Bootstrap Ensemble Neural Networks in MRI of the Breast

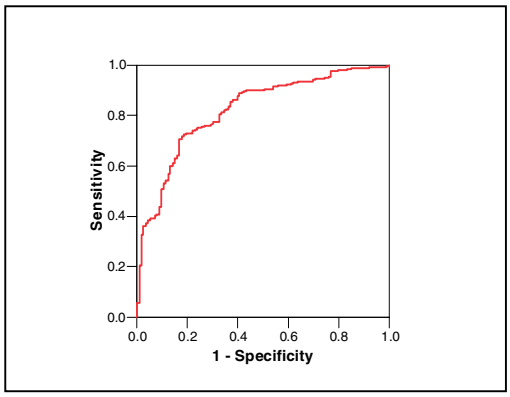
P. Gibbs[1], M. Sreenivas[1], A. Hunter[2], L. Turnbull[1]

[1]Centre for MR Investigations, University of Hull, Hull, United Kingdom, [2]Computing and Informatics, Lincoln University, Lincoln, United Kingdom

Introduction MR imaging of the breast is known to have a high sensitivity but the reported specificity has varied greatly due to a number of factors including overlapping lesion characteristics and the limited availability of experienced radiologists. Whilst the BI-RADS lexicon [1] is an attempt to standardise breast MR reporting it is still open to individual interpretation. Quantification of relevant parameters and subsequent analysis via techniques such as neural networks, are alternative more objective strategies. Pharmacokinetic modelling of contrast uptake, shape descriptors [2], and textural parameters [3] have all been utilised to aid lesion discrimination. Therefore, their combination within a neural network is desirable. However, robust neural network training frequently presents difficulties due to the number of cases required. Heuristically the requisite number of cases is related to the complexity of the network; the simplest suggestion states that there should be 10 times as many cases as connections in the network. Typically this would necessitate thousands of cases, clearly beyond an individual institutions scope. Consequently overlearning on a limited dataset often results, such that on presentation of new cases the neural network performs poorly. This work seeks to address this issue by utilising a bootstrap ensemble technique to combine a large number of neural networks to obtain results with greater confidence.

Methods Data from 311 patients (237 malignant and 113 benign lesions confirmed by histology or clinical follow-up) was retrospectively analysed. All images were acquired using a 1.5 T GE Signa Echo-speed scanner with the patient prone and the breasts suspended in a dedicated breast coil. Dynamic contrast-enhanced images were obtained using a fast spoiled gradient-recalled-echo (FSPGR) sequence (TR/TE 7.6/4.2 ms, flip angle 30°), at 35 time points with a temporal resolution of 11.6 s. High resolution, fat-suppressed, post-contrast images were then obtained, again using an FSPGR sequence (TR/TE 23-28/4.2 ms, flip angle 30°, field of view 20-36 cm, matrix size 512×256, slice thickness 3-5 mm, 1 average). After acquisition an experienced radiologist drew ROIs, on both the dynamic and post-contrast images, encompassing the whole lesion as closely as possible whilst excluding surrounding fat. From the dynamic data pharmacokinetic parameters were obtained using a 2-compartment model similar to that proposed by Brix *et al* [4] for the whole lesion (n=5) and the most enhancing 3×3 pixel square (n=5). Similarly, from the post-contrast images textural parameters (n=14) as defined by Haralick *et al* [5] and shape parameters (n=10), specifically circularity, elongatedness, complexity and Hu invariant moments [6], were calculated. Commercially available software (STATISTICA, StatSoft Inc.) was then used to construct numerous multi-layer perceptron networks with a single hidden layer via bootstrap resampling techniques. Finally, an ensemble of the 50 best performing networks was constructed.

Results All pharmacokinetic modelling derived parameters for both the whole lesion and the most enhancing 3×3 pixel square demonstrated significant differences between the two groups ($p < 0.0002$). Significant differences were also noted for circularity, elongatedness, complexity, and 3, out of 7, Hu invariant moments (see table). For the calculated textural parameters 6, out of 14, revealed significant differences between benign and malignant lesions (see table).



Textural Parameter	*P-value*	*Shape Parameter*	*P-value*
f_1-ASM	<0.001	ϕ_1	0.029
f_2-Contrast	0.686	ϕ_2	0.003
f_3-Correlation	0.717	ϕ_3	0.386
f_4-Variance	0.574	ϕ_4	0.253
f_5-IDM	0.357	ϕ_5	0.995
f_6-Sum Average	0.029	ϕ_6	0.398
f_7-Sum Variance	0.795	ϕ_7	0.037
f_8-Sum Entropy	0.001	Circularity	0.010
f_9-Entropy	<0.001	Elongatedness	0.007
f_{10}-Difference Variance	0.970	Complexity	0.043
f_{11}-Difference Entropy	0.927		
f_{12}-Correlation Measure 1	<0.001		
f_{13}-Correlation Measure 2	<0.001		
f_{14}-Max. Correlation Coeff.	0.172		

On completion of neural network training a 50 network ensemble with performance figures of 79% and 70% for benign and malignant lesions respectively was established. These figures are estimated to be accurate to within ±4% at the 95% confidence level. The overall diagnostic accuracy was calculated to be 0.82±0.02 (see ROC curve illustrated above).

Discussion Using ensemble techniques a robust neural network based discriminator has been developed. Whilst the results obtained preclude neural networks as an exclusive method of lesion discrimination they may find application as a 1st reader, thus only referring the more complicated cases to an experienced radiologist. The incorporation of other important radiological features, including quantification of the degree of architectural distortion of surrounding tissue, may improve results further.

[1] Breast imaging reporting and data system (BI-RADS), American College of Radiology, 1998. [2] P Gibbs et al (2005) Proceedings of the 13th ISMRM Annual Meeting 1870. [3] P Gibbs and LW Turnbull (2003) *Magn. Reson. Med.* 50:92-98. [4] G Brix *et al* (1991) *J Comput. Assist. Tomogr.* 15:6221-628. [5] RM Haralick *et al* (1973) *IEEE Trans. Sys. Man. Cyb.* 3:610-621. [6] MK Hu (1962) *IRE Trans. Info. Theory* IT-8:179-187.