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

1. Introduction

1.1. Radiomics analysis

Radiomics analysis, which was first described by Lambin et al. in 2012 (2), is the process of extracting a large number of quantitative image features, also known as radiomics features, from the regions of interest (ROI) or volumes of interest (VOI) of the scans obtained from medical imaging acquisitions (3). Radiomics analysis is typically combined with machine learning (ML) to identify useful biomarkers among these numerous radiomics features that can be used to predict clinical endpoints. The workflow of radiomics analysis can be divided into the following steps: image acquisition, image post-processing, segmentation, radiomics feature extraction, feature selection, model building, and evaluation. Previous studies demonstrated the potential value of radiomics analysis in the non-invasive diagnostics of various benign and malignant lesions of many different organs (5).

As an example of benign lesions of abdominal organs, we decided to focus our research on liver fibrosis

assessment. The conventional reading of CT scans is not suitable for staging hepatic fibrosis due to its low sensitivity in detecting the early stages of liver fibrosis. Previous studies have suggested that CT may reveal subtle texture differences, even in early stages, that are not visible to the naked eye, and they have shown that computerized CT texture analysis can reveal these subtle textural changes of the liver parenchyma, facilitating the non-invasive staging of liver fibrosis (57-62). These studies, however, only examined selected small parts of the liver – either selected cross-sections or small ROIs – and used a single-center, retrospective study design. Moreover, most of these studies assessed the diagnostic performance of individual radiomics features and did not build multivariate ML models.

For demonstrating the ability of radiomics analysis to differentiate malignant lesions of abdominal organs, we decided to focus our research on renal cell carcinomas. The three most common subtypes of renal cell carcinoma are clear cell (ccRCC), papillary cell (pRCC), and chromophobe cell (chRCC) renal cell carcinoma (68).

These subtypes show significant differences in biological aggressiveness and patient prognosis, therefore the preoperative differentiation between these subtypes is crucial (70, 71). However, in daily clinical practice, the assessment of morphological features varies enormously between observers, and the large number of atypical cases makes it challenging to accurately differentiate between these renal tumor subtypes (74, 75).

Previous studies that investigated the role of radiomics analysis in differentiating between these kidney cancer subtypes (75-77) reported promising results.

These studies used a retrospective, single-center study design and either lacked the validation of the prediction models on independent test cases of external institutions (77, 78), tested the models on cases involving just one external institution from the same country (79), or tested the models on independent cases involving multiple institutions but found that they performed poorly on external cases (76).

2. Objectives

Our first study aimed to investigate the utility of radiomics-based ML analysis in the non-invasive CT-based identifications of patients with advanced-stage liver fibrosis. We investigated the performance of both linear and non-linear ML algorithms for the identification of patients at high risk. The final aim of our study was to prove that radiomics-based ML algorithms could be used for liver fibrosis assessment independent of etiology and scanning protocols.

The aim of our second study was to prove that radiomics analysis combined with ML can facilitate the non-invasive, CT-based preoperative diagnosis of kidney tumors and could be used for distinguishing between clear cell (ccRCC) and non-clear cell (non-ccRCC) renal cell carcinoma subtypes. Our final aim was to build a radiomics ML model that is robust against different institutional imaging protocols and prove its generalizability on external test cases of a publicly available dataset.

3. Methods

The two studies summarized in this dissertation, have been published as original research articles (80, 81). Both studies were approved by Semmelweis University Regional and Institutional Committee of Science and Research Ethics (RKEB: 136/2019). The main steps of radiomics analysis were the same between our two studies, however, the methods showed some differences, which I briefly summarize.

3.1. Patient population

In the first study on liver fibrosis assessment, patients with suspected liver fibrosis who were followed up with shear wave elastography (SWE) since 2017 and had venous phase CT scans at our Institution within six months of the SWE between September 2016 and January 2019 were retrospectively identified. In the second study on kidney cancer diagnostics, patients diagnosed with either clear cell, papillary cell, or chromophobe cell renal cell carcinomas and had available preoperative contrast-enhanced CT scans including both unenhanced, corticomedullary, and excretory phases were

retrospectively collected between January 2008 and May 2021. For external validation of the ML models, cases of the publicly available KITS19 dataset were used (94-96).

3.2. Imaging acquisitions

The CT scans were performed on either a 16-slice or a 64-slice CT scanner in both studies. In the first study, portal venous phase series were retrieved, while in the second study, both unenhanced, corticomedullary, and excretory phases were collected. In the first study, during SWE measurement of the livers, the median stiffness values were given in kPa. The mean density values in Hounsfield Units (HU) of the anatomic liver segments were measured with manually placed circular regions of interest (ROI).

