Short Thesis for the Degree of Doctor of Philosophy (PhD)

Application of texture analysis in medical imaging

Béresová Monika
Supervisors: Dr. Balkay László, PhD
Prof. Dr. Berényi Ervin, PhD

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF MOLECULAR MEDICINE

Debrecen, 2019

Application of texture analysis in medical imaging

Written by Monika Béresová, MSc in Informatics

Doctoral School of Molecular Medicine, University of Debrecen

Supervisors: László Balkay, PhD
Ervin Berényi, PhD

Examination Committee:

Head:	László Csernoch, PhD, DSc
Members:	Balázs Benyó, PhD, DSc
	Katalin Goda, PhD

The Examination takes place at the library of Faculty of Physiology, Faculty of Medicine, University of Debrecen at 12:00 am, 22nd of October, 2019.

Head of Defense Committee:	László Csernoch, PhD, DSc
Reviewers:	Zsolt Bacsó, PhD
	László Czúni , PhD

Members of the Defense Committee:
Balázs Benyó, PhD, DSc
Katalin Goda, PhD

The PhD Defense takes place at Lecture Hall of Department of Dermatology, University of Debrecen at 14:00 pm, 22nd of October, 2019.

1. Background and objectives

Medical imaging is widely used in the oncological patient care to characterize lesions, for tumour staging, for therapy planning and to verify the therapeutic response. In case of tumour grading and staging, the anatomical and functional changes and histological abnormalities may also be specific in medical imaging. For these purposes, either a single modality such as X-ray, US (ultrasound), CT (computerized tomography), PET (positron emission tomography) and MRI (magnetic resonance imaging), or multimodal devices can be used, depending on the tumour type and extension, and the clinical question to be answered. In the general case, the analysis of medical images is based only on visual interpretation, thus reports provide only qualitative results. However, texture analysis (TA) is a promising method for quantitative analysis of tumour diagnosis.

Texture analysis is a calculation process on the whole or just a part of the image, which results in a single number characterizing the pattern, independently whether it is a painting, a medical image or a map. It is a very versatile method for analysing any kind of measured or virtual images. The concept of texture is very useful in different areas, as a macroscopic region, or a ROI (region of interest), which may have a constant or time-varying texture. The characterization of the texture plays an important role in classifying images, since different images may have different texture patterns.

In medical imaging, texture analysis is used to identify tissues, to discover differences between tissues, to diagnose pathological conditions, and a proper way to bring more quantitative information from a given image or a designated region. Nowadays, this goal is also referred as "Radiomics",

which aims at collecting and processing as much quantitative data as possible from medical images, and then discovering new relationships by the tools of statistical analysis and/or artificial intelligence. Utilizing the Radiomics results, personalized therapy can be more easily implemented, and the selection and effectiveness of the therapeutic decision can be improved.

The number of publications on texture analysis in the area of US, PET, CT and MR imaging is growing, and has been over 5,000 over the last 5 years. In year 2019, the number of proposed and recommended texture indices (TI) is greater than 100, thus the selection of appropriate and clinically useful TI is becoming a very difficult and complex task. Another problem of textural analysis is the definition and specific calculation of each TI data, because the mathematical definitions are not uniform. Recently an international consortium has been set up to resolve this issue, which edited a guideline for this purpose.

There are several types of the texture analysis. One of the most accepted classifications is the following:

1. Primary statistical texture analysis (global parameters). Histogram-based calculations are included, such as: minimum, maximum, mean, standard deviation, variance, skewness, kurtosis.
2. Secondary statistical texture analysis (local parameters). These include the calculations based on the so-called gray level co-occurrence matrix (GLCM) describing pixel neighbourhood conditions such as contrast, correlation, dissimilarity, energy, homogeneity and entropy.
3. Higher statistical texture analysis (regional parameters). These parameters are also based on secondary matrix generation, which can be divided into two major matrix

methods. In the GLRLM (gray level run length) method one can get the maximum length (run length) and distribution of pixels of the same value in a given direction. This method is direction dependent. Another method is GLSZM (based on size zone matrix) that finds areas associated with a given pixel value in a texture; in this case the number of areas with given pixels value in the image should be determined. This calculation procedure does not depend on directions.

