Recent Advances in Medical Image Processing for the Evaluation of Chronic Kidney Disease
Israa Alnazer, Pascal Bourdon, Thierry Urruty, Omar Falou, Mohamad Khalil, Ahmad Shahin, Christine Fernandez-Maloigne

To cite this version:

Israa Alnazer, Pascal Bourdon, Thierry Urruty, Omar Falou, Mohamad Khalil, et al.. Recent Advances in Medical Image Processing for the Evaluation of Chronic Kidney Disease. Medical Image Analysis, 2021, pp.101960. 10.1016/j.media.2021.101960. hal-03107259

HAL Id: hal-03107259
https://hal.science/hal-03107259
Submitted on 13 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License
Recent Advances in Medical Image Processing for the Evaluation of Chronic Kidney Disease

Israa Alnazer¹²³, Pascal Bourdon¹², Thierry Urruty¹², Omar Falou³⁴⁵, Mohamad Khalil⁳, Ahmad Shahin³ and Christine Fernandez-Maloigne¹²

¹XLIM-ICONES, UMR CNRS 7252, Université de Poitiers, France
²Laboratoire commune CNRS/SIEMENS I3M, Poitiers, France
³AZM Center for Research in Biotechnology and its Applications, EDST, Beirut, Lebanon
⁴American University of Culture and Education, Koura, Lebanon
⁵Lebanese University, Faculty of Science, Tripoli, Lebanon

Corresponding Author: Israa Alnazer, E-mail: israa_alnazer@hotmail.com

Abstract: Assessment of renal function and structure accurately remains essential in the diagnosis and prognosis of Chronic Kidney Disease (CKD). Advanced imaging, including Magnetic Resonance Imaging (MRI), Ultrasound Elastography (UE), Computed Tomography (CT) and scintigraphy (PET, SPECT) offers the opportunity to non-invasively retrieve structural, functional and molecular information that could detect changes in renal tissue properties and functionality. Currently, the ability of artificial intelligence to turn conventional medical imaging into a full-automated diagnostic tool is widely investigated. In addition to the qualitative analysis performed on renal medical imaging, texture analysis was integrated with machine learning techniques as a quantification of renal tissue heterogeneity, providing a promising complementary tool in renal function decline prediction. Interestingly, deep learning holds the ability to be a novel approach of renal function diagnosis. This paper proposes a survey that covers both qualitative and quantitative analysis applied to novel medical imaging techniques to monitor the decline of renal function. First, we summarize the use of different medical imaging modalities to monitor CKD and then, we show the ability of Artificial Intelligence (AI) to guide renal function evaluation from segmentation to disease prediction, discussing how texture analysis and machine learning techniques have emerged in recent clinical researches in order to improve renal dysfunction monitoring and prediction. The paper gives a summary about the role of AI in renal segmentation.

Keywords: Chronic Kidney Disease, Textural Analysis, Deep Learning, Automatic Renal Segmentation.

Introduction

Chronic Kidney Disease (CKD) is today considered one of the major public health challenges by the World Health Organization (WHO). In France, it affects nearly 82,000 people and the number of patients affected increases by 2% each year (Bayat et al., 2010). CKD remains a widespread and crucial public health problem afflicting more than 12% of population worldwide (Jiang and Lerman, 2019). It is characterized by progressive and irreversible deterioration of renal function accompanied with a low Glomerular Filtration Rate (GFR) leading to end stage renal disease, when kidneys are totally damaged and cannot filter the blood the way they should and either renal dialysis or transplant is needed (Jiang and Lerman, 2019).

Two options are therefore possible when the disease reaches a terminal stage (Chronic Terminal Renal Insufficiency - IRCT). The most common is hemodialysis. This blood filtering technique remains restrictive to date, the patient having to be dialyzed over a period of 4 hours, 3 times a week. The other solution offered to people with CRTI is kidney transplantation. It is widely accepted that the transplant offers a better quality of life compared to dialysis. Indeed, the transplant allows finding an almost normal renal activity while maintaining a lifestyle close to “normal”. The early detection of CKD offers the ability to guide patient management as well as reduce mortality rate by
preventing the progression to the end point, normally associated with many health complications including heart and bone disease and high blood pressure (Saha et al., 2019).

Kidneys filtrate minerals and metabolites from blood as urine (Notohamiprodjo et al., 2010). CKD is accompanied with the accumulation of extra cellular matrix (ECM) or fibrosis-related deposits that change the way the tissue is handling water, as well as tissue stiffness and macromolecule content (Notohamiprodjo et al., 2010). The loss of peritubular capillaries induces a decrease of renal oxygenation and perfusion (Nangaku, 2006). Renal function has been commonly evaluated either by the estimation of Glomerular Filtration Rate (eGFR) based on blood serum Creatine level (sCr) or by the gold standard biopsy, followed by histopathology analysis. However, these medical acts suffer from several drawbacks and limitations. sCr is a late indicator of renal function decline and requires blood specimens to be analyzed. Furthermore, GFR estimation equations lack precision and cannot offer a split renal function measurement (Notohamiprodjo et al., 2010). Although biopsy stands for the gold standard method to assess renal microstructure and quantify the cause of renal dysfunction, it suffers from other drawbacks as well. First, the invasive procedure is associated with bleeding and possible pain to the patient, thus limiting follow-up evaluation. Secondly, biopsy is subject to sampling bias because it samples less than 1% of one of the kidneys where medulla is often not included, as it is limited in terms of spatial resolution (Baues et al., 2017; Leung et al., 2017).

Radiology can provide structural and functional markers that were reported to help in the prediction as well as the follow-up of renal dysfunction. It plays an increasingly more important role in the assessment of renal function. The recently used medical imaging modalities are summarized in Table 1. Magnetic Resonance Imaging (MRI) holds the potential to assess renal microstructure organization, diffusion, perfusion and oxygenation. Additionally, it provides the ability to measure renal hemodynamic, to quantify tissue relaxation time, macromolecule and elasticity and to characterize renal tissue metabolites.

<table>
<thead>
<tr>
<th>Medical Imaging Modality</th>
<th>Description</th>
<th>Biomarker Measured</th>
<th>Informed Biological Marker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffusion weighted imaging (DWI)</td>
<td>Reflects the microstructure of the tissue by describing the water restriction and diffusivity within this tissue</td>
<td>Apparent diffusion coefficient</td>
<td>ADC, ∆-ADC, D, ∆-D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>True diffusion</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pseudo diffusion</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flowing fraction</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interstitial fibrosis, changes in renal perfusion or water handling</td>
</tr>
<tr>
<td>Diffusion tensor imaging (DTI)</td>
<td>Assessment of microstructure organization of a spatial oriented tissue by quantifying the diffusivity direction</td>
<td>Fractional anisotropy</td>
<td>FA, MD, ADC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean diffusivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Changes in renal microstructure that induces change in the water diffusion direction</td>
</tr>
<tr>
<td>Blood oxygenation level dependent (BOLD)</td>
<td>Assessment of tissue oxygenation by using the paramagnetic property of deoxyhemoglobin</td>
<td>Apparent transverse relaxation time T₂* or R₂* (1/ T₂*), R₂* slope</td>
<td>Changes in renal oxygenation, hydration status, sodium balance</td>
</tr>
<tr>
<td>Arterial spin labeling (ASL)</td>
<td>Measurement of tissue perfusion or blood flow using the magnetically labeled water within blood as a contrast agent</td>
<td>Tissue perfusion</td>
<td>f</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Filtration fraction</td>
<td>FF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Renal perfusion alteration</td>
</tr>
<tr>
<td>Dynamic contrast enhanced (DCE)</td>
<td>Measurement of kidney hemodynamic using the administration of contrast agent in the tissue which changes its T₁ relaxation time</td>
<td>Single kidney GFR</td>
<td>SK-GFR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Renal plasma flow</td>
<td>RPF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tubular transit time</td>
<td>Tubular MTT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Plasma mean transit time</td>
<td>Plasma MTT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tubular flow</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Plasma volume</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Split renal function, perfusion, filtration fraction</td>
</tr>
<tr>
<td>T1 mapping</td>
<td>Tissue characterization by measuring its spin-lattice relaxation time</td>
<td>T1</td>
<td>Cortex, medulla, cortico-medullary differentiation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Changes in the molecular (lattice) environment (e.g. water content, fibrosis), changes in oxygenation level</td>
</tr>
</tbody>
</table>
Table 1 Medical imaging in renal function assessment

Ultra sounds Elastography (UE) allows the identification of structural information regarding tissue mechanical properties. Computed Tomography (CT) can provide anatomic and functional information, but it is limited to x-ray exposure and contrast material injection. Scintigraphy combines functional and perfusion measurement using radioactive tracers. Qualitative analysis of those medical images has shown a promising tool in the evaluation of kidney function.

Interestingly, the quantification of medical images that analyses image pixels, also known as textural analysis, was investigated as a practical complementary tool. Obtained from mathematical equations, textural analysis studies spatial arrangement of gray-level pixels and shows their relationships that are generally hidden to the human eye. Given the renal architecture and the influence of renal disease on the distribution of functional markers, textures hold the potential to reflect histopathological heterogeneity (Ding et al., 2019; Shi et al., 2018). Introducing textural analysis with conventional machine learning methods extends the medical imaging capability by offering additional numerical descriptors that could be useful in the diagnosis and prediction of renal dysfunction.

Most importantly, deep learning was recently investigated in clinical applications that aim to diagnose renal function and CKD patients. It is a branch of machine learning inspired by the biological neurons that are composed of many layers of nodes interconnected offering a new robust tool for image feature extraction and patient classification.
In this survey, we first review the use of medical imaging techniques in the evaluation of renal dysfunction, especially CKD. Then, we show the potential of AI to improve the performance of renal dysfunction prediction and diagnosis by summarizing the recent application of textures and machine learning techniques including neural networks on different medical imaging modalities. The role of AI in renal segmentation is also discussed, since accurately identifying kidneys on medical images is an essential step that could save time and turn the renal parenchyma delineation to a subject-independent issue as it is required either prior to human-expert or computer-aided diagnosis.

Magnetic Resonance Imaging

1. Diffusion MRI

   1.1. Diffusion Weighted Imaging

Renal fibrosis, characterized by the deposition of extra cellular matrix (ECM), plays a key role in the development of CKD (Leung et al., 2017; Nangaku, 2006). Water molecule mobility is supposed to be decreased within the tissue due to the involvement of tissue cellularity. Thus, water diffusion caption may reflect tissue microstructure. Diffusion Weighted Imaging (DWI) is a magnetic resonance modality, which uses the motion of water molecules as a contrast providing an in-vivo measurement of water diffusion or Brownian motion. It uses powerful bipolar magnetic gradients in order to create a sensitivity of the received signal to the water movement, thus describing the way the tissue restricts water (Notohamiprodjo et al., 2010). The Apparent Diffusion Coefficient (ADC) is a DWI biomarker corresponding to an overall measure of water diffusion and microcirculation within the tissue (Notohamiprodjo et al., 2010). The Intra-Voxel Incoherent Motion (IVIM) model was introduced by Le Bihan et al. in order to separate the true diffusion or molecular water diffusion from pseudo diffusion or perfusion induced by blood flow in capillary network (Le Bihan et al., 1988). The IVIM derived parameters are the true diffusion (D), associated with extravascular water molecule motion or pure diffusion, pseudo diffusion ($D^*$), associated with intravascular water molecule motion or perfusion, and flowing fraction (f) (Le Bihan et al., 1988).

DWI has been reported as a good predictor of renal changes in diabetic kidneys (Deng et al., 2018; Feng et al., 2018), and a robust technique to monitor renal function after transplantation (Chen et al., 2018; Fan et al., 2019; Ren et al., 2016; Xie et al., 2018). Several studies have shown the potential use of DWI as a promising imaging technique for renal fibrosis evaluation in animal models (Cai et al., 2016; Hennedige et al., 2015), and CKD human studies (Ding et al., 2016; Friedli et al., 2017; Gaggioli et al., 2007; Ichikawa et al., 2013; Q. Li et al., 2014; Mao et al., 2018a, 2018b; Xu, 2010). ADC values were correlated with renal function, usually evaluated by creatine level or eGFR (Ding et al., 2016; Gaggioli et al., 2007; Mao et al., 2018a, 2018b; Xu, 2010), as well as with renal fibrosis and pathology score when biopsy is available (Cai et al., 2016; Ebrahimi et al., 2014a; Friedli et al., 2017; Q. Li et al., 2014; Mao et al., 2018a, 2018b). When associated with kidney function, DWI parameters were found to decrease with severe renal damage (reflected by decreased eGFR) and a strong correlation was found between increasing fibrosis and declined ADCs. The decreased parameters can be attributed to the reduced perfusion, the presence of interstitial fibrosis restricting water as well as the decrease of vascularity (Hennedige et al., 2015). Furthermore, the difference between cortical and medullary DWI parameters ($\Delta$-ADC, $\Delta$-D) were well negatively correlated to the percentage of fibrosis (Friedli et al., 2017). ADCs differed significantly among groups (controls, mildly, moderately and severely impaired) classified according to pathology score based on proliferation fraction in the glomerular segment, which is not the case when classification is based on pathology type (nephropathy, minimal lesion glomerulonephritis, focal segmental proliferative glomerulonephritis, membranous nephropathy, mesangial proliferative Glomerulonephritis, glomerulosclerosis and crescentic glomerulonephritis). This can be referred to the fact that different CKD pathology type share similar pathogenic features contributing ADCs decrease (Q. Li et al., 2014). However, ADC values failed to early detect CKD as no ability to differentiate healthy versus stage I CKD was mentioned, while perfusion-related $D^*$ did in one study supporting that renal impairment is accompanied with reduced perfusion and that IVIM maps are supposed to allow early detection of renal dysfunction more than ADC maps (Ichikawa et al., 2013). On the other hand, performance of the detection of renal damage ($\text{eGFR} < 30 \text{ ml/min/1.73 } \text{ m}^2 \text{ or biopsy-proven}$) was better when using ADCs than IVIM parameters (Ding et al., 2016; Friedli et al., 2017). A significant decline in mean ADC values was
found among patients with advanced stages of Diabetic Nephropathy (DN), a main cause leading to renal failure, compared to diabetic patients with no nephropathy (Jawad, 2019). Recently, DWI was proved to be promising to differentiate healthy and CKD children (glomerulonephritis, hemolytic uremic syndrome, lupus nephritis (LN), nephronophthisis and infantile nephrosis were the main causes of CKD) (Emad-Eldin et al., 2020). ADCs were negatively correlated to CKD stages (stage I to V) (Emad-Eldin et al., 2020). Cortical as well as medullary ADCs were reported to be decreased significantly with the presence of renal allograft dysfunction (mean eGFR of 30 ml/min/1.73 m²) associated with fibrosis deposition (Bane et al., 2020).

