

TEL AVIV UNIVERSITY

The Iby and Aladar Fleischman Faculty of Engineering

The Zandman-Slaner School of Graduate Studies

**Kidney Segmentation and Renal Lesion
Detection in 3D CT**

A thesis submitted toward a degree of
Master of Science in Electrical and Electronic Engineering

by

Neta Blau

October 2017

TEL AVIV UNIVERSITY

The Iby and Aladar Fleischman Faculty of Engineering

The Zandman-Slaner School of Graduate Studies

**Kidney Segmentation and Renal Lesion
Detection in 3D CT**

A thesis submitted toward a degree of
Master of Science in Electrical and Electronic Engineering

by

Neta Blau

This research was carried out in The
School of Electrical and Electronic Engineering

This work was carried out under the supervision of
Prof. Nahum Kiryati

October 2017

Abstract

Renal cysts are common in aging kidneys, and are usually found incidentally in patients undergoing abdominal imaging for other reasons. Although most cysts are benign, they require expert examination as some of them may either indicate the presence of a malignancy or evolve into one. So far, the proposed algorithms for the analysis and detection of renal cysts have been either semi-automatic or evaluated on fairly small data-sets. Here we present a fully automatic method to segment kidneys and to detect simple renal cysts. A fully convolutional neural network (FCN) is employed for segmentation of the kidneys. A combined 3D distance map of the kidneys and surrounding fluids provides initial candidates for cysts. Then, a convolutional neural network (CNN) classifies the candidates as cysts or non-cyst objects. Performance was evaluated on 52 randomly selected volumetric CT scans with more than 70 cysts annotated by an experienced radiologist, with promising results.

Another type of renal lesions are cancerous renal tumors. Though also usually an incidental finding, they are malignant and often fatal. Early detection of such tumors is highly advantageous for recovery. Using the same methods and similar data we attempted to develop a fully automatic system for detection of cancerous renal tumors. We present our experiments and preliminary results and discuss the steps we deem required to undertake this challenge.

Table of Contents

| | |
|--|-----------|
| Acknowledgments | iv |
| List of Figures | 1 |
| List of Tables | 3 |
| 1: Introduction | 4 |
| 1.1 Renal lesions | 4 |
| 1.2 Computed tomography (CT) | 5 |
| 1.3 Kidney segmentation | 6 |
| 1.4 Cyst detection | 7 |
| 1.5 Tumor detection | 8 |
| 1.6 Outline of Thesis | 8 |
| 2: Kidney segmentation | 9 |
| 2.1 Methods | 9 |
| 2.2 Data-set | 9 |
| 2.3 Experiments and results | 10 |
| 2.4 Discussion | 12 |

| | | |
|-----------|---|-----------|
| 3: | Cyst detection | 13 |
| 3.1 | Identification of cyst candidates | 13 |
| 3.2 | Classification of cyst candidates | 14 |
| 3.3 | Data-set | 15 |
| 3.4 | Experiments and results | 15 |
| 3.5 | Discussion | 16 |
| 4: | Tumor detection | 18 |
| 4.1 | Data-set | 18 |
| 4.2 | Experiments and results | 18 |
| 4.3 | Discussion | 20 |
| 5: | Conclusions | 23 |
| | References | 25 |

Acknowledgments

I would like to thank my ...

List of Figures

| | | |
|-----|---|----|
| 2.1 | The kidney segmentation FCN based on the UNet architecture. The figure is adapted from [23]. | 10 |
| 2.2 | Kidney segmentation accuracy measurements of 15 volumetric CT cases | 11 |
| 2.3 | Sample kidney segmentation results using the proposed FCN. The manually annotated ground truth (green) and the segmentation map produced by the FCN (red) are marked on the CT images. Simple renal cysts are evident in all of the examples, demonstrating the network's ability to successfully segment them as part of the kidney. | 12 |
| 3.1 | Main steps of the cyst candidates selection algorithm. | 14 |
| 3.2 | The cyst classification CNN architecture. The network comprises three layers of 3x3 convolution and 2x2 max pooling followed by two fully-connected (FC) layers. | 15 |
| 3.3 | Sample cysts detected using the proposed algorithm. | 16 |
| 3.4 | The trade-off between the TP rate and the average FP rate per patient as a function of (a) the radius threshold and (b) the adjacency threshold. The black line is the maximal TP rate that can be achieved on the given data-set with the proposed algorithm. | 17 |
| 4.1 | Sample tumors from our data-set demonstrating the high variability of cancerous renal tumors. | 19 |

| | | |
|-----|---|----|
| 4.2 | Recall and precision of tumor segmentation on training and validation data-sets using a 2D UNet-based architecture. | 20 |
| 4.3 | Recall and precision of tumor segmentation on training and validation data-sets using a 3D UNet-based architecture. | 21 |

List of Tables

| | | |
|-----|--|----|
| 2.1 | Mean kidney segmentation accuracy evaluated on 15 volumetric CT cases | 11 |
|-----|--|----|

1 Introduction

In a radiologist's line of work time is of the essence. The flux of imaging cases that require human attention is often overwhelming, forcing radiologists to increase their rate of diagnosis. This, in turn, inevitably reduces the quality of their work, causing some pathologies to go undetected or misrepresented. The consequences may vary between suboptimal treatment and considerable damage to patients' health. Automated tools that aid radiologists with their tasks are therefore highly desirable. Automation of medical image analysis have been the focus of much attention in recent years, supported by the rapid development of computer vision technologies.

1.1 Renal lesions

Among the overlooked pathologies are renal lesions. Their asymptomatic nature makes their detection largely incidental, but not less important. Renal masses can be cystic or solid, benign or malignant, and they range from simple renal cysts to cancerous tumors.

Simple renal cysts are benign and asymptomatic fluid collections in the kidney, commonly occurring in aging kidneys. They are usually found incidentally in patients undergoing abdominal imaging for other reasons [27]. In most cases, no treatment is required. However, cysts must be reported by the radiologist and sometimes require follow-up as they may evolve in time.

At the other end of the scale, renal cancer is a lethal disease that manifests as malignant renal lesions. Renal cancer causes approximately 14,000 deaths in the United States every year [26]. The majority of renal cell carcinomas, which ac-

count for 90% of kidney cancers, are asymptomatic [14], and they as well are usually found incidentally on noninvasive imaging performed for unrelated problems [12][19][10]. Early diagnosis of kidney cancer significantly improves survival rates, which fall dramatically once the cancer has spread beyond the kidney [15]. In this work, we address the detection of both types of renal lesions as incidental findings in a real clinical setting.

1.2 Computed tomography (CT)

Among the most frequent imaging procedures carried out in hospitals is the three-dimensional (3D) computerized tomography (CT). CT scans are widely used for preventive medicine and screening for diseases in different parts of the human body. Almost 79 million CT scans were performed in the U.S. during the year of 2015, 245 scans per 1000 population [21].

