MAGMA (2004) 16: 194–201
DOI 10.1007/s10334-003-0023-7

RESEARCH ARTICLE

# Benign/malignant classifier of soft tissue tumors using MR imaging

**Juan M García-Gómez**
**César Vidal**
**Luis Martí-Bonmatí**
**Joaquín Galant**
**Nicolas Sans**
**Montserrat Robles**
**Francisco Casacuberta**

Received: 28 May 2003
Accepted: 20 October 2003
Published online: 1 March 2004


Dr. L. Martí-Bonmatí (✉)
Resonancia Magnética,
Servicio de Radiología,
Hospital Universitario Dr. Peset,
Avda Gaspar Aguilar 90,
46017 Valencia, Spain
E-mail: Luis.Marti@uv.es

C. Vidal · L. Martí-Bonmatí
M. Robles · J. M. García-Gómez
BET, Informática Médica,
Universidad Politécnica de Valencia,
Spain

L. Martí-Bonmatí
Servicio de Radiología,
Hospital U. Dr. Peset,
Valencia, Spain

J. Galant
Servicio de Radiología,
Hospital San Juan de Alicante,
Spain

N. Sans
Department of Radiology,
CHU Purpan, Toulouse, France

F. Casacuberta
ITI-DSIC, Universidad Politécnica de
Valencia, Spain

**Abstract** This article presents a pattern-recognition approach to the soft tissue tumors (STT) benign/malignant character diagnosis using magnetic resonance (MR) imaging applied to a large multicenter database.
*Objective*: To develop and test an automatic classifier of STT into benign or malignant by using classical MR imaging findings and epidemiological information.
*Materials and methods*: A database of 430 patients (62% benign and 38% malignant) from several European multicenter registers. There were 61 different histologies (36 with benign and 25 with malignant nature). Three pattern-recognition methods (artificial neural networks, support vector machine, k-nearest neighbor) were applied to learn the discrimination between benignity and malignancy based on a defined MR imaging findings protocol. After the systems had learned by using training samples (with 302 cases), the clinical decision support system was tested in the diagnosis of 128 new STT cases.
*Results*: An 88–92% efficacy was obtained in a not-viewed set of tumors using the pattern-recognition techniques. The best results were obtained with a back-propagation artificial neural network.
*Conclusion*: Benign vs. malignant STT discrimination is accurate by using pattern-recognition methods based on classical MR image findings. This objective tool will assist radiologists in STT grading.

**Keywords** Magnetic resonance imaging · Soft tissue tumor · Pattern recognition · Clinical decision support systems · Artificial neural networks · Support vector machine · K-Nearest neighbor

## Introduction

Benign tumor constitutes by far the most common soft tissue tumor (STT). Most of these tumors can be characterized with magnetic resonance (MR) imaging because of their main textural and growth pattern. Soft tissue malignant tumors represent approximately 33% of them [1]. Although, in clinical practice, most superficial tumors are diagnosed solely on the basis of clinical exploration and no further studies are required, more deeply located lesions are nowadays studied with MRI. A clear discrimination between benign and malignant tumors becomes crucial in planning the proper diagnostic and surgical procedures.

As STTs are infrequent in the general population, most radiologists and clinicians are not familiar with their appearance and differential diagnosis. Furthermore, their appearance is sometimes misleading. Well-defined malignant tumors are placed on the less aggressive side of the spectrum and have an appearance usually similar to benign lesions, making it difficult on many occasions to distinguish between them. On the other side, some benign tumors can deviate from their typical homogeneous appearance, complicating the differentiation. A diagnostic support tool developed to help in making an accurate diagnosis of tumor aggressiveness would improve the correct management of these patients.

To create such a diagnostic support tool for STTs, a large amount of data is required, usually larger than those disposable in an isolated center. The solution to the lack of data has been solved by the merger of several hospital institution cases. The Information and Communication Technologies allow radiologists to share relevant information about the problem and special cases. In this way, other groups, like the Belgian Soft Tissue Neoplasm Registry [2], also recruit MR imaging cases in order to offer support services to professionals and scientists.

To our knowledge, there is no published report in which automatic classification methods, such as artificial neural networks (ANN), k-nearest neighbor (k-NN), or support vector machine (SVM), have been applied successfully to classify the STT tumors with a MR imaging-findings dataset. A recent article related to soft tissue sarcoma classification using gene expression profile and SVM has been presented by Segal NH et al [3]. However, all of the abovementioned methods have been applied in many medical classification systems with good performance. In tumor classification, k-NN has been used in combination with two other classifiers for melanoma diagnosis [4]; SVM has been successfully used for prediction of malignancy of ovarian tumors [5], in analyzing microcalcifications in digital mammograms [6], and in discriminating breast tumors [7–9]; ANN has also been widely and successfully applied in discriminating breast tumors [10–12], in evaluating different features based on defined imaging criteria [13–16]; and in cervical cancer screening [17]. An interesting review of the applications of ANN to the diagnosis, prognosis, and survival analysis in the medical domains of oncology, critical care, and cardiovascular medicine has been presented by Lisboa [18] to assess the evidence of healthcare benefits involving the application of ANN. In this review, the author presents important recommendations for the design and evaluation of ANN in medicine.

