, and Y. Hamamoto, "Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection," *The Lancet* **361**(9361), 923–929 (2003).
77. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016), 770–778.
78. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2018), 4510–4520.
79. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, and M. Isard, "Tensorflow: A system for large-scale machine learning," in *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, (2016), 265–283.
80. J. Deng, W. Dong, R. Socher, L. Li, L. Kai, and F.-F. Li, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, (2009), 248–255.
81. F. Chollet, Keras, 2015.
82. A. Géron, *Hands-on Machine Learning with Scikit-learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (O'Reilly Media, 2019).
83. D. Krstajic, L. J. Buturovic, D. E. Leahy, and S. Thomas, "Cross-validation pitfalls when selecting and assessing regression and classification models," *J. Cheminform.* **6**(1), 10 (2014).
84. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, (2017), 618–626.
85. F. Chollet, *Deep Learning with Python* (Manning, 2018), Vol. 361.
86. Y. Chen, P. M. Andrews, A. D. Aguirre, J. M. Schmitt, and J. G. Fujimoto, "High-resolution three-dimensional optical coherence tomography imaging of kidney microanatomy ex vivo," *J. Biomed. Opt.* **12**(3), 034008 (2007).
87. C. J. Lote, *Principles of Renal Physiology*, 4th ed. (Kluwer Academic Publishers, 2000), pp. x, 203 p.
88. R. B. Dyer, J. D. Regan, P. V. Kavanagh, E. G. Khatod, M. Y. Chen, and R. J. Zagoria, "Percutaneous nephrostomy with extensions of the technique: Step by step," *Radiographics* **22**(3), 503–525 (2002).
89. K. R. Kruthika, Rajeswari, and H. D. Maheshappa, "CBIR system using Capsule Networks and 3D CNN for Alzheimer's disease diagnosis," *Inf. Med. Unlocked* **14**, 59–68 (2019).