

The use of texture analysis in the grading of breast cancer on MR images: preliminary findings

Shelley Waugh^{1,2}, Richard Lerski^{1,2}, Luc Bidaut², and Alastair Thompson^{2,3}

¹Medical Physics, Ninewells Hospital, Dundee, Angus, United Kingdom, ²University of Dundee, Dundee, United Kingdom, ³Department of Surgery, Ninewells Hospital, Dundee, United Kingdom

Introduction and Aims: Breast Magnetic Resonance Imaging (MRI) is becoming more routinely used as a complementary imaging modality to conventional x-ray based mammography and ultrasound in the identification and diagnosis of breast cancer. The main reasons for its increase in use are the excellent sensitivity and negative predictive value of the dynamic contrast-enhanced images. While the sensitivity is around 100% [1], the reported specificity in the differentiation of malignant from benign disease is still unclear, with reports varying from 40-92% [1,2].

Texture analysis is a computer-assisted method of statistically evaluating the grey-level pixel intensity distribution within an image using various models in order to infer information that may not be visually apparent. It has been successfully applied in many branches of medical diagnostics and more recently it has been used to increase the specificity of breast MRI examinations [3,4]. The stage of cancer growth is diagnosed using histological patterns such as how variable the cell characteristics are, the number of cells in a given area and the degree of glandular formation [5]. As these are all properties that should affect the microscopic texture, we proposed that it should therefore be possible to differentiate between the grades of cancer using computer-aided texture analysis.

This study considers a group of patients who were referred from the breast clinic for a breast MRI examination on the basis of a known malignancy, which had been pathologically confirmed and staged prior to the examination. Based on these data, we aimed to identify how well texture analysis was able to differentiate between different grades in both ductal and lobular cancers.

Methods: A total of 57 patients were imaged on a 16-channel 1.5 Tesla (T) MRI scanner (Avanto; Siemens, Erlangen) using a 4-channel breast matrix coil. A 3-dimensional Fast Low Angle SHot (FLASH) sequence was acquired in the axial orientation through both breasts with an acquisition time of 62s (repetition time= 3.52 ms, echo time= 1.24 ms, flip angle= 6°, slice thickness= 0.83mm, field of view= 320x320 mm, matrix= 384x384, bandwidth= 650 Hertz/pixel, parallel imaging factor x2, 192 slices). The complete dynamic acquisition consisted of 8 volumes (total imaging time 8 minutes 17 seconds), with contrast administered after the second volume acquisition. All patients were injected with a 0.1mmol/kg dose of Dotarem (Guerbet Laboratories, France) at a rate of 2.0 ml/s, followed by a 20ml saline solution injection at the same rate. Regions of contrast uptake were demonstrated by producing subtracted volumes for each acquisition. All analysis was performed using the two-minute post contrast subtracted volume images.

Texture analysis was carried out using MaZda version 4.7 [6]. The three slices on which the lesion was best visualised were first identified before drawing circular regions of interest (ROI) within the lesion. The ROI's were drawn such that they were as large as possible without straying from the region of enhancement, in order to maximise the counting statistics. Grey level normalisation was carried out using a $\mu \pm 3\sigma$ regime (μ - grey level mean, σ - grey level standard deviation) to minimise the effect of image brightness and contrast on the outcome of texture analysis.

Texture features were calculated as derived from the auto-regressive model (ARM), co-occurrence matrix (COM), absolute gradient (GRA), run-length matrix (RLM) and wavelet transform (WAV). The best 30 features were automatically calculated by MaZda using a combination of the Fischer coefficient, Mutual Information and the Probability of Error. The COM and RLM model features were also considered separately as these models were found to contribute most to the 30 best features (accounting for over 96% of the contributing features).

The ductal and lobular cancers were considered separately, and classification and statistical testing were carried out in order to determine whether there were measurable differences in the texture characteristics between each grading of cancer. Classification was carried out by Weka, version 3.6.2 [7], using a 10-folds cross validation routine and k-nearest neighbour (k-NN) classification with k=1. The number of incorrectly identified vectors was represented by the percentage of misclassified vectors. The actual feature values calculated by MaZda were also considered and statistical testing was carried out using a Mann Whitney U test in SPSS, version 18 (IBM corp.; Armonk, NY). A p-value of less than 0.05 was considered to be significant.

