Applied Soft Computing 12 (2012) 2188–2198

Contents lists available at SciVerse ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Classification of EMG signals using combined features and soft computing techniques

Abdulhamit Subasi *

International Burch University, Faculty of Engineering and Information Technologies, Francuske Revolucije bb. Ilidza, Sarajevo, 71000, Bosnia and Herzegovina

ARTICLE INFO

Article history:
Received 25 January 2010
Received in revised form 28 August 2011
Accepted 18 March 2012
Available online 3 April 2012

Keywords:
Adaptive neuro-fuzzy inference system (ANFIS)
Dynamic Fuzzy Neural Network (DFNN)
Electromyography (EMG)
Motor unit action potentials (MUPs)
Discrete Wavelet Transform (DWT)

ABSTRACT

The motor unit action potentials (MUPs) in an electromyographic (EMG) signal provide a significant source of information for the assessment of neuromuscular disorders. Since recently there were different types of developments in computer-aided EMG equipment, different methodologies in the time domain and frequency domain has been followed for quantitative analysis of EMG signals. In this study, the usefulness of the different feature extraction methods for describing MUP morphology is investigated. Besides, soft computing techniques were presented for the classification of intramuscular EMG signals. The proposed method automatically classifies the EMG signals into normal, neurogenic or myopathic. Also, multilayer perceptron neural networks (MLPNN), dynamic fuzzy neural network (DFNN) and adaptive neuro-fuzzy inference system (ANFIS) based classifiers were compared in relation to their accuracy in the classification of EMG signals. Concerning the impacts of features on the EMG signal classification, different results were obtained through analysis of the soft computing techniques. The comparative analysis suggests that the ANFIS modelling is superior to the DFNN and MLPNN in at least three points: slightly higher recognition rate; insensitivity to overtraining; and consistent outputs demonstrating higher reliability.

1. Introduction

The electromyogram (EMG) recorded from a contracting muscle using a needle electrode is used to diagnose neuromuscular disorders in clinical neurology. EMG consists of discrete waveforms called motor unit action potentials (MUPs), which result from the recurring discharges of groups of muscle fibres called motor units (MUs). The usual needle EMG treatment is performed using a concentric needle electrode. EMG signals can be recorded in varying degrees of voluntary muscle activity. When the force of contraction increased, additional motor units start to fire and the EMG signal becomes more and more complex, and as a result, individual MUPs cannot be identified easily. This signal is recognized as the interference pattern (IP). Analysis of the EMG IP is used in the description of muscle activity, muscle fatigue, chronic muscle pain, and in the diagnosis of patients with neuromuscular disorders [1]. MUPs from different MUs are in distinct shapes that remain nearly the same for each discharge. Hence, the MUPs can be identified using different pattern recognition techniques. The resulting information can be used to find out the origin of the diseases, i.e. neurogenic or myopathic [2–4]. It is typical clinical practice to examine the MUPs from visual inspection and from listening to their audio characteristics. Hence abnormalities can be detected easily by an experienced electrophysiologist with reasonable accuracy. Due to loss of individual muscle fibres, MUPs are low in amplitude and short in duration in myogenic disorders. In order to compensate for the low amplitude, a larger number of motor units are recruited at lower than normal levels of muscle contraction. Full interference pattern is developed at less than maximal contraction with increasing force [5]. Nevertheless, in neurogenic disorders, the excited motor neurons are decreased in number and muscle fibres are reorganized to give larger MUs, this causes MUPs having higher amplitude and longer duration than normal. In order to keep a certain force of contraction, the available motor neurons must fire at a higher rate than normal to balance the motor neuron loss. These apparent frequency changes to the signal in neuromuscular disease conditions may be the reason for some of the earliest quantitative methods being based on frequency domain analysis [6–8]. Nevertheless, subjective MUP assessment, even if satisfactory for the detection of obvious abnormalities, may not be sufficient to describe less apparent deviations or mixed patterns of abnormalities [9]. Consequently, for an effective automated EMG signal classification, a systematic treatment of EMG signals must be carried out. For this reason, a number of computer-based quantitative EMG analysis algorithms have been developed [8–23].

