SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PhD)

Examining the applicability of texture analysis on MR images

by Gergő Veres

Supervisors: Dr. László Balkay
Prof. Dr. Ervin Berényi

UNIVERSITY OF DEBRECEN
Doctoral School of Neuroscience

Debrecen, 2023

Examining the applicability of texture analysis on MR images

By Gergő Veres, Msc

Supervisors: László Balkay,PhD
Prof. Ervin Berényi, PhD

Doctoral School of Neuroscience, University of Debrecen

Head of the **Defense Committee**:	Prof. Béla Fülesdi, PhD, DSc
Reviewers:	Balázs Harangi, PhD
	Tamás Spisák, PhD
Members of the Defense Committee:	Prof. Csaba Cserháti, PhD,DSc
	Dávid Sipos, PhD

The PhD Defense takes place at the lecture Hall of Bldg. A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen, 26th of Sept, 2023 at 13:00

1. Background and objectives of the doctoral dissertation

Radiological imaging is widely used to diagnose neurological pathologies, characterise and confirm lesions in tumour staging, therapy planning, and to assess therapeutic response. Different anatomical lesions, pathological processes and functional changes can be assessed and characterised by imaging. However, in most cases, the analysis of imaging images is based on visual inspection only, and thus qualitative results are obtained. It would be much more appropriate to use testing methods and protocols that also provide quantitative results, because this would allow the objective characterisation of certain diseases or conditions in numerical terms. One of the current research goals of medical imaging is to extract as much quantitative information as possible from the diagnostic image obtained. This effort is called radiomics.

A radiomics data or index (RI) can be, for example, a statistical parameter such as skewness or entropy derived from a histogram of voxels associated with a given lesion. However, RIs that attempt to describe the spatial correlation or relationship of voxel values may be even more advantageous, and thus more directly related to tissue area heterogeneity. The latter parameters are also called texture indices (TI) or features, and typical representatives are, for example, the so-called gray level co-occurrence matrix (GLCM)-based parameters. In contrast to histological studies, which have specific difficulties and limitations in determining the heterogeneity of the entire lesion volume, radiomic analysis is a non-invasive method and can provide insight into the texture of the entire lesion. The analysis

of quantitative methods may be inherently given for a single modality, such as PET and CT scans using the SUV and Hounsfield Unit scales. However, in the case of MR imaging, the possibility of quantitative characterisation is relatively limited because the pixel values of the images obtained are highly dependent on the specific examination protocol, the sequence used and its settings.

The last 5 years have seen an explosion of research into the texture characterisation and analysis of various diseases and their diagnostic value. Although texture analysis (TA) is a promising image analysis method that quantitatively determines the 2D or 3D intensity patterns and correlations of voxels, it is currently not yet part of routine imaging diagnostics. Texture analysis may be able to identify pixel patterns, including possibly those that cannot be easily detected by the human eye. These methods have previously been successfully used on MR images for the scientific treatment of a number of neurological diseases, including brain tumours, epilepsy, Alzheimer's disease and multiple sclerosis. Statistical analysis of numerical data from more patient studies may allow us to develop a more accurate diagnostic picture and to discover correlations that may provide new information on differential diagnosis. A limiting factor in the implementation of radiomics is that the value of the calculated parameters can be strongly influenced by a number of measurement and examination conditions, such as the spatial resolution of the imaging modality. In this context, the resolution of histological specimens and state-of-the-art in vivo imaging differ in magnitude (histological sections: 10-4-10-3 mm; in vivo diagnostic imaging: 0.5-5 mm). In

addition, radiomics values can be affected by image reconstruction methods, the segmentation of the lesion used, and the so-called discretisation and normalisation procedures.

A number of published studies in the field of CT, MR and PET imaging have highlighted the challenges related to the reproducibility and reliability of radiomics features when different manufacturers and imaging devices, scan and reconstruction setups have been investigated. Recently, the Image Biomarker Standardization Initiative (IBSI) was introduced to standardize the definition of radiomic features. Currently, close to a thousand radiomics parameters have been proposed and defined in the IBSI guidelines. Their usability and reliability for diagnostic analysis for a given disease is challenging, even on imaging devices of the same type (e.g. MR) but with different technical capabilities (e.g. 1.5 - 3 T field). This is partly because subsets of RIs can be useful in diagnosing lesions in one organ (e.g. the brain), while they may perform poorly in other body regions.

