

# Discrimination between glioblastoma multiforme and solitary metastasis using morphological features derived from the *p:q* tensor decomposition of diffusion tensor imaging

Guang Yang\*, Timothy L. Jones, Thomas R. Barrick and Franklyn A. Howe

**The management and treatment of high-grade glioblastoma multiforme (GBM) and solitary metastasis (MET) are very different and influence the prognosis and subsequent clinical outcomes. In the case of a solitary MET, diagnosis using conventional radiology can be equivocal. Currently, a definitive diagnosis is based on histopathological analysis on a biopsy sample. Here, we present a computerised decision support framework for discrimination between GBM and solitary MET using MRI, which includes: (i) a semi-automatic segmentation method based on diffusion tensor imaging; (ii) two-dimensional morphological feature extraction and selection; and (iii) a pattern recognition module for automated tumour classification. Ground truth was provided by histopathological analysis from pre-treatment stereotactic biopsy or at surgical resection. Our two-dimensional morphological analysis outperforms previous methods with high cross-validation accuracy of 97.9% and area under the receiver operating characteristic curve of 0.975 using a neural networks-based classifier. Copyright © 2014 John Wiley & Sons, Ltd.**

**Keywords:** brain tumour classification; brain tumour segmentation; computer-aided diagnosis; feature selection; pattern recognition and classification; morphological shape analysis; MRI; diffusion tensor imaging

## INTRODUCTION

The radiological appearance of solitary necrotic intracranial metastases and high-grade gliomas, e.g. glioblastoma multiforme (GBM), is often similar (1). However, clinical management of patients with these tumours is significantly different, such that misdiagnosis or delayed treatment can affect the prognosis and subsequent clinical outcomes (2). GBMs are the most common and lethal primary intracranial tumours (3). GBM typically results in a patient's death within 2 years from the initial diagnosis because of its strong resistance to conventional therapies, including surgery, chemotherapy and radiation (4,5). Recent population-based analysis has reported a median survival time of 12.6 months (6). Even with advances in surgery and oncology, the prognosis of GBMs remains poor.

Intracranial metastasis (MET) is another common brain neoplasm in adults. For MET to occur, cancer cells are released or break off from their primary site and migrate to the central nervous system, where there can be significant growth before clinical symptoms arise (7). For brain METs, a recent study has reported that the median survival time can be up to 10.7 months with whole-brain radiation therapy (8).

Currently, the definitive diagnosis of GBM or MET is based on the histopathological analysis of a tissue sample obtained by stereotactic biopsy pre-treatment or during surgical resection. However, the differentiation of these neoplasms non-invasively is beneficial when the mass is located near an eloquent area, or when a patient's advanced age is unsuitable for the biopsy procedure; hence, various neuroradiological methods have been investigated for improved discrimination of these tumour types (9–16). For some cases, GBM

and MET can be differentiated using conventional anatomical MRI. Nevertheless, these tumours may pose a diagnostic dilemma on conventional MRI and a biopsy is often performed for histopathological confirmation, even when the history of a known primary malignancy is available (10,11). Advanced MRI techniques, such as MR Spectroscopy (MRS), perfusion MR and diffusion tensor imaging (DTI), have been reported to be of value in differential diagnosis. However, the varied appearance by MRI and the heterogeneous tumour growth patterns lead to very high accuracy for the differentiation of some tumour types, e.g. glioma *versus* meningioma, but much lower classification rates [60–85% accuracy or <0.8 average area under the receiver operating characteristic (ROC) curve (AUC)] reported in distinguishing between high-grade GBM and metastatic tumours (1,12–16).

\* Correspondence to: G. Yang, Room 31b, Neuroscience Research Centre, Cardiovascular and Cell Sciences Institute, St. George's, University of London, Cranmer Terrace, London SW17 0RE, UK.  
E-mail: gyang@sgul.ac.uk

G. Yang, T. L. Jones, T. R. Barrick, F. A. Howe  
Neuroscience Research Centre, Cardiovascular and Cell Sciences Institute, St. George's, University of London, London, UK

**Abbreviations used:** AUC, area under the ROC curve; BER, balanced error rate; CV, cross-validation; D-SEG, diffusion segmentation; DTI, diffusion tensor imaging; GBM, glioblastoma multiforme; kNN, k-nearest neighbour algorithm; LOO, leave-one-out; MET, metastasis; NB, naïve Bayes; NNW, neural networks; QDA, quadratic discriminant analysis; RBF, radial basis function; ROC, receiver operating characteristic; ROI, region of interest; SVM, support vector machine.

GBMs possess aggressive proliferation and their invasiveness can affect the morphological appearance, e.g. the shape of the tumours, as observed by MRI. In terms of the tissue microenvironment, when proliferation rates are higher, for example in the case of GBM, sub-tumours may emanate from the main tumour, separating off from the tumour as clusters of cells (17). In addition, tumour growth tends to follow particular directions and depends on fluctuations in local oxygen and nutrient supply at a macroscopic scale (17,18). GBMs generally grow along white matter tracts, with less common cases growing along the edge of cerebrospinal fluid-filled spaces or meninges (19). Consequently, their resultant growth is irregular and it has been proposed that their shape is more likely to be complex (20). In contrast, METs are anticipated to expand more homogeneously when proliferation rates are relatively low and cell adhesion is sufficient to maintain a compact tumour shape, potentially displacing rather than infiltrating white matter tracts (17). Thus, the global shape of METs is expected to be more spherical.

Therefore, we hypothesise that parameters that describe tumour morphology and its irregularities can be adopted to differentiate GBM from solitary MET. The semi-automatic diffusion segmentation (D-SEG) method derived from DTI data was used to extract the tumour core shape (21). Semi-automated segmentation allows a more detailed delineation of the tumour perimeter than manual drawing, and may improve the quantitative assessment of the morphological aspects of invasive *versus* encapsulated tumour growth.

## MATERIALS AND METHODS

### Data acquisition

#### Scanner A (2005–2006)

MR data were acquired on a GE Signa LX 1.5 T MRI system (GE Healthcare, Milwaukee, WI, USA) equipped with a maximum field gradient strength of 22 mT/m and using a quadrature head coil. Axial DTI datasets were obtained using a single-shot, spin echo, echo planar imaging sequence, that is one acquisition without diffusion sensitisation, i.e.  $b=0$  s/mm<sup>2</sup>, and 12 acquisitions with diffusion weighting, i.e.  $b=1000$  s/mm<sup>2</sup>, in 12 diffusion gradient directions. Whole-brain coverage was provided by 50 contiguous slices that were acquired as two interleaved series of four repeats (2.8 mm thick slices with 2.8 mm slice gap; TR/TE=8000/88 ms; acquisition matrix, 96 × 96; field of view, 24 cm<sup>2</sup>; in-plane resolution, 2.5 mm<sup>2</sup>).