Artificial neural networks (ANN) have been used to combine the best time and frequency based measures in order to try to improve the diagnostic yield, but have yet to provide sufficient

* Tel.: +387 33 782 130; fax: +387 33 782 131.
E-mail address: asubasi@ibu.edu.ba

1568-4946/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2012.03.035

classification for clinical use [11,16]. Fuzzy set theory plays an imperative role in dealing with uncertainty when making decisions in different applications. First introduced by Zadeh [24], fuzzy logic and fuzzy set theory are employed to illustrate human thinking and reasoning in a mathematical framework. Fuzzy-rule based modelling is a qualitative modelling method where the system behaviour is expressed using a natural language [25]. Fuzzy sets have attracted the growing interest in modern information technology, production technique, decision making, pattern recognition, diagnostics, data analysis, etc. [25–28]. Recently, the integration of neural networks and fuzzy logic has given birth to new research into neuro-fuzzy systems which have the potential to capture the benefits of both these fields in a single framework. Neuro-fuzzy systems reduce the basic problem in fuzzy system design effectively using the learning capability of an ANN for automatic fuzzy if–then rule generation and parameter optimisation. Such applications have been developed for signal processing, automatic control, information retrieval, database management, computer vision and data classification [29–31]. Chan et al. [22] proposed a fuzzy approach to classify single-site electromyograph (EMG) signals for multifunctional prosthesis control. In their study, time segmented features were fed to a fuzzy system for training and classification. Weir and Ajiboye [23] developed a new multifunctional prosthetic hand mechanism that would be interfaced to the user using a four site myoelectric (EMG) controller based on fuzzy-logic techniques. In their work, three to four independent EMG sites were used to control three or four degrees-of-freedom in the prosthetic hand.

Neuromuscular diseases are a group of disorders whose most important characteristic is that they cause muscular weakness and/or muscle tissue wasting. These disorders have an effect on the motor nuclei of the cranial nerves, the anterior horn cells of the spinal cord, the nerve roots and spinal nerves, the peripheral nerves, the neuromuscular junction, and the muscle itself. Pathological changes in the signal are often subtle, especially at the beginning stages of a disease or in less severe cases, which has made quantification for diagnostic purposes difficult to realize. The MUPs are transients that are not similar for each recurrence of the same motor unit. Wavelet transforms can be used to characterize the localized frequency content of each MUP [8,13]. To contribute to the quantification of the routine needle EMG examination, we have developed a methodology for EMG signal classification into normal (NOR), myopathic (MYO) and neurogenic (NEU). The proposed methodology consists of two steps. In the first step, the EMG signal is preprocessed using autoregressive (AR), discrete wavelet transform (DWT) and wavelet packet energy (WPE) to extract features from EMG. In the second step, an unknown EMG signal is classified as NOR, MYO or NEU using soft computing techniques. The remainder of the paper is organized as follows. In the next section, we give information about the subjects and present the methods applied in each step of the EMG signal classification process. Section 3 provides a complete experimental study of the neuro-fuzzy models based EMG signal classification scheme, in which the impact of feature set and neuro-fuzzy algorithmic issues are examined with respect to classification performance. Finally, the conclusions are summarized in Section 4.

2. Materials and methods

2.1. Subjects and data acquisition

An EMG system (*Keypoint*; *Medtronic Functional Diagnostics*, *Skovlunde*, *Denmark*) with standard settings was used. The EMG signal was acquired from the biceps brachii muscle using a concentric needle electrode (0.45 mm diameter with a recording surface area 0.07 mm^2; impedance at 20 Hz below 200 kΩ). The signal was band-pass filtered at 5 Hz to 10 kHz and sampled at 20 kHz for 5 s with 12-bit resolution. The signal was then lowpass filtered at 8 kHz. All the measurements from patients and control group were done in Neurology Department of University of Gaziantep. The recording points within the muscle are standardized, with MUPs recorded from three to five different needle insertions. MUPs were not recorded close to the surface of the muscle. The electrode was usually advanced at least 3–5 mm into the muscle before recording. The electrode was also moved at least 3–5 mm between recordings to make sure that different MUPs are recorded. The electrode was advanced until the medial or posterior border of the muscle is reached. The electrode was then pulled out and inserted to a new radial direction.

The EMG signals were recorded from the biceps brachii muscle at force levels approximately 30% of maximum voluntary contraction (MVC) under isometric conditions. Patient diagnoses were based on a range of clinical information included a general examination and clinical history of the patient, and EMG and nerve conduction tests. Muscle biopsies were not taken in the majority of cases, on the other hand they are only considered in EMG clinic in cases where diagnosis is uncertain or for specific clinical reasons. In this study, EMG data collected from 27 subjects was analysed. Data were recorded from seven healthy subjects (three males, four females) with ages ranging from 10 to 43 years (mean age ± standard deviation (S.D.): 30.2 ± 10.8 years), seven myopathic subjects (four males, three females) with ages ranging from 7 to 46 years (mean age ± standard deviation (S.D.): 21.5 ± 13.3 years) and thirteen neurogenic subjects (eight males, five females) with ages ranging from 7 to 55 years (mean age ± standard deviation (S.D.): 25.1 ± 17.2 years) as in [21].