To produce a CT scan, an X-ray generator rotates around the bed where the patient is lying. The x-ray beams shot by the generator are picked up by X-ray detectors positioned on the opposite side of the circle. When a full rotation is completed, a 2D image is constructed. By moving the bed along the rotation axis, a series of cross-sectional images is obtained, composing the 3D scan. The voxels in the scan are displayed according to the mean attenuation of the tissues that they correspond to on a scale from +3071 (most attenuating) to -1024 (least attenuating). This scale is known as the Hounsfield unit (HU) scale. Water, for example, has an attenuation of 0 HU, while air is approximately -1000 HU. Depending on the diagnostic task, the scan can be viewed as images in each of the anatomical planes - axial, sagittal or coronal. The scans are often anisotropic, and the resolution of each slice (also referred to as the pixel spacing), and the spacing between slices vary from scan to scan.

In order to allow higher visibility, an intravenous (IV) contrast agent is often administered. The contrast material can help enhancing the contrast between a lesion and its surrounding tissue. The time interval between the contrast administration and the image acquisition is determined according to the purpose of the imaging and the investigated organs. To see renal lesions better and detect them more easily, the images are usually acquired at the portal phase of enhancement,

70-80 seconds after the initiation of contrast injection.

1.3 Kidney segmentation

When a radiologist goes over a CT scan and processes it, he first identifies the patient's organs and then looks for any irregularities or abnormalities in them. Adopting the radiologist's strategy, our first step towards detecting renal lesions is to identify and segment the kidneys. Automatic kidney segmentation is a challenging task due to the high inter-patient variability of the kidney shape and location. This task is made harder still by the gray-levels inhomogeneity of the kidney, its similarity to surrounding organs such as the liver, and the possible presence of pathologies.

Numerous methods for kidney segmentation in CT scans have been investigated in the literature. Several atlas-based approaches were proposed in which the segmentation is obtained by registering one or more labeled atlases to the unlabelled target volume [30][5][31]. However, volume registration is very time consuming, and the computation time increases when a multi-atlas approach is used to improve the segmentation accuracy. The increasing availability of medical images enabled the utilization of supervised-learning approaches for organ detection and segmentation. Several proposed methods employ regression forests for organ localization [7][6], classification forests for initial segmentation [6] and joint classification-regression forests for spatially-structured segmentation [9]. These methods are much more efficient than the atlas-based approaches, but their drawback is their dependency on the selected input features.

In recent years, deep convolutional neural networks gained great success in various computer vision tasks, including object detection and segmentation. Given sufficient labeled data, deep networks can learn complex features and representations of the objects to be segmented, thus eliminating the need to manually define and select features. With modern GPUs, the convolution computations have been accelerated significantly, making deep networks a viable option for real-life segmentation tasks. Several studies utilized fully convolutional networks (FCN) for end-to-end segmentation in medical images, and specifically in CT scans. Zhou et al. [32] trained an FCN to segment 19 different organs in 2D slices from all

three anatomical planes of a CT scan. The resulting segmentations were later fused by majority voting into a 3D segmentation. Hu et. al [11] employed a 3D FCN for liver, spleen and kidney segmentation. This step was followed by a time-implicit multi-phase level-set algorithm for segmentation refinement. Cascaded FCNs were also applied to segment the liver and its lesions [4][8]. Thong et. al [28] applied convolutional networks for kidney segmentation in contrast-enhanced CT scans and reported promising results.

1.4 Cyst detection

Provided with a high quality segmentation of the kidneys, we can address the problem of detecting renal lesions, starting with simple renal cysts. Due to their random origins, lesions vary significantly in size, location and quantity. Owing to the homogenous fluid consistency of simple cysts, one may think that simple thresholding of the kidney area can suffice for cyst detection. However, being part of the urinary system, other fluid collections are normally present in the kidneys, making it difficult to distinguish renal cysts from healthy functioning parts of the kidneys such as the renal pelvis and renal calyces.

Only few studies were published on automatic tools for detection or segmentation of renal cysts. Battiato et al. [3] proposed a method to refine a cyst segmentation given an initial boundary marked by the radiologist. The refinement is done in the annotated slice using a bank of filters and is then propagated to the other slices. However, in this method, the cysts must be detected by the users, and their intervention is required for the segmentation. Piao et al. [22] proposed to segment cysts using fuzzy C-mean clustering, but did not report any quantitative results. An algorithm proposed by Badura et. al [2] employs thresholding and morphological operations for the detection of cyst candidates. The candidates are then classified into cysts and non-cysts based on 3D shape-related features. This approach produced promising results, but the algorithm was evaluated on a fairly small data set of 16 patients with 25 cysts.

1.5 Tumor detection

Unlike renal cysts, cancerous tumors are inhomogeneous in their consistency and are more similar to the kidneys in their HU values, thus making it harder to detect them. Summers et al. [25] employed thresholding and morphological operations to segment renal cysts and solid tumors and estimate their volume. However, the work focused on the quantitation of the lesions and detection rates were not reported. Kim and Park [13] detected renal lesions using texture analysis and homogeneous region growing, but their method was evaluated on 12 tumors only. Linguraru et al. [16] proposed a semi-automated algorithm for segmentation and quantification of renal tumors, based on fast marching and geodesic level sets. This algorithm was later extended to also classify the lesion type using support vector machine [17]. This method generated accurate lesion segmentations but user intervention was still required for their detection. Another solution, proposed by Liu et al. [18], relies on geometric changes on the kidney surface for cyst candidate detection. Shape features are then utilized for the classification of the candidates. This algorithm was validated on a large data-set but is limited to detecting only exophytic lesions, and the reported false-positive rate (15 per patient) was high.

1.6 Outline of Thesis

So far, no fully automatic algorithms have been shown to detect simple renal cysts reliably on a large data set. Here, for the first time, we report an algorithm that achieves this goal and demonstrate its performance. The algorithm is constructed from two main modules, a kidney segmentation module and a cyst detection module. In chapter 2 the kidney segmentation module is described, followed by a quantitative evaluation of its performance. Using its output as a starting point, the cyst detection module is then detailed in chapter 3, followed by the overall performance of the algorithm. Additionally, in chapter 4 we relate our efforts to automatically detect cancerous renal tumors. We describe several experiments we conducted and the obtained results, followed by a discussion and suggestions for next steps. Finally, in section 5 we conclude our work.

2 Kidney segmentation

2.1 Methods

Kidney segmentation is undertaken using an FCN. First, the HU values of the CT scan are clipped to the range $[-100, 400]$ to exclude irrelevant organs, followed by normalization to a range of $[-1, 1]$. Then, the 2D axial slices of the scan are fed into the network which classifies each voxel as *kidney* or *non-kidney*. Finally, the 2D segmentation results are assembled back into a 3D segmentation map and the two largest connected components (if exist) are identified as the kidneys.

