CZECH TECNICAL UNIVERSITY IN PRAGUE

FACULTY OF BIOMEDICAL ENGINEERING
Department of Biomedical Technology

Extrakce Příznaků a Rozpoznávání v Ultrazvukovém Obraze se Zvýšeným Kontrastem

Feature Extraction and Pattern Recognition for Contrast-Enhanced Ultrasound Images

MASTER THESIS

Study Program: Biomedical and Clinical Technology
Study Branch: Biomedical Engineering

Author: Zuzanna Anna Magnuska

Supervisors: doc. Ing. Martin Rožánek, Ph.D., CTU Prague
Univ.-Prof. Dr. med. Fabian Kiessling, RWTH Aachen

Kladno, 2018

Czech Technical University in Prague, Faculty of Biomedical Engineering

Department of Biomedical Technology

Academic year: 2017/2018

Diploma thesis assignment

(Master project thesis assignment)

Student:	**Zuzanna Magnuska**
Study branch:	Biomedical Engineering (CEMACUBE)
Title:	**Feature Extraction and Pattern Recognition for Contrast - Enhanced Ultrasound Imageým konstrastems**
Title in Czech:	Extrakce příznaků a rozpoznávání v ultrazvukovém obraze se zvýšeným kontrastem

Instructions for processing:

Process pictures from ultrasound measurement. Realize a reduction of noise in contrast - enhanced ultrasound images and implement an automated tumour detection and segmentation. Extract imaging bio-markers. Realize automated tumour model differentiation. Discuss achieved results.

References:

[1] I Elyasi and S. Zarmehi, Elimination Noise by Adaptive Wavelet Thresholding, International Journal of Electrical, Compute, Energetic, Electronic and Communication Engineering, ročník 3, číslo 8, 2009
[2] P. Viola and M. Jones, Robust Real-time Face Detection, International Journal of Computer Vision, ročník 57, číslo 1, 2004

Validity of assignment until date:	20.09.2019
Supervisor of diploma thesis:	doc. Ing. Martin Rožánek, Ph.D.
Consultant of diploma thesis:	Univ.-Prof. Dr. med. Fabian Kiessling

....
Head of Department

....
Dean

In Kladno, 19.02.2018

Declaration

I hereby declare that I have completed this thesis with the topic: *Feature Extraction and Pattern Recognition for Contrast-Enhanced Ultrasound Images* independently and I have attached a full list of used references and citations. I do not have a compelling reason against the use of the thesis within the meaning of Section 60 of the Act No.121/2000 Coll., on copyright, rights related to copyright and amending some laws (Copyright Act).

In Kladno, date 18.05.2018

…..…………………………………
Zuzanna Anna Magnuska

Acknowledgements

I would like to extend my thanks to people who made the writing of this Master thesis possible and to those who supported me in its creation process:

To Univ.-Prof. Dr. med. Fabian Kiessling, thank you for the opportunity to do my Master thesis in the Institute for Experimental Molecular Imaging and the possibility to continue my work at this excellent Institute during my Ph.D. studies.

To Benjamin Theek, thank you for your constant guidance and supervision, for supporting me at every stage of this Master thesis with your experience and broad knowledge of many scientific fields and for the openness to new ideas and to challenges.

To Susanne Golombek, thank you for your willingness and the time that you contribute in creation of this master thesis.

To Karolin and Milita thank you for your help and assistance during all this time and for the support at the moments of doubt.

To Members of CEMACUBE, thank you for giving me the opportunity to take part in this Master study program and for the two years of great adventure.

To my beloved friend, Dominika, thank you for being always there for me, for your time and support even from the great distance.

And finally, to my beloved parents, Stanisław and Dorota and my grandmothers, Bożenna and Teresa. Thank you all for being always there for me in the good and bad moments, for sharing my successes, worries and doubts, for your understanding and support for each new challenge I set for myself. Thank you because without you all of this would not be possible.

Abstrakt

Vývoj v oblasti výpočetní techniky přispěl ke zlepšení v oblasti rychlé extrakce četných kvantitativních příznaků z medicínských obrazů. Radiomika je poměrně nový vědní obor, který se zabývá identifikací korelace mezi zobrazováním a klinickými daty. Primárním cílem radiomiky je zpřesnění diagnostiky v medicíně pomocí vývoje klasifikačních modelů, které podpoří lékaře při hodnocení diagnózy a terapie. V současné době je radiomická analýza prováděna zejména u obrazů z výpočetní tomografie (CT) a magnetické rezonance (MR). Pouze několik studií bylo provedeno s ultrazvukovými obrazy. Cílem této práce je návrh pracovního postupu, nezávislého na uživateli, k provedení radiomické analýzy ultrazvukových obrazů, který by měl odlišit tři typy nádorů (MLS, A431, and A549). Navržený algoritmus byl rozdělen na následující kroky: odstranění šumu z ultrazvukových obrazů použitím diskrétní Waveletovy transformace (DWT), vytvoření algoritmu automatické detekce nádoru a jeho segmentace, extrakce biomarkerů z obrazů a realizace automatické diferenciace modelových nádorů. Odstranění šumu bylo nejefektivnější při použití bior3.7 Waveletovy vlnky a měkkého prahování s individuálním Bayesovským prahem (BS) aplikovaným pouze do třetí úrovně dekompozice (DL), (MSE=6.87 and PSNR=87.96). Navržený algoritmus pro automatickou detekci a segmentaci nádorů, který je založen na Viola-Jonesově algoritmu a aktivním konturování, dosáhl vysoké přesnosti detekce (89 % správně detekovaných a extrahovaných nádorů). Navržený algoritmus extrakce a selekce obrazových biomarkerů určuje radiomickou signaturu (RS), která se skládá z následujících tří nezávislých příznaků: medián (založený na intenzitě), korelace (textový příznak) a zdůraznění krátkých sekvencí (waveletův příznak). Algoritmy založené na stromovém učení a k nejbližších sousedů dosáhly nejlepších výsledků klasifikace nádorů (p=0.780 [95 % CI 0.550-0.910]). Dosažené výsledky naznačují, že radiomická analýza ultrazvukových obrazů může být použita ke klasifikaci nádorů a měla by být uvažována i pro klinické použití.

Klíčová slova: radiomika, automatická segmentace, umělá intelligence, strojové učenie, obrazová analýza

Abstract

Through the years, development in computing contributed to the improvement in rapid extraction of numerous quantitative features from medical images. Radiomics is a newly established scientific field that concentrates on the identification of a correlation between imaging and clinical data. The primary goal of radiomics is to improve precision medicine by developing the classification models, which are supporting physicians in diagnosis and therapy evaluation. Nowadays, a radiomic analysis is carried out mostly with the use of images obtained with computed tomography (CT) and magnetic resonance imaging (MRI). Only a few similar studies have been conducted with the use of ultrasound (US) data. The aim of this thesis was to develop a user-independent workflow to conduct a radiomic analysis of US, which should differentiate three tumor models (MLS, A431, and A549). The developed algorithm was subdivided into the following subitems: denoising of US images with the use of discrete wavelet transform (DWT), create an algorithm for the automated tumor detection and segmentation, extraction of imaging biomarkers and perform an automated tumor model differentiation. The most efficient noise removal was achieved with the use of bior3.7 wavelet and soft thresholding with individual Bayes Shrink (BS) applied only to 3rd decomposition level (DL) (MSE=6.87 and PSNR=87.96). Furthermore, the proposed algorithm for the automated detection and segmentation of tumors, which is based on the Viola-Jones algorithm and an active contour segmentation, achieved a high accuracy (89% of correct tumor detections and extractions). The developed imaging biomarkers extraction and selection algorithm identified the radiomic signature (RS) consisting of the following three independent features: median (intensity-based feature), correlation (textural feature) and short run emphasis (wavelet feature). The Simple Tree and k-NN learning algorithms achieved the best results of tumor model classification (p=0.780 [95% CI 0.550-0.910]). Altogether, the results indicate that a radiomic analysis of US images can be performed to classify tumors, and should be considered for clinical translation.

Keywords: radiomics, automated segmentation, artificial intelligence, machine learning, image analysis

Table of Contents

List of Abbreviations

AI	Artificial Intelligence
ANNs	Artificial Neural Networks
AS	Automated Segmentation
BB	Bounding Box
BS	Bayes Shrink
CCs	Cascade Classifiers
CCC	Concordance Correlation Coefficient
CEUS	Contrast-Enhanced Ultrasound
CI	Confidence Interval
CT	Computed Tomography
CWT	Continuous Wavelet Transform
DL	Decomposition Level
DWT	Discrete Wavelet Transform
FOS	First Order Statistics
FT	Fourier Transform
gBS	Global Bayes Shrink
GLCM	Grey Level Co-occurrence Matrix
GLRLM	Grey Level Run Length Matrix
HH	Diagonal Detail Coefficients Matrix
HEM	Horizontal Edge Map
HG	Horizontal Gradient
HL	Horizontal Detail Coefficient Matrix
HOG	Histogram of Oriented Gradients

HPF	High-pass Filter
iBS	Individual Bayes Shrink
IDWT	Inverse Discrete Wavelet Transform
k-NN	Nearest Neighbors classification algorithm
LL	Approximation Coefficients Matrix
LBP	Local Binary Patterns
LH	Vertical Detail Coefficients Matrix
LPF	Low-pass filter
MB	Microbubbles
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
NS	Normal Shrink
OCT	Optical Coherent Tomography
PPs	Positive Pixels
PSNR	Peak Signal to Noise Ratio
ROI	Region of Interest
RMS	Root Mean Square
RS	Radiomic Signature
SAR	Synthetic Aperture Radar
sBB	Small Bounding Box
SOS	Second Order Statistics
ST	Sensitivity Threshold
SVM	Supported Vector Machine
UCA	Ultrasound Contrast Agents

US	Ultrasound
VEM	Vertical Edge Map
VG	Vertical Gradient
WT	Wavelet Transform

1. Introduction

'Use a picture. It is worth a thousand words' – this quote published in one of the American newspapers at the beginning of the 20th century refers to a case that a simple image can represent a complex idea**.** Imaging is a standard clinical practice used in cancer diagnostics. However, sometimes the visual interpretation of a lesion is not enough. Often to get better characteristic of a tumor tissue the procedures like a biopsy or invasive surgeries are performed. The recent studies have proven that cancer diagnostics could be supported by exploring the content of medical images [1-3]. Therefore, the image is treated as a source of information about the lesion. Images are composed of small units called pixels. Each of them carries valuable traits that can be extracted, analyzed and classified (Figure 1). Pixels merit the particular attention since the interpretation of their content makes an image more than a "picture." The data characterization-algorithms, like image processing and image analysis, are used to explain the information enclosed in medical images. Comprising processing and analysis of medical images in standard diagnostics procedures could make them less invasive and more individual patient oriented. Through the years, development in computing contributed to the improvement in the extraction of quantitative features from medical images. Hence, it resulted in the creation of a new scientific field known as radiomics.

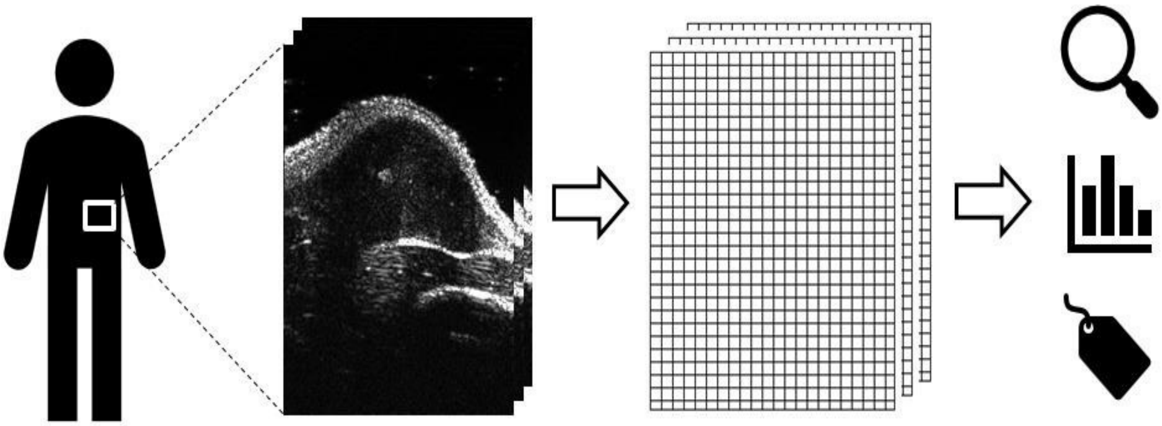


Figure 1: Exploring the content of medical images. Medical imaging is a common practice in modern diagnostics process. Next, to the qualitative evaluation of anatomical features, the quantitative analysis can be performed as well. Therefore, the images are handled as a matrix of pixels to calculate distinct imaging biomarkers. These imaging biomarkers can be used for lesion characterization or classification problems

Radiomics is the process of converting biomedical images into minable data sets. The term itself arises from the connection of the word radiology with the ending omics, which refers to the process resulting in the generation of high throughput qualitative data from a single sample [4, 5]. The power of the radiomics approach is hidden in digital radiological images generated for almost every patient. Identifying correlation between extracted radiomics features and other clinical data could improve the accuracy of cancer diagnostics. Therefore, classification models of specific tumor types can be created to support the characterization and prediction of cancer at low costs [2]. The general radiomics procedure consists of four consecutive steps (Figure 2). Each part constitutes independent procedure; however, together they create a robust differentiating framework.

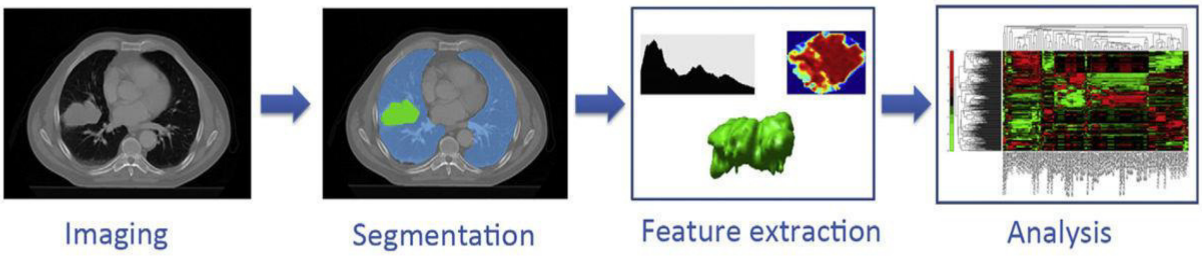


Figure 2: Stages of radiomic analysis. At first, the imaging of tumor is done. Secondly, the region of interest is segmented in the acquired images. In the next step is the features connected with lesion shape, intensity and texture are extracted. In the end, the mined data is used for building the tumor classification model [5].

1.1. Image Acquisition and Processing

Imaging is the first component of the radiomic analysis. A vast variety of medical imaging modalities like computed tomography (CT), magnetic resonance imaging (MRI), or ultrasound (US) are frequently used in nowadays clinical practice**.** Currently, the radiomic analysis is mostly done for CT and MRI images due to their high stability and reproducibility in comparison to US [6]. Even though, a few authors conducted studies concerning the full radiomics [1, 3] or the textural features analysis for tumors in US images [7, 8]. US imaging is used in many branches of medicine (e.g., cardiology). It gained its popularity due to safety, feasibility and low-cost. Moreover, portability and real-time or functional imaging make it a necessary medical device [9]. US imaging is based on the propagation of consecutive US waves through body tissues and reconstruction of the reflected US echoes as the images [10].

The US echo pulses are generated and collected by the transducer. The main structural elements of US probe are piezoelectric crystals. Piezoelectric crystals change their length while being subjected to the electric current of different polarity (Figure 3 A). This reciprocal phenomenon is called piezoelectric effect. The application of the high-frequency alternating current causes oscillation of the piezoelectric crystals, thus, the ultrasound waves are generated (Figure 3 B). The penetration depth and resolution of US are strongly dependent on the frequency of the US (Figure 3 C). One can see, that the US with a higher frequency has higher resolution. However, in this case, the penetration depth is low. Therefore, visualization of only superficial structures is possible. On the other hand, the US with low frequency and low resolution enable imaging of deeper located body tissues. The standard frequencies of medical US varies typically between 1 to 20 [MHz] [10]. The body tissues are characterized by a physical property called acoustic impedance. It depends on the density of tissue and the velocity of US wave propagation in it [10]. The difference in acoustic impedance among different tissue interfaces causes partial US reflection. Therefore, various tissues can be discriminated in US images. After being reflected, the returning echoes fall on the piezoelectric crystals generating an analog signal. However, the obtained signal is of a low amplitude. Thus it is amplified with a high-frequency amplifier and consequently passed through signal conversion, processing, and visualization modules. The collected signals are coded as gray-level images. Currently, B-mode ultrasound is mainly used in clinical practice [10]. The US, due to its high sensitivity to contrast molecules and high spatial and temporal resolution, is a reliable tool not only for diagnostic purposes. Introducing ultrasound contrast agents (UCA) broaden the spectrum of US applications creating contrast-enhanced ultrasound (CEUS) imaging [11]. Microbubbles (MB) are widely used as the UCA for CEUS in the clinics [12]. They can be used for visualization of blood flow in vessels with a diameter below 200 µm [13]. In this case, the US images are generated because of the difference in acoustic impedance between the gas inside the MB and the blood [11, 13]. Furthermore, MB can be used for molecular imaging. For this purpose, their shell must be functionalized by incorporation of specific ligands or peptides, that enables binding of MB to, e.g. the intravascular targets [14].

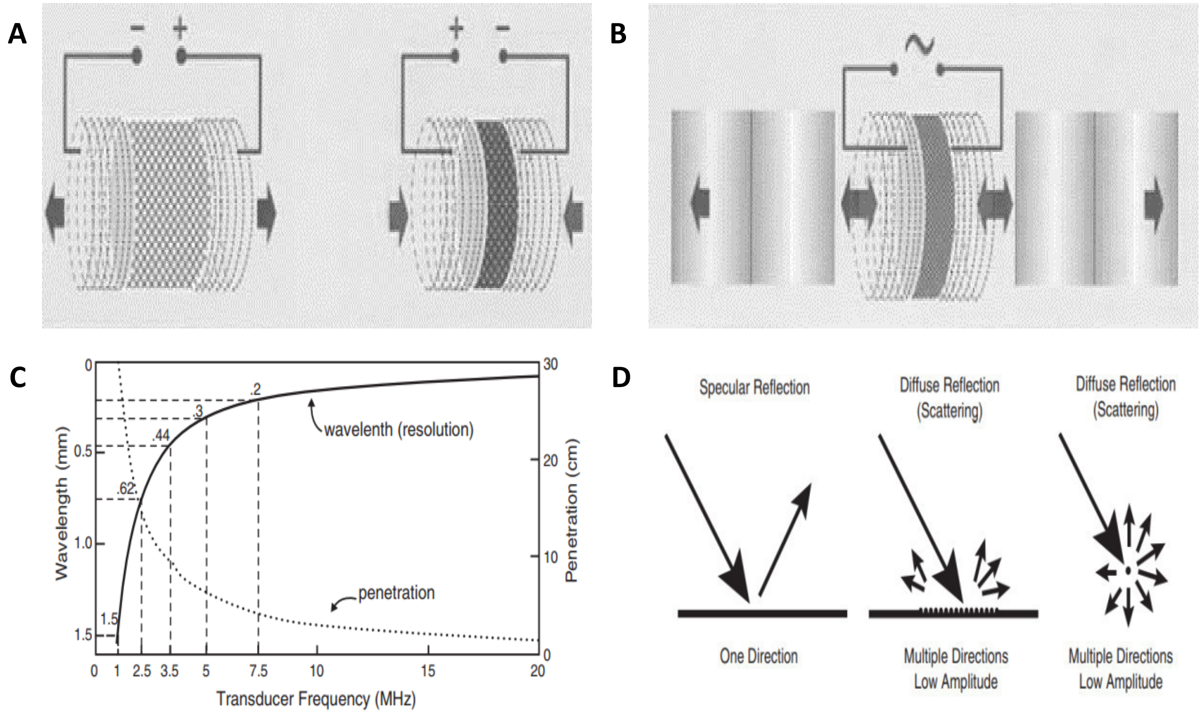


Figure 3: Principles of US imaging. (A) Principles of the piezoelectric effect. The element length changes, depending on the polarity of the applied current. (B) Subjecting the piezoelectric element to high-frequency alternating voltage results in US wave generation. (C) The relationship between resolution and penetration shows a strong frequency dependency [10]. (D) US wave and tissue interactions. Specular reflection is characteristic for interfaces of tissues with different acoustic impedances and is most important for the reconstruction of US images. Scattering occurs when US wave hits irregular tissue connections or objects with dimensions smaller than the wavelength [10].

Despite its broad applications, US imaging also has drawbacks. The first problem relates to the difference in sound propagation among tissues. Such disparity causes refraction artifacts [10]. The algorithm for image reconstruction, comprised in the US, has difficulties with overcoming this problem. Therefore, the lesion is falsely localized in the US image. As it was discussed before, the visualization of different tissues depends on the difference in their acoustic impedances. For instance, structures with very low (i.e., lungs) or very high (i.e., bones) acoustic impedance cannot be imaged with echography [9]. Furthermore, it is difficult to distinguish tissues with very similar or identical acoustic impedance. When the sound wave interacts with irregular surfaces or structures, which are smaller than the wavelength, scattering, may occur [10]. The scattered US pulse spreads as echoes of low intensities in multiple directions (Figure 3 D). Part of the reflected echoes returns to the transducer; thereby, the reconstructed images are rich in granular structures. These structures reflect the texture of imaged tissue. However, scattered reflection contributes as well to the formation of structural distortions in the US, which are referred as the speckle noise. It is a multiplicative type of noise [15]. In signal processing, speckle noise is described as a redundant random signal that is reconstructed together with relevant information. In US imaging, it is generated through the interference of backscattered wave-fronts, which are multiplied in the incident beam voxel [16]. Speckle noise carries corrupted and vital signals together. Therefore, processing it without losing necessary data is quite tricky. Moreover, it is characteristic for speckle noise to contain abrupt changes, which are the sudden changes of intensities in the noise pattern [17]. The goal of filtering an image corrupted with speckle noise is to remove the noise and preserve vital information. Usually, filtering methods treat images as uniform integrity. It can lead to excessive smoothing by using, e.g. median filter with the big kernel. Thereby, the noise is only being suppressed rather than being reduced. Speckle noise corrupts 2D signal at the pixel level. Hence, a filtering method that can reach each pixel separately is needed. There are many different approaches for speckle noise removal like adaptive filtering or the Fourier Transform (FT). Even though FT is a powerful tool for data analysis, it does not represent the abrupt changes efficiently [18]. Furthermore, data processed with FT are depicted as the sum of sinusoidal waves, which are not well localized in time or space. In the case of speckle noise removal, wavelet transform (WT) is the more suitable solution. In contrary to Fourier Transform, WT can detect abrupt changes in images corrupted by speckle noise [19]. The core of WT is a new class of functions called wavelets. They are referred as brief oscillations that can correlate with unknown signals of similar frequency [20]. Unlike sinusoidal waves, wavelets exist for a finite time. They come in a vast variety of sizes and shapes gathered in so-called wavelet families, and they divide into two main groups orthogonal and biorthogonal. Each wavelet is composed of scaling and wavelet function [20]. In the WT they are used as low-pass filter (LPF) and high-pass filter (HPF) respectively. The orthogonal wavelets comprise one scaling and one wavelet function. While, the biorthogonal wavelets are composed of two pairs of mentioned functions, one for signal analysis and the other for signal synthesis (Figure 4). There are two different kinds of wavelet analysis. The first one is continuous wavelet transform (CWT), and it is used for time-frequency analysis and filtering of time localized frequency components [20]. The second type of WT is discrete wavelet transform (DWT). DWT is used for compression and denoising of images and signals containing abrupt changes [20]. Therefore, it is more suitable in the case of speckle noise removal than CWT. Moreover, DWT is used for speckle noise removal from synthetic aperture radar (SAR) and optical coherent tomography (OCT) images [21-23]. In the view of described speckle noise characteristic and its removal methods, the DWT appears to be the most efficient and robust approach [15, 24].

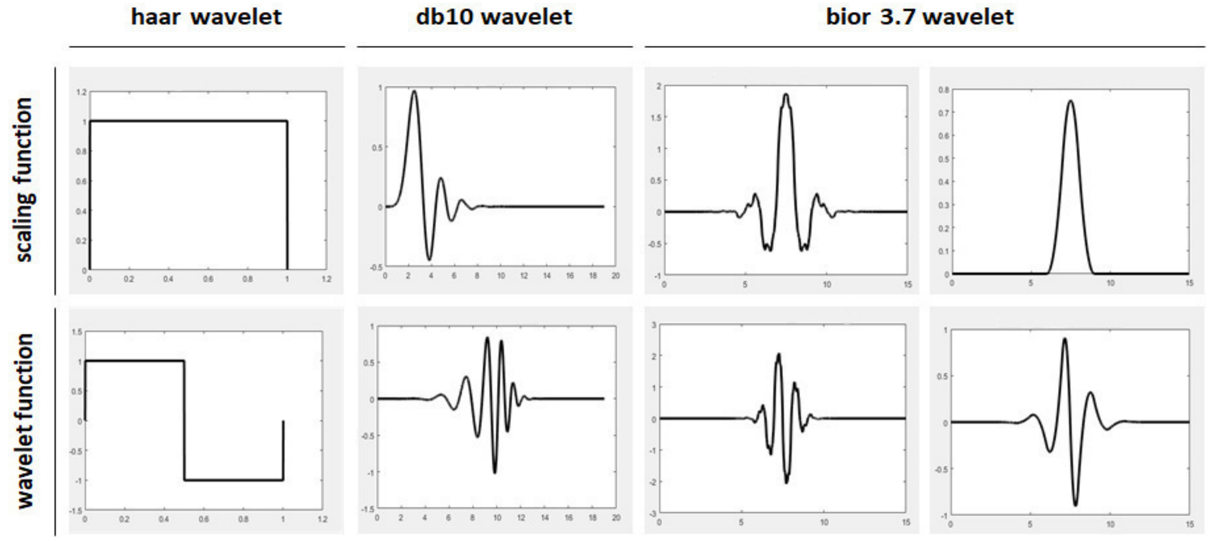


Figure 4: Wavelet types. Haar (db1) and db10 wavelet as an example of orthogonal wavelets. Both wavelets belong to Daubechies wavelet family [25]. Biorthogonal 3.7 wavelet as an example of biorthogonal wavelets. The scaling works as an LPF and wavelet function as an HPF in WT.

To perform discrete wavelet transform, the following operations must be done. Firstly, one must decide if the single or multilevel decomposition should be performed. The second step is choosing the wavelet. The goal of DWT performed with single or multilevel decomposition is to execute fine-scale analysis, which leads to the generation of four detail coefficient matrices. [20] They result from filtering rows and columns respectively in the noisy image, with HPF and LPF (Figure 5). The outcome of this transform is four matrices: diagonal detail coefficient matrix (HH), horizontal detail coefficient matrix (HL), vertical detail coefficient matrix (LH) and approximation coefficient matrix (LL). The last one approximates the original image. Another DWT, during the same fine-scale analysis, is performed always on the approximation images.

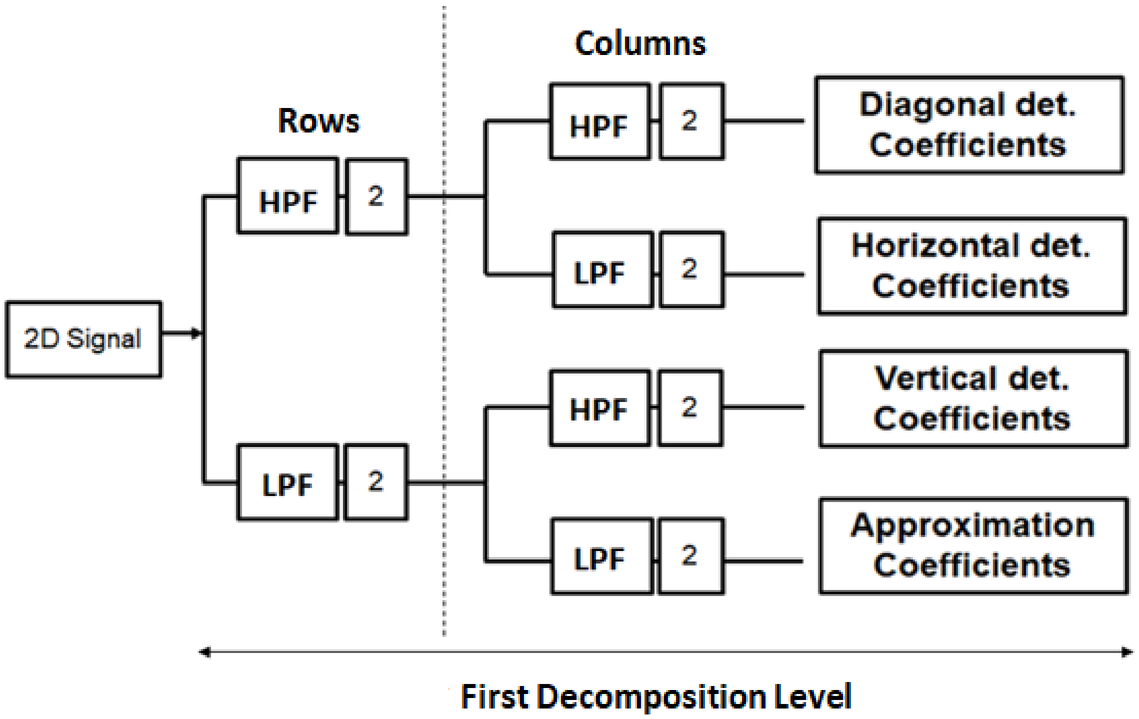


Figure 5: Schematic depiction of the discrete wavelet transform. The principle is explained based on a single level decomposition. At first rows of the input image (2D signal) are filtered with HPF and LPF. The values in the resulting matrices are then down sampled by the factor of 2. The second step is the filtering along the columns. Each column of the matrix, obtained during the first step, is filtered with HPF and LPF respectively. All resulting matrices are again downsampled. The first decomposition level results in 3 details matrices and one approximation matrix, which is the input for higher level DWT.