Our main purpose was to obtain an automatic classifier of STT aggressiveness. The pattern-recognition discipline allows the adaptation of models to a concrete problem. This adaptation consists of an inference process in which a mathematical mechanism changes its parameters based on a set of supervised training examples. Pattern-recognition methods (ANN, k-NN, and SVM) have been applied to design this software because they allow the adaptation of the system to the problem by the use of knowledge contained in the study samples. The classification tool will be used by radiologists and other medical professionals in a clinical decision-support system to aid in the diagnostic report. Our article presents a multidisciplinary study, grouping MRI expertise radiologists and computer science engineers on pattern-recognition techniques, to discriminate between benign/malignant grading of STT.

## Materials and methods

### Patients and equipment

Patients with confirmed musculoskeletal STT examined with MRI were retrospectively selected. Tumors with a biopsy procedure before the imaging studies were excluded. Patients were consecutively recruited at five different European hospital centres.

All malignant and some benign, other than classical hemangioma and lipomas, had histopathological confirmation of their nature. Characteristic multimodality imaging appearance and follow-up studies were employed to confirm the diagnosis on these benign tumors. Finally, there were 430 patients, 267 patients (62.1%) with a benign lesion and 163 patients (37.9%) with a malignant neoplasm. The histological grouping distribution of the cases is shown in Fig. 1.

Imaging studies were performed on 0.5 T and 1.5 T Gyroscan NT (Philips Medical System, Eindhoven, Netherlands) and 0.5 T Signal (General Electric Medical System, Milwaukee, WI) units. T1-weighted (T1W, TR/TE, 450–650 ms/15–30 ms) and T2-weighted fat-suppression (T2W-FS, TR/TE, 2,800-3,500 ms/80–120 ms) or STIR (TR/TE/TI, 1,400-1,600 ms/40 ms/100–120 ms) images were systematically obtained (Fig. 2).

### Database

The following epidemiological and MR image findings [1, 2, 19–21] were obtained from the clinical records and radiological examinations:

1. Age: The age of the patient in years.
2. Clinical presentation: Reason for the patient's consultation (mass, pain, growth, neurological symptoms, skin alterations, asymptomatic).

