

Journal of Biomedical Optics

SPIDigitalLibrary.org/jbo

Neuromuscular disease classification system

Aurora Sáez
Begoña Acha
Adoración Montero-Sánchez
Eloy Rivas
Luis M. Escudero
Carmen Serrano



SPIE

Neuromuscular disease classification system

Aurora Sáez,^a Begoña Acha,^a Adoración Montero-Sánchez,^b Eloy Rivas,^c Luis M. Escudero,^b and Carmen Serrano^a

^aUniversity of Seville, Department of Signal Theory and Communications, ETSI, 41092, Seville, Spain

^bRocío/CSIC/Universidad de Sevilla, Hospital Universitario Virgen del, Instituto de Biomedicina de Sevilla, 41013, Seville, Spain

^cHospital Universitario Virgen del Rocío, Department of Pathology, 41013, Seville, Spain

Abstract. Diagnosis of neuromuscular diseases is based on subjective visual assessment of biopsies from patients by the pathologist specialist. A system for objective analysis and classification of muscular dystrophies and neurogenic atrophies through muscle biopsy images of fluorescence microscopy is presented. The procedure starts with an accurate segmentation of the muscle fibers using mathematical morphology and a watershed transform. A feature extraction step is carried out in two parts: 24 features that pathologists take into account to diagnose the diseases and 58 structural features that the human eye cannot see, based on the assumption that the biopsy is considered as a graph, where the nodes are represented by each fiber, and two nodes are connected if two fibers are adjacent. A feature selection using sequential forward selection and sequential backward selection methods, a classification using a Fuzzy ARTMAP neural network, and a study of grading the severity are performed on these two sets of features. A database consisting of 91 images was used: 71 images for the training step and 20 as the test. A classification error of 0% was obtained. It is concluded that the addition of features undetectable by the human visual inspection improves the categorization of atrophic patterns. © 2013 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.18.6.066017]

Keywords: segmentation; watershed; fuzzy classification; feature extraction; neuromuscular disease; graph theory.

Paper 130052R received Jan. 30, 2013; revised manuscript received May 13, 2013; accepted for publication May 20, 2013; published online Jun. 26, 2013.

1 Introduction

Neuromuscular diseases cover a large group of pathologies with a huge heterogeneous etiology and course. Although all neuromuscular diseases are progressive in nature, they manifest in a wide age range with different degrees of severity. The most common classification relies on the component of the neuromuscular system that is affected. Hence, we can find myopathies, in which the muscle is primarily affected, or neurogenic atrophies (NA), in which the motor neuron is affected. To study the patient affections, the pathologist examines muscular biopsies under microscope focusing into a morphological muscular fiber analysis. The evaluation of the changes in the morphological characteristics of a given biopsy with respect to the normal muscle is one of the main features for the diagnostics of a neuromuscular disease. However, the manual morphometric approach and interpretation of muscle biopsy material is a subjective, tedious, and time-consuming task.¹

Currently, a tissue histopathology slide can be digitized and stored in digital-image form. This progress has allowed the development of computer-assisted diagnosis for disease detection, diagnosis, and prognosis prediction to complement the opinion of the pathologist.² In this sense, in the recent literature works related to grading of prostate cancer,³ detection of cervical cancer,⁴ classification of hepatocellular carcinoma,⁵ detection of cervical cell nuclei,⁶ or simple detection of different types of cells⁷⁻⁹ can be found. Focusing on the studies of

muscular fibers, we can find some works that address the segmentation of fibers in muscular biopsies,¹⁰⁻¹² the classification of muscle-fiber type,¹³⁻¹⁶ and the extraction of morphometric features.¹⁷ However, studies of the characterization of neuromuscular disease based on image processing have not been found in the current literature.

This paper presents an analysis of fluorescence microscopy images of muscular biopsy to obtain useful information for the diagnosis and the severity grading of different neuromuscular diseases. To achieve this, our database consists of 91 images belonging to control biopsies (no disease), biopsies affected by muscular dystrophies (MD), and biopsies affected by NA. The study is based not only on the extraction of features related to the characteristics that the pathologist takes into account for diagnosis, but also on the search of features with inherent properties that escape the evaluation of the pathologist and which could be more efficient for the classification of the different muscular images. In this sense, the paper proposes an extraction of morphometric and structural information based on the assumption that the biopsy is considered as a graph, where the nodes are represented by each fiber, and two nodes are connected if two fibers are adjacent. This was motivated by the work of Escudero et al.,¹⁸ in which the introduction of a network allowed one to describe the epithelial organization objectively. To get the feature extraction, an accurate segmentation was required. In this paper, mathematical morphology and a watershed transform are used to address this task. A fuzzy classification based on a neural network architecture is presented. Finally, a study of the severity grading is carried out to analyze the results. In Sáez et al.,¹⁹ some of these results, analyzed from the biological point of view using different training data sets, are presented.

Address all correspondence to: Aurora Sáez, University of Seville, Department of Signal Theory and Communications, ETSI, 41092, Seville, Spain. Tel: +34 954 486091; Fax: +34 95 448 7341; E-mail: aurorasaez@us.es or Luis M. Escudero, Rocío/CSIC/Universidad de Sevilla, Hospital Universitario Virgen del, Instituto de Biomedicina de Sevilla, 41013, Seville, Spain. Tel: +34 955923048; Fax: +34 95 461 7301; E-mail: lmescudero-ibis@us.es

The rest of the paper is organized as follows: the type of images and the neuromuscular diseases are presented in Sec. 2; in Sec. 3 the procedure followed is explained, which includes a segmentation method, a feature extraction step, a feature selection step, a classification, and the severity grading. In turn, each section contains a subsection of results. Finally, a discussion of the results is presented.

2 Muscle Biopsy Images and Neuromuscular Diseases

The muscular fibers are organized in fascicles, which are surrounded by a layer of connective tissue named the perimysium. Between the fibers within a fascicle appears the endomysium, a mesh of loose connective tissue composed of fine collagen and reticular fibers. Skeletal muscles fibers can be classified into two main types: type I or slow fibers (they have a slow contraction velocity) and type II or fast fibers (fast contraction velocity). They are distributed in a disordered mosaic pattern, and appear with a similar size along the fascicles. The transversal section of a normal muscle represents the fibers with a polygonal shape surrounded by a thin mesh of collagen.²⁰

Muscular biopsies were processed by the standard methods of freezing and cutting with cryostat. The muscular fiber and the collagen content were detected by fluorescence microscopy. The antibodies mouse anti-myosin heavy chain (slow), mouse anti-myosin heavy chain (fast), and rabbit anti-collagen type VI were used using a standard protocol for immunostaining. The sections were incubated with Alexa fluor 488 and Alexa fluor 568 secondary antibodies. All the slides were analyzed under a fluorescence microscope (BX-61 Olympus with a DP70 camera) using

a mercury lamp through a 470 to 490 nm or 560 to 579 nm band-pass filter to excite Alexa fluor 488 or Alexa fluor 568, respectively. The stained cells were photographed, and two high-resolution images (size of 4080×3072 pixels) were obtained. A total of 91 images from 70 muscle biopsies, stored in the Tissue Bank of the Hospital Universitario Virgen del Rocío, Seville, Spain, were processed.

An RGB image was created from the two images obtained. The red component is the image obtained when the biopsy is excited at 560 to 579 nm [Fig. 1(b)], and the green component is the image obtained when the biopsy is excited at 470 to 490 nm [Fig. 1(c)]. Figure 1(a) shows a sample of the resulting RGB image. Slow fibers in a dark color, fast fibers in a reddish color, collagen in a greenish color, and capillaries as small dark structures among the collagen can be observed.