However, several questions are still open relating the applicability, reliability and clinical relevance of texture parameters. In this regard one of the challenging areas is that of the TA calculating software used. Many programs are freely available for texture analysis (for example: MaZda, Matlab-based CGITA, GLCM-textureToo - Java development, ImageJ's Texture Analyzer implementation, TexRad, InterviewFusion, ABAQUS, and FRAGSTAT and Matlab), but there is no information about the reliability of these programs; the implemented algorithms may be different.

In addition, the texture calculations always require some re-sampling method (discretization or quantization) as a first step, which can be considered as a distinct normalization process. More discretization methods exist for heterogeneity analysis, but there is no adequate data about how the different quantization methods affect the reliability and prognostic value of the texture indices. During discretization, the original pixel values are converted to new values, where the new range (max-min value) and the number of new pixel values will be much smaller than those of the original pixel distribution. For example, if an original range is defined by the following interval [-1024, 3072], then by a discretization the scale can be re-sampled to the interval of [1, 64]. The run time (the

computing time) of the texture calculations is also reduced by discretization, because there is a quadratic interrelationship between the size of the texture matrices and the number of distinct pixels values. As a discretization step, two different approaches are the most common in the literature, the fixed bin number and the fixed bin width methods; although there are further rarely used methods.

When analysing PET images, the evaluation of tumour characteristics is based on the [^{18}F] FDG accumulation pattern, but regardless of which specific TI is calculated, the pre-processing discretization step plays an important role. This question has been examined more frequently for PET texture analysis, and less data are available for CT and MRI. In recent years, three fundamentally different resampling methods have been proposed and tested in the field of PET, posing a new challenge to determine the diagnostic and prognostic value of each TI based on different calculation. These techniques are the „lesion relative resampling" (LRR), "lesion absolute resampling" (LAR) and "absolute resampling" (AR) discretization methods.

2. Aims

Considering that the role and importance of texture analysis is increasing as well as the fact that many conceptual and methodological questions have been raised in the literature regarding the analysis of the texture, the following questions were addressed in the doctoral thesis:

1. It is unknown how the heterogeneity parameters are calculated by the programs intended to do texture analysis; that is, whether the implemented algorithms are similar or different. How do the TIs calculated from the same images depend on the software applied?
Our studies used Matlab, MaZda, CGITA and InterviewFusion software, and we compared the results with the manually calculated (gold standard) texture indices using several specially selected images.

2. Another aim was to find texture indexes that can be used to classify histologically diagnosed metastasis types from brain MR studies. Post-contrast MRI-T1 images containing brain metastases of lung and breast tumours were available for this purpose.

3. Currently one of the most challenging issues in TA is the proper selection and application of the discretization method. To further investigate this question, [^{18}F] FDG PET studies were performed to analyse the effect of three fundamentally different re-sampling methods (LRR, LAR, and AR) on several global, local, and regional texture parameters.

3. Materials and Methods

3.1. Reliability of heterogeneity parameters

3.1.1. Software for texture analysis, and the synthetic images used

First, we generated five different images of 12x12 matrix size. Then we arbitrarily inserted each matrix into a 128*128 image, and then converted to DICOM format. We also created a homogeneous constant matrix A0 (value=100) with the same matrix size. First, we calculated global (histogram-based) and local heterogeneity parameters of the A0, A1, A2, A3 and A4 matrices by four software packages (Matlab, MaZda, CGITA, InterviewFusion), and we compared the results with manual calculation.
For the GLCM matrix-based calculation we used the LRR discretization method with 8 and 64 bin values, in 2 preferred directions at 0° and 90°. Manual calculation was the "gold standard" method.

3.1.2. Texture analysis of images with the same appearance

For these tests 2 images were selected: a sagittal histological section of the brain, and Michelangelo's famous painting, the Creation of Adam. While the appearance and shape of these images is remarkably similar, the actual patterns and textures are different. For proper comparison we resampled the original photos to the same pixel size. The input of the TI evaluation programs is 2D image matrix, but the shape of a specific biologically segmented ROI area (the actual texture of which we are interested in) will never be rectangular.