#### Scanner B (2008–2010)

MR data were acquired on a GE Signa HD 1.5 T MRI system (GE Healthcare) equipped with gradients of 33 mT/m and using an eight-channel head coil. Axial DTI datasets were obtained using a similar echo planar imaging sequence as aforementioned, but with 61 diffusion gradient directions. Whole-brain coverage was provided by 55 contiguous slices (TR/TE = 14,150/93.6 ms; acquisition matrix, 96 × 96; field of view, 24 cm<sup>2</sup>; isotropic voxel resolution, 2.5 mm<sup>3</sup>). Parallel imaging using the Array Spatial Sensitivity Encoding Technique was applied to accelerate image acquisitions, and 10 volumes of  $b=0$  s/mm<sup>2</sup> scans were acquired and averaged.

### Patients

In agreement with the local regional ethics committee, 48 patients were recruited in two blocks between 2005–2006 (11 GBM and 9 MET) and 2008–2010 (19 GBM and 9 MET), and retrospectively entered into this study. No patients had been subject to any treatment prior to radiological examination and none had a previous history of surgery, chemotherapy or radiation therapy applied to intracranial structures. Histopathological analysis of resected tissue confirmed the diagnosis of GBM in 30 patients. For the 18 patients with MET, histopathological analysis revealed six cases that originated from the lung, two cases from the breast, two from squamous cells, two from bowel adenocarcinoma, one from melanoma, one from renal cancer, one neuroendocrine case, one from prostate cancer and two from an unknown source.

### DTI segmentation

DTI allows the quantitative assessment of isotropic ( $p$ ) and anisotropic ( $q$ ) tissue water diffusion components. Disruption of brain tissue microstructure by tumour growth and infiltration alters tissue water diffusion characteristics, and a tumour region can be delineated from the abnormal regions apparent on parametric  $p$  and  $q$  maps. To minimise observer subjectivity in defining the tumour region, Jones *et al.* (21,22) developed the D-SEG method which combines  $p$  and  $q$  diffusion data to automatically segment brain images into regions of similar diffusion characteristics. Essentially, D-SEG applies a  $k$ -medians clustering algorithm to whole-brain DTI datasets from patients with tumours, and iteratively segments the two-dimensional ( $p, q$ ) space into  $\hat{k}$  non-overlapping clusters.  $\hat{k} = 16$  initial centroids have been found to be appropriate for brain tumour data to enable the segmentation of the range of potential tissue compartments present within healthy and tumour-affected brain, e.g. normal-appearing white matter, normal-appearing grey matter, cerebrospinal fluid spaces, solid tumour, regional tumour necrosis, tumour-associated cystic regions, perilesional oedema, perilesional tumour infiltrations and distant oedema. It should be noted that multiple clusters are needed to represent grey and white matter because of the heterogeneity of tissue diffusion characteristics throughout normal brain. The ( $p, q$ ) space was initially clustered into 16 subsets by separating the two-dimensional  $p$  and  $q$  data distribution of the full dataset into quartiles. Initial cluster centroids were determined as the median coordinates in ( $p, q$ ) space for each cluster. Next, the distance was calculated from each voxel to each cluster centroid in ( $p, q$ ) space. Each voxel was then assigned to its nearest cluster in this space, and cluster centroids (i.e. medians) were recalculated based on the new data within these clusters. This procedure was iteratively repeated until no voxel changed its cluster, the centroids remained the same or a defined iterative limit (in this case 250 iterations) was reached, which is expressed as:

$$\min_{j \in \{1, 2, \dots, 16\}} \sqrt{(p_i - m_{pj})^2 + (q_i - m_{qj})^2} \quad [1]$$

where ( $p_i, q_i$ ) represents the voxel in ( $p, q$ ) space and ( $m_{pj}, m_{qj}$ ) denotes its nearest cluster median. Each  $p$ – $q$  segment was colour coded and mapped back onto the original image to provide visual demarcation of the tumour region and surrounding oedema. A flood-fill method was then employed by placing seed voxels inside regions characterised as tumour, as defined by particular  $p$ – $q$  segments, to semi-automatically generate regions of interest (ROI) of the tumour on a slice-by-slice basis. Binary masks of the tumour core region, referred to as the tumour ROI, were then computed.

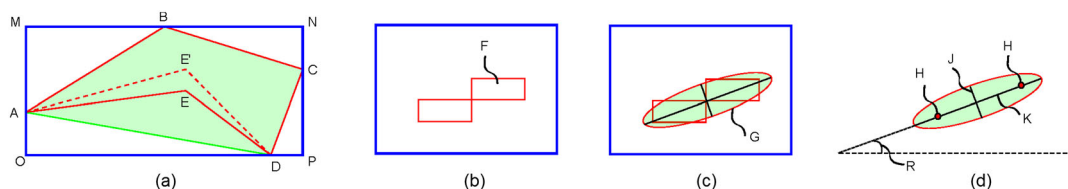
# Morphological feature extraction and selection

Once the tumour ROI had been established using D-SEG, we used in-house software to compute two-dimensional shape features over each imaging slice of each tumour: area, rectangular bounding box area, convex area, eccentricity, equivalent diameter, perimeter, extent, solidity, ellipse major and minor axis lengths and their orientation. The chosen features are easily computed and were chosen as they potentially distinguish whether the tumour shape is more like a solid ellipsoid (as expected for the encapsulated growth of metastases) or more irregular with convex and concave features and projections (as expected for infiltrative growth of glioblastomas). We then computed the median, upper and lower quartiles and standard deviation of these shape parameters over all tumour slices as measures of the irregularity of the shape of each tumour. Our hypothesis is that an optimal combination of these parameters may be used to aid the discrimination of METs from GBMs.

The two-dimensional shape measures for the tumour ROI in each image slice are defined as follows.