The next step of DWT is a selection of optimal thresholding technique for conducted transformation. The goal of this operation is to keep the significant image components (i.e., edges) and filter out the noise. There are two strategies for detail coefficients scaling known as soft and hard thresholding. Both thresholding techniques set the intensities values lower than threshold magnitude to zero [26]. However, soft thresholding additionally shrinks the remaining values of frequencies. From the noise removal point of view, soft thresholding is the most convenient strategy for detail subbands coefficients scaling [27]. In the case of DWT, choosing optimal threshold calculation is an essential aspect of noise removal. Some authors propose their computations [28, 29], however, majority use previously introduced solutions [30, 31]. One of the most frequently used threshold calculation techniques is Bayes Shrink (BS). It reduces effectively the noise with a Gaussian distribution. The magnitude of BS is calculated cumulatively basing on all the detail coefficients generated at each scale of DWT [32]. Thus, the value of the threshold is calculated once and applied to every subband.

Normal Shrink (NS) is the next popular threshold calculation technique used in DWT. NS threshold is obtained concerning every single detail coefficients matrix at each decomposition level (DL). Its calculation depends on the length of subband at specific DL [32]. Therefore, NS is a more flexible method, in comparison to BS. Deriving more adaptable strategy for calculation of the threshold magnitude is crucial in obtaining better noise removal results [24, 33]. The third and the last part of DWT is scaling of all detail coefficients with chosen thresholding technique. Afterwards, the signal is reconstructed with inverse DWT (IDWT). The denoised and recomposed image is then ready for further processing.

1.2. Tumor Detection and Segmentation

For the extraction of radiomics features the region of interest (ROI), i.e. tumor, needs to be segmented after the image acquisition. Recognizing and extracting proper lesions is an essential part of the diagnostic process. The manual lesion segmentation is the prevailing standard in current clinical practice. Nonetheless, it is a rather biased approach, since it depends strongly on the radiologists and their experience [34]. The growing interest in processing and analysis of medical images resulted in the creation of frameworks for segmentation of e.g., particular lesions or even whole organs [35, 36]. The desired ROI can be segmented with the use of standard methods like active contour (snake method), supportive vector machine or k-means clustering [37, 38]. However, in multiple studies, these methods are implemented to support non-automated or semi-automated segmentation, hence, the ROI segmentation is done by the user. [39, 40]. For the radiomic analysis purposes, the segmentations must be robust and should involve minimal user input [4, 34]. The automated segmentation (AS) can contribute consistency and promptness to the normal segmentation process [6]. In the process of current medical diagnostics, the manifold of medical images is generated. They can be used to build artificial intelligence (AI) based functions to generate frameworks for detection of tumors, classification of histopathological data and, furthermore, for the support of automated segmentation [1, 41, 42]. Machine Learning based methods like Cascade Classifiers (CCs) or Artificial Neural Networks (ANNs) are the tools that enable the creation of AI-based functions. For instance, interpretation and identification of lesions, in US images, is rather difficult and required experience [34]. Thereby, providing AI based functions, to already existing software could make the detection of lesions more feasible. Namely, the developed function would detect the ROI in provided images (Figure 6). Hence, the user interaction might have been excluded. Object detection and classification is the core of AI-based solutions. However, one must remember that every AI system has no previous knowledge of the subject, i.e. shape, look or size. Thus, providing big data set with various examples of the desired object is the important case.

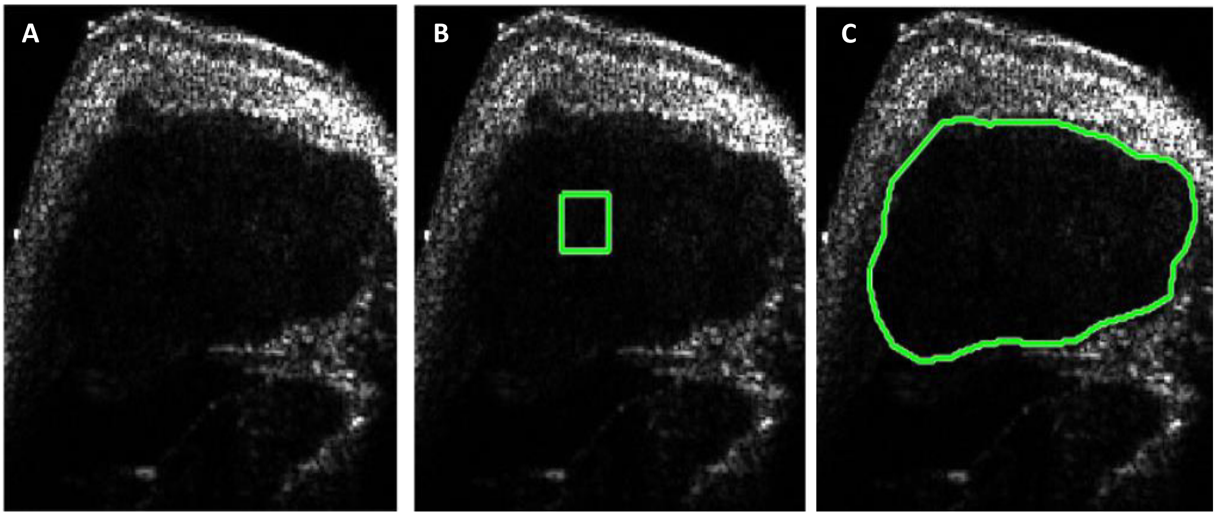


Figure 6: Automated Segmentation of the tumor. (A) Shows the original B-mode US image. The tumor detection is based on the AI-model is marked by the detection box (B). The function was prepared with the use of AI-based solutions. (C) The detection box is used as the starting point for the following tumor segmentation, performed with Active contour method. The final tumor segmentation is shown in (C).

In 2001 Viola et al. published a paper in which they presented new machine learning based framework for faces detection [43]. Even though their work is devoted to improvement in face detection field, the proposed method gives excellent results in objects detection as well. Authors of Viola-Jones rapid object detection framework are using the Haar-like features for description of image content [44]. They represent well the fine-scale characteristic of facial details. Local Binary Patterns (LBP) have similar output to Haar-like features [45]. However, training a classifier basing on LBP or Haar descriptors is rather computationally demanding. Histogram of oriented gradients (HOG) is another type of visual descriptors. It represents well the overall shape of ROIs. The study presented by Hbali et al. compare two CCs for face detection [46]. One of them was built with Haar-like features, and the second one with HOG features. The presented results showed that the CC developed with HOG features has better accuracy in face detection than the one trained with Haar-like features. Moreover, it performs the face detection three times faster than the CC developed with the use of Haar-like features. HOG features can also be used to develop AI based function for object detection [47]. The framework presented by Viola et al. makes a face and object detection way faster and more accurate than before, due to the introduction of several new contributions [43]. The first one is the integral image. Integral image is the new image representation that allows robust feature evaluation. The second one is the use of the AdaBoost approach for building classifiers. It enables the small number of critical attributes being selected in the image. Thereby, each created classifier represents newly selected feature (Figure 7).

Figure 7: Haar-like features selection in Viola-Jones framework. The figure presents the work of the AdaBoost framework [43]. The features are chosen based on the intensity of the selected region. The first selected features (row no. 2, second image) are differentiating eyes from the rest of the face. The second feature (row no. 2, third image) distinguishes between the eyes and the nose bridge.

The last proposed novelty is building cascades from successively created classifiers [43]. Thus, the processing speed is increased, and the detection function concentrates only on parts of the image, that is most likely to be recognized. CCs are well established and fast tool for face detection [48]. However, due to its broad use in this area, they are neglected in the cases of object detection. Many authors use ANNs for object detection and recognition [41]. Nevertheless, CCs are as well the robust and straightforward solution.

1.3. *Feature Extraction and Analysis*

Feature extraction and analysis are the last steps of a radiomic analysis (Figure 2). The mined features combine information for further classification of, e.g. a tumor. Gathered features can be either semantic or agnostic (Table 1) [4]. The first group refers to ones, which describe shape, location or spiculation of the lesion. The second comprises textural features that are computed with the use of first, second and higher order mathematical statistics. They based on intensities values of individual pixels or voxels.

Table 1: Summary of quantitative image features. The table contains types of features, exemplary traits for each group and methods for their extraction [4].

FEATURE TYPES		EXAMPLES	EXTRACTION METHODS
SEMANTIC	Shape-based features	Compactness, Surface area, Sphericity	Calculation of volume and surface area of tumor
AGNOSTIC	Intensity-based features	Range, Kurtosis, Skewness	Histogram Computation
	Textural features	Contrast, Correlation, Homogeneity	GLCM/GLRLM calculation or wavelet filtering

Computation of a gray-level co-occurrence matrix (GLCM) and a gray-level run length matrix (GLRLM) is standard practice for the extraction of textural features [49, 50]. These methods are enabling extraction of microtextural traits [51]. Moreover, performing undecimated wavelet filtering on the extracted segment can provide additional information about a multiresolution tumor pattern [2, 6]. In the case of WT, frequencies comprised in generated detail and approximation matrices (Figure 5) are not down-sampled. GLCM and GLRLM are calculated correspondingly for each resultant subband, yielding in a new group of traits. Currently, statistical analysis is used for the qualification of extracted features. They can be ranked according to their stability or importance, and each of them must be: highly reproducible, informative and non-redundant [34]. The precision of the created analytical model strongly depends on the number of mined features and the variety of data sets, from which these traits were extracted. Obtaining large quantities of attributes is not enough for creating the reliable differential model [52]. Furthermore, researchers must possess experience in the field of designing and conducting studies supported by the statistical analysis. Radiomic analysis is a promising field of the medical related studies. For the good of patients and future research, it should be investigated with high accuracy and careful attention [52].

1.4. Aims of Presented Work

Radiomics is a newly developing field, aiming to improve the diagnostic potential of medical imaging. The combination of new and already established biomarkers might support physicians in tumor classification or therapy evaluation in the future [4]. Currently, the radiomic analysis is mostly done for CT and MRI, due to their high reproducibility and abundant use in clinical practice. Despite the multiple diagnostic applications and manifold of images generated with US systems, only a few radiomics framework has been prepared for this type of data [1, 3, 7]. Those studies focused mainly on analyzing all the different aspects of textural features, and not the combination of the different imaging biomarkers classes. Therefore, presented work aimed to:

1. Perform denoising of US images with the use of DWT
2. Create algorithm for automated tumor detection and segmentation
3. Extract imaging biomarkers
4. Execute automated tumor model differentiation

2. *Materials and Methods*

2.1. *Hardware and Software*

The whole project presented in this master thesis was written in MATLAB environment, with The MathWorks Inc. R2017a release, on Dell OptiPlex 760 with Intel® Core™2 Duo CPU and 64-bit Windows 7 Professional.

2.2. *Tumors Images Data Set*

The provided images data set was taken from a study presented by Opacic et al. [53]. Images were collected from 9 different lesions from 3 different tumor models: MLS, A431, and A549 (Supplementary Figure 1). Each tumor was measured at two different positions. Thus, the provided image data set included, in total, 18 images. The images were collected with the use of the Vevo2100 imaging system (FujiFilm VisualSonics Inc., Toronto, Canada) and the MS550 transducer, operating at 40 MHz. The reconstructed 16-bit US images had a pixel size of 55 µm x 22 µm.

2.3. *Speckle Noise Removal*

The multiplicative nature of speckle noise in US images requires a sophisticated noise removal algorithm. Currently, DWT seems to be the best strategy for speckle noise removal [54]. The sequence of functions used to develop denoising algorithm is presented in Table 2.

Table 2: Algorithm for speckle noise removal in US images.

Speckle Noise Removal

1. **Log transform - conversion of multiplicative noise to additive noise**
2. **DWT - image decompositions**
3. **Thresholding**
4. **IDWT - image reconstruction**
5. **Inverse log transform**
6. **MSE and PSNR calculation – implemented denoising evaluation**

Firstly, the input image was subject to the logarithmic transform calculated with the following equations:

$$a = e/\log\big(1 + abs(R)\big) \tag{1}$$

where ***e*** is the number of gray-level intensities in the image and ***R*** is the maximum intensity value in the image,

$$TI = a \cdot log(1 + I) \tag{2}$$

where ***a*** is the scaling constant, ***I*** is the input image and ***TI*** the transformed image**.** Secondly, DWT was applied to the image and conducted 3 times (Figure 8). This procedure was repeated 5 times and each the time image was decomposed with different wavelet: bior3.7, db10, haar, rbio1.5, sym5.

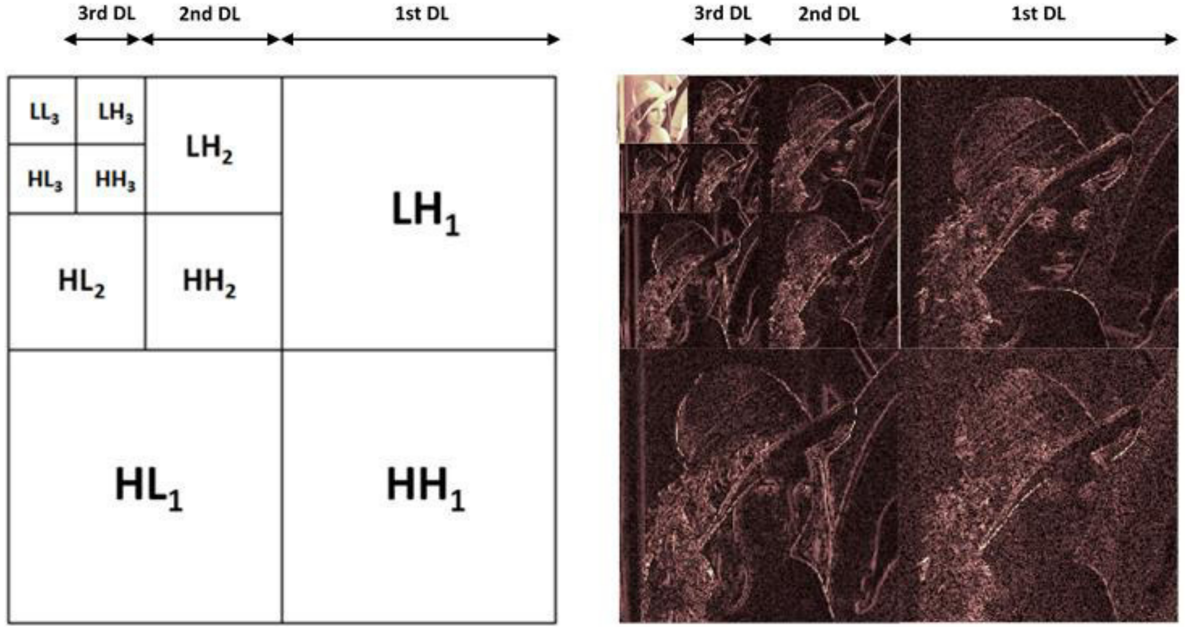


Figure 8: DWT: Triple decomposition of Lena image. Every DL comprises HH, HL and LH [51]. The following DWT are performed LL.

Thresholding of detail coefficients matrices was the next processing step. Frequencies were scaled with soft thresholding, and the threshold magnitude was calculated with two different techniques, namely, BS and NS. The following equations presents the threshold calculations, based on the method proposed by Elyasi et al. [32].

$$Y = X + V \tag{3}$$

where $\boldsymbol{Y}$ is the WT of a noisy image, $\boldsymbol{X}$ is the WT of original image and $\boldsymbol{V}$ is the WT of the degraded components, i.e. noise. It is assumed that speckle noise has a normal distribution. Thus the variances of each component from (3) are as follows:

$$\sigma_y^2 = \sigma_x^2 + \sigma_v^2 \tag{4}$$

$$\sigma_v^2 = \frac{median(HH_1)}{0.6745} \tag{5}$$

$$\sigma_y^2 = \frac{1}{M}\sum_{m=1}^{M} A_m^2 \tag{6}$$

$$\sigma_x^2 = \max(\sigma_y^2 - \sigma_v^2) \tag{7}$$

where, $\boldsymbol{\sigma_v^2}$ is the noise variance, $\boldsymbol{HH_1}$ is the diagonal detail coefficient matrix at the 1st DL, $\boldsymbol{\sigma_y^2}$ is the variance of a noisy image, $\boldsymbol{M}$ is the total number of coefficients at all DLs, $\boldsymbol{A_m}$ are the coefficients in each subband and $\boldsymbol{\sigma_x^2}$ is the variance of the original image. In this framework, BS was calculated in two different ways. The first one is referred as global BS (gBS) was computed according to:

$$gBS = \frac{\sigma_v^2}{\sigma_x} \tag{8}$$

$$\sigma_x = \sqrt{max\left(\sigma_y^2 - \sigma_v^2\right)} \tag{9}$$

where, $\boldsymbol{\sigma_x}$ is the standard deviation of the WT performed on the original image. The second threshold, derived from BS, was called individual BS (iBS). It was calculated with a similar equation to (5). However, this time for each detail coefficient matrix individually:

$$iBS_{HHn} = \frac{\sigma_v^2}{\sigma_{x_{HHn}}}; \quad iBS_{HLn} = \frac{\sigma_v^2}{\sigma_{x_{HLn}}}; \quad iBS_{LHn} = \frac{\sigma_v^2}{\sigma_{x_{LHn}}}; \tag{10}$$

$\boldsymbol{\sigma_{x_{HHn}}}$is the standard deviation of HH at the nth DL, $\boldsymbol{\sigma_{x_{HLn}}}$is the standard deviation of the HL at the nth DL, $\boldsymbol{\sigma_{x_{lHn}}}$is the standard deviation of the LH at the nth DL. The threshold value of iBS was calculated separately for each DL. NS magnitude was computed with the following formula:

$$\alpha = \sqrt{log\left(\frac{L_k}{J}\right)} \tag{11}$$

$$NS = \alpha \frac{\sigma_v^2}{\sigma_y} \tag{12}$$

where $\boldsymbol{L_k}$ is the length of the subband at $\boldsymbol{k^{th}}$ DL, J is the total number of decompositions and $\boldsymbol{\sigma_y}$ is the standard deviation of every detail decomposition matrix at each DL. Calculated gBS, iBS and NS were applied to either all DLs or just the 3rd DL. The next steps of the denoising procedure were image reconstruction with the use of IDWT and inverse log transform respectively. In the end, the mean squared error (MSE) and peak signal to noise ratio (PSNR) were calculated to evaluate the performance of all variations of noise removal algorithm. MSE and PSNR were computed with respect to selected tumor region.

$$MSE = \frac{1}{ROI}\sum_{x,y}\big(I(x,y) - RI(x,y)\big)^2 \tag{13}$$

where ***ROI*** is a selected tumor region, $\boldsymbol{RI}$ is image reconstructed with IDWT, ***x*** is all the columns and ***y*** is all the rows in both $\boldsymbol{I}$ and $\boldsymbol{RI}$**.**

$$PSNR = 10\log(e^2/MSE) \tag{14}$$

where ***e*** is the number of gray-levels in the image. For following automated tumor detection and segmentation, denoised images contrast was adjusted.

2.4. Automated Tumor Detection

As to perform the automated segmentation of given tumors, the object detection function had to be developed. Thus, the following steps had to be performed.

Table 3: Algorithm for automated tumor detection.

Automated Tumor Detection

1. **Preparation of images data sets - positive and negative samples**
2. **Computation of tumor detection functions - cross-out validation technique**
3. **Tumor detection**

One must remember that the artificial intelligence based functions have no previous knowledge about the appearance and form of the detectable object. Hence, the samples containing different positions and sizes of i.e. tumors are desirable. In this study, tumor images previously subjected to the denoising algorithm were doubled and quadrupled (Figure 9). In the end, the data set composed of 4 different groups of images was obtained (see in Table 4). The next step was to extract positive and negative samples from prepared images. With the use of the MATLAB built-in application (Training Image Labeler App), positive and negative ROIs were selected for each image and stored in a table file.

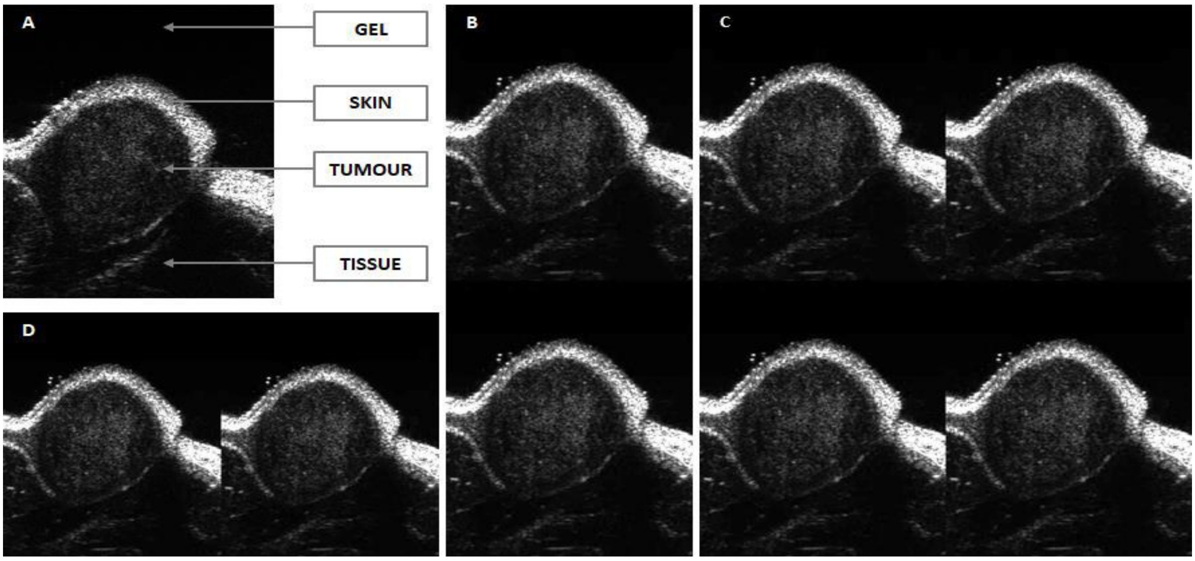


Figure 9: Training data set for automated tumor detection. (A) Each image was divided into 4 regions: GEL, SKIN, TUMOR, and TISSUE. Data was composed of: (A) single, multiplied and concatenated images (B) vertically, (D) horizontally and (C) quadrupled images.

Table 4: Image groups used in building image data set for building cascade classifiers.

IMAGE GROUP	NO. OF IMAGES
Original Tumor Images	18
Duplicated and Horizontally Concatenated Images	18
Duplicated and Vertically Concatenated Images	18
Quadrupled Images	18

Positive ROIs were marked inside the tumor. In total, the set of positive images combined 900 images. The negative ROIs were extracted from 3 different areas: Gel, Skin, and Tissue (Figure 9 A). For each area, 5 negative ROIs were selected per tumor. The negative set of images contained 1890 images (Table 6). Finally, the prepared data set was used for developing the cascade classifiers for tumors detection. Each function was trained with the use of HOG features and contained 15 classification stages. Training was done using leave-one-out cross-validation technique. Thus, the whole procedure resulted in building 18 cascade classifiers, one for every tumor.

Table 5: Positive ROIs marked in the image groups. The number of images per group are indicated in brackets.

IMAGE GROUP	ROIs per TUMOR	ROIs per IMAGE	TOTAL
Single (18)	10	10	180
Doubled (36)	5	10	360
Quadrupled (18)	5	20	360
TOTAL	-	-	**900**

Table 6: Negative ROIs marked in image groups. The number of images per group are indicated in brackets.

IMAGE GROUP	ROIs per AREA	ROIs per IMAGE	TOTAL
Single (18)	5	15	270
Doubled (36)	5	30	540
Quadrupled (18)	5	60	1080
TOTAL	-	-	**1890**

2.5. *Automated Tumor Segmentation*

The developed segmentation algorithm is composed of five following parts. The stream of tumor segmentation algorithm is depicted in Table 7.

Table 7: Algorithm for automatic tumor segmentation.

Automated Tumor Segmentation

1. **Tumor detection and masks preparation**
2. **Pre – processing of the input image**
3. **Edge map calculation**
4. **1st Details removal**
5. **Edge repair**
6. **2nd Details removal**
7. **Morphological operations**
8. **Active Contour segmentation and user refinements**

2.5.1. *Tumor Detection and Masks Preparation*

At first, the tumor was detected in the image with the use of the trained CCs. The detected object was marked with a square box within approximate size of 20x20 pixels (Figure 10 A). If several objects were detected, only the most stable one was considered for the following steps of the algorithm. The detection box was used as a seeding mask for the next step of algorithm, which was the active contour segmentation. Furthermore, a second mask was generated, with the use of image region properties function. A bounding box (BB) was drawn around the region with intensity higher than 2500, which in this case indicates the tumor (Figure 10 B). The second mask was prepared for the following extraction of the tumor region.

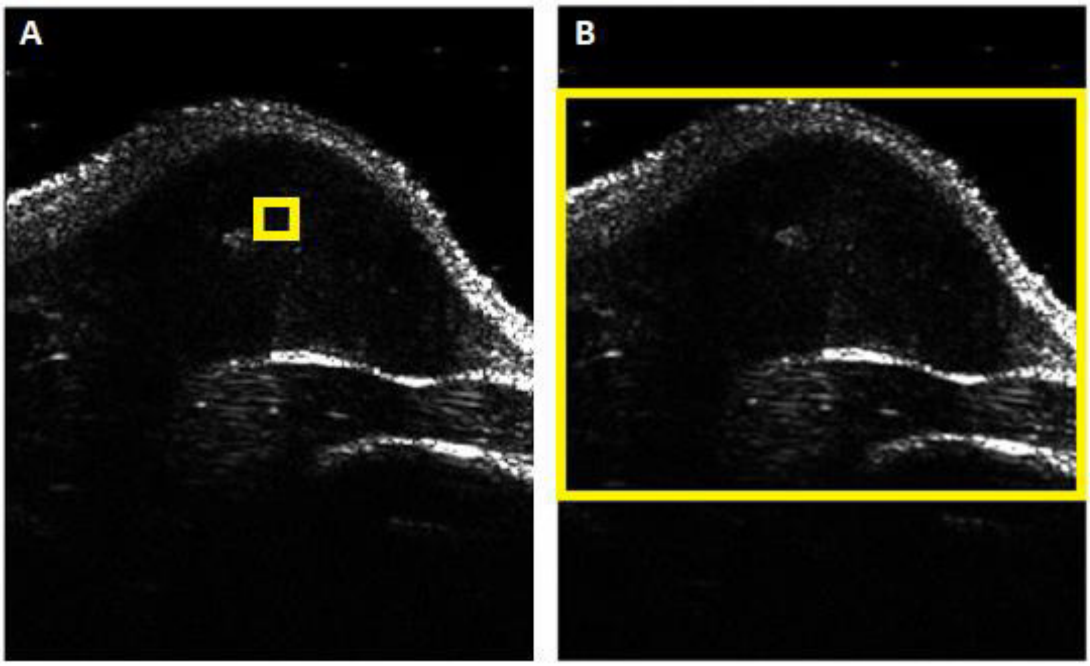


Figure 10: Generated Masks. Mask computed from (A) tumor detection marked by BB with an approximate size of 20x20 pixels and (B) thresholding of the image region enclosing pixel intensities higher than 2500. The extracted region was marked with BB.

2.5.2. Pre-processing of Input Image

In the image processing part of the segmentation algorithm, each image was at first smoothed with a median filter with a kernel size of 10x10 (Figure 11). Secondly, the horizontal gradient (HG) and vertical gradient (VG) of the smoothed image were obtained. The last part of image processing was the computation of a horizontal edge map (HEM) and vertical edge map (VEM) of the tumor. At first, the initial HEM and VEM were generated from HG and VG images using Canny Edge detection. The predefined sensitivity threshold (ST) influence the number of detected edges in both initial edge maps. To calculate the final HEM, the adjustment method was established (Figure 11 Adjustment of HEM). Firstly, the content of positive pixels (PPs) in initial HEM was calculated. If it was lower than 5%, the predefined ST was decreased, and another Canny Edge Detection was performed. This procedure was executed until the content of PPs in calculated HEM was equal or higher than 5%. To compute the final VEM, the ratio of initial VEM and final HEM was calculated according to (15). Afterwards, the final VEM was calculated by using equation presented in (16).

$$ratio = \frac{intitial\ VEM}{final\ HEM} \tag{15}$$

$$final\ VEM = ratio \cdot initial\ VEM \tag{16}$$

The generated horizontal and vertical edge maps were then passed to further steps of the segmentation algorithm.