Therefore, the ROI must be embedded in the original 2D image size so that the TI data calculated in this new image will be the same as the TI value of the pixel range within the ROI. For this, the discretization of the ROI was performed so that if we wanted to use *BN* bin values for the TI calculations, then the values inside the ROI were re-scaled to *BN-1* bin values. Then the pixel values outside the ROI (background area) were substituted by 0, making them different from the values within the ROI (between *1* and *BN-1*).
The element (1,1) of the GLCM, GLRLM and GLSZM matrices created from the re-scaled image contains the frequency of the background value, thus setting it to zero will limit the result of the TA to the area within the ROI.
In the reliability analysis, several other TIs were also calculated using a Matlab-based software tool (invoked by *GLCM_feature.m*). This tool provides the following 22 TIs: autocorrelation (autoc), contrast (contr), correlation_m (corrm), correlation_p (corrp), cluster prominence (cprom), cluster shade (cshad), dissimilarity (dissi), energy (energy), entropy (entro), homogeneity_m (homom), homogeneity_p (homop), maximum probability (maxpr), sum of squares (sosvh), sum average (savgh), sum variance (svarh), sum entropy (senth), difference variance (dvarh), difference entropy (denth), information measure of correlation1 (inf1h), information measure of correlation2 (inf2h), inverse difference normalized (indnc), and inverse difference moment normalized (indmnc). We also investigated how the parameters depend on the gray scale normalization (8, 16, 32, 64, …, or 1024 levels) in the case of 6 selected TIs (contrast, correlation, energy, homogeneity, dissimilarity, and entropy).

3.2. Texture analysis of the metastases of lung and breast tumours

In collaboration with the University of Valencia (Center for Biomaterials and Tissue Engineering Universitat Politècnica de València), we retrospectively processed the brain metastases of a total of 58 patients; twenty-six lesions originating from breast, and thirty-two from lung cancer. All MRI examinations were performed using a 1.5 Tesla magnet with a multichannel phased-array coil. The MRI protocol included T1-weighted axial images with gadolinium, and T2-weighted FLAIR axial images.

Altogether 846 slices of 58 brain lesions were manually delineated with Creaseg software. OIs were categorized into four subgroups, according to the following size ranges: 0–1935, 1936–3845, 3846–7700, and 7701–11,540 mm^2. 2D binary masks were created slice-by-slice from the segmented volumes of every metastasis. As a next step, 2D masks and MRI images were loaded to and displayed in Matlab; then TA was performed inside an appropriate ROI. The background value was set to *ROImin*-1. To analyse heterogeneity, histogram-based [minimum (Min), maximum (Max), mean, standard deviation (SD), variance, SD/mean, and median], and co-occurrence matrix features (contrast, correlation, energy, homogeneity and entropy) were calculated. The re-sampling of the segmented areas was done by LRR method using 64 bins, calculated in four directions (0, 45, 90, and 135°) at 1-pixel distance. Finally, with the average of the four COM elements, we derived the matrix on which texture analysis was applied. The above-mentioned calculations are based on both the post-contrast enhanced T1, and the so-called local binary pattern

(LBP) images. The pixel range of each LBP image is in the interval [1, 256] after the transformation.
One advantage of the calculation detailed above is that it is not sensitive to contrast changes.
Most often, the 3x3 pixel neighbourhood is tested, where the value of the central pixel is compared to the neighbouring pixels. Two new 3x3 matrices - the binary threshold and the weight matrices - are used to calculate the new value for the central pixel. The *binary threshold matrix* contains 1 or 0 for pixel values in the original matrix larger (≥) or less than the central element, respectively. The elements of the *weight matrix* contain the increasing powers of the 2 starting from the left upper element in a clockwise direction.
The value of the central pixel will be equal to the sum of the products of the corresponding threshold and weight matrix elements; which may be considered as if the threshold matrix represented a binary number (from the top left element clockwise).
First, a weighted 3D (2D converted to 3D) TI was calculated from the 2D Tis, using the number of pixels in each slice as weighting factor. Second, we computed the true 3D texture indexes in Matlab with the *cooc3d.m* function in 13 directions. The Kolmogorov–Smirnov test was performed to check the normality of the calculated data. For normally distributed data with similar SDs of both groups, the 2-sample t test, while for the rest the Mann–Whitney test was applied. We applied the Bonferroni correction for multiple comparisons. To assess the diagnostic power of each TI, ROC (receiver operating characteristic) curve analysis was used. Discriminant function analysis (DFA) was done separately for the histogram-based and co-occurrence matrix-based parameters.