All original images have the same resolution (4080×3072 pixels). However, the figures shown in this paper represent only a part of the complete image (1150×1150 pixels) to correctly visualize the details. The yellow bar at the top of Fig. 1(a) represents the image scale corresponding to $200 \mu\text{m}$.

In Fig. 1(d)–1(f), examples of the variability in the images are shown. They belong to the patients with different ages and different diseases.

In this paper, three morphological patterns presented in muscle biopsies are studied. The first one is the normal pattern (no disease). The second is the dystrophic pattern. MD are a type of myopathy, which are characterized by a wide variation in fiber size, a rounded shape of atrophic fibers, and fibrosis (an increase of endomysial collagen). This contrasts with the third pattern, NA that present angulated atrophic fibers, large groups of atrophic fibers, fascicular atrophy, and loss of the random

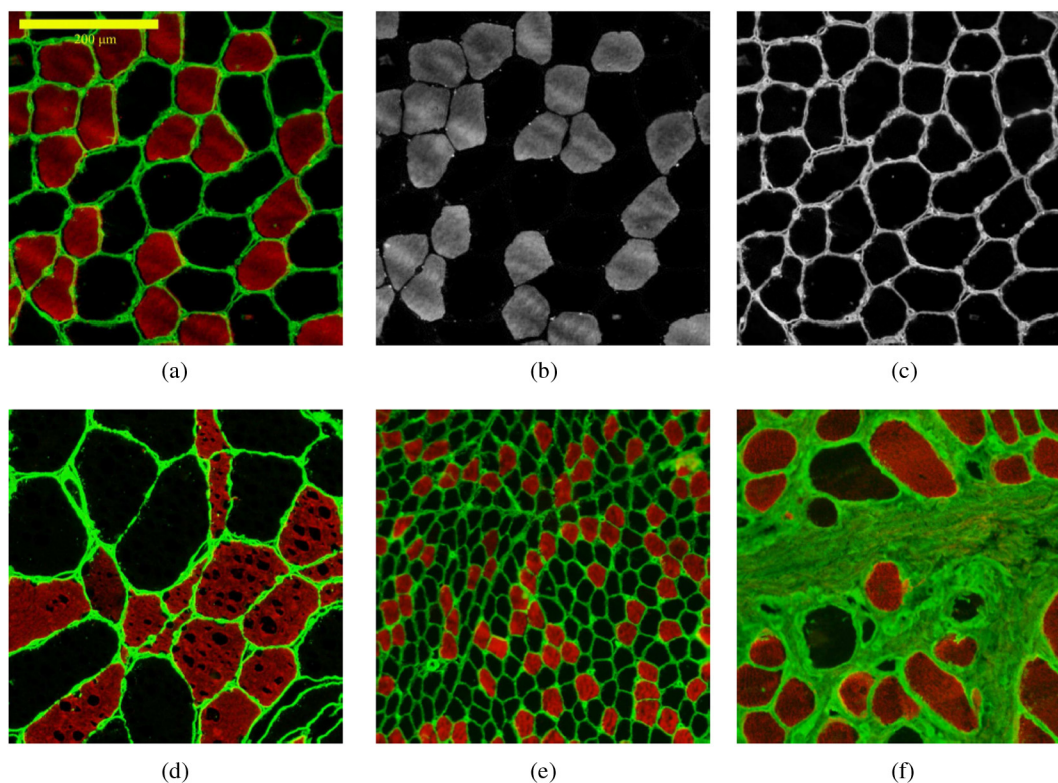


Fig. 1 (a) Muscle biopsy image. The yellow bar at the top represents the image scale. It corresponds to $200 \mu\text{m}$, and it is the same for the rest of the images. (b) R-component. (c) G-component. (d) Example of muscle biopsy image affected by neurogenic atrophy (NA). (e) Example of muscle biopsy image belonging to a child. (f) Example of muscle biopsy image affected by muscular dystrophy (MD).

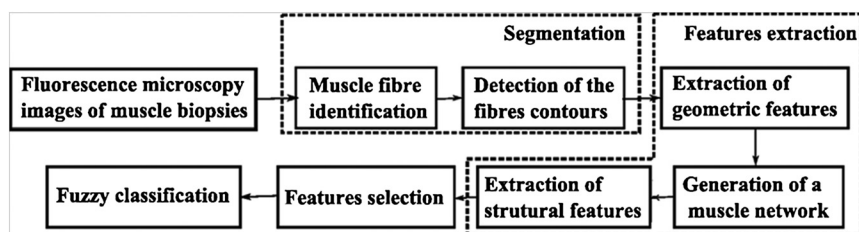


Fig. 2 Flow diagram of the system.

checkerboard distribution of fiber types or fiber-type grouping.²¹ In addition, it is possible to find a muscle that can show a combination of neurogenic and myopathic features. For the study, we have 91 images from 70 subjects; 41 control images, 27 dystrophy images, and 23 atrophy images.

3 Methodology

The morphological analysis is a fundamental tool for the diagnosis of neuromuscular disorders.²² The aim of this paper is to analyze muscle biopsies in a morphological and structural way, which allows one to develop a diagnosis help tool for the neuromuscular diseases explained in the previous section. The procedure followed is described in Fig. 2.

Each block of the flow diagram in Fig. 2 is explained in the following sections.

3.1 Segmentation

A prerequisite to classify any disease is the ability to automatically identify the structures present in the image.² Moreover, to achieve a robust morphological analysis, an accurate segmentation of the fibers in the biopsy image is required. Shape description and accurate segmentation is possible if the initial localization of the muscle fibers is known.²³ Therefore, in this paper the segmentation process is divided into two steps: identification of the muscle fiber localization by applying morphological operators, and accurate detection of the fiber contours by a watershed transformation. It is important to note that the segmentation method must be automatic and valid to all types of images [see Fig. 1(a) and 1(d)–1(f)].

3.1.1 Fiber localization

The biopsy images we work with present different structures such as collagen, muscle fibers, capillaries, and artefacts. The aim of this step is the correct identification of the muscle fibers.

The processing is performed on the G-component of the image due to the high contrast between the muscle fibers and the collagen [see Fig. 1(c)]. Considering that the fibers are darker than the surrounding collagen, intensity valleys in the image are searched. For this reason, the H-minima transform²⁴ is applied to the G-component image in order to get homogeneous minima valleys. This transform has been successfully used in different medical applications.^{25,26} The H-minima or H-maxima transform is a powerful mathematical tool to suppress undesired minima or maxima. It allows one to extract regional minima whose depth is lower than or equal to the given h -value. Regional minima are connected components of pixels with a constant intensity value, and whose external boundary pixels all have a higher value. The H -minima transform²⁴ is performed by

$$H_h(G) = R_G^\varepsilon(G + h), \quad (1)$$

where h represents the given depth, and R and ε represent the reconstruction and erosion operators, respectively. The h -value has a direct influence on the number of segmented regions. The larger the h -value, the fewer segmented regions. As the darkest regions of the G -component image represent the muscles fibers, we can intuitively assume that the required h -value should be lower than the average intensity of the image. Three different values that took into account the average intensity were tested: two thirds of the average intensity of G (h_1), half of the average intensity of G (h_2), and one third of the average intensity of G (h_3). They are calculated as

$$h = K \frac{1}{N \cdot M} \sum_{i=1}^N \sum_{j=1}^M G(i, j), \quad K = \frac{2}{3}, \frac{1}{2}, \frac{1}{3}, \quad (2)$$

where $G(i, j)$ is the intensity value of G-component at the pixel (i, j) and N and M are the image pixel dimensions.

The resulting image is a binary image, in which regions with a pixel value of 1.0, displayed as white, represent candidate muscle fibers. Figure 3 shows the influence of the three h -values on the number of segmented regions.