The FCN that we implemented is based on the UNet architecture [23] which consists of a contracting path to extract features and capture context, and a symmetric expanding path that enables precise localization. The original architecture as was proposed by Ronneberger et al. [23] is comprised of layers of 3×3 convolution interleaved with maximum-pooling layers (in the contracting path) and deconvolution layers (in the expanding path). The 3×3 convolutions were done without zero-padding, leading to segmentation maps smaller than the input data. To get full-size segmentation maps, we perform a 1-pixel width zero padding. Batch normalization (BN) layers were also added before each activation layer to improve the convergence speed of the network. The UNet architecture with the incorporated changes is shown in Fig. 2.1.

2.2 Data-set

The data-set for training and evaluating the kidney segmentation algorithm included 46 abdominal CT scans with contrast agent, taken at the portal phase of

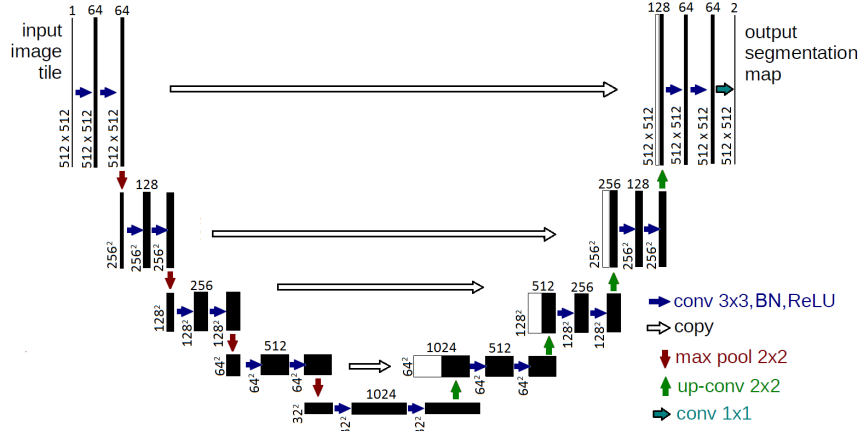


Figure 2.1: The kidney segmentation FCN based on the UNet architecture. The figure is adapted from [23].

enhancement. The CT cases were randomly selected within a one-month period and the kidneys were manually annotated by a senior radiologist. Due to the random selection of the cases, some of them included sick kidneys presenting various pathologies. With the goal of detecting renal lesions in mind, to improve the segmentation robustness to abnormal kidneys, any renal cyst or other pathology observed in the training set was annotated as *kidney* as well.

2.3 Experiments and results

The network was trained on 31 scans for 20 epochs with a mini-batch size of 20 images, learning rate of 0.01 for the first 10 epochs and 0.001 for the next 10 epochs, momentum of 0.9 and weight decay of 0.0005. Segmentation accuracy was then evaluated on the remaining 15 CT cases. Training was performed on an NVIDIA GTX Titan GPU, using the deep learning framework MatConvNet [29].

The accuracy of the 3D segmentation was measured separately for the right and left kidneys in each CT case. Table 2.1 and Fig. 2.2 quantitatively compares between the segmentation results and the ground truth. The network obtained mean dice coefficient of 0.90 for both kidneys. Qualitative examples of the segmentation output and its comparison to the ground truth labelling are shown in Fig. 2.3. The average running time of kidney segmentation was 51 seconds.

Table 2.1: Mean kidney segmentation accuracy evaluated on 15 volumetric CT cases

| Organ | Mean recall | Mean precision | Mean IoU | Mean Dice coefficient |
|--------------|------------------|------------------|------------------|-----------------------|
| Right Kidney | 0.91 ± 0.064 | 0.90 ± 0.049 | 0.82 ± 0.063 | 0.90 ± 0.041 |
| Left Kidney | 0.90 ± 0.084 | 0.92 ± 0.026 | 0.83 ± 0.067 | 0.90 ± 0.044 |