2.2. Feature extraction methods

The success of any signal classification system depends on the choice of features used to characterize the raw signals [6]. It is desirable to use multiple feature parameters for EMG pattern classification since it is very difficult to extract a feature parameter that reflects the distinctive feature of the measured signals. In order to focus on this problem, in the analysis of EMG signals (Fig. 1(a), (c), and (e)), the autocorrelation function; wavelet transform (WT) and frequency domain features have been used [2–6]. The results of the studies in the literature have demonstrated that the WT is one of the most promising methods to extract features from the EMG signals [13,30,31]. In this respect, in the present study for EMG signal classification, we used autoregressive (AR) method and the WT as a comparative work for feature extraction from the EMG.

2.2.1. Autoregressive modelling of EMG signals

Autoregressive (AR) method makes use of a linear process, commonly referred to as a model to estimate the power spectrum. The model parameters adjusted for the best match between model output and the waveform of interest. When the best match is obtained, the model's frequency characteristics give the best estimate of the waveform's spectrum, given the restriction of the model. The equation of a classical AR complex process, in a non-stationary context, is given by:

$$y(n) = -\sum_{i=1}^{p} a(i)y(n-i) + u(n) \quad (1)$$

where $y(n)$ is the model output, $a(i)$ are AR complex parameters, p is the AR model order, $u(n)$ is the input or noise function and n is the sample time [32]. One of the most popular methods for estimating

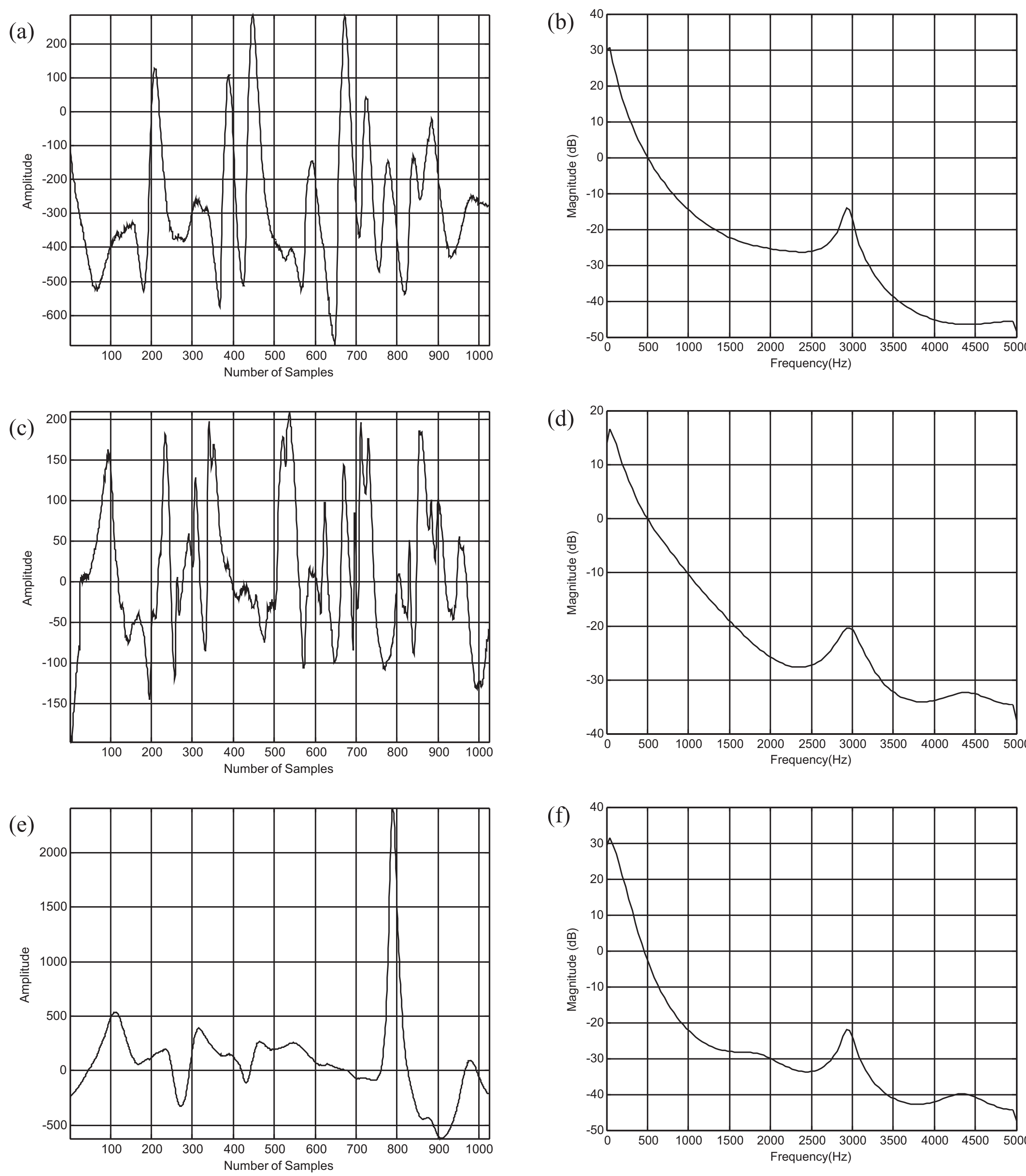

Fig. 1. EMG signal taken from (a) normal subject and (b) its AR PSD, (c) myopathic and (d) its AR PSD, (e) neurogenic and (f) its AR PSD.

