Texture-Based Analysis of 100 MR Examinations of Head and Neck Tumors – Is It Possible to Discriminate Between Benign and Malignant Masses in a Multicenter Trial?

Texturanalyse von 100 Kopf-Hals-MRT-Untersuchungen in verschiedenen Institutionen – ist es möglich zwischen benignen und malignen Raumforderungen zu unterschieden?

Authors J. Fruehwald-Pallamar[1], J. R. Hesselink[2], M. F. Mafee[2], L. Holzer-Fruehwald[3], C. Czerny[1], M. E. Mayerhoefer[3]

Affiliations
[1] Department of Biomedical Imaging und Image-guided Therapy, Subdivision of Neuroradiology and Musculoskeletal Radiology, Medical University of Vienna, Austria
[2] Department of Radiology, UCSD Medical Center, San Diego, United States
[3] Department of Biomedical Imaging und Image-guided Therapy, Medical University of Vienna, Austria

Key words
- head/neck
- tissue characterization
- MR imaging
- technology assessment

received 27.3.2015
accepted 29.7.2015

Bibliography
DOI http://dx.doi.org/10.1055/s-0041-106066
Published online: 30.9.2015
Fortschr Röntgenstr 2016; 188: 195–202 © Georg Thieme Verlag KG Stuttgart · New York · ISSN 1438-9029

Correspondence
Frau Dr. Julia Fruehwald-Pallamar
Department of Biomedical Imaging und Image-guided Therapy, Subdivision of Neuroradiology and Musculoskeletal Radiology, Medical University Vienna
Waehringer Guertel 18–20
1090 Vienna
Austria
Tel.: ++ 43/1/40 40 04 81 80
Fax: ++ 43/1/40 40 04 89 80
julia.fruehwald-pallamar@meduniwien.ac.at

Zusammenfassung

Ziel: Ziel der Studie war die Auswertung der Texturanalyse in Bezug auf eine mögliche Unterscheidung zwischen benignen und malignen Kopf-Hals-Raumforderungen mittels konventioneller MRT-Sequenzen.

Material und Methoden: Die MRT-Daten von 100 Patienten mit histologisch verifizierten Kopf-Hals-Raumforderungen aus zwei Institutionen wurden mit einer Texturanalyse-Software untersucht. Dafür wurden 2D- und 3D-Messfelder auf allen axialen Sequenzen eingezeichnet. Folgende Texturparameter wurden für alle Messfelder berechnet: COC, RUN, GRA, ARM und WAV. Benigne und maligne Raumforderungen wurden anhand von zehn Untergruppen einer linearen Diskriminanzanalyse mit einer k-nearest-neighbor-Klassifikation zugeführt.

Ergebnisse: Die Bilder unterschieden sich aufgrund des Fabrikats und der Feldstärke der MRT-Geräte voneinander. Es war bei folgenden Sequenzen möglich zwischen benignen und malignen RF mittels TA zu differenzieren: auf den axialen STIR und T2-gewichteten-Bildern mit 2D-Messfeldern, und auf den kontrastmittelverstärkten T1-gewichteten Bilder mit Fettunterdrückung für 3D-Messfelder. In einer Subgruppenanalyse für 1,5 T- und 3T-Feldstärke konnten weitere diskriminierende Parameter erarbeitet werden.

Schlussfolgerung: Es ist möglich benigne und maligne Kopf-Hals-Raumforderungen anhand von Texturparametern zu unterscheiden, falls diese mit einem einheitlichen Protokoll auf einem Gerät untersucht werden. Wir können diese Methode allerdings nicht für eine Multicenterstudie empfehlen.

Kernaussagen:
1. Kopf-Hals-Raumforderungen können mittels 2D/3D-Texturanalyse untersucht werden

Abstract

Aim: To evaluate whether texture-based analysis of standard MRI sequences can help in the discrimination between benign and malignant head and neck tumors.

Materials and Methods: The MR images of 100 patients with a histologically clarified head or neck mass, from two different institutions, were analyzed. Texture-based analysis was performed using texture analysis software, with region of interest measurements for 2 D and 3 D evaluation independently for all axial sequences. COC, RUN, GRA, ARM, and WAV features were calculated for all ROIs. 10 texture feature subsets were used for a linear discriminant analysis, in combination with k-nearest-neighbor classification. Benign and malignant tumors were compared with regard to texture-based values.

Results: There were differences in the images from different field-strength scanners, as well as from different vendors. For the differentiation of benign and malignant tumors, we found differences on STIR and T2-weighted images for 2 D, and on contrast-enhanced T1-TSE with fat saturation for 3 D evaluation. In a separate analysis of the subgroups 1.5 and 3 Tesla, more discriminating features were found.

Conclusion: Texture-based analysis is a useful tool in the discrimination of benign and malignant tumors when performed on one scanner with the same protocol. We cannot recommend this technique for the use of multicenter studies with clinical data.

Key Points:
1. 2 D/3 D texture-based analysis can be performed in head and neck tumors
2. Texture-based analysis can differentiate between benign and malignant masses
3. Analyzed MR images should originate from one scanner with an identical protocol

2. Es ist möglich benigne und maligne Raumforderungen anhand von Texturparametern zu unterschieden.
3. Die MRT Untersuchung sollte mit gleichem Protokoll auf einem Gerät stattfinden.

Citation Format:

▶ Fruehwald-Pallamar J., Hesselink J. R., Mafee M. F. et al. Texture-Based Analysis of 100 MR Examinations of Head and Neck Tumors – Is It Possible to Discriminate Between Benign and Malignant Masses in a Multicenter Trial?. Fortschr Röntgenstr 2016; 188: 195–202

Introduction

Malignant head and neck tumors account for approximately 3 % of all malignancies [1]. Over 85 % of head and neck tumors are squamous cell carcinomas (SCC). They are strongly related to tobacco and alcohol abuse, more common in the male population, and mostly located on the mucosal surface. Tumor extent is usually evaluated with cross-sectional imaging.