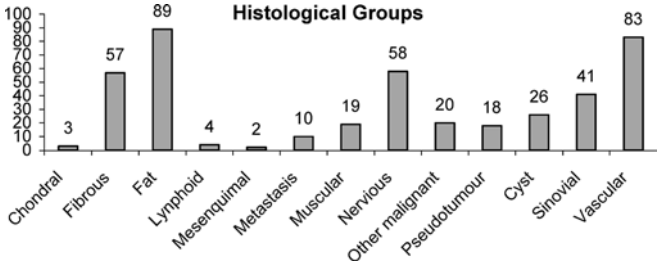


**Fig. 1** Distribution of cases into histological groups

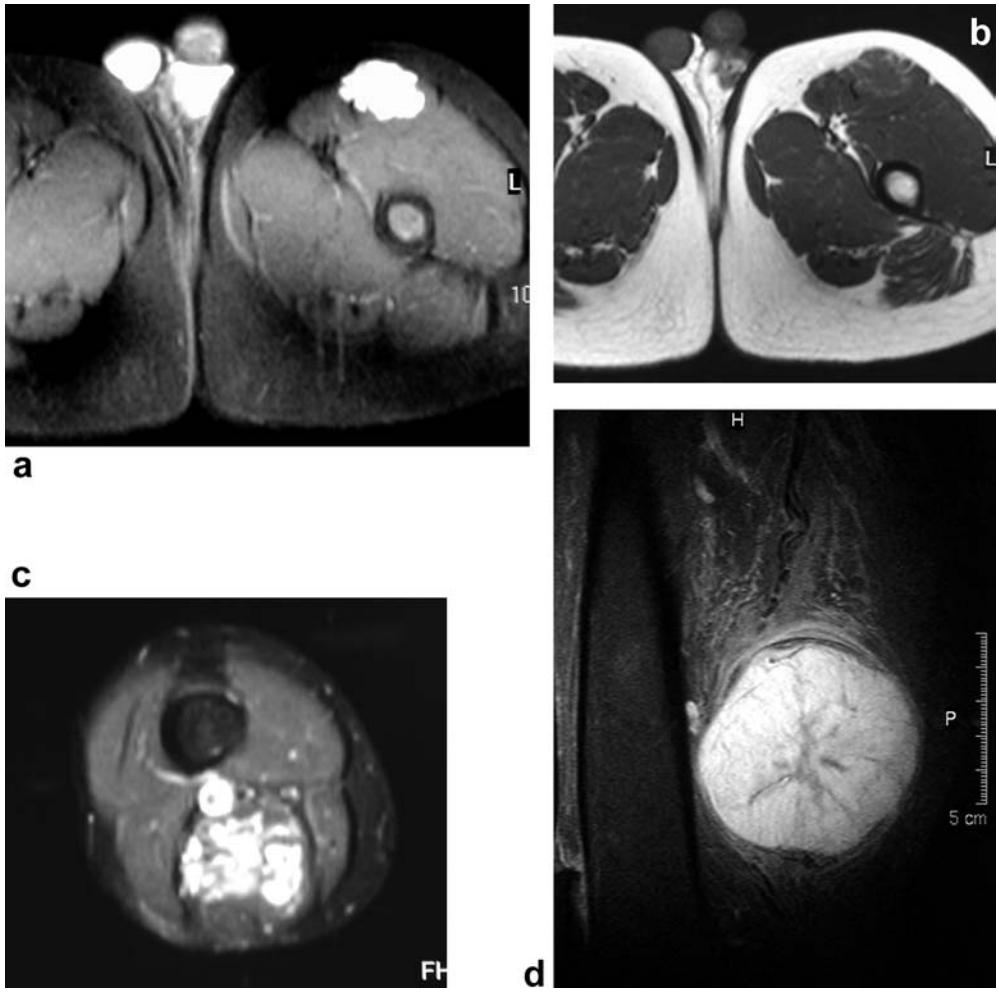


**Fig. 2 a–d** Representative MR images. Homogeneous highly hyperintense hemangioma in T2 weighted image **a** with hyperintense tracts in the T1 weighted image **b**. Serpiginous appearance of an intramuscular hemangioma (**c**, fat suppression T2 weighted image). Oval hyperintense heterogeneous neurinoma with inner areas of fibrosis (**d**, fat suppression T2 weighted image)

3. Localization: The anatomic compartment where the STT was located (subcutaneous, intramuscular, intermuscular, or intraarticular). Lesions extending to more than one compartment were classified within the compartment where the largest part of the tumor was located.
4. Size: The maximum diameter of the lesion, expressed in centimeters (rounded to the nearest unit).
5. Shape: The external morphology of the lesion (rounded, oval, lobulated, serpiginous, fusiform, strand, or irregular). Serpiginous lesions were those with tubular structures within them and polylobular external surface (Fig. 2).
6. Signal intensity: Represents the relative signal intensity of the lesions in the images obtained with different weightings (T1 and T2/STIR). Muscle was selected as the reference signal intensity tissue, with an intermediate-low signal intensity in most sequences (T1-weighted, T2-weighted fat suppression, STIR images). There were four categories in the T1-weighted images (hypointense, isointense, hyperintense, and very hyperintense) and three in the T2-weighted images (isointense, slightly hyperintense, and highly hyperintense) (Fig. 2).
7. Margins: The external borders of the lesion. There were three categories: infiltrative (when most of the margins were blurred or the tumor clearly extended at any point to the surrounding tissue), mainly well-defined lesions (with a partial sector of the margin with unclear borders), and noninfiltrative (when the tumor borders were clearly visible without any peripheral infiltration).
8. Homogeneity: Presence of different areas of signal intensity variation within the lesion in the different image weightings. Taking into consideration the proportion and signal-intensity differences, there were four categories: very homogeneous (only one signal intensity constituent), homogeneous (an area not more than 25% of the lesion with a slight change in its signal intensity compared with the rest of the lesion), heterogeneous (an area between 25 and 75% with different signal intensities), and very heterogeneous (lesions with more than a 75% of them with different components showing large differences in their signal intensity) (Fig. 2).
9. Edema: A peripheral ill-defined area or halo, hypointense in T1-weighted images but hyperintense in STIR/T2-weighted images with a width larger than 5 mm (no, yes).
10. T1 hyperintense tracts: Linear or reticular areas with a high signal intensity in the T1-weighted images without fat suppression (no, yes). (Fig. 2).
11. Multiplicity: Whether the patient had previously or currently presented with similar STT (no, yes).
12. Target appearance: Whether the lesion had an inner well-defined centre and concentric peripheral rings with different signal intensities (no, yes).
13. Muscular atrophy: A decrease in the diameter of the muscle or group of muscles related to the lesion, with enlargement of the fat planes between the muscular fascicles, and especially if the contralateral extremity was present for comparison (no, yes).