Another fact is that the pixel discretization used, such as the fixed bin size (FBS) or fixed bin number (FBN) parameters, can significantly affect the RI values. It is already established in the literature that, for example, FBS discretisation may be preferable for PET imaging, but this is not yet clear for MR. It is well known that although CT and PET are quantitative methods, MRI is inherently not. The number and size of lesions in MRI scans is the most common quantitative measure used by radiologists, yet there is growing interest in measuring and analysing radiomic features. However, there is as yet no consensus on how different MRI systems and data acquisition protocols affect the robustness and

reliability of radiomic parameters for different pathologies.

Indeed, radiomics analysis is more difficult on MRI images because the intensity of the tissue imaging signal is affected by many more data acquisition settings than with other imaging modalities. Most radiomics parameters can be influenced not only by the magnetic field strength of the MR scanners, but also by setup parameters such as field of view, spatial resolution, reconstruction algorithm, number of repetitions, echo time (TE) and number of excitations (NEX or NSA), as well as signal-to-noise ratio (SNR). In addition, the images may contain a number of MR imaging-related artifacts (field distortion or so-called Gibbs artifacts) that are unknown in, for example, CT or PET.

In medical imaging, data acquisition, image reconstruction and optimal adjustment of image processing steps can usually be investigated with specific phantoms, but creating a reproducible heterogeneous phantom is a non-trivial task. Nevertheless, in recent years, several studies have attempted to investigate the reproducibility and reliability of texture parameters using some biological and some simpler physical phantoms. In the field of medical imaging, it is important to note that phantom construction requires the use of materials with stable properties over time, which can be easily reproduced and, if necessary, easily transported between imaging centres. These requirements are generally not met by the textured phantoms produced to date for MRI studies. 3D printing technology has recently become a promising phantom fabrication technique for almost all medical imaging modalities. It has the potential to enable rapid,

reproducible and cost-effective manufacturing. A wide variety of phantoms are now available, ranging from 3D models of actual patient data to mathematical (geometric) models. These phantoms have a fixed shape and are highly reproducible, giving the possibility to emphasise specific small spatial details in a shape. Despite this, dedicated 3D designed and printed MR radiomics phantoms have not yet appeared in the MR literature, which may be due to the fact that they do not have sufficient accurate geometry but must have sufficient heterogeneous contrast.
In the present work, using human MR images, we investigated which types and settings should be chosen for the most critical steps of image processing - segmentation, discretization and normalization options - in order to achieve the most reliable radiomic analysis. In addition, biological phantoms (kiwifruit, onion and tomato) were used to realistically test the robustness and reproducibility of the texture indices. In addition, we developed 3D printed models of two specific textures that can be loaded with MRI contrast agents to compare which phantom type is most suitable for testing the radiomics characteristics of MR images.

Objectives

In our research, we used MRI images to investigate the following specific problems:

1. texture indices can be sensitive to a number of factors, such as the protocol parameters of the imaging study used, as well as the segmentation of the selected pathological area. To analyse this, we

investigated the difference between free-hand and semi-automated elliptical segmentation in terms of derived texture values for human patient groups with glioblastoma, ischaemia and multiple sclerosis.

2. We also analysed the effect of discretisation of voxel values in segmented VOI using human and phantom studies. Three discretization methods were analyzed: the LRR, LAR and AR algorithms. We also analyzed whether there is an optimal bin value choice for each discretization.

3. Due to the non-quantitative nature of MRI, even for the same patient and anatomical region, voxel values will not be the same for repeated scans. For comparative cranial MRI studies, normalization methods are attempted, but their impact on radiomic analysis is not clear. Therefore, we also aimed at comparative texture analysis of normalized and non-normalized human MRI images.

