Tissue discrimination in magnetic resonance imaging of the rotator cuff

Tissue discrimination in magnetic resonance imaging of the rotator cuff

G J Meschino[1, 3], D S Comas[2, 4], M A González[2], C Capiel[3, 5] and V L Ballarin[2, 3]

[1] Bioengineering Lab, ICyTE, Universidad Nacional de Mar del Plata, Argentina.

[2] Digital Image Processing Lab, ICyTE, Universidad Nacional de Mar del Plata, Argentina.

[3] Group of Informatics and Health, Universidad FASTA, Mar del Plata, Argentina.

[4] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.

[5] Instituto Radiológico, Mar del Plata, Argentina.

E-mail: gmeschin@fi.mdp.edu.ar

Abstract. Evaluation and diagnosis of diseases of the muscles within the rotator cuff can be done using different modalities, being the Magnetic Resonance the method more widely used. There are criteria to evaluate the degree of fat infiltration and muscle atrophy, but these have low accuracy and show great variability inter and intra observer. In this paper, an analysis of the texture features of the rotator cuff muscles is performed to classify them and other tissues. A general supervised classification approach was used, combining forward-search as feature selection method with kNN as classification rule. Sections of Magnetic Resonance Images of the tissues of interest were selected by specialist doctors and they were considered as Gold Standard. Accuracies obtained were of 93% for T1-weighted images and 92% for T2-weighted images. As an immediate future work, the combination of both sequences of images will be considered, expecting to improve the results, as well as the use of other sequences of Magnetic Resonance Images. This work represents an initial point for the classification and quantification of fat infiltration and muscle atrophy degree. From this initial point, it is expected to make an accurate and objective system which will result in benefits for future research and for patients' health.

1. Introduction

Evaluation and diagnosis of diseases of the muscles within the rotator cuff can be done using different medical imaging modalities. Magnetic Resonance Imaging (MRI) is the method more widely used, though other types are also chosen, such as Computed Tomography (TC) and Ultrasound Imaging (US) [1-2]. In order to assess intramuscular fatty infiltration and muscular atrophy, there exist currently different classification criteria, which try to quantize the stage of such pathologies. These criteria have a lack of precision, since they present inter- and intra-observer variability. This is mainly because scales used to describe the stages are not sensitive to little changes in the muscle characteristics and by the other hand, scales are subjective and qualitative [3]. A wide used classification that has shown to be useful for clinical diagnosis comes from Goutallier [4], where a degree of fatty infiltration is defined, assigning 0 when no fat is found, 1 when there are only a few fat lines, 2 if there are more muscle than fat, 3 if the same amount of fat than muscle is observed, and 4

when fat exceeds muscle. Fatty infiltration correlates with the risk of failure in tendon regeneration when a shoulder arthroplasty is practiced [5].

Authors have proposed methodologies to improve the described stage classification with good results, but presenting the drawback of requiring a complex task for the study operator (generally the technician or the image specialist) [6]. Considering this drawback, in this work we propose a preliminary design and implementation of a software tool to contribute to the rotator cuff muscle assessment, in order to make this task relatively easy, reproducible and reliable. We propose MRI texture descriptors to discriminate tissues, including muscles having different stages of fatty infiltration.

By direct observation, the muscles of the rotator cuff present different textures in MRI. Therefore, grey level histograms of these images are not bimodal or multimodal for tissue discrimination. As a consequence, thresholding techniques and other global methods are not applicable in this case. It is necessary to consider a processing pipeline including texture feature extraction and pattern recognition in order to analyse the image in a pixel wise basis [7-9].

In this paper we present an analysis of the features that better represent the texture of the rotator cuff muscles and other tissues that will be applied in later stages of the software development. Selected texture features will be used in the ultimate method that will classify or quantify fatty infiltration and muscle atrophy. A good and reliable performance of this method strongly depends on the robustness of the selected features, so the importance of this preliminary study is well established.

We propose a typical supervised classification pipeline, where feature selection was accomplished by *forward-search* and tissue classification was performed by *k-Nearest Neighbours* (kNN) method [10-11]. In this preliminary study, we considered features extracted in T1-weighted and T2-weighted MRI separately, i.e., we do not combine features of both type of MRI simultaneously. We applied different texture features paradigms, such as grey level co-occurrence matrices, granulometry, spectral features and fractal dimensions [9]. We processed small regions of interest (samples) selected by Imaging Specialists, who marked them in the image and labelled each one. These samples were used as Gold-Standard for assessing the performance of the classification algorithm in the feature selection process. In addition, we segmented MRI images in order to visually analyse the discrimination abilities of the selected features.