3.1.2 Results of the fiber localization

Results were tested in only 10 1150×1150 pixel images. The reason for this is that the manual segmentation of each one is necessary in order to check the quality of the method, and each image has hundreds of cells. So, the manual delineation of them is a very time-consuming and tedious task. It should be noted that the number of cells segmented in these test images represents only a portion of the number of cells of the entire image. Table 1 shows the number of cells segmented by the specialist, and the number of the regions detected by the H -minima transform with the three different h -values. The number of detected regions should be close to the number of manually segmented cells, taking into account that the number of the detected regions must be higher than or equal to the number of segmented cells. If the value is lower, it will involve the loss of detected cells.

Although h_1 provides a number of detected regions more similar to the number of cells estimated by the specialist (see Table 1), in some cases the number of detected regions is lower. This means that some cells are not detected. Since the number shown here represents only a portion of the number of cells in the entire image, this number of lost cells could increase. Furthermore, the oversegmentation is due to the existence of artefacts and capillaries in the image. These regions can be removed by using morphological operators. For these

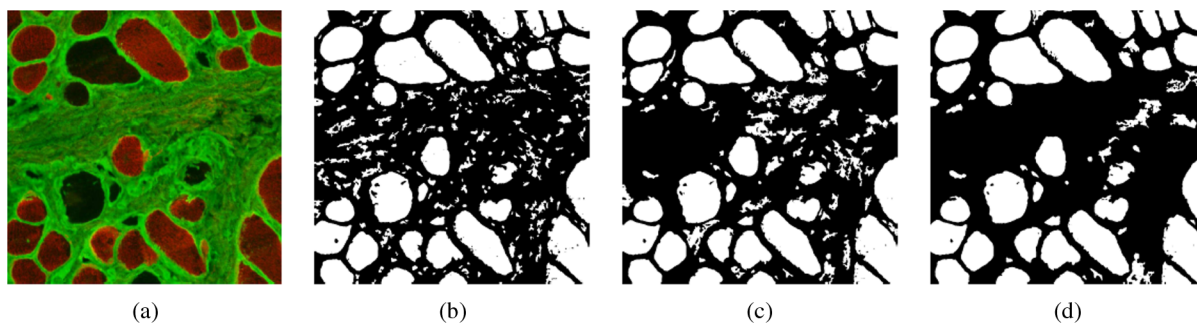


Fig. 3 (a) Muscle biopsy image. (b) H-minima transform with h -value = h_1 ($h_1 = 60.85$). (c) H-minima transform with h -value = h_2 ($h_2 = 45.63$). (d) H-minima transform with h -value = h_3 ($h_3 = 30.42$).

reasons, the chosen h -value is h_2 followed by a processing of the resulting image using mathematical morphology.

The morphological operators used in this step are summarized here. Regions with an area smaller than 0.25% of the biggest dimension in pixels of the original image (taking into account the resolution indicated in Sec. 2) were removed by applying morphological opening.²⁴ Presumably, small regions correspond to capillaries. Irregularities (holes) within the regions detected were refilled by morphological reconstruction.²⁴ Finally, an erosion operator with a 3×3 pixel structural element was applied to prevent that adjacent cells were joined. The results of the number of regions detected after this step are shown in the fourth column of Table 2.

Although the number of regions decreased, it could be desirable to better adjust the number of detected cells and the number of segmented cells. To this aim, finally, two independent color conditions were imposed on the detected regions in order to remove those that did not truly correspond with cells.

First condition: those regions whose average intensity of G-component had a value higher than 25% of the maximum intensity value of G in the database were rejected as candidate fibers, as the muscle fibers present a low green value.

Second condition: a histogram equalization of the G-component was performed, the average intensity of this new image was calculated for all regions, and the maximum of these values was computed. Finally, those regions whose average intensity was higher than the 90% of the maximum and an average value of R -component less than 20% of the maximum intensity value of R in the database were removed.

The thresholds mentioned were experimentally fixed, and provide a correct segmentation of the 91 images analysed.

With these two conditions, capillaries and artefacts present among the collagen are removed. The fifth column of Table 2 presents the final results.

3.1.3 Detection of fiber contours

Once the localizations of the muscle fibers are identified, an accurate detection of their contours is required for a later robust morphological analysis. For this objective, two well-known techniques were applied: level set and a watershed transform. Both methods are briefly explained below.

Level set methods²⁷ have been widely used as a global approach toward the optimization of active contours for the

Table 1 Number of manually segmented cells and number of 5 detected regions by H-minima transform with different h -values.

Image	N cells manual	h_1	h_2	h_3
1	77	76	79	83
2	103	108	128	144
3	40	250	295	366
4	45	256	530	1000
5	56	371	467	574
6	40	61	113	245
7	92	90	105	121
8	49	48	56	60
9	310	377	419	483
10	63	136	155	184

Note: The bold values indicate that number of the regions detected is lower than number of manually segmented cells, this will involve loss of detected cells.

Table 2 Number of manually segmented cells and number of detected regions by H-minima transform with h -value = half of the average intensity of G , after the application of morphological operators and two color conditions.

Image	N cells manual	h_2	Morphological operators	Color conditions
1	77	79	77	77
2	103	128	103	103
3	40	295	59	40
4	45	530	104	48
5	56	467	59	56
6	40	113	47	41
7	92	105	91	91
8	49	56	52	49
9	310	419	333	310
10	63	155	74	65

segmentation of objects of interest from the background. In the literature, numerous works related to this technique have been proposed. In this paper, the method developed by Li et al.²⁸ is used. In the reported work, an energy function tries to maintain the level set function near the signed distance function, thus avoiding the need for re-initialization of the level set function.²⁸ The initial curve required is the contour of the binary image resulting from the muscle fiber localization detection, explained in the previous section.

In the watershed procedure,²⁹ an image is viewed as a topographic surface: the higher the value of a pixel, the higher the altitude at the corresponding point on the topographic surface or relief. The watershed transform is usually applied to the gradient image. The minima in the gradient image will correspond to the sites within homogeneous regions in the original image. However, the watershed algorithm yields results with substantial oversegmentation; that is, the number of segmented regions

could be much larger than desired, with regions being broken into multiple smaller regions. This undesirable result is due to the fact that the gradient image used in the process is sensitive to noise. The problem of oversegmentation can be overcome with the use of markers that identify the objects. The object contours in the gradient image can be seen as the highest crest-lines around the object markers. In our case, the image gradient is calculated in the G -component, and the internal and external markers are derived from the previous section, where the aim was to identify the muscle fibers. The binary image resulting from the previous section constitutes the internal markers. Meanwhile, the external markers were extracted by applying another watershed transform to the binary image used as internal markers.

Table 3 Segmentation results for the watershed transform and level set segmentation evaluated by the Dice coefficient and the Jaccard coefficient.

Image	Dice coefficient		Jaccard index	
	Watershed	Level sets	Watershed	Level sets
1	0.963	0.94	0.927	0.8969
2	0.966	0.93	0.933	0.868
3	0.97	0.95	0.942	0.912
4	0.969	0.952	0.939	0.908
5	0.969	0.949	0.94	0.903
6	0.956	0.938	0.915	0.881
7	0.96	0.917	0.922	0.843
8	0.975	0.947	0.951	0.89
9	0.971	0.917	0.94	0.84
10	0.973	0.922	0.948	0.854
Average	0.967	0.936	0.934	0.879

Note: The bold values indicate the best result.