4. Reliability and reproducibility of texture counting is a major issue when MR scanners with different spatial resolutions are used in a multicentre study. Accordingly, we have designed texture analysis of the same phantoms on both 1.5 and 3 Tesla MR scanners.

5. We also sought to answer the question of which phantom constructs might be appropriate to determine the reliability and reproducibility of texture analyses on MRI images as accurately as

possible. This has involved the use of texture phantoms produced by biological and 3D print techniques.

We planned to address the problems and questions raised through two independent pilot projects. The first was based on human MRI studies, while the second used only biological and 3D printed phantoms.

2. Materials and methods

2.1. MR methodology

In the human study project, we were able to retrospectively select MRI images from 71 patients who underwent contrast-weighted 3D T1 and T2 measurements on a 1.5 Tesla Siemens Magnetom Essenza scanner. Patients were collected into three subgroups according to their pathologies: ischemic stroke (N = 22), multiple sclerosis (N = 22) and neurological tumor (N = 27). Both T2-weighted axial and 3D T1-weighted axial measurements after contrast administration were performed according to local standard protocols for the diseases. In the project with phantoms, MR studies were carried out with biological (kiwi, tomato and onion) and 3D printed objects. All phantoms were examined in two clinical MRI machines in 2021, using MRI imaging equipment at the University of Debrecen Clinical Centre: one was a 3 T Philips Achieva and the other a 1.5 T Siemens Magnetom Essenza system. Three RF coils were used: an 8-channel cranial and a 32-channel neurovascular coil at 3 T, and a 6-channel cranial coil at 1.5 T field strength. For both

devices and for each coil, T2- and T1-weighted 3D coronal isotropic voxel measurements were performed according to clinical routine, with FOV, matrix size and resolution converted to the respective volume. Each measurement was repeated three times to test reproducibility. Each time, a new table position was set before each repetition, so that the phantoms under investigation were always positioned in the isocenter, the most homogeneous magnetic field. All tests were also performed at 1x1x1 and 2x2x2 mm isotropic voxel resolution.

3.2 Image processing

In the human studies part of our study, retrospective MR images were processed from patients with multiple scelrosis, some primary or secondary brain tumour or ischaemia. The imaging plane of all primary MR images was axial. We chose to process axial slices because they were acquired in the same plane aligned/designed to the same anatomical region in all patients: the imaginary plane connecting the anterior and posterior comissura was measured parallel to the imaginary plane.
To process the images of patients, we used the "Carimas 2.10" software for the evaluation of medical images, developed by the Turku PET Centre, for the Windows operating system. The segmentation of the affected areas was performed separately for each disease group. In all cases, a free manual or semi-automated elliptical segmentation method was chosen for the definition of VOIs placed on pathological areas. In addition, within the same brain volume, a healthy brain area of the same size as the pathological VOI was segmented on the opposite

side. For brain tumors or metastases, 1 VOI was placed in the pathology-affected area and 1 in the healthy area. However, in patients with ischaemia and multiple sclerosis, there are more, as these are often multifocal lesions. Here, we considered fresh ischaemic areas, previous vascular lesions and healthy brain in the case of ischaemic lesions. VOIs were placed in each of these areas with a size of at least 1 cm3. For SM patients, VOIs were similarly placed on active and inactive foci and healthy brain tissue. The voxel coordinates and values associated with the segments were saved in a text file format, which was read and processed in a subsequent step in the Matlab environment for texture analysis. The MR images of 3D printed and biological phantoms were segmented semi-automatically using the open source software platform "3D Slicer". In this case, Carimas was not chosen because of the different MR settings (different field strengths, spatial resolutions, weights and RF coils) and the large number of objects (360) to be segmented in three iterations, which would have made manual segmentation of each slice very time-consuming. Furthermore, since the phantom objects are bounded by air, automatic segmentation algorithms could be highly inefficient. An additional advantage of 3D Slicer is its ability to select any number of objects within a volume if segmentation can be easily performed automatically due to the adjacency conditions. Using this, after loading each MRI image, we placed a small sphere VOI in the same type of vegetable/fruit using the Slicer's segment editor, as well as a few additional small sphere VOIs in the background without objects. Then, using the "grow from seeds" built-in algorithm, the program automatically segmented the same fruit or vegetable. If

necessary, the resulting VOIs were manually corrected to exclude artificial products from part-volume or effect effects. For example, if the boundary zone between the fruit/veggie and the surrounding air was incorrectly defined, or if the most apical surfaces of some fruits/veggies were in contact with each other and segmentation artefacts resulted.