Studies that incorporated DWI in clinical applications are still limited. Most of these studies evaluated the association between ADCs and fibrosis or pathology score. Few studies were focusing on using DWI to discriminate between healthy and CKD or DN patients as well as to detect renal allograft dysfunction. Although, unlike biopsy, DWI offers the ability to assess whole-kidney perfusion and diffusion, it is still limited to respiratory motion artifacts, protocol variability, inter-subject variability (e.g. patient preparation) and the contribution of other factors besides fibrosis to water mobility (e.g. Urine flow rate, medications, vascular volume), therefore no clear cut-off value of ADC can be used (Gaggioli et al., 2007; Notohamiprodjo et al., 2010; Sulkowska et al., 2015). Future studies that address these limitations and which further explore the potential of DWI in the evaluation of renal dysfunction in CKD patients as well as the comparison of ADCs and IVIM classification performance are clearly needed.

### 1.2. Diffusion Tensor Imaging

Renal fibrosis or matrix deposition was expected to not only decrease water molecules mobility within renal tissue, but also to disrupt the parenchymal ordered structures that favor the mobility of water in a specific direction (Leung et al., 2017). Diffusion Tensor Imaging (DTI) is an advanced diffusion MR technique that allows the measurement of water molecules mobility along different axes, giving the opportunity to assess the renal microarchitecture organization and to capture the ordered structures disruption (Leung et al., 2017). DTI features the diffusion measurement in at least six directions, from which the Fractional Anisotropy (FA) as well as the mean diffusivity (MD) can be derived. FA represents a measure of diffusion anisotropy and ranges from 0 to 1 as an index of the degree of anisotropy or diffusion directionality, whereas MD is equivalent to the ADC derived from DWI technique which reflects the diffusion magnitude (Notohamiprodjo et al., 2010). Moreover, 3D reconstruction of diffusion direction can be performed using tractography technique that combines the direction and the anisotropy index as shown in Fig. 1 (Notohamiprodjo et al., 2010).

FA of renal pyramids has been reported to be higher than cortical FA suggesting packed fibers and reflecting tubular structural arrangement or tubular flow within the medulla which is normally highly structured (Feasibility et al., 2010; Kataoka et al., 2009; Nicolaescu et al., 2017; Notohamiprodjo et al., 2010; Ries et al., 2001). DTI was used to evaluate renal disease in mouse models and was found to be a good identifier of renal pathology and renal fibrosis (Hueper et al., 2012; J.-Y. Kaimori et al., 2017; J. Y. Kaimori et al., 2017). In the prediction of DN, a leading cause to end stage renal disease, DTI has shown promise in early detection of renal infection even when eGFR remained normal (>60 mL/min/1.73 m²). FA was correlated well with eGFR and significantly lower in DN patients (eGFR >or<60 mL/min/1.73 m²) compared to controls (Lu et al., 2011). Moreover, Medullary FA has been reported as a good biomarker of renal function in kidney allograft where it has revealed a good correlation with eGFR as well as the amount of renal fibrosis (Hueper et al., 2016; Lanzman et al., 2013; Palmucci et al., 2015). When histology is performed, FA was correlated with percent of glomerulosclerosis and the area of interstitial fibrosis as well as the degree of renal impairment (stages 1 to S) (Feng et al., 2015).
In CKD, a significant decline in medullary FA was reported in patients with chronic renal impairment (eGFR <60 mL/min/1.73 m$^2$) (Saini et al., 2018), renal failure caused by: nephroangiosclerosis, renal artery stenosis, medullary kidney cystic disease, DN, pyelonephritis, lithiasis, acute glomerulonephritis, analgesic abuse, interstitial nephritis, Wegener granulomatosis, LN, diuretic abuse (Gaudiano et al., 2013)) compared to control with normal renal function, and strongly correlated to eGFR (Gaudiano et al., 2013; Saini et al., 2018; W. Wang et al., 2014). Tractography has also shown a good visual discrimination between the two subjects, where reduced number of tracts which do not have preferential orientation was presented with worsening renal function compared to a regular tracts arrangement with normal renal function (Gaudiano et al., 2013). Ye et al., DTI has succeeded in early detection of CKD progression in diabetes mellitus (DM). Cortical and medullary FA have shown a significant decline in stage 1 CKD compared to control group and a good correlation with eGFR, demonstrating the ability of DTI to predict CKD progression (Ye et al., 2019).

However, DTI was not able to identify the change in medullary diffusion direction in early stage of DN (Feng et al., 2019). Medullary FA and mean track length were reported to be significantly decreased in children with autosomal recessive polycystic kidney disease (ARPKD) compared to healthy controls (Serai et al., 2019). Recently, DTI parameters, FA and track length were shown to be significantly different between kidneys with and without Uretero Pelvic Junction (UPJ) obstruction, supporting that DTI is able to assess parenchymal damage (Otero et al., 2020). Moreover, Gadolinium-based DTI was reported to be more accurate in measuring renal pathological characteristics, renal fibrosis and renal blood flow in CKD patients stage I and II (60<=eGFR<=90 mL/min/1.73 m$^2$) (Liu et al., 2020). Alterations in diffusion parameters caused by Renal Artery Stenosis (RAS), a renal disease promoting fibrosis by collagen deposition, could be detected by DTI were medullary FA was significantly reduced in patients with RAS (Gaudiano et al., 2020). In (Mrđanin et al., 2020), medullary FA were found to be lower in DM patients and positively correlated with the eGFR. Tractography has shown distributed structure in patients of renal impairment. Cortical as well as Medullary FA were correlated to allograft function and were significantly higher in transplants with good function (stage I, II) than those with dysfunction (stages III-V) (S. et al., 2020). Medullary and Cortico-Medullary Differentiation (CMD) of FA were reported to be well associated with eGFR for healthy and transplanted kidneys (Adams et al., 2020).

DTI was reported to reflect renal impairment severity. In summary, clinical applications that employed DTI aimed to discover the association between DTI parameters and eGFR or fibrosis, evaluate the use of DTI parameters as well as tractography to discriminate healthy volunteers from patients with renal impairment (DN, DM, UPJ obstruction, RAS), evaluate the potential use of DTI to early detect DN and DM, and test the value of DTI in reflecting renal allograft function. Medullary FA was found to be the main DTI marker of renal damage showing a decrease with the worsening of renal function. Many factors could be involved during renal impairment that can lead to the fall of FA, such as reduced tubular flow rate, tubular damage and vascular abnormalities (Lu et al., 2011).

2. Blood Oxygenation Level Dependent Imaging

Renal hypoxia has been recognized to take an important role in the progression of CKD (Nangaku, 2006). Blood Oxygenation Level Dependent MR imaging (BOLD) gives the opportunity to measure the tissue oxygenation level without using contrast materials. Briefly, BOLD-MRI uses the de-oxyhemoglobin paramagnetic properties in order to assess tissue oxygenation: the higher the local de-oxyhemoglobin levels are, the higher the apparent relaxation rate $R2^*$ ($s^{-1}$) is and the lower local tissue oxygen content is.
Several animal studies have shown the linear relationship between $R_2^*$ values and directly measured renal partial blood pressure ($pO_2$) which reflects the amount of oxygen gas dissolved in blood, using oxygen-sensitive electrodes. Thus, BOLD accurately measures the tissue $pO_2$ (Pruijm et al., 2018a). Many factors could alter renal oxygenation ($R_2^*$ values), such as hydration status, sodium balance and use of medication that might affect oxygen delivery or consumption. Thus, standardization of patient preparation is required, including fasting overnight 4-6h, a constant hourly water intake, control of salt intake, monitor of drug intake, record of blood pressure and medication before MRI examination (Pruijm et al., 2018a), (Neugarten, 2012).

In order to analyze BOLD images, four methods are used: Region of Interest (ROI) technique, twelve layer concentric object (TLCO) or ‘onion peel’ technique, the fractional tissue hypoxia and the compartmental method as shown in Fig. 2. ROI technique, as seen, requires the placement of different small ROIs (20-40 voxels) in the cortex and medulla regions to provide separate information. TLCO technique divides the renal parenchyma into 12 layers of equal thickness, where the $R_2^*$ mean value is calculated. From the manually selected kidney parenchyma, the fractional tissue hypoxia reports the percentage of $R_2^*$ values above a certain threshold. The compartmental method analyses the distribution of $R_2^*$ values in renal regions.

Hypoxia hypothesis assumed that renal tissue hypoxia or the decrease of tissue oxygenation is a final common pathway to CKD. At first, studies have failed to find a relationship between $R_2^*$ values and renal function estimated by the eGFR (Michaely et al., 2012; Pruijm et al., 2014). Even more, few studies have reported a decline or no change in cortical $R_2^*$ associated with renal function impairment (Djamali et al., 2007; Khatir et al., 2015; Textor et al., 2008; Z. J. Wang et al., 2011). Contrary, recent researchers have shown a good correlation between cortical $R_2^*$ or $R_2^*$ slope and eGFR, where greater $R_2^*$ values in patients with renal impairment (glomerulonephritis, hypertensive nephropathy, DN, Acute kidney Injury (AKI) and others) were limited to the cortex, supporting the hypoxia hypothesis (Inoue et al., 2011; Milani et al., 2017; Prasad et al., 2018; Pruijm et al., 2018b; Xin-Long et al., 2012; Yin et al., 2012). This dissimilarity within results can be referred to the lack of and even to the dependence of $R_2^*$ not only on intravascular de-oxyhemoglobin but on other factors such as renal edema and perfusion, changes in vascular and tubular volume fraction (Milani et al., 2017; Pruijm et al., 2017).

The BOLD images analysis technique can impact results and that was illustrated by Milani et al. where no changes of $R_2^*$ was shown between controls and CKD patients [eGFR<60 mL/min/1.73 m$^2$ or albuminuria <300 mg/24h for at least 3 months] using ROI technique, while there was a significant difference of $R_2^*$ slope using TLCO technique with the same group of patients (Milani et al., 2017). Similarly, regional $R_2^*$ were strongly correlated between the ROI and TLCO techniques (L.-P. Li et al., 2020). However, in terms of discriminating controls from CKD patients, the TLCO parameter $R_2^*$ slope has shown the largest sensitivity. The association between ADCs and $R_2^*$ demonstrated by Prasad et al. as well as the increase of $R_2^*$ with the increase of fibrosis staging found by Inoue et al. support the hypothesis of fibrosis being a way to decrease local oxygen consumption (Inoue et al., 2011; Prasad et al., 2018). The
evaluation of fibrosis by BOLD imaging was also verified in animal models, where R2* values correlated well with the percentage of fibrosis (Woo et al., 2018; Zha et al., 2019). The administration of furosemide inducing increase of R2* values has improved the differentiation of healthy and affected kidneys in the study of Pruijm et al. (Pruijm et al., 2014), but no positive impact was noted later (Milani et al., 2017; Pruijm et al., 2018b). It is worthy of note that the cortical R2* values were found to be strongly correlated with fluid fraction (FF) compared to GFR, having that FF=GFR/renal plasma flow in healthy volunteers. Thus, making the measurement of FF recommended in future studies (Van Der Bel et al., 2016). Reduced oxygenation reflected by increased R2* values were reported to be a clinically useful marker of CKD progression (Sugiyama et al., 2018). In their study, controls and DM patients were enrolled in a retrospective study over 5 years, where the rate of decline in eGFR was significantly correlated to R2* values. Medullary R2* values were shown to be sensitive for the prediction of early DN (Feng et al., 2019). Moreover, medullary R2* were reported to distinguish controls with mild renal impairment (CKD stages I, II stages) with a high sensitivity and specificity (92.3 and 85.2% respectively) (Li et al., 2019). Incorporating R2* values into machine learning algorithms was able to discriminate rejected from non-rejected transplants with a sensitivity of 100% (Shehata et al., 2019). Recently, cortical and medullary R2* values were positively correlated to serum creatinine level and negatively correlated to eGFR (Luo et al., 2020). In their study, cortical R2* value was ranked as CKD stage IV, V > CKD stage I-III > healthy controls and medullary R2* value was ranked as CKD stage IV, V > CKD stage I-III while no significant difference between controls and patients with CKD stage I-III.