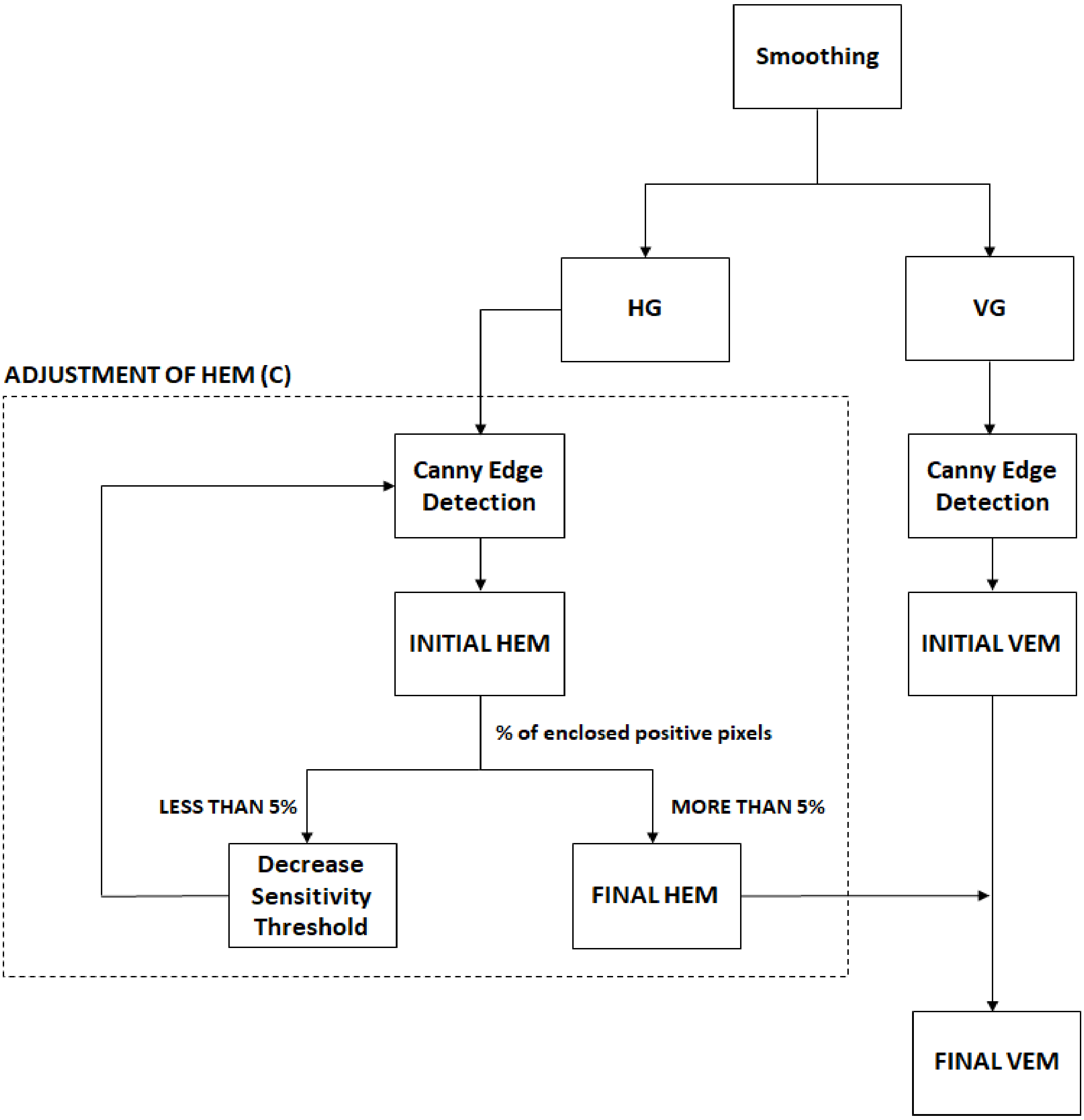


Figure 11: Stream of image processing in automated tumor segmentation algorithm depicted in a block diagram.

2.5.3. Details removal

Some of the prepared HEMs contained too many details inside the tumor region. This may have resulted in slower and less accurate segmentation of the tumor. Thus, a details removal step was introduced. The tumor detection BB was the starting point of the whole procedure. Firstly, the content of positive pixels enclosed in the detection BB was analyzed (Figure 12 A). It had to be smaller than 10% for the details removal to begin. When that condition was fulfilled, the positive pixels inside the detection BB were erased, and 16 small bounding boxes (sBB) were built on its corners and edges (Figure 12 B). Their width and height were equal to 1/3 of the width and height of the tumor detection BB. The next step was to analyze the content of PPs in the 16 sBB. If any of the small boxes enclosed PPs, then the enclosed white pixels were erased, and the position of this particular sBB was saved (Figure 12 C). Flowingly, four sBB were drawn around saved sBB (Figure 12 D). The details removal was repeated if any of them enclosed positive pixels. The function was executed until all the redundant details were removed from input the HEM. The details removal function was used twice in the developed algorithm for automated tumor segmentation (Table 7), before and after the edge repair step.

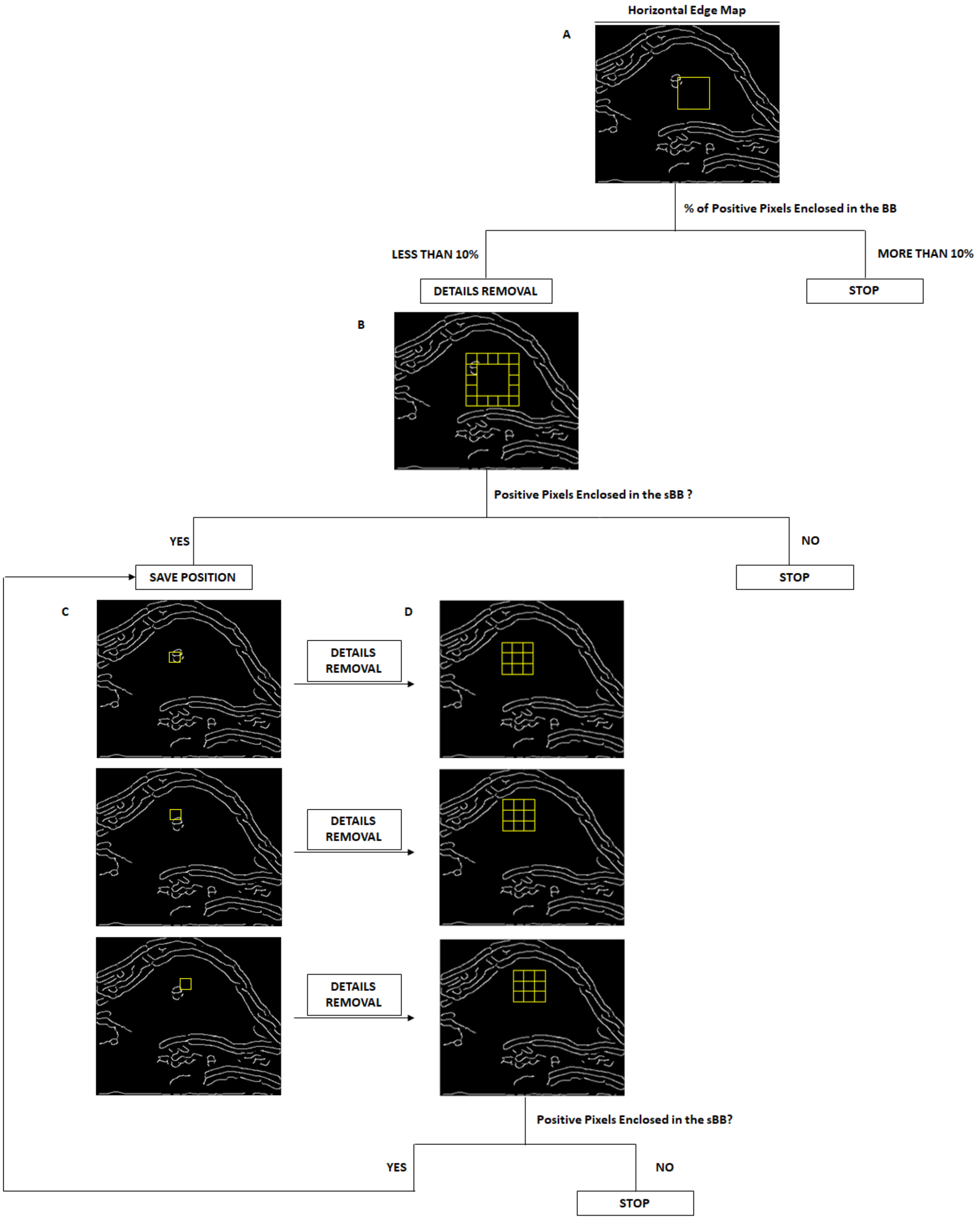


Figure 12: Principal of details removal algorithm. The calculated HEM contains redundant details in the middle of tumor region. Thus, number of PPs enclosed in tumor detection box is analysed. If it is lower than 10%, the details removal begins. The function is executed until all the redundant details are removed.

2.5.4. Edge Repair

The skin above the tumor is the only well-emphasized region in the images (Supplementary Figure 1). Therefore, the edge of the skin is the most uniform part in the calculated HEMs of the tumors. To refine every calculated HEM, the repair function was developed concerning the method presented by Micong et al. [55]. It relays on both the vertical and horizontal edge map. During the image pre-processing flow, the adjusted value of vertical ST was set to be lower than the horizontal one; hence, the VEM contains more details than HEM (Figure 13 A). In the edge repair procedure, HEM was the input image, and the VEM was used to repair it (Figure 13 A, B). Firstly, the algorithm was looking for the edge endpoints in the calculated HEM. When it was found, the number of positive pixels in its neighborhood was analyzed. If it was less than 3, the edge was repaired in such region (Figure 13 C). Afterwards, the next positive point was found. However, if the distance between the new and the previously detected points was less than 10, the edge repair algorithm was not performed. This circumvented too many repairs in the same region. To remove details which were added to the tumor center, the details removal algorithm was performed a second time.

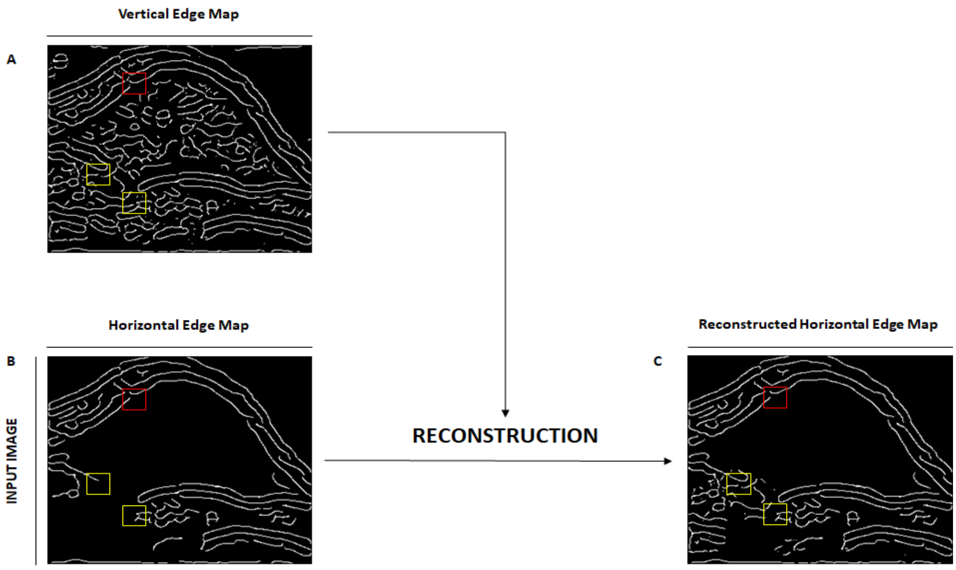


Figure 13: Principle of edge repair. The vertical and horizontal edge maps were used as input for the edge repair. The gaps in HEM were detected (A; yellow squares) and repaired by cropping the respective region from VEM (B) and pasting it into the HEM(C). The red boxes indicate gaps which are also detected by the algorithm, but do not fulfill the conditions for an edge repair.

2.5.5. Morphological Operations

The final edge map was modified with the use of morphological operations. The point of this operation was to decrease the number of objects in the image. The details of the final edge map were thickened and then connected diagonally and horizontally. The area of connected contour was then decreased with thinning operation. The final mask was overlaid on the tumor in the original tumor image.

2.5.6. Tumor Segmentation and User Refinements

The active contour segmentation algorithm was applied to the image prepared in the previous steps of this algorithm. Tumor ROIs were firstly segmented and then passed to the user refinements section. Lastly, the user could refine a created segment by changing shape and size of its contour. Generated segments were refined by 3 different users (User 1, User 2 and User 3). Each group of extracted ROIs constitutes a separate data set. In the end, refined segments were saved and preceded to the next parts of presented radiomic analysis to test the influence of multiple tumor delineations.

2.6. Feature Extraction

The proposed feature extraction follows the framework presented by Aerts et al. [2]. However, in this master thesis only, agnostic features were mined (Table 1). The stream of the algorithm presented in Table 8 was used to extract features from each dataset comprising delineated tumor segments by User 1, User 2 and User 3.

Table 8: Algorithm for tumor features extraction.

Feature Extraction

1. **Histogram computation**
2. **Generation of GLCM and GLRLM**
3. **Undecimated wavelet filtering**
    3.1. Generation of GLCM and GLRLM for each resultant sub - band

To perform tumor feature extraction, previously prepared segments were extracted from the denoised raw images, without contrast enhancement, and overlaid on the matrix of zeros with the same size as the original image. Firstly, 14 intensity-based features were calculated from the computed intensity histogram. Then GLCM and GLRLM matrices were calculated for every input segment, to extract textural features. Both matrices were generated for adjacent intensities in four directions: horizontal (0°), vertical (90°) and along the image diagonal under 45° and 135°. All results, in both GLCM and GLRLM, calculated concerning 0 intensity, were excluded for the following feature extraction. The computation of GLCM and GLRLM resulted in the extraction of 21 and 11 textural features respectively. The last part of the feature extraction algorithm was filtering the tumor segments with undecimated wavelet transform. The image was decomposed one time with the use of coif1 wavelet. In the end, image histogram, GLCM, and GLRLM matrices were computed for each resultant detail subband, what yielded in the extraction of 146 new features. The extracted features were assigned to three groups: first order statistic (FOS), second order statistics (SOS) and wavelet transform (WT). The list of all 230 extracted features is enclosed in Supplementary Figure 2-4

2.7. Automated Tumor Differentiation

Table 9: Algorithm for automated tumor differentiation

Automated tumor differentiation

1. **Feature selection and dimensions reduction**
    - **1.1 CCC calculation (test – retest)**
    - **1.2 P – value computation (Friedman test)**
    - **1.3 ANOVA**
    - **1.4 Radiomic signatures generation**
        - **1.4.1 RS1: 3 highest ranked features**
        - **1.4.2 RS2: 3 highest ranked not correlating features**
2. **Supervised and unsupervised tumor differentiation**

Developed algorithm for the automated tumor differentiation was prepared concerning the method presented by Aerts et al. [2]. Consecutive steps listed in the figure above led to generation of two radiomic signatures (RS), which determined classification models for differentiation of three tumor types comprised in this study. To assess the accuracy of both, supervised and unsupervised learning algorithms were used.

2.7.1. Feature Selection

The first part of automated tumor differentiation was devoted to the creation of the RS. It is a sequence of 3 the most discriminative features that were defined by carrying out multiple statistical tests on previously extracted traits (Supplementary Figure 2). Each of conducted calculations had to rank 230 features according to their stability or discriminative power. Firstly, the test-retest was performed. It relies on the calculation of concordance correlation coefficient (CCC), which indicates variation between measurements. The value of CCC is in the numerical range from 0 to 1, where 1 refers to the most and 0 to the least stable feature. This test investigated the stability of features between pairs of images. The Friedman test was the second conducted traits stability ranking. This test explored if the stability of the particular feature is user-independent in between multiple tumor delineations. In this case, p-values for individual features were calculated. Interpretation of results is the same as in the case of test-retest. One-way ANOVA analysis was the last conducted test. With this analysis, the mean of individual features was studied based on their conformity. To perform ANOVA, every image together with a value of a particular feature was assigned to one of three tumor models (MLS, A431, A549). The more similar the mean of examined features, in between different tumor models, the less discriminative the feature was. P-values calculated from ANOVA were comprised in the numerical range from 0 to 1, where 0 denotes the most and 1 the lest discriminative feature. In the end, ranks obtained from every test were averaged to create a final ranking of an individual attribute. The highest rank feature was assigned with the lowest number. The last and the most essential part of feature selection was selecting traits for RS. Components of RS were chosen from every group of features, namely FOS, SOS, and WT. Traits with their final ranks were assigned to their native group. Afterwards, they were arranged according to their ranks, from the highest to the lowest. In the end, two RS were prepared. The first one (RS no.1) was composed of traits from each group, which yielded the highest rank. Furthermore, the correlation between features in RS no.1 was checked. To generate the second RS (RS no.2), the lack of linear relationship between three chosen features was investigated. If any two features were similar to each other, which meant that their correlation coefficient was close to 1, then the trait with a lower rank in this pair was exchanged. This procedure has been executed until features with low or without linear dependency were found. In the edn RS no.2 was composed of the first highest ranked features, from each feature group, with low or no correlation.

2.7.2. Supervised and Unsupervised Tumor Differentiation

Both RS were validated with the use of supervised and unsupervised classification algorithms. Once again, every image together with its RS features was assigned to one of three tumor models. The values of features were standardized with z-score before further analysis. Afterwards, normalized values were clustered with supervised and unsupervised learning algorithms. In the case of unsupervised learning, the features were grouped according to the distance between them calculated with Manhattan method. To present the results of features clustering, the heat map with dendrogram was computed. The presentation of final results was generated with the use of open source CIMminer software. For supervised training of prepared tumor differentiation model, Classification Lerner App, MATLAB built-in application was used. It provides a vast variety of classification algorithms that helped in training and validation of both RS. In this study, we compared the accuracy of three following training algorithms: Simple Tree, Linear Support Vector Machine (SVM) and fine Nearest Neighbors (k-NN). Each classification model was trained with the leaving-one-out cross-validation technique.

2.8. *Overall Image Processing and Analysis Algorithm*

Table 10: Final algorithm of radiomic analysis. Summary of all the steps performed to conduct feature extraction and pattern recognition of US images.

1. **Speckle Noise Removal**
    1.1. **Log transform - conversion of multiplicative noise to additive noise**
    1.2. **DWT – image decompositions**
    1.3. **Thresholding**
    1.4. **Inverse DWT – image reconstruction**
    1.5. **Inverse log transform**
    1.6. **MSE and PSNR calculation – denoising evaluation**
2. **Automated Tumor Detection**
    2.1. **Preparation of images data sets – positive and negative**
    2.2. **Computation of tumor detection functions – the cross-out validation technique**
    2.3. **Tumors detection**
3. **Automated Tumor Segmentation**
    3.1. **Tumor detection and masks preparation**
    3.2. **Pre-processing of input image**
    3.3. **Edge map execution**
    3.4. **1st Details removal**
    3.5. **Edge repair**
    3.6. **2nd Details removal**
    3.7. **Morphological operations**
    3.8. **Active Contour Segmentation**
    3.9. **User Refinements**
4. **Feature extraction**
    4.1. **Histogram computation**
    4.2. **Generation of GLCM and GLRLM**
    4.3. **Undecimated wavelet filtering**
        4.3.1. **Generation of GLCM and GLRLM for each resultant subband**
5. **Automated tumor differentiation**
    5.1. **Feature selection and dimensions reduction**
        5.1.1. **Test-retest**
        5.1.2. **Friedman test**
        5.1.3. **ANOVA**
    5.2. **Radiomic signatures generation**
        5.2.1. **RS1: 3 highest ranked features**
        5.2.2. **RS2: 3 highest ranked not correlating features**
    5.3. **Supervised and unsupervised tumor differentiation**

3. Results and Discussion

3.1. CEUS Images Denoising with the Use of DWT

The results presented in the following section conclude which form of denoising was the most optimal solution for the processing of all images, from the given data set (Supplementary Figure 1). Each of 18 images, was filtered with 30 speckle removal variants. Each of the developed despeckling variants comprised different: wavelet, threshold magnitude calculation and threshold application (Figure 14 A, B, C).

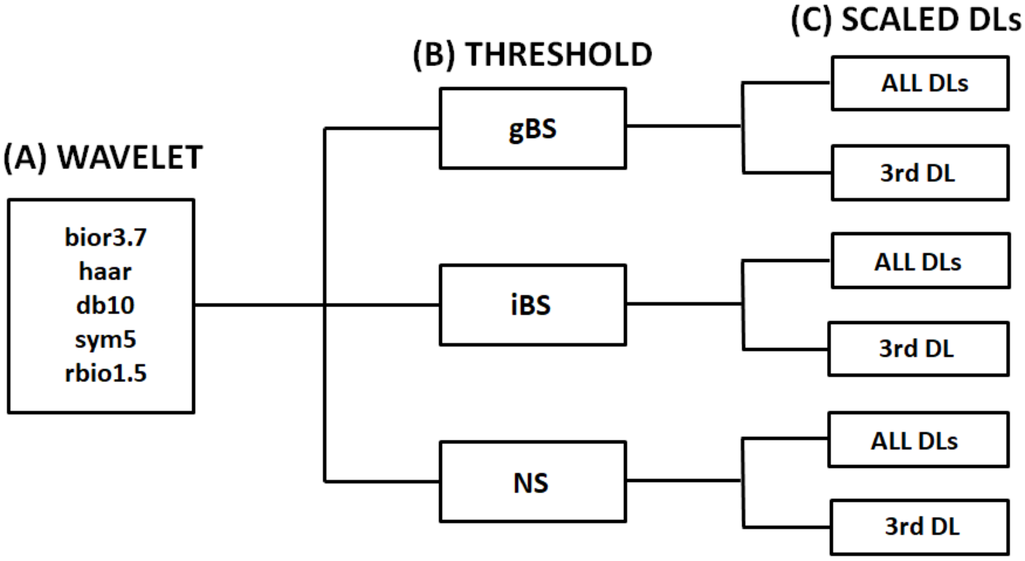


Figure 14: Framework of speckle noise removal. This figure presents the different variants which were used for speckle noise removal. Five different wavelet functions were used for the wavelet decomposition (A), then three different types of threshold magnitude calculation were applied (B). The threshold was applied either to all DLs or only to subband coefficients at the 3rd DL (C). In total, 30 different speckle noise removal algorithms were evaluated.

MSE and PSNR were used to evaluate the performance of developed speckle noise removal variants. These quality criteria were computed with respect to the tumor ROI previously marked with a polygon (Figure 15 B). The quality measurement of the executed despeckling variants was conducted in a described manner, since the region of the lesion is part of the greatest interest. The interpretation of quality measurement is such that the lowest MSE and highest PSNR magnitude characterize the best performance of noise removal algorithm.

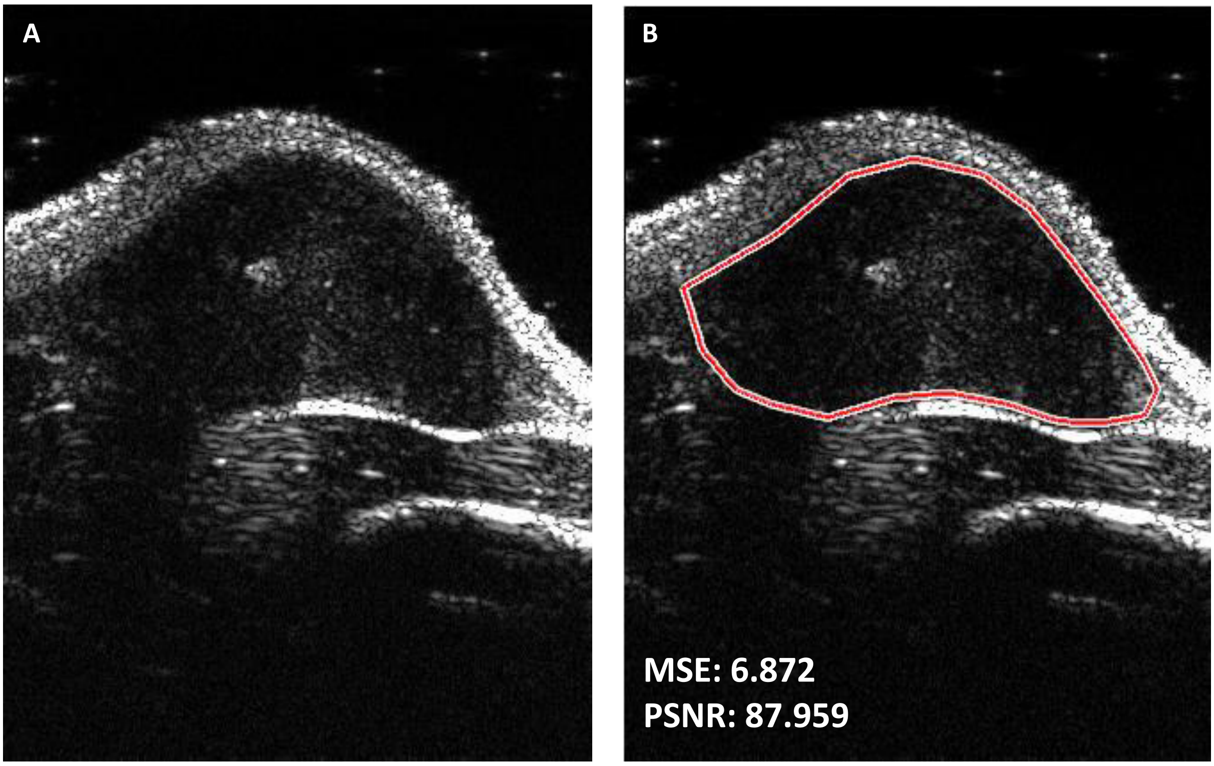


Figure 15: The principle of evaluating the speckle noise removal algorithms. Original US image (A) and image filtered with bior3.7 wavelet with iBS threshold applied only to 3rd DL (B) are exemplarily depicted. The values of MSE and PSNR were calculated with reference to the tumor ROI, marked by the red polygon (B).

Choosing the most suitable wavelet function was the first step of speckle noise removal framework. Selected wavelets are one of the most commonly used functions in studies devoted to speckle noise removal with the use of DWT [26, 56, 57]. The least satisfactory results (MSE = 3479.496 and PSNR = 60.914) were obtained for rbio1.5 wavelet. MSE and PSNR values, computed for images decomposed with haar, db10 and sym5 wavelets, are quite similar (Figure 16 C, D, E). Analyzing the results obtained for every filtered image, one can clearly state, that image decomposed with bior3.7 wavelet (Figure 16 A) yields the best quality (MSE = 6.872 and PSNR = 87.959). The value of MSE obtained in our study for majority of wavelets are higher in comparison to other studies. The initial image data set was composed of 16-bit images. Hence, it may have resulted in obtaining the higher values of MSE. The authors of other studies devoted to the image filtering with DWT did not enclose the information about the color depth of the used images [26, 32, 33, 56]. The values of PSNR values, among filtered images (Figure 16), is not considerably different. However, the best result is still obtained for bior3.7 wavelet. The better performance of bior3.7 in comparison to other wavelets was also presented study by Mustafa et al. [56]. It is possible that having separate sets of functions for signal decomposition (Figure 4) might have resulted in better localization of abrupt changes. However, in the presented study the quality of image denoising with the rbio1.5 wavelet, which belongs to the same wavelet family as bior3.7, was tested and it expressed the least satisfactory results. Therefore, the influence of wavelet composition on quality of noise removal should be further investigated.

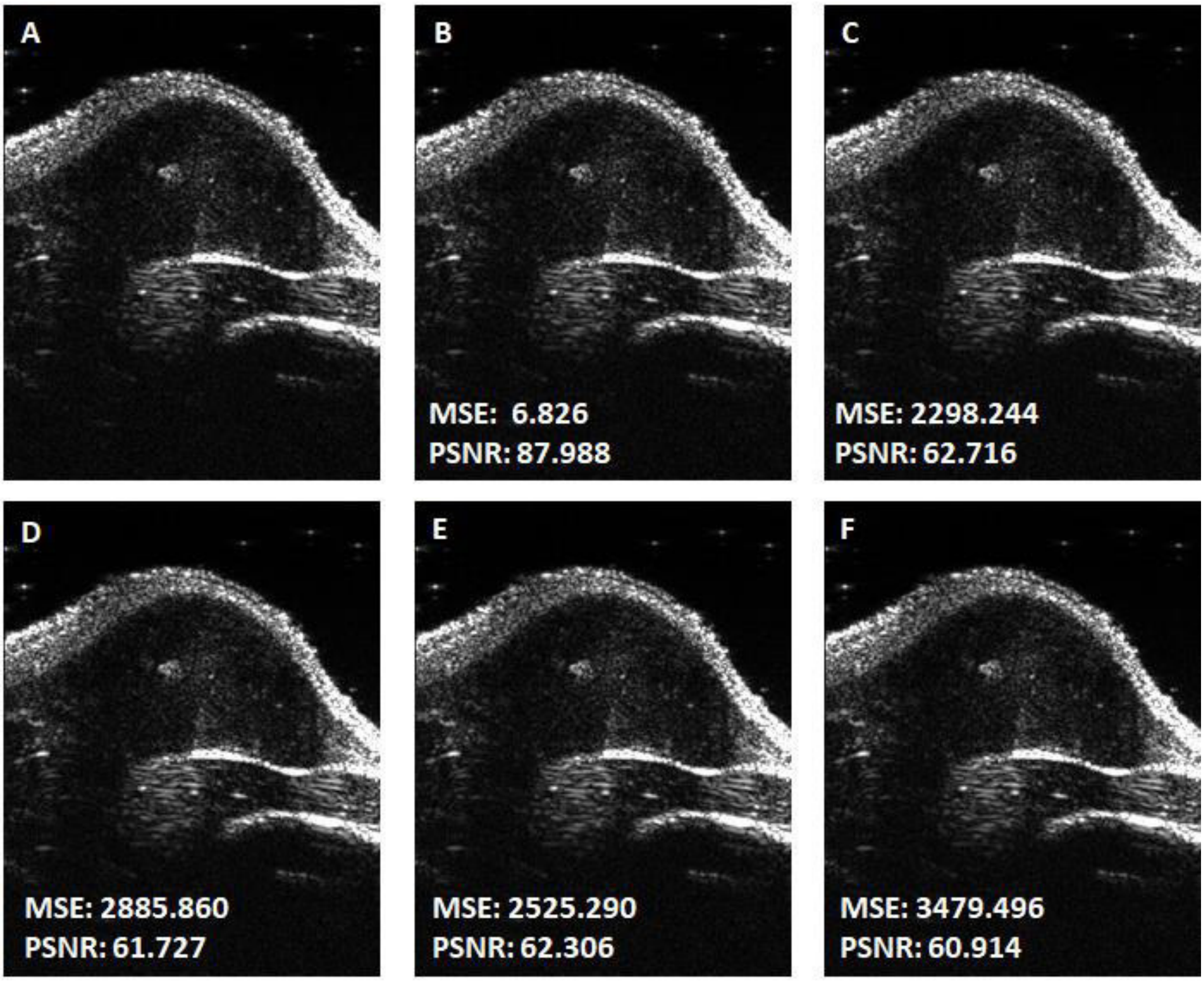


Figure 16: Speckle noise removal using different wavelets - quality comparison. Images in the figure present respectively: (A) original US image, (B) image decomposed with bior3.7 wavelet, (C) image decomposed with haar wavelet, (D) image decomposed with db10 wavelet, (E) image decomposed with sym5 wavelet and (F) image decomposed with rbio1.5 wavelet. The decomposition with each wavelet was followed by application of iBS threshold to the only 3rd DL. The computed values of MSE and PSNR are presented in the left bottom corner of filtered images.