3.3 Normalization, discretization and calculation of texture indices

For the studies with phantoms, two different MRI machines were used, so due to the non-quantitative nature of MRI, the need to normalise the image data arose. Normalisation can also be considered as a necessary harmonisation for MRI that can improve the reliability of radiomic data. To define μ and σ as the mean and standard deviation of the image, the so-called $\mu \pm 3\sigma$ normalization was applied, which centred the voxel values with the defined standard deviation on the corresponding (zero) mean. Furthermore, by this procedure voxels outside the range $[\mu - 3\sigma, \mu + 3\sigma]$ were excluded from the original values. Next, all three discretization methods were applied to the segmented volumes for both human and phantom MRI images. We discretized using the LAR, LRR and AR algorithms, using several different bin widths (B) and bin values (D). For the LRR, D was defined as 8, 16, 32, 64, 128, 256, 512 and 1024, while for the AR and LAR procedures, B was defined as the following set: {1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100}. The values of D and B were defined to include all the possibilities that have been found in the literature so far.

The calculation of the radiomics parameters was the last image processing step. The number of radiomics data documented and applied by IBSI is currently close to two hundred. In our work, we investigated the most commonly used and promising radiomics indices in the literature. The 40 features selected for the processing of human MRI images were 18 GLCM, 11 GLSZM and 11 GLRLM-based texture indices, all calculated according to the IBSI definition. For the phantom MRI images, normalization was also applied, so that simpler statistical features of the segments could be compared, accordingly, the aforementioned 40 TI parameters were complemented with the most commonly used histogram-based statistical feature. The histogram indices included in the study were: minimum, maximum, mean, median and volume.

3.4 Statistical analysis

For the human studies, a non-parametric Wilcoxon rank-sum hypothesis test was also performed to compare TI between healthy and diseased areas using Matlab's rank-sum function. As null hypothesis, we assumed that the TI values of healthy and diseased areas are from the same median distribution. The test assumes that the two samples are independent, and also that the two samples may have different item counts. This test is also called the Mann-Whitney U test.
In addition, we calculated the Spearman correlation coefficient (R2) between the volume of each TI and the associated lesion for each patient and healthy VOI group and for both MRI contrasts.

For our work on biological and 3D printed phantoms, we calculated the mean and standard deviation of the RIs for each modality and for three repeated measurements, and from these we also determined the coefficient of variation (CV). For phantom measurements, we also determined the relative parameter difference (RPD) for measuring the relative RI difference of a given phantom between two different measurement settings. All calculations were performed using Matlab or Microsoft Office Excel software.

4. New scientific results of the thesis

4.1 Relationship between texture analysis and image-discretization methods on T1 and T2 weighted MR images of different patient groups

In our work based on human MR scans, we have investigated in detail the effect of discretization procedures transforming the voxel values of images on texture indices in T1- and T2-weighted MR scans and in three different types of neurological diseases (ischemia, tumor and metastasis). To the best of our knowledge, this is the first study comparing all three discretization methods (AR, LAR and LRR) for MR images and analysing the reliability of the TIs as a function of these methods. Our study revealed that the majority of the selected texture indices varied significantly when the discretization parameters (bin width or bin number) were changed within a predefined range. In this thesis, we selected 18 GLCM, 11 GLSZM and 11 GLRLM based texture features, those that have been previously identified as relevant parameters in the literature. Of the