3.1.4 Results of the detection of the fiber contours

To evaluate the performance of both methods, 10 images manually segmented by the specialist were used. The results were evaluated with the Jaccard coefficient and the Dice coefficient. The Jaccard index, also known as the Jaccard similarity coefficient, is a statistical parameter used for comparing the similarity and diversity of sample sets. It is defined as the size of the intersection divided by the size of the union of the segmented cells. Assuming two sets corresponding to the segmented pixels obtained from the manual segmentation (P_m) and the segmented pixels obtained from the automatic method (P_s), Jaccard's coefficient is defined as

$$J = \frac{|P_m \cap P_s|}{|P_m \cup P_s|}. \quad (3)$$

The Dice coefficient, D , is also a similarity measure, which is defined as

$$D = \frac{2|P_m \cap P_s|}{|P_m| + |P_s|}. \quad (4)$$

Results are presented in Table 3. As can be seen, the watershed transform outperforms level set technique in all images. Some image examples of both segmentations are shown in Fig. 4. It is important to note that besides the accuracy in detecting the contours of the fibers is higher in watershed [Fig. 4(b)] than in level sets [Fig. 4(c)], the watershed transform is able to separate two independent cells although they seem to be linked (see yellow rectangles).

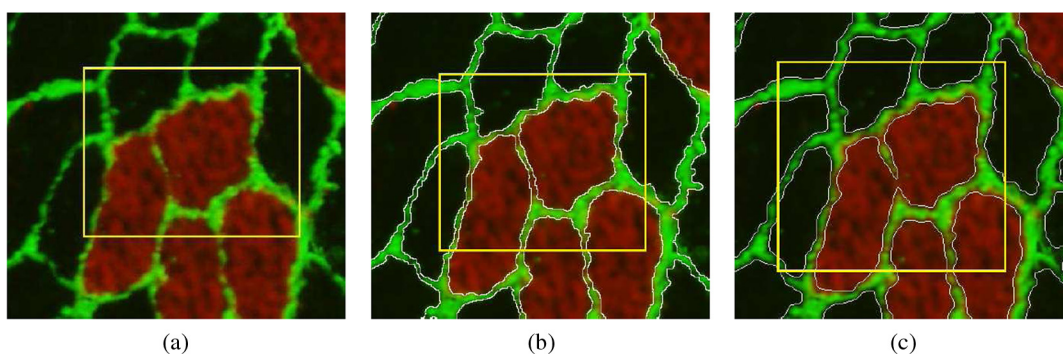


Fig. 4 (a) Muscle biopsy image. The yellow rectangle indicates linked cells. (b) Watershed segmentation result. The yellow rectangle indicates a good result. (c) Level set segmentation result. The yellow rectangle indicates a bad result, because two linked cells have been segmented as only one.

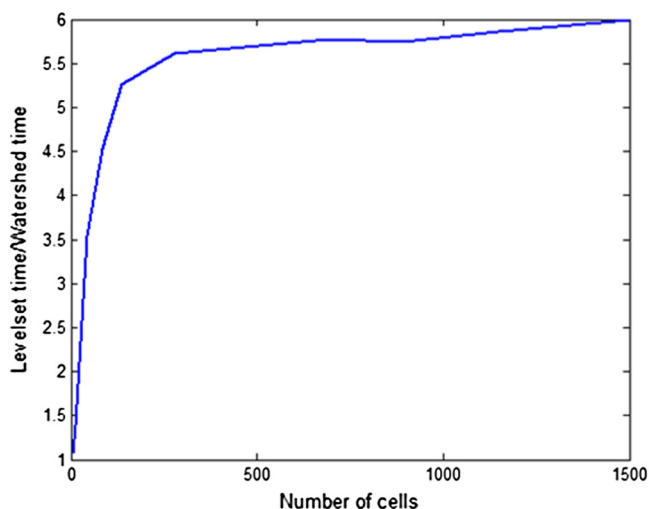


Fig. 5 Ratio between computational cost of both segmentation methods and number of cells segmented.

Furthermore, the computational cost is also lower for watershed. Figure 5 shows that the ratio of computational time between both segmentations (consumed time by level sets/consumed time by watershed) increases logarithmically with the number of cells to segment in the image. Furthermore, it should be noted that the time spent on segmenting the images by the proposed method is significantly less than that needed by the specialist (see Table 4). The steps followed in final segmentation process are shown in Fig. 6.

3.2 Feature Extraction

To identify if a biopsy is affected by a pathology, an objective analysis of the biopsy is needed. The morphological and structural characteristics of the whole biopsy constitute a vector, also called biosignature.²³ The objective of this section is to extract the features that correspond to the visual attributes defined by clinicians

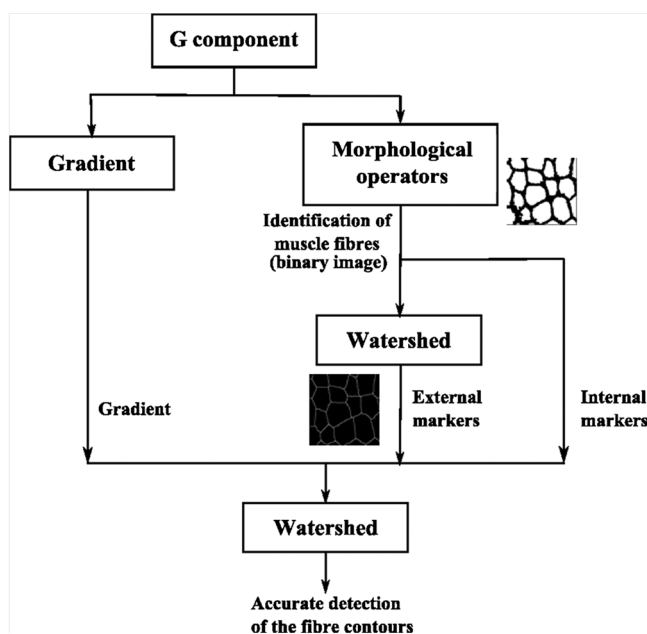


Fig. 6 Steps followed in the fiber segmentation.

Table 4 Time spent on segmenting the images by the proposed method and by the specialist.

Image	Manual segmentation (min)	Proposed method (s)
1	20	33
2	33	35.5
3	14	33
4	22	33.7
5	11	30.14
6	8	30
7	29	41
8	15	34
9	40	45
10	20	33
Average	21.2	31.5

as particularly important for the mentioned pathology grading and diagnosis as well as a study of new features with inherent properties that escape from the evaluation of the pathologist and that could be useful for the characterization of these diseases. In this sense, both morphological and structural features are proposed.

3.2.1 Morphological features

The morphological characteristics can be described by shape or geometry.²³ Formulation of morphological features is an easy

Table 5 Fourteen morphological features of the cells.

1	Average area
2	Std. dev. area
3	Average area of slow cells
4	Std. dev. area of slow cells
5	Average area of fast cells
6	Std. dev. area of fast cells
7	Average major axis
8	Average minor axis
9	Average ratio axis
10	Std. Dev. ratio axis
11	Average convex hull
12	Std. Dev. convex hull
13	Average angles
14	Std. dev. angles

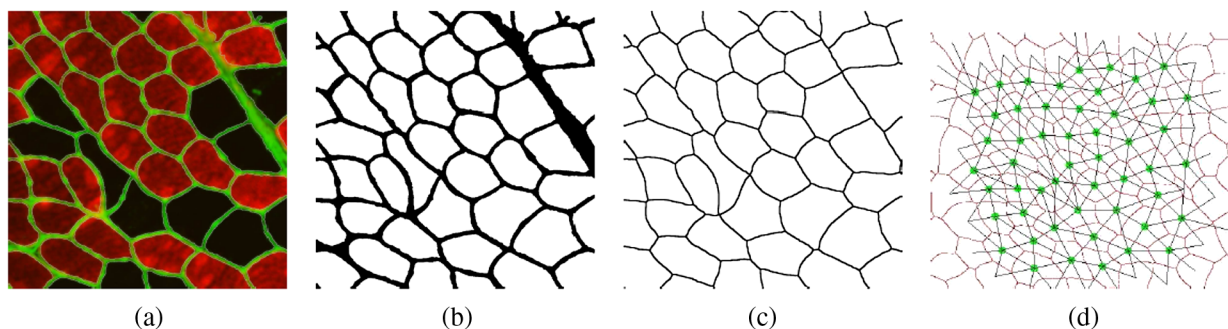


Fig. 7 (a) Muscle biopsy image. (b) Mask of the watershed segmentation result. (c) Mosaic, where each fiber contour is expanded to reach the expanded contour of the adjacent fiber. (d) Graph, where each node is represented by the mass center of each fiber, and the neighborhood relations are mapped into the edges.

and fast way to automatize the manual morphological quantification, which is very laborious and subjective.