Texture analysis is a computer-assisted technique that is conductive to the detection and quantification of mathematical patterns called texture features. Those texture features exist in the gray-level distribution of pixels of digital images, which the human eye can only recognize to a limited degree and is unable to quantify. With texture-based analysis, it is possible to distinguish between different types of tissues, and also between healthy and pathologically altered tissues. Texture features derived from the gray-level histogram, the co-occurrence matrix (COC; information about the gray-level value distribution of pairs of pixels, separated by a defined distance, in a given direction), the run-length matrix (RUN; information about runs of pixels with the same gray-level values, in a given direction), the absolute gradient (GRA; information about sudden signal intensity changes in the gray-level values), the auto-regressive model (ARM; description of texture based on the statistical correlation between neighboring pixels), and the wavelet transform (WAV; information about the frequency content of an image within different scales of that image) are used for tissue differentiation [2 – 5].

Texture analysis of MR images is of particular interest, because this imaging modality offers excellent depiction of anatomic details, high soft-tissue contrast, and allows enhancement of different types of tissues through the use of different pulse sequences, even without the application of contrast media [6, 7].

The aim of the study was to determine whether texture analysis of native, non-contrast-enhanced T1- and T2-weighted MR images or contrast-enhanced T1-weighted images, obtained in routine clinical practice, can provide sufficiently low rates of misclassification (high rates of correctly classified data vectors) of benign and malignant head and neck lesions.

Materials and Methods

This retrospective study included patients from two institutions who underwent a routine MR examination, and who subsequently had surgery or biopsy for pathologic correlation. The study was approved by both institutional review boards. A search for patients was performed using a full-text search in the radiology information system between January 2008 and August 2011. The single criterion for inclusion in our study was the presence of one head and neck mass, proven through fine-needle aspiration, biopsy, or operation. Exclusion criteria were the presence of motion artifacts, or a maximal lesion diameter < 5 mm. This 5 mm cut-off value was chosen to minimize the influence of partial volume effects, which might distort the true tissue-specific image texture. Based on the above-defined criteria, 100 patients (50 per institution; 55 male, 45 female) with a mean age of 50.9 years (ranging from 12 to 83 years) were included.

The MR examinations were performed on different MR scanners; 1.5 T GE (General Electric) and Siemens scanners and a 3 T GE scanner at site A, and a 1.5 T and 3 T Siemens scanner and a 3 T Philips scanner at site B. The standard examination protocols also varied somewhat. At site A, the standard protocol was: axial T1 TSE (turbo spin echo); T2 TSE fs (fat saturation) and coronal T1; and axial and coronal T1 TSE with fat saturation after gadolinium administration. The standard protocol of site B was: axial and coronal STIR (short tau inversion recovery) or T2-TIRM; axial T1-TSE before and after contrast administration; and coronal T1-TSE fs after contrast administration (sequence parameters are listed in **Table 1**). We decided to evaluate only the axial sequences (T1-TSE, STIR, T2-TSE fs, contrast-enhanced T1-TSE, and T1-TSE fs). Due to the retrospective character of the study, some sequences were not available in some patients. In total, we evaluated the following sequences: 97 T1-TSE, 22 T2-TSE, 46 T2-TSE fs, 31 STIR, 48 contrast-enhanced (ce) T1-TSE, and 43 with fat saturation.

Texture analysis was performed independently for all axial sequences (STIR, T2-TSE fs, T1-TSE before and after i. v. contrast administration (with fat saturation)). First, for the 2 D texture analysis, a manually drawn region-of-interest (ROI) was defined independently for each lesion and MRI sequence, covering the

manufacturer	field strength	sequence (axial)	TR (ms)	TE (ms)	IT (ms)	matrix
Siemens	1.5 T	(ce) T1-TSE	700	11		384 × 250
		T2-TIRM	4000 – 6000	73	150	320 × 256
	3 T	(ce) T1-TSE	600	11		448 × 310
		T2-TIRM	6000	74	180	256 × 230
Philips	3 T	(ce) T1-TSE	630	13.5		400 × 282
		STIR	6000 – 8000	15	150	400 × 256
GE	1.5 T	T1-FSE	700	13		320 × 224
		T2-FSE fs	4000 – 6000	85		256 × 224
		Ce T1-FSE fs	700	13		320 × 224
	3 T	T1-FSE	625	7		384 × 192
		T2-FSE fs	4000 – 5000	85		384 × 256
		Ce T1-FSE fs	700	8		384 × 192

Table 1 Sequence parameters.

Tab. 1 Sequenz Parameter.

entire lesion on the image that depicted the lesion at its greatest diameter. A semi-automatic, active computer algorithm using the default setting and automatic fill-in defined the ROI. Then, for the 3 D texture analysis, the tumor was encircled on every slice visible on every axial sequence in the same manner. Gray-level histogram features, co-occurrence matrix (COC) and run-length matrix (RUN) features were calculated for all 2 D and 3 D ROIs, using the popular texture analysis software MaZda (version 4.7, available at http://www.eletel.p.lodz.pl/mazda/) [8]. Additionally for all 2 D ROIs absolute gradient (GRA), autoregressive model (ARM), and wavelet transform (WAV) features were calculated by the same program (**Table 2**).

First, we wanted to investigate whether field strength influenced the performance of texture analysis on 1.5 T versus 3 T. In order to analyze this influence, we chose to evaluate and compare the texture features within our largest homogeneous group, the squamous cell carcinomas. With this group, we also analyzed whether there were differences when performing texture analysis on the MR images from different vendors (GE, Siemens, Philips).