2. Materials

Imaging Specialists selected 20 IRM typical studies considering T1- and T2-weighted images and other sequences generally used for diagnosis purposes (PD, TIR, TRA y SAG), with typical acquisition parameters. Different sequences were taken considering slices exactly in the same patient position, trying the patient to be as still as possible. Images were acquired following the direction of the rotator cuff muscles. In figure 1, a typical case is shown, considering T1- and T2-weighted images.

A simple software tool was developed to be used by the Imaging Specialists to make as simple as possible the regions of interest and their labelling. Elliptical regions were considered, as they naturally adapt to the intuitive concept of "region of interest" that Specialists mark when they refer to a part of an image. While selecting a region of interest in a T1-weighted image, exactly the same region was taken in the corresponding T2-weighted image and both were equally labelled. Region selection was done assuring that each one contained only one type of tissue (class). In addition, more than one region was selected for each class if it was possible. As a result, a set of small labelled images (T1 and T2) was stored. As we did not restricted the region size, we got different sized images, with width and height varying between 20 and 100 pixels, containing the useful elliptical region centred and the remaining pixels in black (zero grey level).
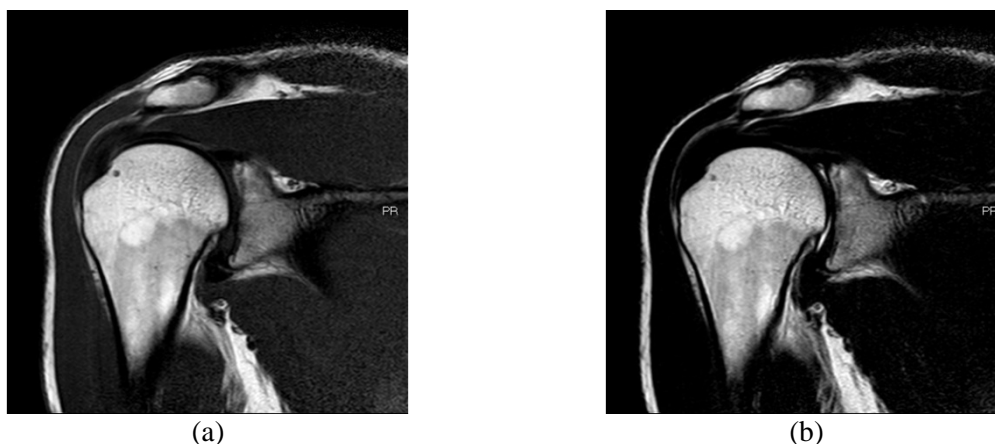
| (a) | (b) |

Figure 1. Typical rotator cuff IRM. a) T1-weighted image. b) T2-weighted image.

These images represent the study object for the analysis we are presenting. Each one contains pixels (and their neighbourhood) corresponding only to one tissue. Imaging Specialists indicated 9 tissues, including muscles with different fatty infiltration, described in table 1. In this sense, as it can be noted, classes #6 to #9 are particularly of interest.

Table 1. Different rotator cuff tissues, selected by Imaging Specialists.

Class	Description
#1	Fat.
#2	Cortical Bone.
#3	Medullary bone.
#4	Generic fluids.
#5	Marrow bone.
#6	Muscle.
#7	Muscle with fatty infiltration.
#8	Muscle with discrete fatty infiltration.
#9	Muscle with low fatty infiltration.

Two different databases were arranged, one containing regions coming from T1-weighted images and the other regions coming from T2-weighted images. Feature selection for optimal discrimination of tissues was made independently for each database, so assessing indirectly which sequence is better in order to recognize each tissue.

In Section 3 we present details of the methods proposed for feature selection and for estimation of the generalization error in the classification.

3. Method
As we described in previous sections, texture features were preliminary studied in order to find the best ones to discriminate IRM pixels corresponding to rotator cuff tissues, including muscles with different fatty infiltration.