Features such as the area of each fiber, the average of the areas of type I (slow) and type II (fast) fibers, identified by the intensity average of R-component, or the major and the minor axis lengths of the cells are calculated. In Table 5, the 14 characteristics computed are shown.

3.2.2 Structural features

To address this approach, each biopsy image was interpreted as a graph. Graphs are efficient data structures to represent spatial data, and an effective way to represent structural information by defining a large set of topological features.² Formally, a simple graph $G = (V, E)$ is an undirected and unweighed graph without self-loops, with V and E being the node and edge set of graph G , respectively. Applying this concept to our problem, we generated a cellular network in which the muscle fibers are represented by nodes, and two nodes are connected if two fibers are adjacent.

To identify the neighborhood of each fiber, we generated a mosaic such that each fiber contour is expanded to reach the expanded contour of the adjacent fiber. This concept was addressed by applying a watershed transform to a binary image resulting from the fiber detection step (Sec. 3.1.3). An example is shown in Fig. 7(c).

Table 6 Structural features.

15	Average ratio A1/A2
16	Std. dev. ratio A1/A2
17	Average neighbors
18	Std. Dev. neighbors
19	Std. dev. neighbors of slow fibers
20	Std. dev. neighbors of fast fibers
21	Slow neighbors of slow fibers
22	Fast neighbors of slow fibers
23	Slow neighbors of fast fibers
24	Fast neighbors of fast fibers

It should be noted that an important feature was extracted from this mosaic. The ratio between the area of a fiber (A2) and the area when its contour is expanded (A1) (features 15 and 16 in Table 6) is indicative of the amount of the collagen, and therefore indicative of the existence of fibrosis, one of the main attributes of the MD.

In this point, a vector of 24 features (included the 14 geometrical ones) was fixed. Characteristics such as the number of neighbors or the number of neighbors for a determined type of fiber were added. Table 6 shows the added features.

These 24 features try to emulate the characteristics that the pathologist takes into account for diagnosis, however, in this paper another set of features that escape from the evaluation of human vision is extracted. For this purpose, a weighted graph derived from the muscular biopsy was generated, where each node was represented by the mass center of the each fiber, and the neighborhood relations are mapped into weighted edges, where each weight corresponds to the Euclidean distance between the nodes [see Fig. 7(d)].

Fifty-eight new characteristics were incorporated. The first 14 characteristics out of 58 ones were obtained from the parameters that were already computed (area, axis, convex hull, angles, and ratio A1/A2), but taking into account the neighborhood of each fiber, i.e., the ratio between the value of the parameters of each fiber and the average of the values of the corresponding adjacent fibers constituting these new 14 features (features 25 to 38 in Table 7). The 44 remaining features are computed from graphs theory (features 39 to 82 in Table 7) when it is applied to an undirected and weighted graph. Table 7 shows these 58 new features.

To avoid errors due to the lack of neighbors of the fibers at the image edge, the characteristics are calculated on a region of interest (ROI) chosen by the users, such that at least one row of fibers around the ROI exists [see Fig. 7(d)].

3.3 Feature Selection

Feature selection has two benefits: it reduces the cost of data collection and computational cost of recognition, and it usually improves the generalization performance of the classifier. Actually, a large set of features may possibly be detrimental to the classification performance, a phenomenon known as “the curse of dimensionality.” Feature selection is a means to select the relevant and important features from a large set of features. An optimal feature selection method would require an exhaustive search, which is not practical for a large set of features generated from a large dataset. Therefore, several heuristic

Table 7 Fifty-eight new structural features.

		60	Std. dev. betweenness centrality of fast cells
		61	Average betweenness centrality of slow cells
25	Average relation neighbors area	62	Std. dev. betweenness centrality of slow cells
26	Std. dev. relation neighbors area	63	Average shortest paths lengths
27	Average relation neighbors major axis	64	Std. dev. shortest paths lengths
28	Std. dev. relation neighbors major axis	65	Average shortest paths lengths from fast cells to fast cells
29	Average relation neighbors minor axis	66	Std. dev. shortest paths lengths from fast cells to fast cells
30	Std. dev. relation neighbors minor axis	67	Average shortest paths lengths from fast cells to slow cells
31	Average relation neighbors relation axis	68	Std. dev. shortest paths lengths from fast cells to slow cells
32	Std. dev. relation neighbors relation axis	69	Average shortest paths lengths from slow cells to slow cells
33	Average relation neighbors convex hull	70	Std. dev. shortest paths lengths from slow cells to slow cells
34	Std. dev. relation neighbors convex hull	71	Average shortest paths lengths from slow cells to fast cells
35	Average relation neighbors angles	72	Std. dev. shortest paths lengths from slow cells to fast cells
36	Std. dev. relation neighbors angles	73	Radius
37	Average relation neighbors ratio A1/A2	74	Diameter
38	Std. dev. relation neighbors ratio A1/A2	75	Efficiency
39	Average strengths	76	Pearson correlation
40	Std. dev. strengths	77	Algebraic connectivity
41	Average strengths of fast cells	78	S metric
42	Std. dev. strengths of fast cells	79	Assortativity
43	Average strengths of slow cells	80	Density
44	Std. dev. strengths of slow cells	81	Transitivity
45	Average clustering coefficient	82	Modularity
46	Std. dev. clustering coefficient		
47	Average clustering coefficient of fast cells		
48	Std. dev. clustering coefficient of fast cells		
49	Average clustering coefficient of slow cells		
50	Std. dev. clustering coefficient of slow cells		
51	Average eccentricity		
52	Std. dev. eccentricity		
53	Average eccentricity of fast cells		
54	Std. dev. eccentricity of fast cells		
55	Average eccentricity of slow cells		
56	Std. dev. eccentricity of slow cells		
57	Average betweenness centrality		
58	Std. dev. betweenness centrality		
59	Average betweenness centrality of fast cells		

algorithms have been developed which use classification accuracy as the optimality criterion.²

In this paper, the well-known feature selection methods, namely sequential forward selection (SFS) and sequential backward selection (SBS),³⁰ are used. SFS works by sequentially adding the feature that most improves the classification performance; similarly, SBS begins with the entire feature set and sequentially removes the feature that most improves the classification performance. While these methods still cannot guarantee optimality of the selected feature subset, they have been shown to perform very well compared with other feature selection methods³¹ and are, furthermore, much more computationally efficient.³²

As it has been mentioned, these methods use classification accuracy as the optimality criterion. In this case, the classification was performed by a Fuzzy-ARTMAP neural network. It is a neural network architecture developed by Carpenter et al.,³³ and it is based on adaptive resonance theory (ART). Fuzzy-ARTMAP is a supervised learning classification architecture for analogue-value input pairs of patterns, where each individual input is mapped to a class label.

To evaluate the classification accuracy, a set of 71 images was used. Thirty-four control biopsies from quadriceps and

biceps, 20 biopsies images affected by dystrophy, and 17 biopsies images affected by NA. Three studies were carried out:

- Comparison between the three groups of images [controls (C)–MD–NA]
- Comparison between control and dystrophies (biopsies of muscle affected by dystrophy belong to quadriceps)
- Comparison between control and NA (biopsies of muscle affected by atrophy belong to biceps)

For each comparison, the selection performance was evaluated by fourfold cross-validation (XVAL).³⁰ In this sense, the disadvantage of sensitiveness to the order of presentation of the training set that the SBS and SFS methods present was diminished. To perform the XVAL method, four disjoint subsets of each class (control, dystrophy, NA) were used. Three of these subsets served as training sets for the neural network, while the other one was used as a validation set. Then, the procedure was repeated interchanging the validation subset with one of the training subsets, and so on, till all four subsets were used as validation sets. The final classification error was calculated as the mean of the errors for each XVAL run.