We then tested our primary hypothesis regarding whether we could find texture features that would distinguish between benign and malignant head and neck tumors.

To identify the most valuable texture features for distinguishing between benign and malignant lesions, three subsets of 10 texture features were extracted, independently for T1-weighted, T2-TSE (±fs), STIR, and contrast-enhanced T1-weighted images (±fs), based on Fisher coefficients (ratio of between-class to within-class variance), minimization of both classification error probability and average correlation coefficients (POE+ACC), and mutual information (MI) coefficients, which measure the dependence between two or more random variable coefficients [2, 9]. This extraction was performed with the program, B11, which is integrated in the MaZda software.

For computer-assisted differentiation between benign and malignant head and neck tumors, based on the previously calculated texture feature subsets (Fisher, POE+ACC, and MI), we used linear discriminant analysis (LDA) in combination with k-nearest-neighbor classification (k-NN). K-NN calculates the rate of misclassified data vectors by comparison with the previously defined true class affiliations. The k-NN classifier implemented in the MaZda software uses the leave-one-out testing technique for this task. Thus, no separation of datasets into a training and test dataset is required. The rates of correctly classified data vectors were used as the primary outcome variables. This strategy of texture-based lesion classification has been used in previously published papers [3 – 5, 10].

Results

Of the 100 patients in our study with a head and neck mass, 54 patients (33 male, 21 female) were diagnosed with a malignant mass and 46 patients with a benign mass (24 female, 22 male). The largest histologically homogeneous group was the patients

Table 2 List of texture features and abbreviations (as used in **Table 4, 5**).

Tab. 2 Texturparameter und Abkürzungen (wie in **Tab. 4, 5** verwendet).

Gray-level histogram 2 D and 3 D	mean, variance, skewness, kurtosis, percentiles (1, 10, 50, 90, 99 %)	mean (3 D) variance (3 D) skewness (3 D) kurtosis (3 D) perc.01 %, Perc.10 %, Perc.50 %, Perc.90 %, Perc.99 % (3 D)
co-occurrence matrix (COC) 2 D and 3 D	angular second moment, contrast, correlation, entropy, sum entropy, sum of squares, sum average, sum variance, inverse difference moment, difference entropy, difference variance; (for four directions and five interpixel distances (offsets; n = 1 to 5)	S(2,-2)AngScMom, S(3,-3)AngScMom, S(1,-1)AngScMom, S(5.5)AngScMom, S(0.2)AngScMom S(3.0)Contrast, S(3.3)Contrast, S(2.2)Contrast, S(4.4)Contrast, S(5.0)Contrast, S(0.1)Contrast S(5.0)Correlat, S(5,-5)Correlat, S(2,-2)Correlat, S(1,-1)Correlat, S(0.1)Correlat, S(3,-3)Correlat S(3.0)Entropy, S(5,-5)Entropy, S(2,-2)Entropy, S(4,-4)Entropy, S(0.5)Entrop, S(3,-3)Entropy S(1.0)SumEntrp, S(0.2)Entropy, S(0.5)SumEntrp, S(4.0)SumEntrp, S(3.0)SumEntrp, S(3.0)SumOfSqs, S(4.0)SumOfSqs, S(0.5)SumOfSqs, S(5,-5)SumOfSqs S(2,-2)SumAverg, S(3,-3)SumAverg, S(0.3)SumAverg, S(0.5)SumAverg, S(0.4)SumAverg S(1.1)SumVarnc, S(3.0)SumVarnc, S(5,-5)SumVarnc, S(0.3)InvDfMom, S(1.0)InvDfMom, S(4,-4)InvDfMom, S(5.0)InvDfMom, S(0.1)InvDfMom, S(5.5)DifEntrp, S(4.4)DifEntrp, S(1,-1)DifEntrp, S(0.1)DifEntrp, S(2.0)DifEntrp, S(1.0)DifEntrp S(3.3)DifVarnc, S(1,-1)DifVarnc, S(2,-2)DifVarnc, S(5,-5)DifVarnc, S(2.0)DifVarnc
run-length matrix (RUN) 2 D and 3 D	run-length non-uniformity, gray-level non-uniformity, long run emphasis, short run emphasis, fraction of image in runs; (for four angles)	Vertl_GLevNonU,,45dgr_GLevNonU, Horzl_GLevNonU, 135dr_GLevNonU, Z_GLevNonU Vertl_RLNonUni, 45dgr_RLNonUni,135dr_RLNonUni Vertl_LngREmph, 45dgr_LngREmph, Horzl_LngREmph, 135dr_LngREmph Vertl_ShrtREmp, 45dgr_ShrtREmp, Horzl_ShrtREmp, 135dr_ShrtREmp Vertl_Fraction, 45dgr_Fraction, Horzl_Fraction, 135dr_Fraction, Z_Fraction
absolute gradient (GRA) 2 D	gradient mean, variance, skewness, kurtosis, non-zeros	GrMean GrVariance GrSkewness GrKurtosis GrNonZeros
autoregressive model (ARM) 2 D	theta 1 to 4, sigma	Teta1, Teta2, Teta3, Teta4 Sigma
wavelet transform (WAV) 2 D	energies of wavelet transform coefficients in sub-bands LL,LH, HL,HH; (for three subsampling factors)	WavEnLL_s-1, WavEnLL_s-2,, WavEnLL_s-3, WavEnLL_s-4 WavEnLH_s-1, WavEnLH_s-2, WavEnLH_s-3 WavEnHL_s-1, WavEnHL_s-2, WavEnHL_s-3, WavEnHL_s-4 WavEnHH_s-1, WavEnHH_s-2, WavEnHH_s-3

Table 3 Histologic subgroups.