The next part of establishing the best form of speckle noise removal algorithm was devoted to the selection of the threshold magnitude calculation and application. Threshold values were computed with 3 different algorithms. Each of them was used for the scaling coefficients of the detail subbands. The performance of gBS, iBS, and NS was evaluated with the calculation of mean values of MSE and PSNR derived for every thresholding variant. The results of this test were compared in Figure 17. Judging by the obtained values of mean MSE and PSNR, when all DLs were thresholded, the best findings were generated for gBS threshold used with every wavelet (Figure 17 A, C). However, the lowest MSE = 5231.711 and highest PSNR = 59.987 were obtained for bior3.7 wavelet. Thresholding of only the 3rd decomposition level yielded the best results for iBS threshold used with every wavelet (Figure 17 B, D). Furthermore, once again, the most suitable results (MSE = 38.923 and PSNR = 82.126) were obtained for bior3.7. Comparing the mean values of MSE and PSNR for image decomposed with bior3.7 wavelet followed by scaling with gBS at all DLs or iBS at only 3rd DL the best performance was demonstrated for iBS threshold applied to the 3rd DL. The presented findings prove that the best variant for speckle noise removal is composed of image decomposition with bior3.7 followed by scaling the detail coefficients at 3rd DL with iBS threshold (MSE = 38.923 and PSNR = 82.126). The BS threshold is frequently used for scaling of the detail coefficients. Furthermore, in the majority of studies BS threshold generates very good results, the images threholded with BS express low MSE and high PSNR values [32, 56]. It was proven in the presented study that the use of adaptively obtained BS (i.e. iBS) in the chosen denoising variant generates the best quality results. The similar findings were presented by Chang et al. [33]. The gold standard for detail coefficients thresholding in DWT has not been yet established. Therefore, multiple studies in the field of DWT concentrates on testing the new derivations of already established methods or introducing the new ones. The application of thresholding to all DL or just 3rd DL (i.e. last DL) has not been discussed yet in the literature. However, the authors of studies devoted to DWT investigate the number of plausible decompositions of the filtered images [18].

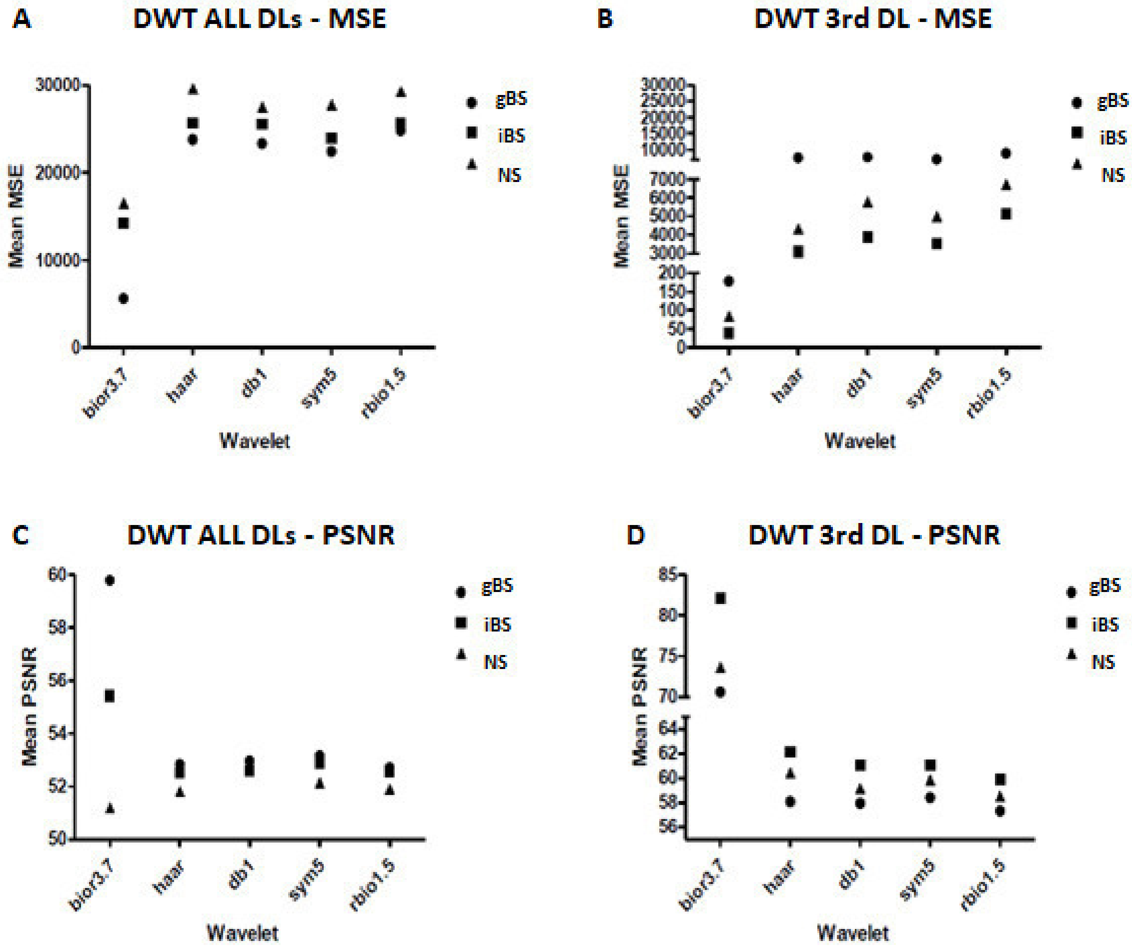


Figure 17: Evaluation of denoising algorithms. The figure shows the juxtaposition of mean MSE values computed for 3 different thresholds (gBS, iBS and NS), which were used for scaling the detail subbands at all DLs (A) or only the 3rd DL (B), for 5 different wavelets: bior3.7, haar, db1, sym5 and rbio1.5. The same comparison is shown for the PSNR calculation.

3.2. Cascade Classifier for Tumor Detection in CEUS Images

The main goal of the discussed procedure was to obtain AI based functions for the detection of tumors. The all generated CCs were prepared in MATLAB environment with 18-folded leaving-one-out cross-validation technique.

The study presented in this thesis concentrates on feature extraction and pattern recognition of subcutaneous tumors in CEUS images. Thus, functions for the detection and recognition of specific lesions had to be generated based on the given database. Wherefore, the algorithm for automated tumor detection based on the CCs was developed. The CCs enable fast and precise object detection. Furthermore, they can be generated for any type of object. The CCs were created and are mainly used for face detection purposes [43, 46, 48, 58]. However, they have never been developed for detection of subcutaneous tumors. Providing the diverse and large training images database is the condition one must fulfill to create the robust CCs. Initially, the image database composed of only 18 samples was given (Supplementary Figure 1). The tumors in this data set varies in size and shape. Nevertheless, every image encloses only one tumor that is always situated in the middle of the image. Therefore, the initial image database constitutes a week learning capacity for the CCs. The above-mentioned problem was solved by preparing the new training data set (Table 4). The positive and negative samples were extracted from single, horizontally and vertically concatenated doubled and quadrupled images (Figure 9). Thereby, in the end the training data set was composed of 2700 images in total (900 positive and 1890 negative). The built-in MATLAB functions were used to develop the CCs. Unfortunately, it occurred, that these functions have restrictions concerning the size of training samples. Namely, they cannot be smaller than 20x20 pixels. Therefore, the samples from prepared training data set had to be excluded. To overcome above-mentioned restrictions, development of CCs without the use of built-in MATLAB functions must be considered. Thus, the smallest size of training sample could be defined by the user. Mostly the negative images were excluded due to the size restriction. The image regions (Gel, Skin and Tissue - Figure 9 A) from which the negative samples were cropped have smaller dimensions in comparison to the whole image. Therefore, some of the extracted samples were smaller than 20x20 pixels. In the end, the training data set consisted of 1735 images in total (900 positive and 835 negative). Naturally, the negative images from widely available image libraries (e.g. The MIT-CSAIL Database) could have been used to train the CCs. The external image library was not used because the main purpose of the developed CCs was to recognize the ROI inside the MLS, A431 or A459 tumor type. Hence, the set of negative images, had to be composed of the particular samples. The results of CCs performance are presented in Figure 18. Each detection function was running until only one object was detected. In 16 out of 18 images the tumor was detected correctly, like in the example presented in Figure 18 A The false detections were obtained for only one pair of images (see Supplementary Figure 1), and one of them is presented as an example of unsuccessful tumor recognition in Figure 18 B.

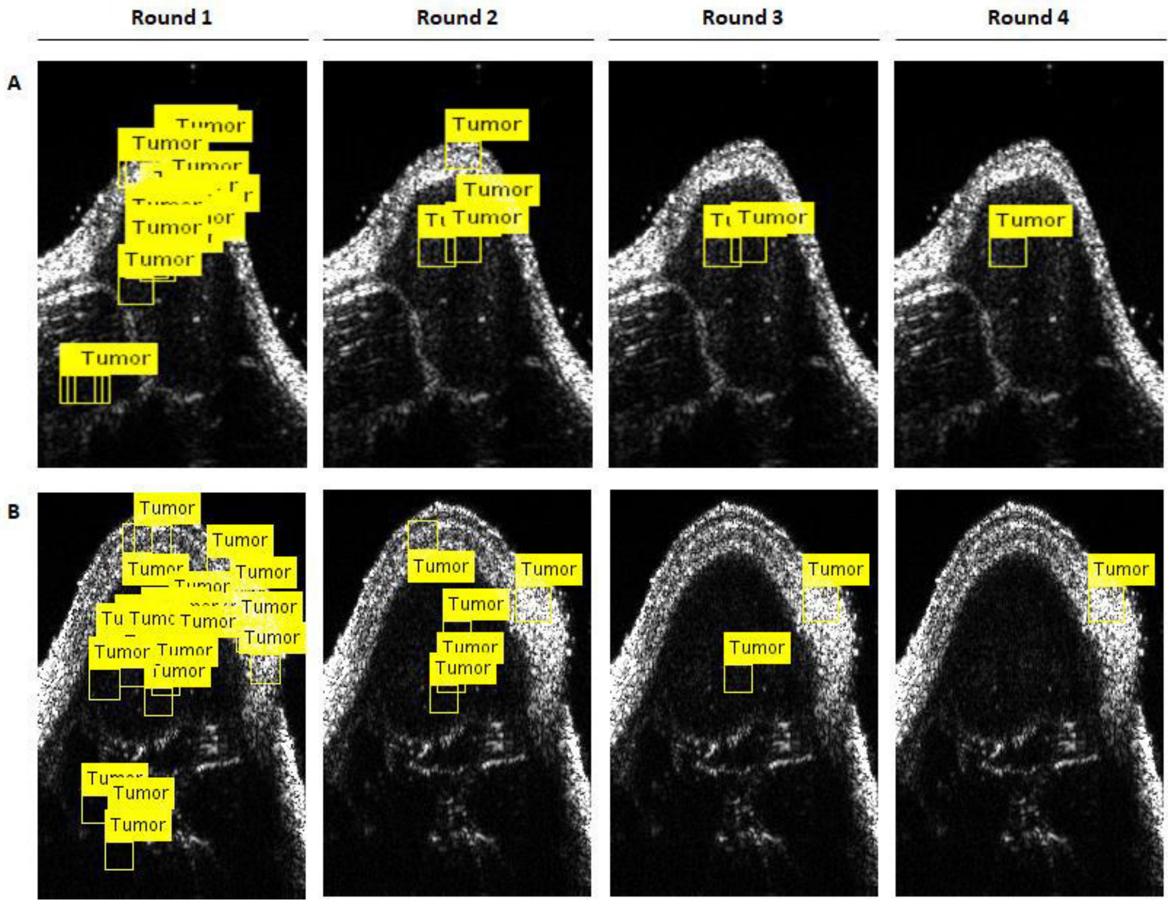


Figure 18: Automated tumor detection. The tumor detection with the use of CC was executed multiple times. Two examples are shown, one for positive (A) and and for false (B) detection of tumors. At each round the number of detected ROIs decreased (Round 1-4). The weak false positive detections were rejected after each detection round. In the end, the strongest true positive detection remained(A - Round 4) and was used as seed point for the following tumor segmentation.

This situation may have been caused by providing an insufficient number of positive and negative examples in training data set. In the end, using presented framework for automated tumor detection yielded in good results. The performance of created functions was highly satisfying since 89% of provided samples were recognized correctly. Furthermore, the concatenation of the images enabled the developed CCs to recognize multiple tumors placed in various locations (Figure 19). The detection results for the rest of the tumors is presented in Supplementary Figure 1.

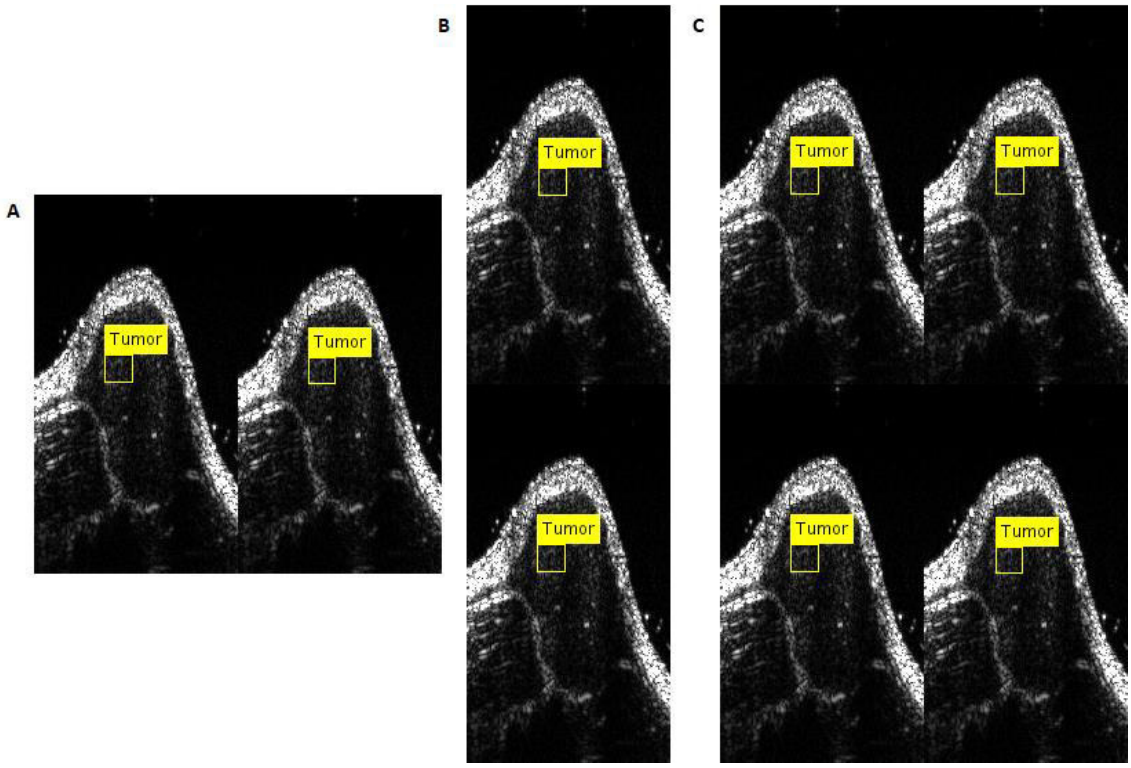


Figure 19: Tumor detection. Exemplary detections of multiple tumors in different positions. Every developed CC can detect two (A, B) or more (C) tumors in different locations.

Taking into consideration further improvements, which can be implemented in the presented automated tumor detection framework, the most important one is the preparation of a diverse and broad data set of images. It was already mentioned that the initial data set was not enough to train the robust CCs. Other researchers, who did similar studies, were equipped with databases comprising radiological images of e.g. 500 individuals obtained with various imaging modalities [2, 59]. Even though, the developed tumor detection functions showed high accuracy in lesions recognition (89% of positive detections), they were developed only for the specific tumor examples. Addressing the problem of small database of images would allow us to reduce the false discovery rate among different lesions types [60]. Furthermore, we will be able to design our study more accurately by dividing images into training and validation datasets. The development of tumor detection function with different programming language than MATLAB (e.g. C++) should be considered as well.

3.3. Automated Segmentation of Tumor in CEUS Images

The first and the most important step was to exclude the user interaction from the tumor segmentation procedure. Promising findings obtained with developed automated tumor detection (89% of positive detections), allowed for fully automated tumor ROI detection. Therefore, the base for development of automated segmentation algorithm was constituted.

After detecting the tumor in the input image, the drawn box was converted into a binary mask (Figure 10 A). This mask was further used as a starting point of the lesion segmentation. Furthermore, in this part the second mask was created for reduction of image area to only the tumor region (Figure 10 B). The developed functions for edge details removal (Figure 12) and edge repair (Figure 13) rely on pixel-based operations. Hence, the smaller the image the faster was its processing. Moreover, the provided images were reached in low intensities, what could have influenced any adaptively obtained values (i.e. ST for HEM calculation). Therefore, cropping out the tumor region from original images helped in faster and more efficient image pre-processing. Strong smoothing with median filter was done in the view of following edge map generation (Figure 11). Having a more blurred picture resulted in drawing the more uniform tumor edge. Since the provided dataset was composed of CEUS images of subcutaneous tumors, they have very well emphasized skin region in contrast to other segments. Therefore, generated edge masks contained well-structured boundary only on the top of the lesion. Taking into consideration following tumor segmentation, the prepared edge map should be as uniform as possible, to prevent any possible leakage of growing region. Thus, the function for edge reconstruction was developed. The HEM was an input image for the following edge repair, since it contained less details, its ST was higher (Figure 20 B). On the other side was VEM calculated with lower ST, thus it contains more details (Figure 20 C). The computed VEM was used to repair missing parts of HEM (Figure 13). However, before the tumor boundary was reconstructed, the details, which were not a part of the surrounding tumor edge, had to be removed. Thereby, the details removal algorithm was created. Without erasing those details, the future tumor segmentation could have been disturbed. The results of all operations conducted for the refinement of HEM are presented in Figure 21.

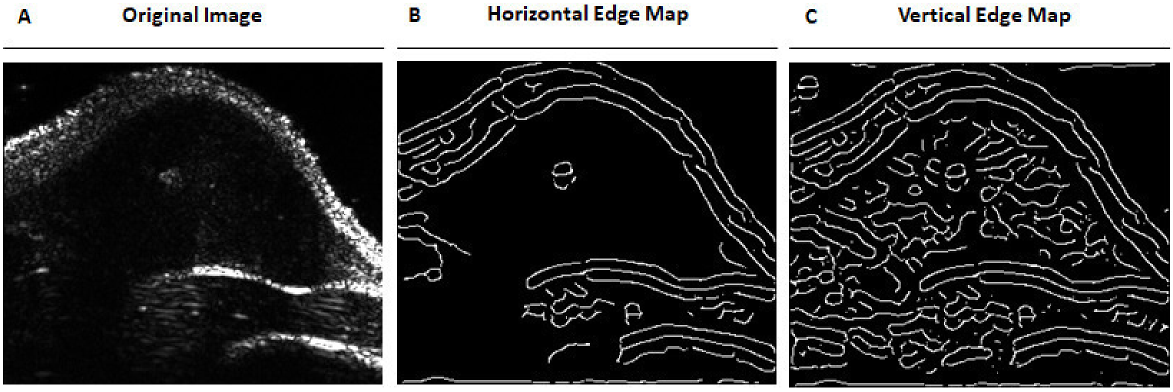


Figure 20: Exemplary edge maps. Based on the US image (A) a horizontal edge map (B) and vertical edge map (C), with different sensitivity threshold values were generated.

One can see, that the original edge of both examples contained gaps in the tumor bottom boundary and some redundant details in the middle of ROI (Figure 21 A). Therefore, both inputs were subjected to a 1st round of details removal. The redundant edge part was successfully removed in both cases (Figure 21 B). Afterwards, processed HEMs underwent edge repair procedure. In the case of the Example 1, the gap present in lower lesion boundary was well reconstructed (Figure 21 C). However, one can clearly see that HEM of the Example 2 was already well prepared after the 1st detail removal step. Hence, edge repair procedure was unnecessary. Moreover, it resulted in computing additional details, which covered some ROI area (Figure 21 C). The last part of edge refinement was subjecting the repaired edge maps to the second round of details removal. This step was implemented because in some cases the edge reconstruction might have resulted in drawing too many details, as it is presented with the Example 2 (Figure 21 C). One can see, that only minor details were removed in the Example 1 (Figure 21 D). Even though, the edge details were removed for the second time, it did not result in final improvement of tumor boundary in the Example 2 (Figure 21 D). Thereby, one can clearly see that the function for controlling the state of refined edge map should be introduced to the developed algorithm. It will help to avoid subjecting already well-established HEMs to the multiple edge repair or details removal steps.

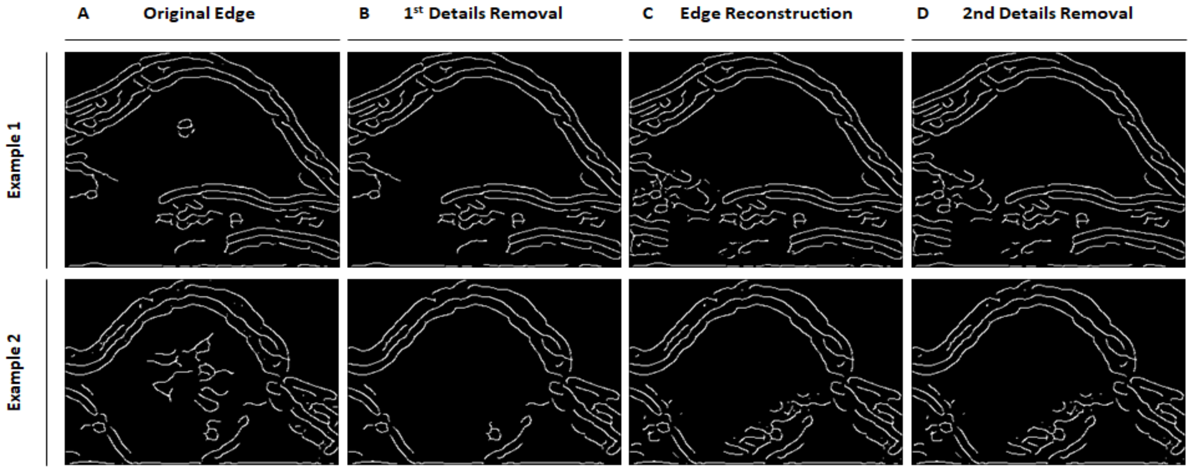


Figure 21: Exemplary edge map refinements. Presented figure enclosed examples of two calculated HEMs (Example 1 and 2), which underwent: 1st Details Removal (B), Edge Reconstruction (C) and 2nd Details Removal (D) respectively. All the steps procedures depicted above are comprised in the established HEM refinement procedure.

Before the actual tumor segmentation, morphological operations were used to create final tumor mask from the final HEM. Afterwards, the final mask was superimposed on the original tumor image. Every tumor was automatically segmented by the developed algorithm, and afterwards each lesion was delineated by three different users. The comparison of the tumor segments extracted by different parties are presented in the Figure 22, Figure 23 and in the supplement (Supplementary Figure 5-20). The following results enclose the positive segmentation of tumor in case of using the automated segmentation algorithm (Figure 22 A). It was achieved, because of the positive detection of ROI. The developed algorithm succeeded in extraction of the tumor region very similar to the ones delineated by the users (Figure 22 B, C and D). The difference in the look among presented segments is not substantial. However, it was already mentioned that not all developed CCs succeeded in the correct recognition of the tumor. Therefore, in the case of false region detection, the segments extracted by the developed AS algorithm does not reflect the actual appearance of the tumor (Figure 23 A). For a better insight of the proposed AS framework performance, the study for comparison of extracted segments area among the users (Figure 24 A) and between users and AS algorithm (Figure 24 B) was done. One can see that the areas of the tumors segments extracted by the users are very similar. The range of overlap percentage varies between 85 % and 100 % (Figure 24 A).

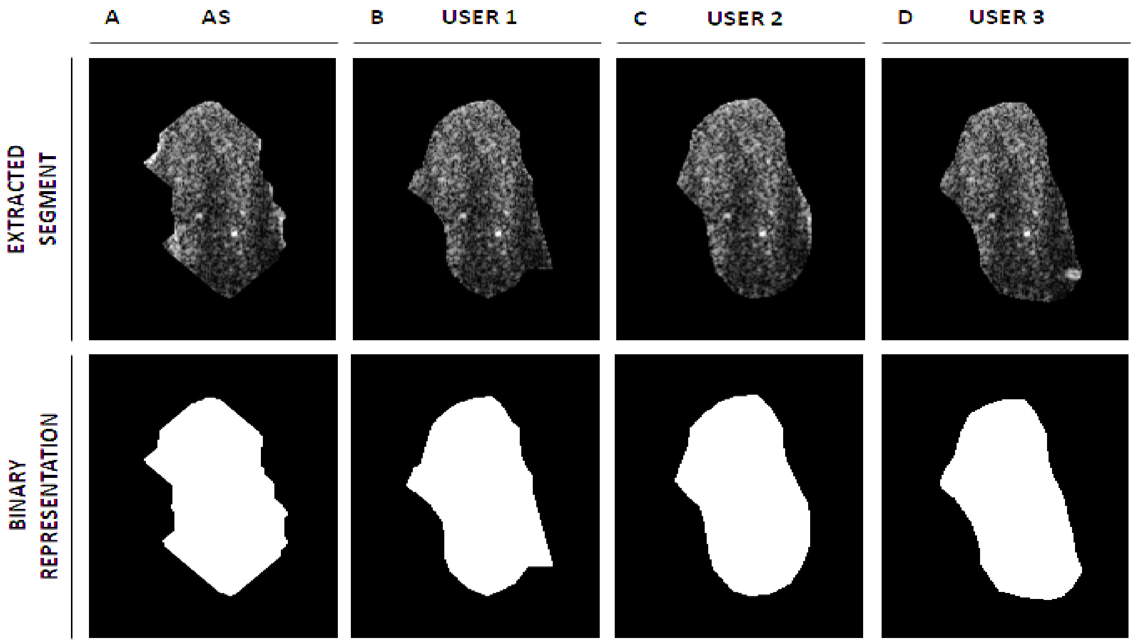


Figure 22: The comparison of extracted tumor segments - positive example. The presented segments were extracted by AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D). One can see that all enclosed segments are very similar.

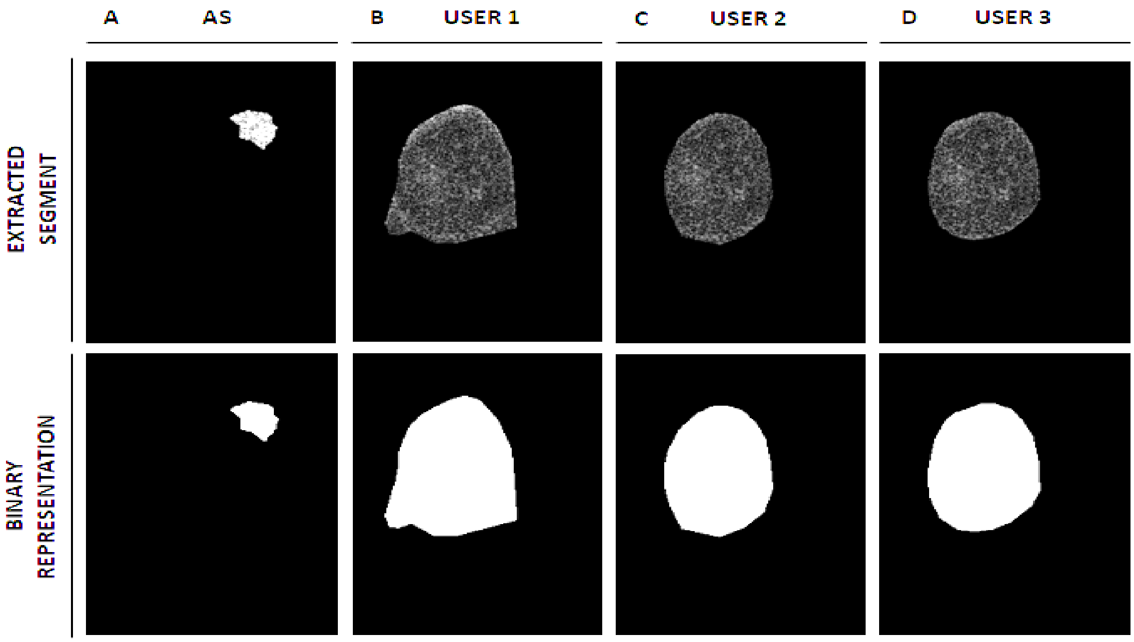


Figure 23: The comparison of extracted tumor segments - negative example. The presented segments were extracted by AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D). One can see that the segment extracted by AS algorithm (A) is different in comparison to segments delineated by the users (B, C and D).