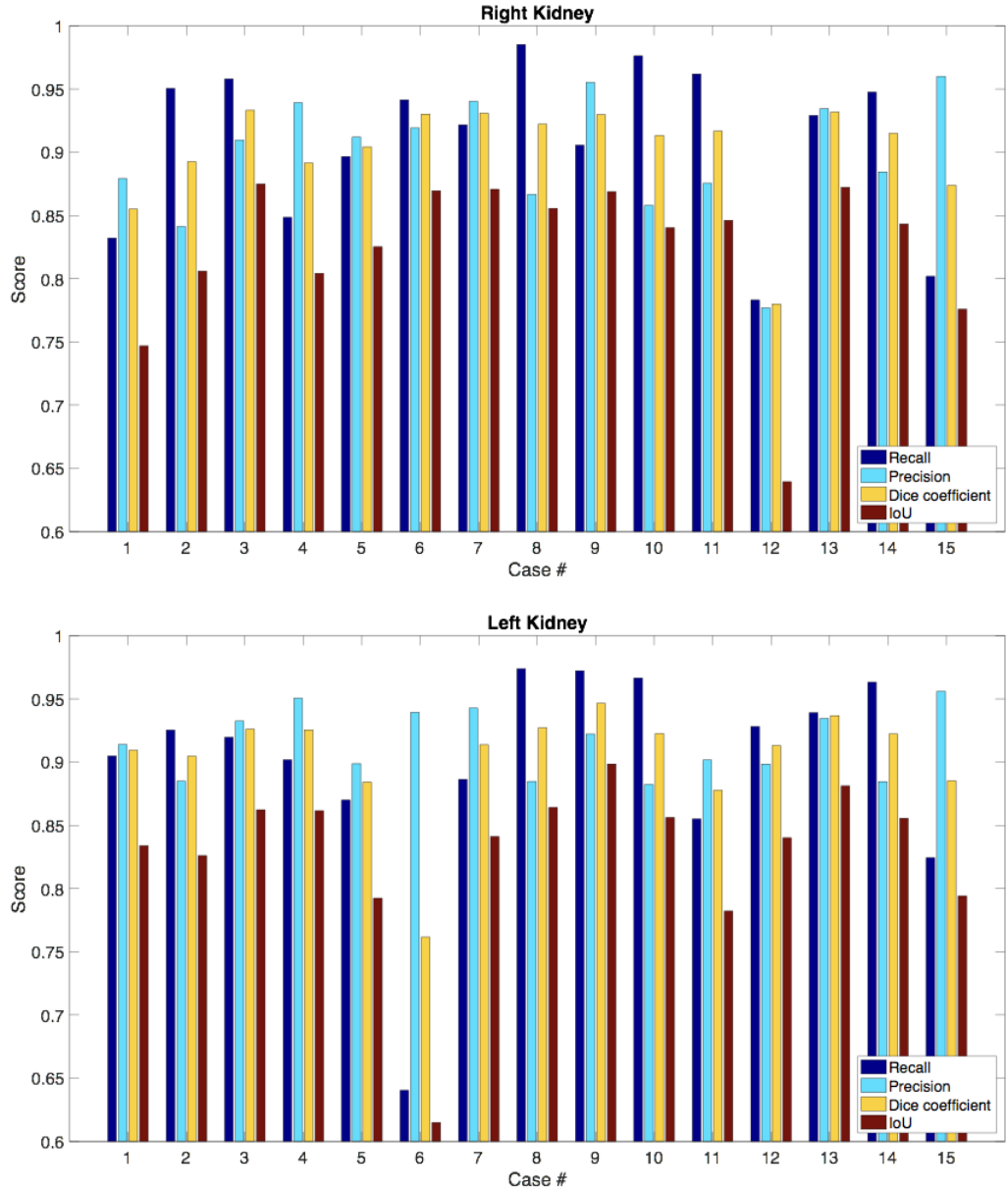


Figure 2.2: Kidney segmentation accuracy measurements of 15 volumetric CT cases

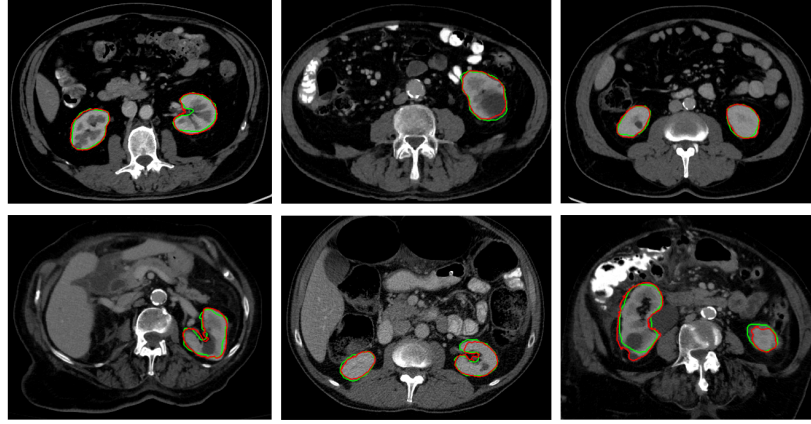


Figure 2.3: Sample kidney segmentation results using the proposed FCN. The manually annotated ground truth (green) and the segmentation map produced by the FCN (red) are marked on the CT images. Simple renal cysts are evident in all of the examples, demonstrating the network’s ability to successfully segment them as part of the kidney.

2.4 Discussion

The kidney segmentation is performed end-to-end by an FCN. Training the network on abdomen CT scans with kidneys presenting multiple cysts and other pathologies achieved promising results, competitive to state-of-the-art methods. When analyzing the segmentation results we noticed that a low recall is often caused by slices in which a renal cyst appears but the kidney is missing. In such cases the cyst was annotated as *kidney* in the ground truth labels but was missed by the FCN. Thus, I believe that the segmentation can be further improved by exploiting 2D slices from all orientations pooled by a majority vote as done in [32], or alternatively, applying a 3D FCN. It is my opinion that significantly enhanced results can be obtained with the same algorithm, by gathering more data, and by improving the quality of the annotated ground truth.

3 Cyst detection

3.1 Identification of cyst candidates

A cyst candidate is the center of a fluid collection, physically attached to one of the kidneys. To locate cyst candidates we generate two 3D euclidian distance maps by applying a 3D distance transform (DT) [20]. An anisotropic DT is necessary to account for the various scan pixel spacing and spacing between slices.

In the first map, the value of each voxel is its distance from the closest *non-fluid* voxel, where a fluid voxel is one with HU in the range $[-50, 80]$. As demonstrated in Fig. 3.1, *non-fluid* voxels obtain zero value, whereas voxels set in the center of a fluid collection (also referred to as seeds) become regional maxima. For simple renal cysts, which are approximately spherical, the seed's distance to *non-fluid* would be the sphere's radius. We denote the seed's distance to *non-fluid* as D_{nf} , and refer to it as the estimated radius of the seed. To reduce the number of potential seeds and considering the scan's resolution, seeds with estimated radius smaller than the spacing between slices are eliminated. In the second distance map, the value of each voxel is its distance from the closest kidney voxel, denoted D_k . We define the adjacency of a seed to the kidney as $D_{nf} - D_k$. For spherical objects, a non-negative adjacency indicates the object and the kidney share a common wall. A negative value indicates they are separate. Provided with a perfect segmentation of the kidneys, cyst candidates would be seeds with a non-negative adjacency value. However, to account for cases of under-segmentation of the kidneys, we decided to lighten the adjacency requirement and set the adjacency threshold to 2.5 mm (the spacing between slices).

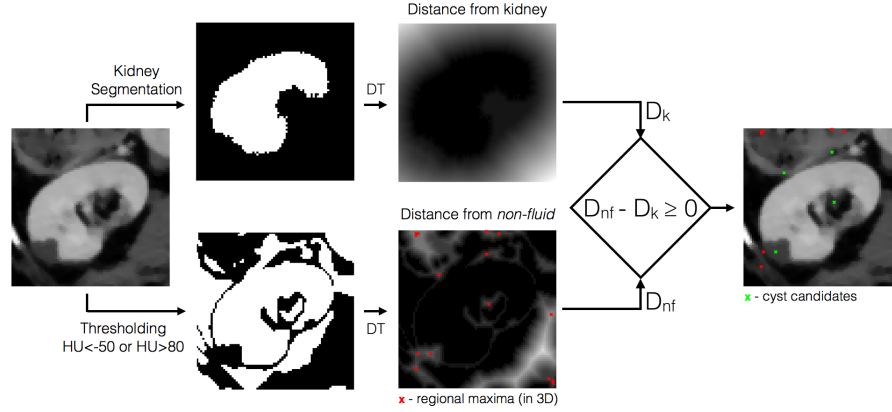


Figure 3.1: Main steps of the cyst candidates selection algorithm.

