SEMMELWEIS EGYETEM

DOKTORI ISKOLA

Ph.D. értekezések

2932.

BUDAI BETTINA KATALIN

Gasztroenterológia
című program

Programvezető: Dr. Molnár Béla, kutatóprofesszor

Témavezető: Dr. Kaposi Novák Pál, egyetemi docens

THE ROLE OF CT TEXTURE ANALYSIS IN THE DIAGNOSTICS OF BENIGN AND MALIGNANT LESIONS OF ABDOMINAL ORGANS

Ph.D. thesis

Bettina Katalin Budai, MD

Károly Rácz Doctoral School of Clinical Medicine
Semmelweis University

Supervisor: Pál Novák Kaposi, MD, Ph.D.
Official reviewers: Dávid Tárnoki, MD, Ph.D.
Márton József Kolossváry, MD, Ph.D.

Head of the Complex Examination Committee: Béla Molnár, MD, D.Sc.

Members of the Complex Examination Committee: Ádám Tárnoki, MD, Ph.D.
István Hritz, MD, Ph.D.

Budapest
2023

Table of Contents

LIST OF ABBREVIATIONS

AI	artificial intelligence
ANN	artificial neural network
AUC	area under the ROC curve
ccRCC	clear cell renal cell carcinoma
chRCC	chromophobe cell renal cell carcinoma
CI	confidence interval
CT	computed tomography
CTTA	CT texture analysis
DICOM	digital imaging and communication in medicine
EFSUMB	European Federation of Societies for Ultrasound in Medicine and Biology
FDR	false discovery rate
GLCM	gray-level co-occurrence matrix
GLDM	gray-level dependence matrix
GLRLM	gray-level run-length matrix (
GLSZM	gray-level size zone matrix
HBV	hepatitis B virus
HCV	hepatitis C virus
HU	Hounsfield unit
IBSI	Image Biomarker Standardization Initiative
ICC	intraclass correlation coefficient
KITS19	kidney and kidney tumor segmentation challenge 2019
LASSO	least absolute shrinkage and selection operator
LI-RADS	liver reporting and data system

ML	machine learning
MRI	magnetic resonance imaging
NAFLD	non-alcoholic fatty liver disease
NGTDM	neighboring gray-tone difference matrix
OMAR	orthopedic metal artifact reduction
PACS	picture archiving and communication system
PCA	principal component analysis
PET	positron emission tomography
pRCC	papillary renal cell carcinoma
RCC	renal cell carcinoma
RFC	random forest classifier
RFE	recursive feature elimination
ROC	receiver operating characteristic curve
ROI	region of interest
SVC	support vector classifier
SWE	shear wave elastography
TCGA	The Cancer Genome Atlas database
US	ultrasound
VOI	volume of interest
WHO	World Health Organization
WFUMB	World Federation for Ultrasound in Medicine and Biology

1. INTRODUCTION

1.1. Introduction to radiomics analysis

Imaging biomarkers can be defined as tissue characteristics that are assessed in images obtained during medical imaging procedures and that carry information about biological processes. Among imaging biomarkers, a distinction can be made between qualitative markers, which allow the subjective characterization of the given tissue, and quantitative markers, which can be measured in a quantitative manner and thus allow the objective characterization of tissues.

In daily clinical practice, radiological images obtained during medical imaging examinations are evaluated on a subjective, qualitative basis. Even current radiological diagnostic classification systems, such as the Liver Reporting and Data System (LI-RADS) (1) estimate the risk of tumor malignancy based on the descriptive morphological characteristics of the tumor. However, these qualitative imaging features are subjective and strongly depend on the radiologist's expertise. The fact that subjective reading of medical images is constrained by the color depth of monitors and visual resolution, as well as other factors, means that a significant amount of the information contained in the images is lost during evaluation.

As more medical data has become available in digital form through the digitalization of medical imaging modalities, new and more complex software have been developed to process, analyze, and evaluate them. As a result of the computerized analysis of medical images, a new subspecialty of radiology called "radiomics" has emerged inspired by the fields of proteomics and genomics. The term, radiomics has been introduced by Lambin et al. in 2012 (2). Radiomics analysis, as opposed to the traditional reading of medical images, is able to extract a large number of quantitative image features so called "radiomics features" based on well-defined mathematical equations from the regions of interest (ROI) or volumes of interest (VOI) of the scans obtained from medical imaging acquisitions. The aims of radiomics analysis are to explore the relationship between radiomics features and various clinical endpoints, and to identify imaging biomarkers with diagnostic and predictive value in determining tumor types, identifying high-risk patients with poor survival outcomes, and identifying those potential candidate patients who can benefit from a particular oncological treatment (3).

However, as a result of feature extraction, radiomics produce so-called “big data”. Big data is defined as data sets that are too large and complex for researchers to handle using traditional data processing methods. To extract meaningful information with predictive value from the vast amount of multidimensional data, techniques from the field of artificial intelligence (AI) are needed (4). Therefore, machine learning (ML) algorithms are almost always used in radiomics studies in addition to radiomics analysis.

Previous studies showed that radiomics analysis can be used for facilitating the non-invasive diagnostics of both benign and malignant lesions of many different organs (5). While radiomics analysis combined with ML can speed up and improve non-invasive diagnosis by automating the evaluation of medical images, the development of clinical decision support tools that incorporate radiomics data alongside patients’ clinical data can help clinicians in identifying high-risk patients and assist in therapeutic decision-making.

The main steps of radiomics analysis are image acquisition, image post-processing, segmentation, radiomics feature extraction, feature selection, model building, and evaluation (Figure 1), which will be covered individually in the following sections.

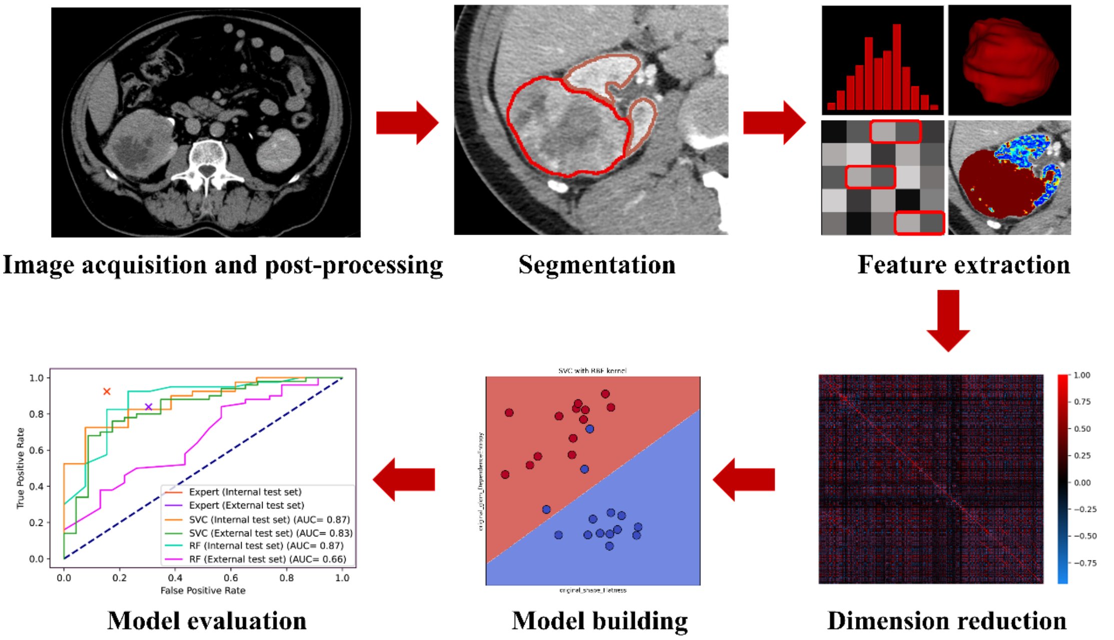


Figure 1. Workflow of radiomics analysis. *The main steps of radiomics analysis are image acquisition followed by post-processing, lesion segmentation, radiomics feature extraction, dimension reduction, model building, and model evaluation. (Own source (6))*

1.1.1. Image acquisition and post-processing

Despite the fact that the majority of studies to date have used CT scans, radiomics analysis can be applied to scans from all types of imaging acquisitions, including computed tomography (CT), positron emission tomography (PET), magnetic resonance imaging (MRI) scans, X-ray, and ultrasound (US) images. The first step in the workflow of a radiomics study is to collect and export medical images. To prevent losing potentially useful image features, images should be exported from the picture archiving and communication system (PACS) as Digital Imaging and Communication in Medicine (DICOM) files. Then, the images must be anonymized to remove patient-identifiable metadata.

The medical images need further post-processing before radiomics analysis to increase the subsequently extracted radiomics features' robustness and reproducibility. The most pivotal post-processing steps are intensity normalization, noise reduction, image registration, isotropic voxel resampling, and gray-level discretization (7).

It is important to highlight that in MRI scans there is no fixed voxel signal intensity range for each tissue, unlike CT, where different tissues have well-defined density values within their more or less specific density range. Therefore, before performing radiomics analysis, it is necessary to standardize the signal intensity range of the selected ROI/VOI of MRI images. Plenty of different methods exist in the literature for standardization, however, this rescaling of voxel intensity values has a marked effect on the extracted radiomics features especially on first-order statistical features (8, 9).

The variations in contrast administration protocols and patient hemodynamics make the contrast enhancement of organs and tumors dependent on these external factors, making it challenging to identify tumor-specific radiomic features (10, 11). However, no universally accepted post-processing method has been proposed so far to deal with this frequently encountered issue during radiomics analysis of post-contrast CT scans.

Denoising is also an important post-processing step to increase the quality of images before feature extraction, since radiomics parameters, especially texture parameters, are sensitive to the signal-to-noise ratio (12). The use of noise filtering is a particularly crucial post-processing step for MRI images because of the presence of Gaussian and Rician noise (13). Meanwhile, in CT scans, noise filtering is most frequently needed for metal implant artifacts (14). However, various artifact reduction algorithms such as the orthopedic metal artifact reduction (OMAR) algorithm now enable effective image evaluation even on CT images acquired with suboptimal quality (15).

Previous studies showed that a significant number of radiomics features show dependency on voxel size that can be corrected with resampling (16). Isotropic resampling is recommended during radiomics feature extraction in order to get rotation-invariant radiomics features. However, whether down-sampling or up-sampling is preferable is not yet clarified (17).

Finally, image registration for motion correction is a fundamental post-processing step during radiomics analysis of multi-phasic post-contrast CT scans and multi-sequence MRI series. In order to ensure that the corresponding voxels from the various phase scans overlap with one another, image registration involves deforming the scan’s geometry using a variety of rigid and flexible algorithms (18).

1.1.2. Segmentation

The radiomics features are always extracted within a selected ROI or VOI. The ROI is frequently the region of the axial slice of the given structure with the largest diameter, whereas the VOI is typically the entire three-dimensional volume of the given structure. The delineation of ROIs/VOIs on the given image is called segmentation, which is most often performed manually (19).

While most of the radiomics studies in the field of oncology define ROIs/VOIs as the area/volume of the lesions, there are examples in the literature investigating the role of radiomics analysis of tumor sub-components (20) and peritumoral areas (21).

To investigate the robustness of radiometric features to inter- and intra-observer differences, either the manual segmentation must be performed by at least two independent readers, or the same reader must perform the manual segmentation twice with an appropriate interval between the two times.

As manual segmentation is considered to be the most time-consuming step in the workflow of radiomics analysis that also suffers from subjectiveness, in order to automate and speed up the process, the need for the development of AI-based segmentation methods has increased in recent years (19). Even novices can easily train the most recently developed convolutional neural network architectures for segmentation tasks if an annotated training data set of sufficient size is available. Moreover, the accuracy of a properly trained neural network can approach that of semi-automatic segmentation tools and the gold standard manual segmentation (Figure 2) while drastically reducing the workflow time (22).

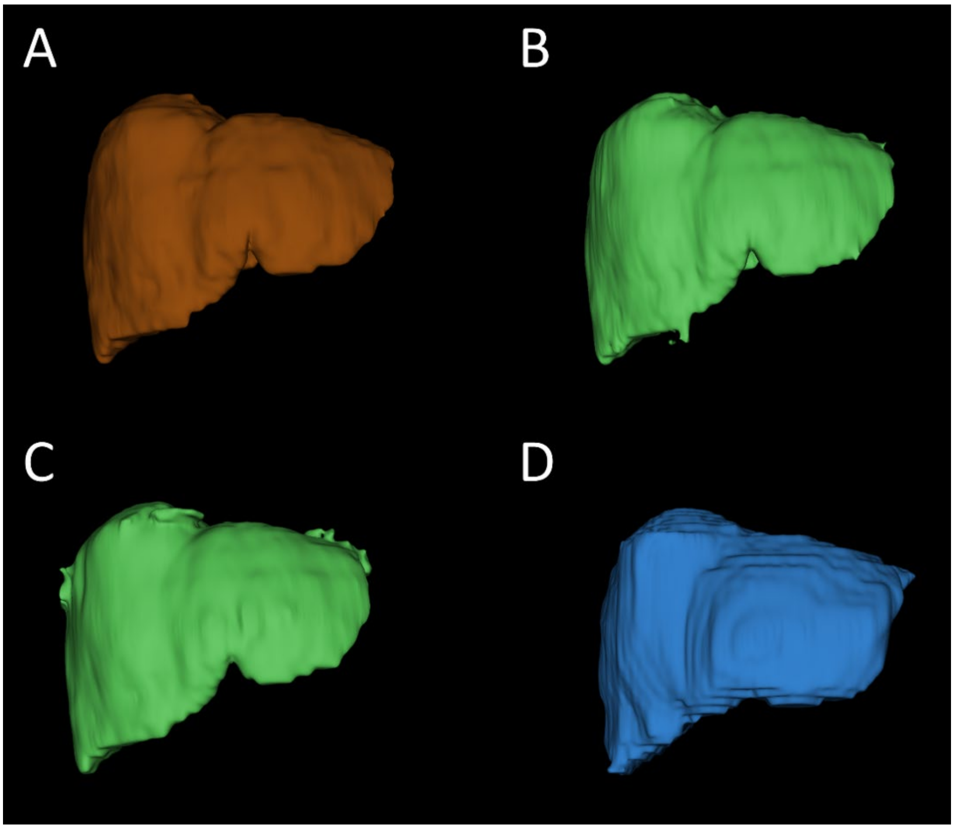


Figure 2. Comparison of manual, semi-automatic, and fully automatic AI-based segmentation of the liver. *The semi-automatic segmentation of the liver completed by the grow from seeds algorithm of the 3DSlicer software showed promising results in both cases when either every 5th (C) or every 3rd (B) slice was segmented manually. Meanwhile, a V-Net convolutional neural network (D) after training on a small number of cases (<50) was able to reduce the time required for segmentation to seconds while reaching comparable accuracy with the manual segmentation (A) (22).*

1.1.3. Radiomics feature extraction

Radiomics feature extraction refers to the automated retrieval of quantitative imaging features from the segmented ROIs of medical images using well-defined mathematical formulas.

The Image Biomarker Standardization Initiative (IBSI), an independent international collaboration, was established in 2016 to standardize the steps of radiomics analysis and the definitions of radiomics features. IBSI's work aims to reduce the diversity of algorithms used in the literature, validate the different software used for radiomics analysis, and thus improve the reproducibility of radiomics research (17).

The radiomics features can be divided into three main classes. The shape-based features assess the morphology of the selected structure including maximum diameter, area, volume, and sphericity among others. First-order statistical features such as mean, entropy, skewness, etc., evaluate the distribution of the intensity values of individual voxels without taking into account the spatial relationship of voxels (23) (Figure 3).

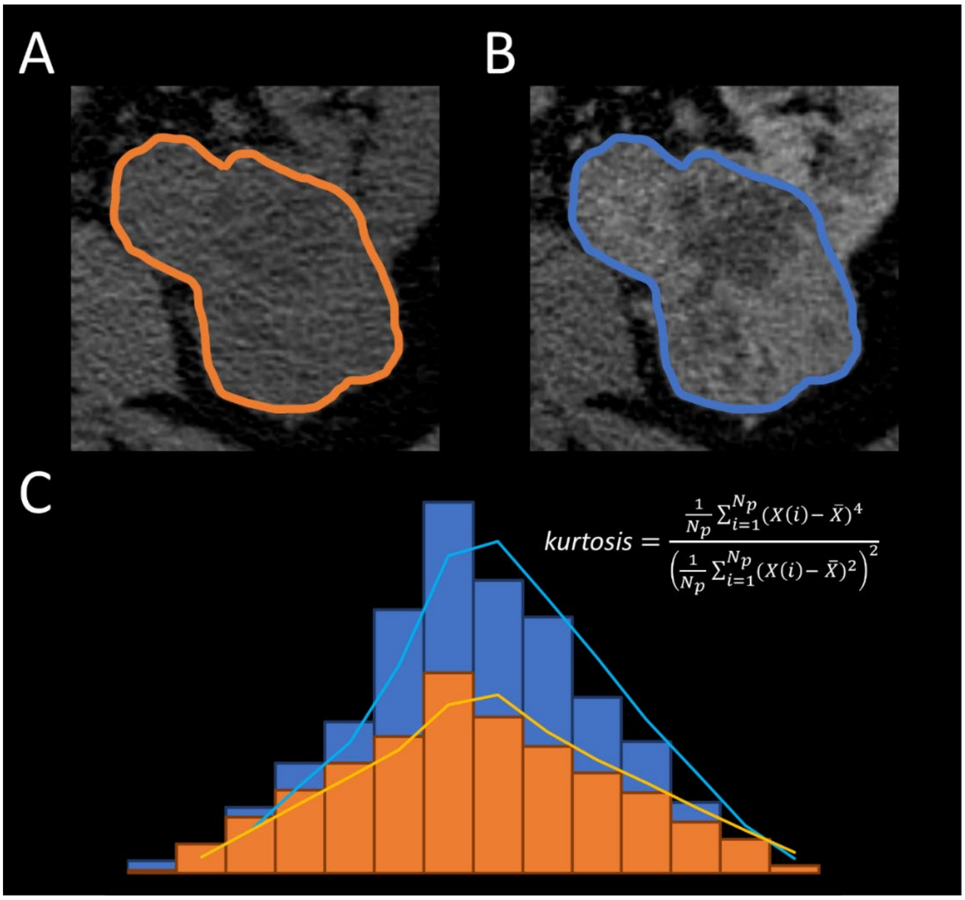


Figure 3. *Schematic illustration of extracting histogram-based first-order statistical features such as kurtosis from unenhanced (A) and corticomedullary phase post-contrast (B) computed tomography scans of a renal cell carcinoma. (Own source (24))*

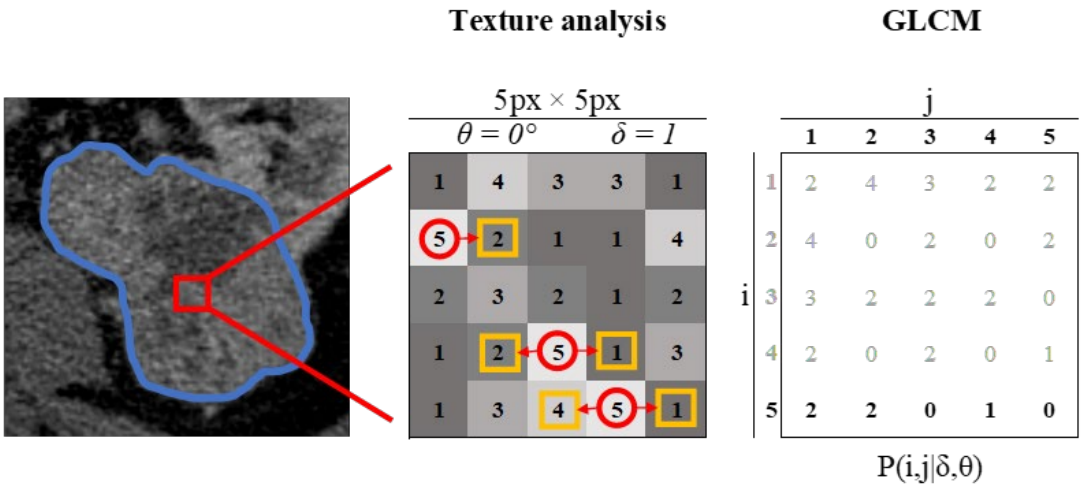
While second-order features also known as texture features quantitatively describe the spatial distribution of voxels with given intensity values and evaluate the voxels' relationship with each other. Texture features are calculated by intensity discretization, which means that each voxel intensity value is assigned to a single bin within a pre-defined range. To improve the reproducibility and robustness of radiomic parameters, the IBSI advises the discretization of voxel intensity values in its reference manual. There are two forms of gray-level discretization, the first is the fixed bin size method which specifies the width of the bins, while the second is the fixed bin count method which specifies the total number of bins used. The authors state that both the fixed bin width and the fixed bin count methods are regarded as acceptable if applied consistently to all images in the study, however in the case of MRI scans due to the arbitrary nature of the signal intensity values, the fixed bin count method is preferable (23).

The main classes of texture parameters are the gray-level co-occurrence matrix (GLCM), the gray-level run-length matrix (GLRLM), the gray-level size zone matrix (GLSZM), the gray-level dependence matrix (GLDM), and the neighboring gray-tone difference matrix (NGTDM) based texture parameters. The detailed description as well as the standardized mathematical formulas of each radiomics feature calculated from these matrices can be found in the IBSI reference manual updated in 2019 (23).

1.1.4. The construction of texture feature matrices

The GLCM-based texture features are the most widely used class of texture features, introduced by Haralick et al. in 1973 (25). For a two-dimensional ROI or a three-dimensional VOI, the GLCM characterizes the co-occurrence of pixels/voxels with given gray values i and j along an axis of angle θ at a given δ distance (23). Most frequently, we characterize the co-occurrence of gray-level values of pixel pairs or voxel pairs that are next to each other ($\delta = 1$) when computing the GLCM-based texture parameters (Figure 4).