3.3.1 Results of feature selection

The feature selection procedure was performed twice for the three comparisons mentioned above. The first selection was carried out on the 24 first features described in the previous section (see Tables 5 and 6), and the second selection was performed on the 82 features (see Table 7). The results are shown in Table 8, in which the selected features and the classification error obtained with this selection are shown. The selected features are presented in ascending order by discrimination power.

As can be seen, when we compare between both quadriceps (dystrophies) and biceps (atrophies), the classification success is 100%, therefore, adding new features does not improve the classification error in this stage of training. In the following section, we will study what classification error is obtained when we classify new biopsies, that were not included in the stage of feature selection. However, in the case of the distinction between the three categories, the classification error decreases when we add structural features. Even so, in the next section we check how these sets of features are good to classify new biopsies.

3.4 Classification

The extracted features represent the inputs to a classification procedure. The classifier used in this paper is a Fuzzy ARTMAP neural network.

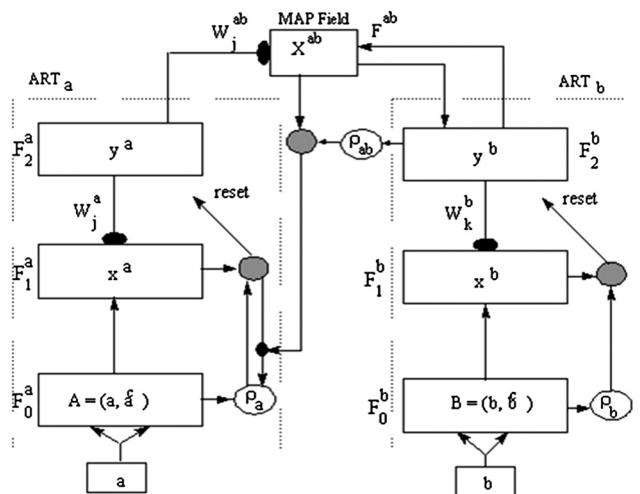


Fig. 8 Fuzzy ARTMAP architecture.

The Fuzzy ARTMAP system incorporates two fuzzy ART modules, ART_a and ART_b , that are linked together via an inter-ART module, F^{ab} , called a map field (see Fig. 8).

In the prediction stage, “ a ” (features) is the input vector, and it is hoped that the system responds with the “ b ” vector (categories). In our implementation, F_2^a is composed by $N \cdot j$ ’th nodes ($j = 1, \dots, N$, $N =$ number of training images) and W_j^{ab} denotes the weight vector associated to the j ’th node of F_2^a . In the prediction stage, the j node of F_2^a is activated throughout the maximum of a choice function T_j .

The W_j^{ab} associated to this node activates a K node of F_2^b . This K node is the category chosen that the system predicts is associated to the “ a ” input. However, in our implementation, the category choice is modified. As each j node in F_2^a is associated to a training image, we have taken into account the average of the values T_j by categories instead of the maximum. Let cl be the category ($cl = 1, \dots, M$, $M =$ number of the categories), in our case, $cl = 1$ represents control, $cl = 2$ represents dystrophy, and $cl = 3$ represents atrophy, respectively. Let n_{cl} be the number of the training images of the cl category. The category choice procedure is:

for $cl = 1:M$
 $k_{cl}(T_j)$, $j = 1: n_{cl}$
 End.

The category chosen fulfills:

$$K_{CL} = \max\{k_{cl} : cl = 1:M\}. \quad (5)$$

Table 8 Feature selection results.

Comparison	24 features		82 features	
	Selected features	Classification error (%)	Selected features	Classification error (%)
Controls-dystrophies (C/MD)	19 18 15	0	25 9	0
Controls-atrophies (C/NA)	12 20 21 22	0	34 16 21 45	0
Controls-dystrophies-atrophies (C/MD/NA)	20 9 19 18 16 21 14 17	13.22	25 32 14 16 58 62 15 30 26	1.48

Table 9 Results of classification 20 images; seven controls, seven dystrophies, five neurogenic atrophies (NA), and one atrophy of pseudo-dystrophic nature.

Comparison	24 features	82 features
	Classification error	Classification error
Controls-dystrophies (C/MD)	100% (15 images/15 images)	100% (15 images/15 images)
Controls-atrophies (C/NA)	92.31% (12 images/13 images)	100% (13 images/13 images)
Controls-dystrophies-atrophies (C/MD/NA)	85% (17 images/20 images)	85% (17 images/20 images)

Each $M k_{cl}$ computed for each input allows one to present a classification called fuzzy, since for each input we obtain a value for each possible category. The system assigns the category (K_{CL}) corresponding to the highest value k_{cl} , however, it also provides an idea of how reliable the output is. It can happen that for a biopsy (input), the response of the system does not provide a clear high value in the k_{cl} 's, and by the contrary provides two similar values for two different categories. This fact implies that the biopsy can present a combination of different pathologies or a preliminary stage of a pathology. Some examples of this concept are presented in the classification results.

3.4.1 Classification results

To evaluate the classification procedure, 20 test images were used: seven controls, seven dystrophy, five NA, and one atrophy with a pseudo-dystrophic component. These images were not part of the training set to select the most discriminant features. Table 9 shows the classification results for the features selected in the previous section (Table 8).

As regards to Table 9, in the classification between C/NA and C/MD/NA, all errors were made in atrophies that were classified as controls. This fact is due to the low degree of severity that the biopsies affected by atrophy present and, therefore, to the great similarity of control biopsies. For the case C/MD/NA, where the three categories are compared, the classification error made with

both sets of features is the same (85%), although in the feature selection procedure the classification error was considerably less for the case of 82 features (see Table 8).

Otherwise, the numeric values (k_{cl}) obtained as output of the neural network could be considered as an indicative of the degree of certainty of belonging to a class. In Table 10, an example of the k_{cl} values for a control biopsy image, a dystrophy with two degrees of severity, an atrophy with zero to one degrees of severity, and an atrophy with pseudo-dystrophic component are shown. The cases that present two similar values for different classes will require a further study of the biopsy.

For these cases, we can give a more reliable result if we apply the classification criterion presented in Table 11. The procedure is to classify each biopsy with the three neural networks trained with C/MD/NA, C/MD, and C/NA. Depending on the category selected by each neural network, the biopsy is classified as Table 11 indicates.

In view of the findings, it seems coherent to think that the use of structural features improves the discrimination between controls and NA (see Table 9). The same conclusions could also be achieved for the case of the triple comparison (C/MD/NA), since the classification error in the feature selection step is lower for the 82 features set. However, when MD are compared with controls, the same classification error for both feature sets is obtained. Therefore, to attain a further analysis, a severity grading is performed as below.

Table 10 Output values of the Fuzzy ARTMAP trained with controls-dystrophies-atrophies from the 82 features when it classifies different biopsies.

Biopsies	Fuzzy ARTMAP output		
	k_1 (controls)	k_2 (dystrophy)	k_3 (atrophy)
Control	0.904	0.757	0.810
Dystrophy with two degrees of severity	0.631	0.813	0.750
Atrophy with zero to one degrees of severity	0.867	0.795	0.855
Atrophy with pseudo-dystrophic nature	0.688	0.860	0.850

Note: The bold values in the two first cases indicate the highest k_{cl} values obtained. This means that the biopsy belongs to this class. For the third and fourth cases there is no value predominantly higher than others. These cases present two similar values for different classes (k_{cl}) indicated in bold. This means that a further study of the biopsy will be required.