Clinical application of the studies employing BOLD MRI imaging were limited to study the association between R2* values with eGFR, serum creatinine level, fibrosis and FF, evaluate the potential use of BOLD in discriminating controls from patients with renal impairment, evaluate the role of R2* values in detecting DN at early stages and discover the value of BOLD in reflecting renal allograft function. Giving that R2* slope derived from TLCO technique is somehow related to the R2* distribution within renal parenchyma, the door to further explore spatial arrangement within BOLD is opened. The usefulness of R2* in staging CKD or predicting renal function decline as well as measuring fibrosis require more validations with follow-up studies.

3. Arterial Spin Labelling

Since CKD is associated with impaired perfusion due to renal fibrosis and a decrease of renal blood flow, assessing renal perfusion might be a powerful tool in the detection of CKD as well as differentiate several CKD staging (Leung et al., 2017; Nangaku, 2006). Arterial Spin Labelling (ASL) in an MRI modality that allows tissue perfusion measurement without administration of gadolinium, using the water within blood as a contrast agent. Blood flow is labeled to have an opposite magnetization compared to destination tissue. The difference between a labeled and a non-labeled image can provide a perfusion-weighted image which signal intensity is proportional to perfusion (Artz et al., 2011; Leung et al., 2017; Odudu et al., 2018). Perfusion map can be then generated by entering the signal from each voxel into a kinetic model, as shown in Fig. 3.

Several studies have supported the feasibility of using ASL to evaluate kidney function either in CKD or in kidney transplant. Artz et al. have found a good correlation between cortical perfusion and eGFR for native and transplanted kidneys (P<0.05) supporting that glomerular filtration rate regulates the renal blood flow. Moreover, perfusion was found to be significantly reduced in transplanted kidneys compared to native kidneys when eGFR >60, which can be referred to the differential regulation of blood flow in transplanted kidneys and no statistical difference of perfusion was found between left and right kidneys (Artz et al., 2011).

Applied to patients with renal impairment, cortical and whole kidney ASL perfusion were significantly lower in CKD patients (staged from I to V according to their eGFR) than healthy subjects, reduced with the increase of renal impairment (Brown et al., 2019; Gillis, 2016; Rossi et al., 2012) with and without matching for age and body mass index (L. Li et al., 2017). Cortical and whole kidney perfusion were well correlated with eGFR (Gillis, 2016; L. Li et al., 2017). No association between renal volume or size and perfusion suggesting that the difference in tissue perfusion cannot be attributable to tissue atrophy (Gillis, 2016). The filtration fraction determined from ASL perfusion (FF=eGFR/ASL) was found to be associated with fibrosis score resulted from biopsy (Brown et al., 2019). Cortical perfusion was reported to be strongly correlated with renal function in healthy and CKD patients (15<=eGFR<60 mL/min/1.73 m²) (Buchanan et al., 2019). Recently, ASL values were found to be significantly different between
patients with primary glomerular disease (PGD) and healthy volunteers and well correlated to eGFR in PGD patients (C. Li et al., 2020).

ASL has been applied in the study of AKI (Dong et al., 2013), LN (Rapacchi et al., 2015; Skeoch et al., 2017), CKD (Breidthardt et al., 2015a; Gillis et al., 2016; L.-P. Li et al., 2017), DN (Mora-Gutiérrez et al., 2017), and renal transplant (Heusch et al., 2014; Hueper et al., 2015). From these studies, consistent outcomes are that cortical perfusion is reduced with the presence of renal impairment, declines with increasing stage of CKD and correlates to eGFR and fibrosis staging. Although its ability to assess renal dysfunction, more validations still warranted to evaluate the association between perfusion and histopathology in order to further explore the relation between reduced renal perfusion and ECM accumulation.

4. Dynamic Contrast Enhanced MRI

Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI), also referred as MR renography, is a MR perfusion technique that uses the administration of contrast chemical agent into the patient’s bloodstream to perfuse inside ilk, precipitated in recreation of flush, to catalyze a discrepancy image of organ (Derle and Dighe, 2015). The interaction between gadolinium-based agent and water protons within the target tissue causes T1 relaxation time to decrease leading to a higher intensity on T1 weighted images (Ebrahimi et al., 2014b). Successive images are acquired during transition of the contrast agent in order to track the dynamic enhancement of the target tissue. Mathematical models are used to estimate the tracer concentration from the MR signal change, as seen in Fig. 4 (Zhang and Lee, 2019), (Octavia et al., 2017). By fitting the model to the concentration vs. time curves of kidneys, perfusion and flow parameters can be estimated, including cortical and medullary renal plasma flow, plasma volume, mean transit times for each individual compartment and the whole kidney, tubular flow and tubular transit time and most importantly the single kidney GFR (SK-GFR) instead of bilateral GFR estimated by the serum creatinine blood level (Elkefjord et al., 2015; Jiang et al., 2019b; Octavia et al., 2017). DCE-MRI requires a proper segmentation and registration process for kidney motion correction (Derle and Dighe, 2015).
Despite its ability of kidney state visualization and functional kidney parameters estimation additionally to its potential for prognosis and diagnosis in living kidney donors (Dujardin et al., 2005; Eikefjord et al., 2016, 2015; Notohamiprodjo et al., 2011), the application of renal DCE-MRI on CKD patient is limited by the use of contrast agent and concerns about patient safety in renal impairment due to the risk of nephrogenic systemic fibrosis even when using low doses (Ebrahimi et al., 2014b), (Zhang and Lee, 2019), (Fraum et al., 2017). Thus, assessment of kidney perfusion in renal insufficiency was performed using ASL technique (Zhang and Lee, 2019), (Conlin et al., 2017).

5. T1 and T2 Mapping

Magnetic resonance relaxometry (MRR) can discriminate tissue composition by measuring the T1 (spin-lattice) and T2 (spin-spin) relaxation time via pixel-wise mapping of true T1 and T2 values of the target tissue without the administration of contrast agent. It was used in order to characterize renal tissue non-invasively.

T1 relaxation time was suggested to increase with progressive tissue scarring. It has shown a good correlation with cardiac fibrosis measured by biopsy and was significantly higher in patient with severe fibrosis staging (Bull et al., 2013), thus the potential of T1 maps in identifying fibrosis burden without contrast agent. Moreover, T1 values were found to be modulated by the tissue and/or blood oxygenation level in several studies where shown the sensitivity of cortical T1 values to oxygenation level changes (the higher the oxygenation level is, the lower the T1 values are) (Wolf et al., 2018). Moreover, T1 native mapping has shown a good to strong inter- and intra-examination reproducibility in both healthy and patients with diabetic nephropathy (Dekkers et al., 2019).

Renal T1 mapping was used as a non-invasive tool in the assessment of kidney function. In their study, Lee et al. have reported a higher T1 relaxation time in the cortex than the medulla due to the higher water mobility within medulla compared to cortex, an increase of cortical T1 with the renal insufficiency which can be referred to increased cortex water content resulted from the more longstanding pathologic changes such as extracellular matrix changes and a good association was found between cortical T1s, the difference between cortical and medullary T1s and the single kidney GFR (SK-GFR) (r=-0.5, P=0.03; r=0.58, P<0.01 respectively) (Lee et al., 2007). Breidthardt et al. have evaluated the association between renal dysfunction, perfusion (ASL) and real parenchymal structure (T1 maps) in healthy volunteers and patients with heart failure (HF) having a different level of renal impairment according to their eGFR (Breidthardt et al., 2015b). Main results have shown a good correlation between T1s and eGFR (r=0.41, P=0.015) and a higher cortical T1s in HF patients with kidney dysfunction compared to those without renal impairment. Similar results were obtained in the study of Gillis et al. where cortical T1s were found higher in clinically evaluated CKD patients compared to normal subjects and correlated to eGFR (r=0.58, P<0.001) (Gillis, 2016). Additionally, the potential use of T1 mapping in the assessment of renal fibrosis was shown by Friedli et al. (Friedli et al., 2016) where CMD of T1s have positively correlated with eGFR and renal fibrosis staging, while cortical or medullary T1s failed to show a correlation which can be explained by the ability of CMD technique to decrease the inter-individual variability resulted from ROI selection from cortex and medulla separately. In their multiparametric study that examined renal perfusion, diffusion, oxygenation and microstructure in CKD patients using ASL, DWI, T2* and T1 mapping, Cox et al. have found the cortical T1s significantly increased contrary with the CMD of T1s which decreased in the presence of renal impairment (CDK stages III and IV) (Cox et al., 2017). Furthermore, T1 native mapping was investigated in renal
transplant studies to evaluate kidney function, where T1s has shown an increase in transplanted compared to native kidneys and in impaired compared to well-functioning kidneys and the CMD of T1s was found to be diminished after transplantation (Huang et al., 2011; Peperhove et al., 2018).

Studies on renal T2 measurement are still scarce. Cortical T2s were found to be significantly increased in transplanted compared to native kidneys, whereas no correlation was found between T2s and allograft function (eGFR) (Mathys et al., 2011). Moreover, T2 mapping was applied to autosomal dominant polycystic kidney disease (ADPKD), characterized by a progressive development of cysts and fibrosis composition, where whole kidney T2s were found to distinguish between healthy and ADPKD kidneys in animal model (Franke et al., 2017). Recently, T1 and T2 were reported to have the ability to assess cystic kidney disease progression in kidneys with ARPKD in a mouse model, where T1s and T2s were significantly higher in affected compared to healthy kidneys (MacAskill et al., 2020). Interestingly, T2s were found to be strongly correlated to cystic fraction in patients with ADPKD (Siedek et al., 2020).

In summary, T1 maps has shown the ability to assess renal damage either in renal impairment reflected by low eGFR, HF associated with renal dysfunction or renal transplant. Consistent findings are T1s were correlated to GFR and fibrosis staging, increased when the presence of renal impairment. The CMD of T1s was reported to be decreased with renal dysfunction and was found to be more efficient than cortical or medullary T1s alone, which gives the opportunity to further study the distribution of the T1s within the complete renal parenchyma. T1 and T2 maps were lightly introduced to clinical studies of polycystic kidney diseases (ADPKD and ARPKD), where T1s and T2s were found to be increased in affected kidneys and T2s were reported to be well correlated to renal cystic fraction. Additionally, T2 maps were not widely explored. Thus, additional studies evaluating their potential in assessment of renal dysfunction are required. Finally, further validation of the correlation between relaxation mapping and renal function reflected as GFR as well as the sensitivity of mapping to detect fibrosis remain in need.

6. Magnetization Transfer Magnetic Resonance Imaging

Magnetization Transfer magnetic resonance imaging (MT) has the potential to assess biological tissue fibrosis. MT is sensitive to immobile macromolecule components within the tissue and can evaluate pathological events accompanied with macromolecule changes (e.g. fibrosis composition). It is sensitive to the interaction between free and immobile protons of large molecules. Using an off-resonance RF pulse, macromolecule protons are saturated and transferred to water protons depending on the exchange rate between the two protons population. Two sets of images need to be acquired, at baseline without the MT pulse and a MT-weighted image. Hence, the percent decrease in water signal following the exchange reflects the MT ratio (MTR) which is an indicative of macromolecule content (Henkelman et al., 2001).

MT has shown its utility in the detection of fibrosis in animal models and humans with renal diseases. In a mouse model of RAS, Ebrahimi et al. have shown the ability of the MT to visually differentiate fibrotic from non-fibrotic tissue where the MT-derived parameters (the magnetization fractional pool and the free magnetization exchange rate) were significantly different between stenotic and the contra-lateral kidney and correlated with fibrosis quantified from trichrome tissue staining, thus their potential use as biomarkers of kidney morphological changes after fibrosis composition (Ebrahimi et al., 2013). Similarly, MT has succeeded to longitudinally monitor renal fibrosis in a mouse model of RAS, where median MTR in cortex and medulla was significantly decreased in RAS kidneys and strongly correlated with ex-fibrosis assessed by histology in addition to the good spatial concordance between MTR maps and renal fibrosis by trichome and Sirius red staining (jiang et al., 2017a), (jiang et al., 2017b). MT was tested on a murine model of ADPKD, as it is accompanied with cysts (decrease in MTR) and fibrosis burden (increase in MTR). Parameters derived from MTR maps (including mean, median, 25th percentile, skewness and kurtosis, ) were closely related to indices of renal pathology and MTR was correlated well with cystic and fibrotic histological index with a good concordance between histological stained slice and MTR maps as shown in Fig. 5 (Kline et al., 2016). Recently, MTR has shown to provide structural and metabolic assessment of renal fibrosis in rats with unilateral ureteral obstruction (UUO), where MTR was significantly different between healthy and the contralateral kidneys with UUO over the obstruction course and well correlated to metabolic markers (A. Li et al., 2020).
Owing that the MTR is affected by many factors related to sequence details and relaxation parameters, quantitative MT (qMT) technique have been developed to provide a more quantitative assessment of macromolecules content within the tissue with a higher sensitivity and specificity (F. Wang et al., 2018). In this technique, the ratio of the macromolecular proton pool to the free water pool (pool ratio size, PSR) is isolated from relaxation and exchange rates and used as a quantitative MT parameter. Using qMT, fibrosis was assessed in murine progressive diabetic nephropathy and tubulointerstitial fibrosis and PSR has been shown to be a useful index of fibrosis (F. Wang et al., 2018), (Wang et al., 2019).

Moreover, one human study has figured out the role of MT in assessment of renal fibrosis. The MTR of cortex was correlated with eGFR and significantly different between subjects with normal renal function compared to patients with renal impairment, classified according to their eGFR (Ito et al., 2013).