3.3. Impact of intensity discretization on PET images

PET data of 58 patients (35 males, 23 females) with confirmed lung lesions were analysed retrospectively. PET images from all patients were acquired by a Philips Gemini scanner; the patients received an average of 325 ± 73 MBq [^{18}F]FDG. The delineation of the 63 tumour lesions was done manually with the Carimas software, controlled by a nuclear medicine specialist. Tumours below 25 ml were excluded from manual segmentation.

Three discretization methods commonly used in the literature were studied (LRR, LAR, AR). LRR is the most frequently used re-sampling method, representing every lesion on a completely new arbitrary scale, so the original uptake values with biological meaning are lost. Five bin numbers (16, 32, 64, 128 and 256) were selected for the LRR method. In the LAR method the *B* value represents a fixed bin width; for our experiment we chose $B \in \{0.05 \ldots 1\}$ in steps of 0.05. In case of the AR method, the number of bins were varied (in the range from 20 to 400) in order to use the same bin widths as investigated with LAR method. Texture analysis was performed using Matlab applications for GLCM, GLSZM and GLRLM data. In GLCM counting, we have focused on five parameters that were previously considered useful, such as contrast, correlation, energy, homogeneity, and entropy. For GLSZM and GLRLM methods, 11 - 11 parameters were calculated. For GLSZM: Zone Percentage (ZP), Small Zones Emphasis (SZE), Large Zones Emphasis (LZE), Gray-Level Non-Uniformity (GLNU), Low Gray-Level Zone Emphasis (LGZE), High Gray-Level Zone Emphasis (HGZE), Small Zone Low Gray Level Emphasis (SZLGE), Small Zone High Gray-Level Emphasis (SZHGE), Large Zone Low Gray Level Emphasis (LZLGE), Large Zone High Gray-Level Emphasis (LZHGE), Zone

Size Non-Uniformity (ZSNU). The GLRLM method: Run Percentage (RP), Short Run Emphasis (SRE), Long Run Emphasis (LRE), Gray-Level Non-Uniformity (GLNU), Low Gray Level Run Emphasis (LGRE), High Gray Level Run Emphasis (HGRE)), Short Run Low Gray-Level Emphasis (SRLGE), Short Run High Gray-Level Emphasis (SRHGE), Long Run Low Gray-Level Emphasis (LRLGE), Long Run High Gray-Level Emphasis (LRHGE), Run Length Non- Uniformity (RLNU).

3.3.1. Impact of intensity discretization on PET images

Lung lesions were delineated by 3 different methods. The first method is the previously mentioned manual one. The two additional methods were threshold-based selection of the voxels in a larger predefined volume using the following criteria: (i) >2.5 SUV value; (ii) >40% of the maximal SUV value. Lesions below 25 ml were usually excluded, since heterogeneity measures of such small lesions have been shown to be unreliable.

4. New scientific results

4.1. Comparison of texture analysis software using synthetic images

To achieve the first point of our goals, we studied the texture calculation mechanisms of four biomedical software packages; we compared the heterogeneity parameters provided by Matlab, MaZda, CGITA, and InterviewFusion to those calculated manually, using synthetic images. In the four inhomogeneous images (A1-A4), the distribution of low or high signal intensities was the same, so global data had to be similar for all four images. The A0 image, a homogeneous one, was only

considered as a reference standard. The structure of the inhomogeneous images A2-A4 was such that, when calculated horizontally, the local (GLCM) parameters had to be the same.