Results and Discussion: In our patient group, a total of 83 lesions were identified on the MRI images. Pathology results showed that there were 46 invasive ductal and 37 infiltrative lobular cancers. Of the ductal cancers, 2 were identified as grade 1, 25 as grade 2 and 19 as grade 3, resulting in a total of 6, 75 and 57 ROI respectively. For the lobular cancers, there were 32 grade 2 lesions (96 ROI) and 5 grade 3 lesions (15 ROI).

Using the best 30 features, the grade 1 ductal cancers were well differentiated from the grade 2 and 3, with a 0% misclassification and a significant difference in feature values demonstrated by the Mann Whitney U test (Table 1). There was a misclassification rate of 1.5% for grade 2 vs 3 data, with two grade 3 cancers classified as grade 2. For the lobular cancers, the 30 best features resulted in a 0.9% misclassification, with one grade 2 cancer classified as a grade 3. The difference in feature values, however, was statistically significant with $p < 0.001$.

Considering the COM only features, classification accuracy was excellent across all grade comparisons and across both cancer types. The individual features were significantly different for all but the grade 2 vs 3 ductal cancers, where $p = 0.108$. A summary of the COM features for each grade and cancer type is shown in Figure 1.

The RLM features performed poorly for differentiating between the cancer grades for both lobular and ductal cancers. There were no significant differences found between feature values and misclassification rates were above around 5%.

This study demonstrates that texture analysis has potential for differentiating between different grades of the two most common breast cancers- invasive ductal and

		Ductal			Lobular
		Gd1&2	Gd1&3	Gd2&3	Gd2&3
30 best	% misclassification	0%	0%	1.5%	0.9%
	Mann Whitney U	p=0.033	p=0.029	p=0.633	p<0.001
COM	% misclassification	0%	0%	0%	0%
	Mann Whitney U	p<0.001	p<0.001	p=0.108	p<0.001
RLM	% misclassification	7.4%	7.9%	18.9%	4.5%
	Mann Whitney U	p=0.481	p=0.496	p=0.906	p=0.589

Table 1- Summary of data classification and statistical testing carried out on different grades of lobular and ductal carcinoma

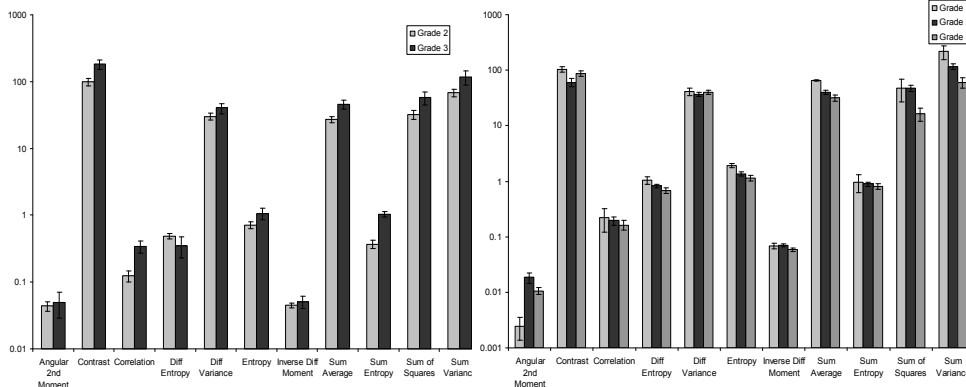


Figure 1- COM model features for different grades of lobular (left) and ductal (right) cancers

References:

- [1] Moy L *et al.* AJR (2009);193:986-993
- [2] Vassiou K *et al.* Eur. J. Radiol. (2009);70:69-76
- [3] Holli K *et al.* Acad. Radiol. (2010);17:135-141
- [4] Gibbs P *et al.* Magn. Reson. Med. (2003) ;50 :92-98
- [5] Tavassoli A & Devilee P. WHO Classification of Tumours. IARC Press. (2003)
- [6] Szczepinski PM *et al.* Comput. Meth. Prog. Bio. (2009);94:66-76
- [7] Hall M *et al.* SIGKDD Explorations (2009);11:10-18

infiltrative lobular carcinoma. Currently our grade 1 ductal cancer and grade 3 lobular cancer groups are sparsely populated and therefore an increase in patient numbers is considered essential; however, preliminary findings are promising for future studies in this area.

Conclusions: Our study presents some very preliminary results suggesting that texture analysis of breast MRI images could potentially be used as a non-invasive method of identifying cancer grade. Our results further suggest that the COM model alone could be sufficient to differentiate between grades, thereby waiving the need to calculate the full range of texture parameters from the different models. If confirmed, this simplification could make the technique even more clinically applicable.