14. Intratumoral hemorrhage: Whether heterogeneous areas with hyperintense and hypointense zones were shown in the T1 and STIR/T2 weighted images (no, yes).
15. Calcification: Very hypointense areas in all the obtained images and weightings, after excluding vessels and hemosiderin. Phebolits were considered only if a plain film or a CT examination showed a rounded calcification with an inner radiolucent centre (no, yes, phebolits).
16. Dependence: Whether an anatomic structure was identified with a very close relationship with the lesion and the lesion seems to originate from that anatomic structure. Special care was taken not to include as dependence a relationship of displacement (none, nerve, tendon, and vessel).
17. Intratumoral fat: When an area of signal intensity equal to that of the subcutaneous fat in all the pulse sequences was observed within the lesion. Lesions with fat were further classified regarding the presence of hyperintense zones in the STIR/T2 weighted images (no, fat without hyperintense zones in T2-STIR, fat with hyperintense zones in T2-STIR).
18. Fibrosis: Defined as areas of very low signal intensity within the tumor in all the images, mainly if the morphology was irregular or ring-like (no, yes) (Fig. 2).
19. Fascial relationship: The relationship of the subcutaneous lesions with the superficial fascia (no contact, small contact, contact with acute angles between the lesion and the fascia, larger contact with obtuse angles, fascia penetration, and fascial origin of the lesion).
20. Bone alterations: If there was an alteration of the underlying bone, it was classified as bone remodeling with periosteal reaction, or bone destruction, including cortical bone permeation (no, yes with erosion-invasion, yes with reshape or reaction).
21. Vessels: Whether large vessels constitute the essential part of the lesion (no, yes).

Two experience radiologists, who were masked to the final histological diagnosis, reviewed the studies. Differences were resolved by consensus with a third radiologist.

## Methods

A clinical decision support system is a tool to improve the quality in the clinical decision, being defined as an active knowledge system that generates specific advice to each new case [22]. It integrates three main features: medical knowledge that solves the disease cases [23, 24], patient data with specific biomedical information of each patient, and specific advice for each case based on the medical knowledge and the patient data.

There are different methods to design clinical decision-support systems using the artificial intelligence approach. In this study, an inductive strategy, more commonly named pattern-recognition strategy, was applied. The conclusions made by the clinical decision-support system were inferred by the knowledge captured from a group of samples representing the problem.

The pattern-recognition methodology used has the functional blocks diagram shown in Fig. 3 [25]. Preprocessing/filtering includes data connection and adaptation to a format recognized by the learning procedures; consequently, it incorporates database connection, queries to get the required registers, and syntactic analysis to transform the formats. The filtering step consisted of a transformation from the Access™ database format to an ASCII file. Another filter task was the normalization of the data to avoid artificial order relationships in qualitative variables with a simple local codification, and range transformation to weight correctly the continuous variables into [0, 1] [25]. As an example, variables like dependence, which takes the values none, nerve, tendon and vessel, that doesn't have an order relation between the values, was codified as (1 0 0 0), (0 1 0 0), (0 0 1 0), (0 0 0 1)"; but variables, such as age or size, that have an important relation in the order of its values, were normalized into [0, 1].

The database was randomly divided into two different sets, one training set used to learn the computational models (70%, 302 cases) and the test set (30%, 128 cases), used to estimate the accuracy of the system. We contemplated the use of the three basic evaluation methods in the experiment planning: cross validation, leaving-one-out, and independent training and test sets. We chose an independent training and test sets because the real independence between both sets will probably give a less optimistic confidence interval of the efficiency (shown in Table 1) than the one obtained with nonindependent test sets.

The learning process is the most important step in the clinical decision support system development and its principal goal is to adapt the software system to the requirements of the problem. An iterative procedure takes the relevant information by studying examples to inference the model of the problem. When the learning process was completed, the adaptive system was released and new cases could be studied. The variables of the unseen new cases were analyzed and the generalization knowledge made from the training examples got the associated prediction.

ANN, k-NN, and SVM were used as pattern-recognition methods to design the clinical classifier. These three techniques are nonparametric because they do not require a prior knowledge of the probability distributions. The ANN model is based on a simple function definition: the perceptron. The multiple connections between many of these mechanisms get the great representation capability. Briefly, the perceptron output is composed of the nonlinear function of a linear combination of the entries [26]. A typical neural connection topology is the multilayer perceptron [25, 26]. In the STT benign and malignant classification problem, the functionality of the neural network can be seen like a nonconnected boundary to separate the regions between each class. The learning process with ANN consisted in the error minimization between the real targets (the class of the training samples) and the net outputs using the back-propagation algorithm. This process has been applied to a set of neural network topologies with one or two hidden layers. In our case, the number of neurons in the hidden

**Table 1** Comparison of artificial neural network (ANN), k-nearest neighbor (k-NN), and support vector machine (SVM) methods. Efficiency is expressed as value ± 95% confidence interval. PPV, positive predictive value; NPV: negative predictive value

| | Efficiency | Sensitivity | Specificity | PPV | NPV |
|---|---|---|---|---|---|
| ANN | 92 ± 5% | 86% | 95% | 91% | 93% |
| KNN | 88 ± 6% | 86% | 90% | 84% | 91% |
| SVM | 90 ± 5% | 84% | 94% | 89% | 89% |