18 GLCM-type TI, only two (NormInvDiff and Correlation) were not dependent on binning parameters, considering the three disease groups and all MRI data acquisition protocols. Although parameters that are independent of lesion volume and MRI data acquisition are desirable, independence also suggests that their use may not be appropriate for disease identification. In addition to this, it is important to note that all other texture indices (N = 38) are monotonic with several orders of magnitude variation depending on the binning values. However, the direction (nature) of the monotonicity and the range of values are typical for all texture indices. These characteristics may depend fundamentally on the mathematical expression used to define the texture indices, so a more specific analysis of the formulae would greatly help to investigate the applicability of the individual radiometric parameters. However, publications on the detailed analysis of mathematical expressions are not yet available in the literature. It is often of great importance to compare the diagnostic performance of texture indices in different MRI centres and in studies obtained using different cohort analyses. However, these comparisons are often clouded by the fact that the diagnostic power and reliability of different TI's can be highly dependent on the specific discretization method and its parameter (number of bins or width). Because of the clinical utility of TI's and the potential dependence between discretization parameters, it is essential that a study should only be compared with other studies that use the same number of bins (or bin width) and discretization technique. Very often, however, this is not the case, with many publications based on very different D values (LRR

technique, D = 256, 128, 32) and B choices (LAR technique, B = 5, 10). Our results showed that the magnitudes of the range of TI values do not depend on the disease groups (IS, MS and TU) or the imaging sequences used for analysis (T1 or T2 image contrast). We also found that, under AR and LRR discretization, all TI have exactly the same range and specific value, although B binning values spanned several orders of magnitude. The possible correlation between the VOI size of the lesions and the TIs is another problem related to the usability of texture indices. Accordingly, using human MR images, we performed a detailed analysis of the actual relationship between the VOI volume of lesions and the TIs, identifying 426 correlation diagrams and parameters. The correlation curves of TIs obtained for AR and LRR methods showed that most indices were weakly correlated with tumor volume. Note that it was not observed that the correlations of TIs and volumes were dependent on the patient groups or the T1 or T2 MR sequence used. For a more transparent presentation, we also calculated Spearman's correlation coefficients (R2) and used these to generate a colour-coded image showing R2 values for all three diseases studied and for T1 or T2 contrast. Cases with $R2 > 0.5$ ($|R| > 0.71$) were considered highly correlated, corresponding to light green, orange and even lighter colours. Thus, from the colour maps, it was clear that these values were mainly associated with the AR and LRR discretization methods and the GLRLM and GLSZM texture index groups for both segmentations. In addition to these, it is also noticeable that when using the LRR method and elliptic segmentation, several GLCM parameters also show higher correlation. Furthermore, the GLCM-based data

are even more highly correlated when LRR discretization and manual segmentation were used for the analysis. The strong correlation is undesirable in most cases, as the changes in TI should depend on texture, not volume. In other words, LRR discretization is detrimental to the correlations between lesion volume and TI, and this finding is consistent with some previous studies. Note that the R2 value for a given TI does not vary significantly as a function of disease or T1 or T2 contrast. The exceptions are the RP, LZE, LZLGE, LZHGE and ZP indices, which all belong to the GLRLM and GLSZM groups. Color-coded p-values of the hypothesis tests performed for LRR and AR discretization were also generated for the relationship between each patient and control area. The hypothesis was that the TI values of healthy areas would differ from those obtained for pathological areas. A higher number of texture traits was observed ($p\text{-value}<0.05$) when the AR technique was used instead of LRR for both segmentation methods. Furthermore, manual segmentation yielded fewer significantly different TI, and this was even more pronounced for LRR discretization. A possible explanation for this phenomenon could be that when using manual segmentation, the tissues outside the lesion also affect the calculation of TI (since the lesion boundary zone is mostly occupied in manual segmentation), as opposed to the case of elliptical VOIs, where only tissues inside the lesion are explicitly selected. Based on the AR discretization and any segmentation, the most promising TI's in MR images are Jvar, Dissim, SumVar, Contrast, NormInvDiff, ClusterProm, LGRE. Another interesting fact is that there are TI's that are correlated with volume (as shown in