Table 11 Classification criterion.

C/MD/NA output	C/MD output	C/NA output	Final categorization
C	C	C	C
C	C	NA	C or low atrophic (requires a study)
NA	C	NA	NA
MD	MD	C	MD
MD	MD	NA	Possible atrophy with dystrophic nature (requires a study)
NA	MD	NA	Possible atrophy with dystrophic nature (requires a study)

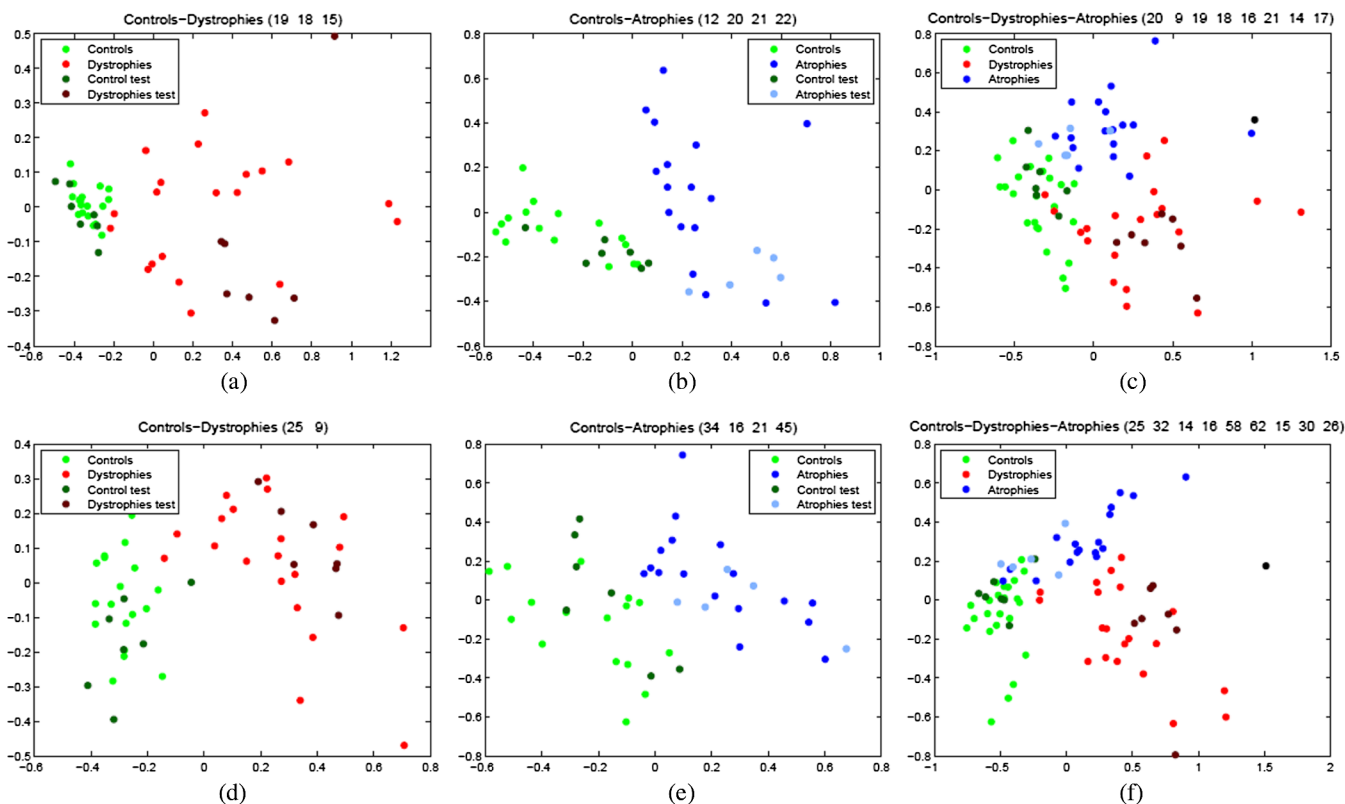


Fig. 9 PCA graphs. (a) Quadriceps controls and dystrophies with the features 19, 18, and 15 as inputs. (b) Biceps controls and atrophies with features 12, 20, 21, and 22 as inputs. (c) Controls, dystrophies, and atrophies with the features 20, 9, 19, 18, 16, 21, 14, and 17 as inputs. (d) Quadriceps controls and dystrophies with features 25 and 9 as inputs. (e) Biceps controls and atrophies with features 34, 16, 21, and 45 as inputs. (f) Controls, dystrophies, and atrophies with features 25, 32, 14, 16, 58, 62, 15, 30, and 26 as inputs.

3.5 Severity Grading

The last study involves the degree of severity that a biopsy with a dystrophic pattern presents. The affection degrees of these biopsies were evaluated by the pathologist. Multivariate statistic analysis is applied to compare multivariate data and establish the quantitative changes and differences between groups under investigation on their characteristics. The kernel approach is to find a high correlation feature set without redundancy.²³ A principal component analysis is applied. Despite not having reliable values of the degrees of severity of biopsies with atrophic patterns, the PCA results for the six cases (the three comparisons with the sets of features of Table 8) are displayed in a bidimensional space representing the two first principal components (see Fig. 9). The green, red, and blue points represent control

Table 12 Pearson correlation between the affection degree of the biopsies and the Euclidean distance between the pathological images and the centroid of the controls in the PCA graph.

Comparison	24 features		82 features	
	Selected features	Pearson coefficient	Selected features	Pearson coefficient
Controls-dystrophies (C/MD)	19 18 15	0.875	25 9	0.7512

Note: The bold value indicates the best result.

biopsies, dystrophies, and atrophies, respectively. The 20 biopsies used to evaluate the classification procedure are also represented; the seven controls are displayed in dark green, the seven dystrophies in dark red, the five NA in light blue, and the atrophy with pseudo-dystrophic nature in black. It is noteworthy that this last case is located between the two big groups of dystrophies and atrophies [see Fig. 9(f)], since this biopsy presents characteristics of the two diseases.

The Pearson correlation coefficient is used to study whether or not these graphs are representative of the affection degree of the biopsies. Since only the severity degree for the dystrophic patterns is known, the Pearson correlation coefficient between these severity degrees, the Euclidean distance between the pathological images, and the centroid of the controls in the PCA graph is computed. This correlation coefficient measures the strength of linear dependence between these two variables (degree and distance). The results are presented in Table 12. We can observe that the set of 24 features is correlated with the pathologist’s evaluation. The high values obtained indicate that the visualization of the data by PCA can be considered as a way of analysis of the affection degree.

4 Conclusions

In this paper a procedure to analyze and classify neuromuscular diseases through biopsy images of fluorescence microscopy is presented. The patterns that muscular biopsies present are dystrophic and atrophic. The procedure begins with an accurate segmentation of the muscle fibers (see Sec. 3.1). Then, two sets of features are extracted: 24 features that physicians take into

account to diagnose the diseases (see Tables 5 and 6) and 58 structural features that the human eye does not see (see Table 7) forming a set of 82 features (24 + 58). A study is derived from each of them to analyse the goodness of these sets and to study if the addition of new features improves the classification of the diseases. This study developed three comparisons that are presented: control/muscular dystrophies (C/MD), control/neurogenic atrophies (C/NA), and controls/muscular dystrophies/neurogenic atrophies (C/MD/NA). A feature selection step is performed for each one and each set of characteristics (see Table 8). Seventy-one training images were used. The classification error obtained for the selected features for the two first comparisons is 0%, however for the triple comparison (C/MD/NA), the classification error decreases with the second set of features. Finally, the classification on 20 test images is carried out by a Fuzzy ARTMAP neural network, whose output is modified to allow the categorization of a new biopsy into a category and an estimation of the certainty of belonging to this category. The results (see Table 9) show that to distinguish between control and atrophies (C/NA), the second set of features improves the classification, however, there is no difference between the other two comparisons. In addition, a study of grading the severity of dystrophic pattern is performed by the application of a principal component analysis for which a correlation coefficient is computed for each set of selected features (see Fig. 9 and Table 12).