Global parameters by all four programs were very similar in case of all four matrices (A1-A4). With Matlab and CGITA the kurtosis values were significantly different from the manually calculated ones. For the rest of the global parameters the discrepancy was below 0.5% in all cases. Differences were even more pronounced in the case of local parameters (contrast, correlation, energy, homogeneity, dissimilarity and entropy) for the 8 and 64-level (D=8 and D=64) LRR discretization methods.

After 8-level discretization, the largest percentage discrepancy was seen for the HI values provided by MaZda and CGITA. For matrix A1 the contrast values showed 75-82% deviations with MaZda (both directions), and ~47% with CGITA. When the correlation HI was calculated from matrices A2, A3 and A4, we also saw a significant difference between software-aided and manual calculations. The energy parameter yielded a difference of ~50% with MaZda, and nearly 70% with CGITA. Homogeneity and dissimilarity could not be calculated by MaZda. In the case of Matlab, all HI parameters had a percentage difference from manual calculation less than 0.5%. The D=8 discretization was not available with Interview Fusion, since only one option D=64 discretization is implemented in the program. Homogeneity and entropy showed 5- 45% and 10-20% biases, respectively, for all matrices when the D=64 discretization was used in InterviewFusion. In case of CGITA, contrast and dissimilarity showed nearly 50% difference for matrix A1, dissimilarity showed a difference greater than 100% for matrix A2, and energy showed a nearly 60% difference for matrices A3 and A4. The largest percentage differences were seen in case of

CGITA. Matlab calculations (in both directions) resulted in a maximum difference of 2%. The D=64 discretization was not available in MaZda, and neither contrast, correlation, energy, nor dissimilarity could be calculated with the built-in modules of InterviewFusion.
In general, the global parameters showed nearly the same (less than 1% compared to the manual) in all 4 software packages, well suited to the original design of the images. In the case of local parameters, the most accurate texture index results were obtained by the parameters calculated in Matlab. Another fact is that when calculating the A2, A3, A4 matrices, the local parameters in Matlab, MaZda, and manual methods had the same values, which is consistent with the definition of synthetic image matrices.

4.1.1. Results of Texture Analysis of the of the same appearance images

When examining the reliability of the texture indices, we selected a sagittal histological section of the brain and Michelangelo's famous painting, the Creation of Adam. Interestingly, from the 22 co-occurrence-based texture parameters 12 indices did not show relevant differences (less than 10%). In addition, 5 indices depicted differences in the range of 10-20%, and only 5 parameters presented more than 20% dissimilarities. We also calculated the percentage differences between the contrast, correlation, energy, homogeneity, dissimilarity, and entropy values obtained from the two images with different D values. Our results confirm that the TI differences did not change when the gray level number was more than 64. In other words, the discretization bin number should be at least 64 for reliable, stable results. These findings suggest that the texture indices obtained from

images with similar overall appearance but with different content may have very similar values, so they are not really suitable for separating the types of images.

4.2. Texture analysis from lung and breast metastases in MR images

In this study we compared the results of texture analysis for metastases from 26 lung and 32 breast tumours, including MRT1 and LBP images, not only with 2D but also with weighted 3D and real 3D-based calculation.

In general, 2D analysis (separately for each slice) provided significant differences in more cases when calculated from the LBP maps than post-contrast MR (CET1) images. When analysing CET1 images, there were no significant differences in the 12 parameters between breast and lung metastases in case of the smallest ROI size 1 (0–1935 mm^2). Considering the histogram-based parameters, there were no statistically different ones in the smallest ROI size (0–1935 mm^2) for the LBP maps either, while only two statistically different parameters were found in the largest ROI size (7701–11,540 mm^2). LBP images resulted in much more significant differences. Contrast, correlation, energy, and homogeneity, calculated from the co-occurrence matrix, varied significantly across all four ROI sizes. The values of correlation, energy, and homogeneity are much lower for metastases from the breast, while contrast values were higher. LBP images resulted in much more significant differences. Co-occurrence matrix-based contrast, correlation, energy, and homogeneity were significantly different in all four ROI sizes. The values of correlation, energy, and homogeneity are much lower in case of metastases from the breast, while contrast values were higher.