Figures 25 and 26) but still can detect differences between patient and control areas based on hypothesis testing. We also compared the effects of two basic segmentation strategies, manual and 3D elliptical volume selection. We used two VOIs to compare healthy and diseased tissue. One VOI was placed in the pathological area and another VOI equal to the previous volume was placed in the healthy tissue on the opposite side. Our analysis showed that fewer statistically significant differences in texture parameters could be obtained with manual VOI segmentation. This can be explained by the fact that voxels at the interface may contain a mixture of surrounding tissue and tissue from the lesion due to inadequate spatial resolution. Accordingly, the calculated texture indices are also distorted. The 3D ellipsoidal segmentation method gave the TI that detects the most tissue discrepancies for any discretization.

4.2 Investigation of the robustness of radiomic characteristics in MR examinations, using 3D printed and biological phantoms

3D printing techniques offer a unique new opportunity to create textures and analyse the reliability of radiomic data derived from MR scans. In general, the results obtained with 3D printed Hilbert and QR code cubes showed good agreement with those obtained with biological phantoms, and could even be used to advantage for the development of optimal MR imaging protocols. This is a very important finding because many radiomics data are significantly influenced by MR field strength and the setup parameters of protocols; thus, harmonization of MRI systems may be critical. It is also

found that for high-resolution MR images (1x1x1 mm3 isotropic volume), it is possible to extract QR code information texturally, thus providing acceptable image quality for radiomic analysis. To the best of our knowledge, this fact has not been previously reported in publications evaluating texture or heterogeneity studies. Based on our results, we propose new 3D printed phantom models to verify and evaluate the applicability and reliability of radiomics methods in MRI scans. Three cube-shaped phantoms have been produced: a simple 5x5x5 cm3 3D Hilbert cube and two 3D QR codes with 5x5x4 and 4x4x3 cm3 dimensions. The 3D QR phantom could be a potential experimental tool for medical imaging, especially for the analysis and investigation of texture indices.

Based on literature data, hundreds of radiomics data can be calculated from a segmented image volume. In research projects, the usefulness and reliability of radiomics features are analysed on the basis of patient examinations and often using machine learning techniques, with appropriate predictive or prognostic models, to select texture indices that may already be of prognostic value. This process would be greatly facilitated if the number of available radiomics data could be reduced and, in addition, the reliability of the radiomics data could be easily checked in advance on several different MR machines. The biological phantoms we use and the 3D printed phantoms we have developed could be suitable for this purpose. In our studies, we used biological phantoms to verify whether the results obtained with 3D phantoms can be translated to real human measurements. We also aimed to use biological phantoms to analyse the robustness of radiomics data

obtained during imaging with different MRI devices. To this end, 3 tomatoes, 3 onions and 3 kiwis were selected and all biological and 3D printed phantoms were subjected to the same MR scans and radiomic analysis.

We have shown that the robust texture of the Hilbert cube does not change significantly at different MR resolutions (1 mm3 or 2 mm3); however, the image structure of the finer patterned QR cube is significantly degraded at 2 mm3 resolution. In the case of biological phantoms, textures characteristic of vegetables/fruits were well identified at both spatial resolutions, although images were definitely more blurred at the lower resolution (2 mm3). This image degradation, or "blurring" effect, is caused by the reduced sampling number in k-space and the associated loss of contrast in MRI. In other words, the performance of RIs can be highly dependent on the specific texture and spatial resolution of the underlying images.

Next, we investigated the effect of image normalization applied to MRI on radiomics calculations. We showed that CVs are generally reduced when normalization is applied, regardless of the biological phantoms involved. Furthermore, normalization increased the reproducibility parameters in almost all cases, expressed as % REP, and this is in good agreement with the results of other publications based on human studies. A similar trend (albeit with different specific values) was seen for each radiomics group (GLCM, GLSZM, GLRM, and histogram-based groups), as the reproducibility was higher in all cases when normalization was performed. Histogram-based parameters had little or negligible dependence on normalization, as would be expected for a first-order

statistical parameter. It was also seen that reproducibility was poor (CV>10%) for all MRI settings and all biological objects for the following RIs: Jmax, Energy, ClusterShade, HGRE, SRHGE, LRHGE, LZE, LZLGE, LZHGE. Due to poor reproducibility, these RIs were ignored for further analyses and only normalized data were used in all further evaluations.