Two databases were achieved for different regions of interest coming from 9 tissue classes, keeping T1 and T2 images separately. The same method was used to data processing, considering the next stages: *1) Texture feature extraction, 2) Arrange dataset to be used in a supervised way (feature vectors and labels), 3) Feature selection and generalization error estimation.*

During the first stage, IRM regions of interest were processed, computing texture features considering different paradigms [7-9], analysing only pixels inside the ellipses. After that, a dataset was built, joining together the feature vectors with the tissue label assigned to the corresponding region of interest. Finally, a feature selection was done, trying to get the lower classification error, by

means of a *forward-search* algorithm, and kNN for classification [10-11].

The remaining parts of this section present these steps in detail.

3.1. Texture feature extraction.

Texture recognition is one of the main approaches for computationally recognizing structures of objects in a digital image [8]. Texture is defined in a region of mutually-related pixels that follow a regular pattern throughout an image. Therefore, textures cannot be descripted only considering the grey intensity of a pixel: the grey levels of pixel neighbourhood must be considered, previously determining some type of neighbourhood.

A lot of techniques to characterize textures can be found. Detailed descriptions of different proposed approaches can be consulted in [8] and [9].

In this work, we characterized T1 and T2 MRI textures by four types of texture descriptors: grey level co-occurrence matrices, granulometry, spectral features and fractal dimensions. T1 and T2 sub-images were processed separately. Features extraction was done in the next way:

1. Each MRI sub-image was processed in a pixel wise basis. Since these sub-images are small, we considered 11x11 squared neighbourhoods surrounding each pixel being characterized. We computed features following next considerations:

   - *Grey level co-occurrence matrices:* we made 32 grey levels compression previous to the co-occurrence computing. We considered grey level transitions between each pixel and its neighbours in a -45° direction. From the co-occurrence matrices we computed 22 descriptors, which definitions can be found in [9].
   - *Granulometry:* we used 7 opening operations with a cross-shaped structure element. We computed the granulometric distribution, considering its grey levels, mean values and variance as features.
   - *Spectral features:* we analysed the entropy within 9 frequency bands computing the frequency transform for each pixel neighbourhood.
   - *Fractal dimensions:* two Hurst exponent estimations were computed, as suggested by Comas *et al.* in a previous work [8].

2. After feature extraction, only pixels inside the ellipses were considered, taking only pixels whose neighbourhood of 11x11 pixels was entirely inside the ellipse.

As mentioned before, same regions of interest were simultaneously taken from T1 and T2 images. As a consequence, both databases had the same number of MRI sections for each tissue, having both the same size and the same features were computed for each one. Considering this all, we achieved 96782 feature vectors for each database, which were linearly normalized in the [-1, 1] interval.

The number of samples for different tissues was not balanced. Relative frequencies for tissues are shown in table 2 for both databases.

Table 2. Frequency of classes in training dataset.

Class	Frequency
#1	8.55%
#2	0.08%
#3	3.20%
#4	0.01%
#5	31.01%
#6	27.25%
#7	5.03%
#8	13.50%
#9	11.38%

3.2. Dataset arrangement, to be used in a supervised way (feature vectors and labels).

Once features were computed, two labelled datasets were built, $\{X_{T1}, Y_{T1}\}$ and $\{X_{T2}, Y_{T2}\}$, where X_{T1} and X_{T2} are the feature vectors extracted from T1 and T2 MRI respectively; Y_{T1} and Y_{T2} are the tissues (classes) for each feature vector in X_{T1} and X_{T2}.

Formally for the T1 dataset: $X_{T1} = \{x_{T1,i}\}_{i=1,...,N}$, $Y_{T1} = \{y_{T1,i}\}_{i=1,...,N}$, $x_{T1,i} \in [-1,1]^{35}$ is the i feature vector in the dataset, $y_{T1,i} \in \{1,2,3,...,9\}$ represents the label for the feature vector i (only one for the 9 tissues considered) y $N=96782$. Similarly, for the T2 dataset: $X_{T2} = \{x_{T2,j}\}_{j=1,...,N}$, $Y_{T2} = \{y_{T2,j}\}_{j=1,...,N}$, $x_{T2,j} \in [-1,1]^{35}$ is the j feature vector in the dataset and $y_{T2,j} \in \{1,2,3,...,9\}$ represents its label, with $N=96782$.

Both dataset $\{X_{T1}, Y_{T1}\}$ and $\{X_{T2}, Y_{T2}\}$ were used in stage 3 for feature selection, in order to find the best features to recognize the different classes.