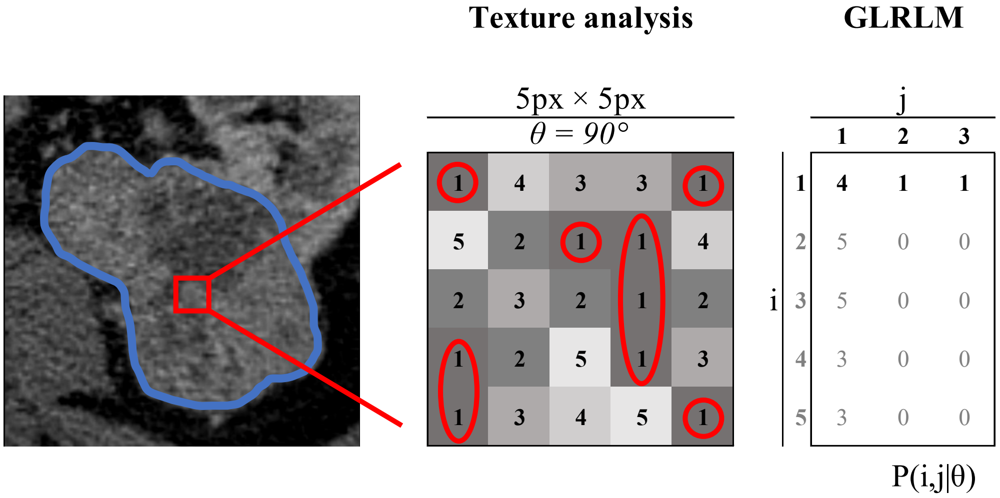
The GLRLM-based texture parameters were created by Galloway in 1975 (26). The GLRLM quantifies the run-length j of connected voxels with the same gray value i along a given θ-axis (23). In other words, it gives the number of those voxels with a given intensity value that are adjacent along a given axis and expresses the frequency of these values within a given segmented VOI (Figure 5).



1	4	3	3	1
5	2	1	1	4
2	3	2	1	2
1	2	5	1	3
1	3	4	5	1

i \ j	1	2	3	4	5
1	2	4	3	2	2
2	4	0	2	0	2
3	3	2	2	2	0
4	2	0	2	0	1
5	2	2	0	1	0

Figure 4. Schematic illustration of the computation of gray-level co-occurrence matrix (GLCM) from a 5-voxel wide region of interest (ROI). *The construction of a GLCM for an angle θ=0° and distance δ=1, where i and j denote the density values of each voxel in the selected ROI. An ROI of 5×5 voxels is defined within an axial CT slice of a renal tumor, where the density values of each voxel are positive integers between 1 and 5 for simplicity. The number of times that neighboring voxels at a given distance δ=1 and along a particular direction θ=0° take on the value of j is calculated for each step of the matrix creation after identifying the voxels with a given value of i. (Own source (27))*



1	4	3	3	1
5	2	1	1	4
2	3	2	1	2
1	2	5	1	3
1	3	4	5	1

i \ j	1	2	3
1	4	1	1
2	5	0	0
3	5	0	0
4	3	0	0
5	3	0	0

Figure 5. Schematic illustration of the computation of gray-level run-length matrix (GLRLM) from a 5-voxel wide region of interest (ROI). *The construction of a GLRLM that summarizes the run-length (j) of adjacent voxels with a given i intensity value along a given direction with angle θ=90°. (Own source (28))*

The GLSZM-based texture features were described by Thibault et al. in 2014 (29). The GLSZM additionally specifies the number of linked voxels with identical gray values. The most notable difference from GRLRM is that GLSZM specifies the number of those given gray-valued voxels that are edge or corner touching within a specific ROI (23). Figure 6 demonstrated the calculation of the GLSZM in a simplified case.

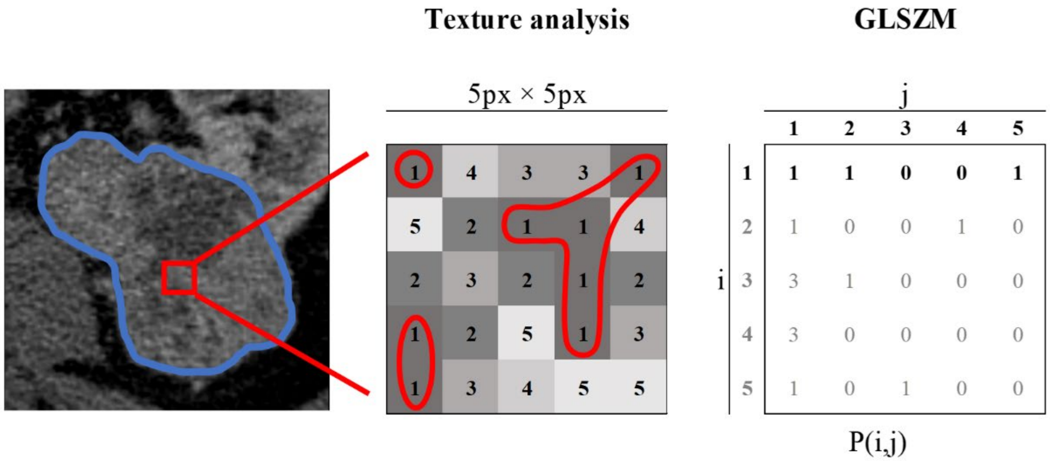


Figure 6. Schematic illustration of the computation of gray-level size zone matrix (GLSZM) from a 5-voxel wide region of interest (ROI). *The construction of a GLSZM, where i denotes the density values of each voxel in the selected ROI, and j means the number of those voxels that have the same intensity value and are either edge or corner touching with each other. An ROI of 5×5 voxels is defined within an axial CT slice of a renal tumor, where the density values of each voxel are positive integers between 1 and 5 for simplicity. (Own source (30))*

The gray-level dependence matrix (GLDM)-based radiomics features were developed by Sun et al. in 1983 (31). During calculating the GLDM, we quantify the number of times that a central voxel with a given *i* intensity value has a neighboring voxel within a given δ distance that has the same *i* intensity ($\alpha = 0$) (23). Figure 7 illustrates the construction of GLSZM in a simplified case.

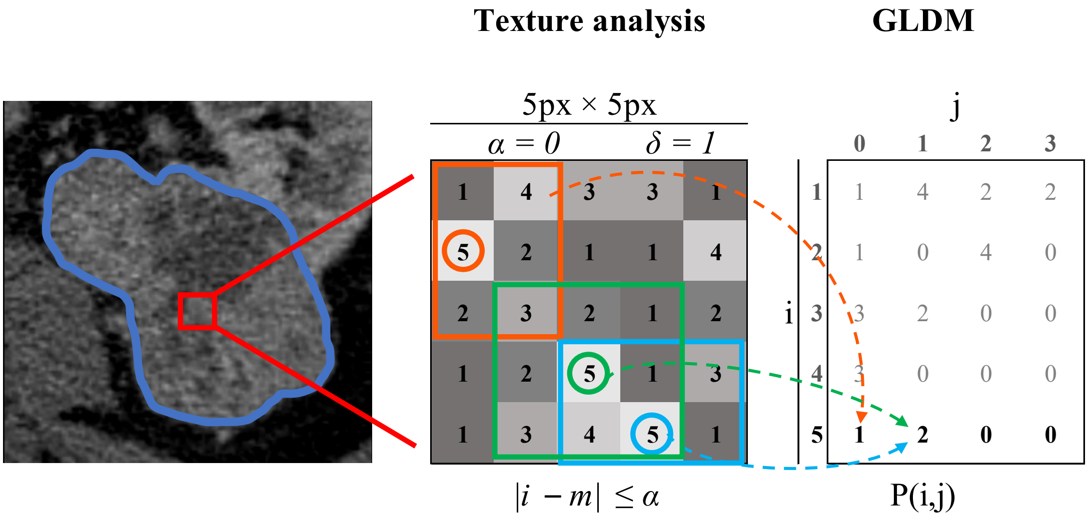
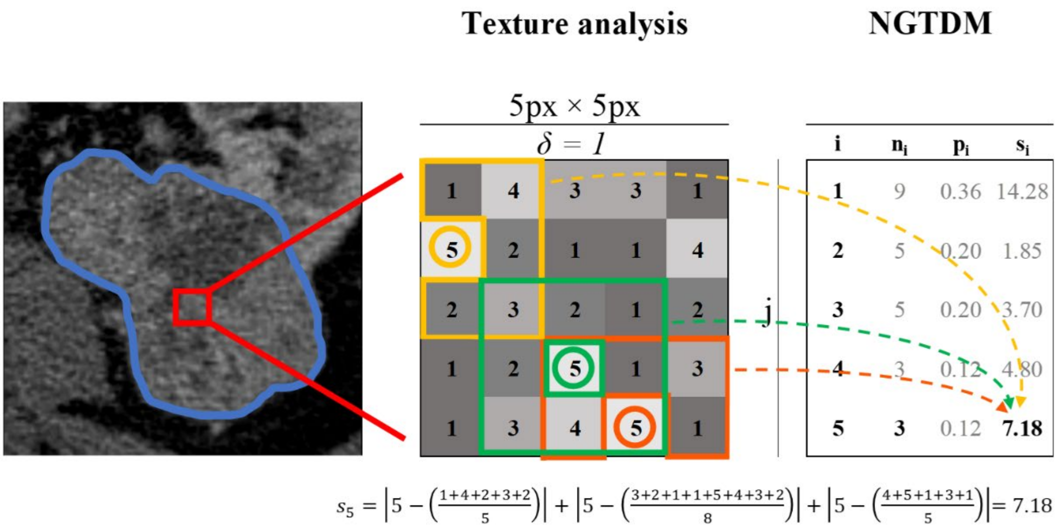


Figure 7. Schematic illustration of the computation of gray-level dependence matrix (GLDM) from a 5-voxel wide region of interest (ROI). *In the GLDM construction, i denotes the density values assigned to each voxel, and j denotes the number of voxels in the neighborhood of the central voxel that are "dependent" on it according to the criterion |i-m|≤α. The value of δ is most often chosen to be 1, in which case only the neighboring voxels directly adjacent to the central voxel are considered. The value of α is most often chosen as 0, in which case |i-m|≤α criterion meets the i=m criterion. An ROI of 5×5 voxels is defined within an axial CT slice of a renal tumor, where the density values of each voxel are positive integers between 1 and 5 for simplicity. (Own source (32))*

The neighboring gray-tone difference matrix (NGTDM) was introduced by Amadasun et al. in 1983 (33). The NGTDM quantifies the difference between the intensity value of a given central voxel and the average intensity value of its neighboring voxels within a particular distance of δ (23). Figure 8 illustrates the calculation of the NGTDM in a simplified case. When constructing the NGTDM, the number of voxels with a given intensity i (n_i) within the selected area is determined, and their probability of occurrence (p_i) is calculated. The average intensity of the neighboring voxels at a distance δ from these voxels is then subtracted from the intensity of the given gray-valued voxel i. Then, by summing their absolute values, we calculate the values of s_i.



i	n_i	p_i	s_i
1	9	0.36	14.28
2	5	0.20	1.85
3	5	0.20	3.70
4	3	0.12	4.80
5	3	0.12	**7.18**

$$s_5 = \left|5-\left(\frac{1+4+2+3+2}{5}\right)\right| + \left|5-\left(\frac{3+2+1+1+5+4+3+2}{8}\right)\right| + \left|5-\left(\frac{4+5+1+3+1}{5}\right)\right| = 7.18$$

Figure 8. Schematic illustration of the computation of neighboring gray-tone difference matrix (NGTDM) from a 5-voxel wide region of interest (ROI). *An ROI of 5×5 voxels is defined within an axial CT slice of a renal tumor, where the density values of each voxel are positive integers between 1 and 5 for simplicity. The construction of the NGTDM, where i is the intensity value of a given voxel,* n_i *is the number of voxels of a given intensity i, and* p_i *is the probability of occurrence of these voxels. From the intensity of a given gray-valued voxel i, the average intensity of the neighboring voxels at a distance δ=1 from this voxel is then subtracted and their absolute values are summed to obtain the matrix-specific values* s_i*. (Own source (34))*

The detailed list of texture features calculated from the above-mentioned GLCM, GLRLM, GLSZM, GLDM, and NGTDM matrices as well as their definitions and mathematical formulas can be found in the reference manual of the IBSI (23). It is important to note that the software used for radiomics analysis may combine the matrix values calculated along each direction vector in different ways, using different mathematical operations. For example, in the case of the pyRadiomics (35) package coded in Python computer language, the radiomics feature extractor calculates each texture parameter symmetrically for each θ-axis separately taking into account both positive and negative direction vectors, and in the final step, it calculates the average value of the texture features.

1.2. Machine learning

As a result of radiomics analysis, a large number of quantitative imaging features are extracted from the ROI/VOI, producing the so-called big data. To evaluate this big data, map the relationships between the large number of variables, and identify imaging features with predictive value, ML methods based on various mathematical models can be applied.

Computer algorithms that can identify patterns and relationships from input data and then make precise decisions based on the knowledge gained during model training are referred to as "machine learning" (36). During ML model building, the algorithm needs a training dataset consisting of input variables such as radiomics features and clinical variables as predictors, and a test dataset whose cases' outcome the algorithm can predict to evaluate its performance. The two main classes of ML algorithms that are commonly used for radiomics analysis are supervised and unsupervised learning algorithms.

In supervised learning, the algorithm is provided with labeled input data containing the outcomes of the cases that the algorithm should learn and be able to predict. Depending on the nature of the output, the algorithm can be a classification algorithm that assigns the cases to classes and results in a categorical outcome, or a regression algorithm that fits the data and results in a continuous outcome (4). Depending on the endpoint of the study, these outcomes could be tumor type, mutation status, recurrence, or survival, to name a few examples.

Meanwhile, in unsupervised learning, the algorithm is not provided with labels for the input cases, we only determine how many output classes the algorithm should classify the cases into. During model fitting, the algorithm itself determines the pattern in the input data and assigns the cases into output classes based on their similarities (4).

Prediction/classification models can be built from either linear or non-linear ML algorithms. The most used linear algorithms in radiomics studies are logistic regression, and linear support vector classifier (SVC). While the most popular non-linear algorithms for classification are the k-nearest neighbor algorithm, the decision tree, the random forest classifier (RFC), the SVC with kernel trick, and artificial neural networks (ANN) (37).

A phenomenon that must be taken into account during model building known as "overfitting" is when an ML algorithm performs exceptionally well on cases of the training dataset but fails to predict the output labels of the cases when tested on an independent test dataset, indicating that the algorithm has learned a noise pattern that is unique to the training cases but not characteristic of the disease in general. On the contrary, a model that is under-fitted performs poorly in both the training and test set cases because it was unable to recognize significant relationships and patterns in the input data (Figure 9).

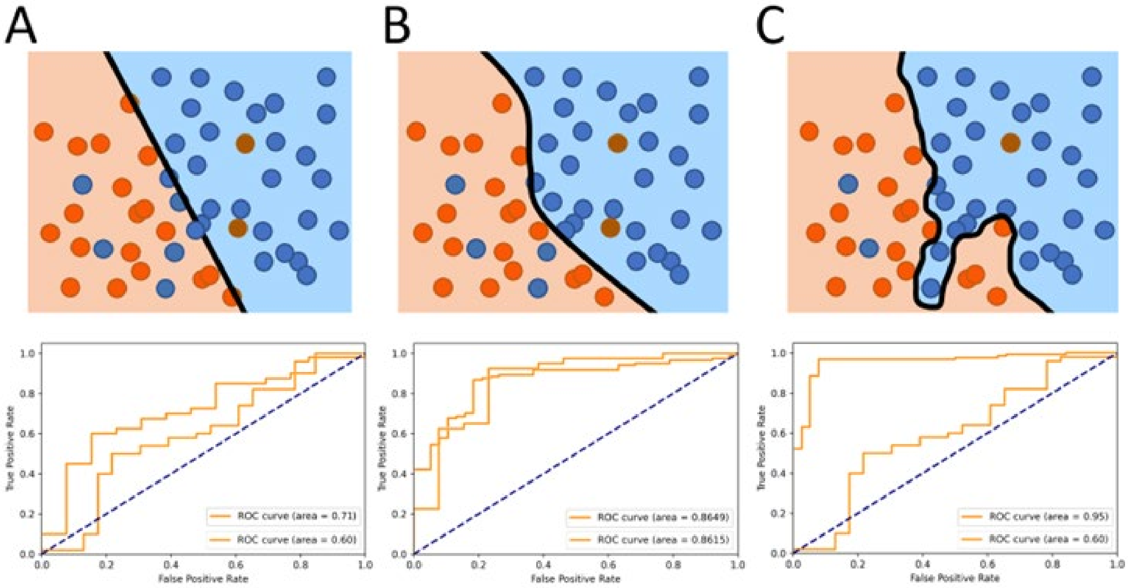


Figure 9. Schematic illustration of fitting machine learning models to the training dataset. *A machine learning model is properly fitted (B) if its performance is similarly high on both the training and the test set. Underfitting (A) and overfitting (B) are two common problems in machine learning model building that impair the generalizability of models. Underfitting means that the model was not able to capture meaningful relationships and patterns in the input data, and therefore the model performs poorly on both the training and test set cases. Conversely, overfitting is the phenomenon where the model captures a specific pattern in the training set that is not characteristic to the disease in general. As a result, the model achieves near-perfect performance on the training set but fails to predict the test set cases. (Own source (38))*

The most important ways to reduce the risk of overfitting are increasing the number of patients, applying regularization to the ML models, optimizing the tunable hyperparameters of ML models, and reducing the number of input variables to a reasonable amount. Both steps should be performed on the training set by means of validation techniques such as k-fold cross-validation (7, 39).

The optimization of ML model hyperparameters is often completed with grid search by using validation datasets. The validation dataset differs from the test set in that it is not independent of the training set cases, instead, it refers to cases decoupled from the training dataset during model optimization and model selection. The grid search is an algorithm that takes into account a wide range of potential hyperparameter properties for an ML model and assesses the model's performance in each potential setting to find the best solution (36). Dimensionality reduction can be carried out by using feature selection and feature extraction algorithms. The three types of feature selection approaches are filters, wrappers, and embedded methods (7). Filters based on univariate statistical tests, such as removing redundant variables with low variance or weak correlation with the outcome, can easily be applied to the dataset to select features. The more complicated wrappers select the most predictive variables with the highest importance score, such as recursive feature elimination (RFE). Embedded methods, which combine the benefits of filters and wrappers, such as the logistic regression-based least absolute shrinkage and selection operator (LASSO) algorithm can also be used to eliminate the redundant features. In contrast to feature selection algorithms, feature extraction algorithms create a new feature set by merging initial features. As an example, principal component analysis (PCA) compresses important features into principal components through the calculation of covariance matrices, eigenvalues, and eigenvectors (40).

The performance of the ML model should be evaluated on independent external test datasets to prove their generalizability. However, in the case of single-center studies, there is no possibility to evaluate models on cases from other institutions. Therefore, it is recommended for single-center retrospective studies to split their own cases into training and hold-out test datasets which can be done either by a random generator or on a temporal basis. However, to ensure the independence of the test cases in studies with this design, it is crucial to take care to prevent data leakage between the training and test datasets (7, 37, 41, 42).

1.3. Radiomics analysis for liver fibrosis staging

1.3.1. The global burden of liver fibrosis

Hepatic fibrosis is a dynamic, progressive-regressive wound healing response process characterized by an imbalance between the synthesis and degradation of the extracellular matrix that involves a number of cell types and mediators. Chronic damaging factors to the liver such as chronic viral infection, drugs, metabolic disorders, or autoimmune diseases activate mechanisms of abnormal matrix synthesis and degradation resulting in changes in the composition and structure of the extracellular matrix around stellate cells. The disrupted homeostasis and impaired cell-matrix interaction lead to the activation of stellate cells, which then transform into myofibroblast-like contractile cells producing proinflammatory cytokines and collagen leading to connective tissue remodeling of the liver parenchyma through chronic accumulation of type I collagen fibers (Figure 10) (43).

Cirrhosis is the most advanced stage of liver fibrosis, which leads to scarring and distortion of the liver tissue through the formation of cirrhotic nodules and connective tissue septa, causing portal hypertension and eventually leading to liver failure (44).

Due to the insufficient reporting quality of certain countries, the estimation of the global burden of liver cirrhosis and the tendency of its prevalence is challenging (45). The Global Burden of Disease Study reported a marked increase in the age-standardized prevalence rate of decompensated cirrhosis between 1990 and 2017 (46). The data showed that while the prevalence of alcohol-associated, hepatitis virus-associated cirrhosis appears to be decreasing in Europe, the prevalence of non-alcoholic fatty liver disease (NAFLD)-associated cirrhosis is increasing (45).

Based on the Global Health Estimates report of the World Health Organization (WHO), the number of deaths due to cirrhosis shows a slightly increasing tendency globally. While in 2000, it was the 11th leading cause of death and accounted for the deaths of 1,098,059 people; in 2019, it was responsible for the deaths of 1,315,359 people. Of these, 311,622 deaths were attributed to hepatitis B virus (HBV) infection, 351,313 to hepatitis C virus (HCV) infection, 323,276 patient deaths were associated with adverse alcohol consumption, and 329,149 deaths were due to cirrhosis of other causes (47).

Meanwhile, in Hungary in 2019, cirrhosis of the liver was responsible for an estimated number of 3059 people's death. While the age-standardized mortality rate for cirrhosis was 19.3/100,000, of which 5.7/100,000 were attributable to HBV infection, 2.9/100,000 to HCV infection, 8.7/100,000 were associated with alcohol consumption, and 2.0/100,000 were due to cirrhosis of other causes. This means that after Moldova, Romania, and Slovakia, Hungary had the 4th highest age-standardized mortality rate and the 3rd highest number of death/100,000 in Europe in 2019 (47).