- (1) 'Area': number of voxels in the ROI (e.g. area ABCDE in Figure 1a).
- (2) 'Rectangular bounding box area': smallest rectangle containing the ROI (e.g. area MNPO in Figure 1a).
- (3) 'Convex area': number of voxels in the convex image; all voxels within the hull are set to a value of unity (e.g. area ABCD in Figure 1a). This allows the possibility of identifying differences between ROIs ABCDE and ABCDE' (see Figure 1) that have different areas but identical rectangular bounding box areas and convex areas.
- (4) 'Eccentricity': eccentricity of the ellipse with the same second moments as the ROI (Figure 1b–d). Here ROI F is shown (Figure 1b) with the ellipse G that has the same second moments (Figure 1c). The major axis K, minor axis J and foci H of the ellipse are shown in Figure 1d. The eccentricity of the ellipse is defined as the ratio of the foci to its major axis length, and has values between zero and unity. Two degenerate cases occur for values of zero and unity which represent a circle and line segment, respectively.
- (5) 'Equivalent diameter': diameter of a circle with the same area as the ROI.
- (6) 'Perimeter': distance around the ROI boundary.
- (7) 'Extent': ratio of ROI area to that in the rectangular bounding box.
- (8) 'Solidity': proportion of voxels in the convex hull that are within the ROI.
- (9) 'Major axis length' of the ellipse.
- (10) 'Minor axis length' of the ellipse.
- (11) 'Ellipse orientation': angle (from  $-90^\circ$  to  $90^\circ$ ) between the image horizontal axis and major axis of the ellipse, as shown by R in Figure 1d. As the ellipse angle of orientation is arbitrary, the standard deviation alone is used for this shape feature.

All the two-dimensional shape features are normalised to zero mean and unity standard deviation.



**Figure 1.** (a) Illustration of the two-dimensional shape features of area, rectangular bounding box area and convex area. (b) Exemplary region of interest F. (c) Ellipse G with the same second moments as ROI F. (d) Illustration of the two-dimensional shape features of eccentricity, equivalent diameter, major axis and minor axis length, and ellipse orientation.

To determine the shape parameters most useful for our classification task, we applied a two-stage feature selection. First, Student's *t*-test (23) was used to determine shape parameters with the greatest difference between GBMs and METs. Features exhibiting a significance level of  $p < 0.00082$  [i.e. Bonferroni correction (24) for 61 parameters] were retained. Second, we investigated which of these features were highly correlated, and redundant features were removed. According to previous research (25), a Pearson's linear correlation coefficient of  $|r| < 0.5$  represents weak correlation and  $|r| > 0.8$  represents strong correlation. A moderate threshold of  $|r| > 0.7$  was applied in this study to define a significant correlation between features. The shape parameter with the lowest *p* value from Student's *t*-test was selected and other parameters correlating with this feature ( $|r| > 0.7$ ) were then removed. The shape parameter with the next lowest *p* value was then selected and the process was iterated, after which only five significantly different features remained.

# Tumour classification

We adopted and compared five of the most widely used supervised learning strategies to classify GBMs and solitary METs based on the selected two-dimensional shape features. The classification methods used were quadratic discriminant analysis (QDA) (26), naïve Bayes (NB) (27), *k*-nearest neighbour algorithm (*k*NN) (28), non-linear support vector machine (SVM) (29), and neural networks (NNW) (30).

QDA is a generative classification model based on the multivariate Gaussian that is similar to linear discriminant analysis, and constructs a statistical framework to calculate the likelihood of target occurrence in future trials given occurrences in previous trials (26). However, QDA incorporates different covariance matrices for each class and hence may provide a better description of individual populations (31).

NB estimates the probability of an object belonging to a specific class when observed feature values of an object and previous class are available (27). The predicted class represents the highest estimated probability assuming that the features are conditionally independent given the class label (31). The model is 'naïve' as features are not expected to be conditionally independent; however, even if the underlying assumptions of NB are not true, the technique often results in accurate classification as the model is simple and relatively immune to overfitting (32).

*k*NN is a non-parametric classifier separating objects based on the nearest training examples in a feature space (28). It is a memory- or instance-based learning method using a distance or similarity function. In this case, the Euclidean distance was applied. *k*NN classifies an object by estimating an empirical fraction (i.e. a majority vote) of its neighbours, with the object being assigned to the most frequent class among its *k*-nearest neighbours (31). For our dataset, we found the optimal number of *k* according to the best cross-validation performance.



SVM is a powerful technique for binary classification (29). Predictions of SVM only depend on a subset of training data (i.e. the support vectors), and SVM finds the hyperplane with the largest margin between the two classes (31). Here, we use a Gaussian radial basis function (RBF) kernel, with scaling factor  $\sigma = 1$ , to map feature vectors into a non-linear feature space where an optimal hyperplane is constructed separating GBM and MET classes.

NNW consist of simple elements operating in parallel and are inspired by the geometry of biological nervous systems (30). NNW are modelled as a weighted graph of connected neurons (33), and the output of each neuron is determined using a transfer function. The logistic function was used as the transfer function and defines the multilayer perceptron network architecture. A series of logistic regression models are stacked on top of each other to form the hidden layers with the final output layer being another logistic regression model to create a feed-forward NNW-based classifier, i.e. the multilayer perceptron classifier (31). In addition, using cross-validation, we determined the optimal number of hidden neurons in the hidden layers ( $h$ ) according to the best cross-validation accuracy.

### Performance measure

The classification performance of our morphological shape analysis framework was evaluated by: (i) leave-one-out (LOO) cross-validation at the patient level after feature selection; (ii) classification accuracy, sensitivity and specificity; and (iii) the balanced error rate (BER) (34). LOO cross-validation was used to provide an unbiased predictor, as our dataset is relatively small for the creation of separate training and test sets as required for  $n$ -fold methods (35). Results obtained for the morphological shape analysis framework were compared with those obtained for the mean extent only (the shape information derived from the bounding box ratio), which has been used in a previous study (36).

## RESULTS

Figure 2 shows example GBMs (Figure 2a–j) and METs (Figure 2k–t). The tumour core was clearly observed in the  $q$  maps (as areas of low anisotropic diffusion indicated by red arrows in Figure 2c, h, m, r), and the surrounding oedema in  $p$  maps (as indicated by the high isotropic diffusion indicated by green arrows in Figure 2b, g, l, q). Figure 2d, i, n, s illustrates whole-brain segmentations generated by D-SEG. By application of the semi-automated flood-fill algorithm, tumour ROIs are segmented as shown in Figure 2e, j, o, t. These were used as binary masks for further analysis. In general, GBMs have more heterogeneous tumour ROI boundaries, e.g. Figure 2a–e shows a GBM with more infiltrative features invading from the right frontal lobe anterior to the Sylvian fissure with surrounding oedema causing compression of the ventricles. In contrast, Figure 2f–j shows a GBM with a more compact tumour core and smaller size near the posterior limb of the internal capsule. The tumour ROIs for METs exhibited a more typically compact pattern. Figure 2k–o shows a MET in the right parietal lobe area and Figure 2p–t shows a MET near the anterior limb of the internal capsule with a large region of oedema across a large region of the right cerebral hemisphere. Although oedema areas may vary, tumour cores of METs are more localised. Figure 3 shows that, although the solitary METs are generally more compact and localised, the quantified two-dimensional shape parameters for area, rectangular bounding box, convex area, equivalent diameter, major and minor axis lengths, and ellipse orientation have a large overlap between tumour types. This indicates that the overall size and orientation of GBMs and solitary METs cannot identify tumour type, as may have been expected, but which were tested for completeness.