The two metastatic groups (26 from lung vs. 32 from breast) were compared by the 3D computation, too. In the case of the weighted 3D TI, the significantly different heterogeneity parameters were the contrast, correlation, energy, homogeneity and entropy, which are almost the same as in the 2D analysis. The exception is the entropy that was significantly different for three of the four size groups by the 2D evaluation. In contrast to this, true 3D TA resulted only two statistically significant parameters from the LBP images, namely the energy and the entropy. In case of CET1 MR images we did not find any significantly different parameters.

In the next step, we performed ROC analysis to visualize the diagnostic value of the parameters of 3D TA, as well as DFA analysis from all three texture calculation methods (2D, weighted 3D, and real 3D). All these ROC curves showed statistically significant discrimination ($p<0.05$). When checking the conditions of DFA, from among the 12 tests (three data sets: 2D, 2D → 3D, 3D; two matrix types: norm, LBP; two sets of variables: co-occurrence based, histogram-based), the linear model could properly classify the origin of the metastases in 44–72% of the cases. For every data set, the highest percentage of correctly classified cases could be obtained from the LBP maps using the combination of co-occurrence-based parameters: 2D: 65.6%, 2D → 3D: 63.8%, 3D: 72.4%.

4.3. Investigation of the effect of discretization using FDG PET images

To achieve the last point of our goals, we first analysed the contrast, correlation, energy, homogeneity, dissimilarity and entropy parameters that we had previously found reliable. The results of the three discretization methods (LRR, LAR, and AR)

showed how the texture parameters change with different bin values and different bin widths. In case of LAR and AR we used box plots, while we demonstrated the results of the LRR method by histograms. The values of investigated textural parameters show two distinct trends: both LAR and AR show a monotonic trend with the applied bin width, but in case of the LRR method, the histogram of TI shows a shift on the x-axis when changing the bin width. All the five GLCM texture parameters changed dramatically (>100%) when the SUV bin width increased from 0.05 to 1 (LAR and AR method), or when the number of bins decreases (LRR method), except the correlation which showed high stability. Analysing the manually delineated data set, we found the same tendency for the 11 GLSZM and the 11 GLRLM-based features: all of them changed greatly when the SUV bin width changed from 0.05 to 1 (LAR and AR), or when the number of bins altered (LRR method - 16, 32, 64, 128, 256). We can state that the application of the LRR method will result in a lesion-dependent bin width.

As already mentioned for the LRR method, the current bin width (B) was different for each lesion, and nor the same as the20 values (0.05, 0.1, 0.15,… 1) used for the for LAR and AR discretization, so we determined separately the B values for LAR re-sampling. Fe found that with increasing bin width, the individual distributions were characterized by a lower pixel value scale. In each case, there were overlaps between the groups.

Pairwise correlation scatter plots were created for the three different quantization methods. The results from AR method showed high correlation with LAR, but not with the LRR method. The 22 regional GLSZM and GLRLM features showed very similar tendency: the LAR and AR values correlated well with each other, but only weekly with the LRR.

We also investigated how the lesion segmentation method affects the results presented above. In addition to manual segmentation, we also analysed two threshold-based methods commonly used in the literature. In every case, a VOI containing the lesion had to be manually selected first, then the pixels >2.5 or SUV(pixel)> 0.4 * SUVmax were selected. Both threshold-based lesion volumes were apparently smaller than the manually delineated lesions. Considering all the lesions, the smallest and the largest volumes were 0.77 ml and 347 ml, respectively. While several small lesions (<20 ml) were generated during the segmentation process, in this study we did not exclude lesions based on volume sizes.
We did this to ensure that in this current discretization analysis the results do not depend on the size of the possible limited lesion.
We can summarize that the individual distributions and correlations were very similar to the cases of manual segmentation, irrespective of the thresholding method.