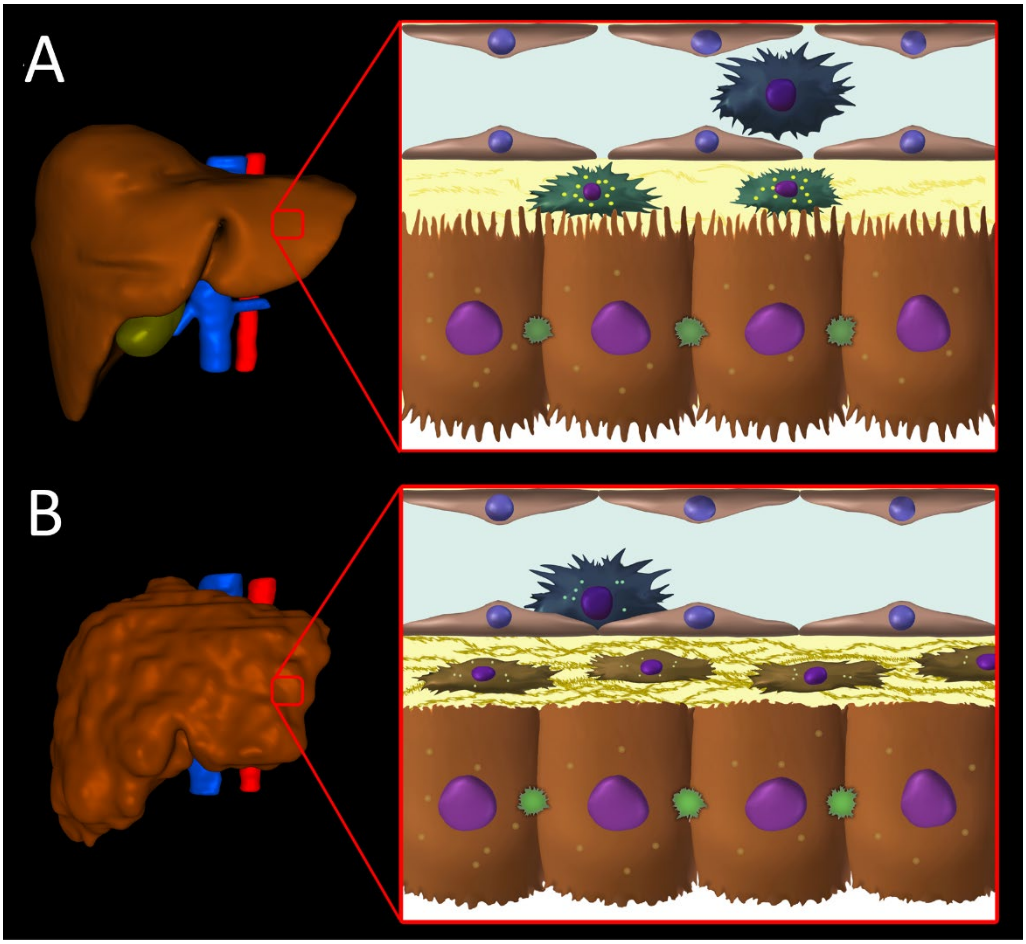


Figure 10. Pathomechanism of liver fibrosis. *In healthy livers, hepatic stellate cells resting in a quiescent state, are located in the space of Disse between the sinusoidal endothelium and hepatocytes (A). In hepatic fibrosis, Kupffer cells, which are activated by liver injury, start to produce cytokines, while stellate cells, which are also activated, transform into myofibroblast-like cells and synthesize large amounts of extracellular matrix. Deposition of collagen fibers and capillarization of sinusoids contribute to the development of cirrhosis and liver failure (B). (Own source (48))*

1.3.2. Diagnosis of liver fibrosis

Early detection of liver fibrosis is critical in clinical practice, as patients who progress into advanced stages of the disease are at increased risk of developing hepatocellular carcinoma, and hepatic decompensation, however, liver fibrosis is considered to be reversible in its early stages.

The current gold standard for the diagnosis and progression of liver fibrosis is percutaneous liver biopsy. Several different systems are used for the histological staging of liver fibrosis, the best-known and most widely used is the METAVIR fibrosis staging system. The METAVIR staging system distinguishes five different stages from F0 to F4. In stage F0, no fibrosis is detected; in F1, portal fibrosis is visible without septa formulation; in F2 stage, some septa are present along with portal fibrosis; in F3, several septa are visible without cirrhosis; finally, the F4 stage represents the state of cirrhosis (49, 50).

The reliability of histopathological examination and staging after biopsy sampling is dependent on the size of the histological sample and the number of portal fields. Sampling error due to the heterogeneous distribution of fibrosis in the liver parenchyma is also a major limitation of the biopsy procedure (51, 52). In addition to this, the reliability of histological diagnosis is further limited by the fact that even with a biopsy specimen of adequate size and quality, the histological result is representative of approximately only 1/50,000th of the total liver (53).

Due to the limitations of liver biopsy, the development of non-invasive alternatives is strongly recommended. Current European practice guidelines published by the European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) suggest shear wave elastography (SWE) as a non-invasive ultrasound-based alternative to liver biopsy to rule out significant liver fibrosis (54). While current guidelines of the World Federation for Ultrasound in Medicine & Biology (WFUMB) recommend SWE as the preferred first-line method for staging liver fibrosis in untreated patients with chronic HCV and rule out advanced stages of liver fibrosis in patients with HBV (55).

The classic morphological signs of end-stage cirrhosis are clearly visible on CT, and also in advanced stages, some morphological signs may be present, including the appearance of regenerative nodules, the surface of the liver becoming nodular, the hilar periportal space becoming dilated, the caudate lobe and the left lobe becoming hypertrophic, while segments VI and VII of the right lobe becoming atrophic. Additionally, CT scans may reveal symptoms of portal hypertension such as splenomegaly, varices, recanalization of the umbilical vein, ascites formation, and portal vein widening (56) (Figure 11).

The conventional reading of CT scans has low sensitivity in the detection of early stages of liver fibrosis and cannot be used for liver fibrosis staging. Previous research has suggested that even in early stages, CT may show subtle differences in texture that are not visible to the naked eye, and illustrated that CT texture analysis can reveal these subtle textural changes of the liver parenchyma facilitating the non-invasive staging of liver fibrosis, however, these studies had a single-center, retrospective study design, were completed on a small number of patients, and assessed either only small regions or selected cross-sections of the liver (57-62).

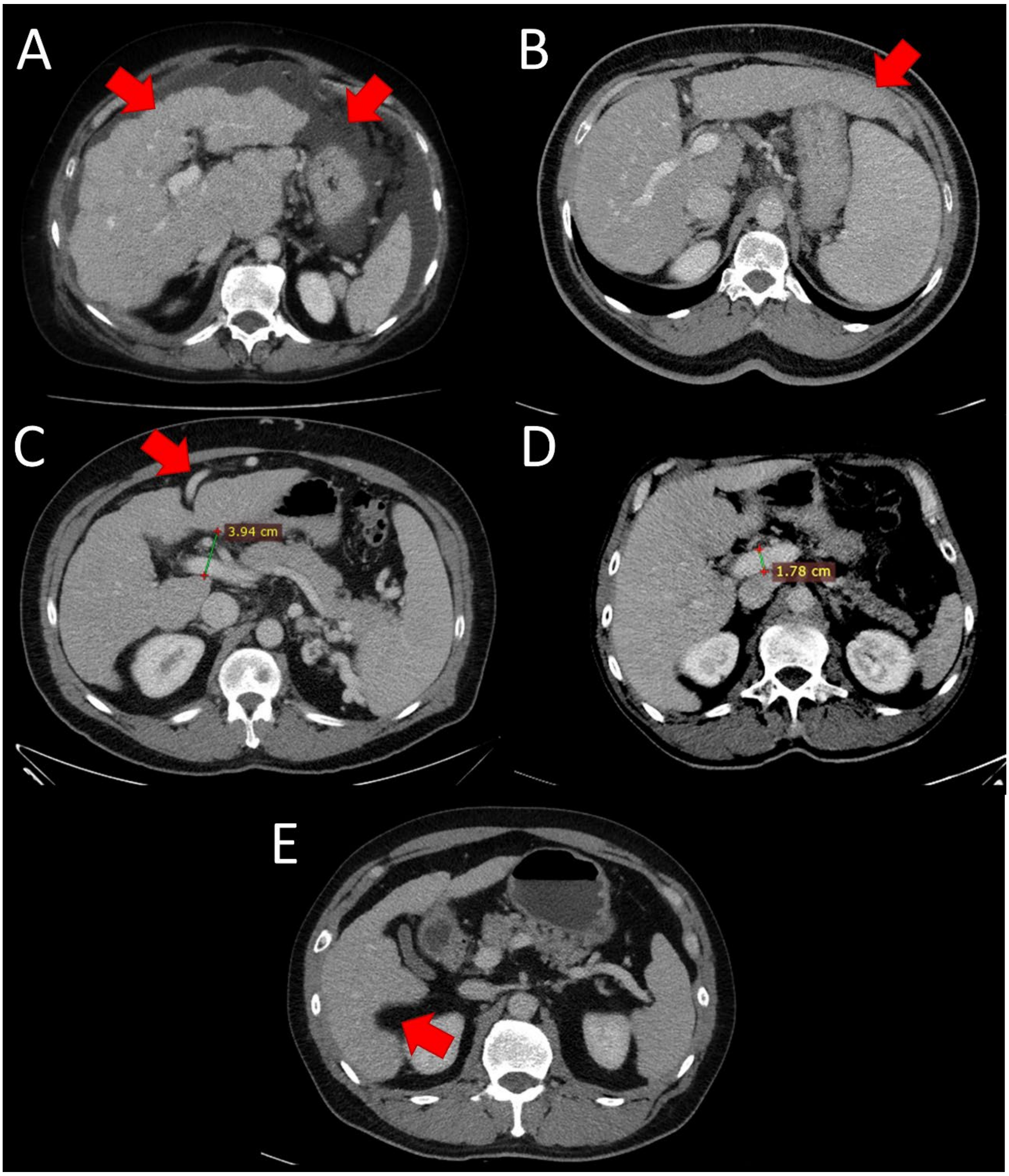


Figure 11. Classic morphological signs of end-stage liver cirrhosis on CT scans. *The morphological signs of cirrhosis on CT scans include liver surface nodularity, ascites, caudate lobe hypertrophy (A), left lobe hypertrophy (B), recanalization of the umbilical vein, widening of the hilar periportal space (C), portal vein widening (D), and hepatic notch sign (E). (Own source (63))*

1.4. Radiomics analysis for differentiating renal cell carcinoma subtypes

1.4.1. The global burden of kidney cancer

Kidney cancer is one of the leading causes of cancer deaths worldwide, accounting for around 179,000 deaths each year. It is almost twice as common in men as in women, making it the 11th and 16th most common malignancy in the world, respectively (64).

In 2020, a total number of 2,057 new cases were diagnosed in Hungary, accounting for 3.1% of all malignancies, while the number of deaths was 720, accounting for 2.4% of all cancer deaths in that year. According to these statistics, kidney cancer was the 8th most common cancer type among all malignancies in terms of incidence and the 11th most lethal cancer type in terms of mortality (65).

1.4.2. The importance of differentiating kidney cancer subtypes

The histological classification and grading of kidney tumors are of paramount importance, as the prognostic and therapeutic implications differ between histological subtypes. The current WHO classification of kidney tumors (66) distinguishes several types of renal neoplasms, but the most common malignant types are renal cell carcinomas (RCC), which account for about 90% of all tumors (67). Its three most common subtypes which contribute to about 75%, 15-20%, and 5% of all cases, are clear cell renal cell carcinoma (ccRCC), papillary cell renal cell carcinoma (pRCC), and chromophobe cell renal cell carcinoma (chRCC) (68). The prognosis of RCC depends mainly on the subtype, tumor stage, histological grade, the presence of rhabdoid or sarcomatoid differentiation, and the presence of necrosis (69). Previous studies have shown that patients with ccRCC have a worse prognosis than patients with pRCC or chRCC, moreover, the histological subtype is an independent predictor of survival (70, 71). There are two forms of papillary subtype, also called type I and type II, the prognosis of type II is worse than type I, but the prognosis of type I is similar to that of the chromophobe subtype which is much better compared to the prognosis of ccRCC (69).

Current clinical practice guidelines state that contrast-enhanced chest, abdominal and pelvic CT acquisition is mandatory for accurate staging of RCCs, which allows for the assessment of local invasiveness, eventual lymph node involvement, and distant metastases (72).

The ccRCC which can occur in both solid and cystic forms, usually appears heterogeneous in the corticomedullary phase and often shows signs of necrosis and intratumoral hemorrhage. The pRCC subtype which also has solid and cystic variations, typically exhibits later contrast enhancement on CT; therefore, this subtype typically has a lower density on the corticomedullary phase compared to the nephrographic phase, which can distinguish it from the ccRCC subtype, which typically has a peak contrast enhancement in the corticomedullary phase. In contrast to the other two, the chRCC subtype is more homogeneous and typically exhibits peripheral contrast enhancement on CT, which is the most prominent during the nephrographic phase (73). Figure 12. shows the main types of RCCs and their typical appearance on contrast-enhanced CT scans.

However, in daily clinical practice, the assessment of morphological features shows huge interobserver differences and the accurate differentiation between these kidney tumor subtypes is challenging due to the huge number of atypical cases. Lee-Felker et al. reported kappa values of 0.33 to 0.76 in the assessment of morphological features on contrast-enhanced CT scans (74). In differentiating ccRCCs from other renal tumors, the authors reported an accuracy of 69-79%, a sensitivity of 90%, and a specificity of 47-67% for radiologists. Meanwhile, Wang et al. reported a sensitivity of 85%, a specificity of 58%, and an AUC of 0.69 for expert radiologists in differentiating ccRCC from non-ccRCC subtypes (75).

In order to overcome the subjective nature of image interpretation, previous studies aimed to investigate the role of radiomics analysis in distinguishing subtypes of RCCs (75-77). Although these prior publications reported promising results, these studies had a retrospective, single-center study design and either did not include validation of the models on independent test cases from external institutions (77, 78), or the models were tested on cases of only one external institution of the same country (79), or although the models were tested on independent cases of several institutions, those showed relevant overfitting resulting in poor performance on the external cases (76).

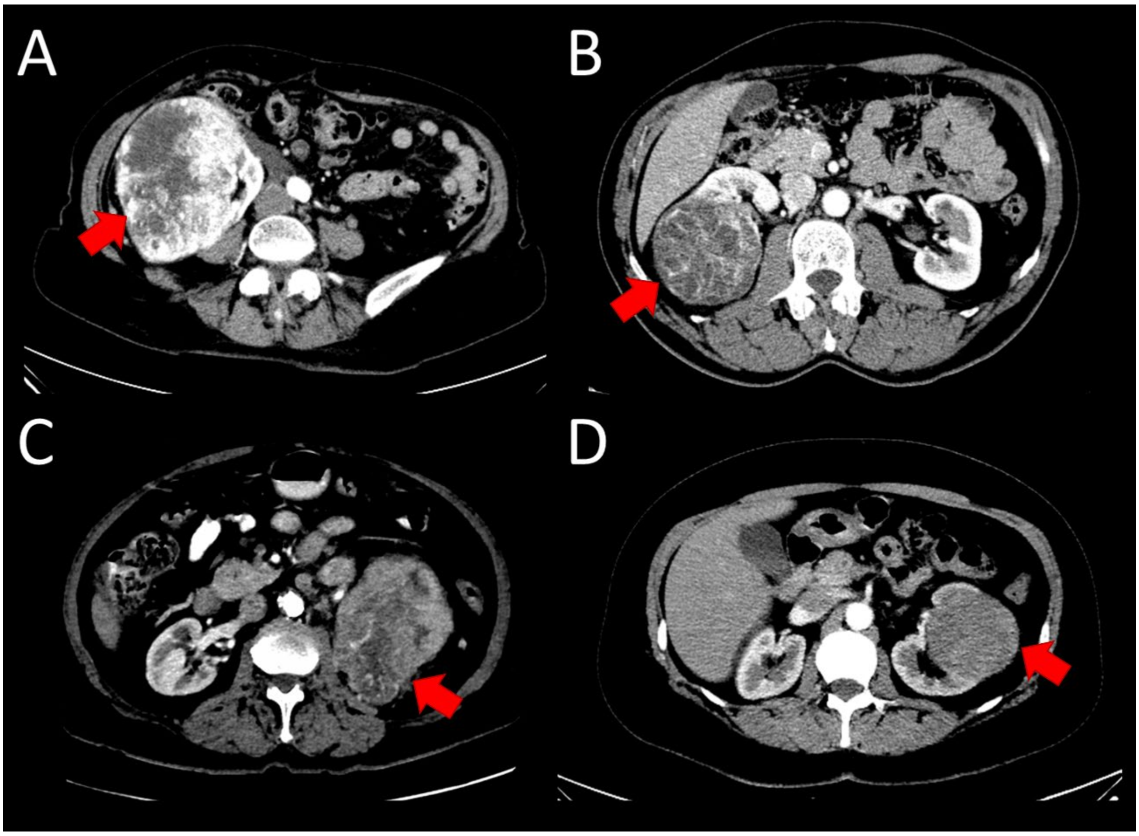


Figure 12. Main types of renal cell carcinomas and their typical appearance on corticomedullary phase post-contrast CT scans. *Clear cell renal cell carcinoma typically shows early contrast-enhancement on CT scans, this subtype has both solid (A) and cystic (B) forms, the solid form usually appears heterogenous on the corticomedullary phase and often shows signs of necrosis and intratumoral hemorrhage. The papillary cell subtype (C) shows later contrast enhancement on CT; therefore, it typically has a lower density in the corticomedullary phase compared to the nephrographic phase which can differentiate it from the clear cell subtype that typically has a peak contrast enhancement in the corticomedullary phase. This subtype also has solid and cystic forms. The chromophobe cell subtype (D) is more homogenous compared to the other two and usually shows peripheral contrast enhancement on CT which is the most pronounced in the nephrographic phase. (Own source (80))*

2. OBJECTIVES

In our first study on CT texture analysis-based liver fibrosis staging, we aimed to diagnose advanced stages of liver fibrosis based on CT images using radiomic analysis and machine learning methods. Furthermore, we aimed to investigate the effect of different CT machines and scanning protocols on radiomics features used for texture quantification. We wanted to investigate the performance of different ML algorithms trained by predictive radiomics features in the differentiation of low-risk and high-risk patients. Our ultimate goal was to build an ML algorithm-based binary classification model for fibrosis assessment, which is independent of etiology and scanning protocols.

In our second study on investigating the utility of radiomics analysis in distinguishing between the histological subtypes of kidney cancer, in order to improve the preoperative non-invasive diagnosis, the aim of our study was to establish a radiomics-based ML model to distinguish ccRCCs from non-ccRCCs based on radiomics analysis of preoperative contrast-enhanced CT scans. To demonstrate that our ML model is generalizable and robust to different institutional imaging protocols, its performance was tested on a publicly available dataset as an external test set.

3. METHODS

The two studies reported below, and their results have been published as original research articles (81, 82).

Both studies were approved by Semmelweis University Regional and Institutional Committee of Science and Research Ethics (RKEB: 136/2019) according to the World Medical Association guidelines and Declaration of Helsinki, revised in 2000 in Edinburgh. As both had a retrospective study design, and all patient data were analyzed anonymously, the ethics committee waived the requirement of written informed consent for participation. However, in compliance with our institutional protocols, written informed consent was obtained before the CT and SWE scans from all patients.

The main steps of radiomics analysis were the same between our two studies, however, the methods showed some differences, which are later discussed in the discussion section of the thesis. For an easier interpretation of the current Methods section, the main differences are briefly summarized hereby. In the first study on liver fibrosis assessment, we used a bin width value of 25 for texture feature extraction, while in the second study on kidney cancers, we set the bin width value to 16 considered as optimal based on the intensity ranges of the segmented tumors. In the first study, Laplacian of Gaussian (LoG) filters and wavelet filters were also applied and higher-order radiomics features were also calculated, while the second study was focused on the original images without image filter application. In the first study, robust scaling was used for data standardization due to outliers, while in the second study, classical standardization was applied. During feature selection, an FDR filter algorithm was applied in the first study, while in the second study, we investigated the radiomics features reproducibility against slight differences in segmentation by intraclass correlation coefficient (ICC) analysis. In both studies, a linear and a non-linear ML analysis design were directly compared to each other, however, the ML-based feature selection algorithms differed between the two studies.

3.1. Radiomics analysis for liver fibrosis staging

The main steps of feature selection and machine learning analysis are illustrated on Figure 13.