Tab. 3 Pathologien.

	histologic subgroup	number of patients
benign tumors n = 46	cysts	8
	inflammatory mass/abscess	5
	parotid tumor	9
	glomus tumor	9
	vascular lesion/malformation	5
	schwannoma	4
	other	6
malignant tumors n = 54	squamous cell carcinoma	31
	lymphoma	8
	adenoid cystic carcinoma	5
	adenocarcinoma	4
	other	6

diagnosed with (SCC) squamous cell carcinoma (n = 31; 21 male, 10 female; ● **Table 3**).

SCC – comparison of vendors and field strength

Considerable differences in terms of texture features were observed between the squamous cell carcinomas at 1.5 and 3 Tesla. Discrimination based on 2 D texture features extracted from T1-weighted sequences produced misclassification numbers of 12.9 % for POE+ACC, and 19.35 % for MI. We found a low misclassification rate only for MI (19.35 %). Similar results were found in the analysis of the T1-weighted sequences with the 3 D ROI (correctly classified data vectors [percentages] and the list of the 10 extracted features are listed in ● **Table 4**). For the discrimination of SCC, mutual information coefficients proved to be the most effective method.

Table 4 List of correctly classified data vectors and the texture feature subset best suited for the discrimination of SCC at 1.5 and 3.0 Tesla.

Tab. 4 Korrekt klassifizierte Datenvektoren und Texturparameter Subset für die Unterscheidung von Plattenepithelkarzinomen bei 1,5 und 3 Tesla.

	sequence	fisher	POE+ACC	MI
2 D	T1-TSE	18/31 or 58.06 %	27/31 or 87.1 %	25/31 or 80.65 %
	features	1 Perc.01 %	1 Vertl_RLNonUni	1 S(4.4)DifEntrp
		2 S(5.5)DifEntrp	2 Perc.50 %	2 S(1,-1)DifEntrp
		3 Perc.10 %	3 Vertl_ShrtREmp	3 S(3.0)Entropy
		4 Mean	4 Mean	4 S(0.3)SumAverg
		5 Perc.90 %	5 Perc.10 %	5 S(0.5)SumAverg
		6 Perc.50 %	6 S(5.5)DifEntrp	6 S(0.3)InvDfMom
		7 Perc.99 %	7 Perc.90 %	7 135dr_Fraction
		8 Vertl_RLNonUni	8 S(4.4)DifEntrp	8 135dr_ShrtREmp
		9 S(4.4)DifEntrp	9 Perc.01 %	9 S(2,-2)SumAverg
		10 Vertl_ShrtREmp	10 Perc.99 %	10 S(3,-3)SumAverg
	without gray-level histogram features	24/31 or 77.42 %	21/31 or 67.7 %	25/31 or 80.65 %
	features	1 S(5.5)DifEntrp	1 Vertl_GLevNonU	1 S(4.4)DifEntrp
		2 Vertl_RLNonUni	2 WavEnHL_s-2	2 S(1,-1)DifEntrp
		3 S(4.4)DifEntrp	3 Vertl_ShrtREmp	3 S(3.0)Entropy
		4 Vertl_ShrtREmp	4 Vertl_RLNonUni	4 S(0.3)SumAverg
		5 S(5.5)Entropy	5 WavEnLH_s-2	5 S(0.5)SumAverg
		6 S(5.0)Correlat	6 Vertl_LngREmph	6 S(0.3)InvDfMom
		7 Vertl_LngREmph	7 S(3.0)Entropy	7 135dr_Fraction
		8 45dgr_ShrtREmp	8 WavEnLL_s-1	8 135dr_ShrtREmp
		9 S(3.3)DifVarnc	9 WavEnHH_s-1	9 S(2,-2)SumAverg
		10 Vertl_Fraction	10 Vertl_Fraction	10 S(3,-3)SumAverg
	T2-TSE fs	14/18 or 77.78 %	12/18 or 66.66 %	16/18 or 88.9 %
	features	1 45dgr_ShrtREmp	1 Vertl_GLevNonU	1 S(3,-3)AngScMom
		2 45dgr_Fraction	2 S(2,-2)AngScMom	2 S(2,-2)Entropy
		3 S(5,-5)Correlat	3 WavEnHL_s-2	3 S(1.0)SumEntrp
		4 S(2,-2)Correlat	4 Variance	4 S(0.2)Entropy
		5 S(5,-5)Entropy	5 Vertl_ShrtREmp	5 S(2,-2)AngScMom
		6 WavEnHL_s-1	6 WavEnHL_s-1	6 S(4,-4)Entropy
		7 S(5,-5)SumEntrp	7 WavEnLH_s-3	7 S(0.5)Entropy
		8 45dgr_LngREmph	8 WavEnLL_s-4	8 S(3,-3)Entropy
		9 S(1,-1)Correlat	9 Vertl_RLNonUni	9 S(0.5)SumEntrp
			10 WavEnHH_s-1	10 S(1,-1)Correlat
3 D	T1-TSE	24/31 or 77.42 %	24/31 or 77.42 %	25/31 or 80.65 %
	features	1 Vertl_ShrtREmp	1 S(0.1.0)InvDfMom	1 Vertl_ShrtREmp
		2 Vertl_GLevNonU	2 Horzl_LngREmph	2 Vertl_Fraction
		3 135dr_GLevNonU	3 135dr_GLevNonU	3 Vertl_LngREmph
		4 Vertl_Fraction	4 GrKurtosis	4 135dr_Fraction
		5 45dgr_GLevNonU	5 GrMean	5 Perc.01 %3 D
		6 Z_GLevNonU	6 GrNonZeros	6 135dr_ShrtREmp
		7 Vertl_LngREmph	7 GrSkewness	7 Perc.90 %3 D
		8 Horzl_GLevNonU	8 GrVariance	8 Mean3 D
		9 S(1,-1.0)DifEntrp	9 Vertl_LngREmph	9 Perc.50 %3 D
		10 S(1,-1.0)DifVarnc	10 Vertl_GLevNonU	10 Perc.99 %3 D