the AR parameters of a sequence of *N* data points is using Burg's algorithm [33–37].

$$\hat{P}_{BURG}(f) = \frac{\hat{e}_p}{\left|1 + \sum_{k=1}^{p} \hat{a}_p(k) e^{-j2\pi fk}\right|^2} \tag{2}$$

A very good tutorial on the theoretical foundations of this algorithm can be found in Kay and Marple [35]. The AR Burg algorithm was applied to a specific EMG signal epochs (1024 samples). One of the most important aspects of the model-based methods is the selection of the model order. One of the better known criteria for selecting the model order has been proposed by Akaike [37], called the Akaike information criterion (AIC). In this study, model order of the AR method was taken as 15 by using AIC and window size is taken as 256. The AR PSDs, which are shown in Fig. 1(b), (d) and (f) for normal, myopathic and neurogenic EMG signals respectively, is used as an input to neuro-fuzzy classifiers.

2.2.2. Feature extraction using discrete wavelet transform

The Wavelet transform (WT) presents the time–frequency representation. WT is able to provide the time and frequency information at the same time, thus giving a time–frequency representation of the signal. As the time-domain signal is passed from various high pass and low pass filters, then, either portion or both are taken, and do the same thing again. This operation is called decomposition. This continues until the signal decomposed to a pre-defined certain level. Then a bunch of signals represent the same signal, but all corresponding to different frequency bands. Higher frequencies are

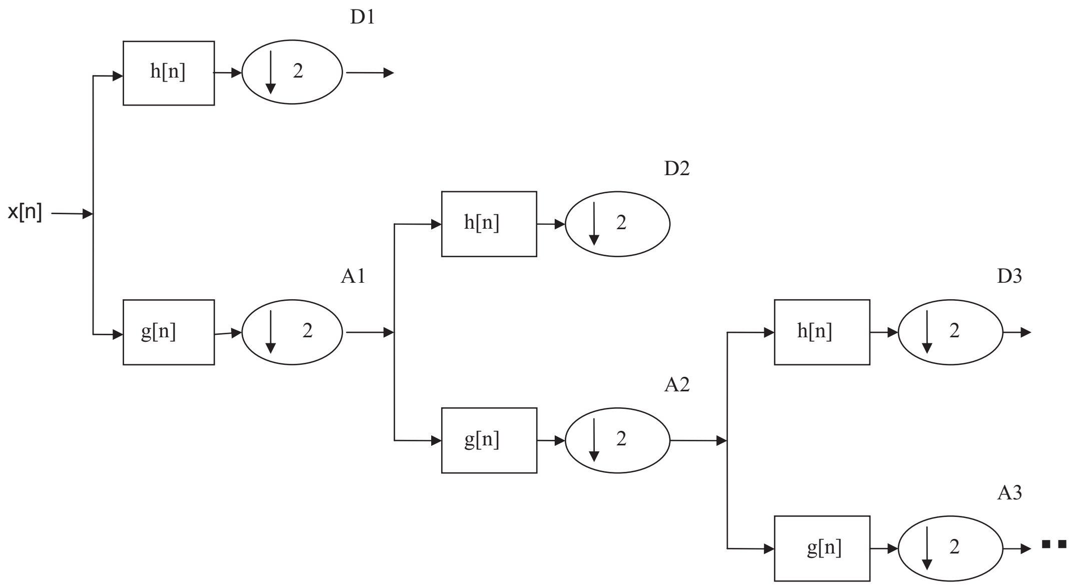

Fig. 2. Sub-band decomposition of DWT implementation; $h[n]$ is the high pass filter, $g[n]$ the low pass filter.

better resolved in time, and lower frequencies are better resolved in frequency. The continuous wavelet transform is defined as follows

$$W(a,b)=\int_{-\infty}^{\infty} x(t)\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)\,dt, \quad (3)$$

As seen in the above equation, the transformed signal is a function of two variables, b and a, the translation and scale parameters, respectively. ψ is the transforming function, and it is called the mother wavelet. High scales correspond to a non-detailed global view (of the signal), and low scales correspond to a detailed view. Scaling either dilates or compresses a given signal. Larger scales correspond to dilated (or stretched out) signals and small scales correspond to compressed signals. This definition of the WT demonstrates that the wavelet analysis is a measure of similarity between the basis functions (wavelets) and the signal itself [38].