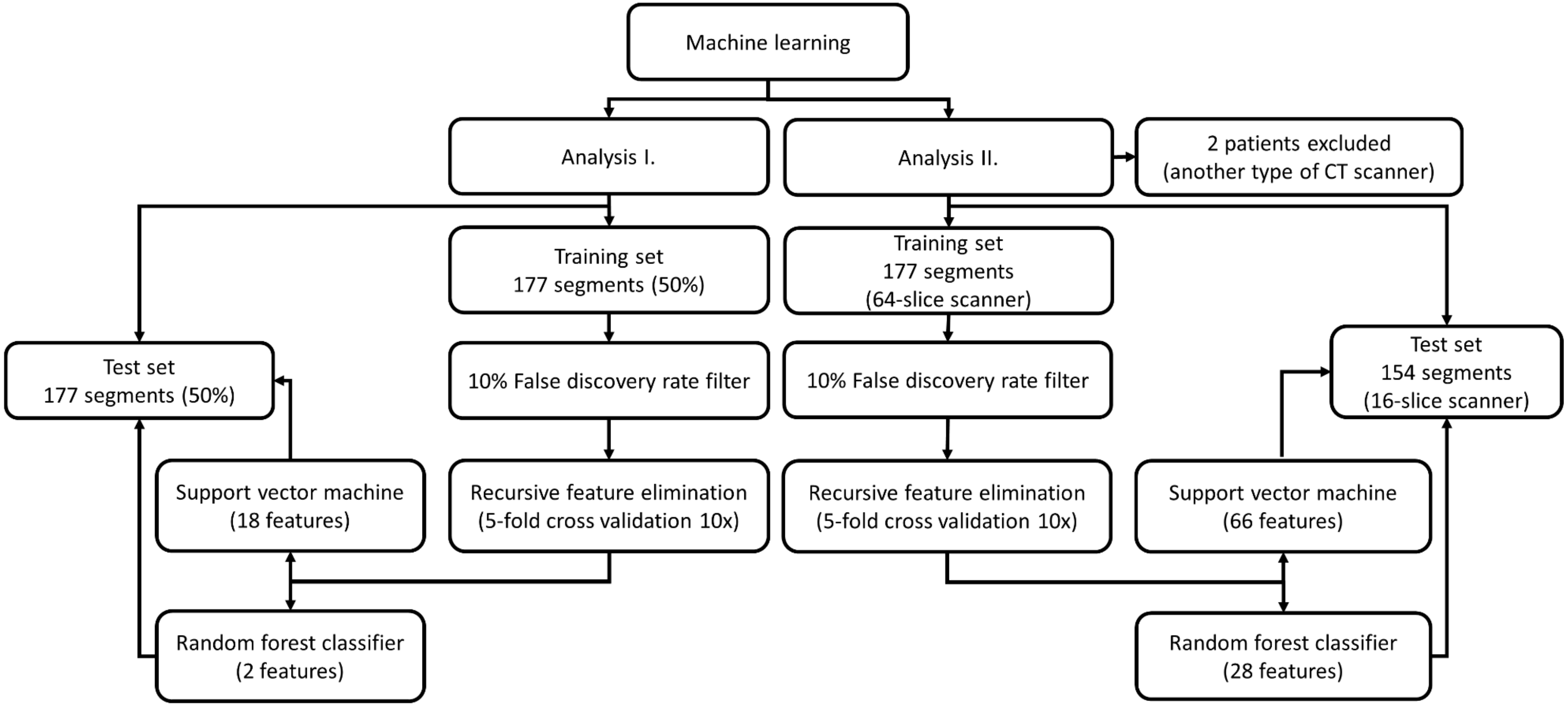


Figure 13. Steps of radiomics feature selection and machine learning model building for liver fibrosis staging. *During analysis I., the liver segments were randomly split into training and test datasets, while in analysis II., the train-test split was carried out based on the scanner types (82).*

3.1.1. Patient population

As US-based SWE measurements are accepted non-invasive clinical alternatives of liver biopsy for ruling out advanced liver fibrosis, in our institution, patients with suspected liver fibrosis are followed-up with SWE since 2017 for research and later for clinical purposes (83-85).

In this study, patients clinically diagnosed and followed for chronic liver disease were retrospectively collected who underwent point-SWE measurements due to suspected fibrosis and had CT scans at our Institution within six months of the SWE between September 2016 and January 2019. Due to the lack of portal venous phase contrast-enhanced CT series, two patients were excluded from the research.

3.1.2. Shear wave elastography examination protocol

The US examinations were completed with an RS85 Prestige ultrasound scanner equipped with a CA1-7A convex probe (Samsung Medison, Hongcheon, Korea) and the point-SWE measurements were performed with its S-Shearwave™ application. Prior to the examination, the patients had to fast for at least two hours. An intercostal view was used to examine the right lobe of the liver. The point-SWE measurement box was placed by the operator at least 1.5 cm from the liver capsule and not deeper than 6 cm from the skin's surface, close to the ultrasound beam's centerline. Measurements of liver stiffness were performed while holding the breath mid-expiration. The results of the median liver stiffness were given in kPa.

3.1.3. CT imaging protocol

The CT examinations were performed according to our routine abdominal diagnostic protocols on either a 16-slice Brilliance or a 64-slice Ingenuity Core 64 CT scanner (Philips Healthcare, Best, the Netherlands). The scanners had the following settings: a collimation of 16 × 1.5 mm, or 64 × 0.625 mm, and a spiral pitch of 0.813 or 0.798 for the two scanners, respectively; a tube voltage of 120 keV with an automatic tube current modulation and a rotation time of 0.5 s were used. The 16-slice scanners' acquisitions were routinely reconstructed with filtered back projection, and the 64-slice scanners' with the iDose4™ hybrid iterative reconstruction algorithm at the axial plane with a slice thickness of 5.0 mm.

For the post-contrast CT series, an ROI was positioned in the lumen of the descending aorta above the diaphragm in order to time the scan using the bolus tracking technique. Using a power injector, a non-ionic, iodinated contrast agent (concentration range: 350-370 mg/ml) was administrated intravenously with an injection rate of 2-3.5 ml/sec for the post-contrast CT series. The dosage (0.5 g iodine/kg) was adjusted for body weight. The portal venous phase scan was started after 60 seconds the ROI’s aortic enhancement had surpassed the 150 Hounsfield Unit (HU) threshold.

3.1.4. Image preprocessing, segmentation, and radiomics feature extraction

The portal venous phase CT scans and the US examinations of the included patients were retrieved from the institutional PACS system and were exported in DICOM format to a clinical workstation using the IntelliSpace Portal (Philips Healthcare, Best, the Netherlands). The DICOM files were then anonymized by using the dcm2nii package of the MRIcron software (86).

The image post-processing was completed with the 3DSlicer software v.4.8.1 (87). The entire liver volumes, as well as the right and left lobes, and the anatomical liver segments were manually segmented slice-by-slice on the portal venous phase axial series. The upper and lower segments of the left and right lobes were divided using the axial slice with the largest cross-section at the portal vein bifurcation, which gave an excellent approximation of the anatomical liver segments (Figure 14).

In total, 30 patients’ portal venous phase scans were manually segmented including the nine anatomic liver segments, the right and left lobes, and the whole liver volume. Six liver segments had to be excluded from the radiomics analysis because one patient had undergone surgical resection of the SVI segment, and two patients had substantial cysts in their SIII and SIVA- SIVB segments. As a result, the final dataset only included 354 liver segment volumes.

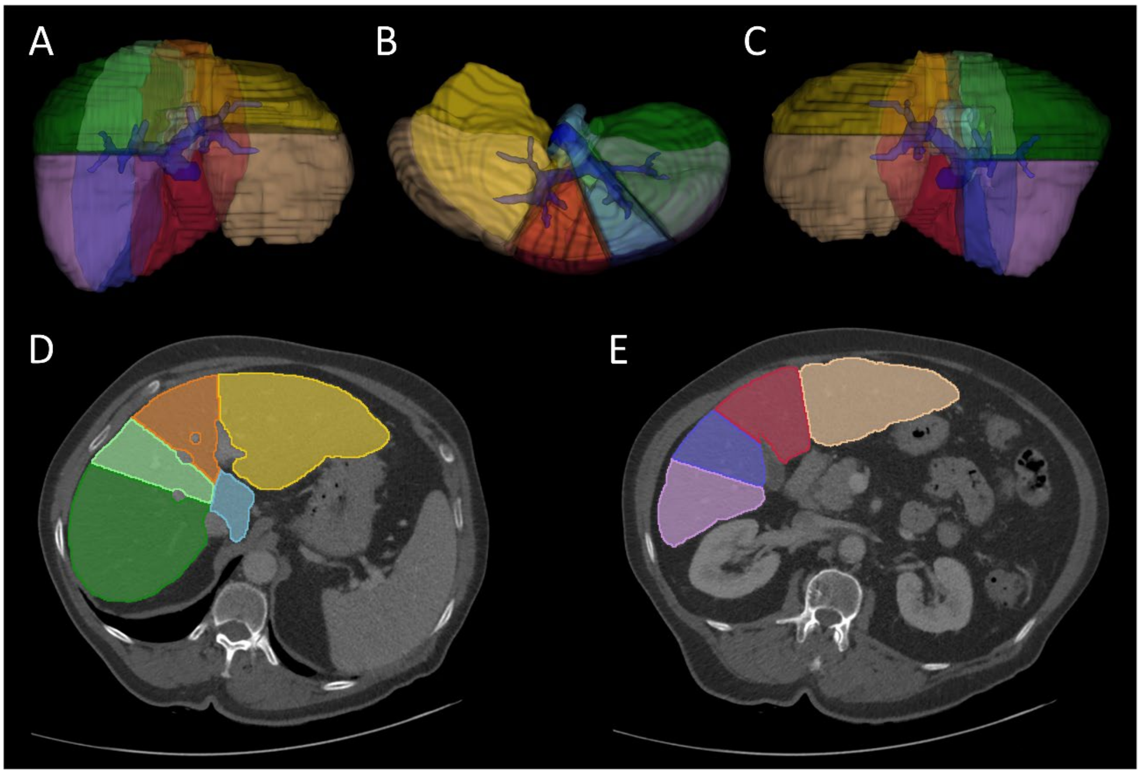


Figure .14. Manual segmentation of the liver and the anatomical liver segments. *The anterior (A), the inferior (B), and the posterior (C) view of the reconstructed segmentation mask of the liver. The upper (E) and the lower (D) Couinaud segments were manually delineated on the portal venous phase axial scans (82).*

During radiomics feature extraction, 13 shape-based features, 18 first-order intensity features, 23 GLCM-based, 16 GLRLM-based, 16 GLSZM-based, 14 GLDM-based, and 5 NGTDM-based texture features were calculated from each segment with the pyRadiomics package using a bin width value of 25 and an isotropic (1x1x1mm) voxel resampling (35). Similarly to the previous research of Kayaalti et al. (57) and Lubner et al. (59, 61), in addition to the analysis of original images, the radiomics features were also extracted after applying LoG filters, and coiflet1 wavelet filters on the scans in order to extract fine, medium and coarse texture features that may alter with fine and coarse textural changes in the liver parenchyma during fibrosis progression and cirrhotic nodule formation. 276 features were calculated with LoG kernel sizes of 3 mm, 6 mm, and 9 mm, and 736 features after discrete wavelet transformations. Thus, a total number of 1117 radiomics features were extracted for each patient.

To assess the association between liver fibrosis and the average density of the liver parenchyma as well as to reveal the correlation between radiomics features and differences in contrast-enhancement, the average CT density was measured using circular ROIs in all anatomic liver segments in the portal venous phase.

3.1.5. Data preprocessing and initial feature selection

The data preprocessing included transformation and dimension reduction steps. First, the radiomics features were log-transformed to deal with the skewness of the data. Then, to exclude those redundant features that show a high correlation with each other, Pearson's correlation coefficient values were calculated, a correlation matrix was computed, and the features with $r > 0.95$ were eliminated as recommended previously (37). The remaining 453 radiomics features were used for further analysis. As the variables contained outliers, the standardization of the features was carried out using the robust data scaling method, which uses the median and quantiles for rescaling and involves subtracting the median from all of the observations before dividing by the interquartile difference so that it is not sensitive to outliers.

3.1.6. Univariate analysis

Univariate linear regression was performed to identify predictors of liver stiffness measured by point-SWE. The 95% confidence interval (CI) values of the R^2 and β coefficients, as well as the p-values were calculated with bootstrapping using 1000 replications.

To assess the diagnostic performance of individual radiomics features in distinguishing between low-grade and high-grade fibrosis, receiver operating characteristic curve (ROC) analysis was performed. The area under the ROC curve (AUC) value and its 95% CI were calculated using 5-fold cross-validation.

Due to multiple comparisons, the significance level had to be adjusted. PCA was used to determine the effective number of measures through eigen value calculation as proposed by Gao et al. (88). Based on the number of principal components that explain the 95% of the total variance of the dataset, the effective number of measures was 45. Then, the adjusted statistical significance level was defined as 0.05/45, $p < 0.0011$.

3.1.7. Unsupervised classification

For the unsupervised classification of liver segments, non-hierarchical clustering namely k-means clustering was performed and combined with hierarchical cluster analysis. Cluster analysis, which aims to cluster data points based on their similarities, is one of the most popular data mining algorithms and the most commonly used unsupervised technique in radiomics studies (89). During k-means clustering, the optimal k-value (the optimal number of clusters) was determined with silhouette analysis as proposed by Rousseeuw (90). Then, hierarchical cluster analysis was carried out according to the k-means clusters using Pearson's correlation as a distance measure and the Ward.D as linkage criteria. The dendrograms and heatmaps were reconstructed and split by rows according to the results of k-means clustering. ANOVA with post hoc Tukey's tests were used for continuous variables to show differences between cases in various clusters, whereas F-test was applied for categorical data.

3.1.8. Machine learning

Before ML model building, a univariate false discovery rate (FDR) filter based-on Benjamini-Hochberg correction was applied to the training dataset with an upper bound of 10% (p-value: 0.1). Then, the optimal number of features was determined with an RFE algorithm by using ten times repeated stratified 5-fold cross-validation. RFE is a popular wrapper algorithm for feature selection (91) which has the advantage to be used with either linear or non-linear machine learning algorithms, those based on the calculation of feature importance scores e.g. RFC or SVC.

In this study, we trained and tested both linear and non-linear ML classification algorithms to distinguish between high-grade and low-grade liver fibrosis. We used the linear SVC and evaluated its performance against that of an RFC from among non-linear algorithms. These classifiers are among the most used ML algorithms in radiomics studies (37, 92) and proved their applicability in the case of various organs (93).

After defining the best set of radiomics features, the classification models' hyperparameters were optimized with a grid-search algorithm by using ten times repeated stratified 5-fold cross-validation.

3.2. Radiomics analysis for differentiating renal cell carcinoma subtypes

3.2.1. Patient population

Patients who underwent either radical or partial nephrectomy between January 2008 and May 2021 at the Department of Urology of our University were retrospectively collected. The preoperative CT scans of patients were retrieved from the PACS of our Medical Imaging Centre. The pathology reports were retrieved from the hospital information system. Out of the 551 patients who had available preoperative CT scans, 107 were excluded because were diagnosed with other types of malignant kidney tumors instead of ccRCC, pRCC, or chRCC, 75 were excluded because the nephrectomy was completed due to other reason than a tumor, 61 were excluded because CT scan including both unenhanced, corticomedullary and excretory phase series was not available, 44 were excluded because of damaged files, 33 were excluded because were diagnosed with benign kidney tumors, 30 were excluded because the histopathologic report was not available, 2 cases were excluded because of preoperative radiofrequency ablation, and 1 case was excluded because of incomplete coverage of the tumor. Therefore, the final patient cohort consisted of 209 patients with 212 tumors including 161 ccRCCs, 34 pRCCs, and 17 chRCCs.

3.2.2. CT imaging protocol

We performed the CT scans with either a 16-slice or a 64-slice CT scanner. The tube voltage was set between 100-140 keV, and by using an automatic tube current modulation, the tube current varied between 80-977 mAs. The reconstructed slice thicknesses varied between 1.25-5.0 mm. For the post-contrast series, a non-ionic iodinated contrast agent was injected at a flow rate of 1.5-3.5 ml/s (volume was adjusted for body weight).

3.2.3. External test dataset

For external validation, cases from the 2019 Kidney and Kidney Tumor Segmentation Challenge (KiTS19) were used therefore, the results shown here are in part based upon data from the C4KC-KiTS dataset of The Cancer Imaging Archive (94-96). This publicly available dataset includes the CT scans, the corresponding segmentation masks as well as the detailed clinical data of 210 cases. Only those cases were included in our external test set that had both unenhanced. corticomedullary and excretory phase

CT series. No cases had available nephrographic phase CT scans, therefore that post-contrast phase could not be included in this study. In total, 69 eligible cases were identified including 50 ccRCCs, 13 pRCCs, and 10 chRCCs. The CT scans were performed by 19 different CT models of 4 different vendors. The slice thickness in this dataset varied between 1.0-7.0 mm, the tube voltage was between 100-140 keV, and the tube current varied between 80-747 mAs.

3.2.4. Image preprocessing, segmentation, and radiomics feature extraction

Figure 15 shows the steps of data analysis. The preoperative CT scans of the included patients were retrieved from the institutional PACS system and were exported in DICOM format to a clinical workstation. The DICOM files were then anonymized by using an in-house developed anonymizer script coded in Python computer language.

The unenhanced and the excretory phase series were co-registered to the corticomedullary phase using the Elastix extension of the 3D Slicer software developed by Kline et al. for medical image co-registration (18). For computing the transformation matrix of the given scan, the tumor was identified, and a cubic VOI was defined around the tumor. The tumor volumes were then cropped, and the moving volumes (unenhanced and excretory phase) were co-registered to the fixed volume (corticomedullary phase). After the image transformation matrix was calculated, it was applied to the original whole volumes of CT scans.

The tumor volumes were segmented manually on the corticomedullary phase axial CT scans. Before radiomics analysis, the voxel density values of the corticomedullary phase CT scans were normalized to the mean density of the healthy renal cortex by shifting the individual voxel intensity values with the average measured cortical density to reduce the impact of differences in hemodynamics, variations in scanners and institutional imaging protocols (Figure 16). The image processing and segmentation steps were completed with the 3D Slicer software v.4.10.2 (87).

The radiomics analysis was performed with the pyRadiomics package (35). During radiomics feature extraction, 1x1x1mm isotropic resampling and a fixed bin width of 16 were used for calculation. 107 radiomics features were extracted from all three CT series including first-order statistical parameters, shape-based features, and texture features (Figure 17). No higher-order features were calculated in this study.

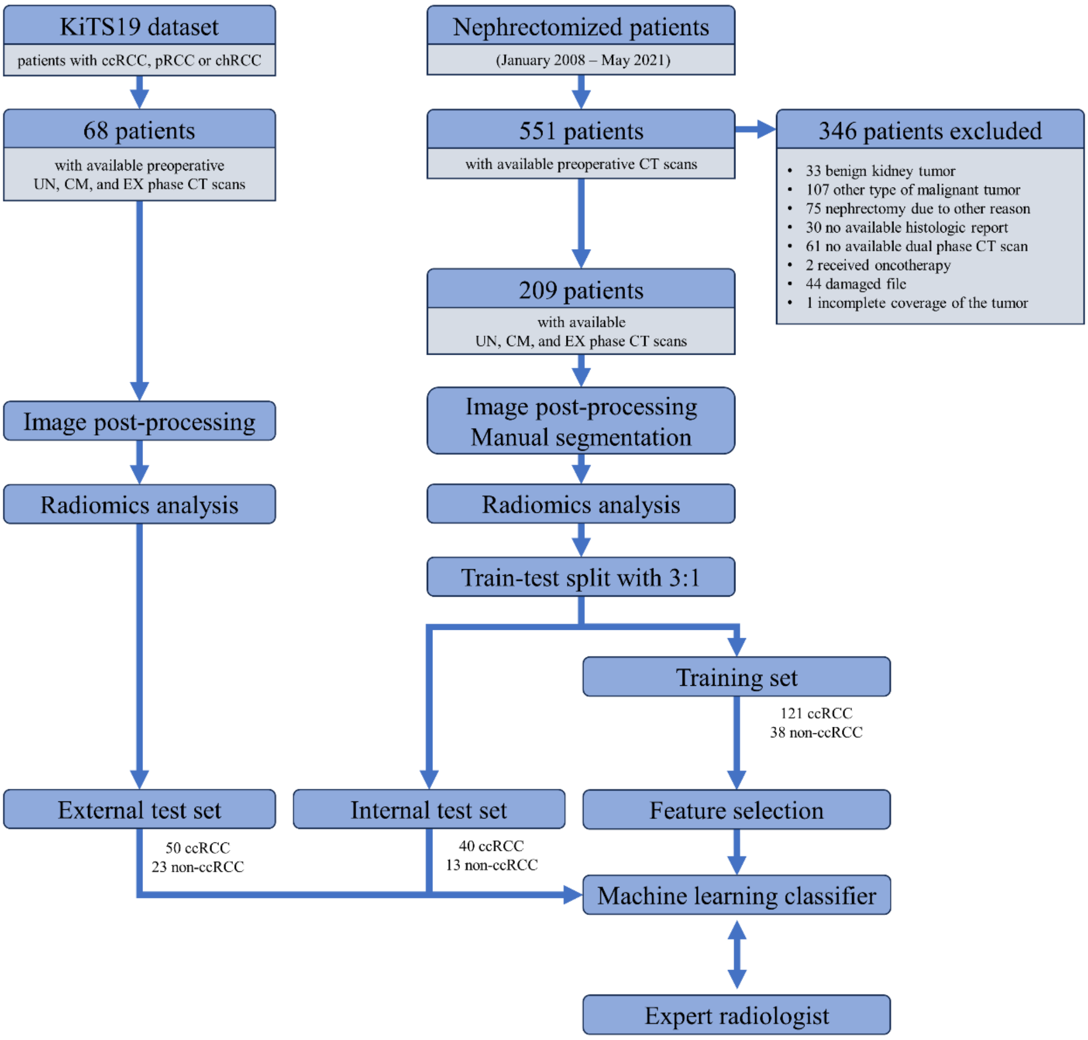


Figure 15. The study design of radiomics analysis of renal cell carcinomas (70).