The highest degree of variability is expressed in the segmentation of tumor no. 5 among different users (Supplementary Figure 9). The segmented area of tumor no. 8 was almost the same for every user (Supplementary Figure 12). In the bottom graph, one can see that the overlap percentage for tumors with false object detection, segments 6 and 17, yields in the lowest concordance in between AS and users (Figure 24 B). The images of above mentioned tumors segments are presented in Supplementary Figure 10 and Figure 23 respectively. Moreover, the overlap percentage results obtained for segments 2 and 5 are not so high, even though, tumors were correctly recognized in these images. Such outcome might have arisen from the process of final mask preparation. In some cases, processing of refined HEM with morphological operations resulted in over-thickening of lesion boundaries. Hence, the final mask comprised smaller than original ROI area and this influenced the look of the extracted segment (Supplementary Figure 6 and 9). The morphological processing of final HEMs had to be introduced to create the masks with uniform tumor boundary. The developed edge repair algorithm managed to reconstruct the computed HEMs fully in the majority of the cases. However, for some examples the full edge repair was not reached. Hence, the morphological processing of computed HEMs had to be introduced. Once again, the need for development of the additional edge state controlling function is arising. Even though, developed AS framework yields in only a few false segmentations, the overall performance of the algorithm is more than satisfactory. The developed AS framework succeeded in obtaining correct segments for 89% of the cases. Moreover, the results of the tumor overlap study showed that 14 out of 18 segments expressed high overlap concordance (from 80% to 100%) in between AS algorithm and users.

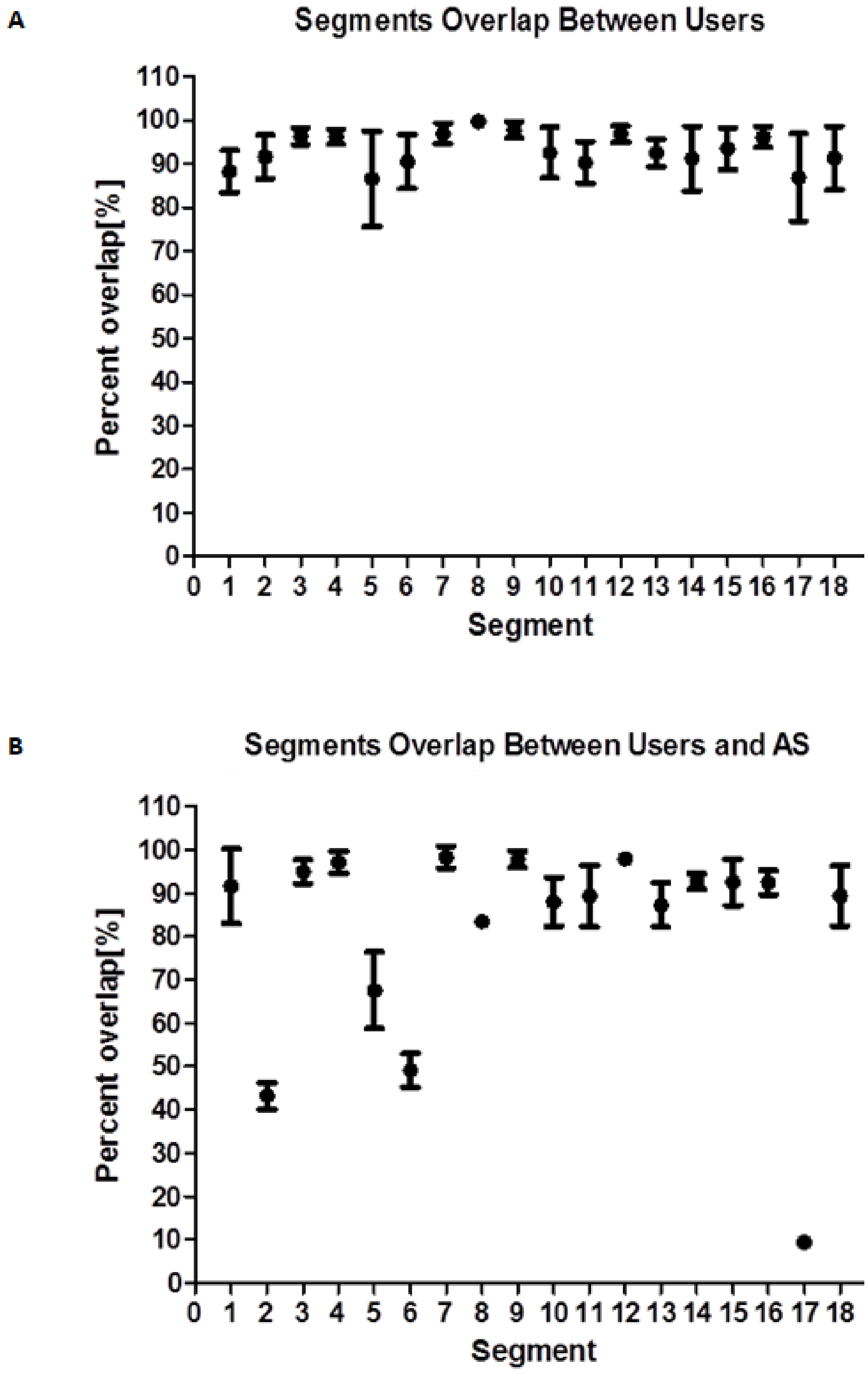


Figure 24: Evaluation of automated segmentation. The enclosed graphs show the percentage overlap (y - axis) of tumor segments (x - axis) between different users (A), and between users and the AS framework (B).

3.4. Tumor Differentiation Model

The main objective of the presented results was to choose the valuable framework for the generation of tumor classification model. Furthermore, to the validate the precision of a chosen RS. The findings presented in this section are discussed with reference to results obtained for segments delineated by User 1. The remaining results for User 2 and User 3 were included in the supplement (Supplementary Figure 21-28).

Firstly, intensity-based, textural and wavelet features were extracted for each US image. The shapes of extracted ROIs were not taken into consideration since the radiomic analysis was done on 2D US images of subcutaneous tumors. The actual shape of the whole lesion could not have been analyzed. Thereby, the shape-based features were not considered as a qualitative feature for the tumor differentiation. Even though, some authors consider morphology of a tumor as a trait for distinguishing its, i.e. malignancy [3], in case of this study usage of the shape-based features could have introduced bias. Recently textural features are claimed to have the greatest differentiating power because of their consistency and ability to reflect the patterns of different parts of the lesion. Furthermore, the textural features may support the recognition of tumor heterogeneity, which is associated with the worse prognosis [4]. We extracted 230 features (14 intensity-based, 32 textural and 184 wavelet). Afterwards, the ranks of extracted features were computed by carrying out multiple statistical tests, namely: test-retest, Friedman test and one-way ANOVA. The final ranking was obtained by averaging results from all tests. In the end the most stable feature was assigned to 1 and the least stable to 230. The final ranking of the features shows that the textural features, from both SOS and WT, have higher stability in comparison to the features from FOS group (Supplementary Figure 2-4). The ranked traits were then used to create the RS no.1 and RS no.2. The RS no.1 was made of the highest ranked features from each group (Table 11). Furthermore, the linear dependency of the 3 selected features was investigated. The test showed that 2 out of 3 features in RS no. 1 are correlating ($p < 0.05$). One can see that Short Run Emphasis feature appears twice in RS no.1 (Table 11). Even though, the first one belongs to the SOS and the second one to the WT feature group, they are still very similar.

Both groups are composed of textural features. The only difference between them is such that textural features in the WT group were extracted from tumor segments, which underwent the undecimated DWT. The values of pixels intensities in resultant detail coefficients matrices are different in comparison to the ones in original segments. However, the difference is not big enough to influence the values of the extracted features.

Table 11: Radiomic Signature no.1.

	RS no.1		
Feature	Uniformity (FOS)	Short Run Emphasis (SOS)	Short Run Emphasis LL (WT)
Final Rank	73	12	6

In contrary to RS no1, the selection of the features for the RS no2 was based on the lack of linear dependency. Thereby, the first highest ranked features, with low or no correlation ($p > 0.05$) were selected for RS no.2 (Table 12). One can see that all the features, comprised in RS no.2, are different. Thus, it is already less probable for them to correlate. Moreover, in comparison to RS no.1, RS no.2 consists of features with lower ranks.

Table 12: Radiomic Signature no.2.

	RS no.2		
Feature	Median (FOS)	Correlation (SOS)	Short Run Emphasis LL (WT)
Final Rank	101	95	6

One can see, the selected features for both RSs comes from different feature groups (Table 11 and 12). It was already proven in other studies devoted to radiological analysis, that combining features, which were extracted with different methods (Table 1), makes the tumor classification model more discriminative [2, 3]. Moreover, the findings of the study presented by Guo et al. [3] states that, the tumor classification model composed of strongly correlating features is highly accurate in predicting the characteristics of the breast invasive ductal carcinoma. However, the discriminative power of the tumor classification model composed of features with low or no correlation has not been investigated yet. The validation of both

RS no.1 and RS no.2 was done with the use of unsupervised and supervised classification algorithms. The first comparison was presented with the use of hierarchical agglomerative clustering (Figure 25). To state, that the prepared classification models discriminate between 3 distinct tumors (MLS, A431 and A459), one should be able to distinguish 3 major clusters in the prepared dendrograms. In both examples the dendrograms are presented at the left side of the heat map (Figure 25 A, B). The unsupervised classification algorithm clusters the tumor models based on the distance between the features values. The algorithm assumes that the similar features values correspond to particular tumor model. In this study Manhattan method was used to calculate the distance between the feature values. The length of the dendrogram branches reflects the distance (i.e. similarity) between classified tumors [61]. The links between branches placed closer to the left side of the dendrogram represents the clusters situated further away from each other. To show the final results of unsupervised classification, the dotted line was drawn between linkages that are the second and the third to the last (Figure 25 A, B). The tumor clusters are easier to distinguish, when the distance between these linkages is bigger. One can see, that in the case of RS no.1 it is difficult to distinguish between 3 tumor models. Mainly, the two clustered groups in the bottom of the dendrogram are very similar. On the other hand, in the dendrogram of RS no.2 one can see that the gap between the second and the third to the last linkage is bigger. Thus, it is easier to identify 3 tumor clusters. The initial clustering of tumor models, at the bottom of both dendrograms, is mostly correct for both RSs (Figure 25 A, B). However, the final clustering is not correct. Results of hierarchical agglomerative clustering are strongly dependent on the choice of subject’s similarity estimation method (i.e. distance between objects) and linkage algorithm [61]. The Manhattan method was implemented since it yielded in the best clustering outcome in case of investigated differentiation models. Nonetheless, obtained results do not discriminate well between the tumor models (Figure 25). In the view of presented findings using hierarchical agglomerative clustering method for validation and comparison of generated RS no.1 and RS no.2 is not a good solution.

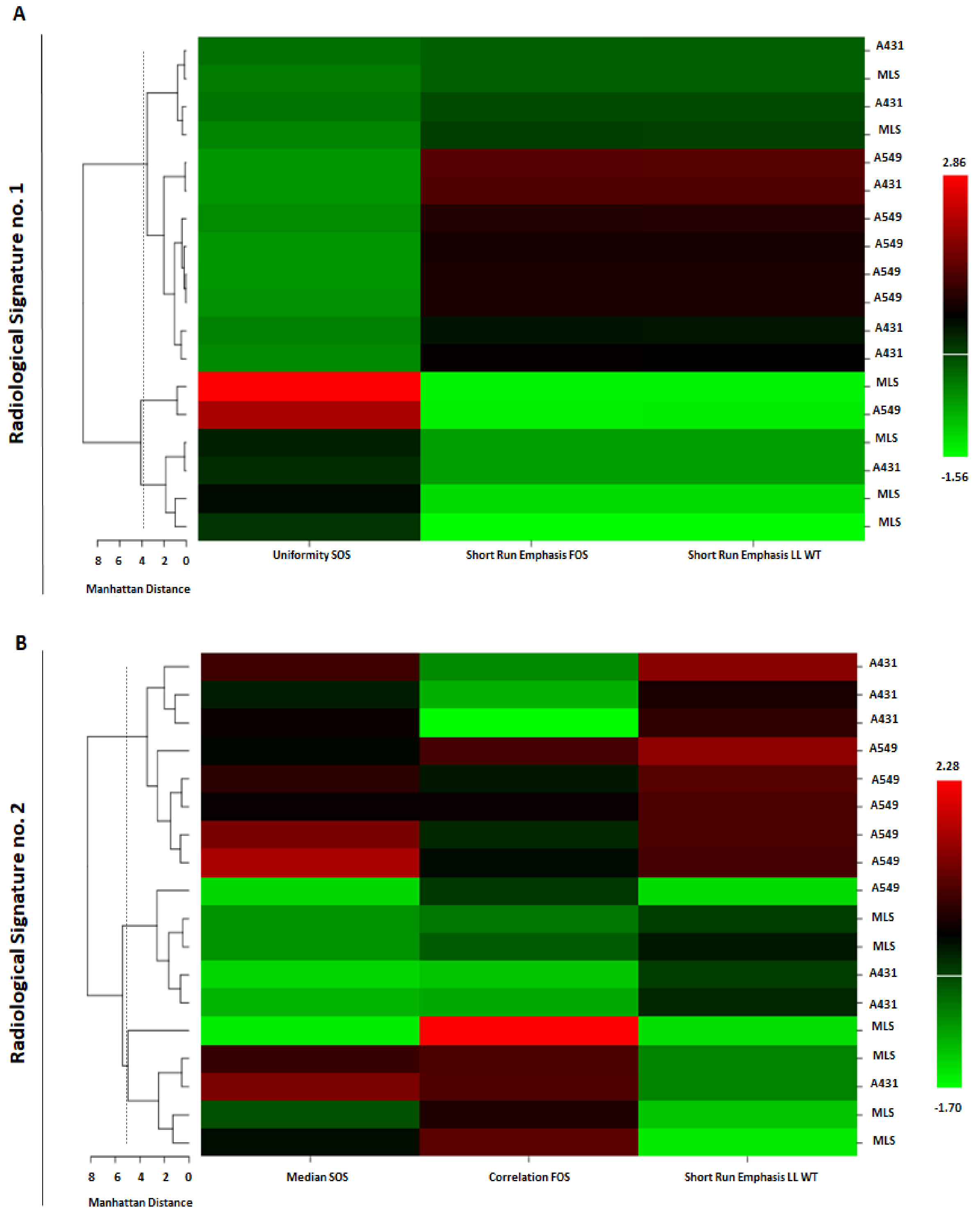
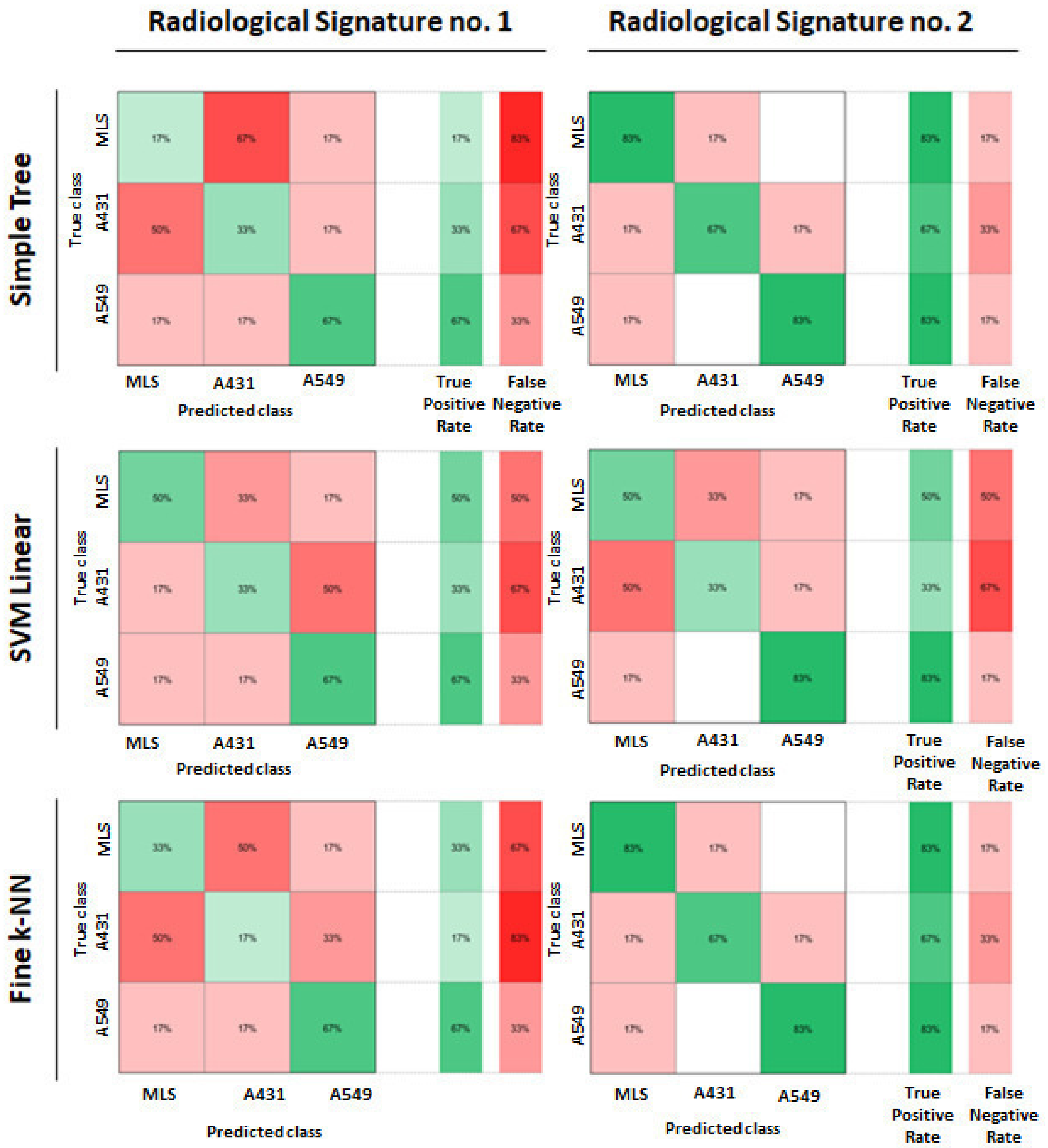


Figure 25: Hierarchical agglomerative clustering - unsupervised validation of tumor classification models. In both (A) and (B) clustering of features is presented on the left side with dendrogram and assigned tumor types on the right side of the computed heat map. The linkages above the drawn dotted line constitutes 3 final clusters.

Another validation of the assembled tumor differentiation models was done with the use of supervised classification algorithms implemented in the MATLAB software. The Classification Lerner App allows testing several groups of supervised learning algorithms to train and estimate the accuracy of the prepared classification model. In the beginning, the application chooses by itself the most robust and suitable algorithms for the input data validation. In this study Simple Tree, linear SVM and fine k-NN models were used, since they lead the highest accuracy out of all algorithms suggested by Classification Lerner App. The results of RS no.1 and RS no.2 validation are presented in Figure 26. The performance of tested classification models was presented in the form of confusion matrices generated for both RS. Each true positive classification was marked with green color and false negative one with red color. In the view of outcome presented for Simple Tree algorithm one can see that in the case of RS no.1 the majority of samples from the first and the second group were misclassified with high false negative rate. On the other hand, outcome presented in the second confusion matrix shows that validation of RS no.2 with Simple Tree yielded to much better results for these tumors groups. Furthermore, in this case samples from all tumor clusters were classified with high true positive rate. The supervised classification with linear SVM gave almost the same accuracy for both RS. However, separation of samples from the third cluster produced preferable output for RS no.2. The nearest neighbors' model was the last tested classification algorithm. Analyzing the content of presented confusion matrices, one can clearly see that separation of samples with the use of RS no.2 yielded to much higher true positive rate in comparison to RS no.1. The confusion matrices, presented in Figure 26, depict what was the percentage of correct classifications of tumor models from certain group, while using different classification algorithms. However, these findings do not reflect the overall probability of samples true positive classification with used learning classifiers. Calculation of Wilson Score Interval [62] will give the probability of how many tumors, out of 18, were classified correctly with certain type of classification algorithm. The results of this test are presented in the form of 95% confidence interval (CI). It points out the numerical range in which the true positive classifications should be comprised with the 95% certainty (Table 13).



Simple Tree — Radiological Signature no. 1

True class	MLS	A431	A549	True Positive Rate	False Negative Rate
MLS	17%	67%	17%	17%	83%
A431	50%	33%	17%	33%	67%
A549	17%	17%	67%	67%	33%

Simple Tree — Radiological Signature no. 2

True class	MLS	A431	A549	True Positive Rate	False Negative Rate
MLS	83%	17%		83%	17%
A431	17%	67%	17%	67%	33%
A549	17%		83%	83%	17%

SVM Linear — Radiological Signature no. 1

True class	MLS	A431	A549		
MLS	50%	33%	17%	50%	50%
A431	17%	33%	50%	33%	67%
A549	17%	17%	67%	67%	33%

SVM Linear — Radiological Signature no. 2

True class	MLS	A431	A549	True Positive Rate	False Negative Rate
MLS	50%	33%	17%	50%	50%
A431	50%	33%	17%	33%	67%
A549	17%		83%	83%	17%

Fine k-NN — Radiological Signature no. 1

True class	MLS	A431	A549		
MLS	33%	50%	17%	33%	67%
A431	50%	17%	33%	17%	83%
A549	17%	17%	67%	67%	33%

Fine k-NN — Radiological Signature no. 2

True class	MLS	A431	A549	True Positive Rate	False Negative Rate
MLS	83%	17%		83%	17%
A431	17%	67%	17%	67%	33%
A549	17%		83%	83%	17%

Figure 26: The comparison of supervised classification algorithms accuracy in training discriminative tumor models. (A) Presented the results obtained for training RS no.1 with Simple Tree, SVM Linear and Fine k-NN algorithms respectively. (B) Presents the results obtained for training RS no.2 with Simple Tree, SVM Linear and Fine k-NN algorithms respectively. The green color assigns true positives classification. The red color assigns false negative classification.

Table 13: Wilson score interval for estimation of tumor model classification probability.

	RS no.1		RS no.2	
	p	**95% CI**	**p**	**95% CI**
Simple Tree	0.400	0.200 - 0.614	0.780	0.550 – 0.910
Linear SVM	0.500	0.290 - 0.710	0.560	0.337 – 0.755
k-NN	0.400	0.203 - 0.614	0.780	0.550 – 0.910

The results presented in Table 13 shows that the probability of correct tumor classification out of 18 provided samples, in case of using all presented learning algorithms, is higher for RS no.2. Moreover, both Simple Tree and k-NN learning algorithms yielded the best results in training the RS no.2. These findings imply that RS no.2 is a more suitable tumor classification model for MLS, A431 and A549 discrimination. The chosen RS is comprised of 3 features: Median (FOS), Correlation (SOS) and Short Run Emphasis (WT). The values of these features may represent the different characteristics of classified tumor models (Figure 27). For instance, the Median (FOS) feature is giving the information about the typical pixel intensity in extracted segments of each tumor model. One can see that A549 comprise the highest intensity of the commonly appearing pixels. While, the lowest median was obtained in the case of MLS (Figure 27 A). The difference in the median values might imply that the cellular density is changing among the different tumor models. However, this has not been further investigated in presented study. The values of Correlation (SOS) feature for analyzed tumor models are presented in Figure 27 B. In comparison to the intensity-based features, the textural features may have ability to describe the composition of the particular tumor. The values of correlation describes the linear dependency among the intensity values in the tumor segment [63]. In the case of particular tumor model, it can indicate how similar are the cells in studied lesions. Thereby, there is a possibility that the higher the correlation the more homogenous the lesion is. On the other hand, the lower the correlation the higher the heterogeneity of the tumor. Following the above assumptions, one can see that the A431 tumor model is the most heterogeneous among the studied tumors (Figure 27 B). The correlation values obtained for MLS and A459 indicate that there are more homogenous in comparison to A431 (Figure 27 B). However, one must take into consideration that it is only assumption based on the interpretation of correlation coefficients values. The last feature comprised in RS no.1 is the Short Run Emphasis LL (WT). These feature was introduced by Galloway et al in [50]. It describes the short sequences of the same intensities. One can see that the value of Short Run Emphasis LL (WT) feature is the highest for A549 and the lowest for MLS tumor model (Figure 27 B). However, the difference in magnitudes of this particular features among analyzed tumor models is not substantial. In the study presented by Tang et al. [64], it is stated that the long gray-level runs would occur more often in the coarse texture patterns than in the fine ones. Thereby, one can conclude that the principal of short gray-level runs appearance is opposite to the mentioned statement. Hence, the finer the texture, the more short gray-level runs is comprise in it. This assumption might imply that the A549 has the finest texture, hence, its cellular density might be lower in comparison to other tumors. Nonetheless, this is only an assumption. To sum up, the presented findings proved that the tumor classification model composed of features with low or no correlation (RS no.2) has greater discriminative power in comparison to the one composed of correlating features (RS no.1). Therefore, the lack of similarity among the radiomics features should be investigated, while preparing the tumor classification models. It was also shown that the values of statistical features comprised in the chosen RS might have a biological interpretation. However, the possible biological meaning of the features values was not investigated further in presented study.

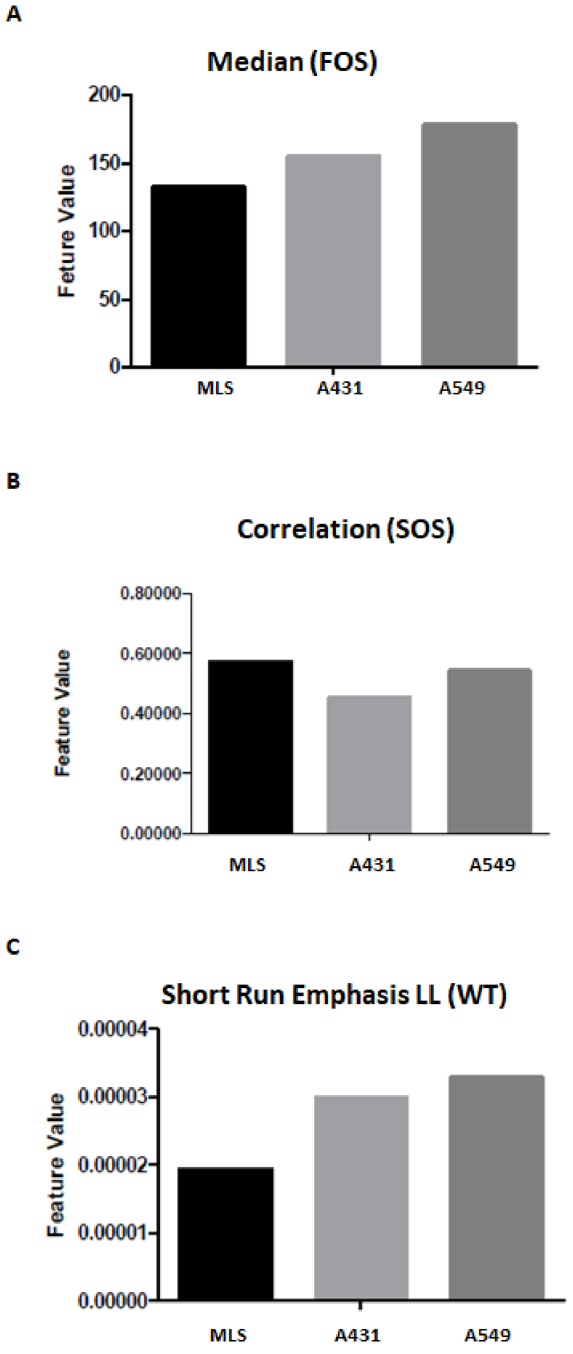


Figure 27: Values of features enclosed in RS no.2 for MLS, A431, and A549 tumor models. The values of Median intensity-based feature (A), Correlation textural feature (B) and Short Run Emphasis LL the wavelet feature (C) are presented for each studied tumor model (MLS, A431, A459).

4. Conclusion and Perspective

4.1. Conclusion

The findings presented in this master thesis showed that proposed framework for pattern recognition and feature extraction in US images is promising. Starting with the speckle noise removal, it has been concluded that DWT is currently the most suitable method for US images denoising. Furthermore, the selected form of filtering (bior3.7 wavelet with iBS applied to the 3rd DL) proved to be the best performing variant of speckle noise filtering in US images. Taking the outcome of the automated tumor detection into consideration, it can be stated that the proposed framework succeeded in recognizing lesions in the provided images. The usage of CCs allowed fast and robust lesion identification, recognizing the desired ROIs in 89% of the analyzed images. A larger dataset of US images might be helpful to improve these results in further studies. Summarizing the outcome of the automated segmentation algorithm, it can be concluded that the developed framework is working when the tumor is detected correctly. When the tumor is detected correctly the segments computed by the AS algorithm and segmented by the users overlap between 80% and 100%. In the cases where the tumors were wrongly detected, the extracted segments have the smallest overlap percentage from all (Figure 24). Furthermore, introducing the function for edge refinements control should be considered to improve the results of the developed AS algorithm. The tumor detection seems to be the most crucial step for the functioning of this algorithm, hence, corrections at this stage will have the biggest impact. More advanced AI algorithms might be able to improve the reliability of lesion detection. Moving to the part of the study devoted to the composition and validation of a tumor classification model, the performed tests lead to the isolation of the most suitable radiomic signature (RS no. 2). It was proven that supervised learning algorithms performed their classification task better when linear independent imaging biomarkers were used (both Simple Tree and k-NN with $p=0.780$ [95% CI 0.550-0.910]). It was emphasized in this work, that outcomes of critical steps like e.g. automated tumor detection are strongly dependent on the number and diversity of provided images since, they constitute its learning capacity. Moreover, in the case of imaging biomarkers extraction, more versatile images samples can lead to computation of more distinctive features comprised in the created classification model. The developed radiomic analysis obtained the best possible results of MLS, A431, and A549 tumor model classification. Nonetheless, we are aware that the performance of the generated AI-based functions could be most likely improved with the broader and more diverse training data.