On the other hand, the discrete wavelet transform (DWT) presents adequate information both for the analysis and synthesis of the original signal, with a significant reduction in the computation time. The DWT analyses the signal at different frequency bands with different resolutions by decomposing the signal into a coarse approximation and detail information. DWT employs two sets of functions, called scaling functions and wavelet functions, which are associated with low pass and high-pass filters, respectively. The decomposition of the signal into different frequency bands can be obtained by successive high-pass and low-pass filtering of the time domain signal. The original signal $x[n]$ is first passed through a half-band high-pass filter $g[n]$ and a low-pass filter $h[n]$. The signal can therefore be sub-sampled by 2, simply by discarding every other sample. This constitutes one level of decomposition and can be expressed as follows:

$$y_{high}[k]=\sum_n x[n].g[2k-n] \quad (4)$$

$$y_{low}[k]=\sum_n x[n].h[2k-n] \quad (5)$$

where $y_{high}[k]$ and $y_{low}[k]$ are the outputs of the high-pass and low-pass filters, respectively, after sub-sampling by 2. This decomposition halves the time resolution when only half the number of samples describes the entire signal. Nevertheless, this procedure doubles the frequency resolution, because the frequency band of the signal now spans only half the previous frequency band, successfully reducing the uncertainty in the frequency by half. The above procedure, which is also known as the sub-band coding, can be repeated for further decomposition. At every level, the filtering and sub-sampling results in half the number of samples and half the frequency band spanned. Fig. 2 shows this procedure, where $x[n]$ is the original signal to be decomposed, and $h[n]$ and $g[n]$ are low-pass and high-pass filters, respectively [38]. The down-sampled outputs of first high-pass and low-pass filters provide the detail, D_1 and the approximation, A_1, respectively. The first approximation, A_1 is further decomposed and this process is continued as in Fig. 3. These approximation and detail records are reconstructed from the Daubechies 4 (DB4) wavelet filter. Details are given in [38–45].

The extracted wavelet coefficients provide a compact representation that shows the energy distribution of the EMG signal in time and frequency. In order to decrease the dimensionality of the extracted feature vectors, statistics over the set of the wavelet coefficients were used [46]. The following statistical features were utilized to represent the time–frequency distribution of the EMG signals:

(1) Mean of the absolute values of the coefficients in each sub-band.
(2) Average power of the wavelet coefficients in each sub-band.
(3) Standard deviation of the coefficients in each sub-band.
(4) Ratio of the absolute mean values of adjacent sub-bands.

Features 1 and 2 represent the frequency distribution of the signal and the features 3 and 4 the amount of changes in frequency distribution. We extracted 6 different values for features (1), (2) and (3); 5 values for feature (4), hence we had total 23 features. These 23 features were calculated from the frequency bands D1, D2, D3, D4, D5 and A5 of DWT and then used as an input to neuro-fuzzy classifiers.

2.2.3. Feature extraction using wavelet packed energy

Wavelet packets were introduced by Coifman and Wickerhauser [47] as a generalized family of multiresolution orthogonal or biorthognal basis. Unlike wavelet transform which is realized only by a low-pass filter bank, wavelet packet transform (WPT) is realized by a two-channel filter bank which can be iterated over