Both, MTR and PSR has been shown to be a promise fibrosis biomarker. The ability of MT in the detection of fibrosis, a common pathway of CKD, has been demonstrated in animal models, while its clinical application in human studies and especially in identifying CKD still limited.

### 7. Magnetic Resonance Elastography

Organ stiffness has been demonstrated to precede the implications of fibrosis, the extra cellular matrix deposition (Georges et al., 2007). Fibrosis decreases tissue elasticity; hence the tissue elasticity measurement could provide a good marker of fibrosis (Hewadikaram et al., 2018). Magnetic Resonance Elastography (MRE) can measure this tissue mechanical property by applying a mechanical vibration over the target organ and capturing the shear waves generated and propagated by motion-synchronized or phase contrast imaging (Muthupillai et al., 1995). More rapid waves with longer wavelength are captured from stiffer or fibrotic compared to soft healthy tissues. Stiffness maps or elastograms are generated by processing these waves (Fig. 6).
MRE was reported as a good liver fibrosis detector in animal and human studies where an increased stiffness was captured with the increased severity of fibrosis and a good fibrosis stages prediction power was found. MRE has even surpassed ADC maps in the fibrosis staging with a better sensitivity and specificity (Kim et al., 2013; Rouvière et al., 2006; Talwalkar et al., 2008; Y. Wang et al., 2011; Yin et al., 2007b, 2007a).

Similarly, alteration in renal mechanical properties, elasticity, and stiffness, have been detected using MRE and were found to be correlated with kidney fibrosis. Shah et al. have shown the ability of MRE to measure cortical renal stiffness induced by nephrocalcinosis with mild fibrosis in mouse model (Shah et al., 2004). Medullary stiffness has been found to be significantly increased, and strongly correlated with the degree of fibrosis histologically evaluated in a swine model of RAS (Korsmo et al., 2013; X. Zhang et al., 2018). Lee et al. have reported a modest increase of stiffness in fibrotic kidneys (Lee et al., 2012). Moreover, renal stiffness was correlated with eGFR and significantly higher in functioning compared to nonfunctioning transplanted kidneys (Garcia et al., 2016). Whole-kidney stiffness was found associated with fibrosis scores and eGFR showing an increase with the eGFR decline in renal allograft patients even during a follow up study (Kim et al., 2017; Kirpalani et al., 2017). Similarly, a negative association was found between eGFR and MRE stiffness (Zhang and Zhang, 2020). Recently, Hodneland et al. have reported a good association between shear waves derived parameters (reflecting the pressure gradients and volumetric and shear deformations) and arteriosclerosis grade determined by biopsy (Hodneland et al., 2019). Surprisingly, stiffness has shown a decrease in all stages CKD DN patients (stages I to V) compared to controls, which can be explained by the fact that MRE is influenced by hemodynamic (blood flow) beside fibrosis deposition (Brown et al., 2019).

MRE has been reported as a good tool for fibrosis detection and staging allowing its potential as renal function predictor. Main results have shown that increased stiffness was associated with increased fibrosis stage. Clinical applications on renal transplants have reported a correlation of MRE stiffness with eGFR and an increase of stiffness with the presence of renal dysfunction. However, renal stiffness can be influenced by other factors other than fibrosis like hydronephrosis, edema formation, renal blood flow, collecting system dilatation, changes in the paramagnetic material composition (e.g. lipid, protein) and structural factors such (anisotropic structure) that can mask fibrosis during MRE interpretation (Leung et al., 2017), (Lee et al., 2012; Warner et al., 2011), (Gennisson et al., 2012). Those results require further studies to test its ability to diagnose CKD patients.

8. Other MRI Techniques

Susceptibility-Weighted Imaging (SWI) has shown good promises in renal function assessment and fibrosis detection. Mie et al. have investigated the feasibility of SWI on human kidneys (Mie et al., 2010). SWI signal was found to be affected by perfusion alterations and tissue fibrosis in animal models (Pan et al., 2017; J. G. Zhang et al., 2018). Pan et al. have demonstrated the sensitivity of SWI after renal reperfusion where the SWI score has decreased and returned to baseline over 48h after reperfusion injury (Pan et al., 2017). Additionally, Zhang et al. have shown a decrease of SWI signal ratio and a strong correlation with fibrosis staging (J. G. Zhang et al., 2018).

Moreover, fat quantification by using Dixon technique also known as fat fraction imaging has been investigated to early detect renal lipid deposition in patients with diabetic nephropathy, a decline in fat fraction was noted with the
The use of Dixon technique in CKD assessment warrants studies to test its ability to detect renal impairment.

Furthermore, MR Spectroscopic Imaging (MRSI) that describes metabolites characterization within the tissue has been investigated in assessing renal diseases. Given that renal failure is associated with a progressive formation of inorganic phosphorus and a loss of adenosine triphosphate, $^{31}$P MR spectroscopy has been emerged by using the phosphomonoesters to inorganic phosphorus ratio as a renal metabolic marker (Ebrahimi et al., 2014b). Metabolomic profile has succeed to reflect different renal allograft function as it was well associated with eGFR (Bassi et al., 2017). CKD metabolic biomarkers are markers of glomerular filtration, tubular function and mitochondrial function, alterations in urea cycle or amino acid that are lost with the increase of impairment severity (Hocher and Adamski, 2017). MRSI involves a powerful tool to assess renal metabolic markers giving the opportunity to early diagnosis and intervention in CKD (Ye and Mao, 2016).

### Other Imaging Modalities

#### 1. Ultrasound Elastography

Ultrasound Elastography (UE) enables the detection of tissue mechanical properties, like MRE. UE techniques are based either on the imaging of the generated shear wave propagating within the tissue or on the strain analysis of the tissue under external compression (Gennisson et al., 2013). Shear wave elastography (SWE or USE) utilizes sound wave to assess tissue elasticity by measuring the Shear Wave Velocity (SWV) reflecting tissue stiffness. Strain Elastography (SE) applies an external compression to the kidney and measures the deformation occurred due to this stress known as cortical or medullary strain and the cortico-medullary strain ratio. Compared to conventional Doppler ultrasound (US) derived parameters, UE-derived parameters have shown a superior potential in the assessment of renal fibrosis (Hu et al., 2015; Leong et al., 2019; Marticorena Garcia et al., 2018).

SE was used to evaluate kidney function after transplantation, as kidneys are close enough to the body surface. Cortico-medullary strain ratio has presented a decline with the increase of cortical fibrosis degree and has shown a significant difference between groups of different fibrosis scores and inversely correlated to the fibrosis stage (Gao et al., 2015, 2013a). When it was normalized, the cortical strain was found to have a good discrimination power between mild and moderate fibrosis stages (Gao et al., 2013b). Moreover, the tissue mean elasticity has been shown an inverse correlation with fibrosis degree (Orlachio et al., 2014). SWE was used by Garcia et al. to detect renal dysfunction in allograft patients and have reported a higher stiffness in functional compared to non-functional kidneys as well as a good prediction performance (sensitivity 90.9%, specificity 85.7%, AUC 0.925 for pyramids stiffness), SWV was associated well with renal blood flow and eGFR (Marticorena Garcia et al., 2018). Furthermore, several studies have reported a good correlation between SWE stiffness and the fibrotic stage (Arndt et al., 2010; Ma et al., 2018; Nakao et al., 2015). In contrast, some studies have reported that SWV or estimated stiffness was not correlated renal dysfunction and did not differ in grafts with different stage of fibrosis (Grenier et al., 2012; Lee et al., 2015; Syversvseen et al., 2012, 2011).

Applied to CKD patients, SWV was found to be well associated with eGFR, significantly higher in healthy compared to each stage of CKD patients (AUC=0.752) but it was not able to differentiate CKD stages (CKD stages based on eGFR) (Guo et al., 2013), (L. Wang et al., 2014). A decline of SWV has been reported with the severity of renal impairment or fibrosis score assessed by histology (Hu et al., 2015). Compared to conventional ultrasound parameters, including renal length, parenchymal thickness, and resistive index, SWV has shown a better differentiation performance (Hu et al., 2015). Similarly, Bob et al. have reported a decrease in SWV associated with the decrease of renal function qualified by eGFR in patients with diabetic kidney disease (DKD) without other renal disease or DM (Bob et al., 2017). SE strain ratio was found to be well correlated with eGFR in patients with DN (Iacob et al., 2019). On the other hand, no correlation was found between renal stiffness and fibrosis scores assessed with biopsy (Cardenas et al., 2019), (L. Wang et al., 2014) and estimated stiffness was found higher in CKD patients (stages III to V) (Lin et al., 2017; Samir et al., 2015). Recently, SWE was found to be able to detect abnormal renal stiffness in patients with early stages of glomerulonephritis and preserved renal function (Grossmann et al., 2019). Young’s Modulus (YM) was measured to assess renal stiffness and has shown the largest discrimination power between healthy volunteers, type 2 DM
without DKD and type 2 DM with DKD patients (Shi et al., 2020). YM was negatively correlated to eGFR and higher in healthy controls compared to patients. SWE has outperformed conventional US parameters and has shown its ability to monitor type 2 DM (Shi et al., 2020).

UE has not shown an ability to objectify stages of renal damage. In terms of clinical applications, SE was employed in renal transplants studies, showing a good correlation between strain ratio or tissue mean elasticity and the fibrosis stage. A higher stiffness was reported in functional compared to unfunctional allografts. SWE was incorporated in CKD, DKD, glomerulonephritis, DM and DN studies leading to conflicting results about the association between estimated stiffness and renal function. Moreover, the shear wave was supposed to be faster propagated in fibrotic compared to healthy tissue which cannot be noticed in mentioned studies. The association between SWV and renal blood flow, reported previously (Marticorena Garcia et al., 2018), supports the fact that many factors can affect tissue stiffness and mask fibrosis which has induced these conflicting results about the potential of UE in fibrosis assessment (Leung et al., 2017), (Lee et al., 2012; Warner et al., 2011), (Gennisson et al., 2012). Interestingly, spectral parameters derived from ultrasound waves are supposed to reflect the frequency content and then the disease severity of CKD (Hewadikaram et al., 2018).

2. Computed Tomography

Computed Tomography (CT) is an imaging technique that uses a motorized x-ray source shooting narrow beams of x-rays and producing tomographic images. CT offers a good temporal and spatial resolution as well as a quantitative power for contrast agents (Zhu et al., 2018).

Split renal length and volume derived from volumetric CT have shown the ability to assess renal function in kidney donors by allowing the measurement of SK-GFR (Gaillard et al., 2017; Jiang et al., 2019a; Patankar et al., 2014; Yanishi et al., 2015). Split parenchymal volume was found strongly correlated to split renal GFR measured by scintigraphy and a better indicator of reduced split renal function when combined with split renal ADC (Li et al., 2018; Mitsui et al., 2018). Additionally, cortical volume was demonstrated to be a powerful tool of renal function estimation and renal outcomes prognosis in kidney donor as well as a good predictor of CKD development after nephrectomy (Gardan et al., 2018; You et al., 2018). Most importantly, CT has shown a good promise in decline renal function assessment as well as fibrosis detection. Owing that CKD is normally associated with renal microvascular rarefaction, Stillfried et al. have shown the potential of CT angiography derived parameters in the assessment of renal function in CKD, where the renal relative Blood Flow (rBF), closely mirrored the renal rarefaction resulted from biopsy, as well as the arterial diameter were found to be significantly reduced in CKD patients (eGFR <=32 mL/min/1.73 m²) (von Stillfried et al., 2016). Furthermore, a new gold nanoparticles conjugated to an anti-collagen-I antibody was involved by Zhu et al. as a CT contrast agent and was able to visualize mouse kidney fibrosis either by multi detector CT (MDCT) or micro-CT matching the renal histology (Zhu et al., 2018).

Although it ability to measure SK-GFR as well as the potential of CT angiography in renal function assessment in CKD patients, CT is still limited to ionizing radiation and contrast agent injection that can injure the kidney (Lerman et al., 1996; Maioli et al., 2012). Recently, Contrast-Enhanced CT was reported to increase the risk of End-Stage Renal Disease (ESRD) (Lim et al., 2020). However, identification of renal fibrosis or perfusion with non-contrast CT remains challenging (Zhu et al., 2018).

3. Scintigraphy Imaging

Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT) also known as scintigraphy imaging are the most common imaging modalities used by nuclear medicine that uses radioactive tracers to assess organ function and perfusion. PET relies on the detection of radiation emitted through the electron-positron annihilation when positron emitting nuclides are applied while SPECT measures gamma rays emitted from tracers (Kohnke et al., 2019).

Nuclear imaging allows split renal function measurement (SK-GFR) (Patankar et al., 2014; SHIMIZU et al., 2016; Yanishi et al., 2015). Moreover, PET and SPECT were proved to be able to detect hepatic, cardiac and pulmonary fibrosis (Désogère et al., 2017; Kim et al., 2016; Li et al., 2011). Additionally, scintigraphy has succeeded to early
predict renal reduced function after transplantation (Yazici et al., 2015, 2013; Yoon et al., 2016), with an improvement of prediction power when coupled with CT (Lovinfosse et al., 2016). Recently, nuclear medicine has been demonstrated to be a noninvasive tool in the quantitative assessment of glomerular function (Qin et al., 2019). PET has shown a potential predictive power for renal recovery in renal transplants undergoing AKI (Pajenda et al., 2020). Results presented support the need for more investigation of nuclear imaging in renal fibrosis and perfusion assessment in CKD, as clinical applications were limited to SK-GFR measurement, fibrosis detection and renal transplants function prediction.