3.3. Feature selection and generalization error estimation.

In this stage, the labelled datasets $\{X_{T1}, Y_{T1}\}$ y $\{X_{T2}, Y_{T2}\}$ generated in stage 2 were separately processed. For each dataset, an optimal-feature search was performed, obtaining a feature subset that achieves the less global error for data (pixels) classification (assignment to one tissue).

The *forward-search* algorithm was used for optimal-feature search, using kNN as classification algorithm [10-11]. During the iterations of this process, we took special care in generalization abilities for the method; that is why we estimated error for each feature combination using *K-fold cross validation* (*K*=10) [10]. For each *K-fold* iteration step, the number of samples for different classes in each fold was balanced by a random sampling with replacement.

Different optimal feature combinations were achieved for each dataset, considering 3, 5, 7 and 9 nearest neighbours for kNN. As a result, we found two optimal feature sets, one for T1 MRI and the other one for T2 MRI.

Considering the best features, we computed the classification results using confusion matrices [8, 12], considering *10-fold cross validation* again, and the optimal number of neighbours in each case. Confusion matrices allow knowing how many pixels have been correctly labelled by the algorithm and in addition, they allow analysing how the errors are propagated for erroneously classified pixels.

The ratio of number of pixels detected for a tissue to the number of pixels corresponding actually to that tissue is computed. The biggest values are expected to be in the main diagonal. These values can be used to compute the classification Accuracy [8]:

$$Accuracy = \frac{\sum_{i=1...\lambda} c_{ii}}{\sum_{i=1...\lambda,\, j=1...\lambda} c_{ij}} \tag{1}$$

where $c_{ij}, i=1,...,\lambda, j=1,...,\lambda$ are the elements of the confusion matrix (number of pixels) and λ represents the number of classes (9 for this work). Each element c_{ij} can be interpreted as "the number of pixels that being i-class were classified as j-class."

4. Results and discussion

In this section, results obtained using optimal features for both T1 and T2 datasets are shown.

Tables 3 and 4 show the confusion matrices obtained after the classification process for T1 and T2 datasets, respectively. They allow analysing the percentage of pixels that belonged to one class that were been classified correctly, and also allow to know to what class they were wrongly assigned. Elements in the main diagonal of the confusion matrices are shaded to make easier the interpretation.

After feature selection, the best combination for T1 dataset had 25 features and the best for T2 dataset had 20 features. For both datasets a value of $k = 3$ for the parameter of kNN algorithm was

obtained. Features selected are detailed in table 5.

We found the highest values in the main diagonal of the confusion matrices. This allows inferring that selected features and the classification algorithm proposed are both successful to solve the problem of tissue recognizing in rotator cuff MRI.

Table 3. Confusion matrix for T1 MRI applying the kNN algorithm with $k = 3$.

	... #1	... #2	... #3	... #4	... #5	... #6	... #7	... #8	... #9	Sum
Pixels being Fat (Class #1) that were classified as...	0.85	0.00	0.01	0.00	0.14	0.00	0.00	0.00	0.00	1.00
Pixels being Cortical Bone (Class #2) that were classified as...	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Pixels being Medullary Bone (Class #3) that were classified as...	0.06	0.00	0.76	0.00	0.18	0.00	0.00	0.00	0.00	1.00
Pixels being Generic fluids (Class #4) that were classified as...	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00
Pixels being Marrow bone (Class #5) that were classified as...	0.02	0.00	0.01	0.00	0.97	0.00	0.00	0.00	0.00	1.00
Pixels being Muscle (Class #6) that were classified as...	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.01	0.00	1.00
Pixels being Muscle with fatty infiltration (Class #7) that were classified as...	0.00	0.00	0.00	0.00	0.05	0.01	0.86	0.03	0.05	1.00
Pixels being Muscle with discrete fatty infiltrations (Class #8) that were classified as...	0.00	0.00	0.00	0.00	0.00	0.04	0.01	0.89	0.06	1.00
Pixels being Muscle with low fatty infiltration (Class #9) that was classified as...	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.08	0.89	1.00

Table 4. Confusion matrix for T2 MRI applying the kNN algorithm with $k = 3$.