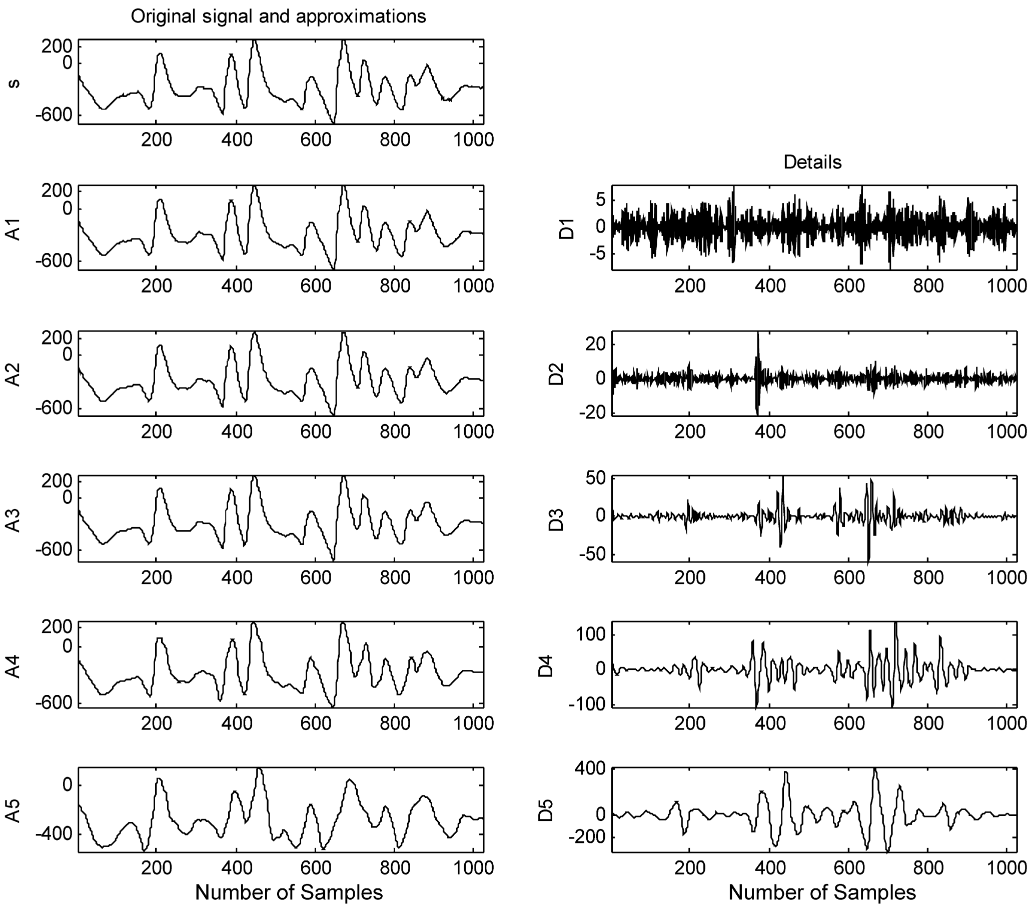

Fig. 3. EMG signal and its wavelet decomposition into approximate (A1–A5) and detailed (D1–D5) coefficients.

either a low-pass or a high-pass branch. Therefore, the information in high frequencies can be analysed as well as that in low frequencies in WPT. As a consequence, finer frequency bands can be gained by WPT than by wavelet transform.

In order to separate the discriminable features from the interferential features effectively, we should concentrate the MUPs on one narrow frequency band as much as possible, rather than make them spread out over one wide frequency band. When $x(t)$ is decomposed into the fourth resolution level ($k = 5$) with wavelet packet transform, the whole scaling space V_0^0 with frequencies in the interval $(0, f_s/2)$ is divided into 32 subspaces, the subsignal at V_k^{n-1}, the nth subspace on the kth level, can be reconstructed [48] by

$$x_k^n(t) = \sum_i D_i^{k,n} \psi_{k,i}(t), \quad i \in Z \tag{6}$$

where $D_i^{k,n}$ is the wavelet packet coefficients at V_k^{n-1}, $\psi_{k,i}(t)$ is the wavelet function.

Thus, we can be write $x(t)$ as

$$x(t) = \sum_{n=1}^{2^{-k}} x_k^n(t) = \sum_{n=1}^{2^{-k}} \sum_i D_i^{k,n} \psi_{j,k}(t) \tag{7}$$

Since the wavelet $\psi_{k,i}(t)$ is an orthogonal basis at $L^2(\Re)$, the energy of the subsignal $x_k^n(t)$ is calculated by

$$E_n = \sum_i \left| D_i^{k,n} \right|^2 \tag{8}$$

The total energy of $x(t)$ is

$$E = \sum_n E_n \tag{9}$$

As a result, wavelet packet energy (WPE) which quantifies the probability distribution of the spectral energy is given by

$$p_n = \frac{E_n}{E} \tag{10}$$

In order to find the general distribution rule of wavelet packet energy for every pattern of EMG signal, we evaluate the average WPE of every EMG signal pattern

$$p_n = \frac{1}{N} \sum_j^N p_n^j \tag{11}$$

where p_n^j is WPE at the nth subspace of the jth subject and N is the total number of subjects [49].

In this work, Daubechies 2 (db2) family of wavelet packets is implemented as the mother wavelet [49]. As a result of decomposition, 32 feature vectors were extracted for each signal frame and used as an input to neuro-fuzzy classifiers for the classification of the EMG signals.

2.3. Neuro-fuzzy classification techniques

Several factors must be considered when choosing a classifier or a recognition method for the present application. Due to the nature of the EMG signal, it is reasonable to expect a large variation in the value of particular feature between individuals. Many factors such as changes in electrode position, EMG signal training will produce changes in feature values over time. A suitable recognition method must be able to accommodate the expected individual differences. Moreover, it must generate reasonably perfect results with the extracted feature parameters by minimizing the required computing time for real time application [6].