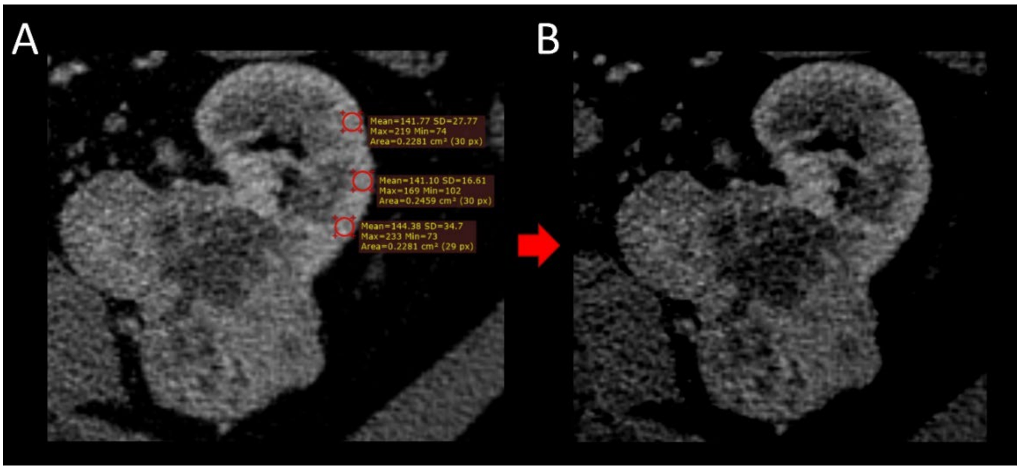


Figure 16. Normalization of the voxel intensity values to the density of the healthy renal cortex on the corticomedullary phase post-contrast CT scans. *(Own source (97))*

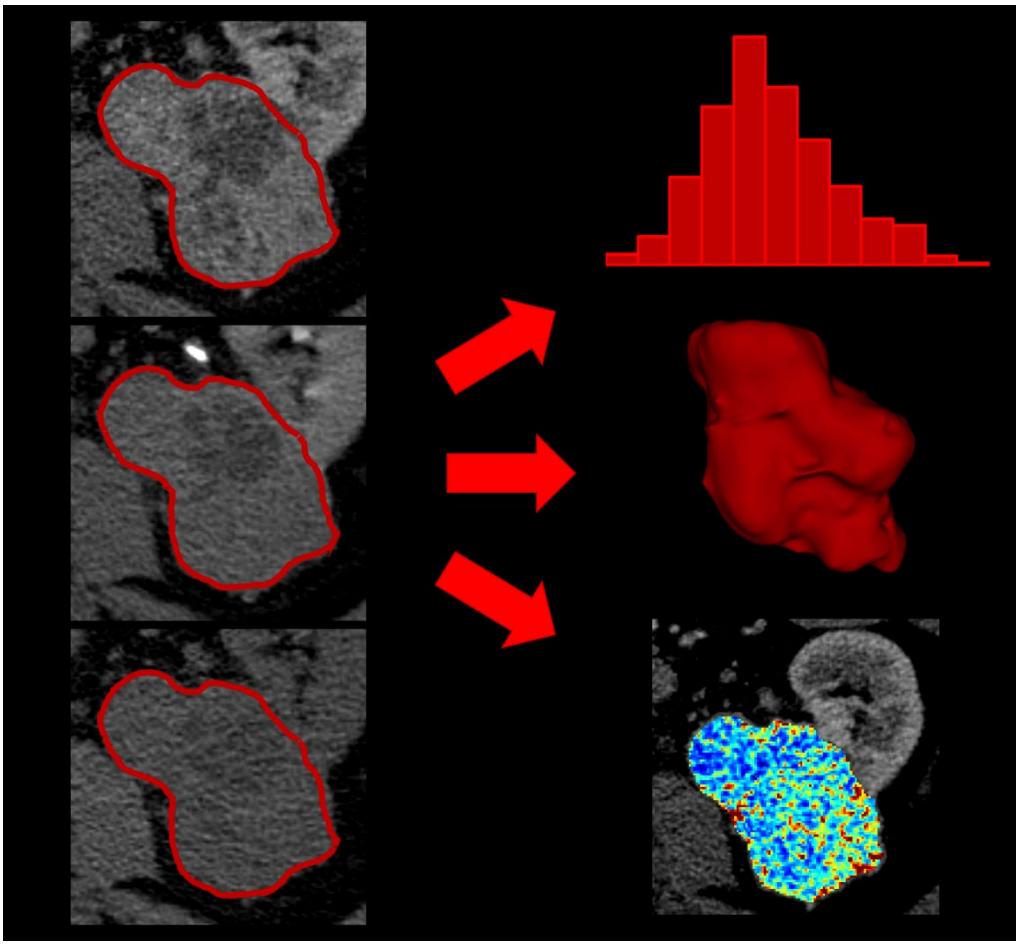


Figure 17. Radiomics feature extraction from the unenhanced, corticomedullary, and excretory phase preoperative CT scans of a renal cell carcinoma. *(Own source (98))*

3.2.5. Feature selection and machine learning model building

The cases were randomly divided into training and internal test datasets with a 3:1 ratio. The feature selection steps were completed on the training set. The features were first standardized, then Pearson's correlation coefficient was used for excluding highly correlated features ($r > 0.95$). The features with poor reproducibility were excluded based on ICC values ($ICC<0.90$). As previously suggested (99, 100), the area of the segmented tumor masks was eroded by 1-1 voxel in each direction, and the extraction of the radiomics feature was repeated. The final feature selection step and the ML model building consisted of either a linear or a non-linear model pipeline. The linear model pipeline included the LASSO as a feature selector and the SVC as a binary classifier, while the non-linear model pipeline included a tuned ReliefF wrapper as a feature selector and an RFC as a binary classifier. The hyperparameters of the classifiers were optimized with grid search using five-times repeated 5-fold stratified cross-validation. The class imbalance problem was handled by using balanced class weights during model fitting. ROC analysis was used to evaluate the diagnostic performance of the ML models on both the training set, the internal test set, and the external test set. The AUC, accuracy, sensitivity, and specificity values were calculated.

The performance of the ML models was compared to that of an expert radiologist with over 10 years of experience in urologic imaging. The specialist classified the tumors of the internal and external test datasets blinded to the patient's history, medical records, and to the results of tumor segmentation.

3.2.6. Statistical analysis

Continuous variables were compared between the two tumor groups with the Mann-Whitney U-test and categorical variables with the chi-squared test. DeLong's method was used for calculating the 95% CIs of the AUC values. The optimal cut-off point of the ROC curve was determined based on the "closest top left" method. The ROC curves were directly compared with the DeLong test. The ICC was calculated for each radiomics feature based on a 2-way, single-rater, absolute agreement model. The threshold of $p < 0.05$ was applied to determine the significance in all comparisons.

4. RESULTS

4.1. Radiomics analysis for liver fibrosis staging

According to the inclusion and exclusion criteria, we retrospectively identified sixteen female (ages 22–72, mean age of 52 years) and sixteen male patients (ages 42–75, mean age of 63 years) who had both SWE measurements and portal venous phase contrast-enhanced CT scans at our institution.

Based on the liver stiffness measurements of the point-SWE, the patients were divided into low-grade and high-grade fibrosis groups. The low-grade group included those 11 patients that had point-SWE < 9.5 kPa indicating F0, F1, and F2 METAVIR fibrosis stages, meanwhile the 21 patients with point-SWE ≥9.5 kPa including F3 and F4 METAVIR fibrosis stages were assigned to the high-grade fibrosis group (84). The various etiologies and the clinical characteristics of the included chronic liver disease patients are summarized in Table 1.

Table 1. Characteristics of the chronic liver disease patient cohort (82).

Sex	**Number**	**Age (range)**
Female	16	52 years (22–72)
Male	16	63 years (42–75)
Etiology of liver fibrosis	**Number**	**Ratio (%)**
chronic HCV	14	14/32 (43.8%)
toxic hepatitis	7	7/32 (21.9%)
PBC, PSC, AIH	3	3/32 (9.4%)
chronic HBV	1	1/32 (3.1%)
Unknown	7	7/32 (21.9%)
Fibrosis stage	**Number**	**Ratio (%)**
Low-grade (< 9.5 kPa)	11	12/32 (34.4%)
High-grade (≥9.5 kPa)	21	20/32 (65.6%)

AIH: autoimmune hepatitis; HBV: hepatitis B virus; HCV: hepatitis C virus; PBC: primary biliary cirrhosis; PSC: primary sclerosing cholangitis

4.1.1. Univariate analysis

After Pearson's correlation filter was applied to the total number of 1117 extracted radiomics features, 453 variables remained.

Univariate linear regression analysis of the radiomics features revealed 38 significant predictors, however, the R^2 values of these features suggested a weak association. Table 2. demonstrates the top 10, significant ($p<0.0001$) radiomics features ranked by their R^2 values. The best predictor of liver stiffness was wavelet-HLH filtered GLCM Imc2 with an R^2 of 0.074 and a β-coefficient of − 2.59 ($p < 0.0001$).

Table 2. The top 10 radiomics features in univariate linear regression analysis for liver stiffness prediction (82).

Feature name	Filter type	AUC	R^2	β
GLCM-Imc2	wavelet-HLH	0.71 [0.65-0.77]	0.07	-2.59
GLCM-Imc1	wavelet-HLH	0.65 [0.60-0.71]	0.07	2.08
GLCM-Imc2	wavelet-LLH	0.68 [0.62-0.74]	0.06	-2.01
GLCM-Imc1	LoG-sigma-9-0-mm-3D	0.66 [0.61-0.72]	0.05	-1.91
GLSZM-ZonePercentage	LoG-sigma-9-0-mm-3D	0.60 [0.54-0.67]	0.05	-1.06
GLCM-Correlation	wavelet-LLH	0.72 [0.66-0.78]	0.05	2.09
GLCM-Imc1	wavelet-LLH	0.57 [0.50-0.63]	0.05	1.56
NGTDM-Busyness	wavelet-HLH	0.63 [0.57-0.69]	0.05	2.17
GLSZM-ZoneVariance	LoG-sigma-9-0-mm-3D	0.57 [0.50-0.63]	0.04	2.03
GLSZM-ZoneEntropy	original	0.62 [0.56-0.69]	0.04	1.77

AUC: area under the receiver operating characteristic curve value; GLCM: gray-level co-occurrence matrix; GLSZM: gray-level run-length matrix; NGTDM: neighboring gray-tone difference matrix

The diagnostic performance of the individual radiomics features was assessed by their cross-validated AUC values (Figure 18.). GLCM-based texture features achieved the highest diagnostic ability in identifying high-grade liver fibrosis, while the shape-based geometric features showed the worst performance.

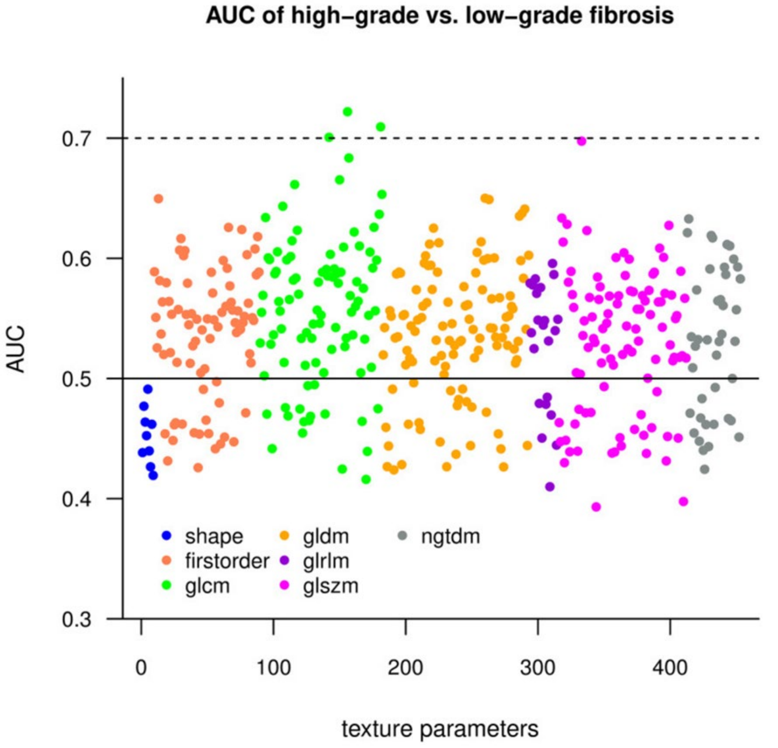


Figure 18. The diagnostic performance of the individual radiomics features (82).
AUC: area under the receiver operating characteristic curve value; GLCM: gray-level co-occurrence matrix; GLRLM: gray-level run-length matrix; GLSZM: gray-level run-length matrix; GLDM: gray-level dependence matrix; NGTDM: neighboring gray-tone difference matrix

The top three radiomics features included higher-order texture features. Wavelet-LLH filtered GLCM Correlation reached the highest diagnostic ability with an AUC of 0.72 [95% CI: 0.66–0.78], wavelet-HLH filtered GLCM Informational Measured Correlation resulted in an AUC of 0.71 [0.65–0.77], while wavelet-LHL filtered GLCM Correlation showed an AUC of 0.70 [0.64–0.76]. As all of these texture features quantify the complexity of the tissue texture by quantitatively describing the correlation between the probability distributions of the GLCM elements, they may be linked to the formation of regenerative nodules during the progression of liver fibrosis.

4.1.2. Unsupervised classification

The optimal value of *k* was found to be 4 during the silhouette analysis of the k-means clustering. The resulting k-means clusters contained 213 segments, 57 segments, 39 segments, and 45 segments, respectively (Figure 19.). The four clusters were compared on the basis of anatomic liver segments, scanner type, liver stiffness, and average CT density value. There were no significant differences in the distribution of either CT scanner type, anatomical liver segments, or measured liver stiffness values, however, a significantly higher ($p < 0.0001$) mean CT density (110HU ± 10.1HU) was found in the fourth cluster compared to the first (96.1HU ± 11.3HU), second (90.8HU ± 16.8HU), and third (93.1HU ± 17.5HU) clusters.

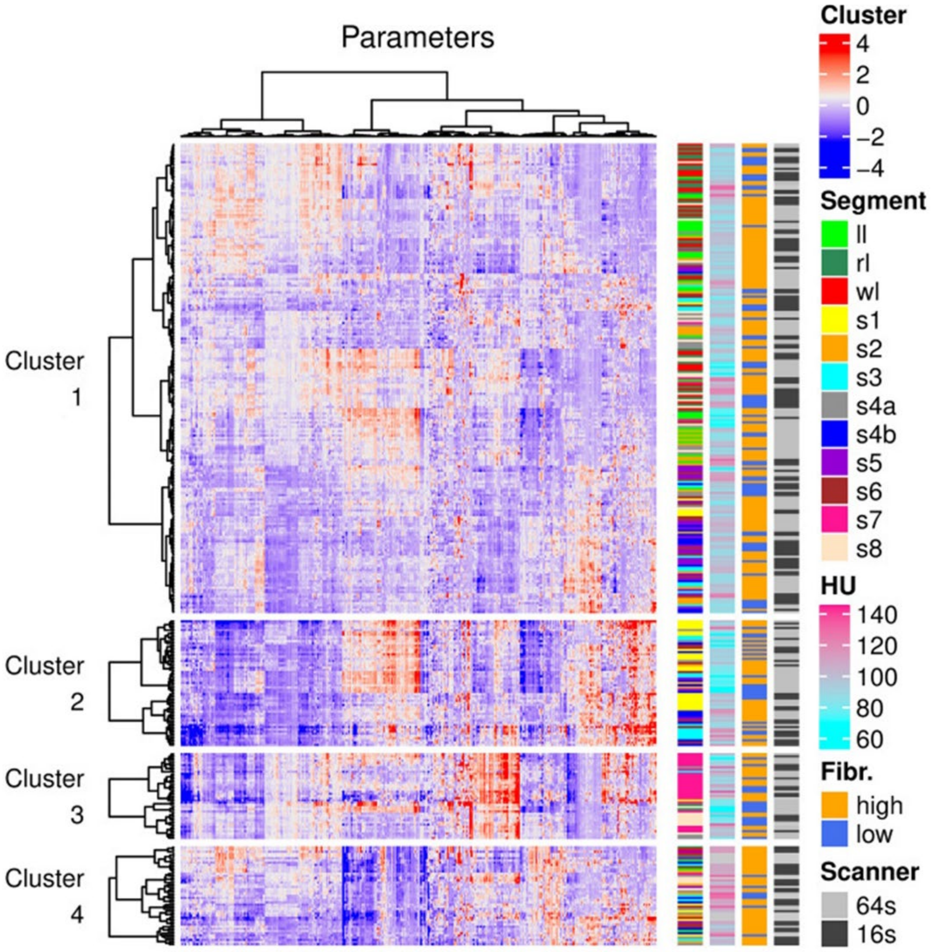


Figure 19. Hierarchical cluster analysis (82). *The clusters showed no association with either the scanner type, fibrosis grade, or liver segments. However, a significantly higher the average CT density value was found the fourth cluster compared to the other three.*

4.1.3. Machine learning

During the first analysis, the liver segments were randomly split into training and test datasets with a 1:1 ratio (177:177 segments) while also setting the ratio of the cirrhotic and non-cirrhotic segments equal in the two datasets (non-cirrhotic: 27% vs. 33%). After applying the FDR filter, the RFE resulted in 2 and 18 radiomics features for the RFC and SVC models, respectively.

The SVC and RFC models trained to discriminate between low-grade and high-grade fibrosis liver segments based on selected features resulted in excellent diagnostic performance on the training set during cross-validation with an AUC of 0.95 [95% CI: 0.91-0.98] vs. AUC of 0.88 [95% CI: 0.81-0.94]. When evaluated on the independent test dataset, the RFC model demonstrated robustness and achieved excellent diagnostic ability with an AUC of 0.90 [95% CI: 0.85-0.95], a sensitivity of 0.86, and a specificity of 0.78, while the SVC showed signs of overfitting as its performance declined to an AUC of 0.76 [95% CI: 0.67-0.84], a sensitivity of 0.93, and a specificity of 0.31 (Figure 20).

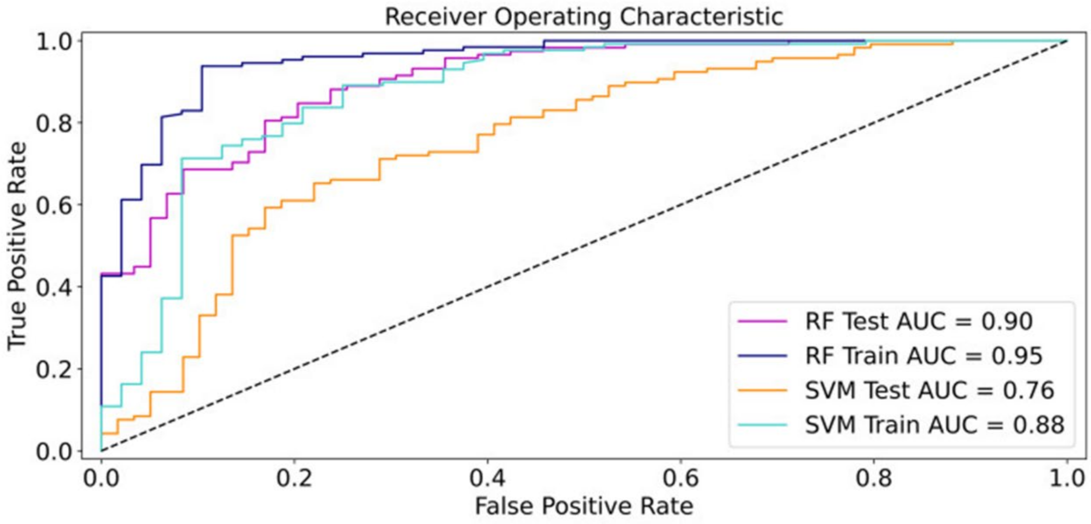


Figure 20. Receiver operating characteristic curve of the machine learning models in the first analysis (82).

RF: random forest classifier; SVM: support vector machine classifier; AUC: area under the receiver operating characteristic curve value

In the second analysis, the liver segments were divided according to the type of CT scanner used: 177 segments from 15 patients scanned with a 64-slice scanner made up the training set, while 154 segments from 13 patients scanned with a 16-slice scanner were included in the test set. Two patients who were scanned with another type of scanner were excluded from this part of the analysis.

After applying the FDR filter, 28 and 66 radiomics features were used for training the RFC and SVC models, respectively. Both the RFC and the SVC models in this second analysis resulted in robust diagnostic abilities. The SVC overperformed the RFC model with an AUC of 0.91 [95% CI: 0.88-0.94] on the training set and a similarly excellent performance on the independent test set with an AUC of 0.90 [95% CI: 0.87-0.93], a sensitivity of 0.83 and a specificity of 0.95. Meanwhile, the RFC model achieved an AUC of 0.84 [95% CI: 0.83-0.85] on the training set and an AUC of 0.88 [95% CI: 0.84-0.91], a sensitivity of 0.86, and a specificity of 0.92 on the test dataset (Figure 21).

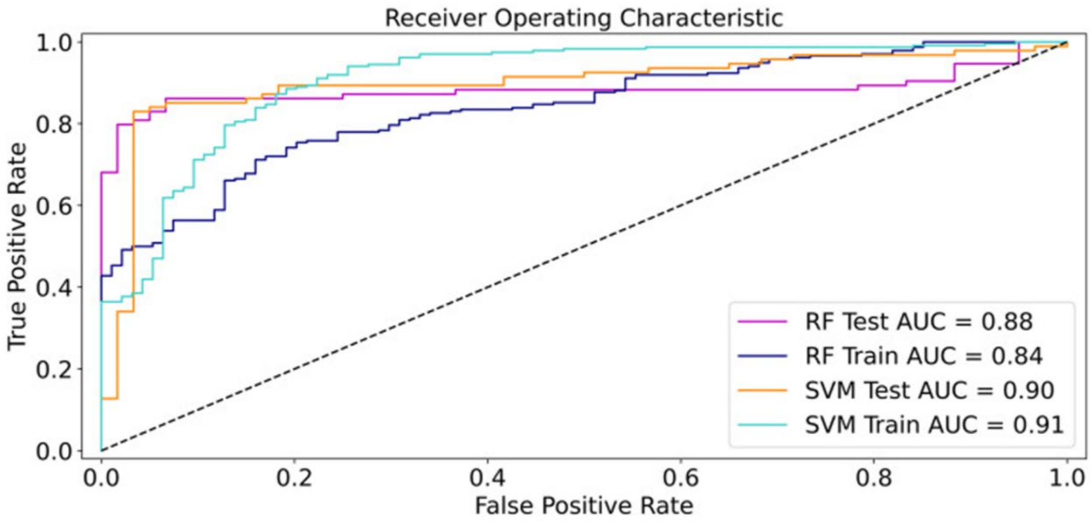


Figure 21. Receiver operating characteristic curve of the machine learning models in the first analysis (82).

RF: random forest classifier; SVM: support vector machine classifier; AUC: area under the receiver operating characteristic curve value

4.2. Radiomics analysis for differentiating renal cell carcinoma subtypes

As the ccRCC subtype has a drastically worse prognosis than pRCC and chRCC, recent European practice guidelines distinguish between two main groups of ccRCC vs. non-ccRCC (72). Therefore, in our study, we also assigned the cases to ccRCC and non-ccRCC groups. The final patient cohort included 212 tumors of 209 cases, 161 ccRCC, and 51 non-ccRCC.