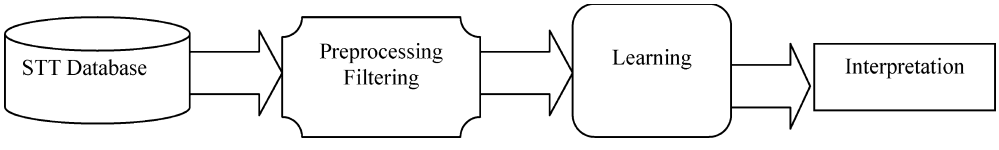


**Fig. 3** Pattern-recognition methodology followed in the STT study

layers varied from 10 to 100 and two neurons in the output layer that represents (1 0) for benign and (0 1) for malignant. Each ANN trained with the STT database was tested with a set of samples that had not been used for training. The stop criteria in all the NN experiments was to repeat the training cycle until one of these conditions was reached: cycles $\geq$ 3,000, MSE (mean square error) $< 0.01$, or differential MSE ($|\mathrm{MSE}_i\text{-}\mathrm{MSE}_{i-1}|$) $< 0.000001$. The test offers an estimate of goodness of each network. We have used the Stuttgart Neural Networks Simulator to train the decision system [27].

K-nearest neighbors is a simple and fast distance-based classifier broadly used as a reference when studying a pattern-recognition problem [4, 28]. The training database of STT establishes the knowledge of systems like a set of points in a multidimensional space. When a new sample needs to be classified, the k-NN chooses by vote the class of the sample. In our STT experiments, the k-NN classifier was proved varying $k$ between 1 and 15 with three distance metrics[1] The best configuration was the five nearest neighbors with the $L_1$ distance, but its efficiency was worse than that obtained with the other two methods, as will be commented upon in the "Results."

Support vector machines define the optimal separating hyperplane with the maximal margin [25]. This margin is the minimum distance of patterns of the training set to the hyperplane. SVMs use the approach of representing the data patterns in a higher dimension space with a nonlinear mapping function, increasing the discrimination power [29]. The choice of the nonlinear mapping function or kernel depends on the problem. Due to the way in which kernel transformation is made, by means of an inner product, the computational cost of managing such a high number of features is not relevant. The problem to solve is reduced to a convex Lagrange optimization problem, minimizing a quadratic function under linear inequality constraints. This choice provides the absence of local minima and allows the training with thousands of samples. If data are not separable with a linear hyperplane, as in most real problems, some new conditions are added to the optimization problem to allow the limitation of errors contribution in the construction of the hyperplane. In our experiments, several kernels have been tried out: polynomial, Gaussian and radial basis function, modifying the inner parameters in a wide range, using SVM-light implementation [30].

The total number of experiments using the explained techniques was about 1,500, most of them due to the high number of different topologies in two hidden layers ANN.

## Results

We created a database with a set of MRI findings of 430 soft tissue tumors (62% benign, 38% malignant) coming from five different European hospital centres, with the purpose of developing an automatic benign/malignant classifier of STT by using a set of pattern-recognition techniques.

The best results obtained classifying a set of 128 samples, corresponding to 30% of the total database obtained with a random selection, with ANN, k-NN, and SVM are shown in Tables 2–4. A comparative table including the main parameters[2] estimated for each technique is in Table 1.

**Table 2** Best results obtained with artificial neural network (ANN). One case was not classified by the network. Distribution of cases into histological groups

| Test/final diagnosis | Malignant | Benign | Unknown | Total |
|---|---|---|---|---|
| Malignant | 42 (86%) | 6 (12%) | 1 (2%) | 49 |
| Benign | 4 (5%) | 75 (95%) | 0 (0%) | 79 |
| Total | 46 (86%) | 81 (86%) | 1 (86% | 128 |

**Table 3** Best results obtained with K-NN. The number of neighbors was $k = 5$

| Test/final diagnosis | Malignant | Benign | Total |
|---|---|---|---|
| Malignant | 42 (86%) | 7 (14%) | 49 |
| Benign | 8 (10%) | 71 (90%) | 79 |
| Total | 50 (90%) | 78 (90%) | 128 |

**Table 4** Best results obtained with SVM. They were obtained using a polynomial kernel (degree 7)

| Test/final diagnosis | Malignant | Benign | Total |
|---|---|---|---|
| Malignant | 41 (84%) | 8 (16%) | 49 |
| Benign | 5 (6%) | 74 (94%) | 79 |
| Total | 46 (94%) | 82 (94%) | 128 |

With ANN, the best efficiency (92%) was obtained with 30 neurons in the first hidden layer and 10 in the second layer, trained with the back-propagation algorithm. Despite the prevalence of benign classes, sensitivity and specificity were quite good (86 and 95%, respectively). The decision thresholds made one malignant case not be classified (Table 2). Best efficiency with k-NN (88%) has been obtained using $k = 5$ neighbors and $L_1$ distance. The sensitivity was 86% (Table 3) and the specificity 90%. Best efficiency with SVM (90%) was obtained using a polynomial kernel of 7th degree. The sensitivity (84%) was worse than that obtained with the other techniques and specificity 94% (Table 4).