2.3.1. *Artificial neural networks (ANN)*

Artificial neural networks (ANNs) have been first studied by Rosenblatt who applied single-layer perceptrons to pattern classification learning [50]. But, later Minsky and Papert pointed out that single-layer systems were limited and expressed pessimism over multilayer systems [51,52]. The number of inputs is fixed by the number of predictors and the number of outputs is fixed by the number of classes. But only one output is required for two classes. A typical ANN has three layers: input; hidden; and output. This kind of network is known as a multi-layer perceptron neural network (MLPNN). Because of this complexity it is possible that the network can be practically tolerant of missing values. This means that the network's model of the relationship between the predictors and the class (the output) is dispersed through its weights. Here the problem is to find proper values for these weights to maximise the accuracy of the predictions given a set of training data. This is achieved by a series of iterative tunings to the weights during a training cycle. One of the most commonly used algorithms to implement this training is known as back-propagation [53]. A back-propagation MLPNN is an adaptive network whose nodes (neurons) carry out the same function on incoming signals. This node function is generally consists of the weighted sum and a differentiable non-linear activation function (transfer function). The net input x of a node is described as the weighted-sum of the incoming signals plus a bias term. For example, the net input and output of node j as

$$x_j = \sum_i w_{ij} x_i + w_j$$
$$x_j = f(x_j) = \frac{1}{1 + \exp(-x_j)} \tag{12}$$

where x_i is the output of node i set in any one of the previous layers, w_{ij} is the weight related to the link connecting nodes i and j, and w_j is the bias of node j. The backward error propagation or simply backpropagation (BP) can be defined as follows. First, a squared error measure for the pth input–output pair is

$$E_p = \sum_k (d_k - x_k)^2 \tag{13}$$

where d_k is the desired output for node k, and x_k is the actual output for node k as the input part of the pth data pair presented. In order to find the gradient vector, an error term ε_i is defined as [52]

$$\varepsilon_i = \frac{\partial^+ E_p}{\partial x_i} \tag{14}$$

The network begins with randomised values for the weight values. Any difference between the actual class and the output is a measure of network's current inaccuracy which comes from inappropriate weights. If the weights are tuned, the error can be corrected. The back-propagation weight adjustment algorithm measures how the error is propagated back through the network's weights. After adjusting the weights the next case is applied to the network and the weight adjustment procedure is repeated. When all of the cases in the training set have passed through the network one epoch has been completed. The entire training cycle is determined by some pre-specified termination criterion such as a minimum error rate. In order to minimise training times, a momentum term is used. The momentum takes account of the magnitude of the preceding weight change and adjusts the current weight change. In essence, this accelerates the rate of change when it is suitable [53,54].

The MLPNN was designed with combined features of EMG signal in the input layer; and the output layer consisted of three nodes representing normal, myopathic or neurogenic disorder. The preliminary architecture of the network was examined using one and two hidden layers with a variable number of hidden nodes in each. It was found that one hidden layer is adequate for the problem at hand. Consequently, the required network will contain three layers of nodes. The training procedure started with one hidden node in the hidden layer, followed by training on the training data (900 distinct data sets), and then by testing on the test data (300 distinct data sets) to examine the network's prediction performance on cases never used in its development. Then, the same procedure was run repetitively each time the network was expanded by adding one more node to the hidden layer, until the best architecture and set of connection weights were obtained. Using the error–backpropagation algorithm for training, a training rate of 0.001 and momentum coefficient of 0.9 was found optimum for training the network with various topologies. The selection of the optimal network was based on monitoring the variation of mean squared error (MSE) of the network and classification accuracy parameters as the network was expanded in the hidden layer size and for each training cycle. The mean of squares of error representing the mean of square of deviations of MLPNN solution (output) from the true (target) values for both the training and test sets was used for selecting the optimal network.

2.3.2. *Dynamic fuzzy neural networks (DFNN)*

Neuro-fuzzy systems are fuzzy systems that use ANNs theory in order to validate their properties by using data examples. In order to characterize inaccurate data and knowledge, fuzzy rules are used instead of exact rules when representing knowledge. Fuzzy systems are rule-based systems based on fuzzy rules and fuzzy inference. Fuzzy rules characterize in a straightforward manner commonsense knowledge. Commonsense knowledge may have been acquired from long-term experience over many years. One of the most important problems in constructing fuzzy systems is the problem of defining fuzzy membership functions in addition to the problem of characterizing fuzzy rules. The dynamic fuzzy neural network (DFNN) [55,56] is an implementation of the Takagi–Sugeno–Kang (TSK) fuzzy system based on an extended radial basis function networks (RBFN) and its learning algorithm. The extended RBFN has five layers and no bias, and the weights may be a linear regression of the input. The system starts with no hidden unit, and neurons can be added or deleted dynamically according to their significance to system performance. Both the parameters and the structure can be adjusted simultaneously [57].

DFNN model is a network with unconstrained connectivity and with dynamic elements in its fuzzy processing units (feuron) [58–60]. The feuron corresponds to a single dynamic neuron with fuzzy activation function. Fuzzy activation function we have used here is a standard fuzzy system that has singleton fuzzifier, product inference engine with Gaussisan membership function and centre average defuzzifier.