We also found textural differences between the three different vendors (GE, Siemens, and Philips) for the 2 D ROIs on the T1-weighted sequence, with the following correctly classified data vectors: 26/31 or 83.87 % for Fisher (41.94 % without gray-level histogram features); 24/31 or 77.42 % for POE+ACC; and 28/31 or 90.32 % for MI (80.65 % without gray-level histogram features).

Discrimination of benign and malignant tumors

For the discrimination of benign and malignant lesions, STIR and T2-weighted images contained the most relevant texture features for 2 D evaluation. For 3 D texture-based analysis, only contrast-enhanced, T1-weighted images with fat saturation had a low misclassification rate. **Table 5** shows the results for the LDA of the comparison between benign and malignant tumors.

Table 5 List of correctly classified data vectors and the texture feature subset best suited for the discrimination of benign and malignant tumors.

Tab. 5 Korrekt klassifizierte Datenvektoren und Texturparameter Subset für die Unterscheidung von benignen und malignen Raumforderungen.

	sequence	fisher	POE+ACC	MI
2 D	T1-TSE	51/97 or 52.58 %	60/97 or 38.14 %	50/97 or 48.45 %
	features	1 S(2,-2)Correlat	1 Vertl_RLNonUni	1 S(2,-2)Correlat
		2 S(1.1)Correlat	2 WavEnLH_s-3	2 S(3.0)Contrast
		3 S(2.2)Correlat	3 Variance	3 Vertl_RLNonUni
		4 S(5.0)Correlat	4 Vertl_LngREmph	4 Vertl_GLevNonU
		5 S(3.3)DifVarnc	5 WavEnLL_s-1	5 S(5,-5)DifVarnc
		6 S(1,-1)Correlat	6 Vertl_GLevNonU	6 WavEnHH_s-2
		7 S(2,-2)DifVarnc	7 WavEnHL_s-1	7 S(1.1)SumVarnc
		8 S(0.1)Correlat	8 Vertl_ShrtREmp	8 S(2.0)DifVarnc
		9 S(3,-3)DifVarnc	9 WavEnHH_s-1	9 S(2.2)DifVarnc
		10 S(5,-5)DifVarnc	10 S(3.3)DifVarnc	10 S(4,-4)InvDfMom
	STIR	19/31 or 61.29 %	20/31 or 64.52 %	25/31 or 80.65 %
	features	1 S(0.1)Correlat	1 Vertl_Fraction	1 WavEnHL_s-4
		2 Vertl_ShrtREmp	2 WavEnHL_s-4	2 S(0.1)Correlat
		3 Vertl_Fraction	3 Variance	3 WavEnLH_s-1
		4 S(0.1)InvDfMom	4 WavEnHH_s-2	4 S(4.0)SumEntrp
		5 WavEnHL_s-2	5 WavEnHL_s-1	5 S(1.1)DifEntrp
		6 S(1.0)InvDfMom	6 Vertl_LngREmph	6 S(1,-1)AngScMom
		7 WavEnHL_s-1	7 Vertl_RLNonUni	7 S(3.0)SumEntrp
		8 S(1,-1)Correlat	8 Vertl_ShrtREmp	8 WavEnHH_s-3
		9 Vertl_LngREmph	9 WavEnLL_s-3	9 S(1,-1)Correlat
		10 WavEnHH_s-3	10 S(3,-3)Correlat	10 S(3.3)Correlat
	Ce T1	35/48 or 77.92 %	36/48 or 75 %	27/48 or 56.25 %
	features	1 Skewness	1 Vertl_GLevNonU	1 S(1,-1)DifEntrp
		2 S(1.1)Correlat	2 Vertl_LngREmph	2 S(2,-2)SumEntrp
		3 S(0.1)Correlat	3 WavEnLL_s-3	3 S(4.0)SumEntrp
		4 S(5.5)AngScMom	4 Variance	4 S(3.0)SumVarnc
		5 Perc.01 %	5 WavEnLL_s-2	5 S(0.3)SumEntrp
		6 WavEnLH_s-3	6 Vertl_ShrtREmp	6 S(0.1)DifEntrp
		7 S(1.1)AngScMom	7 Vertl_RLNonUni	7 S(2.0)DifEntrp
		8 S(0.2)AngScMom	8 WavEnHH_s-3	8 S(4.0)SumOfSqs
		9 S(1,-1)Contrast	9 WavEnHL_s-3	9 S(5.0)InvDfMom
		10 S(3.3)Correlat	10 S(3.0)SumOfSqs	10 S(0.1)InvDfMom
	T2-TSE		18/22 or 81.82 %	22/22 or 100 %
	features	1 Teta4	1 WavEnHH_s-3	1 S(2,-2)InvDfMom
		2 Teta3	2 WavEnHH_s-1	2 S(0.1)DifEntrp
		3 Teta2	3 WavEnHL_s-2	3 WavEnLH_s-1
		4 Teta1	4 WavEnHL_s-1	4 WavEnHH_s-3
		5 Sigma	5 Vertl_LngREmph	5 WavEnHL_s-2
		6 S(3.3)Contrast	6 WavEnHL_s-3	6 S(3.3)InvDfMom
		7 S(1.0)DifEntrp	7 WavEnLH_s-1	7 S(1,-1)DifEntrp
		8 S(4,-4)InvDfMom	8 WavEnLH_s-3	8 S(4.4)InvDfMom
		9 S(0.1)DifEntrp	9 Vertl_RLNonUni	9 S(4.0)DifVarnc
		10 S(1.0)InvDfMom	10 Vertl_GLevNonU	10 S(2.2)Contrast
	T2-TSE fs	36/46 or 78.26 %	33/46 or 71.74 %	34/46 or 73.91 %
	features	1 GrKurtosis	1 WavEnHH_s-1	1 S(4.4)AngScMom
		2 S(5.0)DifVarnc	2 WavEnLL_s-3	2 GrKurtosis
		3 S(5.0)Contrast	3 Variance	3 S(0.2)DifEntrp
		4 S(5,-5)SumVarnc	4 Vertl_GLevNonU	4 S(0.1)AngScMom
		5 S(0.1)DifVarnc	5 Vertl_LngREmph	5 S(0.1)DifVarnc
		6 S(5.0)SumOfSqs	6 WavEnHH_s-3	6 S(0.2)Correlat
		7 S(5,-5)DifVarnc	7 WavEnLL_s-2	7 S(0.1)Contrast
		8 S(4.0)DifVarnc	8 Vertl_RLNonUni	8 S(4.4)DifEntrp
		9 S(5,-5)SumOfSqs	9 WavEnLH_s-3	9 S(3.3)AngScMom
		10 WavEnHH_s-1	10 S(4,-4)Correlat	10 S(5.5)Entropy