It can be concluded that, in view of all the results presented, the best set of features to distinguish between controls and muscular dystrophies (C/MD) is the first one, derived from the features that the physicians take into account in their diagnosis. This fact may be due to clear differences between the biopsies belonging to these two groups. However, when biopsies belonging to atrophies come into play (C/NA and C/MD/NA), the addition of new features undetectable by the human visual inspection improves its categorization. This is because biopsies affected by atrophy present similar characteristics of control biopsies.

Therefore, to classify a new biopsy it would proceed as follows:

- Segmentation of the muscular fibers
- Extraction of features
- Classification by Fuzzy ARTMAP neural network with the second selected features set for the triple comparison C/MD/NA (features numbers 25, 32, 14, 16, 58, 62, 15, 30, and 26) obtaining a category for this biopsy.
- Study of output values (k_{cl}) of the neural network.
- If the output values do not present a clear result, i.e., a k_{cl} significantly higher than the other ones, a classification with the other two comparisons (controls/dystrophies and controls/atrophies) is required.
- Application of the criterion shown in Table 11 (C = control, MD = muscular dystrophy, and NA = neurogenic atrophy).
- Application of the principal component analysis to estimate the severity degree of the dystrophic patterns.

In future works, a further study of different methods of feature selection, classification, and severity grading will be performed focusing into the categorization of biopsies affected by atrophies.

Acknowledgments

Aurora Sáez is funded by the Consejería de Innovación, Ciencia y Empresa of Junta de Andalucía, Spain. Luis M. Escudero and Adoración Montero-Sánchez are supported by the Miguel Servet (Instituto Carlos III) program. This work is supported by the project TEC2010-21619-C04-02, CICYT, Spain.

References

1. K. R. Castleman et al., "Quantitative muscle biopsy analysis," *Monogr. Clin. Cytol.* **9**, 101–116 (1984).
2. M. N. Gurcan et al., "Histopathological image analysis: a review," *IEEE Rev. Biomed. Eng.* **2**, 147–171 (2009).
3. S. Doyle et al., "Automated grading of prostate cancer using architectural and textural image features," in *4th IEEE Int. Symp. Biomed. Imag.*, Metro Washington DC, pp. 1284–1287 (2007).
4. S. W. K. Chan, K. S. Leung, and W. S. F. Wong, "An expert system for the detection of cervical cancer cells using knowledge-based image analyzer," *Artif. Intell. Med.* **8**(1), 67–90 (1996).
5. P. W. Huang and Y. H. Lai, "Effective segmentation and classification for HCC biopsy images," *Pattern Recogn.* **43**(4), 1550–1563 (2010).
6. M. E. Plissiti, C. Nikou, and A. Charchanti, "Automated detection of cell nuclei in Pap smear images using morphological reconstruction and clustering," *IEEE Trans. Inf. Technol. Biomed.* **15**(2), 233–241 (2011).
7. E. Ficarra et al., "Automated segmentation of cells with IHC membrane staining," *IEEE Trans. Biomed. Eng.* **58**(5), 1421–1429 (2011).
8. C. Bergmeir et al., "Segmentation of cervical cell images using mean-shift filtering and morphological operators," *Proc. SPIE* **7623**, 76234C (2010).
9. N. M. Harandi et al., "An automated method for segmentation of epithelial cervical cells in images of ThinPrep," *J. Med. Syst.* **34**(6), 1043–1058 (2010).
10. A. G. Todman and E. Claridge, "Low-level grouping mechanisms for contour completion," *Inf. Sci.* **125**(1–4), 19–35 (2000).
11. Y. J. Kim et al., "Fully automated segmentation and morphometrical analysis of muscle fiber images," *Cytometry* **71**(1), 8–15 (2007).
12. A. Klemencic, S. Kovacic, and F. Pernus, "Automated segmentation of muscle fiber images using active contour models," *Cytometry* **32**(4), 317–326 (1998).
13. O. Sertel et al., "Microscopic image analysis for quantitative characterization of muscle fiber type composition," *Comput. Med. Imag. Graph.* **35**(7–8), 616–628 (2011).
14. B. Meunier et al., "Development of image analysis tool for the classification of muscle fibre type using immunohistochemical staining," *Histochem. Cell Biol.* **134**(3), 307–317 (2010).
15. P. Karen et al., "Software for muscle fibre type classification and analysis," *Eur. J. Histochem.* **53**(2), 87–95 (2009).
16. W. M. Behan et al., "Validation of a simple, rapid, and economical technique for distinguishing type 1 and 2 fibres in fixed and frozen skeletal muscle," *J. Clin. Pathol.* **55**(5), 375–380 (2002).
17. F. Garton et al., "Validation of an automated computational method for skeletal muscle fibre morphometry analysis," *Neuromuscular Disord.* **20**(8), 540–547 (2010).
18. L. M. Escudero et al., "Epithelial organisation revealed by a network of cellular contacts," *Nat. Commun.* **2**(1), 526 (2011).
19. A. Sáez et al., "Quantifiable diagnosis of neuromuscular diseases through network analysis," *BMC Med.* **11**(1), 77 (2013).
20. T. R. Helliwell, "Muscle: part 1—normal structure and function," *Curr. Orthop.* **13**(1), 33–41 (1999).
21. J. Finsterer, L. Papic, and M. Auer-Grumbach, "Motor neuron, nerve, and neuromuscular junction disease," *Curr. Opin. Neurol.* **24**(5), 469–474 (2011).
22. V. Dubowitz and C. A. Sewry, *Muscle Biopsy: A Practical Approach*, 3rd ed., Elsevier, Philadelphia (2007).
23. S. Chen et al., "Recent advances in morphological cell image analysis," *Comput. Math. Methods Med.* **2012**, 101536 (2012).
24. P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer-Verlag, Berlin, Germany (1999).

25. C. Jung and C. Kim, "Segmenting clustered nuclei using H-minima transform-based marker extraction and contour parameterization," *IEEE Trans. Biomed. Eng.* **57**(10 Pt 2), 2600–2604 (2010).
26. K. Ali et al., "Medical image segmentation using h-minima transform and region merging technique," in *Proc. 2011 9th Int. Conf. Frontiers of Information Technology, FIT 2011*, Islamabad, Pakistan, Vol. 6137132, pp. 127–132, IEEE Computer Society (2011).
27. S. Osher and J. A. Sethian, "Fronts propagating with curvature dependent speed algorithms based on Hamilton-Jacobi formulations," *J. Comput. Phys.* **79**(1), 12–49 (1988).
28. C. Li et al., "Level set evolution without re-initialization: a new variational formulation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn. CVPR 2005*, San Diego, CA, Vol. 1, pp. 430–436, IEEE Computer Society (2005).
29. F. Meyer and S. Beucher, "Morphological segmentation," *J. Vis. Commun. Image Represent.* **1**(1), 21–46 (1990).
30. K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, San Diego, CA (1990).
31. A. Jain and D. Zongker, "Feature selection: evaluation, application, and small sample performance," *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(2), 153–158 (1997).
32. P. Pudil, J. Novovicova, and J. Kittler, "Floating search methods in feature selection," *Pattern Recogn. Lett.* **15**(11), 1119–1125 (1994).
33. G. A. Carpenter et al., "Fuzzy ARTMAP: a neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Trans. Neural Netw.* **3**(5), 698–713 (1992).