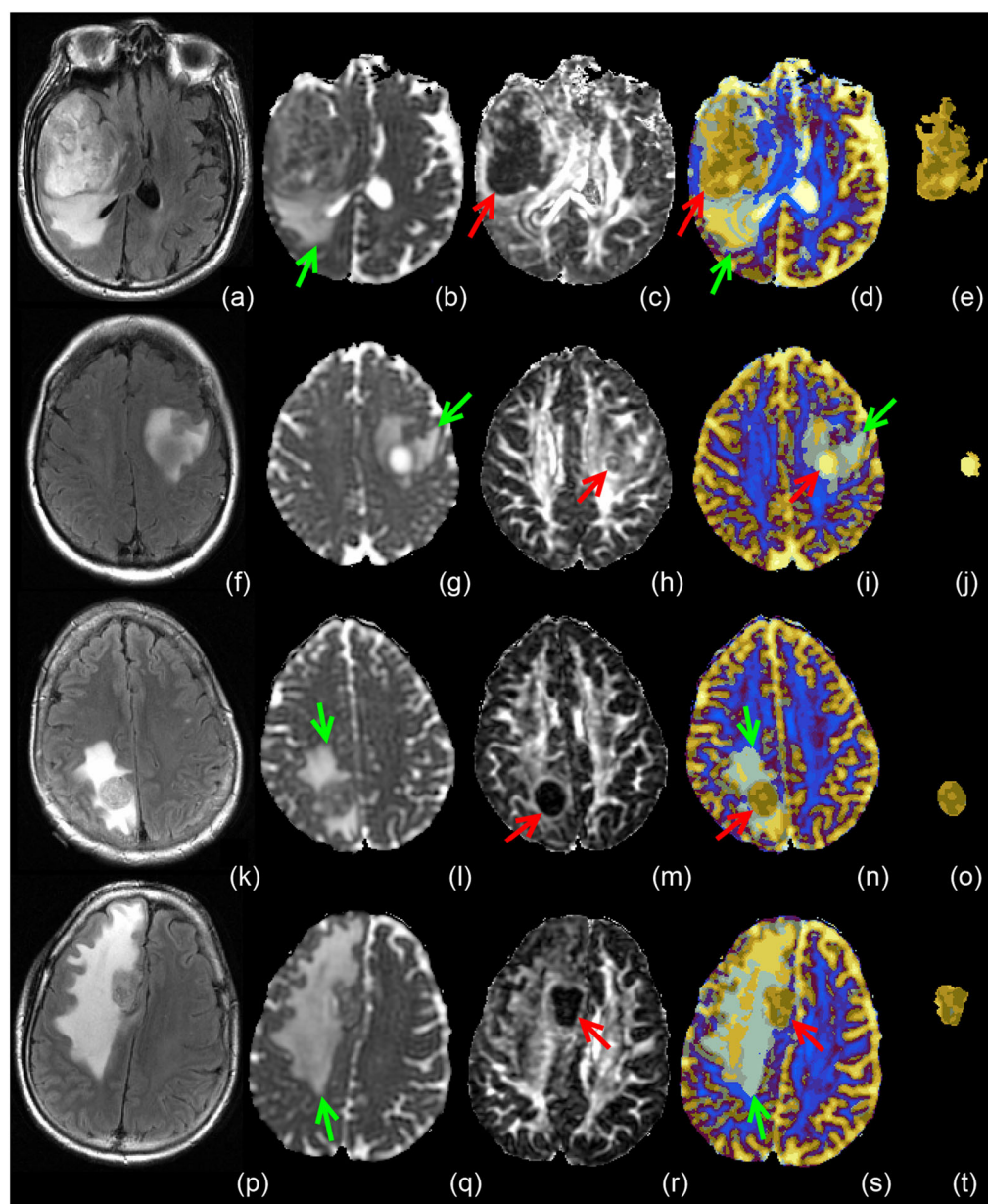
Feature selection identified 33 significant two-dimensional shape features after application of Student's  $t$ -test. Five features remained after removal of highly correlated redundant features. The features selected for pattern recognition analysis were: maximum eccentricity (f19), maximum perimeter (f31), lower quartile (f39) and standard deviation of extent (f42) and maximum solidity (f43), as shown in Figure 3. Student's  $t$ -test and correlation coefficients for the five two-dimensional shape parameters derived by the feature selection process are shown in Table 1. The maximum correlation coefficient of  $-0.6$  was identified between the lower quartile and standard deviation of the extent measure for the GBM data, but, on average over all five parameters, the correlation coefficients were low ( $r = 0.26$ ).

The accuracy, sensitivity, specificity, AUC and BER of the cross-validation (LOO) are summarised in Table 2 with comparison with the results obtained using only the mean extent. For the five selected features, QDA obtained 88% accuracy and an AUC of 0.969, but provided a sensitivity (83%) much lower than the specificity (90%). NB and kNN (with  $k = 10$  according to the best cross-validation accuracy as seen in Figure 4a) obtained the same accuracy (94%), sensitivity (94%) and specificity (93%); however, NB had a slightly higher AUC in comparison with kNN (NB 0.985 versus kNN 0.983). SVM achieved comparable results to NB and kNN with lower specificity (90%) and accuracy (92%). With respect to cross-validation accuracy and BER, NNW outperformed all classifiers and provided 98% accuracy with 100% sensitivity and 97% specificity, with an AUC of 0.975, which is slightly lower than the AUC obtained by NB and kNN. These NNW results were achieved using  $h = 9$  hidden neurons in the hidden layers that were also determined according to the best cross-validation accuracy (Figure 4b). Using only the mean of the extent feature, a maximum accuracy of 92% for the NNW classifier with an AUC of 0.965 and BER of 0.08 achieved, but did not outperform, the NNW classifier that used the five selected shape features (BER of 0.02). As aforementioned, in order to obtain the best parameter settings for our kNN and NNW classifiers, we plotted the cross-validation error with respect to various  $k$  and  $h$ . For the kNN classifier, we tested  $k=1-24$  as higher  $k$  values generate a cluster size greater than half the total number of data. Similarly, we tested  $h = 1-24$  for the NNW classifier. Figure 4a shows that, for both  $k = 10$  and  $k = 11$  we obtained the lowest cross-validation error of 6.25%. Figure 4b illustrates that various  $h$  can achieve an equally low cross-validation error of 2% for this dataset.