Table 5 (Continuation)

	sequence	fisher	POE+ACC	MI
	ce T1-TSE fs	26/43 or 80.47 %	20/43 or 46.51 %	25/43 or 58.14 %
	features	1 S(4.4)Correlat	1 Variance	1 S(0.5)InvDfMom
		2 S(3.0)Correlat	2 WavEnLL_s-1	2 S(4.4)DifEntrp
		3 S(2.0)Correlat	3 Vertl_RLNonUni	3 S(0.3)SumAverg
		4 S(1.0)DifVarnc	4 WavEnHH_s-1	4 S(0.4)SumAverg
		5 S(5.0)Correlat	5 Vertl_LngREmph	5 S(0.5)Entropy
		6 S(3.3)Correlat	6 Vertl_GLevNonU	6 S(0.5)SumAverg
		7 S(1.0)Contrast	7 WavEnLL_s-2	7 S(0.5)SumOfSqs
		8 S(1.0)Correlat	8 Vertl_ShrtREmp	8 S(0.4)InvDfMom
		9 45dgr_Fraction	9 WavEnHL_s-1	9 S(3.0)Correlat
		10 S(4.4)Contrast	10 S(2.0)Correlat	10 S(2.0)SumAverg
3 D	T1-TSE	40/96 or 41.67 %	45/96 or 46.88 %	43/96 or 44.79 %
	features	1 Vertl_ShrtREmp	1 45dgr_LngREmph	1 Horzl_Fraction
		2 Vertl_GLevNonU	2 Vertl_LngREmph	2 Horzl_ShrtREmp
		3 Vertl_LngREmph	3 Horzl_LngREmph	3 Horzl_LngREmph
		4 Horzl_GLevNonU	4 135dr_LngREmph	4 45dgr_RLNonUni
		5 45dgr_LngREmph	5 Vertl_GLevNonU	5 135dr_GLevNonU
		6 45dgr_GLevNonU	6 Horzl_GLevNonU	6 Vertl_GLevNonU
		7 45dgr_Fraction	7 Horzl_Fraction	7 45dgr_LngREmph
		8 135dr_GLevNonU	8 45dgr_ShrtREmp	8 135dr_LngREmph
		9 Vertl_Fraction	9 Vertl_Fraction	9 Horzl_GLevNonU
		10 Vertl_RLNonUni	10 135dr_GLevNonU	10 135dr_RLNonUni
	ce T1-TSE	27/52 or 51.92 %	32/52 or 61.54 %	29/52 or 55.77 %
	features	1 45dgr_ShrtREmp	1 S(1.0.0)Correlat	1 45dgr_Fraction
		2 45dgr_Fraction	2 135dr_RLNonUni	2 45dgr_LngREmph
		3 S(0.0.1)AngScMom	3 Z_LngREmph	3 135dr_ShrtREmp
		4 S(1.0.0)Contrast	4 S(0.0.1)InvDfMom	4 Kurtosis3 D
		5 S(1.0.0)InvDfMom	5 Skewness3 D	5 Horzl_Fraction
		6 S(0.0.1)SumVarnc	6 S(0.0.1)Correlat	6 45dgr_ShrtREmp
		7 S(1.0.0)DifVarnc	7 Vertl_LngREmph	7 Vertl_ShrtREmp
		8 S(1.1.0)Correlat	8 Kurtosis3 D	8 Vertl_Fraction
		9 S(0.0.1)SumOfSqs	9 Mean3 D	9 S(0.0.1)SumAverg
		10 S(0.0.1)InvDfMom	10 45dgr_LngREmph	10 Skewness3 D
	T2-TSE	17/26 or 65.38 %	15/26 or 57.69 %	13/26 or 50.00 %
	features	1 45dgr_LngREmph	1 Horzl_GLevNonU	1 45dgr_LngREmph
		2 45dgr_Fraction	2 Skewness3 D	2 Perc.01 %3 D
		3 Vertl_Fraction	3 Vertl_LngREmph	3 Horzl_Fraction
		4 Vertl_ShrtREmp	4 Z_Fraction	4 45dgr_Fraction
		5 45dgr_ShrtREmp	5 Perc.01 %3 D	5 Z_Fraction
		6 Vertl_LngREmph	6 S(0.0.1)Correlat	6 45dgr_ShrtREmp
		7 Horzl_ShrtREmp	7 Vertl_GLevNonU	7 Z_LngREmph
		8 Horzl_Fraction	8 45dgr_LngREmph	8 Horzl_LngREmph
		9 135dr_LngREmph	9 Z_LngREmph	9 Horzl_ShrtREmp
		10 135dr_Fraction	10 45dgr_Fraction	10 Z_ShrtREmp
	T2-TSE fs	30/44 or 68.18 %	32/44 or 72.73 %	32/44 or 72.73 %
	features	1 Z_LngREmph	1 S(0.1.0)Correlat	1 Variance3 D
		2 Horzl_LngREmph	2 S(0.0.1)Correlat	2 Z_ShrtREmp
		3 135dr_LngREmph	3 S(1.0.0)Contrast	3 Z_Fraction
		4 S(0.0.1)Correlat	4 Vertl_ShrtREmp	4 Horzl_LngREmph
		5 S(0.0.1)SumAverg	5 Skewness3 D	5 Z_LngREmph
		6 GrVariance	6 Horzl_Fraction	6 Vertl_ShrtREmp
		7 GrSkewness	7 S(1.1.0)Correlat	7 Skewness3 D
		8 GrKurtosis	8 Z_GLevNonU	8 Vertl_RLNonUni
		9 GrNonZeros	9 Kurtosis3 D	9 135dr_LngREmph
		10 GrMean	10 Horzl_LngREmph	10 Horzl_Fraction
	ce T1-TSE fs	24/41 or 58.54 %	35/41 or 85.37 %*	29/41 or 70.73 %
	features	1 S(0.1.0)Correlat	1 135dr_ShrtREmp	1 135dr_GLevNonU
		2 S(1.0.0)Correlat	2 S(1,-1.0)Correlat	2 135dr_RLNonUni
		3 S(1.1.0)Correlat	3 Perc.10 %3 D	3 S(0.0.1)Correlat
		4 S(1,-1.0)Correlat	4 Z_GLevNonU	4 Perc.01 %3 D
		5 45dgr_ShrtREmp	5 Kurtosis3 D	5 45dgr_Fraction
		6 S(0.0.1)SumAverg	6 45dgr_ShrtREmp	6 Horzl_ShrtREmp
		7 Z_ShrtREmp	7 Vertl_LngREmph	7 135dr_ShrtREmp
		8 45dgr_Fraction	8 S(1.0.0)Correlat	8 Horzl_Fraction
		9 S(0.0.1)SumEntrp	9 135dr_GLevNonU	9 Z_GLevNonU
		10 S(0.0.1)DifEntrp	10 Perc.01 %3 D	10 S(0.1.0)Correlat