Artificial Intelligence Applied to Medical Imaging

In addition to the qualitative analysis provided by the mean value of gray level pixels within a region of interest, texture analysis derived from mathematical techniques has the potential to give quantitative information that are usually imperceptible to the radiologist eye. Texture analysis techniques characterize the heterogeneity of the distribution of pixels and their spatial arrangement by describing the pixels inter-relationships and the gray level frequencies within an image that cannot sense visually. Several approaches have been involved in order to quantify image texture, including histogram analysis, 2-D Fourier transform, statistical methods (features derived from the gray level co-occurrence matrix (GLCM), the gray level run-length matrix (GLRLM) and the local binary pattern (LBP)), model-based methods (autoregressive and fractal model), and transform-based methods (wavelet transform) (Larroza et al., 2016). The process starts with image acquisition, and involves several steps which usually include ROI definition and preprocessing, feature extraction, feature selection based on parameter statistical significance for discrimination, and classification using simple statistical models or machine learning techniques (Larroza et al., 2016).

Machine learning stands for computer algorithms that learn from observations in order to predict future outcomes. There is a large variety of classifiers used in the medical imaging domain going from traditional statistical methods to more sophisticated algorithms. Classification performance of any model uses many measures such as confusion matrix measures (e.g. sensitivity, specificity, precision and accuracy or Area under the receiver operating characteristic curve (AUC). In traditional machine learning, training and testing the data are supposed to share the same distribution and the real-world input data feature space. The success of prediction depends highly on whether it is the case or not (Weiss et al., 2016). Conventional classifiers could be based on instance (K nearest neighbors), statistical learning theory (support vector machine), decision tree (random forest), feature combination (linear discriminant analysis, linear or polynomial regression) or derived from probability and statistics (Bayes) (Ohata et al., 2019), while deep learning techniques rely on the biological neuron architecture and are characterized by multiple artificial layers composed of nodes related to each other by weights representing the contribution of each node to the output.


Texture, extracted from medical images, reflects the micro- and macro-structure of the selected organ (Materka, 2004). Textural analysis was widely applied to different medical imaging modalities and has shown its ability to be a powerful computer-aided diagnostic tool that helps in clinical decision making (e.g., liver lesions differentiation, breast tumors classification, prediction of non-response to neo-adjvant chemotherapy in breast cancer and renal tumor classification between simple cysts, kidney stones and the complex renal cell carcinoma on CT and MR images) as well as segmentation tasks (e.g., kidney segmentation on CT and Ultrasonography (US) images) (Timothy L Kline et al., 2017; Lubner et al., 2016; Mayerhoefer et al., 2010; Michoux et al., 2015; Raman et al., 2015; Sreelatha and and Ezhilarasi, 2018; Yu et al., 2017).

Quantification of tissue fibrosis has recently emerged based on textural analysis. Textural parameters extracted from CT and MR images have shown a good correlation between liver fibrosis and increased heterogeneity (Daginawala et al., 2016; Yu et al., 2015; Zhang et al., 2015). Owing that fibrosis plays an important factor in the progression of CKD and given the heterogeneity of renal parenchyma (cortex and medulla), texture analysis on renal medical images may be a good predictor of renal dysfunction.
Regarding renal function evaluation, texture analysis was applied on MR, US as well as scintigraphy imaging, as summarized in Table 2. Integrated with machine learning techniques, textures were proven to be a good complementary tool that could serve doctors especially when discovering the CKD progression at early stages of the disease. Studies employing textural analysis were focusing on various clinical applications, including the detection of renal dysfunction at early stage by discriminating between healthy volunteers and patients with mild or non-severe renal impairment, the differentiation between healthy from diseased kidneys (CKD, LN, renal disorder), the evaluation of the association between textures and fibrosis composition as well as eGFR, the discrimination of rejected from non-rejected renal allograft, the prediction of the progression to CKD in ADPKD patients and most importantly, the prediction of CKD five stages.

Applied to different functional MR images, textures have succeeded in evaluating renal dysfunction in recent animal (Zha et al., 2019) and human studies (Rossi et al., 2012), (Alnazer et al., 2019; Ding et al., 2019; Timothy L Kline et al., 2017; Kociolek and Strzelecki Michal and Klepaczko, 2019; Shi et al., 2018). Histogram and GLCM based parameters were extracted from renal parenchyma on three MRI sequences including DWI, BOLD and SWI (Ding et al., 2019). Derived features were well correlated with eGFR. Textures of BOLD and SWI were able to discriminate control and non-severe renal dysfunction groups demonstrating the ability of textures to detect renal failure at early stages when the disease cannot be detected by eGFR (Ding et al., 2019). In line with these recently published results, our preliminary study which have applied textural analysis on DWI MR images, has confirmed that textures were affected by CKD (Alnazer et al., 2019). Despite the modest sample size, wavelet-based parameters in addition to GLCM-based parameters extracted from renal parenchyma were found to be significantly different between the two groups (Controls and CKD patients). In patients with LN a renal scar having CKD as an outcome, GLCM-based parameters succeeded detecting changes in the texture characteristics of BOLD MRI with a good prediction rate of renal pathology patterns (Shi et al., 2018). Authors have shown that renal histopathology changes resulted from LN impairment, led to decreased cortical R2* values as well as affected the normal oxygenation distribution within the kidney (Shi et al., 2018). In an animal model of UUO, histogram features extracted from R2* maps were able to distinguish the degree of induced fibrosis. In a short-term of evaluation of fibrosis, cortical textures declined significantly and correlated closely with the percentage of fibrosis (Zha et al., 2019). Additionally, textural analysis performed on T2 weighted MRI (T2W) in order to predict renal dysfunction in a retrospective study of ADPKD patients (Timothy L Kline et al., 2017). Incorporating stable texture features, that were strongly correlated with changes in eGFR, into a traditional modeling (age, eGFR, total kidney volume (TKV)) has improved the prediction power for progression to CKD. Thus, textural analysis provided additional insights to the existing TKV biomarker of renal function decline prognosis in ADPKD (Timothy L Kline et al., 2017). Kociolek et al. have shown that texture of DCE-MR images extends their possibilities by adding new information about renal functionality (Kociolek and Strzelecki Michal and Klepaczko, 2019). Histogram analysis of perfusion maps (ASL) has confirmed the importance of regional assessment of renal perfusion (Rossi et al., 2012). Authors have demonstrated that CKD is not only associated with changes of perfusion mean values, but also with changes in the distribution of perfusion values within the cortex and the renal parenchyma (Rossi et al., 2012).

Moreover, assessment of fibrosis deposition and renal impairment associated with CKD by textural analysis was performed on ultrasonography (US) renal images (Ardakani et al., 2017; Chen et al., 2019; Iqbal et al., 2017; Sharma and Virmani, 2017). Textures based on Fourier transform that reflect spatial frequencies, succeeded differentiating CKD and healthy kidneys while GLCM-based parameters failed (Iqbal et al., 2017). The use of combination of GLCM features vectors was able to distinguish US images of normal and diseased kidneys (Sharma and Virmani, 2017). A comprehensive approach to analyze and classify CKD stages based on the texture of US images was provided by Chen et al. in (Chen et al., 2019), that has reached a good classification performance. Concerning renal function change after transplantation, textures of US images have shown a good association with sCr as well as a significant difference between subjects (rejected and non-rejected allografts) with a good classification performance (Ardakani et al., 2017).

Textural analysis was applied on renal scintigraphy images as well. Ohata et al. have extracted several textures and used different machine learning techniques in order to hit the best accuracy in CKD stages classification (stage 1, 2 versus stage 3-5) (Ohata et al., 2019). Ardakani et al. have used textures to detect renal status after transplantation and found that textures improved clinical diagnosis by providing a good classification performance (Ardakani et al., 2018).
<table>
<thead>
<tr>
<th>Study</th>
<th>Clinical Application</th>
<th>Sample Size (n)</th>
<th>Imaging Modality</th>
<th>Textural Parameters</th>
<th>Classification</th>
<th>Main Results</th>
<th>Clinical Evaluation</th>
<th>Quantitative Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Ding et al., 2019)</td>
<td>Discrimination between control patients (&gt;= 80 ml/min/1.73 m²), patients with non-severe renal impairment (&gt;= 30 and &lt;80 ml/min/1.73 m²) and severe impairment (&lt; 30 ml/min/1.73 m²)</td>
<td>eGFR≥80 (control, n=17), 30-80 (non-sRI, n=24), &lt;30 (sRI, n=31)</td>
<td>MRI – BOLD, DWI, SWI</td>
<td>Histogram, GLCM (d=7, O=0°)</td>
<td>ANOVA</td>
<td>Textures were correlated well with eGFR, DWI texture was able to detect renal dysfunction, BOLD and SWI texture detected renal failure at early stage (controls Vs. non-sRI)</td>
<td>AUC=0.84 (non-sRI Vs sRI) AUC=0.75 (CG Vs non-sRI)</td>
<td></td>
</tr>
<tr>
<td>(Alnazier et al., 2019)</td>
<td>Differentiation between normal and chronic kidney disease (CKD) kidneys</td>
<td>Healthy (n=2), CKD (n=2)</td>
<td>MRI – DWI</td>
<td>GLCM (d=1-8, all directions), wavelet</td>
<td>Absent</td>
<td>Textures have shown a significant difference between CKD and healthy kidneys</td>
<td>p-value &lt;0.05 between groups</td>
<td></td>
</tr>
<tr>
<td>(Shi et al., 2018)</td>
<td>Distinction of patients with lupus nephritis (LN) staging from CKD-I to CKD-III from healthy volunteers</td>
<td>Healthy (n=11), LN (n=12) divided into 4 subgroups of pathology, having CKD stage 1-3</td>
<td>MRI – BOLD</td>
<td>GLCM (d=1, all directions)</td>
<td>Fisher linear discrimination</td>
<td>Good disease prediction power of GLCM parameters</td>
<td>Correct prediction rate =77.8%</td>
<td></td>
</tr>
<tr>
<td>(Zhang et al., 2019)</td>
<td>Evaluating the change of textures values over unilateral ureteral obstruction (UUO) associated with fibrosis composition</td>
<td>Rabbit study-control (n=6), UUO (n=30)</td>
<td>MRI – BOLD</td>
<td>Histogram</td>
<td>Absent</td>
<td>Histogram parameters of cortical R2* decreased and showed significant differences over the course of UUO</td>
<td>p-value &lt;0.05 between groups</td>
<td></td>
</tr>
<tr>
<td>(Timothy L Kline et al., 2017)</td>
<td>Prediction of the progression to CKD and the reduction of eGFR for patient with autosomal dominant polycystic kidney disease (ADPKD)</td>
<td>3A CKD (n=44), 3B CKD (n=22), 30% reduction of eGFR (n=47)</td>
<td>MRI – T2W</td>
<td>GLCM</td>
<td>Multiple linear regression</td>
<td>Textural parameters have improved the prediction power of progression to CKD and were well correlated to eGFR changes</td>
<td>AUC=0.93, 0.86 and 0.82, progression to 3A, 3B and 30% reduction in eGFR respectively</td>
<td></td>
</tr>
<tr>
<td>(Rossi et al., 2012)</td>
<td>Discriminating patient with mild renal dysfunction staging from CKD-I to CKD-III and healthy volunteers</td>
<td>Healthy (n=8), patients with mild renal impairment – eGFR&gt;30 (n=9)</td>
<td>MRI – ASL</td>
<td>Histogram</td>
<td>Absent</td>
<td>Histogram-based parameters of the cortex were significantly affected by renal dysfunction</td>
<td>p-value &lt;0.05 between groups</td>
<td></td>
</tr>
<tr>
<td>(Iqbal et al., 2017)</td>
<td>Discriminating CKD patients from healthy volunteers</td>
<td>Healthy (n=24), CKD (n=8)</td>
<td>US</td>
<td>Fourier-based parameter, GLCM</td>
<td>Absent</td>
<td>Fourier-based parameters succeeded to distinguish normal and CKD kidneys</td>
<td>p-value &lt;0.05 between groups</td>
<td></td>
</tr>
<tr>
<td>(Sharma and Virmani, 2017)</td>
<td>Discriminating healthy kidneys from kidneys with renal disorders (polyuria, pyuria, proteinuria and hematuria) and no hydronephrosis</td>
<td>Normal kidneys (n=11), diseased (n=8)</td>
<td>US</td>
<td>GLCM ratio feature vector, GLCM additive feature vector, GLCM concatenated feature vector (d=1 to 10)</td>
<td>SVM</td>
<td>Good classification performance accuracy=85.7% between groups indicated for d=1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2 Application of texture analysis and conventional machine learning techniques on medical images to evaluate renal. Linear discriminant analysis (LDA), support vector machines (SVM), multi-layer perceptron (MLP), k-nearest neighbors (KNN), random forest (RF).

Applying texture analysis to clinical practice as well as research studies faces many challenges like the great influence of signal intensity and heterogeneity quantification which requires standardization of image acquisition protocols as well as prior image normalization and intensity correction when required (Timothy L Kline et al., 2017). It is worthy to note that studies evaluating the potential of textural analysis in the assessment of CKD renal impairment (besides their association with histology and pathology) are still limited and yet to be explored. Thus, further validations are recommended to evaluate its potential as a new dimension in CKD management. Proposed approaches, presented in Table 2, do not yet provide complete and robust methods to be used to aid clinical practice. Such approaches have to be evaluated in large and multi-center studies that employed standardized imaging protocols for different modalities, applied a standardized image intensity normalization schema, respected the patient’s preparation prior to image acquisition (e.g. fasting, checking the hydration state) and selected the most reproducible and significant textures that show a strong association with renal tissue histology and pathology more than eGFR. Once such protocols are evaluated and proven to give an excellent CKD detection and staging performance where no misclassification, they will be able to be incorporated to clinical practice in terms of diagnosis and prognosis.