4.2. Future Perspective

The radiomic analysis of medical images might have a great potential to support the classification of tumors in the clinics. However, to develop more accurate and impartial systems, much more images should be provided. Analyzing the clinical imaging data of numerous individuals diagnosed in different medical centers is crucial for generating more precise and impartial tumor classification models. Furthermore, providing the broader and more diverse images data sets will support the development of more robust AI-based functions and automated segmentation algorithms. Combining medical imaging data with other clinical information i.e. tumor vascularity has a potential to make radiomics even more powerful. Moreover, the radiogenomics, which is combination of radiomics and genomics, might be able to support the assessment of disease progression and the patient treatment response [4]. Introducing, the radiomics to the clinics might help to support the current diagnostic procedures and contributed to development of personalized medicine.

5. Index

5.1. List of Figures

5.2. List of Tables

6. References

1. Andrekute, K., et al., *Automatic Differential Diagnosis of Melanocytic Skin Tumors Using Ultrasound Data.* Ultrasound Med Biol, 2016. **42**(12): p. 2834-2843.
2. Aerts, H.J., et al., *Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach.* Nat Commun, 2014. **5**: p. 4006.
3. Guo, Y., et al., *Radiomics Analysis on Ultrasound for Prediction of Biologic Behavior in Breast Invasive Ductal Carcinoma.* Clin Breast Cancer, 2017.
4. Gillies, R.J., P.E. Kinahan, and H. Hricak, *Radiomics: Images Are More than Pictures, They Are Data.* Radiology, 2016. **278**(2): p. 563-77.
5. Lambin, P., et al., *Radiomics: extracting more information from medical images using advanced feature analysis.* Eur J Cancer, 2012. **48**(4): p. 441-6.
6. Larue, R.T., et al., *Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures.* Br J Radiol, 2017. **90**(1070): p. 20160665.
7. Ardakani, A.A., A. Gharbali, and A. Mohammadi, *Classification of breast tumors using sonographic texture analysis.* J Ultrasound Med, 2015. **34**(2): p. 225-31.
8. Song, G., F. Xue, and C. Zhang, *A Model Using Texture Features to Differentiate the Nature of Thyroid Nodules on Sonography.* J Ultrasound Med, 2015. **34**(10): p. 1753-60.
9. Shung, K.K., *Ultrasound: Past, Present and Future*, in *The Third International Conference on the Development of Biomedical Engineering in Vietnam: BME2010, 11–14 January, 2010, Ho Chi Minh City, Vietnam*, V. Van Toi and T.Q.D. Khoa, Editors. 2010, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 10-13.
10. Chan, V. and A. Perlas, *Basics of Ultrasound Imaging*, in *Atlas of Ultrasound-Guided Procedures in Interventional Pain Management*, S.N. Narouze, Editor. 2011, Springer New York: New York, NY. p. 13-19.
11. Kiessling, F., et al., *Recent advances in molecular, multimodal and theranostic ultrasound imaging.* Adv Drug Deliv Rev, 2014. **72**: p. 15-27.
12. Lindner, J.R., *Microbubbles in medical imaging: current applications and future directions.* Nature Reviews Drug Discovery, 2004. **3**(6): p. 527-533.
13. Guvener, N., et al., *Recent advances in ultrasound-based diagnosis and therapy with micro- and nanometer-sized formulations.* Methods, 2017. **130**: p. 4-13.
14. Kaufmann, B.A. and J.R. Lindner, *Molecular imaging with targeted contrast ultrasound.* Curr Opin Biotechnol, 2007. **18**(1): p. 11-6.
15. Singh, S., A. Jain, and V. Bhateja. *A comparative evaluation of various de-speckling algorithms for medical images*. in *Proceedings of the CUBE International Information Technology Conference*. 2012. ACM.
16. Zhang, J., et al., *Speckle filtering of medical ultrasonic images using wavelet and guided filter.* Ultrasonics, 2016. **65**: p. 177-93.
17. Tourneret, J.-Y., *Detection and estimation of abrupt changes contaminated by multiplicative Gaussian noise.* Signal Processing, 1998. **68**(3): p. 259-270.

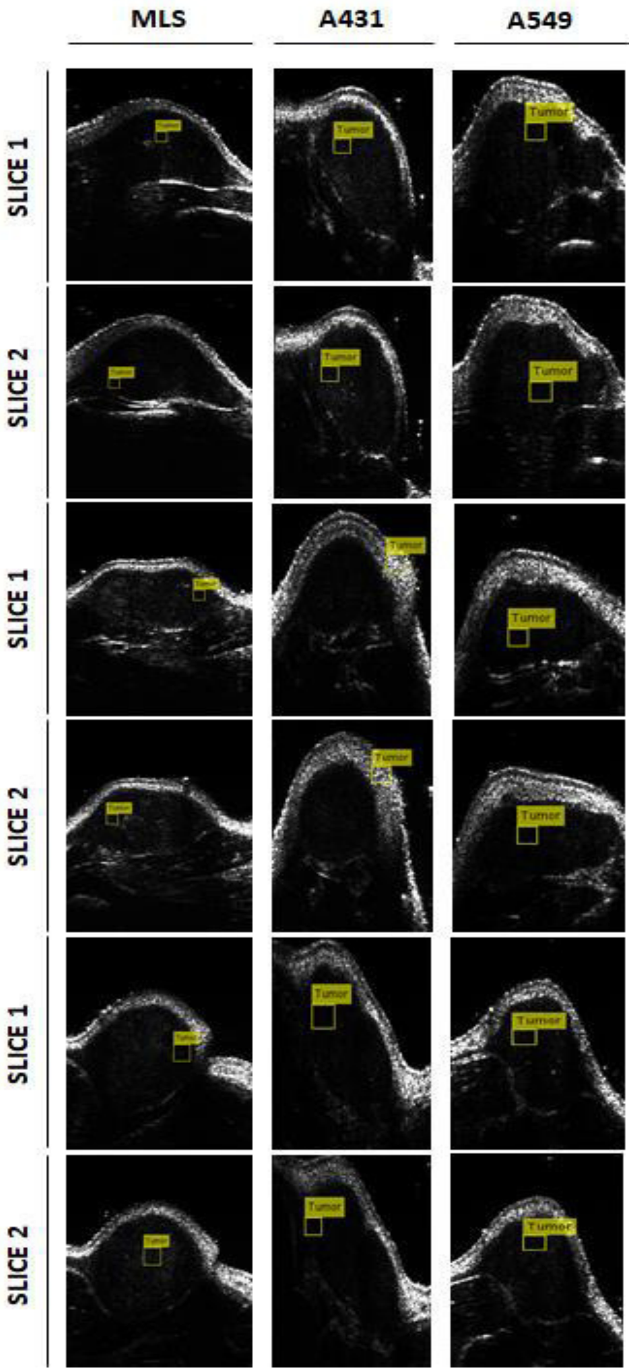
18. Lei, L., C. Wang, and X. Liu, *Discrete Wavelet Transform Decomposition level determination exploiting sparseness measurement.* International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering, 2013. **7**: p. 1182-1185.
19. Chabert, M., J.Y. Tourneret, and F. Castanie. *Non-linear time-scale contrast for abrupt change detection in multiplicative noise*. in *Ninth IEEE Signal Processing Workshop on Statistical Signal and Array Processing (Cat. No.98TH8381)*. 1998.
20. Burrus, C.S., et al., *Introduction to wavelets and wavelet transforms: a primer*. Vol. 1. 1998: Prentice hall New Jersey.
21. Sudarshan, V.K., et al., *Application of wavelet techniques for cancer diagnosis using ultrasound images: A Review.* Comput Biol Med, 2016. **69**: p. 97-111.
22. Acharya, U.R., et al., *Effect of complex wavelet transform filter on thyroid tumor classification in three-dimensional ultrasound.* Proc Inst Mech Eng H, 2013. **227**(3): p. 284-92.
23. Quellec, G., et al., *Optimal wavelet transform for the detection of microaneurysms in retina photographs.* IEEE Trans Med Imaging, 2008. **27**(9): p. 1230-41.
24. Kaur, L., S. Gupta, and R. Chauhan. *Image Denoising Using Wavelet Thresholding*. in *ICVGIP*. 2002.
25. Daubechies, I., *Orthonormal bases of compactly supported wavelets.* Communications on pure and applied mathematics, 1988. **41**(7): p. 909-996.
26. Kother Mohideen, S., P. Arumuga, and M. Sathik, *Image Denoising using Discrete Wavelet transform*. Vol. 8. 2008.
27. Gupta, S., R. Chauhan, and S. Sexana, *Wavelet-based statistical approach for speckle reduction in medical ultrasound images.* Medical and Biological Engineering and computing, 2004. **42**(2): p. 189-192.
28. Hao, X., S. Gao, and X. Gao, *A novel multiscale nonlinear thresholding method for ultrasonic speckle suppressing.* IEEE Transactions on Medical Imaging, 1999. **18**(9): p. 787-794.
29. Suresh, S., S. G R, and R. Sukanesh, *Speckle Noise Reduction in Ultrasound Images by Wavelet Thresholding based on Weighted Variance*. Vol. 1. 2009. 1793-8201.
30. Sendur, L. and I.W. Selesnick. *Multivariate shrinkage functions for wavelet-based denoising*. in *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, 2002.* 2002.
31. Gupta, S., et al. *A wavelet based statistical approach for speckle reduction in medical ultrasound images*. in *TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region*. 2003.
32. Elyasi, I. and S. Zarmehi, *Elimination noise by adaptive wavelet threshold.* World Academy of Science, Engineering and Technology, 2009. **56**: p. 462-466.
33. Chang, S.G., B. Yu, and M. Vetterli, *Adaptive wavelet thresholding for image denoising and compression.* IEEE transactions on image processing, 2000. **9**(9): p. 1532-1546.
34. Kumar, V., et al., *Radiomics: the process and the challenges.* Magn Reson Imaging, 2012. **30**(9): p. 1234-48.
35. Armato, S.G., 3rd and W.F. Sensakovic, *Automated lung segmentation for thoracic CT impact on computer-aided diagnosis.* Acad Radiol, 2004. **11**(9): p. 1011-21.

36. Sciolla, B., et al., *Improved boundary segmentation of skin lesions in high-frequency 3D ultrasound.* Computers in Biology and Medicine, 2017. **87**(Supplement C): p. 302-310.
37. Marcomini, K.D., H. Schiabel, and A.A.O. Carneiro. *Quantitative evaluation of automatic methods for lesions detection in breast ultrasound images*. in *Medical Imaging 2013: Computer-Aided Diagnosis*. 2013. International Society for Optics and Photonics.
38. Kass, M., A. Witkin, and D. Terzopoulos, *Snakes: Active contour models.* International journal of computer vision, 1988. **1**(4): p. 321-331.
39. Rios Velazquez, E., et al., *A semiautomatic CT-based ensemble segmentation of lung tumors: comparison with oncologists' delineations and with the surgical specimen.* Radiother Oncol, 2012. **105**(2): p. 167-73.
40. Ali, R., et al., *Semi-automatic segmentation of subcutaneous tumours from micro-computed tomography images.* Phys Med Biol, 2013. **58**(22): p. 8007-19.
41. Marcomini, K.D., A.A. Carneiro, and H. Schiabel, *Application of Artificial Neural Network Models in Segmentation and Classification of Nodules in Breast Ultrasound Digital Images.* International journal of biomedical imaging, 2016. **2016**.
42. Doyle, S., et al., *A Boosting Cascade for Automated Detection of Prostate Cancer from Digitized Histology*, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006: 9th International Conference, Copenhagen, Denmark, October 1-6, 2006. Proceedings, Part II*, R. Larsen, M. Nielsen, and J. Sporring, Editors. 2006, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 504-511.
43. Viola, P. and M. Jones. *Rapid object detection using a boosted cascade of simple features*. in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. 2001. IEEE.
44. Papageorgiou, C.P., M. Oren, and T. Poggio. *A general framework for object detection*. in *Computer vision, 1998. sixth international conference on*. 1998. IEEE.
45. Trefný, J. and J. Matas. *Extended set of local binary patterns for rapid object detection*. in *Computer Vision Winter Workshop*. 2010.
46. Hbali, Y., M. Sadgal, and A.E. Fazziki. *Object detection based on HOG features: Faces and dual-eyes augmented reality*. in *2013 World Congress on Computer and Information Technology (WCCIT)*. 2013.
47. Felzenszwalb, P.F., et al., *Object Detection with Discriminatively Trained Part-Based Models.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010. **32**(9): p. 1627-1645.
48. Castrillón, M., et al., *A comparison of face and facial feature detectors based on the Viola–Jones general object detection framework.* Machine Vision and Applications, 2011. **22**(3): p. 481-494.
49. Haralick, R.M. and K. Shanmugam, *Textural features for image classification.* IEEE Transactions on systems, man, and cybernetics, 1973(6): p. 610-621.
50. Galloway, M.M., *Texture analysis using grey level run lengths.* NASA STI/Recon Technical Report N, 1974. **75**.
51. Wang, Z.Z. and J.H. Yong, *Texture analysis and classification with linear regression model based on wavelet transform.* IEEE Trans Image Process, 2008. **17**(8): p. 1421-30.
52. Aerts, H.J., *Data Science in Radiology: A Path Forward.* Clin Cancer Res, 2017.

53. Opacic, T., et al., *Super-Resolution Ultrasound Bubble Tracking for Preclinical and Clinical Multiparametric Tumor Characterization.* bioRxiv, 2017: p. 203935.
54. MPHIL, V.K., et al., *Review on Speckle Noise Reduction Techniques for Medical Ultrasound Image Processing.*
55. Mincong, H., L. Weixin, and Z. Jianhui. *Image Segmentation Based on Global Extraction and Local Repair of Boundaries*. in *Intelligent Computation Technology and Automation (ICICTA), 2010 International Conference on*. 2010. IEEE.
56. Mustafa, N., et al. *Medical image De-noising schemes using wavelet transform with fixed form thresholding*. in *Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2014 11th International Computer Conference on*. 2014. IEEE.
57. Sihag, R., R. Sharma, and V. Setia. *Wavelet thresholding for image de-noising*. in *International Conference on VLSI, Communication & Instrumentation (ICVCI)*. 2011.
58. Masood, S., et al., *Prediction of Human Ethnicity from Facial Images Using Neural Networks*, in *Data Engineering and Intelligent Computing*. 2018, Springer. p. 217-226.
59. Parmar, C., et al., *Machine Learning methods for Quantitative Radiomic Biomarkers.* Sci Rep, 2015. **5**: p. 13087.
60. Yip, S.S. and H.J. Aerts, *Applications and limitations of radiomics.* Phys Med Biol, 2016. **61**(13): p. R150-66.
61. Gan, G., C. Ma, and J. Wu, *Data clustering: theory, algorithms, and applications*. 2007: SIAM.
62. Agresti, A. and B.A. Coull, *Approximate is better than "exact" for interval estimation of binomial proportions.* The American Statistician, 1998. **52**(2): p. 119-126.
63. Katz, M.H., *Multivariable analysis: a practical guide for clinicians and public health researchers*. 2011: Cambridge university press.
64. Tang, X., *Texture information in run-length matrices.* IEEE transactions on image processing, 1998. **7**(11): p. 1602-1609.
65. Soh, L.-K. and C. Tsatsoulis, *Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices.* IEEE Transactions on geoscience and remote sensing, 1999. **37**(2): p. 780-795.
66. Clausi, D.A., *An analysis of co-occurrence texture statistics as a function of grey level quantization.* Canadian Journal of remote sensing, 2002. **28**(1): p. 45-62.

7. Supplementary Information

Supplementary Figure 1: The Initial data set of tumor images. The presented figure encloses 18 images of 3 different tumor models (MLS, A431 and A549). Each tumor was measured at 2 different positions (Slice 1 and 2). The presented images enclose the results of tumor detection with developed CCs.



Supplementary Figure 2: Intensity-based features: list of extracted attributes with assigned final rank and source of origin.

Feature Name	Rank	Group	Source
Energy	144	FOS	[2]
Entropy	108	FOS	[2]
Kurtosis	140	FOS	[2]
Max	109	FOS	[2]
Mean	102	FOS	[2]
Mean Absolut Deviation	106	FOS	[2]
Median	101	FOS	[2]
Minimum	156	FOS	[2]
Range	115	FOS	[2]
Root Mean Square	103	FOS	[2]
Skewness	160	FOS	[2]
Standard Deviation	113	FOS	[2]
Uniformity	73	FOS	[2]
Variance	117	FOS	[2]

Supplementary Figure 3: Textural features: list of extracted attributes with assigned final rank and source of origin.

Feature Name	Rank	Group	Source
Auto Correlation	115	SOS	[65]
Cluster Prominence	159	SOS	[65]
Cluster Shade	149	SOS	[65]
Cluster Tendency	164	SOS	[2]
Contrast	124	SOS	[49]
Correlation	95	SOS	[49]
Difference Entropy	134	SOS	[49]
Dissimilarity	127	SOS	[65]
Energy	100	SOS	[49]
Entropy	90	SOS	[49]
Homogeneity 1	107	SOS	[66]
Homogeneity 2	108	SOS	[65]
Informational Measure of Correlation 1	48	SOS	[49]
Informational Measure of Correlation 2	99	SOS	[49]
Inverse Difference Moment Normalized	126	SOS	[49]
Inverse Difference Normalized	114	SOS	[49]
Maximum Probability	119	SOS	[65]
Sum Average	106	SOS	[49]
Sum Entropy	102	SOS	[49]
Sum Variance	119	SOS	[49]
Variance	123	SOS	[49]

Feature Name	Rank	Group	Source
Short Run Emphasis	12	SOS	[50]
Long Run Emphasis	16	SOS	[50]
Gray Level Non-Uniformity	72	SOS	[50]
Run Length Non-Uniformity	70	SOS	[50]
Run Percentage	58	SOS	[50]
Low Gray Level Run Emphasis	157	SOS	[64]
High Gray Level Run Emphasis	129	SOS	[64]
Short Run Low Gray Level Emphasis	142	SOS	[64]
Short Run High Gray Level Emphasis	117	SOS	[64]
Long Run Low Gray Level Emphasis	149	SOS	[64]
Long Run High Gray Level Emphasis	159	SOS	[64]

Supplementary Figure 4: Wavelet-based features: list of extracted attributes with assigned final rank and source of origin.

Feature Name	Rank	Group	Source
Energy FOS LL	158	WT	[2]
Entropy FOS LL	105	WT	[2]
Kurtosis FOS LL	113	WT	[2]
Max FOS LL	115	WT	[2]
Mean FOS LL	111	WT	[2]
Mean Absolut Deviation FOS LL	116	WT	[2]
Median FOS LL	106	WT	[2]
Min FOS LL	151	WT	[2]
Range FOS LL	115	WT	[2]
Root Mean Square FOS LL	108	WT	[2]
Skewness FOS LL	119	WT	[2]
Standard Deviation FOS LL	111	WT	[2]
Uniformity FOS LL	26	WT	[2]
Variance FOS LL	118	WT	[2]
Energy FOS LH	138	WT	[2]
Entropy FOS LH	63	WT	[2]
Kurtosis FOS LH	161	WT	[2]
Max FOS LH	135	WT	[2]
Mean FOS LH	65	WT	[2]
Mean Absolut Deviation FOS LH	64	WT	[2]
Median FOS LH	68	WT	[2]
Min FOS LH	151	WT	[2]
Range FOS LH	135	WT	[2]
Root Mean Square FOS LH	112	WT	[2]
Skewness FOS LH	157	WT	[2]

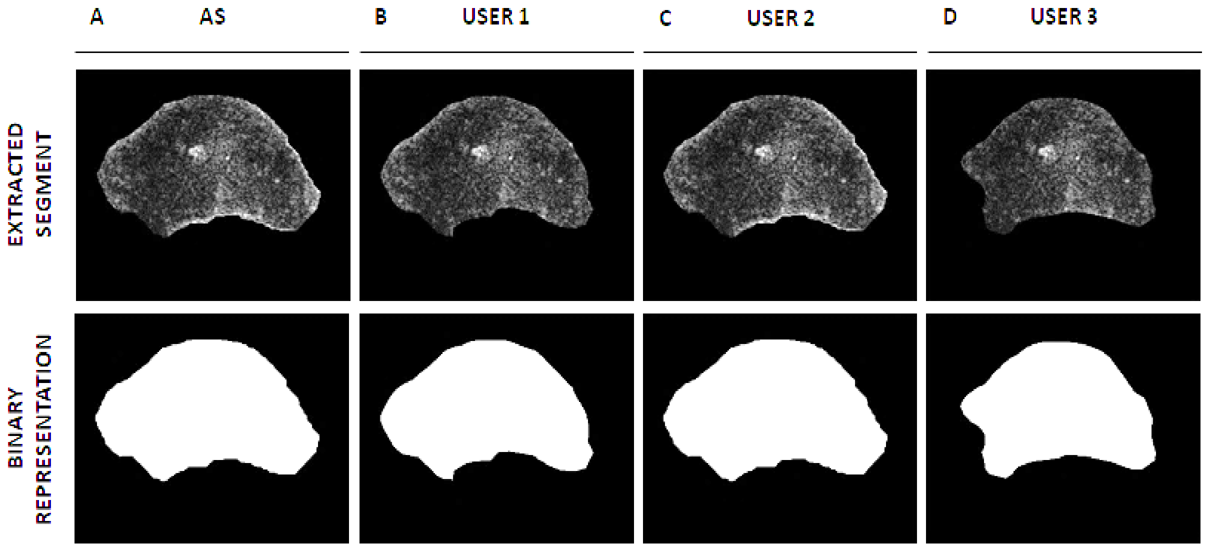
Feature Name	Rank	Group	Source
Standard Deviation FOS LH	69	WT	[2]
Uniformity FOS LH	34	WT	[2]
Variance FOS LH	73	WT	[2]
Energy FOS HL	162	WT	[2]
Entropy FOS HL	115	WT	[2]
Kurtosis FOS HL	200	WT	[2]
Max FOS HL	153	WT	[2]
Mean FOS HL	117	WT	[2]
Mean Absolut Deviation FOS HL	129	WT	[2]
Median FOS HL	131	WT	[2]
Min FOS HL	151	WT	[2]
Range FOS HL	152	WT	[2]
Root Mean Square FOS HL	124	WT	[2]
Skewness FOS HL	198	WT	[2]
Standard Deviation FOS HL	140	WT	[2]
Uniformity FOS HL	30	WT	[2]
Variance FOS HL	143	WT	[2]
Energy FOS HH	178	WT	[2]
Entropy FOS HH	138	WT	[2]
Kurtosis FOS HH	150	WT	[2]
Max FOS HH	152	WT	[2]
Mean FOS HH	129	WT	[2]
Mean Absolut Deviation FOS HH	137	WT	[2]
Median FOS HH	132	WT	[2]
Min FOS HH	151	WT	[2]
Range FOS HH	152	WT	[2]
Root Mean Square FOS HH	105	WT	[2]
Skewness FOS HH	143	WT	[2]
Standard Deviation FOS HH	137	WT	[2]
Uniformity FOS HH	30	WT	[2]
Variance FOS HH	157	WT	[2]
Auto Correlation LL	119	WT	[65]
Cluster Prominence LL	161	WT	[65]
Cluster Shade LL	151	WT	[65]
Cluster Tendency LL	154	WT	[2]
Contrast LL	124	WT	[49]
Correlation LL	90	WT	[49]
Difference Entropy LL	118	WT	[49]
Dissimilarity LL	115	WT	[65]
Energy LL	91	WT	[49]

Feature Name	Rank	Group	Source
Entropy LL	80	WT	[49]
Homogeneity 1 LL	109	WT	[66]
Homogeneity 2 LL	99	WT	[65]
Informational Measure of Correlation 1 LL	47	WT	[49]
Informational Measure of Correlation 2 LL	106	WT	[49]
Inverse Difference Moment Normalized LL	123	WT	[49]
Inverse Difference Normalized LL	104	WT	[49]
Maximum Probability LL	137	WT	[65]
Sum Average LL	108	WT	[49]
Sum Entropy LL	92	WT	[49]
Sum Variance LL	117	WT	[49]
Variance LL	120	WT	[49]
Auto Correlation HH	134	WT	[65]
Cluster Prominence HH	159	WT	[65]
Cluster Shade HH	154	WT	[65]
Cluster Tendency HH	193	WT	[2]
Contrast HH	150	WT	[49]
Correlation HH	177	WT	[49]
Difference Entropy HH	132	WT	[49]
Dissimilarity HH	124	WT	[65]
Energy HH	140	WT	[49]
Entropy HH	144	WT	[49]
Homogeneity 1 HH	130	WT	[66]
Homogeneity 2 HH	125	WT	[65]
Informational Measure of Correlation 1 HH	78	WT	[49]
Informational Measure of Correlation 2 HH	110	WT	[49]
Inverse Difference Moment Normalized HH	167	WT	[49]
Inverse Difference Normalized HH	169	WT	[49]
Maximum Probability HH	146	WT	[65]
Sum Average HH	122	WT	[49]
Sum Entropy HH	113	WT	[49]
Sum Variance HH	145	WT	[49]
Variance HH	143	WT	[49]
Auto Correlation LH	65	WT	[65]
Cluster Prominence LH	144	WT	[65]
Cluster Shade LH	132	WT	[65]
Cluster Tendency LH	173	WT	[2]
Contrast LH	61	WT	[49]
Correlation LH	148	WT	[49]
Difference Entropy LH	76	WT	[49]

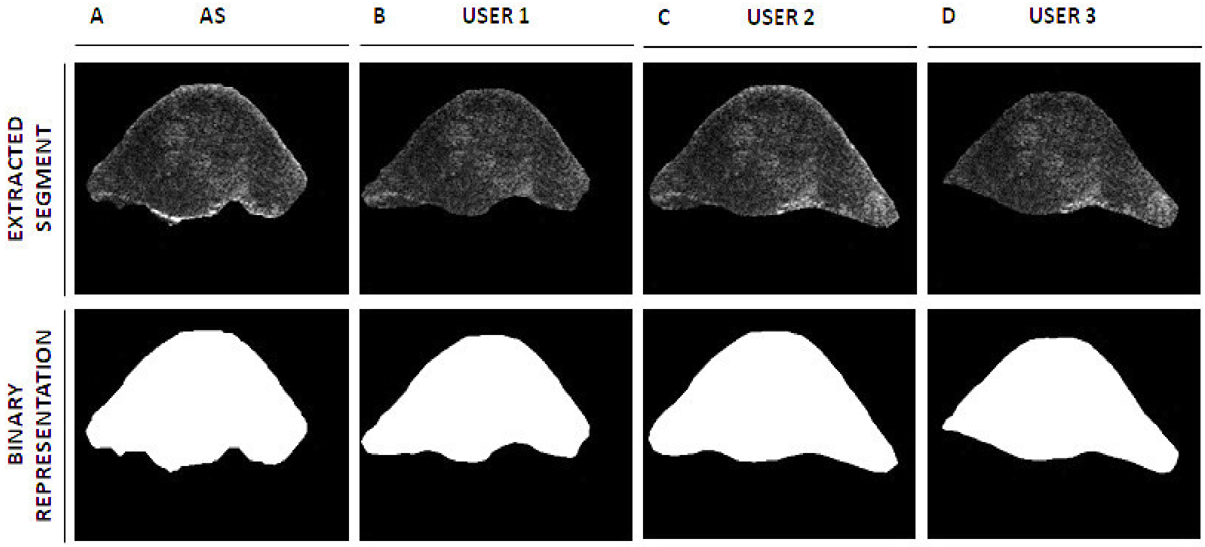
Feature Name	Rank	Group	Source
Dissimilarity LH	53	WT	[65]
Energy LH	129	WT	[49]
Entropy LH	129	WT	[49]
Homogeneity 1 LH	75	WT	[66]
Homogeneity 2 LH	76	WT	[65]
Informational Measure of Correlation 1 LH	44	WT	[49]
Informational Measure of Correlation 2 LH	89	WT	[49]
Inverse Difference Moment Normalized LH	158	WT	[49]
Inverse Difference Normalized LH	159	WT	[49]
Maximum Probability LH	123	WT	[65]
Sum Average LH	62	WT	[49]
Sum Entropy LH	83	WT	[49]
Sum Variance LH	89	WT	[49]
Variance LH	71	WT	[49]
Auto Correlation HL	152	WT	[65]
Cluster Prominence HL	163	WT	[65]
Cluster Shade HL	147	WT	[65]
Cluster Tendency HL	150	WT	[2]
Contrast HL	126	WT	[49]
Correlation HL	121	WT	[49]
Difference Entropy HL	118	WT	[49]
Dissimilarity HL	120	WT	[65]
Energy HL	124	WT	[49]
Entropy HL	99	WT	[49]
Homogeneity 1 HL	131	WT	[66]
Homogeneity 2 HL	127	WT	[65]
Informational Measure of Correlation 1 HL	72	WT	[49]
Informational Measure of Correlation 2 HL	120	WT	[49]
Inverse Difference Moment Normalized HL	155	WT	[49]
Inverse Difference Normalized HL	134	WT	[49]
Maximum Probability HL	103	WT	[65]
Sum Average HL	125	WT	[49]
Sum Entropy HL	121	WT	[49]
Sum Variance HL	154	WT	[49]
Variance HL	113	WT	[49]
Short Run Emphasis LL	6	WT	[50]
Long Run Emphasis LL	14	WT	[50]
Gray Level Non-Uniformity LL	71	WT	[50]
Run Length Non-Uniformity LL	70	WT	[50]
Run Percentage LL	59	WT	[50]