We then decided to analyze the 2 D ROI of benign and malignant tumors at 1.5 and 3.0 Tesla separately. This analysis showed better results in the discrimination between benign and malignant tumors. For the T1-weighted images, we found 41/49 or 83.67 % correctly classified data vectors at 3 T with MI. The contrast-enhanced, T1-weighted images with fat saturation at 1.5 T had 86.67 % (or 26/30) correctly classified data vectors. For the T2-weighted images with fat saturation (fs), all vectors were correctly classified at 3 T, with 81.48 % (Fisher), and, at 1.5 T, 92.59 % (POE +ACC – and MI) were correctly classified (n = 27).

Additionally, we analyzed the 2 D ROI on axial STIR images of a single 3 T unit (n = 28). All vectors were correctly classified in 71.43 % vs. 68.82 % (n = 31) for Fisher, POE+ACC and MI – in detail: 16/28 for Fisher; 18/28 for POE+ACC, and 22/28 for MI.

Discussion

In contrast to the published literature about the application of texture-based analysis in phantoms and in patients examined on one scanner with the same protocol, our data show that, when using different MR protocols on different MR scanners (field-strength, vendor), texture-based analysis is not practical for the differentiation between benign and malignant head and neck tumors. Although we found discriminating features for STIR and T2-weighted images, the number of correctly classified data vectors was not sufficient to implement this technique in the routine evaluation of MR images.

Several studies have shown the successful use of texture-based analysis in the brain [11, 12], the liver [3, 10], and muscle [13]. Herlidu-Meme et al. performed MR imaging at three sites, but with the same brand and field-strength MR unit [14]. Jirak et al. performed phantom measurements on six different MR scanners, all using identical MR parameters for the acquisition protocol [15]. In a recently published study, Fruehwald-Pallamar et al. were able to discriminate between benign and malignant parotid tumors based on texture analysis [5] – notably, using images obtained using a single MR unit and identical protocol. Contrary to the above described successful applications of texture analysis with (more or less) uniform scanning conditions, previous studies indicated considerable influence of protocol heterogeneity. In two articles Lerski et al. reported that texture-based information extracted from foam test objects was susceptible to the choice of MR protocol and scanner [16, 17]. Mayerhoefer et al. showed in a phantom study that texture features (especially COC and RUN) are sensitive to variations in the scanned protocol (matrix size, TR, TE, and number of acquisitions) [7].

To our knowledge, texture-based analysis has not been applied for the analysis of different tissue types (different kinds of head and neck masses) in a multicenter study. In the present study, routine MR images were obtained on several MR units, with similar imaging protocols (plane, and the use of fat saturation), but with each center using its own acquisition parameters. This was done to determine the clinical value of this technique and to analyze which texture feature subset and extraction method would perform best under given conditions. Therefore, we also tried to acquire a large collective of 100 patients.