We finally compared our two-dimensional shape classification with the simple bounding box analysis using a non-parametric statistical test, i.e. analysis of variance by ranks, for the cross-validated accuracy, AUC and BER results. We found that the accuracy and BER were significantly lower for the two-dimensional shape analysis compared with the use of the mean extent only, with  $p = 0.025$ , and AUC was also lower for our two-dimensional shape analysis with  $p = 0.046$ . Overall, we advocate the BER metric as the main criterion for comparison of classifiers because of the unequal numbers of GBM and MET samples.

## DISCUSSION

The aim of this study was to provide high classification accuracy of GBMs and solitary METs using an objective pattern recognition method that exploits tumour morphological structure. By application of a novel segmentation method (D-SEG) to identify tumour core ROIs, we identified a variety of two-dimensional

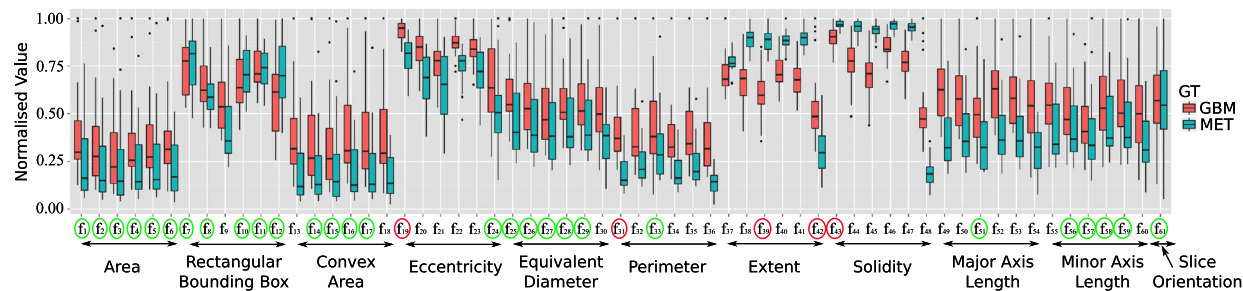


**Figure 2.** (a)–(j) Two cases of glioblastoma multiforme (GBM). (k)–(t) Two cases of metastasis (MET). (a, f, k, p) Fluid-attenuated inversion recovery images. (b, g, l, q)  $p$  maps. (c, h, m, r)  $q$  maps. (d, i, n, s) Whole-brain diffusion segmentation (D-SEG). (e, j, o, t) Flood-filled tumour core region of interest (ROI) segmentation (red arrows) distinguished from surrounding oedema area (green arrows). Tumour ROIs are created from the active (non-zero) voxels in the tumour core segmentation (e, j, o, t).

shape parameters, and determined the best combination of features for tumour classification. Shape parameters relating to tumour size did not discriminate tumour type. For example, Figure 3 shows that the quantified parameters extracted from area, rectangular bounding box, convex area, equivalent diameter, major and minor axis lengths, and slice orientation have a large overlap between GBMs and solitary METs. Although solitary METs are generally more compact and localised, the size (volume) of the GBMs varies widely (Figure 2a–e *versus* Figure 2f–j). In addition, the large variation in ellipse orientation indicates that this parameter is not likely to be related to an infiltrative growth characteristic along a particular direction. In contrast, eccentricity, perimeter, extent and solidity were identified as the most powerful feature discriminators. For instance, maximum perimeter is a major discriminant, indicating that, for a similar range of areas, GBMs have a greater

perimeter than METs, which may be attributed to the infiltrative growth pattern of GBMs producing projective features, when compared with the more encapsulated growth of METs. Similarly, maximum eccentricity, maximum solidity and the lower quartile and standard deviation of extent are also major discriminants. This indicates that GBMs have high cross-sectional variation, when approximated by an ellipsoid, and exhibit greater variation in concave or convex shape when compared with METs. In summary the shape parameters that aid discrimination are those characterising the heterogeneous, infiltrative growth of GBMs when compared with the more ellipsoidal encapsulated growth of METs.

Feature selection for tumour classification was provided using a two-step method based on Student's  $t$ -test and correlation analyses. The use of such a method is supported by a recent study on biomarker discovery that compared 32 feature selection



**Figure 3.** Boxplot of the sixty-one two-dimensional shape features for 30 glioblastomas (GBM) (red) and 18 metastases (MET) (blue) including maximum, median, lower and upper quartiles, mean and standard deviation of the area (f1–f6), rectangular bounding box (f7–f12), convex area (f13–f18), eccentricity (f19–f24), equivalent diameter (f25–f30), perimeter (f31–f36), extent (f37–f42), solidity (f43–f48), major axis length (f49–f54) and minor axis length (f55–f60), and the standard deviation of the ellipse orientation (f61). Green circles highlight the 28 features filtered out by Student’s *t*-test during the first feature selection step. Red circles highlight the five selected features according to Student’s *t*-test and correlation analysis results. These are the maximum eccentricity (f19), maximum perimeter (f31), lower quartile of extent (f39), standard deviation of extent (f42) and maximum solidity (f43).

**Table 1.** Feature selection according to a two-sample Student’s *t*-test and correlation analysis. *p* values and correlation coefficients are shown for the five selected two-dimensional shape features

	Shape features	Lower quartile of extent (lqExt)	Maximum eccentricity (maxEcc)	Maximum perimeter (maxP)	Standard deviation of extent (sdExt)	Maximum solidity (maxSol)
Student’s <i>t</i> -test <i>p</i> values (GBM & MET)	–	3.87E-11	4.21E-07	3.06E-06	1.38E-05	1.58E-05
Correlation coefficients (GBM)	lqExt	1	–	–	–	–
	maxEcc	–0.43	1	–	–	–
	maxP	0.13	–0.13	1	–	–
	sdExt	–0.60	0.46	–0.24	1	–
	maxSol	0.08	0.08	–0.07	0.55	1
Correlation coefficients (MET)	lqExt	1	–	–	–	–
	maxEcc	–0.41	1	–	–	–
	maxP	–0.27	0.30	1	–	–
	sdExt	–0.06	0.14	–0.17	1	–
	maxSol	0.57	–0.13	–0.04	1	–

GBM, glioblastoma multiforme; MET, metastasis.

strategies and found that feature identification using Student’s *t*-test provided the best classification results (23).