5. Summary

Methods of texture analysis can provide further quantitative information for the evaluation of medical images, which can help with the separation and classification of healthy and abnormal areas during the reporting and the design of therapies. Over the past decade, it has become an outstanding aspiration and goal in the process of medical imaging to extract as much additional information and numerical data as possible from the images. As already mentioned in the introduction, this new effort is called “Radiomics”, which also aims at investigating the prognostic power of the new parameters.

Therefore, texture analysis is part of Radiomics in the present years, and it is a critical challenge to prove that the results of the selected TI calculation are reliable and reproducible.
Since TA is a multi-step (segmentation, discretization, TI selection, and 2D or 3D implementation), complex calculation method, so it is very important to examine each step and its effect. In connection with the reliability and use of the texture indices, we investigated three interrelated issues in my doctoral thesis:

- Examining various texture calculation programs with synthetic images
- Comparing the textures of metastases from lung and breast tumours in MR images
- Analysis of the effect of three different pixel discretization methods using FDG PET images.

It is well known that a very large number of different texture indexes have been defined in recent years (> 100), so it is also very important that their mathematical definitions be used in the same way in research applications.
Standardisation was established in 2017, by publishing a document titled “Image biomarker standardisation initiative”, and since then those definitions are expected to be used in this field. When starting to work on my doctoral thesis (in 2015) such document and description did not exist yet, so it was even more relevant to check whether each texture analysis program gives the same results when applied to the same images.

There are several software tools frequently used in the medical field, including Matlab, CGITA, InterviewFusion, MaZda, Textrad and ImageJ, but in many cases the exact formulas that they apply are unclear. We studied the texture

calculation of four biomedical software packages; we compared the heterogeneity parameters provided by Matlab, MaZda, CGITA, and InterviewFusion to those calculated manually, using synthetic images (A0, A1, A2, A3 and A4). For the comparisons, manual calculation following the recommended formulas was selected as the "gold standard".
The following local parameters were included in the study: contrast, correlation, energy, homogeneity and entropy, and the LRR discretization method for bin values of 8 and 64 was generally used in 2 distinct directions, at 0 ° and 90 °.
The histogram-based parameters provided by the four software packages were nearly the same as those of the manual calculation, the differences being less than 0.5%.
When comparing the local parameters, there were larger differences: in the case of CGITA, InterviewFusion and MaZda the maximum deviation in some cases reached 100%.
We obtained the smallest (<0.5%) deviation from the values of the manual calculations for the parameters by the Matlab software. The differences of the results may be explained by the differences between the implemented calculation methods, and the averaging algorithms of multidirectional co-occurrence matrices.

In our further work, we used 2D and 3D texture analysis for the comparison of the metastases of lung and breast tumours, both for T1 post-contrast MRI images and for LBP images.
Our results showed differences between the types of metastases mainly on the LBP maps, but these differences were significantly dependent on the applied texture analysis method (2D or 3D). The results of the 2D analysis showed significant differences for only four parameters (contrast, correlation, energy and homogeneity) in all four previously selected size ranges (0-1935, 1936-3845, 3846-7700, 7701-

114040 mm^2). Correlation, energy, and homogeneity, calculated from LBP images, gave higher values for breast metastases (compared to lung metastases), while the contrast values were lower. Post-contrast MR images showed more significant differences in the 2D analysis compared to the results obtained from the LBP images. Even in the case of weighted 3D texture analysis, LBP-based parameters, namely contrast, correlation, energy, homogeneity and entropy, gave more significant differences.
On the contrary, from real 3D texture calculations only two parameters, energy and entropy showed significant differences.

Thus we found several heterogeneity parameters that showed differences between metastatic brain lesions of different primary tumours.
The parameters calculated from the LBP maps were more sensitive to the origin of the metastases.
Our results show that the relevance of the texture indices may depend on the 2D or 3D technique used, since these methods significantly influence the results of the texture calculations.
In our study, the most promising texture indices were entropy and energy, which showed significant differences between the LBP images by all three computation methods.