	... #1	... #2	... #3	... #4	... #5	... #6	... #7	... #8	... #9	Sum
Pixels being Fat (Class #1) that were classified as...	0.87	0.00	0.01	0.00	0.12	0.00	0.00	0.00	0.00	1.00
Pixels being Cortical Bone (Class #2) that were classified as...	0.00	0.99	0.00	0.00	0.00	0.01	0.00	0.00	0.00	1.00
Pixels being Medullary Bone (Class #3) that were classified as...	0.03	0.00	0.81	0.00	0.16	0.00	0.00	0.00	0.00	1.00
Pixels being Generic fluids (Class #4) that were classified as...	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00
Pixels being Marrow bone (Class #5) that were classified as...	0.02	0.00	0.01	0.00	0.97	0.00	0.00	0.00	0.00	1.00
Pixels being Muscle (Class #6) that were classified as...	0.00	0.00	0.00	0.00	0.00	0.96	0.00	0.04	0.00	1.00
Pixels being Muscle with fatty infiltration (Class #7) that were classified as...	0.01	0.00	0.00	0.00	0.06	0.00	0.85	0.02	0.06	1.00
Pixels being Muscle with discrete fatty infiltrations (Class #8) that were classified as...	0.00	0.00	0.00	0.00	0.00	0.14	0.01	0.80	0.06	1.00
Pixels being Muscle with low fatty infiltration (Class #9) that was classified as...	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.08	0.88	1.00

For T1 images it can be observed that Cortical Bone, Generic fluids and Muscle classes are detected with accuracy higher than 97%. Bone Marrow, Fat and muscle with different fatty infiltration classes are detected with a slightly lower accuracy, but keeping values higher than 85%. The highest error can be observed for the Medullary Bone class (accuracy of 76%), but analysing the propagation of the error to other classes, we found that most of them are assigned to Bone Marrow (18%), being this a similar tissue, and to Fat (6%). These classes are not included in the main objectives of this work.

For T2 images, Cortical Bone, Generic fluids, Bone Marrow and Muscle classes are detected with accuracy higher than 97%. It is quite similar to T1 images, adding the Bone Marrow to the best recognized tissues. Following the same analysis for T1 images, we can note that all the remaining

classes are detected with accuracy higher than 80%. One of the tissues that presents the highest error is again the Medullary Bone (accuracy of 81%), keeping the same behaviour than observed in T1 images: the error propagation is to Bone Marrow class (16%) and Fat class (3%). In this case, the lower accuracy is 80% and corresponds to Muscle with discrete fatty infiltrations.

Results for T2 images seem to be in general better than for T1 images, but it depends on the types of tissues that are priority expected to be low-error detected.

Table 5. Optimal features obtained for T1 and T2 MRI.

MRI	Quantity	Description
T1	25	*Co-occurrence matrix:* Energy, Homogeneity, Entropy, Homogeneity 0.01, Mean value, Variance, Cluster Shade, Cluster Prominence, Correlation, Mean of sum, Entropy of sum, Variance of sum, Variance of difference, Correlation measure, Maximum probability, Order-2 differential inverse moment, Directivity, Dissimilarity, Mean probability. *Granulometry:* Central pixel intensity, Mean value of the granulometric distribution, Variance of the granulometric distribution. *Spectral features:* Lower frequency band entropy. *Fractal models:* both Hurst coefficient estimations.
T2	20	*Co-occurrence matrix:* Energy, Entropy, Homogeneity 0.01, Mean value, Variance, Cluster Shade, Cluster Prominence, Correlation, Mean of sum, Entropy of difference, Variance of sum, Correlation measure, Maximum probability, Order-2 differential inverse moment, Dissimilarity, Mean probability. *Granulometry:* Central pixel intensity, Mean value of the granulometric distribution, Variance of the granulometric distribution. *Spectral features:* no features. *Fractal models:* second Hurst coefficient estimation.

Values observed in the confusion matrices show good accuracy for all classes. Overall accuracy is 93% for T1 images and 92% for T2 images. Analysis of the confusion matrices is more detailed than these general values.

It is reasonably expected that combining what we obtained from both T1 and T2 datasets results would improve, which is the aim for immediate future works. The same is expected considering other MRI sequences.

In future work we will tackle the problem of computing new quantification indices for different stages of fatty infiltration.

Finally, segmentation of whole MRI images are shown for T1 and T2 in figure 2, as examples, with the only objective of visually analyse the discrimination ability of the selected features. Wrong regions can be observed, but we expect to improve these results considering T1 and T2 taken together.