Our clinical data resemble the phantom data from the above-mentioned study by Lerski et al. because the texture features extracted from SCC clearly differed between examinations performed at different field strengths, and also between images obtained from scanners by different vendors [16]. For instance, our classifier was able to differentiate between the SCC at 1.5 and the SCC at 3 Tesla in almost 90 % of the cases, based on texture features extracted from T2 TSE fs images. It is quite possible that the strong influence of acquisition parameters affected the texture-based discrimination of benign and malignant head and neck tumors in our study. When performing the analysis with the data of one scanner (benign vs. malignant lesions on one 3 T unit), we were able to increase the correctly classified data vectors by 3 % for axial STIR images. Our results were generally less favorable than those reported by Fruehwald-Pallamar et al. [5] in a single-center study using 2 D texture features. Notably, our results indicate that 3 D texture features are not superior to a 2 D evaluation of a single slice for head and neck masses—this topic has not been evaluated thus far in a clinical setting. For a possible future introduction into routine clinical practice, and also for clinical trials involving texture analysis, we recommend that only data from a single scanner are used, in order to avoid an incalculable influence from hardware specifications. Furthermore, 2 D evaluation is less time-consuming, and therefore, more practical, while providing results comparable to 3 D evaluation, where every slice of the tumor has to be processed.

We note the limitations of our study: There were minor acquisition parameter differences, even between comparable pulse sequences obtained by the different scanners, as mentioned above.

In conclusion, texture-based analysis has the potential to help with the discrimination of benign and malignant head and neck tumors when performed on the same scanner with an identical protocol. We cannot recommend the application of texture-based analysis in a multicenter study with different types of scanners and varying acquisition parameters.

Clinical relevance

Since morphological sequences alone often do not allow a correct diagnosis of different types of tumors in the head-neck-region, it is of great help for radiologists to have additional information to generate the correct diagnosis. Texture-based analysis is a tool, which has the potential to help with the discrimination between benign and malignant head and neck tumors, provided that the examinations take place on one specific machine with a defined protocol.

Acknowledgments

This Research was funded by the Austrian Science Fund (FWF): J3200-B13.

References

1 *Jemal A, Siegel R, Ward E et al.* Cancer statistics, 2007. CA Cancer J Clin 2007; 57: 43 – 66
2 *Hajek M, Dezortova M, Materka A et al.* Texture analysis for magnetic resonance imaging; 2006
3 *Mayerhoefer ME, Schima W, Trattnig S et al.* Texture-based classification of focal liver lesions on MRI at 3.0 Tesla: a feasibility study in cysts and hemangiomas. J Magn Reson Imaging 2010; 32: 352 – 359
4 *Mayerhoefer ME, Breitenseher M, Amann G et al.* Are signal intensity and homogeneity useful parameters for distinguishing between benign and malignant soft tissue masses on MR images? Objective evalu-

ation by means of texture analysis. Magn Reson Imaging 2008; 26: 1316 – 1322
5 *Fruehwald-Pallamar J*, *Czerny C*, *Holzer-Fruehwald L et al.* Texture-based and diffusion-weighted discrimination of parotid gland lesions on MR images at 3.0 Tesla. NMR in biomedicine 2013; 11: 1372 – 1379
6 *Mayerhoefer ME*, *Szomolanyi P*, *Jirak D et al.* Effects of magnetic resonance image interpolation on the results of texture-based pattern classification: a phantom study. Invest Radiol 2009; 44: 405 – 411
7 *Mayerhoefer ME*, *Szomolanyi P*, *Jirak D et al.* Effects of MRI acquisition parameter variations and protocol heterogeneity on the results of texture analysis and pattern discrimination: an application-oriented study. Med Phys 2009; 36: 1236 – 1243
8 *Szczypinski PM*, *Strzelecki M*, *Materka A et al.* MaZda–a software package for image texture analysis. Computer methods and programs in biomedicine 2009; 94: 66 – 76
9 *Tourassi GD*, *Frederick ED*, *Markey MK et al.* Application of the mutual information criterion for feature selection in computer-aided diagnosis. Med Phys 2001; 28: 2394 – 2402
10 *Jirak D*, *Dezortova M*, *Taimr P et al.* Texture analysis of human liver. J Magn Reson Imaging 2002; 15: 68 – 74
11 *Georgiadis P*, *Cavouras D*, *Kalatzis I et al.* Enhancing the discrimination accuracy between metastases, gliomas and meningiomas on brain MRI by volumetric textural features and ensemble pattern recognition methods. Magn Reson Imaging 2009; 27: 120 – 130
12 *Zacharaki EI*, *Wang S*, *Chawla S et al.* Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med 2009; 62: 1609 – 1618
13 *Herlidou S*, *Rolland Y*, *Bansard JY et al.* Comparison of automated and visual texture analysis in MRI: characterization of normal and diseased skeletal muscle. Magn Reson Imaging 1999; 17: 1393 – 1397
14 *Herlidou-Meme S*, *Constans JM*, *Carsin B et al.* MRI texture analysis on texture test objects, normal brain and intracranial tumors. Magn Reson Imaging 2003; 21: 989 – 993
15 *Jirak D*, *Dezortova M*, *Hajek M.* Phantoms for texture analysis of MR images. Long-term and multi-center study. Med Phys 2004; 31: 616 – 622
16 *Lerski RA*, *Schad LR*, *Luypaert R et al.* Multicentre magnetic resonance texture analysis trial using reticulated foam test objects. Magn Reson Imaging 1999; 17: 1025 – 1031
17 *Lerski RA*, *Schad LR.* The use of reticulated foam in texture test objects for magnetic resonance imaging. Magn Reson Imaging 1998; 16: 1139 – 1144