3.2 Classification of cyst candidates

Following the detection of cyst candidates, we wish to determine which of them are indeed cysts. We note that the cysts are radially symmetrical and hence their 2D projection has similar properties in any orientation. Capitalizing on this notion, patches of the 2D projections of each candidate on three orthogonal planes (axial, sagittal and coronal) are used to construct a 3-channel image. This representation of a volume-of-interest is known as a 2.5D image [24]. The patch size is 32x32 pixels and it is sampled from the original CT scan according to the seed's estimated radius so that the suspected object will occupy the central 10x10 pixels of the patch. The sampling is done using bilinear interpolation, after which the patches are smoothed by a gaussian filter ($\sigma = 0.5$, kernel size = 3x3) and clipped to HU range [-50,80]. The 2.5D images of the candidates are fed into a small convolutional neural network, composed of three layers of 3x3 convolution and two fully-connected layers, which outputs the probability of each seed being a cyst (see Fig. 3.2). Seeds with probability greater than 0.5 are classified as cysts. As a final post-processing step, cysts that are contained in another cyst, i.e. their seed is within the estimated radius of another seed, are considered redundant and eliminated.

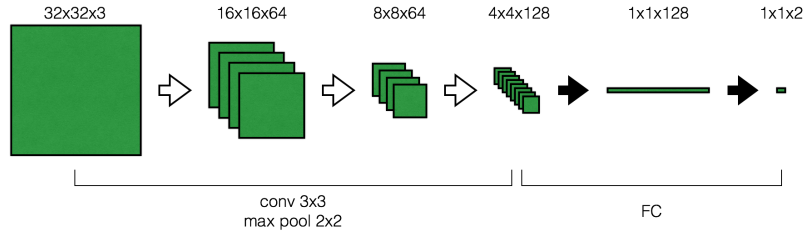


Figure 3.2: The cyst classification CNN architecture. The network comprises three layers of 3x3 convolution and 2x2 max pooling followed by two fully-connected (FC) layers.

3.3 Data-set

Our goal was to detect cysts with a diameter of 10 mm or more. A different data-set was used for this task, with 247 such cysts observed in 173 abdominal CT cases. The cases were randomly selected within a one-month period. All scans were taken with contrast agent injection at the portal phase of enhancement, all have a 2.5 mm spacing between slices and pixel spacing of 0.64-0.98 mm. A line was marked along the diameter of each cyst, allowing the extraction of its center and radius.

3.4 Experiments and results

The classification network was trained on a set of 12,000 seeds extracted from 121 scans for 20 epochs with a mini-batch size of 32 images, learning rate of 0.001 for the first 10 epochs and 0.0001 for the next 10 epochs, momentum of 0.9 and weight decay of 0.001. To increase the number of cysts in the training-set, we augmented the data by rotating and flipping the positive examples. The classification performance was then evaluated on the remaining 52 scans.

The algorithm obtained a true-positive (TP) rate of 83.1% - 59 out of 71 cysts were successfully detected. Qualitative examples of detected cysts are shown in Fig. 3.3. Two cysts were missed due to under-segmentation of the kidney, the rest were incorrectly classified by the network. The algorithm also produced 83 false detections, leading to false-positive rate of 1.6 per patient. However, 5 of these

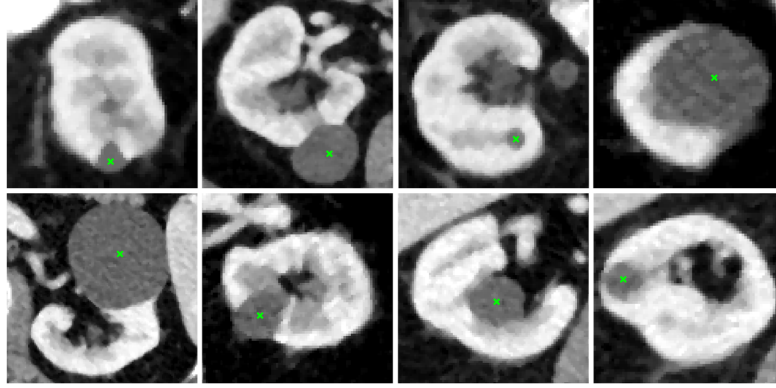


Figure 3.3: Sample cysts detected using the proposed algorithm.

false detections were indications of severe hydronephrosis, still having clinical relevance and requiring radiologist attention. In addition, the algorithm detected 9 more cysts smaller than 10 mm. The average running time of the cyst candidates selection and classification was 22 seconds.

In subsection 3.1 we mentioned that seeds with estimated radius less than 2.5 mm are eliminated. The trade-off between the TP and FP rates can be tuned by changing the radius threshold. As shown in Fig. 3.4 (a), setting the threshold to 3 mm significantly reduces the FP rate from average of 1.6 per patient to 0.88, but at the cost of missing seeds of 4 cysts sized 10-12 mm (5.6% of the cysts). Reducing the adjacency threshold from 2.5 mm to 0.5 mm leads to a similar result (see Fig. 3.4 (b)).

3.5 Discussion

A combined 3D distance map of the kidneys and surrounding fluids provides us with a large number of cyst candidates (almost 20,000 candidates for 52 scans) but the classification network handles them well, outputting clinically relevant results with high sensitivity and specificity rates. In the context of automatic incidental findings detection, 1.6 FP/case means, in average, less than two unnecessary key image per case. This is an acceptable overhead given the detection benefits of the proposed system. In comparison, Liu et al. reported 15 FP/case computer-aided detection of exophytic renal lesions [18].

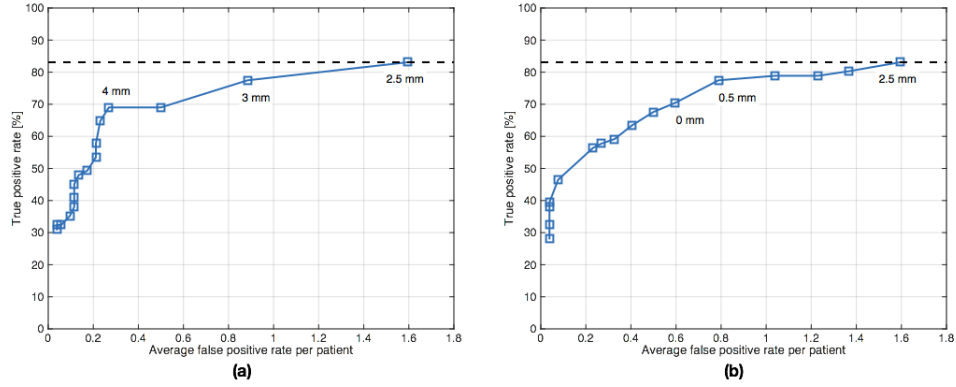


Figure 3.4: The trade-off between the TP rate and the average FP rate per patient as a function of (a) the radius threshold and (b) the adjacency threshold. The black line is the maximal TP rate that can be achieved on the given data-set with the proposed algorithm.