No significant differences were found between ccRCC vs. non-ccRCC groups regarding age and sex distribution in either our own dataset ($p=0.079$; $p=0.978$) or the external test set ($p=0.556$; $p=0.069$). The training set included 121 ccRCCs and 38 non-ccRCCs, the internal test set consisted of 40 ccRCCs and 13 non-ccRCCs, meanwhile, the external test set had 73 tumors of 68 cases, 50 ccRCCs, and 23 non-ccRCCs.

In the first step, the filter-based feature selection resulted in 39 corticomedullary phase, 38 excretory phase, and 35 unenhanced phase radiomics features, in the second step, the LASSO feature selection resulted in 9 unenhanced, 10 corticomedullary, and 5 excretory phase radiomics features, each set included both first-order statistical features, shape-based features, and texture features.

The RFC achieved the best results with the corticomedullary phase features but showed clear signs of overfitting. It achieved an AUC of 1.0, with a sensitivity of 1.0 and a specificity of 1.0 on the training set and an AUC of 0.874 on the internal test set but failed to predict tumor types on the external test set, where it yielded an AUC of 0.663.

Meanwhile, the optimized SVC trained on the 10 corticomedullary features proved to be the most stable, achieving the highest diagnostic performance with AUC values of 0.951, 0.873, and 0.834 and accuracies of 0.925, 0.811, and 0.781 on the training set, internal test set and external test set at a "closest top left" method-based calculated optimal threshold of 0.655, respectively, which proved to be non-inferior compared to the performance of the expert radiologist on either the internal (accuracy of 0.906 vs. 0.811) or the external test set (0.795 vs. 0.781) ($p=0.866$; $p=0.256$) (Figure 22, Table 3). It is important to highlight that the radiologist also had a lower accuracy for external cases compared to the internal test set (accuracy of 0.795 vs. 0.906).

Table 3. The performance of the machine learning models on the test sets (81).

	AUC	Accuracy	Sensitivity	Specificity
SVC - Internal	0.873 [0.774–0.972]	0.811	0.900	0.539
SVC - External	0.834 [0.730–0.938]	0.781	0.800	0.739
RFC - Internal	0.874 [0.755–0.993]	0.868	0.950	0.615
RFC - External	0.663 [0.529–0.796]	0.685	0.900	0.217
Expert – Internal	0.886 [0.776–0.996]	0.906	0.925	0.846
Expert – External	0.768 [0.659–0.877]	0.795	0.840	0.696

AUC: area under the curve; SVC: support vector classifier; RFC: random forest classifier

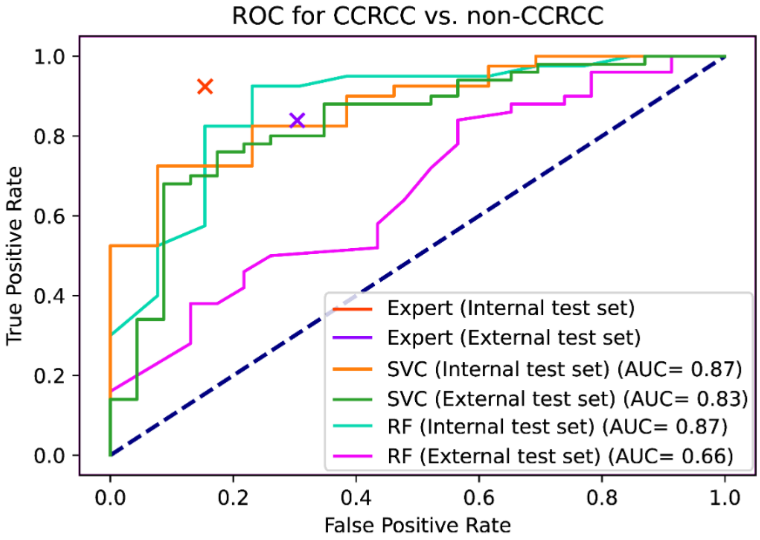


Figure 22. Results of receiver operating characteristic curve analysis (81).

AUC: area under the receiver operating characteristic curve; RFC: random forest classifier; ROC: receiver operating characteristic curve; SVC: support vector classifier

The corticomedullary phase radiomics features proved to be the most valuable, the SVC trained on the radiomics features of the excretory phase showed worse performance with AUC values of 0.719 on the internal test set and 0.64 on the external test set. As expected, the SVC trained on unenhanced phase features performed even worse on both internal and external test sets with an AUC of 0.725 and 0.598, respectively. Furthermore, the addition of unenhanced and excretory phase features did not improve the predictive performance of the corticomedullary SVC model, and in fact increased the overfitting, as the diagnostic performance of the combined SVC model was AUCs of 0.862 vs. 0.711 on the internal and external test sets.

5. DISCUSSION

5.1. Radiomics analysis for liver fibrosis staging

Early detection of liver fibrosis is essential, as it can cause life-threatening complications due to an increased risk of cirrhosis, liver failure, and liver cancer as chronic liver disease progresses. Although the current gold standard method for liver fibrosis staging is percutaneous liver biopsy, its well-known limitation is the sampling variability due to the heterogeneous distribution of fibrotic lesions in the liver parenchyma (52). Therefore, current clinical practice guidelines highlight the importance of developing non-invasive alternative approaches and recommend the usage of US-based elastography for the assessment of advanced-stage liver fibrosis in patients with HBV and HCV infection (54, 55).

In recent years, SWE has become the non-invasive reference standard for the assessment of liver fibrosis severity, our study is based on a direct comparison of SWE and radiomics analysis of CT scans. Our aim was to prove that radiomics analysis can be applied retrospectively to CT scans that were carried out due to other indications and to identify imaging biomarkers that can be used to train ML algorithms for identifying high-risk patients with advanced liver fibrosis. Our goal was to demonstrate that, while radiomics-based ML models have the advantage of the retrospective application, thus our proposed method does not require specific instrumentation and additional imaging studies, it can achieve the accuracy of SWE. Although SWE assesses the stiffness of the parenchyma while CT texture analysis measures architectural distortion, we proved that radiomics analysis accurately predicts the fibrosis stage identified by SWE.

Previous studies assessing the utility of CT texture analysis in the diagnostics of liver fibrosis reported promising results (57-61). However, in these studies, radiomics analysis was applied to either selected axial slices or small ROIs of the liver. On the contrary, we performed a three-dimensional CT texture analysis assessing the entire volume of the liver that can give a more complex insight into the architectural changes of the liver while revealing novel quantitative biomarkers of liver fibrosis. Zhang et al. (58) demonstrated that portal venous phase radiomics features are more effective compared to those extracted from the unenhanced and arterial phase, therefore we decided to perform radiomics analysis on the portal venous phase post-contrast CT series.

Daginawala et al. (60) were among the first, who investigated the role of radiomics analysis in liver fibrosis staging, the authors identified radiomics features extracted from portal venous phase thin slice CT scans that were significantly different between liver fibrosis stages, however when assessing the diagnostic performance of these individual features with ROC curve analysis, they found only fair performance with AUCs 0.50-0.76 that is comparable with our findings during the univariate analysis of radiomics features. This study however did not investigate the potential use of ML model construction and performed radiomics analysis on only 5 consecutive slices at the level of the porta hepatis instead of covering the entire liver volume. In our study, the univariate analysis identified wavelet-LLH filtered GLCM Correlation, wavelet-HLH filtered GLCM Imc2, and wavelet-LHL filtered GLCM Correlation as the best predictors of liver stiffness with performances ranging between AUCs of 0.71-0.72. These texture features describe the correlation between the intensity value of the central voxel and the intensity value of its neighboring voxels within a given VOI so that their increased value indicates a repetitive texture pattern, which in the case of liver parenchyma may indicate the formation of cirrhotic nodules.

Previous studies by Lubner et al. (59, 61) reported that first-order statistical features such as mean gray-level intensity and entropy show a strong association with the severity of liver fibrosis, which is confirmed by the results of our research, as we found that wavelet-LHL filtered first-order Median, original first-order 10Percentile and original first-order 90Percentile achieved high feature importance values during RFE indicating that these intensity-based first-order statistics are highly useful for the correct classification of low-grade vs. high-grade fibrosis. This can be explained by that the increased extracellular compartment and collagen deposition during fibrosis results in increased attenuation and tissue heterogeneity.

However, it is important to highlight, that the average CT density values of the liver parenchyma measured by circular ROIs in our study were significantly different in hierarchical cluster analysis, while no differences could be detected between the clusters in terms of fibrosis stage. Therefore, the changes in tissue density values can only partially be explained by disrupted blood supply due to architectural distortion of the liver parenchyma caused by the deposition of collagen fibers in the extracellular matrix, and it suggests that the voxel intensity values are also dependent on individual differences in contrast-enhancement that can be explained with other internal and external confounding factors such as hemodynamics and differences in contrast-injection protocols. Which underlines the sensitivity of radiomics features to these factors.

In our study, we not just assessed the diagnostic performance of individual radiomics features, but also investigated the added value of multivariable analysis by constructing ML models from the best-performing features. Our RFC model achieved excellent performance on both the training and test set with AUCs of 0.84-0.95 vs. 0.88-0.90, respectively. Similar results were reported in previous studies that also investigated the role of multivariable models in the diagnosis of advanced liver fibrosis using radiomic features extracted from multiple 16- and 64-slice CT scans. Lubner et al. (61) constructed a multivariate logistic regression model from first-order radiomics features for identifying cases with advanced-stage liver fibrosis and reported an AUC of 0.86, and in the later study of the research group (62), the multivariate logistic regression model consisting of CT-based parameters achieved an AUC of 0.94. These results are comparable with the performance of our RFC model that achieved AUCs of 0.88-0.90 on the independent test set, however in these previous studies, the results were reported solely on the training set, and the models were not evaluated on independent test cases.

The diagnostic performance of our SVC models on the test set is also comparable to those SVC models published by Kayaalti et al. (57) and Zhang et al. (58). By extracting radiomics features from cubic ROIs of the liver parenchyma, Kayaalti et al. constructed an SVC-based ML model from the best 5 features that was able to differentiate between stages of liver fibrosis with an average accuracy of 94% during 2-fold cross-validation. Meanwhile, Zhang et al. constructed an SVC model from the best 15 features for differentiating between cirrhosis vs. non-cirrhosis and reported an accuracy of 66.83%.

In contrast to previously published studies, we not only performed a three-dimensional texture analysis by segmenting the entire volume of the liver and its anatomic segments, but in our study, before feature selection and model building, we divided the cases into training and test data sets and thus evaluated the performance of the models on independent test cases to prove the models' robustness. We tested two data-splitting strategies that resulted in similar classification accuracy in detecting advanced-stage fibrosis. In the first analysis, each liver segment was considered as an independent sample and was randomly allocated between the training and test sets, while in the second analysis, the cases were split into training and test sets according to the scanner type.

In the first analysis, the RFE was shown to be an effective method for selecting the best-performing features, as the optimized RFC model trained by only two features achieved excellent performance on both the training and test cases (AUC of 0.95 vs. 0.90) without showing signs of overfitting. However, the linear SVC constructed from the best 18 radiomics features showed overfitting as its performance dropped from 0.88 to 0.76.

The advantage of our study is that we analyzed CT scans performed with different types of CT scanners, allowing our second analysis to split the cases into training and test data sets according to the scanner type. This time, the classification performance showed no decrease between the training and test sets, the RFC yielded AUCs of 0.84 vs. 0.88, while the SVC achieved AUCs of 0.91 vs. 0.90, respectively, indicating that there is no relevant overfitting, thus we were able to demonstrate that our proposed final ML models are robust to scanner types, contrast injection protocols and reconstruction algorithms, proving their generalizability and the suitability of our image preprocessing, data preprocessing, and feature selection steps.

However, our research had some limitations that need to be addressed. Our study is limited by the small number of patients and the mixed etiology of chronic liver disease. The individual liver segments were handled as independent samples as a kind of data augmentation, however during the train-test split, care was taken to ensure that segments belonging to the same patient were assigned to the same dataset. Furthermore, as our study was a retrospective, single-center study, the proposed ML models need to be evaluated on external cases from other institutions in a future multi-center study.

5.2. Radiomics analysis for differentiating renal cell carcinoma subtypes

In our second study, we constructed a radiomics-based ML model for differentiating between ccRCC and non-ccRCC subtypes based on preoperative contrast-enhanced CT scans. Our study aimed to prove the generalizability of the constructed ML models; therefore, those were validated on both independent test cases from our own Institution and on external test cases from the KiTS19 public dataset. Our proposed SVC model built from the corticomedullary phase radiomics features demonstrated its robustness and reproducibility by achieving an AUC of 0.87 on independent internal test cases, and an AUC of 0.83 during external validation. To prove that the model can help radiologists in daily practice, we also compared the performance of the ML model to that of a radiologist expert, which revealed that it performs comparably (accuracy of 0.79 vs. 0.78 on the external dataset), supporting the existing literature and highlighting the usefulness of radiomics analysis in this application.

Only a few studies have attempted to differentiate between RCC subtypes (75-79) while the majority of previously published studies have concentrated on differentiating benign from malignant renal lesions (101-103) or on identifying aggressive tumor characteristics of ccRCC (104-110). Moreover, these previous studies on radiomics-based classification of RCC subtypes had a single-center study design and their models had not been effectively validated on independent external cases from a huge number of institutions.

Yu et al. were among the first who assessed the role of radiomics analysis in differentiating RCC subtypes (78). The authors assessed the diagnostic performance of the individual radiomics features extracted from the nephrographic phase post-contrast CT scans by constructing a linear SVC model for differentiating between RCC subtypes and oncocytomas. Based on the 5-fold cross-validation results, the highest performing radiomics feature was the geometric mean of voxel gray-level values that reached an AUC of 0.809 in differentiating ccRCCs from pRCCs, chRCCs, and oncocytomas. Meanwhile, in differentiating pRCCs from ccRCCs, chRCCs, and oncocytomas, the median of voxel gray-level values achieved the best performance with an AUC of 0.811. Interestingly, the worst performance was found in differentiating between chRCCs vs. pRCCs, ccRCCs, and oncocytomas, in which case the highest AUC was 0.757.

In multivariate analysis, the SVC model trained on the best 43 features achieved excellent diagnostic accuracy in both three tasks with AUCs of 0.85-0.92.

The results of our SVC trained on the corticomedullary phase radiomics features (AUC of 0.87 on the internal test set) are comparable to those reported by Yu et al. (AUC of 0.91 during cross-validation) who constructed an SVC from nephrographic phase radiomics features which may indicate that the performance of the corticomedullary phase features is similar to that of the nephrographic phase features in identifying ccRCCs. However, it is important to note, that in our study, we reported the results on independent test cases, while the single-center study by Yu et al. reported cross-validated results and they also included 10 oncocytomas in their dataset.

Chen et al. aimed to compare the diagnostic value of radiomics features extracted from corticomedullary and nephrographic phase CT series in differentiating ccRCCs from non-ccRCCs (77). By constructing a LASSO-based radiomics score, three logistic regression models were built, one from non-textural features, one from textural features, and one from a combination of the two. Although the model of the corticomedullary phase outperformed the nephrographic phase in both cases, where it was constructed from non-textural features (AUC 0.823 vs. 0.784) and where it was constructed from textural features (AUC 0.887 vs. 0.826), the combined model performance showed similar results for the corticomedullary phase and nephrographic phases with AUC values of 0.891 and 0.900, respectively. The results of this study showed that non-textural features added to a model based on textural features can improve its predictive performance and that radiomics features extracted from the corticomedullary phase and the nephrographic phase have a similar diagnostic ability in distinguishing ccRCCs from non-ccRCCs. Although these models were not validated on an independent test set, their results are comparable with the performance of our SVC that was constructed from the corticomedullary phase features achieving AUCs of 0.951, 0.873, and 0.834 on the training set, internal test set, and external test set, respectively. As the publicly available KiTS19 dataset did not contain nephrographic phase CT scans, we were not able to investigate the value of nephrographic phase radiomics features, but we proved that the unenhanced phase and the excretory phase radiomics features have no added value to those extracted from the corticomedullary phase.

Wang et al. investigated the performance of radiomics-based ML models including an RFC, an SVC, and a logistic regression algorithm (75). By evaluating the performance of the models on independent internal test cases, the authors found that the RFC had the best performance with an AUC of 0.909, while the logistic regression classifier reached an AUC of 0.906 and the SVC resulted in an AUC of 0.841. These results are similar to those that we reported during internal validation of the SVC which achieved an AUC of 0.88, however, we also proved our model's generalizability during external validation. This study also reported the diagnostic performance of an expert radiologist, and the authors were able to show that their proposed ML model can outperform an experienced radiologist's accuracy. However, the performance of the radiologist reported in this study was worse than that of our study (AUC of 0.69 vs. 0.76-0.88, sensitivity of 0.85 vs. 0.84-0.93, and specificity of 0.58 vs. 0.70-0.85).

Externally validated results can be found in only two studies in the relevant literature; however, these studies also have several limitations. Li et al. performed texture analysis on both the unenhanced, corticomedullary, and nephrographic phase CT series and compared two feature selection algorithms to test their effectiveness in selecting the most relevant features (79). After applying either a Boruta algorithm or a minimum redundancy maximum relevance algorithm, two RFC models were created from the selected features whose performance was evaluated on 85 independent external test cases from another hospital The first model trained on the 8 corticomedullary texture features that the Boruta algorithm selected achieved an AUC of 0.949, while the second model built from the combination of the 7 nephrographic and one corticomedullary phase texture features yielded an AUC of 0.851. These results are consistent with those reported by Chen et al. indicating that the corticomedullary phase features have better diagnostic ability compared to the nephrographic phase ones. However, these results also underline the importance of feature selection algorithms. Our results further support these previous findings, however, in our study we also investigated the role of first-order features and shape-based features extracted from the tumor volumes in the corticomedullary phase. Two shape-based, three first-order, and five texture features were chosen by the LASSO algorithm as the most significant, which may point to the value of both first-order statistics and shape-based features in addition to texture-based features. Despite the fact that our results on the external test set were marginally worse than those reported by

Li et al. (AUC of 0.834 vs. 0.949), it is not considered as a sign of overfitting, as our test set had a significant number of atypical cases, which is supported by the fact that our SVC model's accuracy was found to be comparable to that of the radiologist expert (accuracy of 0.78 vs. 0.79).

Kocak et al. were among the first, who reported the results of the proposed ML models on publicly available datasets (76). The authors extracted texture features from both the unenhanced and corticomedullary phase CT scans and trained either an ANN-based or an SVC-based model for differentiating between ccRCCs and non-ccRCCs. For external validation, the authors selected 13 ccRCCs, 7 pRCCs, and 6 chRCCs from three publicly available datasets of The Cancer Genome Atlas database (TCGA) including TCGA-Kidney Renal Clear Cell Carcinoma (94, 111), the TCGA-Kidney Renal Papillary Cell Carcinoma (94, 112), and the TCGA-Kidney Chromophobe datasets (94, 113). By using the corticomedullary phase radiomics features, the ANN model combined with adaptive boosting yielded AUCs of 0.870 and 0.822, while the SVC model combined with adaptive boosting resulted in AUCs of 0.852 and 0.793 on the internal and external test sets, respectively. We reported comparable results with those achieved by Kocak et al., our SVC however achieved marginally better performance on both the internal and external test datasets with AUCs of 0.873 and 0.834, respectively, which suggests the superiority of three-dimensional whole tumor analysis over single slice analysis. However, it is important to highlight that while we used 73 tumors from the KiTS19 dataset for external validation, Kocak et al. validated their models by using 26 selected cases from the TCGA dataset.

Our study has a number of limitations that need to be pointed out. First, this was a retrospective study, where we did not apply strict inclusion and exclusion criteria to avoid selection bias. Thus, the distribution of tumor subtypes in the patient groups was unbalanced, reflecting the real-life circumstances, as the prevalence of these tumors in the population is also unbalanced. Secondly, due to the huge prevalence differences between tumor subtypes, only the three most common RCC subtypes were included in this study. Finally, the analysis of nephrographic phase CT scans was not possible in our study as they were not available in the KiTS19 dataset.

5.3. Discussing the methodological differences between our two studies

The main steps of radiomics analysis were the same between our two studies, however, the methods showed some differences. It must be highlighted, that no universally accepted pipeline exists yet, and a vast number of acceptable algorithms can be found in the literature, therefore, the methods should be chosen to be optimized for the given research question and the characteristics of the analyzed dataset.

Regarding radiomics feature extraction by the pyRadiomics package, in the first study, we performed the radiomics analysis with a bin width value of 25, however in the second study an optimal bin width of 16 was defined based on the density range inside of the segmented VOIs. Furthermore, in the first study on the liver parenchyma, we extracted higher-order radiomics features from the LoG-filtered and wavelet-filtered images to reveal differences between fibrosis stages in fine, medium, and coarse texture features. In the second study, due to the unbalanced distribution and the relatively low number of chRCCs and pRCCs tumors, we decided to limit the analysis to the original unfiltered images to avoid overfitting.

During data preprocessing, we found outliers in the first study, therefore instead of classic standardization according to the mean and standard deviation, we choose a robust scaling algorithm that is not sensitive to outliers, while in the second study on kidney cancer, we used classical standardization to rescale the feature values by setting the mean to zero and the standard deviation to one.

In both studies, a linear and a non-linear ML pipeline were constructed, using either RFC or SVC algorithms in the final classification models. However, the feature selection algorithms showed some differences between the two studies. In both studies, Pearson's correlation coefficient analysis was used to filter out highly correlated features that are redundant. But in the first study on liver fibrosis, we then applied an FDR filter, while in the second analysis, we investigated the radiomics features reproducibility and robustness against slight differences in segmentation by using ICC analysis. The ML-based feature selection algorithms also differed, in the first study we used either an SVC-based or an RFC-based RFE algorithm, while in the second study, RFE was not efficient enough to reduce the number of features and prevent overfitting when applied to the training set, so we used either a LASSO or a tuned ReliefF algorithm as feature selector.