For the images in figure 2, the label assignation method for kNN was modified. Considering $k = 3$, the optimal value obtained for T1 and T2, given a feature vector for labelling, a class was assigned only if the same is observed at least 2 times between its 3 nearest neighbours. If it is not so, the feature vector (and consequently the pixel it represents) are marked as "Unknown pixels."

These images are included in order to show preliminarily a possible future work and they will be analysed by Specialists. There are wrong labelled regions, mainly medullary bone, cortical bone and fat. However, muscle and its different fatty infiltrations were acceptably recognized.

5. Conclusions

In this work we presented a preliminary study of the texture features to best represent rotator cuff muscles with different fatty infiltration and other tissues, considering T1- and T2-weighted MRI.

Starting from a labelled dataset coming from elliptical regions of interest classified by Specialists, an optimal set of features was obtained to discriminate the tissues found in this type of images. At the same time, it was shown that the kNN algorithm is a suitable method for this application.

Using these features and algorithm, we obtained accuracies higher than 90% for the main tissues considered, analysing by means of confusion matrices the error cases.

This work is the starting point for future segmentation of rotator cuff MRI and for the

quantification of the stage of fatty infiltration and muscle atrophy. Efficacy and performance of the future methods will depend on the robustness of the features used, which was the motivation of the present work.
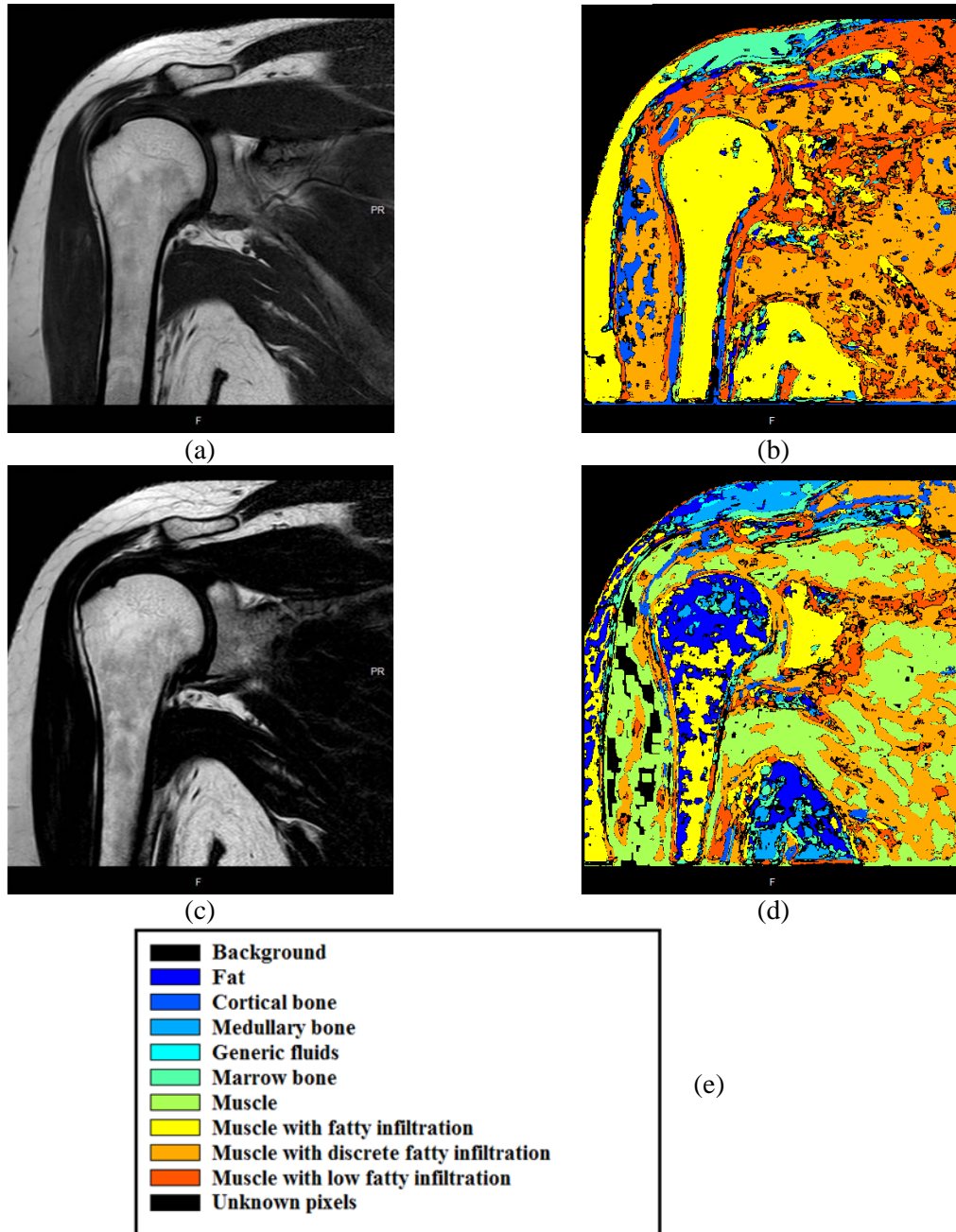