The kidney segmentation algorithm proves to be a good start for the purpose of cyst detection. However, a more accurate segmentation would allow us to reduce the adjacency threshold and eliminate a significant amount of the false seeds without missing true cysts. In addition, performing the CT scan with lower spacing between slices would allow us to increase the radius threshold and reduce the false positive rate as well. While these improvements may eliminate false candidates, the classification network itself is still responsible for the misclassification of 14% of the cysts. Therefore, improving the network’s recall and precision is still a major goal. Training the network with more positive examples and utilizing other augmentations of the data (e.g scaling and noise addition) may help achieve this goal. It is also possible to enrich the data-set by generating synthetic examples using a Generative Adversarial Network.

4 Tumor detection

4.1 *Data-set*

After addressing the problem of cyst detection, we approached the challenge of renal tumor detection. Like cysts, tumors vary in size, location and quantity. But unlike cysts, tumors have an inhomogeneous consistency and an undefined shape and borders, making their detection and segmentation far more complex. For this task, a data-set of 58 CT cases presenting 58 tumors was used. In each scan the tumor was manually annotated by a senior radiologist. Both kidneys were annotated as well. Fig. 4.1 exhibits the high variability of the tumors in our data-set. All scans were taken with contrast agent injection at the portal phase of enhancement. The pixel spacing ranges from 0.6836 to 0.9766 mm and the spacing between slices ranges from 1 to 3 mm.

4.2 *Experiments and results*

Eager/Excited to examine the UNet performance on another segmentation task, we started by training the same architecture to segment tumors in axial 2D slices. We used the exact same parameters as for the kidney segmentation and also the same pre-processing steps. The network learned the training examples pretty well, but the recall and precision on the validation set were low. Moreover, the validation accuracy was not correlated with that of the training set - a clear indication for overfitting.

To handle the overfitting, we tried two approaches. The first approach was to dramatically reduce the capacity of the network. Instead of computing two 3x3

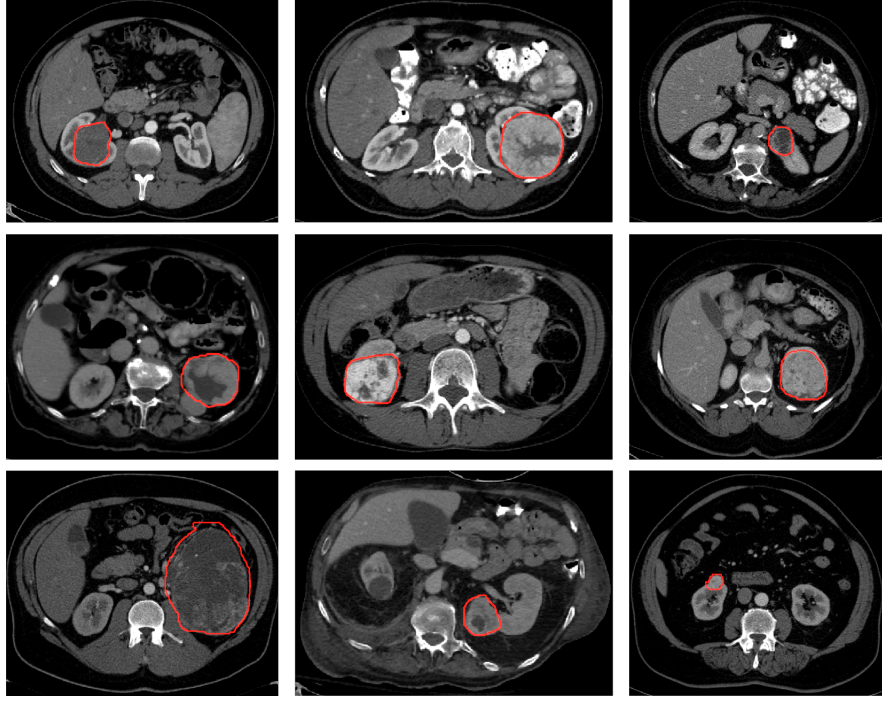


Figure 4.1: Sample tumors from our data-set demonstrating the high variability of cancerous renal tumors.

convolutions between the pooling and deconvolution layers, we computed only one, hence cutting the number of parameters almost in half. The results remained practically the same and the problem of overfitting was not resolved. The second approach was to add more data by exploiting 2D slices from all orientations - axial, sagittal and coronal - adding about 10 times more images that include tumors to the data-set. As a result, the recall and precision on the training set improved, but the network's performance on the validation set remained poor. The above-mentioned experiments were conducted using MatConvNet framework, and the results are presented below in Fig. 4.2.

At this point we decided to implement a 3D UNet architecture. Due to memory limitations, we used tensorflow framework [1]. A 3D volume of 128x128x64 voxels was extracted around the center of each kidney and was fed as an input to the 3D network. Since the number of tumor voxels might be very small compared to the volume size, we separated the loss into tumor loss and background loss and tested different linear combinations of the two. The first combination we

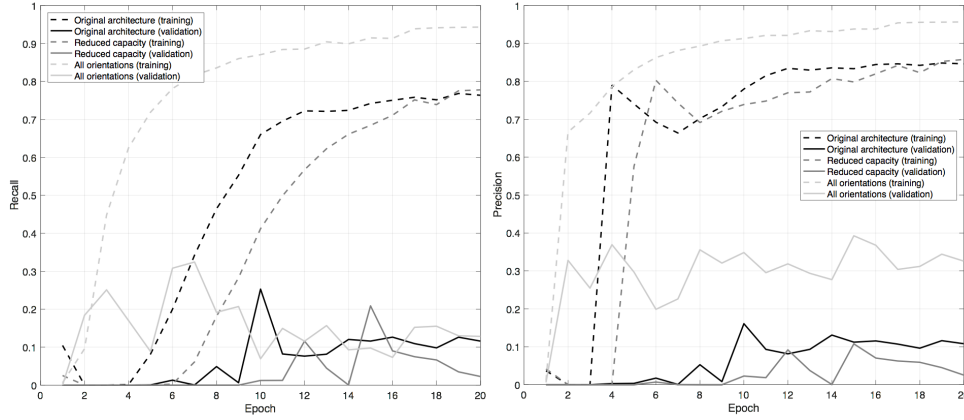


Figure 4.2: Recall and precision of tumor segmentation on training and validation data-sets using a 2D UNet-based architecture.

tested was giving the same weight to the tumor loss and background loss (50-50). The second combination we tested was 20-80, giving lower weight to the tumor loss. As in our 2D experiments, we clipped the HU of the volumes to the range $[-100, 400]$ and normalized them to $[-1, 1]$. We trained the network on 48 kidneys with tumors for 300 epochs with a mini-batch size of 4 volumes. We used Adam optimizer with learning rate of 0.0001. Segmentation accuracy was evaluated at the end of each epoch on the remaining 16 kidneys with tumors. As can be seen in Fig. 4.3, when the loss was composed of 20% tumor loss and 80% background loss, the learning process seemed promising for the first 100 epochs. From this point on, the performance on the training set continued to improve while the performance on the validation set remained the same and later deteriorated. Meaning, we were not able to overcome the overfitting obstacle by extending the architecture to 3D.