2. Deep Learning

2.1. Principle

Deep learning is a part of machine learning inspired by the way the brain fulfills a learning task, trying to imitate biological neural networks (Goodfellow et al., 2016). It employs multilayer artificial neural networks through mathematically interconnected nodes (Goodfellow et al., 2016). Weights connecting these nodes are adjusted based on an optimization equation in the course of training process until the neural network is well learned. Deep learning has changed the traditional pipeline of feature extraction followed by the use of conventional machine learning algorithms to a simple input-output procedure with a complex architecture of deep layers that enable internal deep
features extraction (Kavur et al., 2020). Recent deep Convolutional Neural Networks (ConvNets, CNNs) are structured into 10 to 20 layers of linear units, hundreds of millions of weights, and billions of connections between units (Lecun et al., 2015). CNNs architecture support the process of multiple array data including 1D signals, 2D images, audio and 3D videos (Lecun et al., 2015). The multilayer composition of CNNs enable learning hierarchy of features without relying on handcraft features (Sharma et al., 2017a). In terms of image classification, CNNs take the image as input and transform raw pixels to a class score by passing them throughout convolutional filters (Sharma et al., 2017a).

Deep learning has been used successfully in organ segmentation (Roth et al., 2015; Zheng et al., 2017), total kidney volume determination (Timothy L. Kline et al., 2017; Sharma et al., 2017b), chronic myocardial delineation (Zhang et al., 2019), cerebral micro-bleed detection (Y.-D. Zhang et al., 2018) and pulmonary nodules classification (Ciompi et al., 2015).

Deep networks were initially fed by medical clinical data including attributes such as age, blood pressure, blood glucose, serum creatinine etc., and have outperformed all conventional machine learning techniques (Kriplani et al., 2019; Saha et al., 2019; Shankar et al., 2018). Moreover, CNNs were found to be a powerful tool in glomeruli localization on biopsy slices (Bukowy et al., 2018; Kannan et al., 2019; Marsh et al., 2018) and has outperformed classifiers based on the pathologist-estimated fibrosis score in the prediction of fibrosis staging (Kolachalama et al., 2018).

<table>
<thead>
<tr>
<th>Study</th>
<th>Clinical Application</th>
<th>Sample (n)</th>
<th>Size</th>
<th>Augmented Data</th>
<th>Imaging Modality</th>
<th>Neural Network Architecture</th>
<th>Main Results</th>
<th>Data Accessibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Kuo et al., 2019)</td>
<td>Estimation of artificial intelligence-based eGFR (AI-eGFR) and prediction of kidney function (&lt; or &gt; 60 ml/min/1.73 m²)</td>
<td>1446 images of 1299 patients labeled with eGFR</td>
<td>4505 from 1446 images 1285 for training 161 for testing</td>
<td>US</td>
<td>Transfer learning using pre-trained CNN (ResNet) - Integration of regression layers - Gradient boosting-model classifier</td>
<td>Good correlation between sCr based e-GFR and AI-GFR (r=0.741) and a better CKD detection accuracy compared to nephrologist (85.6% Vs 60.3-80.1%)</td>
<td>Available on request from the corresponding author, CCK</td>
<td></td>
</tr>
<tr>
<td>(Zheng et al., 2019)</td>
<td>Differentiate normal kidneys from kidneys with congenital abnormalities of the kidney and urinary tract (CAUT)</td>
<td>Controls (n=50), children with acute kidney injury (n=50)</td>
<td>No</td>
<td>US</td>
<td>Transfer learning based on a pre-trained CNN (AlexNet)</td>
<td>Combined, transfer learning and conventional images features can provide the best classification performance (accuracy&gt;80%)</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>(Hao et al., 2019)</td>
<td>Discrimination between normal patients and chronic kidney disease (CKD) patients staging from CKD-I to CKD-V</td>
<td>Normal (n=46), CKD stage I-V (n=180)</td>
<td>864 for training</td>
<td>US</td>
<td>Transfer learning using pre-trained CNN (ResNet-34) - Texture branch as a residual structure (GLCM, HOG) - Features fusion - Dense layers for classification</td>
<td>Fusion of deep features and textures features provided an accuracy of 95.13% and a sensitivity of 99.44% outperforming the classification based only on textural features or deep features</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>(Vasant Hselvakumar et al., 2020)</td>
<td>Automatic detection and classification of renal diseases including kidney stones, cysts and cancer tumors</td>
<td>90 images of non-diseased and diseased kidneys</td>
<td>No</td>
<td>US</td>
<td>AdaBoost Classifier for the disease detection - CNN for the disease classification: - 5 convolutional layers - 3 max pools - 2 average pooling - One flatten layer - 2 dense layers</td>
<td>The proposed CNN has outperformed conventional machine learning techniques with an accuracy of 86.76% and a precision of 84.3%</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

Table 3 Deep learning in renal function evaluation
In terms of renal function evaluation, CKD prediction and diagnosis, deep learning or more specifically transfer learning has shown a good promise either alone or served by a texture branch network, as summarized in Table 3. Recent studies were focusing on using deep networks in different clinical applications including distinction of healthy kidneys from kidneys with CKD, discriminating normal kidneys from kidneys with congenital abnormalities, classification of different renal diseases such as renal stones, cysts and tumors and most interestingly, predicting the eGFR.

2.2. Data Availability

When training data is limited or hard and expensive to collect, there is a need for neural networks that are able to overcome lack of data-induced issues. Data augmentation is a necessary step before training the model. It increases the diversity of training samples and prevents trained models from overfitting (Sharma et al., 2017a). To expand the limited data, several approaches were adopted. Here, we discuss the use of transfer learning and data augmentation either by applying image transformation on the existing ones or by using a high topic of deep learning, the deep Generative Adversarial Networks (GAN).

Transfer learning aims at ameliorating learning performance on a target domain by transferring information from a related source domain (Weiss et al., 2016). As an example, a person who has a good musical knowledge through playing guitar and want to learn to play the piano would learn in a more efficient way, compared to another person who had no musical background, by transferring his knowledge to the task of learning to play the piano (Weiss et al., 2016). Thus, transfer learning is possible.

Considering the huge data size required during the training phase of deep networks, data augmentation has been extensively used to enrich training dataset. Data augmentation is an artificial expanding of dataset samples by applying classical image transformation to the existing samples including random gray-level transformation of pixels (<3%), cropping, rotating, zooming and shifting (Hao et al., 2019; Pavinkurve et al., 2019), applying a low frequent intensity variation (Sharma et al., 2017a), or much more complex image transformation algorithms like the radial transform sampling (Hojjat Salehinejad et al., 2018; Salehinejad et al., 2018). Data can be augmented as well by applying deformable image registration (Yin et al., 2019).

Transfer learning was used by several studies that carry out the renal status evaluation through imaging. Researchers have re-trained CNNs, previously trained on photographic images from the ImageNet challenge, to classify ultrasound renal images (Cheng and Malhi, 2017; Hao et al., 2019; Kuo et al., 2019; Zheng et al., 2019). Data augmentation was applied as well in order to get the larger sample size required by deep networks. Cheng et al. demonstrated the effectiveness of transfer learning in classifying abdominal images with different diseases including end-stage renal disease, liver and bladder diseases (Cheng and Malhi, 2017). Neural networks have sometimes outperformed radiologists in terms of diseases discrimination performance (Cheng and Malhi, 2017). In their recent study, Kuo et al. suggest a deep neural network model in order to estimate artificial intelligence based GFR (AI-GFR) and to detect CKD from renal US images (Kuo et al., 2019). Authors have extended their data by augmentation and have adopted the transfer learning approach of a pre-trained CNN model (ResNet) that has exceeded nephrologists classification between CKD and non-CKD patients and has achieved good correlation between AI and sCr-based eGFR (Kuo et al., 2019).

GANs are a part of deep generative modeling that take input training samples from some distribution and learn a model that represents that distribution. GANs rely on a generator with multiple encoding and decoding layers and a discriminator. During GAN training, the generator tries to create imitations of data (synthetic images) while the discriminator tries to identify real data from fake data created by the generator. Training continues until the discriminator cannot find any difference between fake and real data. GANs have gained a great interest in the field of medical imaging and have been used in different applications including image synthesis (Lutnick et al., 2020; Sivanesan et al., 2019), image translation (Murali et al., 2020), image super resolution (Mahapatra and Bozorgtabar, 2019), and image transformation (Chandrashekar et al., 2019). Regarding renal images, GAN were used to generate real-like micro-anatomic renal images (Murali et al., 2020). Authors have used a cyclic GAN to create an artificial effect of staining without physically tampering the histopathological slide. Their proposed GAN was demonstrated to be able to translate different stain styles of renal pathology (e.g. transforming hematoxylin and eosin stain to periodic acid-Schiff stain). The same team has used the concept of GAN in order to generate realistic looking
synthetic images of renal biopsy (Lutnick et al., 2020). Most interestingly, GAN was employed in (Sivanesan et al., 2019), to create synthetic US images as seen in Fig. 7, therefore expanding their limited dataset.

![Image of real vs synthetic renal biopsy images](image)

**Fig. 7**

### 2.3. Discussion on Texture Descriptors

Interestingly, textures were integrated with typical CNNs. Novel models were developed, combining both deep networks and textural features as a residual structure. The models hold a CNN served from a textural branch in order to extract both deep and textural features from images and use the fused information for classification.

Zheng et al. have mixed transfer learning and conventional imaging features, including geometrical and histogram of oriented gradient (HOG) features (Zheng et al., 2019). Their proposed model has succeeded to accurately discriminate between normal and kidneys with urinary track (Zheng et al., 2019). Similarly, in the aim of improving decision making and classification performance, HAO et al. suggested adding textural features to the CNNs (ResNet), a way to provide additional descriptors (Hao et al., 2019). Authors have used data augmentation prior to transfer learning approach and residual textural branch in order to get a model of multi-level descriptors that mixed deep features with domain texture features (Hao et al., 2019). Their developed model has outperformed the use of textural features or CNNs features alone with a high accuracy and an excellent sensitivity, and shown its ability as a computer aided for CKD screening (Hao et al., 2019).

Artificial intelligence that applies machine learning and deep learning algorithms into clinical research was found to be a useful noninvasive assessment tool that succeeded to transform renal imaging into a real time screening tool and renal function estimator by the investigation of AI-GFR. The combination of transfer learning with conventional images features was reported to provide a robust classifier with a high performance. Further studies investigating the deep learning in other imaging modalities is required and additional validation to test its usefulness to not only detect CKD but to also predict its stages still warranted.

### 3. Artificial Intelligence in Renal Segmentation

#### 3.1. Automatic Renal Segmentation

Automatic renal segmentation is a crucial task since the manual delineation of renal tissue is time consuming and subject-dependent. Segmentation is a key step in the analysis of abdominal images with different applications including surgery planning, computer-aided monitoring, extraction of qualitative or quantitative features, image-guided interventions (Conze et al., 2020). The segmentation of the kidney could serve the radiologist prior to his qualitative assessment and it is needed before feature extracting from the input image in order to get to a fully automated software tool of kidney detection and status evaluation. Different segmentation approaches were adopted in clinical researches aiming to avoid manual renal delineation and observer intervention efforts.
Kidney identification is generally performed semi-automatically or automatically. Semi-automatic segmentation is usually used due to the similar gray level intensities and the high variation of organ shapes and positions in the abdominal images. It requires various intervention mechanisms, such as identification of initial seeds, localization of samples within the kidney tissue and the background, pre-segmentation of the kidney with a circular contour, defining a parameter range (Hammon et al., 2016; Hu et al., 2012; Mortensen and Barrett, 1998; Sandmair et al., 2016; Torres et al., 2018). However, these interventions are operator-dependent and additional analysis is required to test the segmentation repeatability and to reveal a significant inter and intra-observer agreement and reliability. A limited number of experts is usually presented to perform such analysis, which prevent the generalization over consistency and repeatability (Kavur et al., 2020). Furthermore, the user interaction requires time, which may be tedious in challenging tasks (Kavur et al., 2020). Consequently, many studies aim to convert the user intervention to automatic interactions by applying image processing and conventional machine learning algorithms that require extraction of handcraft features (Vasanthselvakumar et al., 2020), (Gao and Boliang, 2010), (Akbarian and Fei, 2013). Refer to (Torres et al., 2018) for more details about renal semi-automatic segmentation strategies applied on MRI, US and CT imaging.