Figure 2. Preliminary segmentation for MRI obtained using the dataset coming from regions of interest as prototype feature vectors. a) Original T1-weighted MRI. b) Segmentation obtained for T1-weighted MRI, *k*=3. c) Original T2-weighted MRI. d) Segmentation obtained for T2-weighted MRI, *k*=3. e) Colour reference for tissues.

By achieving the future objective, we will accomplish an objective and easy-to-use clinical decision support system. It will be useful for assessing the grade of muscular atrophy induced by fat growing in different stages. This will result in benefits for future research and for patients' health.

References

[1] van de Sande M A, Stoel B C, Obermann W R, Tjong a Lieng J G, Rozing P M 2005 Quantitative Assessment of Fatty Degeneration in Rotator Cuff Muscles Determined With Computed Tomography *Invest. Radiol.* 40 **5** 313-9

[2] Khoury V, Cardinal E, Brassard P 2008 Atrophy and Fatty Infiltration of the Supraspinatus Muscle: Sonography Versus MRI *Am. J. Roentgenol.* 190 **4** 1105-11

[3] Spencer E, Warren R, Dunn R, Wright R, Wolf B, Spindler B, Kurt P, McCarty E, Benjamin E, Grant J, Safran M, Brian H, Kuhn J 2008 Interobserver Agreement in the Classification of Rotator Cuff Tears Using Magnetic Resonance Imaging *Am. J. Sports Med.* 36 **1** 99-103

[4] Santoro Belangero P, Ejnisman B, Arce G 2013 A Review of Rotator Cuff Classifications in Current Use in: *Proc. of the ISAKOS Upper Extremity Committees 2009-2013* pp 5-13

[5] Zanneti M, Gerber C, Hodler J 1998 Quantitative Assessment of the Muscles of the Rotator Cuff with Magnetic Resonance Imaging *Invest. Radiol.* 33 **3** 163-170

[6] Gerber C, Schneeberger A, Hoppeler H, Meyer D 2007 Correlation of atrophy and fatty infiltration on strength and integrity of rotator cuff repairs: A study in thirteen patients *J. Shoulder Elbow Surg.* 16 **6** 691-6

[7] Gonzalez R C and Woods R E 2002 *Digital image processing* (Upper Saddle River, N. J.) vol 1 (Prentice Hall)

[8] Comas D S, Meschino G J Ballarin V L 2010 Framework de segmentación y análisis de imágenes mediante reconocimiento de texturas in: *Proc. Argentinian Symposium of Technology (AST 2010)* (Buenos Aires, Argentina) pp. 1529-41

[9] Moler E G, Ballarin V L, Gonzalez M L 1999 Características Estadísticos, Espectrales y Morfológicos para Clasificación de Texturas: Un Análisis Comparativo en: *Proc. VIII RPIC Reunión de Trabajo en Procesamiento de la Información y Control* (Mar del Plata, Argentina) pp. 59-62

[10] Duda R O, Hart P E, Stork D G 2001 *Pattern Classification* (Wiley-Interscience)

[11] Rajini N H 2011 Classification of MRI brain images using k-nearest neighbor and artificial neural network in: *Proc. 2011 International Conference on Recent Trends in Information Technology*, (Chennai, India) pp. 563-568

[12] Kohavi R and Provost F 1998 Glossary of Terms *Mach. Learn.* 30 **2-3** 271-4