Finally, we examined the role of different discretization methods using three completely different TI counting methods (GLCM, GLSZM and GLRLM based calculations) for the same [18F] FDG PET images. The discretization methods were LRR, LAR, and AR. We showed that a given TI can vary within a very wide range (up to 100%) by changing the actual SUV bin width (B) or bin value (D) in the discretization step. Comparative analysis was performed on the same population by examining a total of 27 texture parameters (GLCM-5, GLSZM and GLRLM-

11-11), using different bin widths and bin values. The TIs showed a monotonous dependence on B and D values in each case. Interestingly, at a given bin width or at a given bin number, the values of the lesions were in a relatively narrow range. In addition, the value of the 27 texture parameters changed dramatically (> 100%) if the bin width of SUV increased from 0.05 to 1 (LAR and AR method), or the number of bin values decreased from 256 to 16 (LRR method). It seems that the calculated values of the given texture characteristics can vary over a wide range, depending on how you change the bin width or the value of the bin.

In principle, the segmentation method may fundamentally influence the volume and the texture of the lesion, but when a fixed-threshold segmentation (SUV_2.5 or 40%_SUVmax) was used before the LAR, AR, LRR discretization, the tendency of the resulting values did not change compared to the cases of manual segmentation.

In the case of the LRR method, the actual bin width varied depending on the lesions, while for the AR and LAR methods, the B value was the same, within the range of [0.05,1].

However, this fact does not mean that one of the quantization methods would be better in this respect than the others, but rather refers to their different behaviour.

Generally we can say that the different types of discretization may have a big impact on the characteristics of the calculated texture, as well as the reliability and reproducibility of the given TI.

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/212/2019.PL
Subject: PhD Publikációs Lista

Candidate: Mónika Béres
Neptun ID: QM6NHY
Doctoral School: Doctoral School of Molecular Medicine
MTMT ID: 10054298

List of publications related to the dissertation

1. Forgács, A., **Béres, M.**, Garai, I., Lassen, M. L., Beyer, T., DiFranco, M. D., Berényi, E., Balkay, L.: Impact of intensity discretization on textural indices of [18F]FDG-PET tumour heterogeneity in lung cancer patients.
   Phys. Med. Biol. [Epub ahead of print], 1-20, 2019.
   DOI: http://dx.doi.org/10.1088/1361-6560/ab2328
   IF: 2.665 (2017)
2. **Béres, M.**, Larroza, A., Arana, E., Varga, J., Balkay, L., Moratal, D.: 2D and 3D texture analysis to differentiate brain metastases on MR images: proceed with caution.
   Magn. Reson. Mat. Phys. Biol. Med. 31 (2), 285-294, 2018.
   DOI: http://dx.doi.org/10.1007/s10334-017-0653-9
   IF: 1.832 (2017)
3. **Béres, M.**, Forgács, A., Bujdosó, B., Székely, A., Varga, J., Berényi, E., Balkay, L.: Comparing the reliability of biomedical texture analysis tools on different image types.
   Acta polytechn. Hung. 15 (7), 29-48, 2018.
   IF: 0.909 (2017)

DEBRECENI EGYETEM
Egyetemi és Nemzeti Könyvtár

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

List of other publications

4. Jakab, A., Emri, M., Spisák, T., Szemán-Nagy, A., **Béres, M.**, Kis, S. A., Molnár, P., Berényi, E.: Autistic Traits in Neurotypical Adults: correlates of Graph Theoretical Functional Network Topology and White Matter Anisotropy Patterns.
PLoS One. 8 (4), e60982, 2013.
DOI: http://dx.doi.org/10.1371/journal.pone.0060982
IF: 3.534

5. Jakab, A., Molnár, P. P., Bogner, P., **Béres, M.**, Berényi, E.: Connectivity-based parcellation reveals interhemispheric differences in the insula.
Brain Topogr. 25 (3), 264-271, 2012.
DOI: http://dx.doi.org/10.1007/s10548-011-0205-y
IF: 3.671

Total IF of journals (all publications): 12,611
Total IF of journals (publications related to the dissertation): 5,406

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.

30 May, 2019

DEBRECENI EGYETEM
Egyetemi és Nemzeti Könyvtár