3.2. Traditional Approaches for Automatic Renal Segmentation

In terms of developing a fully automated renal delineation framework, several traditional segmentation approaches were proposed during the last decade. Concerning kidney segmentation on MRI images, thresholding and shape detection, probabilistic shape model, Bayesian probability maps, unsupervised classification and deformable models were used. In (Will et al., 2014), authors have proposed the use of thresholding followed by shape detection techniques to segment renal cortex on T1 and T2 MR images efficiently and simply. However, this method is limited to the dependence on MR images quality and visibility. (Shehata et al., 2018) employed the use of a 3D probabilistic shape model on DWI images in order to evolve a 3D geometric deformable model reaching a high similarity coefficient with manual segmentation (DICE score or DSC). Despite their promising results, the developed model is limited to the lack of reliability and the high running-time. A DSC of 0.90/0.89 for right and left kidney parenchyma on DCE images was achieved by (Gloger et al., 2012). Authors have used refined probability maps and incorporation of external cortex edge alignment to develop a 3D segmentation framework for fully automatic renal parenchyma volumetry that exclude parenchymal cysts and could be used in clinical applications and epidemiological studies. However, the proposed approach was specifically designed for epidemiological studies, that require adaptation of data and domain knowledge in the training phase prior to extending the framework to perform kidney volumetry. Additionally, clustering methods were used on DCE images in order to identify multiple renal structures (e.g. cortex, medulla and pelvis) (Chevaillier et al., 2011; Li et al., 2012; Yang et al., 2016, 2015; Zöllner et al., 2011). Actually, differentiating renal structures could be served from the knowledge the temporal behavior oh the contrast agent and the information on the change pixels intensity over time. In their approaches, pixels were classified according to their intensity evolution by clusters, including K-means (Chevaillier et al., 2011; Yang et al., 2016, 2015; Zöllner et al., 2011), growing neural gas (Chevaillier et al., 2011), wavelet-based (Li et al., 2012; Zöllner et al., 2011) and Gaussian mixture clusters (Zöllner et al., 2011). However, clustering is based on the pixel behavior under the contrast agent, which is not necessarily used in other imaging sequences or modalities. Furthermore, active contour (AC) or deformable models that use energy constraints and forces in the image for separation of regions of interest were employed (Al-Shamasneh et al., 2020; L. Li et al., 2014). AC models are based on partial differential equations and variational models. The main idea behind AC in image segmentation is to start with an arbitrary boundary (closed curve), the curve is then updated iteratively by shrink and moved by image-driven forces to accurately detect the objects boundaries within the image (Hoang Ngan Le et al., 2020). Li et al. have succeeded to segment the complex renal contour on DCE MRI images by using a geometric active contours served from multi-scale edge detection algorithm that segment inhomogeneous regions (L. Li et al., 2014). Recently, a new active contour model was proposed by Al-Shamasneh et al. to segment kidneys on low-contrast MR images (Al-Shamasneh et al., 2020). Their model, which used a novel fractional function (Mittag–Leffler’s function) for energy minimization, has outperformed other approaches such as the Chan-Vese active contour model (C-V model), the model proposed by Ibrahim et al. that uses the wright fractional function and the segmentation by deep serial networks (Chan and Vese, 2001; Ibrahim et al., 2018). Their results have shown a high segmentation accuracy and DSC (98.95% and 0.93 respectively).
Moreover, renal segmentation on CT images was performed based on supervised and unsupervised conventional classification, multi-Atlas registration and C-V active contour models. In (Khalifa et al., 2017), authors proposed the extraction of appearance-based, voxel-appearance and higher order spatial features from CT scans followed by a Random Forest (RF) classifier schema resulting in a mean segmentation accuracy of 97.27%. Atlas-based method was proposed in (Yang et al., 2018) for renal segmentation. Their approach was based on multi-Atlas registration in a coarse-to-fine methodology. Song et al. have used a Fuzzy logic C-means clustering served from spatial information which was refined by an improved grow-cut technique and have achieved a high segmentation sensitivity and specificity (95.46 and 99.82% respectively) (Song et al., 2015). In (Huang et al., 2009), authors introduced a dynamic shape model and a connected component analysis as a multi-level set for multi objects segmentation into their C-V model resulting in a better segmentation performance compared to using a simple C-V active contour model.

Furthermore, shape detection techniques and deformable models were adopted in order to perform renal segmentation on US images. In (Marsousi et al., 2017), kidneys were detected by fitting a 3D shape kidney model on 3D US volumes. Then, the fitted model was used to evolve a level set function in order to delineate kidney boundaries. Authors approach have shown an accuracy of 97.48% and a DSC of 0.81 outperforming other methods (Ardon et al., 2015; Marsousi et al., 2014). This approach failed to detect the kidney volume in low-quality US images. A distance regularized level set deformable model followed by a prior shape for smooth boundary detection was applied by (Yang et al., 2012). Their model has shown a sensitivity of 95% and a specificity of 95%. In (Huang et al., 2013), a new active contour framework was proposed where a fast segmentation with an error of 0.028 was achieved by performing a convex relaxation of the energy function. While supervised classification was shown to be a successful segmentation method, it is still limited to their need for handcraft features extraction. Although deformable models can be considered as one of the most widely used approaches in medical image segmentation in last few decades, they impose several limitations. AC needs no learning properties from the training images. Therefore, it has a difficulty to deal with occlusions and noise. Moreover, it represents an unsupervised framework that lack a way to work with labeled images in a supervised approach. Thus, it gives unpredictable segmentation results. Lastly, deformable models are strongly depend on several parameters which are selected by experimental results (Hoang Ngan Le et al., 2020).

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size (n)</th>
<th>Augmented Data</th>
<th>Imaging Modality</th>
<th>Neural Network Architecture</th>
<th>Segmentation Evaluation</th>
</tr>
</thead>
</table>

Fig. 8
<table>
<thead>
<tr>
<th>(Conze et al., 2020)</th>
<th>40 normal patients</th>
<th>Yes (scaling, rotation, shifting), 100 augmented images for each slice</th>
<th>MRI – T1-DUAL MRI – T2-SPIR</th>
<th>CED with the following architecture were tested:</th>
<th>- No robustness when using CEDs alone</th>
<th>- Improvement of performance when using pre-trained CEDs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• UNet</td>
<td>- Cascaded CEDs</td>
<td>- Cascaded fashion of CEDs improved the performance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Pre-trained UNet</td>
<td>- Cascaded pre-trained CEDs</td>
<td>- Cascaded pre-trained cGAN gave the higher dice score for the right kidney (0.9, 0.93 on T1, T2) and a good score for the left kidney (0.93, 0.92 on T1, T2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Deep UNet (UNet extended with VGG)</td>
<td>- cGAN with CED as generator</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Pre-trained deep UNet</td>
<td>- Cascaded generators</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Yin et al., 2019)</th>
<th>Controls (n=50), CAXUT (n=50), 185 images</th>
<th>Yes (deformable image registration), no details about the final data size</th>
<th>US</th>
<th>Feature extraction by VGG-16, pre-trained on ImageNet</th>
<th>- Accuracy of 98% and a dice score of 0.94</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Kidney distance map generation by a boundary regression network (convolutional and deconvolutional layers)</td>
<td>- The proposed model outperformed the state of the art models: FCN and Deeplab</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Pixel wise segmentation based on the distance using a network that has the same architecture of the previous network with some modification</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Bevilacqua et al., 2018)</th>
<th>ADPKD (n=32), 57 images</th>
<th>Yes (horizontal flip and shift, scaling), no details about the final data size</th>
<th>MRI – T2</th>
<th>1st approach</th>
<th>- 2 architectures were tested</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• SegNet (a CED having VGG16 as encoder) (Badrinarayanan et al., 2017)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Fully convolutional network (Long et al., 2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ADAM optimizer</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Sharma et al., 2017a)</th>
<th>ADPKD (n=125)</th>
<th>Yes (image shift and application of intensity variation), 48000 from 16000 images</th>
<th>CT</th>
<th>CED network with VGG-16 as encoder</th>
<th>DICE score of 0.86</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AdaGrad as optimizer</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Xavier initialization to initialize weights</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Thong et al., 2016)</th>
<th>Anonymized data from 79 patients</th>
<th>No</th>
<th>Contrast-enhanced CT</th>
<th>A proposed CNN architecture learned patch-wise (2 convolutional layers followed by maxpooling layer and 2 fully connected layers)</th>
<th>DICE score over than 0.94 and 0.92 for left and right kidneys respectively</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CNN architecture modified after training (transforming fully connected layers into convolutional operations and fragmenting the maxpooling layers while keeping the same learned weights.</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 Deep learning in automatic renal segmentation

3.3. Deep Networks for Automatic Renal Segmentation

On the other hand, recent researches are exposing deep networks that offer the construction of systems able to turn the segmentation to a fully automatic procedure with a high accuracy and repeatability where neither interventions nor handcraft features are required. In addition to the success of CNNs in renal failure detection and patient
classification, the entire kidney segmentation presents an essential issue that was recently addressed, as shown in Table 4. Most of recent approaches were based on semantic segmentation that employs a fully convolutional network (FCN) or a Convolutional Encoder-Decoder network (CED). The encoding and decoding parts of a CED allows discriminating image pixels as it belongs or not to a specified region of interest (pixel wise segmentation). An example of a CED architecture is given in Fig. 8 (left panel).

Kidney segmentation using deep learning was tested on MR, US and CT images. An overall segmentation accuracy of 86% on MR images of kidneys with ADPKD was reached by Bevilacqua et al. in their study (Bevilacqua et al., 2018). Authors have tested different CNNs-based segmentation approaches on their augmented dataset. The full-image segmentation based on an FCN has exceeded the CED (SegNet with VGG-16 as encoder) segmentation and the introduction of a region-based CNN (R-CNN) prior to the semantic segmentation did not enhance the delineation performance. Moreover, a novel boundary distance deep network served by transfer learning was developed by Yin et al. in order to segment US kidney images of children with congenital abnormalities of the kidney and urinary tract (CAKUT) and controls, achieving an accuracy of 98.9% (Yin et al., 2019). A CNN (VGG-16), pre-trained on ImageNet, was employed to extract deep features that are used to learn boundary maps by a boundary distance regression network. A pixel classification network was then used to classify predicted maps as kidney pixels or not. Furthermore, Sharma et al. have employed a CNN that follows the VGG-16 architecture to generate score maps (see Fig. 8) based on pixel wise classification and achieved a DICE score of 0.86 between proposed and manual delineation of ADPKD kidneys on CT images (Sharma et al., 2017a). Similarly, Thong et al. proposed a fully automatic framework for kidney segmentation on contrast-enhanced CT scans based on CNN trained patch-wise. The proposed model achieved a similarity score between manually and automatically renal delineation, over than 0.9 for both left and right kidneys (Thong et al., 2016).

Most interestingly, the Combined Healthy Abdominal Organ Segmentation (CHAOS) benchmark aims to develop deep networks that are able to segment multiple abdominal organ including kidneys, liver and spleen on CT and MRI images (T1-DUAL and T2-SPIR sequences) of a public dataset, provided with a groundtruth organ delineation (Conze et al., n.d.). Different tasks were involved under the CHAOS challenge, going from liver segmentation on only CT or MRI modality to the most complicated task that tackles abdominal organ segmentation on both CT and MRI modalities using a single network architecture (pre-trained or not) (Conze et al., 2020). All participating teams, except for one, have used an extension of the U-Net deep network and all have achieved a high segmentation process. A recent cross-modality multi-organ segmentation study employed for CHAOS has shown a good performance for multi-organ segmentation task and received the first rank for three competition categories: liver MR, liver CT and multi-organ MR segmentation (see Fig. 9) (Conze et al., 2020). Authors have tested different pipelines based on different pre-trained networks. The highest right and left kidney segmentation score was achieved when a cascaded pre-trained U-Net associated with a conditional Generative Adversarial Network (cGAN) was used. The proposed architecture has led to strengthen the ability of the network to segment multiple organs with a good generalization capability. Briefly, in a conditional GAN, the generator learns to create synthetic images with some specific conditions. In terms of segmentation task, the generator, as its name shows, will generate masks through its layers of encoding and decoding while the discriminator assesses either the generated mask is realistic or not. Therefore, the adversarial network will be able to differentiate realistic and synthetic delineations and enforce the generator to create masks as real as possible. Cascaded CEDs was used in order to exploit multi-level contextual information (Conze et al., 2020).
Performing automatic renal segmentation based on deep networks was reported to be promising. Deep networks take the advantage of being supervised methods that do not require handcraft features. Interventions are limited to set the network hyper parameters (batch size, number of iterations) that could affect the model convergence. Most interestingly, such networks could offer an adequate solution for cross-modality problem. In another way, a single model could be trained to segment MRI as well as CT images.

Adopting deep learning for renal segmentation imposes several limitations to be mentioned. Deep networks are supervised methods that require a hug amount of training ground-truth data, which can be surpassed by using various data augmentation schemas as well as GANs for synthetic data generation. Additionally, transfer learning offers a way to accelerate the model convergence during the training phase (Conze et al., 2020). Moreover, it is well known that determining an adequate architecture for a specific semantic segmentation problem is hard. Thus, testing several network architectures and comparing the segmentation results is the best way to precise which model could offer the highest segmentation performance. One should be taken into account is the need of deep networks to large memories and powerful graphics card, as well as the need of experience in order to find optimal parameters for convergence (Kavur et al., 2020).

Although many kidney segmentation methods and deep models have been developed, studies in this field are still limited and more models have to be introduced in future works. Deep segmentation frameworks such as DeepMedic (Kamnitsas et al., 2017) and NiftyNet (Gibson et al., 2018) must be explored. Networks such as VNet (Milletari et al., 2016), ScaleNet (Fidon et al., 2017) and HighRes3dNet (W. Li et al., 2017) have to be evaluated on renal segmentation. Interestingly, incorporating active contour models in deep network framework takes the advantage of inheriting the merits of both. Several approaches have been proposed, (i) AC can be used as post-processing tool after deep network, (ii) within the deep network, it can replace the fully connected layer for segmentation, (iii) the energy minimization of AC can be used as the network loss function, (vi) deep network can learn parameters from AC. Such approaches are yet to be explored in the field of kidney segmentation (Hoang Ngan Le et al., 2020). Most importantly, deep networks could work in parallel if necessary computational power is available. Therefore, their results can be combined to obtain superior performance through an ensemble system (Kavur et al., 2020). The Ensembles of Multiple Models and Architectures (EMMA) model is a great example of such fusion systems (Kamnitsas et al., 2018). Applying such promising approach on renal segmentation is yet to be introduced. Lastly, performing segmentation of renal structures (e.g. medulla, cortex, pelvis) based on deep learning is yet to be explored.