4.3 Discussion

At the beginning of this chapter, in Fig. 4.1, 9 examples of cancerous renal tumors were presented. These 9 examples account for 15% of the samples we have, and yet they differ greatly from one another. It is difficult to find one clear and accurate characteristic to describe them all. They are not of the same size, or at the same position, they don't have the same gray-level intensities or the same texture,

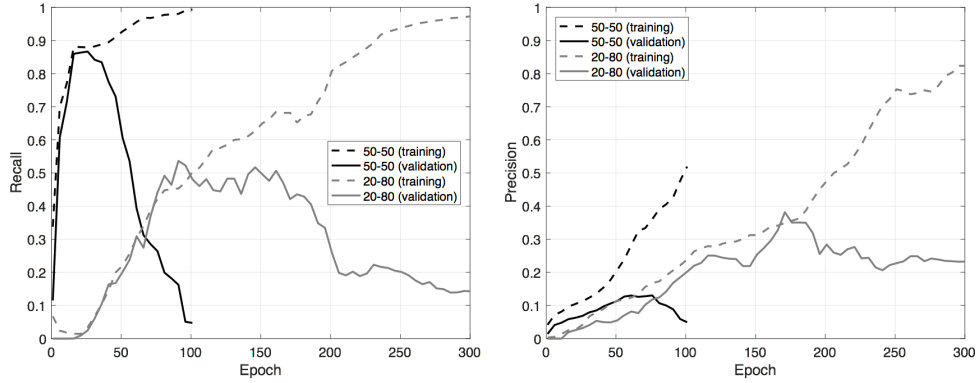


Figure 4.3: Recall and precision of tumor segmentation on training and validation data-sets using a 3D UNet-based architecture.

and they have different shapes and structures. Certainly, we haven't exhausted all research directions regarding detection and segmentation of renal tumors, yet in my opinion it is obvious that our current data-set is not large enough in order to learn from examples. So first and foremost I would recommend to continue collecting more cases to be annotated by the radiologists and perhaps even classified into sub-categories.

To my non-professional eyes, the main common denominator to all examples is the strong deformation of the kidney caused by their presence. In the "lighter" cases, the kidney boundaries become less smooth, or the kidney's homogeneous texture is interrupted. In other cases, the deformation is so severe, making it hard to even recognize that there is a kidney hiding somewhere beneath the tumor. It is often difficult to identify structural changes in the kidney based on a single CT slice, surely if the kidney is not even evident in that slice. Therefore, I think that utilizing a 3D network is worthwhile, and perhaps even essential for the purpose of tumor detection. Same as the doctor usually examines several consecutive slices of the scan together - using a 3D input can provide contextual information that will allow the network to tell where a kidney was supposed to be even if it isn't, and identify the tumor as an anomaly. Of course, causing deformation is not unique to tumors alone, but it can help detect all kinds of lesions. Their classification can then be handled separately.

In my experiments to train the network to segment the tumors, I did not at

all use the kidney annotations provided by the radiologist. My naive approach was that if identifying the kidney is essential for detecting renal tumors, then the network will learn to identify kidneys even without their annotations. However, it is possible that learning a separate *kidney* class would have contributed to learning the *tumor* class. Another way to achieve that is to employ transfer-learning and use the kidney segmentation network as a feature extractor or a starting point for the tumor segmentation network. A network that is able to segment kidneys may differentiate between healthy and sick kidney voxels more easily.

One more aspect that is worth discussing, is the network architecture. In my experiments, only the UNet architecture was tested, in 2D and 3D versions. There are many other known architectures that excel in computer vision tasks in general and organ segmentation in medical images in particular. One of these architectures may be more suitable for the task of tumor detection and outperform the UNet. Yet at this point I do not believe that the network architecture is our limitation - the small size of our data-set is.

5 Conclusions

We proposed a fully automatic method for kidney segmentation and simple renal cyst detection. Overall, the proposed algorithm is robust and efficient with an average run time of 73 seconds per scan. The algorithm is composed of two separate and independent modules. The first module is responsible for the kidney segmentation. For this purpose we employed an FCN which proved to be suitable for the task. With a larger data-set for training and more accurate annotations, we believe this method could achieve state-of-the-art performance. The kidney segmentation module can provide an infrastructure and a starting point for other automated tools for medical diagnosis of renal diseases.

The second module automatically detects simple renal cysts. It can be easily extended by automatic follow-up, quantification and measurement tools that so far required radiologist intervention for the detection of the cysts. The cyst detection module depends on the kidney segmentation as an input. However, it does not depend on the proposed kidney segmentation module. If another method for kidney segmentation is available, which provides a faster or more accurate segmentation, replacing it would be transparent to the cyst detection module.

We also began exploring the problem of cancerous renal tumor detection, but we were only scratching the surface. Certainly, detecting cancerous tumors is an important objective of high medical relevance. We trained our kidney segmentation network to segment renal tumors, and tested several modifications of the 2D architecture. We tried extending the same architecture to three dimensions as well. All experiments resulted in severe overfitting. These results, along with the visible high variability of the data led us to believe that the data-set must be significantly larger. We still believe that deep convolutional networks may be applicable to this problem, but given the low availability of examples to learn from, devising a more

traditional computer vision algorithm might result in more satisfactory detection rates.

References

- [1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL <https://www.tensorflow.org/>. Software available from tensorflow.org.
- [2] Pawel Badura, Wojciech Wieclawek, and Bartlomiej Pycinski. Automatic 3d segmentation of renal cysts in ct. In *Information Technologies in Medicine*, pages 149–163. Springer, 2016.
- [3] S Battiato, GM Farinella, G Gallo, O Garretto, and C Privitera. Objective analysis of simple kidney cysts from ct images. In *2009 IEEE International Workshop on Medical Measurements and Applications*, pages 146–149, 2009.
- [4] Patrick Ferdinand Christ, Mohamed Ezzeldin A Elshaer, Florian Ettlinger, Sunil Tataavarty, Marc Bickel, Patrick Bilic, Markus Rempfler, Marco Armbruster, Felix Hofmann, Melvin D’Anastasi, et al. Automatic liver and lesion segmentation in ct using cascaded fully convolutional neural networks and 3d conditional random fields. In *MICCAI*, pages 415–423. Springer, 2016.