ANN reached the best results, with more than 90% in efficacy and specificity coefficients (Table 1). Classifier fusion has been considered, but we have studied the concrete error cases and only one ANN error is correctly classified by k-NN and SVM (Table 5). This result

[1] $L_0(x,y) = \max_{1\leq i\leq d}|x_i - y_i|$

$L_1(x,y) = \sum_{i=1}^{d}|x_i - y_i| \quad L_2(x,y) = \sqrt{\sum_{i=1}^{d}(x_i - y_i)^2}.$

[2] Efficiency, (TP + TN)/N; TP, true positive; FN, false negative; prevalence, (TP + FN)/N (a priori probability of malignant); FP, false positive; TN, true negative; specificity, TN/(TN + FP + UN) (probability of true benign); UN, unknown; sensitivity, TP/(TP + FN + UN) (probability of true malignant); PPV, positive predictive value, TP/(TP + FP) (reliability of positive); NPV, negative predictive value, TN/(TN + FN), reliability of negative

**Table 5** Errors in benign vs. malignant classification with the artificial neural network (ANN), k-nearest neighbor (k-NN), and support vector machine (SVM) methods

| Definitive diagnosis | ANN | KNN | SVM | Histology |
|---|---|---|---|---|
| Benign | Benign | Malignant | Malignant | Lymphomatoid granulomatosis |
| Benign | Malignant | Malignant | Malignant | Granuloma |
| Benign | Benign | Malignant | Benign | Hemangioma |
| Benign | Malignant | Malignant | Malignant | Localized pigmented villonodular synovitis |
| Benign | Benign | Malignant | Benign | Pseudotumor |
| Benign | Malignant | Malignant | Malignant | Myxoma |
| Benign | Malignant | Malignant | Malignant | Morton's neuroma |
| Benign | Benign | Malignant | Benign | Desmiod |
| Malignant | Benign | Malignant | Malignant | Malignant fibrous histiocytoma |
| Malignant | Unknown | Malignant | Benign | Malignant fibrous histiocytoma |
| Malignant | Malignant | Malignant | Benign | Malignant fibrous histiocytoma |
| Malignant | Benign | Benign | Benign | Metastases |
| Malignant | Benign | Benign | Benign | Metastases |
| Malignant | Malignant | Benign | Malignant | Malignant schwannoma |
| Malignant | Malignant | Malignant | Benign | Myxoid liposarcoma |
| Malignant | Malignant | Benign | Malignant | Liposarcoma |
| Malignant | Malignant | Benign | Malignant | Synovial sarcoma |
| Malignant | Benign | Malignant | Benign | Synovial sarcoma |
| Malignant | Benign | Benign | Benign | Fibrosarcoma |
| Malignant | Benign | Benign | Benign | Fibrosarcoma |

indicates that it is not necessary for the classifier fusion, but it must be considered in future improvements.

## Discussion

Magnetic resonance imaging clearly depicts STT due to its high-contrast tissue resolution and multiplanar capability. Correct diagnosis includes the detection, characterization, and staging of these tumors. The MR differentiation between benign and malignant tumors is complicated by the low prevalence of these lesions, radiologist's low experience in nondedicated hospitals, indirect information of the MR signal intensities, high diversity of histologies, and natural evolution of the lesions. There is much controversy regarding the MR value in the differentiation of benign and malignant STTs. Although a subjective analysis has a high sensitivity for malignancy (78%) [21], the statistical combination of individual MR parameters provides both a high sensitivity and specificity, close to 81% (2). Our approach was to develop an automatic technique to help radiologist characterize benign and malignant STTs with high accuracy.

The use of automatic techniques to help in the characterization of STT requires the homogenization of patient databases. A standard MR protocol that incorporates the relevant information to characterize STT is very important in an automatic classification task. For this reason, a standard protocol based on MRI T1-weighted and T2/STIR-weighted images has been used by MR expert radiologists in this study. The automatic systems can retrieve the relevant information of the cases, and good classification was reached because of the quality of the variables extracted by the radiologist. We also did not use the information generated by contrast-enhanced images because of the diversity of the data within centers. The usefulness of gadolinium-enhanced MR images, although widely accepted, did not assist in further narrowing the differential diagnosis compared with plain MR scans in a large series of cases [31]. Moreover, the lack of uniform imaging sequences and parameters after contrast administration within the different centers made standardization difficult.