3.3. Segmentation and Radiomics analysis

In the first study, the entire liver volumes, the right and left lobes, as well as the anatomical liver segments, while in the second study, the entire tumor volumes were manually segmented using the 3D Slicer software. During radiomics feature extraction, isotropic voxel resampling with 1x1x1mm was applied, and the optimal bin width

value was defined based on the density range inside the analyzed volumes of interest (VOI) and was set to 25 in the first study and to 16 in the second study on kidney tumors. In the first study, Laplacian of Gaussian filters and wavelet filters were also applied to the images and higher-order radiomics features were also extracted to reveal fine and coarse textural changes of the liver parenchyma, while the second study was focused on the original images without image filter application. As a result, a total number of 1117 radiomics features were calculated in the first study and 107 in the second study.

3.4. Data preprocessing

In the first study, log-transformation was applied to the data to handle its skewness, and robust scaling was used for data standardization due to outliers, while in the second study, classical standardization was applied.

3.5. Feature selection

Both studies had multiple feature selection steps including filters and ML-based algorithms. In both studies, the highly correlated radiomics features were filtered out based on Pearson's correlation coefficients ($r > 0.95$).

In the first study, an FDR filter algorithm was also applied, while in the second study, the features that showed poor reproducibility against segmentation based on intraclass correlation coefficients (ICC) were excluded. In both studies, linear and non-linear ML-based feature selection algorithms were used, in the first study either a support vector classifier (SVC)-based, or a random forest (RFC)-based recursive feature elimination (RFE) algorithm was applied, while in the second study, either a least absolute shrinkage and selection operator (LASSO) or a tuned Relief (TuRF) algorithm was used to select the most predictive features.

3.6. Machine learning

In both studies, a linear and a non-linear ML pipeline was carried out and directly compared against each other including an RFC and an SVC model. The hyperparameters of the classifiers were tuned with a grid search algorithm using repeated five-fold cross-validation. The diagnostic performance of the models was evaluated based on the receiver operating characteristic curve (ROC) analysis.

4. Results

4.1. Patient population

In the first study, the final patient cohort consisted of 32 patients (16 females and 16 males), 11 with low-grade (SWE < 9.5 kPa) and 21 with high-grade liver fibrosis (SWE ≥9.5 kPa). In the second study, the final patient cohort of our Institution included 209 patients with 161 ccRCCs, 34 pRCCs, and 17 chRCCs. From the KiTS19 dataset, 69 eligible cases were identified including 50 ccRCCs, 13 pRCCs, and 10 chRCCs. The CT scans were performed by 19 different CT models.

4.2. Machine learning

During the first study on liver fibrosis, two sub-analyses were carried out. In analysis I., the cases were randomly split into training and test datasets with 1:1 ratio, while in analysis II., the train-test split was carried out based on the scanner types. The 16-slice scans were used as the training, and the 64-slice scans as the test set. In both analyses, the RFC and the SVC models showed good-to-excellent diagnostic performances. In analysis I., the RFC model constructed from 2 features had an area

under the ROC curve (AUC) value of 0.90, a sensitivity of 0.86, and a specificity of 0.78 on the test dataset, while the SVC trained on 18 features yielded an AUC of 0.76, a sensitivity of 0.93, and a specificity of 0.31. In analysis II., the SVC trained on 66 features slightly surpassed the RFC model built from 28 features on the test dataset with an AUC of 0.90 vs. 0.88, a sensitivity of 0.83 vs. 0.86, and a specificity of 0.95 vs. 0.92.

In the second study on differentiating between ccRCC and non-ccRCC subtypes, the RFC achieved the best performance with the 10 best corticomedullary phase radiomics features but showed clear signs of overfitting and failed to classify the external test cases (AUC of 0.66). Meanwhile, the SVC trained on the 10 most predictive corticomedullary phase radiomics features proved to be reliable and achieved AUCs of 0.873 and 0.834, and accuracies of 0.811 and 0.781, on the internal and external test datasets. Moreover, the performance of this model also proved to be comparable to that of an expert radiologist who achieved an accuracy of 0.795 (vs. 0.781; p=0.256) on the external test set.

5. Conclusions

Our first research on radiomics-based liver fibrosis assessment was among the first studies in the literature that assessed the feasibility of ML classification model building from three-dimensional radiomics features for the diagnosis of advanced-stage liver fibrosis. By carrying out the train-test split based on the types of CT scanners, we demonstrated the reliability of our ML models by testing them on independent internal test cases, and 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.

In our second study on the radiomics-based differentiation of kidney cancer subtypes, we successfully built an SVC-based ML model from the radiomic features of the corticomedullary phase CT scans that could distinguish ccRCCs from non-ccRCCs with an accuracy comparable to that of an expert radiologist. Additionally, we were able to demonstrate the generalizability of our ML model by successfully validating its performance on external test cases from the KiTS19 public dataset.

6. Bibliography of the candidate's publications

6.1. 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)**

6.2. 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**