In general, the CVs of all vegetables and fruits are lower at the lower field strength (1.5 T), which is in agreement with other observations. Another fact is that for all three biological phantoms, there were several MR test protocols at 3T field strength for which the CV was greater than 10%. This could be due to two reasons: first, the homogeneity of the magnetic field in 3T MRI is generally worse than at 1.5 T field strength. Second, the structure of biological phantoms can change slightly within a few days (this was the typical time of repetition between the two MRI machines). It is also noticeable that this effect is only significant for the 32-channel RF coil when using tomato and onion, while it is detectable for both RF coils when using kiwi. Probably due to the finer structure of kiwifruit, the T1 weighting or the lower spatial resolution parameters gave a different reproducibility compared to the T2 contrast or the 2 mm resolution setting.

It was also found that there was no difference between AR and LRR discretization after normalization of the MRI data.

However, AR discretisation was found to be better in a previous study analysing human MRI scans. However, the protocol of the previous study differed in some respects from the current study, as only one MR machine was used; thus, no normalization of the images was

necessary. Without normalisation, AR discretisation may therefore be preferable. Several studies also emphasise that AR discretisation can lead to more reliable radiomic analysis, but there is no longer necessarily a difference between the effect of AR and LRR for normalised MR images. However, it can be concluded from our present results and other published data that normalized MRI images can facilitate the selection of the most appropriate binning parameters.
For the statistical analysis, we also defined an RPD metric to determine the relative parameter difference value, for a given phantom and between two different measurement settings, to measure the difference in relative RI. We performed this analysis on both 3D printed and biological phantoms and found that RPD can vary by more than 20% for the same object when comparing MR systems with different field strengths. This fact did not depend on whether the spatial resolution was 1 mm3 or 2 mm3, nor on whether the MRI scan was T1 or T2 contrast. In addition, some GLSM-type RIs (GLNU, ZSNU and ZP) very often have high RPD values. It was also observed that the RPD "patterns" for 3D phantoms were not characteristically different from the image patterns of biological phantoms. So QR cubes perform similarly well to biological phantoms. In addition, the RPD parameters of the more robust Hilbert cube typically have a smaller value distribution than the RPD values of QR and all other biological phantoms, which means a higher degree of reproducibility of the Hilbert cube from a radiomics point of view. In other words, radiomics analyses suggest that the Hilbert cube is a simpler shape than any biological phantom.

The phantoms developed and applied in this thesis have several textural advantages and may be useful in radiomics analyses using MRI images. The phantoms have been designed and selected to ensure reproducibility and reliability not only during the measurements of the scan, but also to keep the objects in an unchanged physical state for as long as possible. The biological phantoms were considered stable for approximately two weeks, while the 3D phantoms did not change their physical properties during the entire duration of the study (approximately 1 year). Based on these results, it can be concluded that the phantoms used (both 3D printed and biological) could be beneficial to the radiomics community, which is striving to standardize both imaging protocols and radiomics analysis strategies. For the analysis of radiomics characteristics (reproducibility, reliability), the flexibility of 3D printing is a favourable method for producing new types of texture phantoms, as the use of identical materials and printing settings can ensure that imaging centres located far apart can produce fully equivalent objects to be imaged.

5. Summary

From the point of view of radiomic analysis, we compared the effect of three voxel discretization methods on brain MRI images of three different types of diseases (ischaemia, tumor and metastasis). We found that the values of all 40 TIs characteristically depend on the applied binning parameters, thus appropriate binning selection is not a trivial task. Therefore, it is very critical that comparative and large multicentre studies should only be performed with the same discretization and binning strategies. In addition, the above-mentioned

characteristic does not change if it is examined with a different type of disease or with a different MR sequence. The AR and LAR based methods give TI values that are similar to each other, but significantly different compared to LRR, when we consider the calculations of control and pathological brain areas. We also found that in general, TI's are weakly correlated with the volume of lesions, however, when LRR discretization method was chosen several GLCM based texture parameters showed higher correlation. Therefore, AR or LAR discretisation is recommended instead of LRR for brain MRI images. Furthermore, using semi-automated elliptical VOIs, we can obtain more significantly different TI between control and patient areas compared to using manual segmentation.