Feature Name	Rank	Group	Source
Low Gray Level Run Emphasis LL	127	WT	[64]
High Gray Level Run Emphasis LL	124	WT	[64]
Short Run Low Gray Level Emphasis LL	134	WT	[64]
Short Run High Gray Level Emphasis LL	110	WT	[64]
Long Run Low Gray Level Emphasis LL	144	WT	[64]
Long Run High Gray Level Emphasis LL	150	WT	[64]
Short Run Emphasis LH	9	WT	[50]
Long Run Emphasis LH	15	WT	[50]
Gray Level Non-Uniformity LH	65	WT	[50]
Run Length Non-Uniformity LH	75	WT	[50]
Run Percentage LH	62	WT	[50]
Low Gray Level Run Emphasis LH	146	WT	[64]
High Gray Level Run Emphasis LH	146	WT	[64]
Short Run Low Gray Level Emphasis LH	150	WT	[64]
Short Run High Gray Level Emphasis LH	133	WT	[64]
Long Run Low Gray Level Emphasis LH	130	WT	[64]
Long Run High Gray Level Emphasis LH	165	WT	[64]
Short Run Emphasis HL	11	WT	[50]
Long Run Emphasis HL	14	WT	[50]
Gray Level Non-Uniformity HL	77	WT	[50]
Run Length Non-Uniformity HL	63	WT	[50]
Run Percentage HL	56	WT	[50]
Low Gray Level Run Emphasis HL	148	WT	[64]
High Gray Level Run Emphasis HL	172	WT	[64]
Short Run Low Gray Level Emphasis HL	130	WT	[64]
Short Run High Gray Level Emphasis HL	145	WT	[64]
Long Run Low Gray Level Emphasis HL	154	WT	[64]
Long Run High Gray Level Emphasis HL	139	WT	[64]
Short Run Emphasis HH	7	WT	[50]
Long Run Emphasis HH	14	WT	[50]
Gray Level Non-Uniformity HH	79	WT	[50]
Run Length Non-Uniformity HH	71	WT	[50]
Run Percentage HH	57	WT	[50]
Low Gray Level Run Emphasis HH	166	WT	[64]
High Gray Level Run Emphasis HH	157	WT	[64]
Short Run Low Gray Level Emphasis HH	146	WT	[64]
Short Run High Gray Level Emphasis HH	131	WT	[64]
Long Run Low Gray Level Emphasis HH	188	WT	[64]
Long Run High Gray Level Emphasis HH	166	WT	[64]

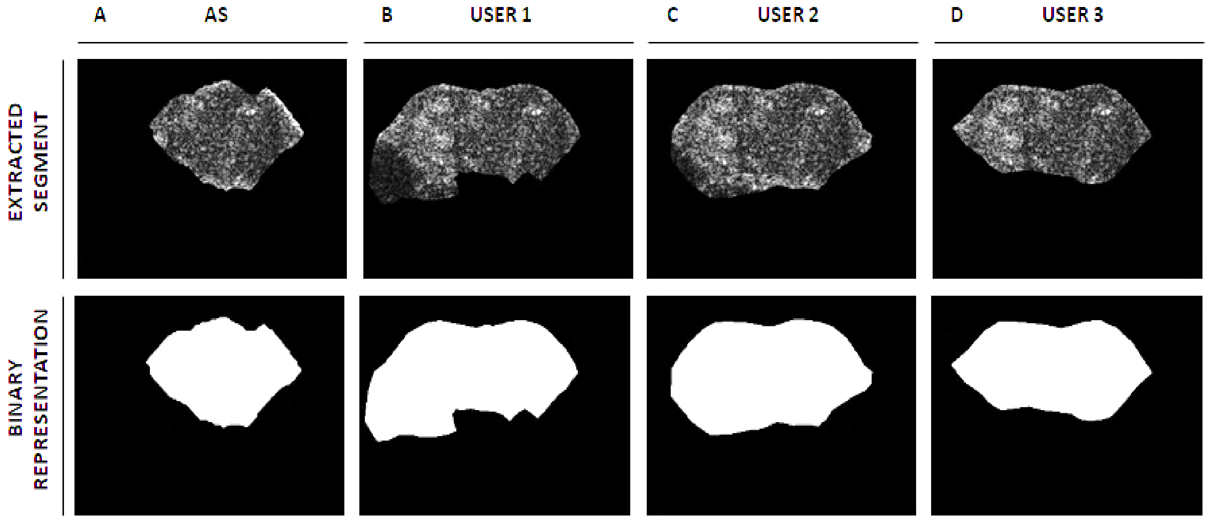
Supplementary Figure 5: The comparison of extracted MLS tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



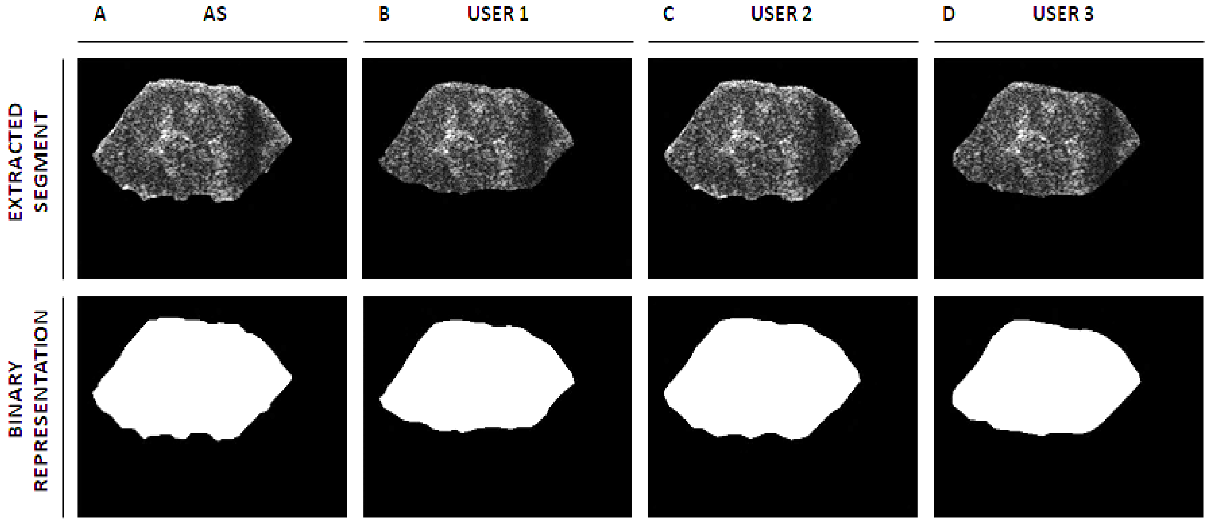
Supplementary Figure 6: The comparison of extracted MLS tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



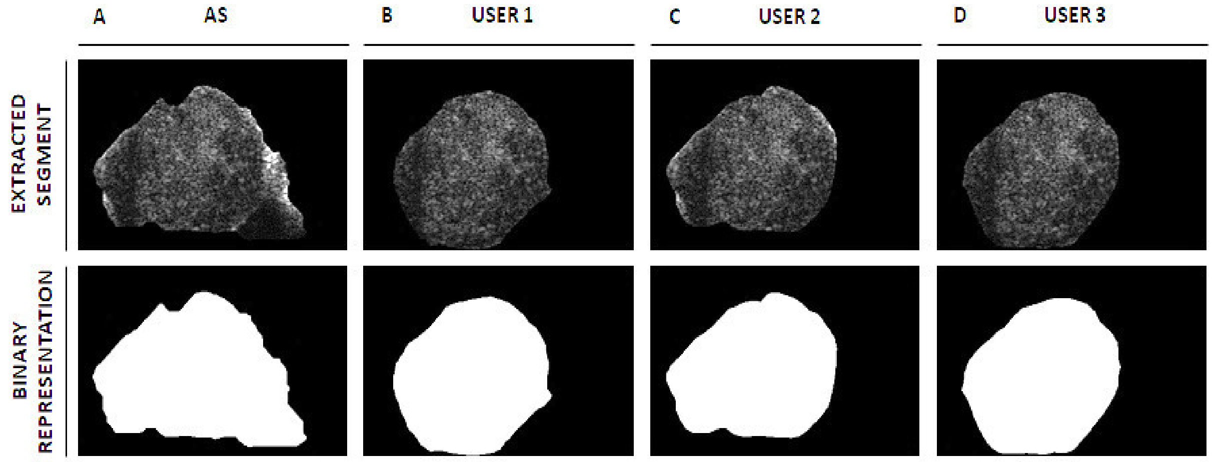
Supplementary Figure 7: The comparison of extracted MLS tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that the segment extracted with the use of AS (A) has smaller area in comparison to segments extracted by the users (B, C and D).



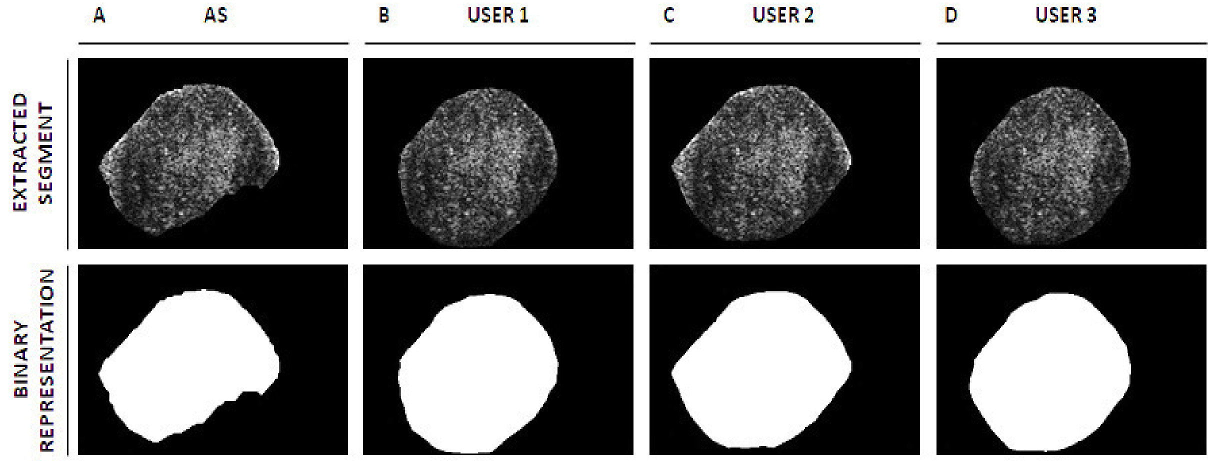
Supplementary Figure 8: The comparison of extracted MLS tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



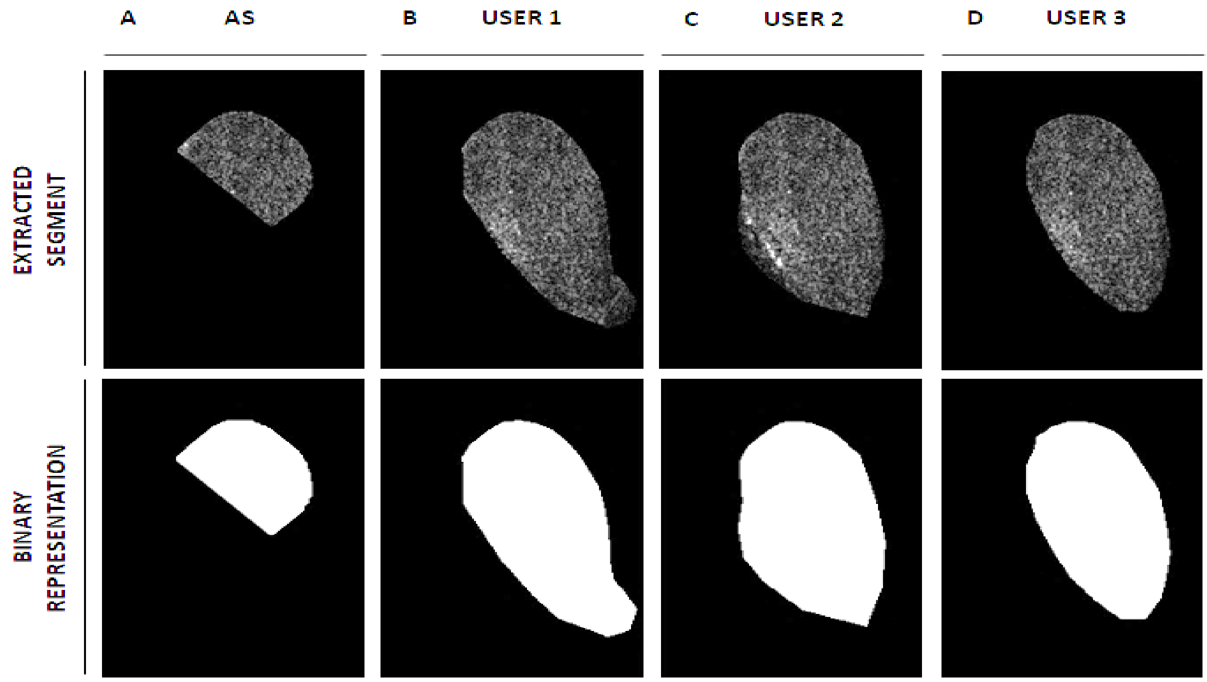
Supplementary Figure 9: The comparison of extracted MLS tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that the segment extracted with the use of AS (A) has bigger area in comparison to segments extracted by the users (B, C and D).



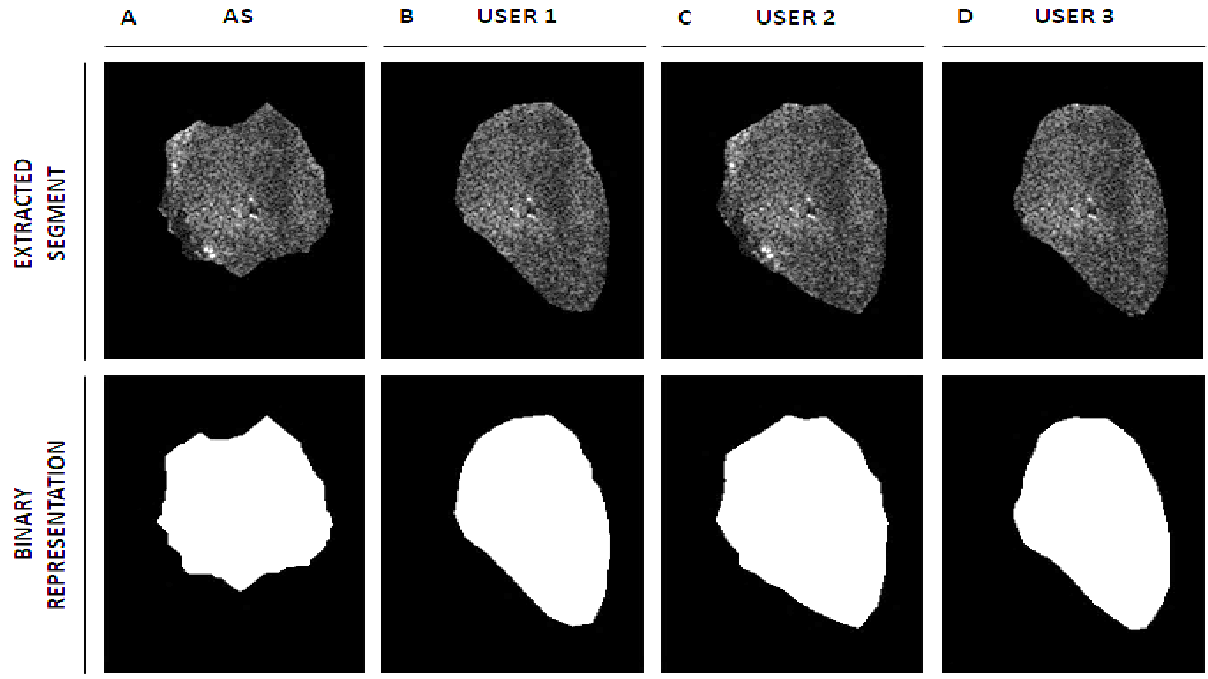
Supplementary Figure 10: The comparison of extracted MLS tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



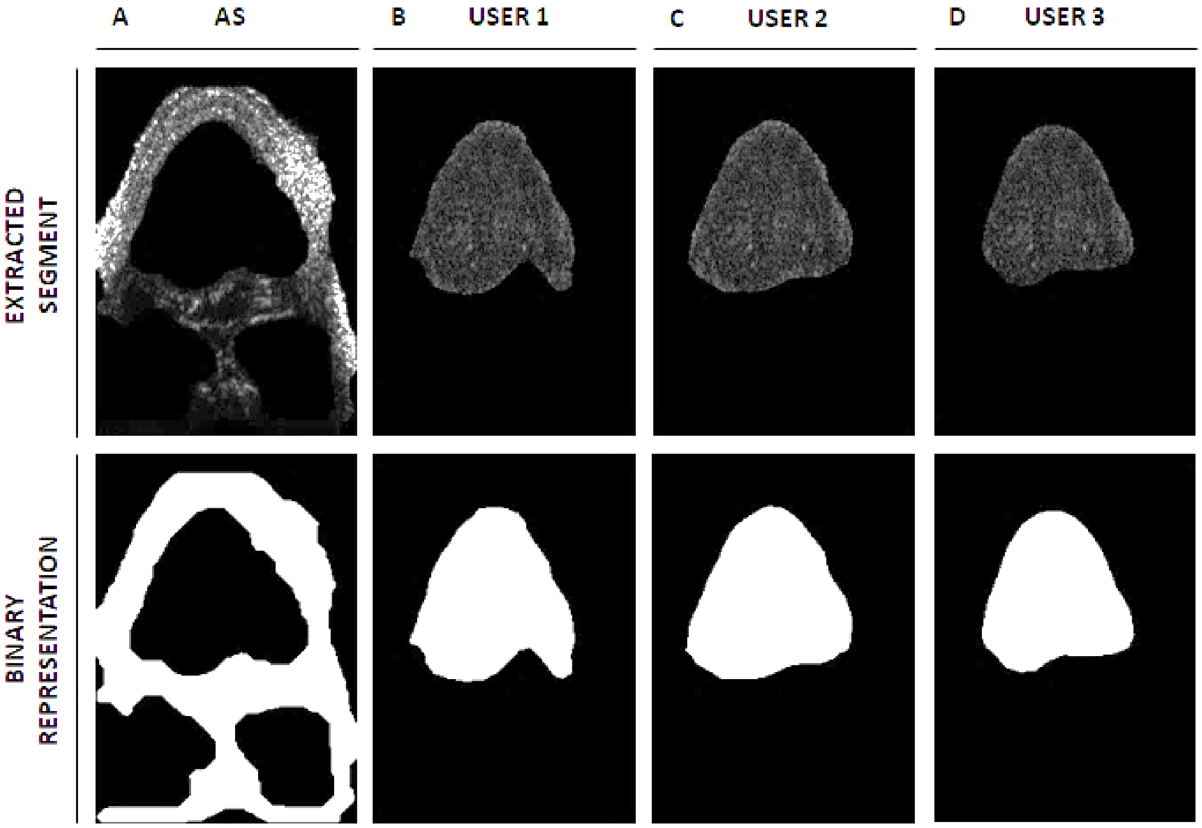
Supplementary Figure 11: The comparison of extracted A431 tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that the segment extracted with the use of AS (A) has smaller area in comparison to segments extracted by the users (B, C and D).



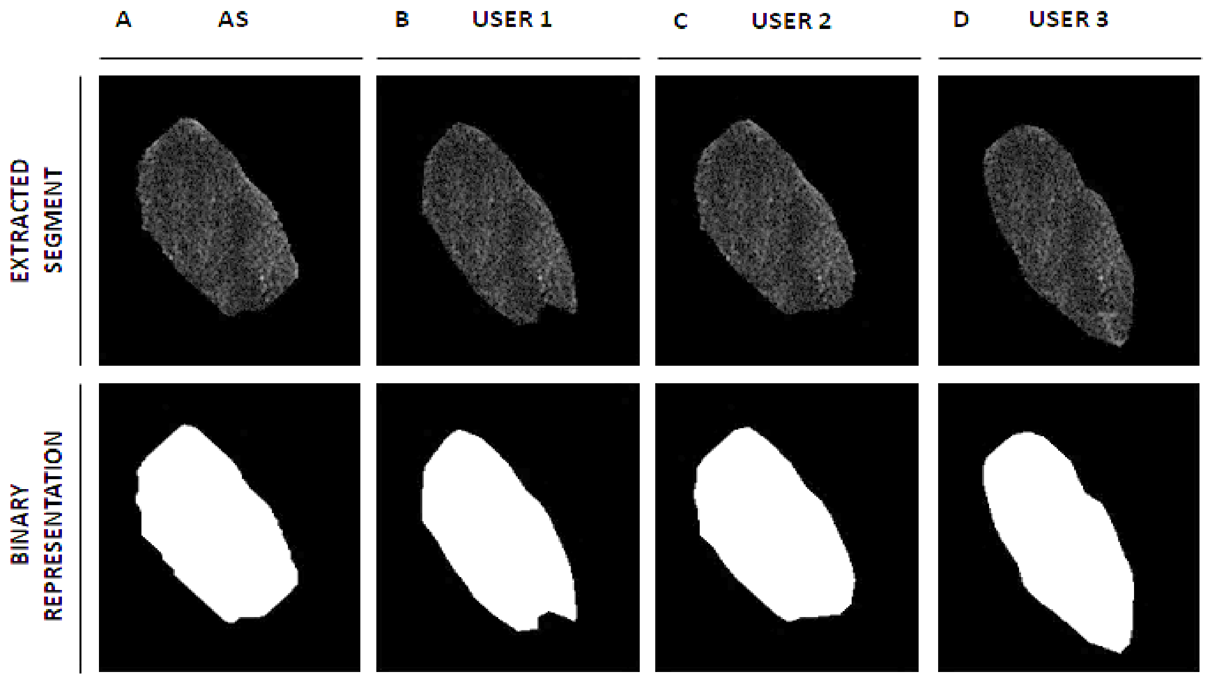
Supplementary Figure 12: The comparison of extracted A431 tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



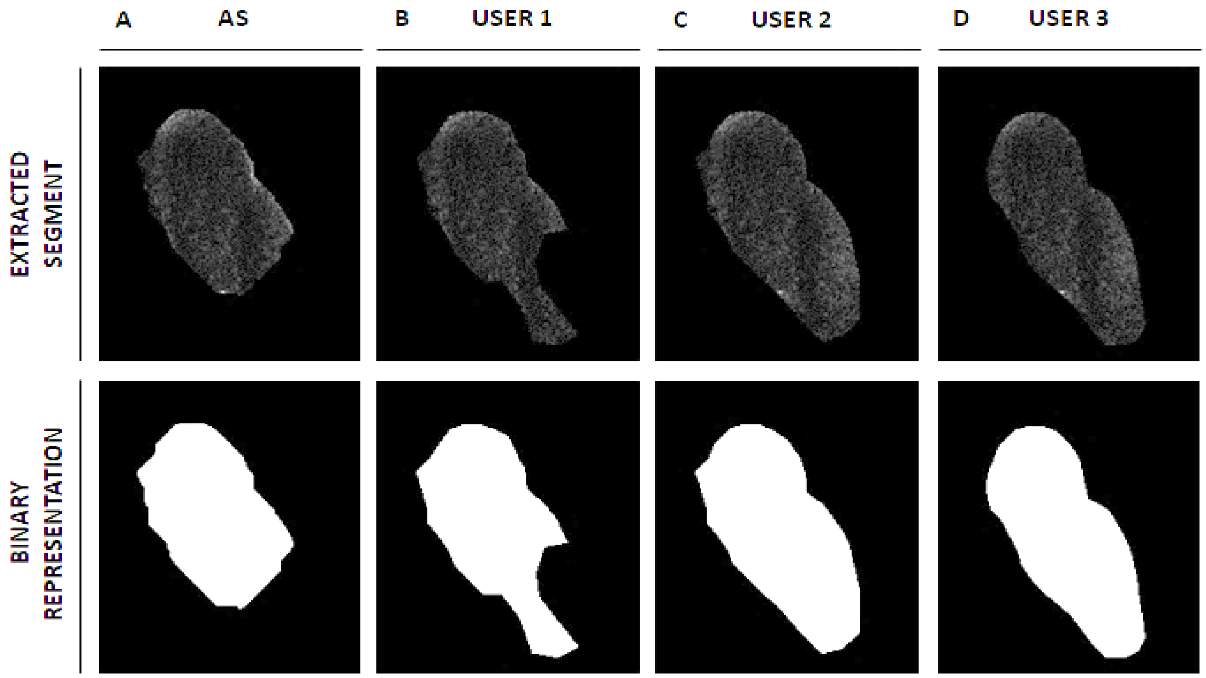
Supplementary Figure 13: The comparison of extracted A431 tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. The incorrect tumor detection resulted in segmentation of the lesion boundary (A). One can see that segment extracted with the use of AS (A) is much different in comparison to segments extracted by the users (B, C and D).