- [5] Chengwen Chu, Masahiro Oda, Takayuki Kitasaka, Kazunari Misawa, Michitaka Fujiwara, Yuichiro Hayashi, Yukitaka Nimura, Daniel Rueckert, and Kensaku Mori. Multi-organ segmentation based on spatially-divided probabilistic atlas from 3d abdominal ct images. In *MICCAI (2)*, pages 165–172, 2013.
- [6] Antonio Criminisi, Duncan Robertson, Ender Konukoglu, Jamie Shotton, Sayan Pathak, Steve White, and Khan Siddiqui. Regression forests for efficient anatomy detection and localization in computed tomography scans. *Medical image analysis*, 17(8):1293–1303, 2013.
- [7] Rémi Cuingnet, Raphael Prevost, David Lesage, Laurent D Cohen, Benoît Mory, and Roberto Ardon. Automatic detection and segmentation of kidneys in 3d ct images using random forests. In *MICCAI*, pages 66–74. Springer, 2012.
- [8] Qi Dou, Hao Chen, Yueming Jin, Lequan Yu, Jing Qin, and Pheng-Ann Heng. 3d deeply supervised network for automatic liver segmentation from ct volumes. In *MICCAI*, pages 149–157. Springer, 2016.
- [9] Ben Glocker, Olivier Pauly, Ender Konukoglu, and Antonio Criminisi. Joint classification-regression forests for spatially structured multi-object segmentation. *Computer Vision–ECCV 2012*, pages 870–881, 2012.
- [10] Yukio Homma, Kazuki Kawabe, Tadaichi Kitamura, Yoji Nishimura, Mitsuru Shinohara, Yasushi Kondo, Isao Saito, Shigeru Minowada, and Yasuyuki Asakage. Increased incidental detection and reduced mortality in renal cancer? recent retrospective analysis at eight institutions. *International journal of urology*, 2(2):77–80, 1995.
- [11] Peijun Hu, Fa Wu, Jialin Peng, Yuanyuan Bao, Feng Chen, and Dexing Kong. Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets. *Int. J. Comput. Assist. Radiol. Surg.*, pages 1–13, 2016.

- [12] Mary Jayson and Holt Sanders. Increased incidence of serendipitously discovered renal cell carcinoma. *Urology*, 51(2):203–205, 1998.
- [13] D-Y Kim and J-W Park. Computer-aided detection of kidney tumor on abdominal computed tomography scans. *Acta radiologica*, 45(7):791–795, 2004.
- [14] Cheryl T Lee, Jared Katz, Paul A Fearn, and Paul Russo. Mode of presentation of renal cell carcinoma provides prognostic information. In *Urologic Oncology: Seminars and Original Investigations*, volume 7, pages 135–140. Elsevier, 2002.
- [15] Gareth Lewis and Alexander P. Maxwell. Early diagnosis improves survival in kidney cancer. *Practitioner*, 256(1748):13–16, 2012.
- [16] Marius George Linguraru, Jianhua Yao, Rabindra Gautam, James Peterson, Zhixi Li, W Marston Linehan, and Ronald M Summers. Renal tumor quantification and classification in contrast-enhanced abdominal ct. *Pattern recognition*, 42(6):1149–1161, 2009.
- [17] Marius George Linguraru, Shijun Wang, Furhawn Shah, Rabindra Gautam, James Peterson, W Marston Linehan, and Ronald M Summers. Automated noninvasive classification of renal cancer on multiphase ct. *Medical physics*, 38(10):5738–5746, 2011.
- [18] Jianfei Liu, Shijun Wang, Marius George Linguraru, Jianhua Yao, and Ronald M Summers. Computer-aided detection of exophytic renal lesions on non-contrast ct images. *Medical image analysis*, 19(1):15–29, 2015.
- [19] Lorenzo G Luciani, Roberto Cestari, and Carlo Tallarigo. Incidental renal cell carcinoma? age and stage characterization and clinical implications: study of 1092 patients (1982–1997). *Urology*, 56(1):58–62, 2000.
- [20] Calvin R Maurer, Rensheng Qi, and Vijay Raghavan. A linear time algorithm for computing exact euclidean distance transforms of binary images in

arbitrary dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):265–270, 2003.

- [21] OECD. Health care utilisation. *OECD Health Statistics (database)*, 2017. doi: <http://dx.doi.org/10.1787/data-00542-en>. URL [/content/data/data-00542-en](http://content/data/data-00542-en).
- [22] Nanzhou Piao, Jong-Gun Kim, and Rae-Hong Park. Segmentation of cysts in kidney and 3-d volume calculation from ct images. *Int. J. Comput. Graph. Animation*, 5(1):1, 2015.
- [23] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, pages 234–241. Springer, 2015.
- [24] Holger R Roth, Le Lu, Ari Seff, Kevin M Cherry, Joanne Hoffman, Shijun Wang, Jiamin Liu, Evrim Turkbey, and Ronald M Summers. A new 2.5 d representation for lymph node detection using random sets of deep convolutional neural network observations. In *MICCAI*, pages 520–527. Springer, 2014.
- [25] Ronald M Summers, Cecily ML Agcaoili, Matthew J McAuliffe, Sarang S Dalal, Peter J Yim, Peter L Choyke, MM Walther, and W Marston Linehan. Helical ct of von hippel-lindau: semi-automated segmentation of renal lesions. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 2, pages 293–296. IEEE, 2001.
- [26] Epidemiology Surveillance and End Results Program. Kidney and renal pelvis cancer. *National Cancer Institute*, 2017. URL <https://seer.cancer.gov/statfacts/html/kidrp.html>.
- [27] Maarten W Taal, Glenn M Chertow, Philip A Marsden, Karl Skorecki, SL Alan, and Barry M Brenner. *Brenner and Rector’s the Kidney*. Elsevier Health Sciences, 2011.

- [28] William Thong, Samuel Kadoury, Nicolas Piché, and Christopher J Pal. Convolutional networks for kidney segmentation in contrast-enhanced ct scans. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, pages 1–6, 2016.
- [29] A. Vedaldi and K. Lenc. Matconvnet – convolutional neural networks for matlab. In *Proceeding of the ACM Int. Conf. on Multimedia*, 2015.
- [30] Robin Wolz, Chengwen Chu, Kazunari Misawa, Kensaku Mori, and Daniel Rueckert. Multi-organ abdominal ct segmentation using hierarchically weighted subject-specific atlases. In *MICCAI*, pages 10–17. Springer, 2012.
- [31] Guanyu Yang, Jinjin Gu, Yang Chen, Wangyan Liu, Lijun Tang, Huazhong Shu, and Christine Toumoulin. Automatic kidney segmentation in ct images based on multi-atlas image registration. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 5538–5541. IEEE, 2014.
- [32] Xiangrong Zhou, Takaaki Ito, Ryosuke Takayama, Song Wang, Takeshi Hara, and Hiroshi Fujita. Three-dimensional ct image segmentation by combining 2d fully convolutional network with 3d majority voting. In *International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*, pages 111–120. Springer, 2016.