6. CONCLUSIONS

Our first research investigating the role of radiomics analysis in the non-invasive staging of liver fibrosis was among the first studies in the literature that assessed the feasibility of three-dimensional liver segmentation and three-dimensional radiomics feature extraction from the liver parenchyma. Comparing our results to those of previously published studies, an important advancement of our study was that we validated the diagnostic performance of our models on independent test cases. In addition to that we demonstrated that texture features extracted from portal venous phase CT scans that quantitatively evaluate the tissue heterogeneity showed a strong association with the progression of liver fibrosis; we also proved that our proposed feature extraction, data-post-processing, and feature selection pipeline is suitable for extracting meaningful imaging biomarkers from CT scans that are robust against the type of CT scanners and acquisition protocols. And finally, by testing two data splitting strategies, we demonstrated that radiomics analysis combined with machine learning model building can be used for differentiating between low-grade and high-grade fibrosis, thus the proposed non-invasive approach could successfully be used for screening chronic liver disease patients and identifying patients at high risk.

In our second study, we assessed the role of radiomics analysis in the CT-based diagnostics of kidney cancers. After extracting radiomics features from the unenhanced, corticomedullary phase and excretory phase preoperative CT scans, we successfully built an SVC-based ML model from the corticomedullary phase radiomics features that was able to distinguish between ccRCCs and non-ccRCCs with an accuracy comparable to that of an expert radiologist. However, we were not able to prove the added value of unenhanced and excretory phase radiomics features over the corticomedullary phase ones. The performance of our corticomedullary phase model was successfully validated on both hold-out internal test cases, and external test cases from the KiTS19 dataset, demonstrating the reliability and generalizability of our ML model. Therefore, we conclude that radiomics analysis combined with machine learning as a quantitative and objective method can facilitate the non-invasive diagnosis of RCCs.

7. SUMMARY

In our first study, we retrospectively identified patients with chronic liver disease who had undergone both SWE measurement of liver stiffness and portal venous phase contrast-enhanced CT scan. A total number of 30 cases were identified and assigned to low-grade and high-grade fibrosis groups according to the stiffness measurements. After the entire volume of the liver, as well as the lobes and the anatomic liver segments were manually segmented, three-dimensional radiomics analysis was performed on these volumes including a total number of 1117 radiomics features. In this study, the cases were first split into equal portions of training and test set, while in a second analysis, the cases were divided based on the scanner types. After feature selection, the best-performing features were used for building RFC and SVC models. In the first analysis, the RFC model achieved excellent performance on both the training and test cases with AUCs of 0.95 vs. 0.90, while SVC showed overfitting as its performance dropped from 0.88 to 0.76. However, in the second analysis, the classification performance showed no decrease between the training and test sets, the RFC yielded AUCs of 0.84 vs. 0.88, while the SVC achieved AUCs of 0.91 vs. 0.90, respectively, indicating no relevant overfitting.

In our second study, we retrospectively collected preoperative CT scans from patients who had undergone nephrectomy and had been diagnosed with RCC at our Institution. After retrieving the unenhanced, corticomedullary, and excretory phase series of 209 RCC patients, the three-dimensional tumor volumes were manually segmented for radiomics analysis. A total number of 107 radiomics features were then extracted from the tumor volumes from each phase separately. After feature selection, the best-performing features were used for building an SVC-based and an RFC-based machine learning model for distinguishing between ccRCCs and non-ccRCCs. The effectiveness of the models built was first tested on cases from our own institution and then was also validated externally on cases from the KiTS19 dataset. The corticomedullary phase SVC achieving an AUC of 0.87 vs. 0.83 demonstrated its reliability and generalizability. Its accuracy was also compared on the external test cases to that of an expert radiologist which proved the effectiveness of the radiomics-based approach (accuracy of 0.79 vs. 0.78). Thus, we draw the conclusion that radiomic analysis combined with machine learning can be a useful tool in the non-invasive diagnosis of RCC in daily clinical practice.

8. REFERENCES

1. Chernyak V, Fowler KJ, Kamaya A, Kielar AZ, Elsayes KM, Bashir MR, Kono Y, Do RK, Mitchell DG, Singal AG, Tang A, Sirlin CB. Liver Imaging Reporting and Data System (LI-RADS) Version 2018: Imaging of Hepatocellular Carcinoma in At-Risk Patients. Radiology. 2018;289(3):816-830. doi: 10.1148/radiol.2018181494

2. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48(4):441-446. doi: 10.1016/j.ejca.2011.11.036

3. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016;278(2):563-577. doi: 10.1148/radiol.2015151169

4. L'Heureux A, Grolinger K, Elyamany HF, Capretz MAM. Machine Learning With Big Data: Challenges and Approaches. IEEE Access. 2017;5:7776-7797. doi: 10.1109/ACCESS.2017.2696365

5. Lubner MG, Smith AD, Sandrasegaran K, Sahani DV, Pickhardt PJ. CT Texture Analysis: Definitions, Applications, Biologic Correlates, and Challenges. Radiographics. 2017;37(5):1483-1503. doi: 10.1148/rg.2017170056

6. Budai B. Figure 1. Workflow of radiomics analysis. 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

7. Shur JD, Doran SJ, Kumar S, Ap Dafydd D, Downey K, O'Connor JPB, Papanikolaou N, Messiou C, Koh DM, Orton MR. Radiomics in Oncology: A Practical Guide. Radiographics. 2021;41(6):1717-1732. doi: 10.1148/rg.2021210037

8. Carré A, Klausner G, Edjlali M, Lerousseau M, Briend-Diop J, Sun R, Ammari S, Reuzé S, Alvarez Andres E, Estienne T, Niyoteka S, Battistella E, Vakalopoulou M, Dhermain F, Paragios N, Deutsch E, Oppenheim C, Pallud J, Robert C. Standardization of brain MR images across machines and protocols: bridging the gap for MRI-based radiomics. Sci Rep. 2020;10(1):12340. doi: 10.1038/s41598-020-69298-z

9. Hoebel KV, Patel JB, Beers AL, Chang K, Singh P, Brown JM, Pinho MC, Batchelor TT, Gerstner ER, Rosen BR, Kalpathy-Cramer J. Radiomics Repeatability Pitfalls in a Scan-Rescan MRI Study of Glioblastoma. Radiol Artif. 2021;3(1):e190199. doi: 10.1148/ryai.2020190199

10. Dercle L, Ma J, Xie C, Chen AP, Wang D, Luk L, Revel-Mouroz P, Otal P, Peron JM, Rousseau H, Lu L, Schwartz LH, Mokrane FZ, Zhao B. Using a single abdominal computed tomography image to differentiate five contrast-enhancement phases: A machine-learning algorithm for radiomics-based precision medicine. Eur J Radiol. 2020;125:108850. doi: 10.1016/j.ejrad.2020.108850

11. Perrin T, Midya A, Yamashita R, Chakraborty J, Saidon T, Jarnagin WR, Gonen M, Simpson AL, Do RKG. Short-term reproducibility of radiomic features in liver parenchyma and liver malignancies on contrast-enhanced CT imaging. Abdom Radiol. 2018;43(12):3271-3278. doi: 10.1007/s00261-018-1600-6

12. Roy S, Whitehead TD, Quirk JD, Salter A, Ademuyiwa FO, Li S, An H, Shoghi KI. Optimal co-clinical radiomics: Sensitivity of radiomic features to tumour volume, image noise and resolution in co-clinical T1-weighted and T2-weighted magnetic resonance imaging. EBioMedicine. 2020;59:102963. doi: 10.1016/j.ebiom.2020.102963

13. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. Magn Reson Med. 1995;34(6):910-914. doi: 10.1002/mrm.1910340618

14. Gjesteby L, Man BD, Jin Y, Paganetti H, Verburg J, Giantsoudi D, Wang G. Metal Artifact Reduction in CT: Where Are We After Four Decades? IEEE Access. 2016;4:5826-5849. doi: 10.1109/ACCESS.2016.2608621

15. Kaposi PN, Youn T, Tóth A, Frank VG, Shariati S, Szendrői A, Magyar P, Bérczi V. Orthopaedic metallic artefact reduction algorithm facilitates CT evaluation of the urinary tract after hip prosthesis. Clin Radiol. 2020;75(1):78.e17-78.e24. doi: 10.1016/j.crad.2019.08.027

16. Shafiq-Ul-Hassan M, Zhang GG, Latifi K, Ullah G, Hunt DC, Balagurunathan Y, Abdalah MA, Schabath MB, Goldgof DG, Mackin D, Court LE, Gillies RJ, Moros EG. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. Medical Phys. 2017;44(3):1050-1062. doi: 10.1002/mp.12123

17. Zwanenburg A, Vallières M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, Ashrafinia S, Bakas S, Beukinga RJ, Boellaard R, Bogowicz M, Boldrini L, Buvat I, Cook GJR, Davatzikos C, Depeursinge A, Desseroit MC, Dinapoli N, Dinh CV, Echegaray S, El Naqa I, Fedorov AY, Gatta R, Gillies RJ, Goh V, Götz M, Guckenberger M, Ha SM, Hatt M, Isensee F, Lambin P, Leger S, Leijenaar RTH, Lenkowicz J, Lippert F, Losnegård A, Maier-Hein KH, Morin O, Müller H, Napel S, Nioche C, Orlhac F, Pati S, Pfaehler

EAG, Rahmim A, Rao AUK, Scherer J, Siddique MM, Sijtsema NM, Socarras Fernandez J, Spezi E, Steenbakkers R, Tanadini-Lang S, Thorwarth D, Troost EGC, Upadhaya T, Valentini V, van Dijk LV, van Griethuysen J, van Velden FHP, Whybra P, Richter C, Löck S. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. Radiology. 2020;295(2):328-338. doi: 10.1148/radiol.2020191145

18. Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging. 2010;29(1):196-205. doi: 10.1109/tmi.2009.2035616

19. Kocak B, Durmaz ES, Erdim C, Ates E, Kaya OK, Kilickesmez O. Radiomics of Renal Masses: Systematic Review of Reproducibility and Validation Strategies. AJR Am J Roentgenol. 2020;214(1):129-136. doi: 10.2214/ajr.19.21709

20. Xie C, Yang P, Zhang X, Xu L, Wang X, Li X, Zhang L, Xie R, Yang L, Jing Z, Zhang H, Ding L, Kuang Y, Niu T, Wu S. Sub-region based radiomics analysis for survival prediction in oesophageal tumours treated by definitive concurrent chemoradiotherapy. EBioMedicine. 2019;44:289-297. doi: 10.1016/j.ebiom.2019.05.023

21. Li N, Wan X, Zhang H, Zhang Z, Guo Y, Hong D. Tumor and peritumor radiomics analysis based on contrast-enhanced CT for predicting early and late recurrence of hepatocellular carcinoma after liver resection. BMC Cancer. 2022;22(1):664. doi: 10.1186/s12885-022-09743-6

22. Budai B, Borsos P, Magyar P, Bérczi V, Kaposi N. ECR 2020 Book of Abstracts. S 8-4. Convolutional neural network-based volumetric segmentation of the liver compared with semi-automatic and manual methods. Insights Imaging. 2020;11(Suppl 1):34. doi: 10.1186/s13244-020-00851-0

23. Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative. 2019. doi: 10.48550/arXiv.1612.07003

24. Budai B. Figure 3. Schematic illustration of extracting histogram-based first-order statistical features. 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

25. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. IEEE Trans Syst Man Cybern. 1973;SMC-3(6):610-621. doi: 10.1109/TSMC.1973.4309314

26. Galloway MM. Texture analysis using gray level run lengths. Comput Graph Image Process. 1975;4(2):172-179. doi: 10.1016/S0146-664X(75)80008-6

27. Budai B. Figure 4. Schematic illustration of the computation of gray-level co-occurrence matrix from a 5-voxel wide region of interest. 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

28. Budai B. Figure 5. Schematic illustration of the computation of gray-level run-length matrix from a 5-voxel wide region of interest 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

29. Thibault G, Angulo J, Meyer F. Advanced statistical matrices for texture characterization: application to cell classification. IEEE Trans Biomed Eng. 2014;61(3):630-637. doi: 10.1109/tbme.2013.2284600

30. Budai B. Figure 6. Schematic illustration of the computation of gray-level size zone matrix from a 5-voxel wide region of interest 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

31. Sun C, Wee WG. Neighboring gray level dependence matrix for texture classification. Computer Vision, Graphics, and Image Processing. 1983;23(3):341-352. doi: 10.1016/0734-189X(83)90032-4

32. Budai B. Figure 7. Schematic illustration of the computation of gray-level dependence matrix from a 5-voxel wide region of interest 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

33. Amadasun M, King R. Textural features corresponding to textural properties. IEEE Trans Syst Man Cybern. 1989;19(5):1264-1274. doi: 10.1109/21.44046

34. Budai B. Figure 8. Schematic illustration of the computation of neighboring gray-tone difference matrix from a 5-voxel wide region of interest 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

35. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S, Aerts H. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Res. 2017;77(21):e104-e107. doi: 10.1158/0008-5472.Can-17-0339

36. Wang S, Summers RM. Machine learning and radiology. Med Image Anal. 2012;16(5):933-951. doi: 10.1016/j.media.2012.02.005

37. Koçak B, Durmaz E, Ateş E, Kılıçkesmez Ö. Radiomics with artificial intelligence: a practical guide for beginners. Diagn Intervl Radiol. 2019;25(6):485-495. doi: 10.5152/dir.2019.19321

38. Budai B. Figure 9. Schematic illustration of fitting machine learning models to the training dataset. 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

39. van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging-"how-to" guide and critical reflection. Insights Imaging. 2020;11(1):91. doi: 10.1186/s13244-020-00887-2

40. Anowar F, Sadaoui S, Selim B. Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE). Comput Sci Rev. 2021;40:100378. doi: 10.1016/j.cosrev.2021.100378

41. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ Brit Med J. 2015;350:g7594. doi: 10.1136/bmj.g7594

42. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, Sanduleanu S, Larue R, Even AJG, Jochems A, van Wijk Y, Woodruff H, van Soest J, Lustberg T, Roelofs E, van Elmpt W, Dekker A, Mottaghy FM, Wildberger JE, Walsh S. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14(12):749-762. doi: 10.1038/nrclinonc.2017.141

43. Friedman SL. Mechanisms of hepatic fibrogenesis. Gastroenterology. 2008;134(6):1655-1669. doi: 10.1053/j.gastro.2008.03.003

44. Schuppan D, Afdhal NH. Liver cirrhosis. Lancet. 2008;371(9615):838-851. doi: 10.1016/s0140-6736(08)60383-9

45. Huang DQ, Terrault NA, Tacke F, Gluud LL, Arrese M, Bugianesi E, Loomba R. Global epidemiology of cirrhosis - aetiology, trends and predictions. Nat Rev Gastroenterol Hepatol. 2023:1-11. doi: 10.1038/s41575-023-00759-2

46. Sepanlou SG, Safiri S, Bisignano C, Ikuta KS, Merat S, Saberifiroozi M, Poustchi H, Tsoi D, Colombara DV, Abdoli A, Adedoyin RA, Afarideh M, Agrawal S, Ahmad S, Ahmadian E, Ahmadpour E, Akinyemiju T, Akunna CJ, Alipour V, Almasi-Hashiani A, Almulhim AM, Al-Raddadi RM, Alvis-Guzman N, Anber NH, Angus C, Anoushiravani A, Arabloo J, Araya EM, Asmelash D, Ataeinia B, Ataro Z, Atout MMdW, Ausloos F, Awasthi A, Badawi A, Banach M, Bejarano Ramirez DF, Bhagavathula AS, Bhala N,

Bhattacharyya K, Biondi A, Bolla SR, Boloor A, Borzì AM, Butt ZA, Cámera LLAA, Campos-Nonato IR, Carvalho F, Chu D-T, Chung S-C, Cortesi PA, Costa VM, Cowie BC, Daryani A, de Courten B, Demoz GT, Desai R, Dharmaratne SD, Djalalinia S, Do HT, Dorostkar F, Drake TM, Dubey M, Duncan BB, Effiong A, Eftekhari A, Elsharkawy A, Etemadi A, Farahmand M, Farzadfar F, Fernandes E, Filip I, Fischer F, Gebremedhin KBB, Geta B, Gilani SA, Gill PS, Gutirrez RA, Haile MT, Haj-Mirzaian A, Hamid SS, Hasankhani M, Hasanzadeh A, Hashemian M, Hassen HY, Hay SI, Hayat K, Heidari B, Henok A, Hoang CL, Hostiuc M, Hostiuc S, Hsieh VC-r, Igumbor EU, Ilesanmi OS, Irvani SSN, Jafari Balalami N, James SL, Jeemon P, Jha RP, Jonas JB, Jozwiak JJ, Kabir A, Kasaeian A, Kassaye HG, Kefale AT, Khalilov R, Khan MA, Khan EA, Khater A, Kim YJ, Koyanagi A, La Vecchia C, Lim L-L, Lopez AD, Lorkowski S, Lotufo PA, Lozano R, Magdy Abd El Razek M, Mai HT, Manafi N, Manafi A, Mansournia MA, Mantovani LG, Mazzaglia G, Mehta D, Mendoza W, Menezes RG, Mengesha MM, Meretoja TJ, Mestrovic T, Miazgowski B, Miller TR, Mirrakhimov EM, Mithra P, Moazen B, Moghadaszadeh M, Mohammadian-Hafshejani A, Mohammed S, Mokdad AH, Montero-Zamora PA, Moradi G, Naimzada MD, Nayak V, Negoi I, Nguyen TH, Ofori-Asenso R, Oh I-H, Olagunju TO, Padubidri JR, Pakshir K, Pana A, Pathak M, Pourshams A, Rabiee N, Radfar A, Rafiei A, Ramezanzadeh K, Rana SMM, Rawaf S, Rawaf DL, Reiner RC, Jr., Roever L, Room R, Roshandel G, Safari S, Samy AM, Sanabria J, Sartorius B, Schmidt MI, Senthilkumaran S, Shaikh MA, Sharif M, Sharifi A, Shigematsu M, Singh JA, Soheili A, Suleria HAR, Teklehaimanot BF, Tesfay BE, Vacante M, Vahedian-Azimi A, Valdez PR, Vasankari TJ, Vu GT, Waheed Y, Weldegwergs KG, Werdecker A, Westerman R, Wondafrash DZ, Wondmieneh AB, Yeshitila YG, Yonemoto N, Yu C, Zaidi Z, Zarghi A, Zelber-Sagi S, Zewdie KA, Zhang Z-J, Zhao X-J, Naghavi M, Malekzadeh R. The global, regional, and national burden of cirrhosis by cause in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Gastroenterol Hepatol. 2020;5(3):245-266. doi: 10.1016/S2468-1253(19)30349-8

47. Global Health Estimates 2020: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019. Geneva, World Health Organization. 2020. doi:

48. Budai B. Figure 10. Pathomechanism of liver fibrosis. 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

49. The French METAVIR Cooperative Study Group, Bedossa P. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. The French METAVIR Cooperative Study Group. Hepatology. 1994;20(1):15-20. doi: 10.1002/hep.1840200104

50. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. Hepatology. 1996;24(2):289-293. doi: 10.1002/hep.510240201

51. Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, Grimaldi A, Capron F, Poynard T. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology. 2005;128(7):1898-1906. doi: 10.1053/j.gastro.2005.03.084

52. Bedossa P, Dargère D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology. 2003;38(6):1449-1457. doi: 10.1016/j.hep.2003.09.022

53. Bravo AA, Sheth SG, Chopra S. Liver biopsy. N Eng J Med. 2001;344(7):495-500. doi: 10.1056/nejm200102153440706

54. Dietrich CF, Bamber J, Berzigotti A, Bota S, Cantisani V, Castera L, Cosgrove D, Ferraioli G, Friedrich-Rust M, Gilja OH, Goertz RS, Karlas T, de Knegt R, de Ledinghen V, Piscaglia F, Procopet B, Saftoiu A, Sidhu PS, Sporea I, Thiele M. EFSUMB Guidelines and Recommendations on the Clinical Use of Liver Ultrasound Elastography, Update 2017 (Long Version). Ultraschall in der Medizin. 2017;38(4):e16-e47. doi: 10.1055/s-0043-103952

55. Ferraioli G, Wong VW, Castera L, Berzigotti A, Sporea I, Dietrich CF, Choi BI, Wilson SR, Kudo M, Barr RG. Liver Ultrasound Elastography: An Update to the World Federation for Ultrasound in Medicine and Biology Guidelines and Recommendations. Ultrasound Med Biol. 2018;44(12):2419-2440. doi: 10.1016/j.ultrasmedbio.2018.07.008

56. Huber A, Ebner L, Heverhagen JT, Christe A. State-of-the-art imaging of liver fibrosis and cirrhosis: A comprehensive review of current applications and future perspectives. Eur J Radiol open. 2015;2:90-100. doi: 10.1016/j.ejro.2015.05.002

57. Kayaaltı Ö, Aksebzeci BH, Karahan İÖ, Deniz K, Öztürk M, Yılmaz B, Kara S, Asyalı MH. Liver fibrosis staging using CT image texture analysis and soft computing. Appl Soft Comput. 2014;25:399-413. doi: 10.1016/j.asoc.2014.08.065

58. Zhang X, Gao X, Liu BJ, Ma K, Yan W, Liling L, Yuhong H, Fujita H. Effective staging of fibrosis by the selected texture features of liver: Which one is better, CT or MR

imaging? Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society. 2015;46 Pt 2:227-236. doi: 10.1016/j.compmedimag.2015.09.003

59. Lubner MG, Malecki K, Kloke J, Ganeshan B, Pickhardt PJ. Texture analysis of the liver at MDCT for assessing hepatic fibrosis. Abdom Radiol. 2017;42(8):2069-2078. doi: 10.1007/s00261-017-1096-5

60. Daginawala N, Li B, Buch K, Yu H, Tischler B, Qureshi MM, Soto JA, Anderson S. Using texture analyses of contrast enhanced CT to assess hepatic fibrosis. Eur J Radiol. 2016;85(3):511-517. doi: 10.1016/j.ejrad.2015.12.009

61. Lubner MG, Jones D, Kloke J, Said A, Pickhardt PJ. CT texture analysis of the liver for assessing hepatic fibrosis in patients with hepatitis C virus. Br J Radiol. 2019;92(1093):20180153. doi: 10.1259/bjr.20180153

62. Pickhardt PJ, Graffy PM, Said A, Jones D, Welsh B, Zea R, Lubner MG. Multiparametric CT for Noninvasive Staging of Hepatitis C Virus-Related Liver Fibrosis: Correlation With the Histopathologic Fibrosis Score. AJR Am J Roentgenol. 2019;212(3):547-553. doi: 10.2214/ajr.18.20284

63. Budai B. Figure 11. Classic morphological signs of end-stage liver cirrhosis on CT scans. 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

64. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: Cancer J Clin. 2021;71(3):209-249. doi: 10.3322/caac.21660

65. Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F. Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. 2020 [Available from: https://gco.iarc.fr/today/data/factsheets/populations/348-hungary-fact-sheets.pdf.