Our combination of feature selection followed by classification consistently outperformed classification based on the single bounding box ratio (mean extent) feature alone. Furthermore, the NNW classifier achieved the greatest accuracy in our study and achieved a greater accuracy than previously published results for tumour classification using the bounding box ratio and simple threshold of the feature space (36) (93.9% *versus* 97.9% using our method). Furthermore, the current study was applied to a larger dataset than previously used (36) (33 patients *versus* 48 patients in our study). We only achieved 92% accuracy when tumours were classified using NNW on the single bounding box ratio feature alone, which further validates the effectiveness of our selected shape features. In addition, we also investigated five of the most widely used classifiers for comparison. In general, as the feature selection procedure filtered out the least discriminating two-dimensional shape features, all classifiers produced reasonably high cross-validation accuracy.

Most classification methods require parameter tuning before obtaining the best performance. For example, the number of nearest neighbours *k* in *k*NN, the scaling factor of the Gaussian

RBF kernel in SVM, and the number of hidden neurons *h* in the hidden layers used in NNW. In this study, we optimised *k* and *h* according to the cross-validation performance (Figure 4). *k*NN is a local non-linear method as only a limited number of nearest neighbours are used and, as the number of *k* increases, the model becomes less local (37). The accuracy of the *k*NN method can be severely degraded in the presence of redundant, irrelevant, interacting or noisy features (38). Therefore, the choice of too many neighbours would incorporate noisy data that are irrelevant to the query object, and thus the classifier would be under-fitted. In contrast, the model would be over-fitted if too few neighbours were selected, as the prediction would be based on insufficient data points. Consequently, we can observe that, at *k* = 1, 2 or 3, the cross-validation error is relatively larger than at *k* = 10 or 11, where we obtained the best performance of *k*NN. Similarly, we performed the cross-validation for NNW with a set of *h*, and the results (Figure 4b) suggest that various *h* perform equally well and the error curve oscillated between 2.08% and 14.58% (excluding the value when *h* = 1). This may be attributed to the fact that NNW is robust when applied to our data and the average cross-validation error is 5.5%. However, no elevated error was observed when *h* was large, and this might be because



**Table 2.** Quantitative classification (LOO cross-validation) results for the five selected two-dimensional shape features compared with classification (LOO cross-validation) results using the mean extent feature only

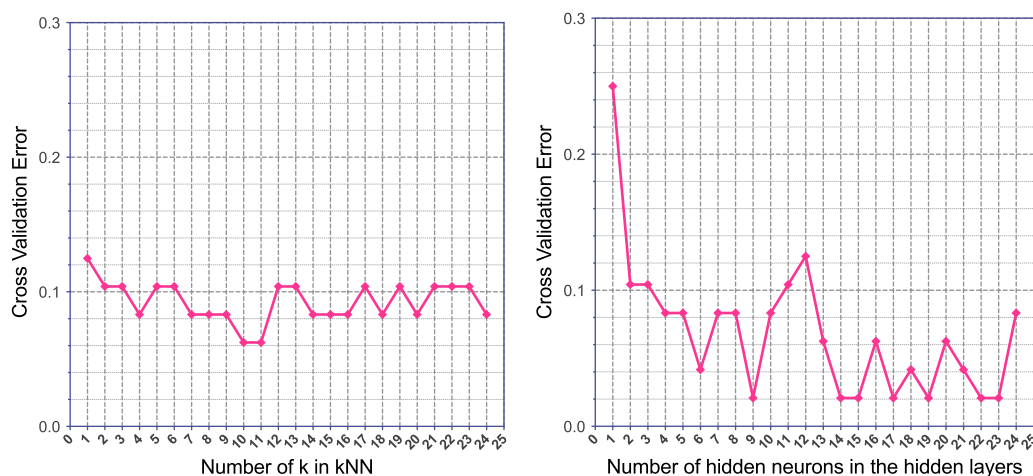
Results for selected two-dimensional shape features					
Classifiers	Accuracy	Sensitivity	Specificity	AUC	BER
QDA	87.5	83.33	90	0.969	0.13
NB	93.75	94.44	93.33	0.985	0.06
kNN	93.75	94.44	93.33	0.983	0.06
SVM	91.67	94.44	90	0.959	0.08
NNW	97.92	100	96.55	0.975	0.02
Results for mean extent feature only					
Classifiers	Accuracy	Sensitivity	Specificity	AUC	BER
QDA	85.42	83.33	86.67	0.954	0.15
NB	89.58	94.44	86.67	0.944	0.09
kNN	85.42	94.44	80	0.954	0.13
SVM	87.5	100	80	0.959	0.1
NNW	91.67	87.5	95.83	0.965	0.08

AUC, area under the receiver operating characteristic curve; BER, balanced error rate; kNN, *k*-nearest neighbour algorithm; NB, naïve Bayes; NNW, neural networks; QDA, quadratic discriminant analysis; SVM, support vector machine.

the model is still not over-fitted. As  $h$  becomes larger, both the computational burden of the model and memory usage increase; therefore, we chose  $h = 9$  to ensure the use of the simplest model with the best classification performance. There are many variations of SVM, e.g. we can plug in linear, quadratic, Gaussian RBF and multilayer perception kernels, and model selection is beyond the scope of this study. In order to make our work available to compare with other research, we used a standard RBF kernel with a default scaling factor of  $\sigma = 1$  for the SVM classifier. However, we can envisage a further boost of the classification performance by the incorporation of ensemble learning methods (39) into the current morphological learning framework.