Using biological phantoms, we have investigated in detail how radiomics analysis in a multicentre MRI environment depends on the object (kiwi, onion, tomato), the field strength B0, the T1 or T2 weighting, the RF coil used and the voxel discretisation. Under test-retest experimental conditions, we found that normalization of MRI images significantly reduces the variability of radiomics indices. We show that 3D T1-weighted imaging, with lower field strength (1.5 T) and better spatial resolution (1 mm3), provides the most robust radiomics features. We also generated phantoms using 3D printing techniques, and the results obtained with the presented 3D printed Hilbert and QR code cubes showed good agreement with the results obtained with biological phantoms. With the Hilbert phantom, which has a more robust and therefore simpler texture than real tissue, fewer mismatched texture parameters can be filtered out, but significantly under-performing

radiomics parameters can be successfully identified. The analysis with the REP parameter also showed that QR code-like phantoms can be equivalent to biological phantoms when radiomic analysis is investigated.

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/174/2023.PL
Subject: PhD Publication List

Candidate: Gergő Veres
Doctoral School: Doctoral School of Neurosciences
MTMT ID: 10078931

List of publications related to the dissertation

1. **Veres, G.**, Kiss, J., Vas, N. F., Kallos-Balogh, P., Máthé, N. B., Lassen, M. L., Berényi, E., Balkay, L.: Phantom Study on the Robustness of MR Radiomics Features: comparing the Applicability of 3D Printed and Biological Phantoms.
Diagnostics. 12 (9), 1-24, 2022.
DOI: http://dx.doi.org/10.3390/diagnostics12092196
IF: 3.992 (2021)

2. **Veres, G.**, Vas, N. F., Lyngby Lassen, M., Béres, M., Krizsán, Á. K., Forgács, A., Berényi, E., Balkay, L.: Effect of grey-level discretization on texture feature on different weighted MRI images of diverse disease groups.
PLoS One. 16 (6), 1-18, 2021.
DOI: http://dx.doi.org/10.1371/journal.pone.0253419
IF: 3.752

List of other publications

3. Csoma, S. L., Bedekovics, J., **Veres, G.**, Árokszállási, A., András, C., Méhes, G., Mokánszki, A.: Circulating Cell-Free DNA-Based Comprehensive Molecular Analysis of Biliary Tract Cancers Using Next-Generation Sequencing.
Cancers (Basel). 14 (1), 1-13, 2022.
DOI: http://dx.doi.org/10.3390/cancers14010233
IF: 6.575 (2021)

4. Farkasinszky, G., Dénes, N., Rácz, S., Kis, A., Péli-Szabó, J., Opposits, G., **Veres, G.**, Balkay, L., Kertész, I., Mező, G., Hunyadi, J., Trencsényi, G.: In Vivo imaging of Ischemia/Reperfusion-mediated Aminopeptidase N Expression in Surgical Rat Model Using Ga-NOTA-c(NGR).
In Vivo. 36 (2), 657-666, 2022.
DOI: http://dx.doi.org/10.21873/invivo.12750
IF: 2.406 (2021)

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

5. **Veres, G.**, Dankó, Z., Balkay, L., Bágyi, P.: Diagnosztikai kijelzők, monitorok jellemzői és minőségellenőrzése.
Magy. Radiol. Online. 10 (2), 1-10, 2019.
DOI: http://dx.doi.org/10.5374/mro.2019.2.3.

Total IF of journals (all publications): 16,725
Total IF of journals (publications related to the dissertation): 7,744

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

25 May, 2023

DEBRECENI EGYETEM
Egyetemi és Nemzeti Könyvtár