66. Moch H, Cubilla AL, Humphrey PA, Reuter VE, Ulbright TM. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. Eur Urol. 2016;70(1):93-105. doi: 10.1016/j.eururo.2016.02.029

67. Hsieh JJ, Purdue MP, Signoretti S, Swanton C, Albiges L, Schmidinger M, Heng DY, Larkin J, Ficarra V. Renal cell carcinoma. Nat Rev Dis Primers. 2017;3:17009. doi: 10.1038/nrdp.2017.9

68. Ricketts CJ, De Cubas AA, Fan H, Smith CC, Lang M, Reznik E, Bowlby R, Gibb EA, Akbani R, Beroukhim R, Bottaro DP, Choueiri TK, Gibbs RA, Godwin AK, Haake S, Hakimi AA, Henske EP, Hsieh JJ, Ho TH, Kanchi RS, Krishnan B, Kwiatkowski DJ, Lui W, Merino MJ, Mills GB, Myers J, Nickerson ML, Reuter VE, Schmidt LS, Shelley CS, Shen H, Shuch B, Signoretti S, Srinivasan R, Tamboli P, Thomas G, Vincent BG, Vocke CD, Wheeler DA, Yang L, Kim WY, Robertson AG, Spellman PT, Rathmell WK, Linehan WM. The Cancer Genome Atlas Comprehensive Molecular Characterization of Renal Cell Carcinoma. Cell Rep. 2018;23(1):313-326.e5. doi: 10.1016/j.celrep.2018.03.075

69. Delahunt B, Cheville JC, Martignoni G, Humphrey PA, Magi-Galluzzi C, McKenney J, Egevad L, Algaba F, Moch H, Grignon DJ, Montironi R, Srigley JR. The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters. Am J Surg Pathol. 2013;37(10):1490-1504. doi: 10.1097/PAS.0b013e318299f0fb

70. Leibovich BC, Lohse CM, Crispen PL, Boorjian SA, Thompson RH, Blute ML, Cheville JC. Histological subtype is an independent predictor of outcome for patients with renal cell carcinoma. J Urol. 2010;183(4):1309-1315. doi: 10.1016/j.juro.2009.12.035

71. Capitanio U, Cloutier V, Zini L, Isbarn H, Jeldres C, Shariat SF, Perrotte P, Antebi E, Patard JJ, Montorsi F, Karakiewicz PI. A critical assessment of the prognostic value of clear cell, papillary and chromophobe histological subtypes in renal cell carcinoma: a population-based study. BJU Int. 2009;103(11):1496-1500. doi: 10.1111/j.1464-410X.2008.08259.x

72. Escudier B, Porta C, Schmidinger M, Rioux-Leclercq N, Bex A, Khoo V, Grünwald V, Gillessen S, Horwich A. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. Ann Oncol. 2019;30(5):706-720. doi: 10.1093/annonc/mdz056

73. Bazzocchi MV, Zilioli C, Gallone VI, Commisso C, Bertolotti L, Pagnini F, Ziglioli F, Maestroni U, Aliprandi A, Buti S, Procopio G, Ascenti G, Martini C, De Filippo

M. The Role of CT Imaging in Characterization of Small Renal Masses. Diagnostics [Internet]. 2023; 13(3).

74. Lee-Felker SA, Felker ER, Tan N, Margolis DJ, Young JR, Sayre J, Raman SS. Qualitative and quantitative MDCT features for differentiating clear cell renal cell carcinoma from other solid renal cortical masses. AJR Am J Roentgenol. 2014;203(5):W516-24. doi: 10.2214/ajr.14.12460

75. Wang P, Pei X, Yin XP, Ren JL, Wang Y, Ma LY, Du XG, Gao BL. Radiomics models based on enhanced computed tomography to distinguish clear cell from non-clear cell renal cell carcinomas. Sci Rep. 2021;11(1):13729. doi: 10.1038/s41598-021-93069-z

76. Kocak B, Yardimci AH, Bektas CT, Turkcanoglu MH, Erdim C, Yucetas U, Koca SB, Kilickesmez O. Textural differences between renal cell carcinoma subtypes: Machine learning-based quantitative computed tomography texture analysis with independent external validation. Eur J Radiol. 2018;107:149-157. doi: 10.1016/j.ejrad.2018.08.014

77. Chen M, Yin F, Yu Y, Zhang H, Wen G. CT-based multi-phase Radiomic models for differentiating clear cell renal cell carcinoma. Cancer Imaging. 2021;21(1):42. doi: 10.1186/s40644-021-00412-8

78. Yu H, Scalera J, Khalid M, Touret AS, Bloch N, Li B, Qureshi MM, Soto JA, Anderson SW. Texture analysis as a radiomic marker for differentiating renal tumors. Abdom Radiol. 2017;42(10):2470-2478. doi: 10.1007/s00261-017-1144-1

79. Li ZC, Zhai G, Zhang J, Wang Z, Liu G, Wu GY, Liang D, Zheng H. Differentiation of clear cell and non-clear cell renal cell carcinomas by all-relevant radiomics features from multiphase CT: a VHL mutation perspective. Eur Radiol. 2019;29(8):3996-4007. doi: 10.1007/s00330-018-5872-6

80. Budai B. Figure 12. Main types of renal cell carcinomas and their typical appearance on corticomedullary phase post-contrast CT scans. 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

81. Budai BK, Stollmayer R, Rónaszéki AD, Körmendy B, Zsombor Z, Palotás L, Fejér B, Szendrõi A, Székely E, Maurovich-Horvat P, Kaposi PN. Radiomics analysis of contrast-enhanced CT scans can distinguish between clear cell and non-clear cell renal cell carcinoma in different imaging protocols. Front Med. 2022;9:974485. doi: 10.3389/fmed.2022.974485

82. Budai BK, Tóth A, Borsos P, Frank VG, Shariati S, Fejér B, Folhoffer A, Szalay F, Bérczi V, Kaposi PN. Three-dimensional CT texture analysis of anatomic liver segments can differentiate between low-grade and high-grade fibrosis. BMC Med Imaging. 2020;20(1):108. doi: 10.1186/s12880-020-00508-w

83. Fejér B, Kucsa A, Unger Z, Youn T, Abonyi M, Bérczi V, Kaposi P. Validation of interobserver reliabilty of the S-Shearwave liver elastography application. Poster C-1265. ECR 2018; Vienna 2018.

84. Kaposi PN, Unger Z, Fejér B, Kucsa A, Tóth A, Folhoffer A, Szalay F, Bérczi V. Interobserver agreement and diagnostic accuracy of shearwave elastography for the staging of hepatitis C virus-associated liver fibrosis. J Clin Ultrasound. 2020;48(2):67-74. doi: 10.1002/jcu.22771

85. Folhoffer A, Rónaszéki AD, Budai BK, Borsos P, Orbán V, Győri G, Szalay F, Kaposi PN. Follow-Up of Liver Stiffness with Shear Wave Elastography in Chronic Hepatitis C Patients in Sustained Virological Response Augments Clinical Risk Assessment. Processes [Internet]. 2021; 9(5).

86. Li X, Morgan PS, Ashburner J, Smith J, Rorden C. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. J Neurosci Methods. 2016;264:47-56. doi: 10.1016/j.jneumeth.2016.03.001

87. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, Buatti J, Aylward S, Miller JV, Pieper S, Kikinis R. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging. 2012;30(9):1323-1341. doi: 10.1016/j.mri.2012.05.001

88. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet Epidemiol. 2008;32(4):361-369. doi: 10.1002/gepi.20310

89. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, Bellomi M. Radiomics: the facts and the challenges of image analysis. Eur Radiol Experimental. 2018;2(1):36. doi: 10.1186/s41747-018-0068-z

90. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20:53-65. doi: 10.1016/0377-0427(87)90125-7

91. Mwangi B, Tian TS, Soares JC. A review of feature reduction techniques in neuroimaging. Neuroinformatics. 2014;12(2):229-244. doi: 10.1007/s12021-013-9204-3

92. Wagner MW, Namdar K, Biswas A, Monah S, Khalvati F, Ertl-Wagner BB. Radiomics, machine learning, and artificial intelligence-what the neuroradiologist needs to know. Neuroradiology. 2021;63(12):1957-1967. doi: 10.1007/s00234-021-02813-9

93. Avanzo M, Wei L, Stancanello J, Vallières M, Rao A, Morin O, Mattonen SA, El Naqa I. Machine and deep learning methods for radiomics. Med Phys. 2020;47(5):e185-e202. doi: 10.1002/mp.13678

94. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging. 2013;26(6):1045-1057. doi: 10.1007/s10278-013-9622-7

95. Heller N, Isensee F, Maier-Hein KH, Hou X, Xie C, Li F, Nan Y, Mu G, Lin Z, Han M, Yao G, Gao Y, Zhang Y, Wang Y, Hou F, Yang J, Xiong G, Tian J, Zhong C, Ma J, Rickman J, Dean J, Stai B, Tejpaul R, Oestreich M, Blake P, Kaluzniak H, Raza S, Rosenberg J, Moore K, Walczak E, Rengel Z, Edgerton Z, Vasdev R, Peterson M, McSweeney S, Peterson S, Kalapara A, Sathianathen N, Papanikolopoulos N, Weight C. The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: Results of the KiTS19 challenge. Med Image Anal. 2021;67:101821. doi: 10.1016/j.media.2020.101821

96. Heller N, Sathianathen N, Kalapara A, Walczak E, Moore K, Kaluzniak H, Rosenberg J, Blake P, Rengel Z, Oestreich M, Dean J, Tradewell M, Shah A, Tejpaul R, Edgerton Z, Peterson M, Raza S, Regmi S, Papanikolopoulos N, Weight C. Data from C4KC-KiTS [Data set]. The Cancer Imaging Archive. 2019. doi: 10.7937/TCIA.2019.IX49E8NX

97. Budai B. Figure 16. Normalization of the voxel intensity values to the density of the healthy renal cortex on the corticomedullary phase post-contrast CT scans. 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

98. Budai B. Figure 17. Radiomics feature extraction from the unenhanced, corticomedullary, and excretory phase preoperative CT scans of a renal cell carcinoma. 2023. Unpublished. Medical Imaging Centre, Semmelweis University.

99. Cattell R, Chen S, Huang C. Robustness of radiomic features in magnetic resonance imaging: review and a phantom study. Visual computing for industry, biomedicine, and art. 2019;2(1):19. doi: 10.1186/s42492-019-0025-6

100. Dercle L, Lu L, Schwartz LH, Qian M, Tejpar S, Eggleton P, Zhao B, Piessevaux H. Radiomics Response Signature for Identification of Metastatic Colorectal Cancer Sensitive to Therapies Targeting EGFR Pathway. J Natl Cancer Inst. 2020;112(9):902-912. doi: 10.1093/jnci/djaa017

101. Feng Z, Rong P, Cao P, Zhou Q, Zhu W, Yan Z, Liu Q, Wang W. Machine learning-based quantitative texture analysis of CT images of small renal masses: Differentiation of angiomyolipoma without visible fat from renal cell carcinoma. Eur Radiol. 2018;28(4):1625-1633. doi: 10.1007/s00330-017-5118-z

102. Raman SP, Chen Y, Schroeder JL, Huang P, Fishman EK. CT texture analysis of renal masses: pilot study using random forest classification for prediction of pathology. Acad Radiol. 2014;21(12):1587-1596. doi: 10.1016/j.acra.2014.07.023

103. Dana J, Lefebvre TL, Savadjiev P, Bodard S, Gauvin S, Bhatnagar SR, Forghani R, Hélénon O, Reinhold C. Malignancy risk stratification of cystic renal lesions based on a contrast-enhanced CT-based machine learning model and a clinical decision algorithm. Eur Radiol. 2022;32(6):4116-4127. doi: 10.1007/s00330-021-08449-w

104. Schieda N, Thornhill RE, Al-Subhi M, McInnes MD, Shabana WM, van der Pol CB, Flood TA. Diagnosis of Sarcomatoid Renal Cell Carcinoma With CT: Evaluation by Qualitative Imaging Features and Texture Analysis. AJR Am J Roentgenol. 2015;204(5):1013-1023. doi: 10.2214/ajr.14.13279

105. Gurbani S, Morgan D, Jog V, Dreyfuss L, Shen M, Das A, Abel EJ, Lubner MG. Evaluation of radiomics and machine learning in identification of aggressive tumor features in renal cell carcinoma (RCC). Abdom Radiol. 2021;46(9):4278-4288. doi: 10.1007/s00261-021-03083-y

106. Meng X, Shu J, Xia Y, Yang R. A CT-Based Radiomics Approach for the Differential Diagnosis of Sarcomatoid and Clear Cell Renal Cell Carcinoma. BioMed Res Int. 2020;2020:7103647. doi: 10.1155/2020/7103647

107. Yi X, Xiao Q, Zeng F, Yin H, Li Z, Qian C, Wang C, Lei G, Xu Q, Li C, Li M, Gong G, Zee C, Guan X, Liu L, Chen BT. Computed Tomography Radiomics for

Predicting Pathological Grade of Renal Cell Carcinoma. Front Oncol. 2020;10:570396. doi: 10.3389/fonc.2020.570396

108. Xv Y, Lv F, Guo H, Zhou X, Tan H, Xiao M, Zheng Y. Machine learning-based CT radiomics approach for predicting WHO/ISUP nuclear grade of clear cell renal cell carcinoma: an exploratory and comparative study. Insights Imaging. 2021;12(1):170. doi: 10.1186/s13244-021-01107-1

109. Yang L, Gao L, Arefan D, Tan Y, Dan H, Zhang J. A CT-based radiomics model for predicting renal capsule invasion in renal cell carcinoma. BMC Med Imaging. 2022;22(1):15. doi: 10.1186/s12880-022-00741-5

110. Bektas CT, Kocak B, Yardimci AH, Turkcanoglu MH, Yucetas U, Koca SB, Erdim C, Kilickesmez O. Clear Cell Renal Cell Carcinoma: Machine Learning-Based Quantitative Computed Tomography Texture Analysis for Prediction of Fuhrman Nuclear Grade. Eur Radiol. 2019;29(3):1153-1163. doi: 10.1007/s00330-018-5698-2

111. Akin O, Elnajjar P, Heller M, Jarosz R, Erickson BJ, Kirk S, Lee Y, Linehan MW, Gautam R, Vikram R, Garcia KM, Roche C, Bonaccio E, Filippini J. The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma Collection (TCGA-KIRC) (Version 3) [Data set]. The Cancer Imaging Archive. 2016. doi: 10.7937/K9/TCIA.2016.V6PBVTDR

112. Linehan M, Gautam R, Kirk S, Lee Y, Roche C, Bonaccio E, Filippini J, Rieger-Christ K, Lemmerman J, Jarosz R. The Cancer Genome Atlas Cervical Kidney Renal Papillary Cell Carcinoma Collection (TCGA-KIRP) (Version 4) [Data set]. The Cancer Imaging Archive. 2016. doi: 10.7937/K9/TCIA.2016.ACWOGBEF

113. Linehan MW, Gautam R, Sadow CA, Levine S. The Cancer Genome Atlas Kidney Chromophobe Collection (TCGA-KICH) (Version 3) [Data set]. The Cancer Imaging Archive. 2016. doi: 10.7937/K9/TCIA.2016.YU3RBCZN

9. BIBLIOGRAPHY OF THE CANDIDATE’S PUBLICATIONS

9.1. Publications related to the present thesis

1. **Budai BK**, Stollmayer R, Rónaszéki AD, Körmendy B, Zsombor Z, Palotás L, Fejér B, Szendrői A, Székely E., Maurovich-Horvat P, Kaposi PN. Radiomics analysis of contrast-enhanced CT scans can distinguish between clear cell and non-clear cell renal cell carcinoma in different imaging protocols. Front Med. 2022;9:974485. **(IF: 3.9, 2022)**

2. **Budai BK**, Tóth A, Borsos P, Frank VG, Shariati S, Fejér B, Folhoffer A, Szalay F, Bérczi V, Kaposi PN. Three-dimensional CT texture analysis of anatomic liver segments can differentiate between low-grade and high-grade fibrosis. BMC Med Imag. 2020;20(1):108. **(IF: 1.93, 2020)**

9.2. Publications not related to the present thesis

1. Stollmayer R, **Budai BK**, Rónaszéki A, Zsombor Z, Kalina I, Hartmann E, Tóth G, Szoldán P, Bérczi V, Maurovich-Horvat P, Kaposi PN. Focal Liver Lesion MRI Feature Identification Using Efficientnet and MONAI: A Feasibility Study. Cells. 2022;11(9). **(IF:6.0, 2022)**

2. Stollmayer R, **Budai BK**, Tóth A, Kalina I, Hartmann E, Szoldán P, Bérczi V, Maurovich-Horvat P, Kaposi PN. Diagnosis of focal liver lesions with deep learning-based multi-channel analysis of hepatocyte-specific contrast-enhanced magnetic resonance imaging. World J Gastroenterol. 2021;27(35):5978-88. **(IF:5.374, 2021)**

3. Zsombor Z, Rónaszéki AD, Csongrády B, Stollmayer R, **Budai BK**, Folhoffer A, Kalina I, Győri G, Bérczi V, Maurovich-Horvat P., Hagymási K, Kaposi PN. Evaluation of Artificial Intelligence-Calculated Hepatorenal Index for Diagnosing Mild and Moderate Hepatic Steatosis in Non-Alcoholic Fatty Liver Disease. Med Lith. 2023;59(3). **(IF: 2.6, 2022)**

4. Rónaszéki AD, **Budai BK**, Csongrády B, Stollmayer R, Hagymási K, Werling K, Fodor T, Folhoffer A, Kalina I, Győri G, Maurovich-Horvat P, Kaposi PN. Tissue attenuation imaging and tissue scatter imaging for quantitative ultrasound

evaluation of hepatic steatosis. Medicine (Baltimore). 2022;101(33):e29708. **(IF: 1.6, 2022)**

5. Folhoffer A, Rónaszéki AD, **Budai BK**, Borsos P, Orbán V, Győri G, Szalay F, Kaposi PN. Follow-Up of Liver Stiffness with Shear Wave Elastography in Chronic Hepatitis C Patients in Sustained Virological Response Augments Clinical Risk Assessment. Processes. 2021;9(5):753. **(IF:3.352, 2021)**

6. Rónaszéki AD, Dudás I, Zsély B, **Budai BK**, Stollmayer R, Hahn O, Csongrády B, Park B, Maurovich-Horvat P, Győri G, Kaposi PN. Microvascular flow imaging to differentiate focal hepatic lesions: the spoke-wheel pattern as a specific sign of focal nodular hyperplasia. Ultrasonography (Seoul, Korea). 2023;42(1):172-81. **(IF:3.1, 2022)**

7. 10. Fazekas S, **Budai BK**, Stollmayer R, Kaposi PN, Bérczi V. Artificial intelligence and neural networks in radiology – Basics that all radiology residents should know. IMAGING. 2022;14(2):73-81. **(IF: 0.4, 2022)**

8. **Budai BK**, Tánczik Zs, Körmendy B, Zsombor Z, Palotás L, Stollmayer R, Rónaszéki AD, Maurovich-Horvat P, Kaposi PN. Radiomikai elemzés – ami igazán lényeges, az a szemnek sokszor láthatatlan: Metodikai közlemény. MAGYAR RADIOLÓGIA ONLINE. 2022; 13(2): 3/1-16. **(IF: 0.0, 2022)**

9. **Budai BK**, Frank V, Shariati S, Fejér B, Tóth A, Orbán V, Bérczi V, Kaposi PN. CT texture analysis of abdominal lesions – Part I.: Liver lesions. IMAGING. 2021;13(1):13-24. **(IF: 0.0, 2021)**

10. Frank V, Shariati S, **Budai BK**, Fejér B, Tóth A, Orbán V, Bérczi V, Kaposi PN. CT texture analysis of abdominal lesions – Part II: Tumors of the Kidney and Pancreas. IMAGING. 2021;13(1):25-36. **(IF: 0.0, 2021)**

Cumulative impact factor of the candidate's publications related to the thesis: **5.83**

Total cumulative impact factor of the candidate's publications: **24.756**

10. ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Dr. Pál Novák Kaposi, whose unwavering support and expert guidance made it possible to conduct this research and write this diploma work.

I would also like to thank Prof. Pál Maurovich Horvat, who as the Director of the Medical Imaging Centre gave me the opportunity of doing research as a PhD student at his Institution and has supported me consistently over the years.

As the former Director of the Department of Radiology, Prof. Viktor Bérczi enabled me to join the student scientific community of the Institute and supported my research during my undergraduate years. For this, I would also like to express my gratitude.

I also want to express my thanks to Professor Dr. Kinga Karlinger, who served as the institutional leader of the scientific student research community and provided supervision and encouragement for my activities as a TDK and later as a Ph.D. student.

Regarding the radiomic analysis of liver fibrosis, I am grateful to Dr. Zsuzsanna Unger and Dr. Bence Fejér, specialists in radiology, for their assistance in the elastography measurements of chronic liver disease patients, and to Prof. Ferenc Szalay and Dr. Anikó Folhoffer specialists in internal medicine for the follow-up and documentation of the patients.

I would like to thank Dr. Bence Fejér for his assistance in the radiological evaluation of renal tumors and Dr. Attila Szendőri, urologist, for the patient follow-up and documentation in the research on the radiomic analysis of renal tumors.

Finally, I would like to express my sincere gratitude to the radiographers and staff of the Medical Imaging Centre, my students, and everyone who supported my research.