Other MRI classification studies have reported lower accuracies than those reported in the current study for discrimination between GBMs and METs: short-echo  $^1\text{H}$  MRS with SVM has demonstrated AUCs of 0.607 (12) and 0.884 (1); 71% classification accuracy was achieved using texture analysis on  $T_1$ -weighted gadolinium-enhanced MR images (40); 85% accuracy was achieved using DTI metrics (14). Many of the methods used in previous studies included a more subjective manual definition of the tumour ROI than that provided by our semi-automated D-SEG procedure, potentially limiting comparisons between studies. For example, in a recent study, the tumour ROI was visually located by the human operator in each two-dimensional axial plane ( $T_1$ ,  $T_2$  and  $T_1$  contrast-enhanced images were used) before application of a segmentation process (36). Furthermore, when using  $^1\text{H}$  MRS, there is significant operator input in planning the voxel size and location for single-voxel MRS or in the selection of appropriate voxels from a multi-voxel dataset of the scan.  $^1\text{H}$  MRS is also technically more challenging (e.g. there is a need to obtain good water suppression and magnetic field homogeneity across the voxel) and may not adequately encompass the whole tumour region. In contrast, our D-SEG method requires minimal user intervention and is based on a whole-brain segmentation of DTI data.

A limitation of our proposed decision support system is that, currently, the segmentation is semi-automated, i.e. it includes manual placement of seed voxels inside the tumour ROI that are used for flood filling to identify the tumour core. The flood-filling method is reproducible across multiple users if there is common agreement on the  $p, q$  segments selected during tumour ROI extraction. We also assume that the D-SEG map defined an accurate representation of the border between tumour and unaffected healthy tissue. Further studies to include histopathological analysis at the edges between the tumour core, oedema and apparently healthy tissue are required to validate this segmentation technique. In addition, further work is required to provide a more fully automated selection of the segmented regions that are assigned to the tumour core ROI. The current method uses two-dimensional shape features; hence, for a given tumour, the derived parameters will depend on the orientation of the image slice through the tumour. However, the image slice thickness in this study is greater than the in-plane resolution and is not ideal for three-dimensional shape analysis. Consequently, future work is required to investigate tumour classification accuracy between two- and three-dimensional shape



**Figure 4.** (a) Plot of various  $k$  used in the  $k$ -nearest neighbour algorithm (kNN) with respect to the cross-validation errors. (b) Plot of various numbers of hidden neurons in the hidden layers with respect to the cross-validation errors.

parameters. Finally, as our dataset is relatively small, a statistical comparison between different pattern recognition methods cannot be used to unequivocally indicate the optimal classification method. However, BER was significantly lower for all our selected two-dimensional shape features than for the simple bounding box method (36). We have developed a pattern recognition method that exploits tumour morphological structure which can be applied to larger prospective studies. Our objectively derived shape parameters may also be relevant for the monitoring of the tumour growth and treatment response.

The classification accuracy of our method is comparable with that from histopathological diagnosis (41), and thus could provide a diagnostic adjunct to guide the radiologist on the most probable diagnosis prior to considering further intervention or treatment. However, a larger prospective trial is needed to fully evaluate the accuracy of this method for its application in routine clinical use.

## CONCLUSIONS

The presented study proposes a two-dimensional shape-based decision support system to aid in the clinical diagnosis of GBMs *versus* solitary METs. The use of the D-SEG method provides an objective segmentation of the tumour boundary that includes subtle details of tumour shape that would be difficult to obtain by manual delineation. In comparison with other classification schemes, the distinct morphological appearances of GBMs and METs appear to provide features for improved classification accuracy.

## ACKNOWLEDGEMENTS

Guang Yang acknowledges the support of Cancer Research UK project grant C1459/A13303. Data were acquired during projects funded by Cancer Research UK (Grant number C8807/A3870) and EU FP7 [Grant number: LSHC-CT-2004-503094 (eTUMOUR)]. The authors would like to thank Tiernan Byrnes, Chris Clark and Alan Wright for help with MRI data acquisition.