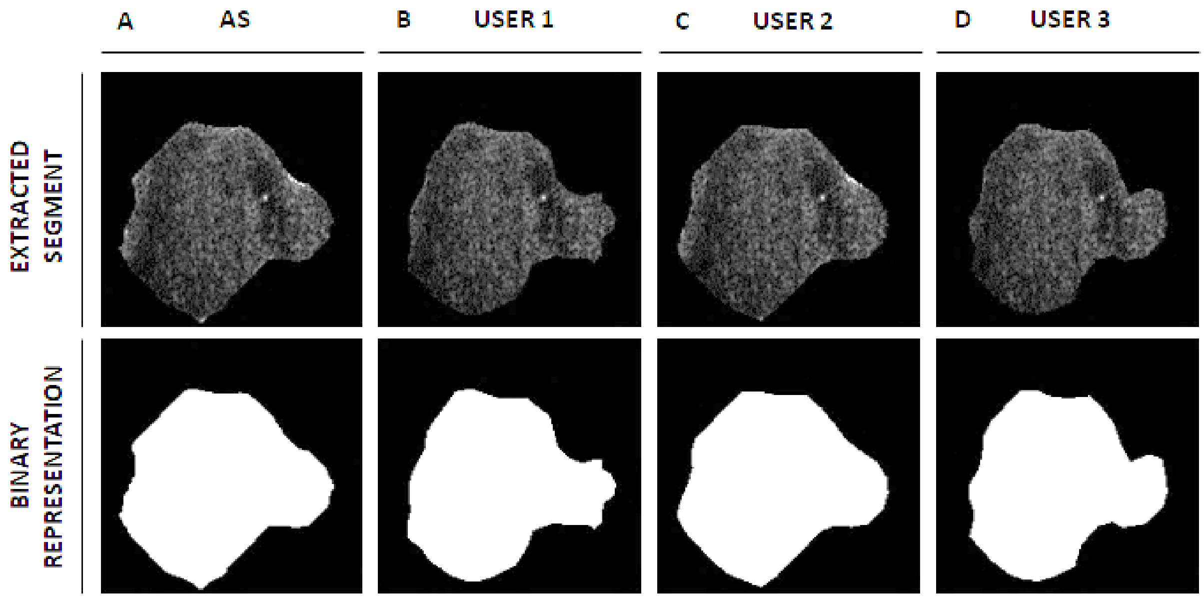
Supplementary Figure 14: The comparison of extracted A431 tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



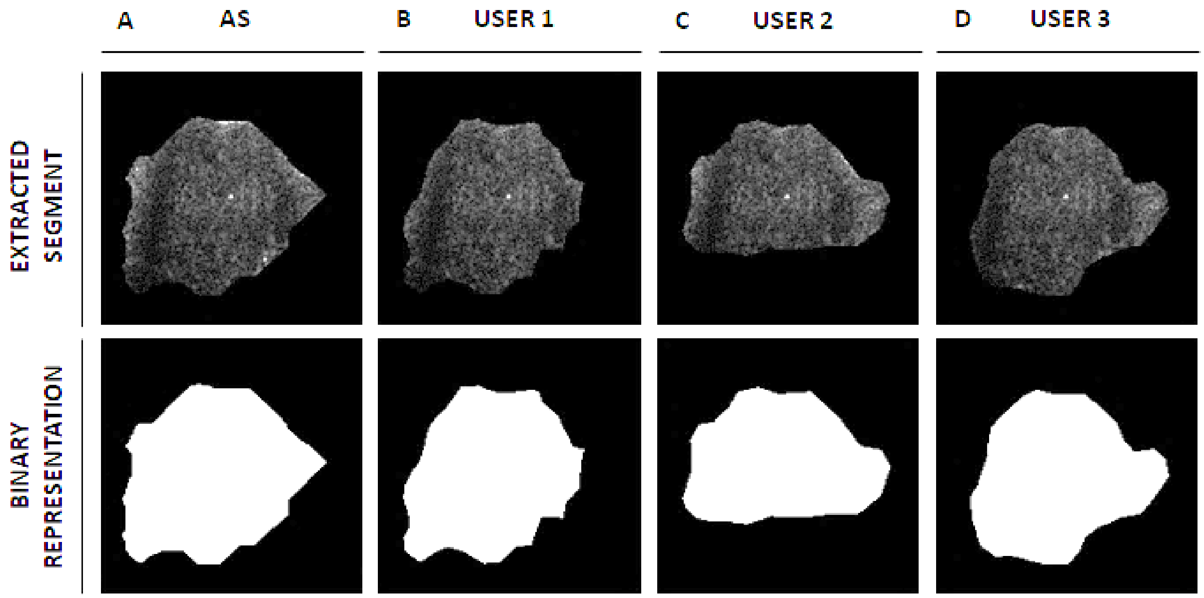
Supplementary Figure 15: The comparison of extracted A431 tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



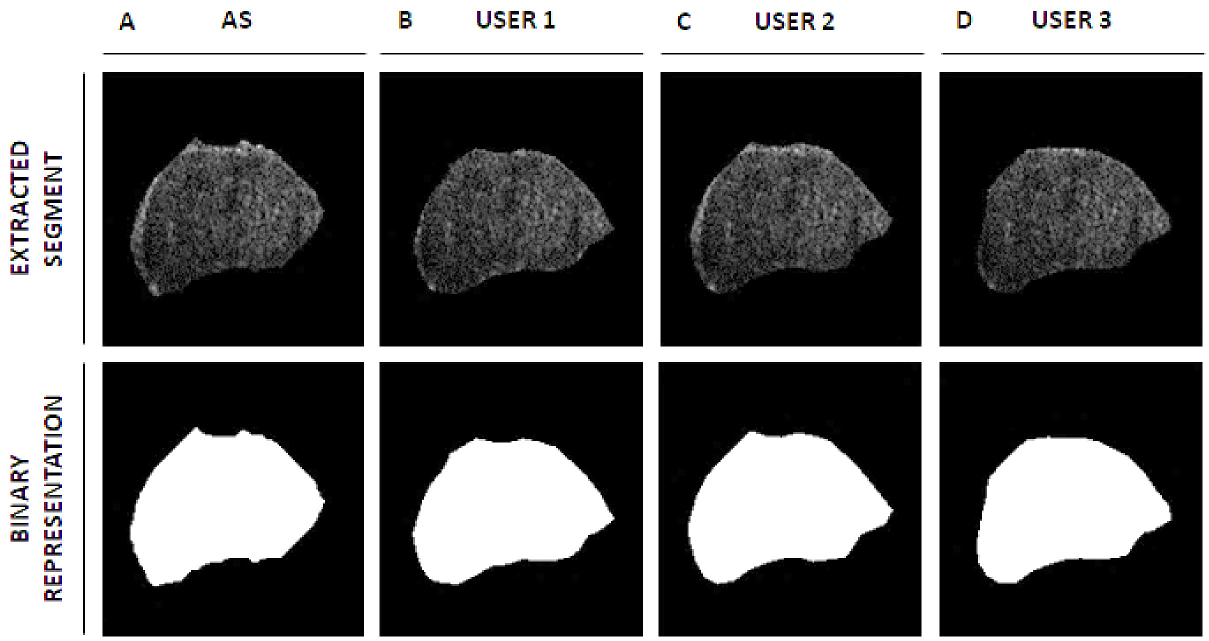
Supplementary Figure 16: The comparison of extracted A549 tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



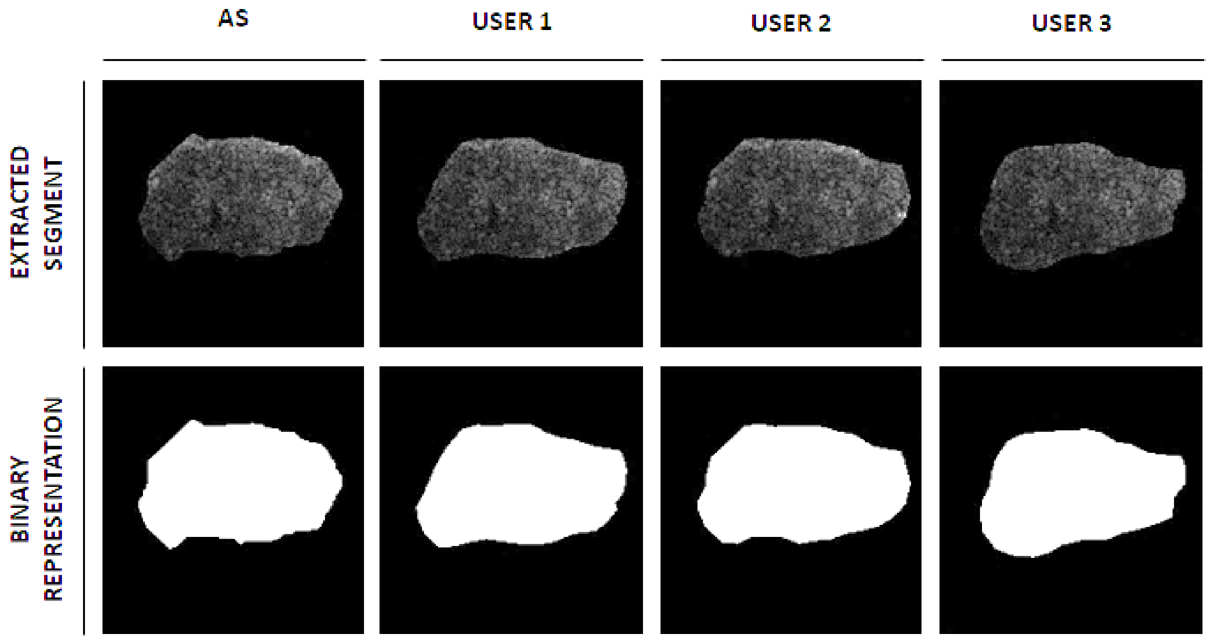
Supplementary Figure 17: The comparison of extracted A549 tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



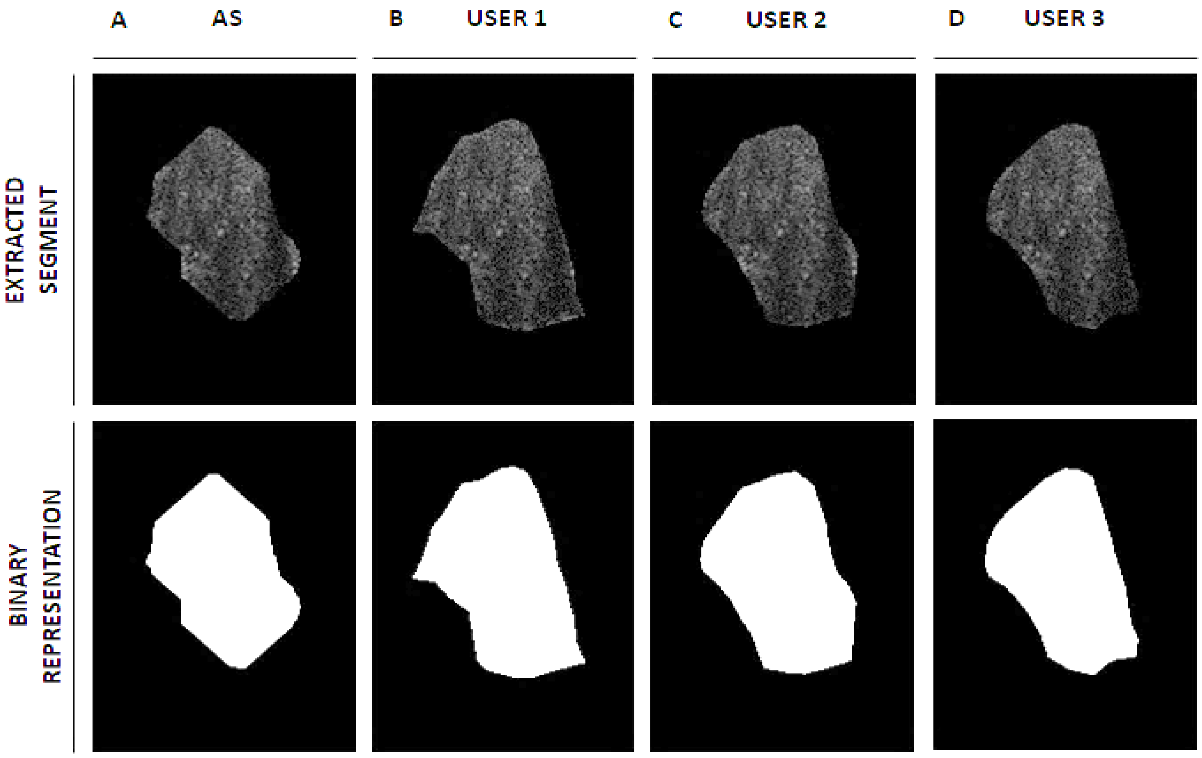
Supplementary Figure 18: The comparison of extracted A549 tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



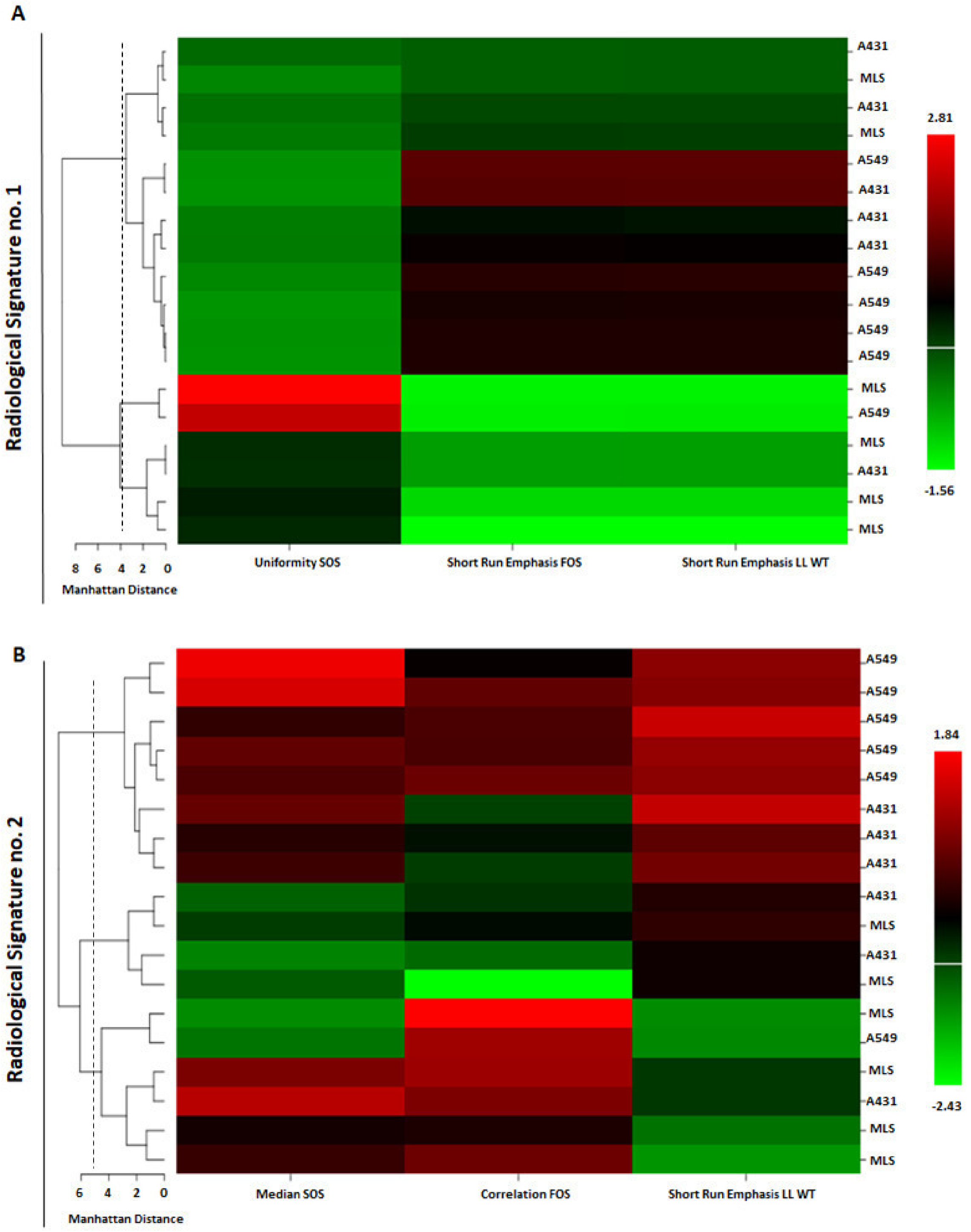
Supplementary Figure 19: The comparison of extracted A549 tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



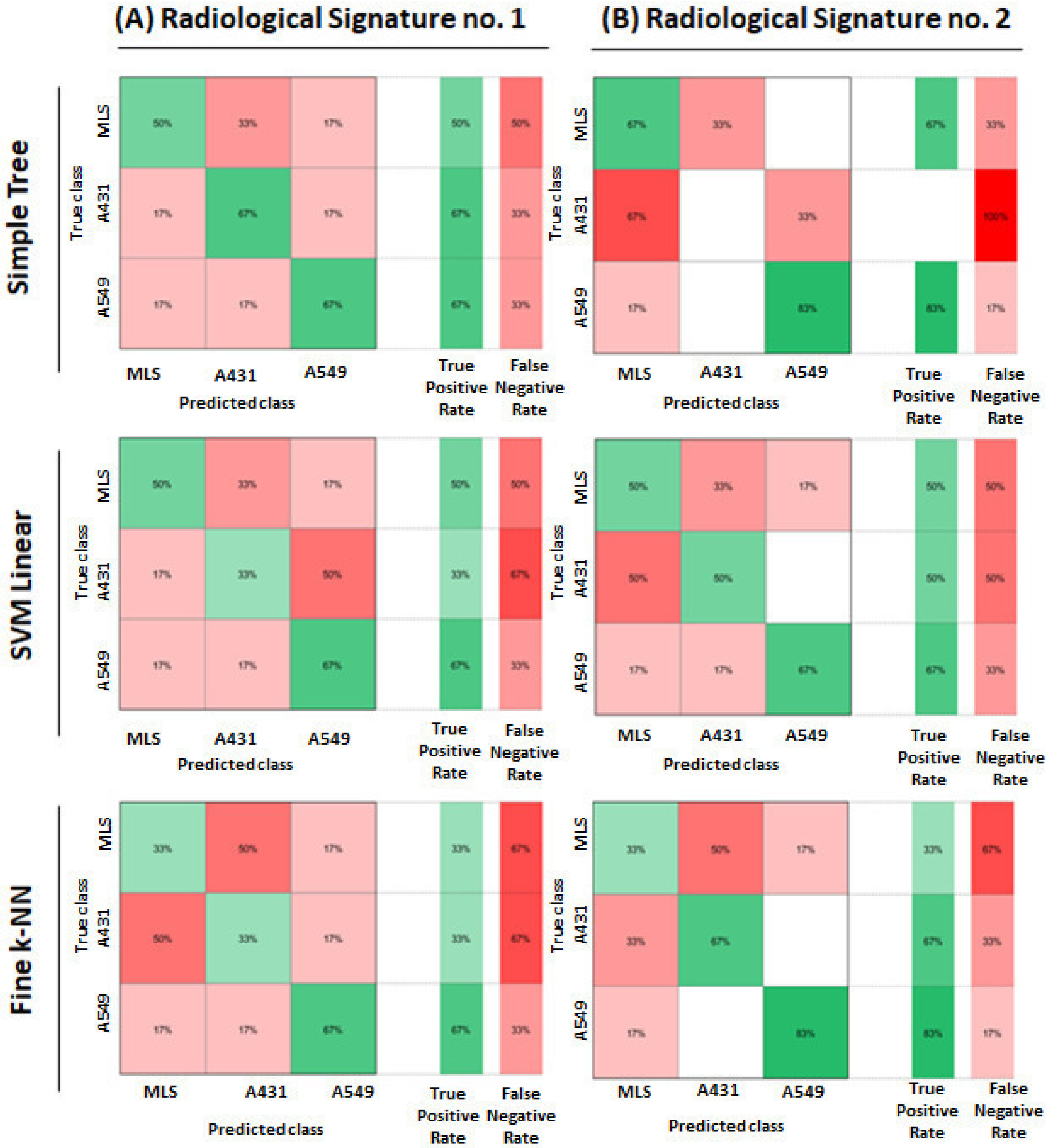
Supplementary Figure 20: The comparison of extracted A549 tumor model segments. The presented figure comprise segments extracted by the AS algorithm (A), User 1 (B), User 2 (C) and User 3 (D), and their binary representations as well. One can see that all enclosed segments are very similar.



Supplementary Figure 21: Hierarchical agglomerative clustering: unsupervised validation of tumor classification models. In both (A) and (B) clustering of features is presented on the left side with dendrogram and assigned tumor types on the right side of the computed heat map. The linkages above the drawn dotted line constitutes the 3 final clusters. The results presented in this figure refers to the validation of RS no.1 and RS no.2 developed with reference to the segments delineated by User 2.



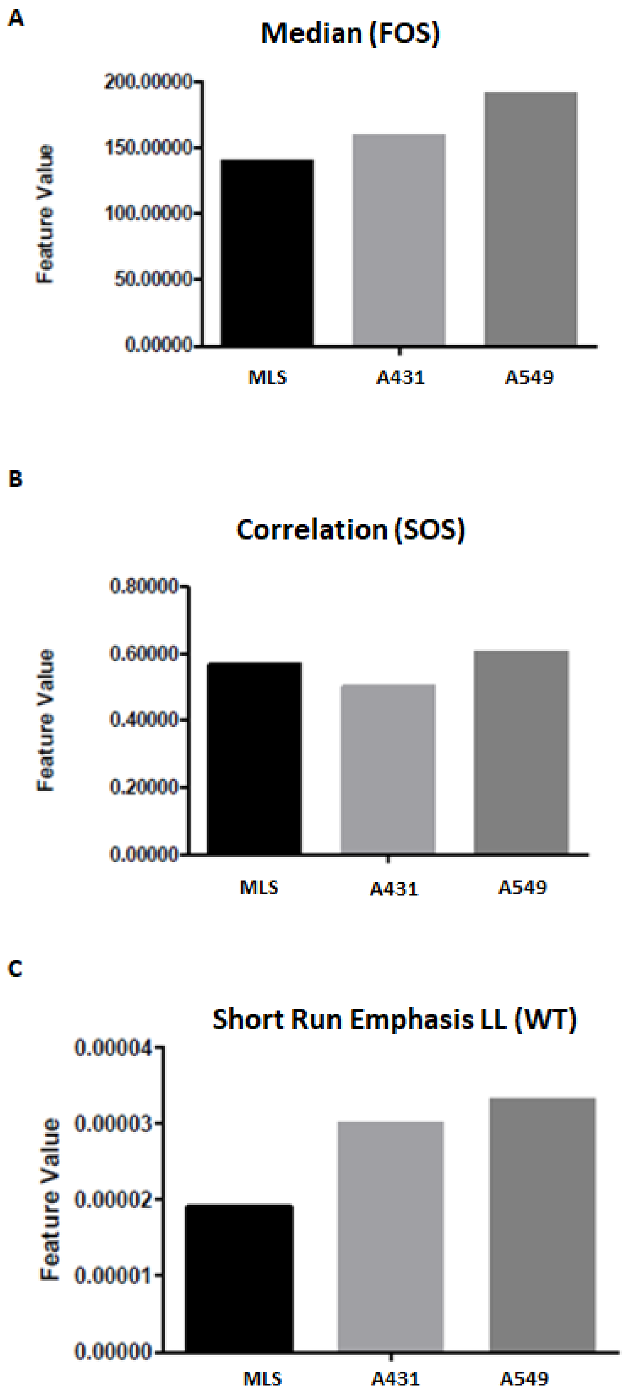
Supplementary Figure 22: The comparison of supervised classification algorithms accuracy in training discriminative tumor models. (A) Presented the results obtained for training RS no.1 with Simple Tree, SVM Linear and Fine k-NN algorithms respectively. (B) Presents the results obtained for training RS no.2 with Simple Tree, SVM Linear and Fine k-NN algorithms respectively. The green color assigns true positives classification. The red color assigns false negative classification. The results presented in this figure refers to the validation of the RS no.1 and RS no.2 developed with reference to the segments delineated by User 2.



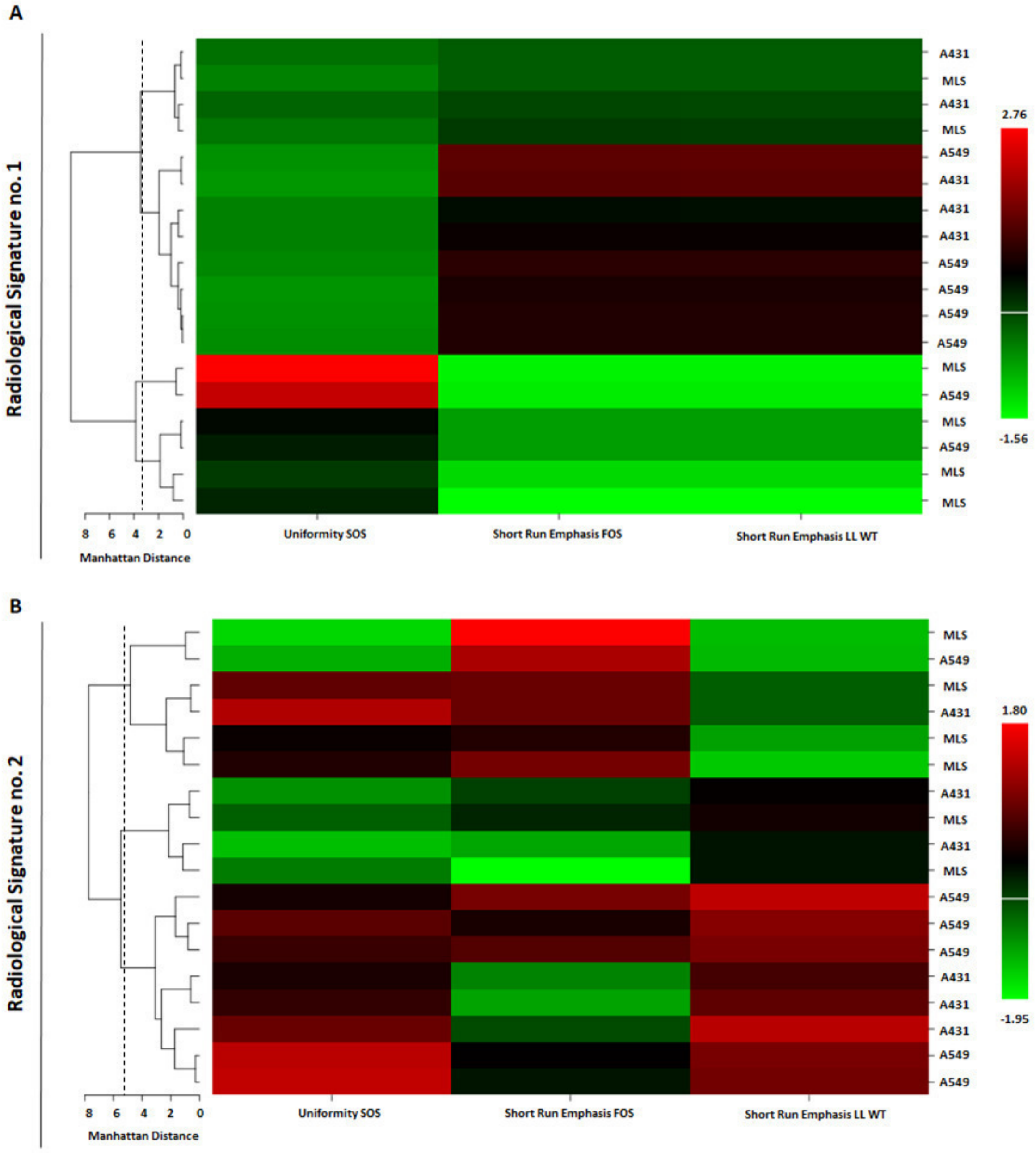
Supplementary Figure 23: Wilson score interval for estimation of tumor model classification probability (User 2).

	RS no.1		**RS no.2**	
	p	**95% CI**	**p**	**95% CI**
Simple Tree	0.611	0.386 - 0.797	0.500	0.290 - 0.710
Linear SVM	0.500	0.290 - 0.710	0.556	0.337 - 0.754
k-NN	0.444	0.246 - 0.663	0.611	0.386 - 0.797

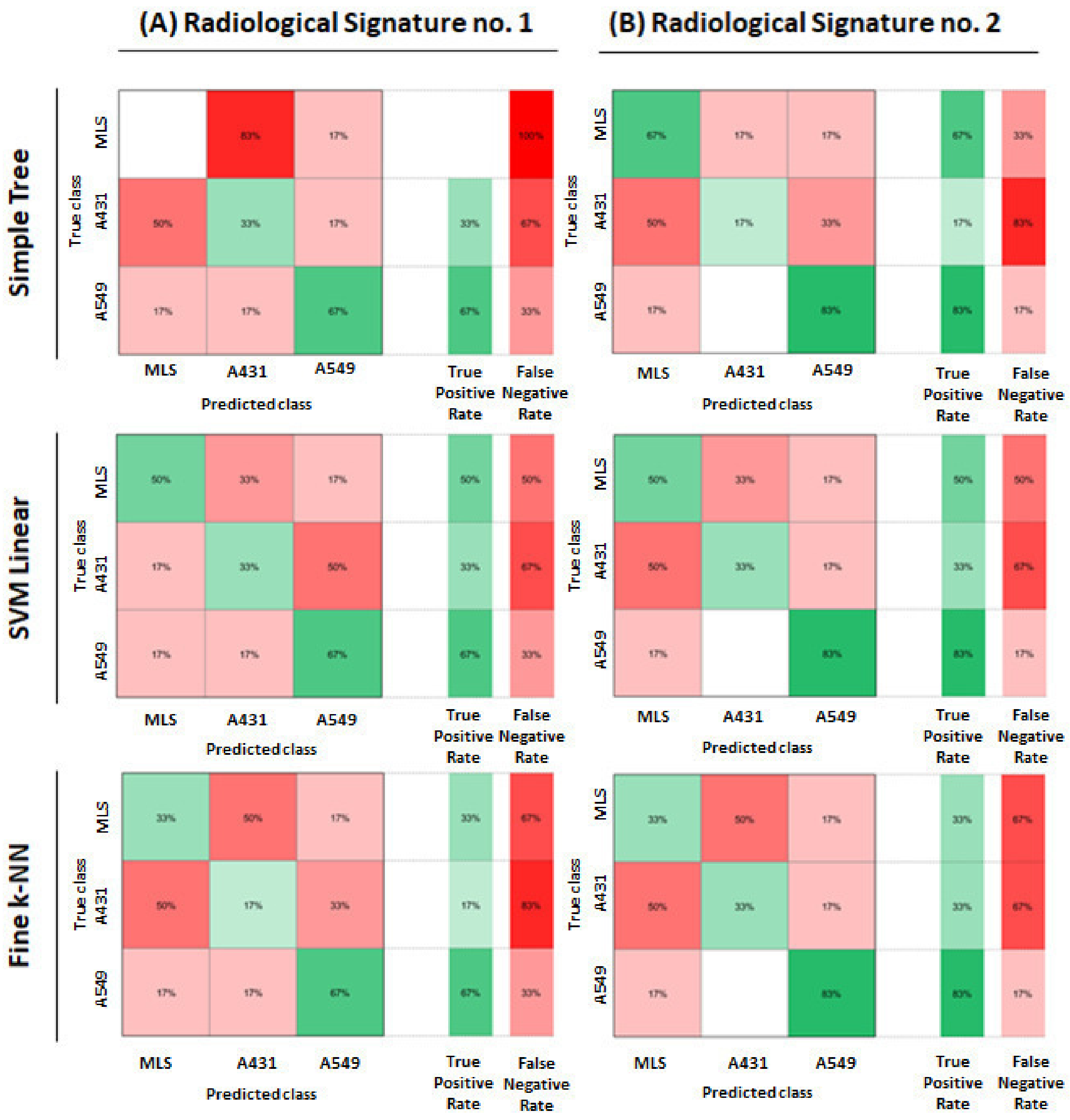
Supplementary Figure 24: Values of features enclosed in RS no.2 for each tumor mode. The values of median intensity-based feature (A), correlation textural feature (B) and short run emphasis LL the wavelet feature (C) are presented for each studied tumor model (MLS, A431, A459). The presented results were computed for the tumor segments delineated by User 2.



Supplementary Figure 25: Hierarchical agglomerative clustering-unsupervised validation of tumor classification models. In both (A) and (B) clustering of features is presented on the left side with dendrogram and assigned tumor types on the right side of the computed heat map. The linkages above the drawn dotted line constitutes the 3 final clusters. The results presented in this figure refers to the validation of the RS no.1 and RS no.2 developed with reference to the segments delineated by User 3.



Supplementary Figure 26: The comparison of supervised classification algorithms accuracy in training tumor discriminative models. (A) Presented the results obtained for training the RS no.1 with Simple Tree, SVM Linear and Fine k-NN algorithms respectively. (B) Presents the results obtained for training the RS no.2 with Simple Tree, SVM Linear and Fine k-NN algorithms respectively. The green color assigns true positives classification. The red color assigns false negative classification. The results presented in this figure refers to the validation of the RS no.1 and RS no.2 developed with reference to the segments delineated by User 3.



Supplementary Figure 27: Wilson score interval for estimation of tumor model classification probability (User 3).

	RS no.1		**RS no.2**	
	p	**95% CI**	**p**	**95% CI**
Simple Tree	0.333	0.163 - 0.563	0.556	0.337 - 0.754
Linear SVM	0.500	0.290 - 0.710	0.556	0.337 - 0.754
k-NN	0.389	0.203 - 0.614	0.500	0.290 - 0.710

Supplementary Figure 28: Values of features enclosed in RS no.2 for each tumor mode. The values of median intensity-based feature (A), correlation textural feature (B) and short run emphasis LL the wavelet feature (C) are presented for each studied tumor model (MLS, A431, A459). The presented results were computed for the tumor segments delineated by User 3.