It has been important to make a multicenter consortium to recruit the STT database. A large database was needed to properly develop a general classifier. The use of cases from different places to train and test the decision-support systems makes it useful to generalized our results. However, although the series of STTs used in our study was large, with good results in a multicenter database, the incorporation of new STT categories and presentations to the study will increase the quality of the decision-task process. All classifiers yielded significant prediction efficiency (88-92%), with great consensus between them. The best technique was the ANN with 10 errors over 128 cases (92% of hits).

The similar results of efficiency, sensitivity, and specificity obtained with the three evaluated techniques

may prove that some bias may be due to an inherent biological overlap. One possible source of bias is the fact that radiologists' readings were used to train the clinical decision support system. The use of simple MRI variables, as the ones used in this study, by the STT classifier with the adaptive approach offered by the pattern recognition discipline will minimize the influence of reading errors from different radiologists.

The practical result of the pattern-recognition experiments should be a clinical decision-support system in STT discrimination. This tool is useful because it is an objective method to confirm the benign/malignant characterization, allows the investigation of suspicious cases, and has the capability to assist the radiologist's decision in a new case and aids in the education of new radiologists' expertise in STT.

The inductive approximation used in this study enables the system to learn the important features of the cases to make a classification into two categories (benign vs. malignant) defined by the values of the variable or character. This STT problem is a dichotomy classification between two exclusive classes. The main difficulty of this problem is the heterogeneous constitution of the classes. Each class (benign and malignant) is formed by a group of histologies of very different origin. Consequently, there is not just one big cluster per class, making the decision boundary easy to discriminate.

The number of cases in which all the techniques failed is 8 (6% of test set) (Table 5). There are various reasons to explain this circumstance: these cases can be very abnormal and the learning process has not enough information to extract the discriminative pattern correctly; the cases are really atypical in their appearance; or the registries have incorrectly filled variables.

The use of the pattern-recognition approach in medical research is growing more and more because of the new possibilities opened by the digitalization of biomedical information. The disposability of biomedical information in electronic repositories [32] enables the data-mining studies and research by automatic methods to get new and interesting correlations to improve human health. The pattern-recognition approach can help the search for biomedical pointers of important diseases (like tumors or degenerative diseases) and the development of technological tools applied to clinical and basic medicine research [3, 4, 7, 8, 33, 34].

The practical result of the pattern-recognition experiments should be a clinical decision-support system in STT discrimination.

Our benign/malignant classification is part of a larger study project on STT computer-aided diagnosis, including visualization of tumor data, and feature selection and classification into different histologies [35, 36]. A Web Services layer and a graphical application are being implemented in order to allow the use of clinical decision-support systems developed by the group from distributed application around Internet such as a Web Site, Desktop Application, Clinical Electronic History, or Telemedicine applications.

**Acknowledgements** The authors thank the involved hospitals in the project (Hospital Universitario San Juan de Alicante, Hospital Universitario Dr Peset de Valencia, Hospital Cruces de Baracaldo, Hospital Juan Canalejo de La Coruña, and Hôpital Universitaire de Toulouse) for their recruiting support; ADIRM (Asociación para el Desarrollo y la Investigación en Resonancia Magnética) for their members' scientific advice; Ministerio de Sanidad y Consumo supporting grant INBIOMED; grant IM3 and Universidad Politécnica de Valencia for its supporting grant 20010690.

## References

1. Weatherall P (1995) Benign and malignant masses. MR imaging differentiation. Magn Reson Imaging Clin N Am 3:669–694
2. De Schepper AM, De Beuckeleer L, Vandevenne J, Somville J (2000) Magnetic resonance imaging of soft tissue tumors. Eur Radiol 10:213–223
3. Segal NH, Pavlidis P, Antonescu CR et al (2003) Classification and subtype prediction of adult soft tissue sarcoma by functional genomics. Am J Pathol 163:691–700
4. Sboner A, Eccher C, Blanzieri E et al (2003) A multiple classifier system for early melanoma diagnosis. Artif Intell Med 27:29–34
5. Lu C, Gestel V, Suykens JAK et al (2003) Preoperative prediction of malignancy of ovarian tumors using least squares support vector machines. Artif Intell Med 28:281–306
6. Bazzani A, Bevilacqua A, Bollini D et al (2001) A SVM classifier to separate false signals from microcalcifications in digital mammograms. Phys Med Biol 46:1651–1663
7. Chang RF, Wu WJ, Moon WK, Chen DR (2003) Improvement in breast tumor discrimination by support vector machines and speckle-emphasis texture analysis. Ultrasound Med Biol 29:679–686
8. Chang RF, Wu WJ, Moon WK, Chou YH, Chen DR (2003) Support vector machines for diagnosis of breast tumor on US images. Ultrasound Med Biol 10:189–197
9. Liu HX, Zhang RS, Luan F et al (2003) Diagnosis breast cancer based on support vector machines. J Chem Inf Comput Sci 43:900–907
10. Dhawan P, Chitre Y, Kaiser-Bnoasso C, Moskowitz M (1996) Analysis of mammographic using gray-level image structure features. IEEE Trans Med Imag 15:246–259
11. Sahiner B, Chan HP, Petrick N et al (1996) Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. IEEE Trans Med Imag 15:598–610