## REFERENCES

- Opstad KS, Murphy MM, Wilkins PR, Bell BA, Griffiths JR, Howe FA. Differentiation of metastases from high-grade gliomas using short echo time  $^1\text{H}$  spectroscopy. *J. Magn. Reson. Imaging*. 2004; 20(2): 187–192.
- Giese A, Westphal M. Treatment of malignant glioma: a problem beyond the margins of resection. *J. Cancer Res. Clin. Oncol.* 2001; 127(4): 217–225.
- Furnari FB, Fenton T, Bachoo RM, Mukasa A, Stommel JM, Stegh A, Hahn WC, Ligon KL, Louis DN, Brennan C, Chin L, DePinho RA, Cavenee WK. Malignant astrocytic glioma: genetics, biology, and paths to treatment. *Genes Dev.* 2007; 21(21): 2683–2710.
- Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJB, Belanger K, Brandes A, Marosi C, Bogdahn U, Curschmann J, Janzer RC, Ludwin SK, Gorlia T, Allgeier A, Lacombe D, Cairncross JG, Eisenhauer E, Mirimanoff RO. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N. Engl. J. Med.* 2005; 352(10): 987–996.
- Marumoto T, Saya H. Molecular biology of glioma. In: Yamanaka R (ed.). *Glioma*, Vol. 746. Springer: New York; 2012; 2–11.
- Johnson DR, Ma DJ, Buckner JC, Hammack JE. Conditional probability of long-term survival in glioblastoma: a population-based analysis. *Cancer* 2012; 118(22): 5608–5613.
- Gavrilovic I, Posner J. Brain metastases: epidemiology and pathophysiology. *J. Neurooncol* 2005; 75(1): 5–14.
- Kocher M, Soffietti R, Abacioglu U, Villà S, Fauchon F, Baumert BG, Fariselli L, Tzuk-Shina T, Kortmann R-D, Carrie C, Ben Hassel M, Kouri M, Valeinis E, van den Berge D, Collette S, Collette L, Mueller R-P. Adjuvant whole-brain radiotherapy versus observation after radiosurgery or surgical resection of one to three cerebral metastases: results of the EORTC 22952-26001 study. *J. Clin. Oncol.* 2011; 29(2): 134–141.
- Fellows GA, Wright AJ, Sibbain NA, Rich P, Opstad KS, McIntyre DJO, Bell BA, Griffiths JR, Howe FA. Combined use of neuroradiology and  $^1\text{H}$ -MR spectroscopy may provide an intervention limiting diagnosis of glioblastoma multiforme. *J. Magn. Reson. Imaging*. 2010; 32(5): 1038–1044.
- Cha S, Lupo JM, Chen M-H, Lamborn KR, McDermott MW, Berger MS, Nelson SJ, Dillon WP. Differentiation of glioblastoma multiforme and single brain metastasis by peak height and percentage of signal intensity recovery derived from dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging. *Am. J. Neuroradiol.* 2007; 28(6): 1078–1084.
- Tang YM, Ngai S, Stuckey S. The solitary enhancing cerebral lesion: can FLAIR aid the differentiation between glioma and metastasis? *Am. J. Neuroradiol.* 2006; 27(3): 609–611.
- Devos A, Lukas L, Suykens JAK, Vanhamme L, Tate AR, Howe FA, Majós C, Moreno-Torres A, van der Graaf M, Arús C, Van Huffel S, Van Huffel S. Classification of brain tumours using short echo time  $^1\text{H}$  MR spectra. *J. Magn. Reson.* 2004; 170(1): 164–175.
- Georgiadis P, Cavouras D, Kalatzis I, Glotsos D, Athanasiadis E, Kostopoulos S, Sifaki K, Malamas M, Nikiforidis G, Solomou E. 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(1): 120–130.
- Byrnes TJD, Barrick TR, Bell BA, Clark CA. Diffusion tensor imaging discriminates between glioblastoma and cerebral metastases in vivo. *NMR Biomed.* 2011; 24(1): 54–60.
- Chen XZ, Yin XM, Ai L, Chen Q, Li SW, Dai JP. Differentiation between brain glioblastoma multiforme and solitary metastasis: qualitative and quantitative analysis based on routine MR imaging. *Am. J. Neuroradiol.* 2012; 33(10): 1907–1912.
- Crisi G, Orsingher L, Filice S. Lipid and macromolecules quantitation in differentiating glioblastoma from solitary metastasis: a short-echo time single-voxel magnetic resonance spectroscopy study at 3 T. *J. Comput. Assist. Tomogr.* 2013; 37(2): 265–271.
- Frieboes HB, Zheng X, Sun C-H, Tromberg B, Gatenby R, Cristini V. An integrated computational/experimental model of tumor invasion. *Cancer Res.* 2006; 66(3): 1597–1604.
- Frieboes HB, Lowengrub JS, Wise S, Zheng X, Macklin P, Bearer EL, Cristini V. Computer simulation of glioma growth and morphology. *Neuroimage* 2007; 37(Suppl. no. 0): S59–S70.
- Claes A, Idema A, Wesseling P. Diffuse glioma growth: a guerilla war. *Acta Neuropathol.* 2007; 114(5): 443–458.
- Kim Y, Lawler S, Nowicki MO, Chiocca EA, Friedman A. A mathematical model for pattern formation of glioma cells outside the tumor spheroid core. *J. Theor. Biol.* 2009; 260(3): 359–371.
- Jones TL, Bell BA, Barrick TR. A novel whole-brain DTI segmentation technique for brain tumour delineation and diagnosis. *Proceedings of the 20th Annual Meeting ISMRM*, Town, State, Country, 2012; 188.
- Jones TL, Byrnes TJ, Yang G, Howe FA, Bell BA, Barrick TR. Brain tumour classification using the diffusion tensor image segmentation (D-SEG) technique. *Neuro Oncol.* 2014; In press.
- Haury A-C, Gestraud P, Vert J-P. The influence of feature selection methods on accuracy, stability and interpretability of molecular signatures. *PLoS One*. 2011; 6(12): e28210.
- Abdi H. The Bonferroni and Šidák corrections for multiple comparisons. In: Salkind NJ (ed.). *Encyclopedia of Measurement and Statistics*, 1st edn. SAGE Publications, Inc.: Town; 2007; 103–107.
- Peck R, Devore JL. *Statistics: The Exploration and Analysis of Data*, 7th edn. Publisher: Town; 2011, 213–219.
- Lachenbruch PA, Goldstein M. Discriminant analysis. *Biometrics*. 1979; 35(1): 69–85.
- John G, Langley P. Estimating continuous distributions in Bayesian classifiers. *Proc. Elev. Conf. Uncertain. Artif. Intell.* 1995; XX: 338–345.
- Friedman J, Bentley J, Finkel R. An algorithm for finding best matches in logarithmic expected time. *ACM Trans. Math. Softw.* 1977; 1549 (July): 209–226.
- Cristianini N, Shawe-Taylor J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, 1<sup>st</sup> edn. Cambridge University Press: Town; 2000.
- Hagan MT, Demuth HB, Beale MH. *Neural Network Design*. PWS Pub. Co.: Town; 1996.
- Murphy K. *Machine Learning: A Probabilistic Perspective*. MIT Press: Town; 2012, 387–391.
- Domingos P, Pazzani M. On the optimality of the simple Bayesian classifier under zero-one loss. *Mach. Learn.* 1997; 29(2–3): 103–130.



- 33 Karabatak M, Ince MC. Expert systems with applications an expert system for detection of breast cancer based on association rules and neural network. *Expert Syst. Appl.* 2009; 36(2): 3465–3469.
- 34 Chen Y, Lin C. Combining SVMs with various feature selection strategies. In: Guyon I, Nikravesh M, Gunn S, Zadeh LA (eds). *Feature Extraction: Foundations and Applications*. Springer: Berlin Heidelberg; 2006, 315–324.
- 35 Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. *J. Am. Stat. Assoc.* 1983; 78(382): 316–331.
- 36 Blanchet L, Krooshof PWT, Postma GJ, Idema AJ, Goraj B, Heerschap A, Buydens LMC. Discrimination between metastasis and glioblastoma multiforme based on morphometric analysis of MR images. *Am. J. Neuroradiol.* 2011; 32(1): 67–73.
- 37 Nigsch F, Bender A, van Buuren B, Tissen J, Nigsch E, Mitchell JBO. Melting point prediction employing k-nearest neighbor algorithms and genetic parameter optimization. *J. Chem. Inf. Model.* 2006; 46(6): 2412–2422.
- 38 Wetschereck D, Aha D, Mohri T. A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. *Artif. Intell. Rev.* 1997; 11(1–5): 273–314.
- 39 Rokach L. Ensemble-based classifiers. *Artif. Intell. Rev.* 2009; 33(1–2): 1–39.
- 40 Georgiadis P, Cavouras D, Kalatzis I, Daskalakis A, Kagadis GC, Sifaki K, Malamas M, Nikiforidis G, Solomou E. Improving brain tumor characterization on MRI by probabilistic neural networks and non-linear transformation of textural features. *Comput. Methods Programs Biomed.* 2008; 89(1): 24–32.
- 41 Aker FV, Hakan T, Karadereler S, Erkan M. Accuracy and diagnostic yield of stereotactic biopsy in the diagnosis of brain masses: comparison of results of biopsy and resected surgical specimens. *Neuropathology*, 2005; 25(3): 207–213.