**Conclusion**

Renal disease is characterized by alteration in renal macrostructure including kidney volume and CMD, and microstructure including fibrosis composition and lipid fraction. Developing methodologies to diagnose and predict CKD at early stage still a challenging issue that can reduce the cost of disease treatment and slow the kidney damage progression.
MRI was demonstrated to be a powerful tool to evaluate renal tissue by assessing both renal function and structure for both kidneys. MRI embeds several sequences reflecting different renal properties and function including diffusion, perfusion, oxygenation, tissue elastography, hemodynamics and others. Although MRI sequences offer the opportunity to estimate microvascular and microstructural integrity of both kidneys but are still limited to the fact that other factors could contribute to the marker measured (e.g. urine flow rate, edema and medications for DWI, intravascular volume and tubular dysfunction for BOLD, intravascular volume and medications for ASL etc.). Further, human studies evaluating the effectiveness of MT and T1 as well T2 mapping is required. Multi-parametric MRI studies would give a better reflection of the association between all renal functionality measurements and could select the most effective MRI sequences to detect renal failure. Taking the advantage of its low cost and availability, UE also has offered promising results to monitor CKD through measuring the renal stiffness as MRE does, but the conflicting results obtained from UE studies make the MRE more suitable for renal stiffness assessment. Despite their effectiveness in renal dysfunction detection, CT and scintigraphy are still limited to the radiation exposure and contrast agent injection.

In terms of routine check-up for renal function in order to prevent kidneys from being affected, US as well as UE and CT present suitable imaging techniques that could replace biopsy and GFR test. Those modalities offer a low cost, non-invasive, clinically available and short time examination techniques. Those imaging techniques could provide a rapid assessment tool when strong CKD suspicions appear. One should be aware of is the radiation exposure caused by CT. Given that renal tissue changes occur at the microscopic level that often need a biopsy for assessment, nephrology should take the advantage from functional MRI for risky patients. MRI could predict renal failure that is associated with function disturbance as it is able to give functional and structural parameters safely. For CKD patients or risky patients, contrast-enhanced and radiation exposure imaging should be eliminated due to their side effects especially when kidneys are already affected.

Texture analysis takes the advantage of reducing inter-observer variability since the analysis covers the whole renal parenchyma as it is a spatially heterogeneous tissue due to its internal structure. Textures were able to extend the possibility of medical imaging and could function as a complementary evidence to conventional qualitative markers. Artificial intelligence detects earlier the renal deformation. It carries out the use of machine learning techniques and deep neural networks that has performed well in earlier CKD detection and GFR estimation through AI-GFR. Artificial intelligence has demonstrated its power to turn conventional medical imaging into a real-time screening tool that could help doctors in clinical decision-making. The ResNet network served from a textural branch has outperformed other networks and shown to be a promising architecture that has to be tested on other imaging modalities. Although the ability of deep networks to hold a fully automated strategy of renal segmentation and disease classification associated with a high performance, CNNs lack the huge amount of data for training that should be overcome by using the transfer learning approach and data augmentation, and require a high computational power for their implementation (e.g. larger memory, powerful graphical cards). Furthermore, deep networks need a strong experience in order to pick the optimal parameters for the convergence and to set the parameters of optimization.

Even their promising results in CKD early detection and GFR estimation, proposed models based on textural analysis and deep networks are still not enough to be incorporated to clinical practice until an excellent detection and estimation performance is achieved. Having an accuracy <99% means that misclassification of some cases still exists. Future research work should investigate the successful literatute models in a large follow-up and multi-center studies that uses standardized imaging protocols and homogeneous patient preparation scene (e.g. fasting, standardized hydration status, verification of salt and medication intake prior to imaging) with a balanced distribution of positive and negative samples. The selection of the most reproducible and discriminative features is needed. The association between textures as well as deep features must be evaluated to profoundly understand their role in renal impairment detection and staging.

An effort must be made by the researchers to apply automatic kidney segmentation prior to the disease diagnosis using different imaging modality in order to rich a fully automated renal dysfunction detection network with a high accuracy and efficiency. The VGG-16 as well as the cGAN/U-Net, pre-trained on ImageNet challenge, have shown promise in the segmentation task. Achieving a multi-modality segmentation and CKD evaluation using a single framework must be addressed since cross-modality training still more challenging than individual learning. Future work must include evaluation of different deep networks in cross-modality renal segmentation (DeepMedic,
ScaleNet, VNet and HighRes3dNet). Additionally, the use of deep networks incorporated with active contour in the field of renal segmentation have to be investigated. Furthermore, the efficacy of ensemble fusion system in this context has to be tested. Incorporating kidney segmentation deep models in clinical applications has to have an excellent delineation performance (e.g. DICE score, accuracy, sensitivity).

CKD staging challenge must be further evaluated with deep neural networks. Incorporating textures from other imaging modalities into these networks needs further validation to test its effectiveness in clinical decision-making and outcome prediction.

Lastly, in terms of clinical workflow, the gap between research and real-life tools needs to be filled. The successful segmentation as well as renal disease evaluation solutions proposed by researchers have to be implemented in real-world applications. Such framework should be accessible to clinicians during their daily clinical routine.

References


Cheng, P.M., Malhi, H.S., 2017. Transfer Learning with Convolutional Neural Networks for Classification of Abdominal

12, 1–14. https://doi.org/10.1371/journal.pone.0169077


Dujardin, M., Sourbron, S., Luypaert, R., Verbeeelen, D., Stadnik, T., 2005. Quantification of renal perfusion and


Kim, J.K., Yuen, D.A., Leung, G., Jothy, S., Zaltzman, J., Ramesh Prasad, G. V., Prabhudesai, V., Mnatzakanian, G.,


Collected Figure Captions

For all figures: Color online

Fig. 1 Fractional anisotropy (FA) map and tractography of healthy kidney. FA map shows a higher diffusion anisotropy of the medulla (empty arrow) compared to the cortex (full arrow). Direction coded tractography reveals a radial oriented diffusion in the medulla (Notohamiprodjo et al., 2010).

Fig. 2 Examples of different ways to analyze the images obtained with BOLD-MRI. (A) Small ROI technique. (B) TLCO technique. (C) R2* profile: the curve links the mean R2* value of each of 12 layers shown under B, and the red line depicts its steepness. (D) Fractional tissue hypoxia. (E) Compartmental method (Pruijm et al., 2018a).
Fig. 3 Perfusion map generation. Tag (A), control (B) and difference (C) images acquired in a sagittal plane for a healthy transplant kidney subject with negligible motion (eGFR=74 ml/min per 1.73 m²) along with the resulting perfusion map (D) shown in units of ml/min per 100 g (Artz et al., 2011).

Fig. 4 Right: cortex and medulla segmentation of the right kidney. Left: Fitted cortex and medulla concentration curves, with arterial input function (AIF) (Octavia et al., 2017).

Fig. 5 (a) Magnetic transfer ratio (MTR) map, (b) cystic probability map derived from MTR map based on Gaussian mixture model (GMM), (c) fibrotic probability, (d) MTR histogram with GMM classification overlaid, cystic tissue is conveyed by the green component, parenchyma by the red and fibrotic tissue by the blue, (e) corresponding histological slice stained with hematoxylin and eosin, (f) picrosirius red staining (Kline et al., 2016).

Fig. 6 (A) Vibrational wave applied to the kidney. MR scanner acquires images at the same vibrational frequency to detect small displacements in the vibrating organ and calculate renal stiffness. Sample pseudo-colorized displacement images are shown to the left and right of the magnetic resonance scanner in (A), with blue and red depicting alternating phases of the displacement waves. Below each displacement image is the corresponding pseudo-colorized elastogram. (B) In these representative images of healthy (left panels) and fibrotic (right panels) transplant kidneys, standard transverse relaxation time (T2)-weighted images are shown in the top row. In the bottom row, pseudo-colorized elastogram stiffness maps are shown of the same two kidneys, with red color reflecting stiffer tissue. The color bar is scaled from 0 kPa (blue) to red (8 kPa) (Leung et al., 2017).

Fig. 7 Ultrasound images synthetization using deep networks. Real (left side) and generated (right side) Ultrasound Images (Sivanesan et al., 2019).

Fig. 8 Right panel: CNN predictions of ADPKD Kidneys. Segmentations (red contour) of ADPKD kidneys from CT acquisitions are shown. The corresponding CNN-generated probability maps are shown in pseudo colors. Left panel: fully convolutional neural network architecture (CNN). Convolutional and pooling layers were used for feature extraction. Deconvolutional and unpooling layers were used for upsampling feature maps in order to achieve pixelwise segmentation (Sharma et al., 2017a).

Fig. 9 Abdominal multi-organ segmentation on magnetic resonance images (MRI). Segmentation on T1-DUAL in and T2-SPIR sequences using several deep networks; DeepMedic, denseVNet, UNet, v19pUNet, v19pUNet+, v19pUNet 1-1 and cGv19pUNet 1-1. Groundtruth delineation of liver, right kidney, left kidney and spleen are superimposed in red, green, blue and yellow respectively (Conze et al., 2020).

Vitae

Israa Alnazer received here Electrical and Electronics Engineering degree from the Lebanese University, Faculty of Engineering, Lebanon in 2018. Currently, she is a PhD student at the Lebanese University and the University of Poitiers. She is a member of the common laboratory between CNRS (CHU Poitiers, Univ Poitiers, XLIM Lab) and Siemens in France. She is also a member of the Laboratory of Informatics and their Applications (LIA Lab, AZM Center, DSST) in Lebanon. Here research interests are in the field of medical image processing and machine learning technologies for clinical applications.

Pascal Bourdon is an Associate Professor at the University of Poitiers, France. His research interests include medical imaging, computer vision, machine learning and biological data analysis. Dr. Bourdon holds a M.S. degree in Electrical Engineering. He completed his PhD in Signal/Image Processing and Digital Communications in 2004. The basis of his doctoral thesis was on the application of Partial Differential Equations-based image processing methods for source coding and joint source/channel coding. From 2005-2013, he has worked as a researcher for the Thomson R&D and Technicolor R&I companies, covering topics such as color science and automatic face analysis for digital cinema.

Thierry Urruty is assistant professor at university of Poitiers since 2011. His research interests are image processing, analysis, indexing and retrieval which includes an expertise in computer vision and machine learning technologies for different applications as cultural heritage and healthcare. He has participated in various French and European projects with multiple industrial and academic partners. He is currently working in a French national project DIGIPI.
which aims at providing metrology foundation for spectral images. He is also an active member of the common laboratory between CNRS (CHU Poitiers, Univ Poitiers, XLIM) and Siemens.

Omar Falou received his PhD in Electrical and Computer Engineering from Ryerson University, Canada in 2011. He is currently an Assistant Professor in the Faculty of Science at the Lebanese University and an instructor at the Lebanese American University and the American University of Culture and Education. His research interests are in the field of medical imaging, particularly quantitative ultrasound, photoacoustics and magnetic resonance imaging. He is a founding officer and treasurer of the Institute of Electrical and Electronics Engineers—Engineering in Medicine and Biology Society (IEEE—EMBS), Lebanon Section.

Mohamad Khalil is currently professor, teacher and researcher at Lebanese University, faculty of engineering. He received the DEA in biomedical engineering from the University of Technology of Compiègne (UTC) in France in 1996. He received his PhD from the University of Technology of Troyes in France in 1999. He received his HDR (Habilitation a diriger des recherches) from UTC in 2006. He is the chair of the EMBS chapter in Lebanon, chair of ICABME international Conference. His current interests are the signal and image processing problems: detection, classification, analysis, representation and modeling of non stationary signals.

Ahmad Shahin has BSc degree in Mathematics and Computer Science from Lebanese University (LU), Master’s Degree in Automatic and Computer Engineering from La Rochelle University – France (LRUF), and PhD in Computer Science from LRUF. Ahmad has worked on Doppler Color aliasing correction with CHU-Poitiers – France, for several years. Currently, Mr. Shahin is professor at LU. He is also Director of Laboratory of Informatics and their Applications, Doctoral School in Science and Technology (DSST), LU. Dr. Shahin’s research focuses on medical image processing, biometrics, and data mining for medical predictions. For several years, he was head of CIS department at LU.

Christine Fernandez-Maloigne is currently Vice-Rector of the University of Poitiers since June 2016. Professor in signal and image processing, specialized in digital color and medical imaging, she is co-director of the laboratory I3M, on multi-organ metabolic imaging between CNRS, SIEMENS, the University and the Hospital of Poitiers. She received the Augustin Fresnel award in 2012. She is member of the National Council of Universities and expert for many European organizations. She is secretary of the Division Image Technology of the International Commission for Lighting since 2015 where she represents France. She is also Deputy Editor in Chief of JOSA A.
Kidney Segmentation

Qualitative Analysis
Mean Value, Cortico-Medullary Difference, etc.

Quantitative Analysis
Textures, Geometricals

Feature Extraction
Conventional Classification
LR, SVM, K-NN, etc.

Deep Networks
Deep Features Extraction and Classification

MRI
DWI, DT, BOLD, ASL, DCE, T1-T2 Mapping, MT, MRE, SWI, DIXON, MRSI

UE
SE, SWE

CT
Scintigraphy
PET, SPECT