12. Chen DR, Chang RF, Kuo WJ et al (2002) Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks. Ultrasound Med Biol 28:1301–1310
13. Floyd CE, Lo JY, Yun AJ, Sullivan DC, Kornguth PJ (1994) Prediction of breast cancer malignancy using an artificial neural network. Cancer 74:2944–2948
14. Gurney JW, Swensen SJ (1995) Solitary pulmonary nodules: determining the likelihood of malignancy with neural network analysis. Radiology 196:823–829
15. Baker JA, Kornguth, PJ, Lo JY, Floyd CE (1996) Artificial neural network: improving the quality of breast biopsy recommendations. Radiology 198:131–135
16. Abdolmaleki P, Buadu LD, Murayaama, S et al (1997) Neural network analysis of breast cancer from MRI findings. Radiat Med 15:283–293
17. Prismatic Project Management Team (1999) Assessment of automated primary screening on PAPNET of cervical smears in the PRISMATIC Trial. Lancet 353:1381–1385
18. Lisboa PJG (2002). A review of evidence of health benefit from artificial neural networks in medical intervention. Neural Net 15:11–39
19. Galant J (1998) Magnetic Resonance contribution to the characterization of soft tissue tumor lesions. PhD Thesis. Facultad de Medicina, Universidad Miguel Hernández de Alicante. (Actual text in Spanish)
20. Galant J, Marti-Bonmatí L, Sáez F, Soler R, Alcala-Santaella R, Navarro M (2002) The value of fat-suppressed T2 or STIR sequences in distinguishing lipoma from well-differentiated liposarcoma. Eur Radiol 13:337–343
21. Moulton JS, Blebea JS, Dunco DM, Braley SE, Bisset GS 3rd, Emery KH (1995) MR imaging of soft-tissue masses: diagnostic efficacy and value of distinguishing between benign and malignant lesions. Am J Roentgenol 164:1191–1199
22. Van der Lei J, Talmon JL (1997) Clinical decision-support systems. In: van Bemmel JH, Musen MA (eds) Handbook of medical informatics. Springer, Berlin Heidelberg New York, pp 261–276
23. Van Bemmel JH (1997) Methods for decision support. In: van Bemmel JH, Musen MA (eds) (1997) Handbook of medical informatics. Springer, Berlin Heidelberg New York, pp 239–260
24. Shortliffe EH, Perrealt LE (2000) Medical informatics, computer applications in health care and biomedicine. Springer, Berlin Heidelberg New York
25. Duda RO, Hart Peter E, Stork David G (2001) Pattern classification. Wiley, New York
26. Bishop CM (1995) Neural networks for pattern recognition. Oxford University Press, New York
27. Zell A et al. Stuttgart Neural Networks Simulator (1995–1998). University of Stuttgart
28. Cover TM, Hart PE (1967) Nearest neighbor pattern classification. IEEE Trans Inform Theory 13:21–27
29. Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, Cambridge
30. Joachims T (2002) SVM-light V4.00: support vector machine classification and learning modules.
31. Maldonado JA, Robles M, Cano C (2001) Integration of healthcare information systems: application of CEN/TC251 ENV13606. Proceedings of the 23rd annual international conference IEEE Engineering in Medicine and Biology Society
32. May DA, Good RB, Smith DK, Parsons TW (1997) MR imaging of musculoskeletal tumors and tumors mimickers with intravenous gadolinium: experience with 242 patients. Skeletal Radiol 26:2–15
33. Underwood J, Tate AR, Luckin R, Majos C, Capdevila A, Howe F, Griffiths J, Arus, C (2001) A prototype decision support system for MR spectroscopy-assisted diagnosis of brain tumors. Medinfo 2001:561–565
34. Abdolmaleki P, Buadu LD, Naderimansh H (2001) Feature extraction and classification of breast cancer on dynamic magnetic resonance imaging using artificial neural network. Cancer Lett 171:183–191
35. García-Gómez JM., Vidal C, Juan A, Martí-Bonmatí L, Galant J, Robles M (2002) Pattern classification techniques applied to soft tissue tumors using MR and epidemiological data. Proceedings 19th annual meeting of the European Society for Magnetic Resonance in Medicine and Biology, pp 274–275
36. Vidal C, García-Gómez JM, Martí-Bonmatí L, Robles M, Galant J, Millet J (2002) Study of vascular histological group tumors using pattern recognition techniques. Proceedings of the 9th congress of the *Sociedad Española de Informática de la Salud*. (Actual text in Spanish), pp 207–211