Because the feuron's activation model has a single input, the product inference output becomes the membership functions itself. The activation function of the ith feuron can be expressed as:

$$\phi_i(x_i) = \frac{\sum_{j=1}^{R_i} a_{ij}\mu_j(x_i)}{\sum_{j=1}^{R_i} \mu_j(x_i)} = \frac{\sum_{k=1}^{R_i} a_{ij} \ \exp\left(-\frac{1}{2}\left(\frac{x_i - c_{ij}}{\sigma_{ij}}\right)^2\right)}{\sum_{j=1}^{R_i} \exp\left(-\frac{1}{2}\left(\frac{x_i - c_{ij}}{\sigma_{ij}}\right)^2\right)} \tag{15}$$

where c_{ij} is centre and σ_{ij} is spread of the jth receptive field unit of the its ith feuron.

The boundary membership functions of the feuron may be represented by hard constraints in the form of:

$$\mu_1(x_i) = 1, \quad \text{if } x_i = x_{iL} \quad \text{and} \quad \mu_{Ri}(x_i) = 1, \quad \text{if } x_i = x_{iU} \tag{16}$$

where x_{iL} and x_{iU} are the lower and upper limits of the state values of the ith feuron, respectively. The general computational model that has been used for DFNN can be represented as:

$$z_i = \sum_{j=1}^{n} q_{ij} y_j, \quad i = 1, 2, \ldots, M \tag{17}$$

$$y_i = \phi_i(x_i, \pi_i) = \frac{\sum_{k=1}^{R_i} a_{ij} \exp\left(-\frac{1}{2}\left(\frac{x_i - c_{ij}}{\sigma_{ij}}\right)^2\right)}{\sum_{j=1}^{R_i} \exp\left(-\frac{1}{2}\left(\frac{x_i - c_{ij}}{\sigma_{ij}}\right)^2\right)}, \quad i = 1, 2, \ldots, n \tag{18}$$

$$\dot{x}_i = f_i(x_i, p) = \frac{1}{T_i}\left[-x_i + \sum_{j=1}^{n} w_{ij} y_j + \sum_{j=1}^{L} p_{ij} u_j + b_i\right];$$
$$x_i(0) = x_{i0}, \quad i = 1, 2, \ldots, n \tag{19}$$

There are L input signals, which can be time-varying, n dynamic feuron units, n bias terms, and M output signals. The units have dynamics associated with them, and they get input from themselves, bias term and from all other units. The output of a unit y_i is a fuzzy activation function $\phi(x_i)$ of a state variable x_i relevant to the unit. The output of the overall network is a linear weighted sum of the unit outputs. The bias term b_i is added to the unit inputs. p_{ij} is the input connection weights from jth input to ith feuron, w_{ij} is the interconnection weights from jth feuron to ith feuron and q_{ij} is the output connection weights from jth feuron to ith output. T_i is the dynamic constant of ith feuron and b_i is the bias term of ith feuron. π is the parameters of fuzzy activation function which are centres c, spreads σ and output centres a. The initial conditions on the state variables $x_i(0)$ should be specified. In DFNN, parameters are extracted from a data set that describes the system behaviour. The combination of fuzzy concept in DFNN considerably improves the transparency for a superior understanding of their inner workings. Since each membership function and fuzzy rule used in DFNN has comprehensible semantic implication, the neural network-like structure of DFNN becomes more and more apparent once they are trained. Details of this model has encountered in the literature [55–67].

By using the DFNN algorithm several sets of fuzzy rules were extracted from the EMG data set. First, three fuzzy predicates were used to represent each of the input attributes and each of the three output variables. A DFNN structure was trained with fuzzified EMG training data for 1500 training cycles, a learning rate of 0.1 and momentum of 0.2 until an RMS error of 0.01. Fuzzy rules were then extracted. Several experiments with test examples were carried out, which is the confusion matrix of the desired classification and the classification produced by the network.

2.3.3. *The adaptive neuro-fuzzy inference systems (ANFIS)*

The adaptive network-based fuzzy inference system (ANFIS) first introduced by Jang in 1993 [29]. ANFIS composed of three abstract components: a fuzzy rule base, which includes a set of fuzzy if–then rules; a database, which identifies the membership functions used in the fuzzy rules; and a reasoning mechanism, which carries out the inference procedure upon the rules to get a reasonable output or conclusion [67]. The ANFIS under consideration has two inputs x and y and one output f. For a first-order Sugeno fuzzy model [68], a general rule set with two fuzzy if–then rules is the following:

Rule 1: if x is A_1 and y is B_1 then $f = p_1x + q_1y + r_1$,
Rule 2: if x is A_2 and y is B_2 then $f = p_2x + q_2y + r_2$.

The corresponding equivalent ANFIS architecture is as shown in Fig. 4, where nodes of the same layer have similar functions, and the node function in each layer can be described as follows:

Layer 1: Each node i in this layer is an adaptive node, indicated by square node, with node function

$$O_i^1 = \mu_{A_i}(x), \quad i = 1, 2, \tag{20}$$

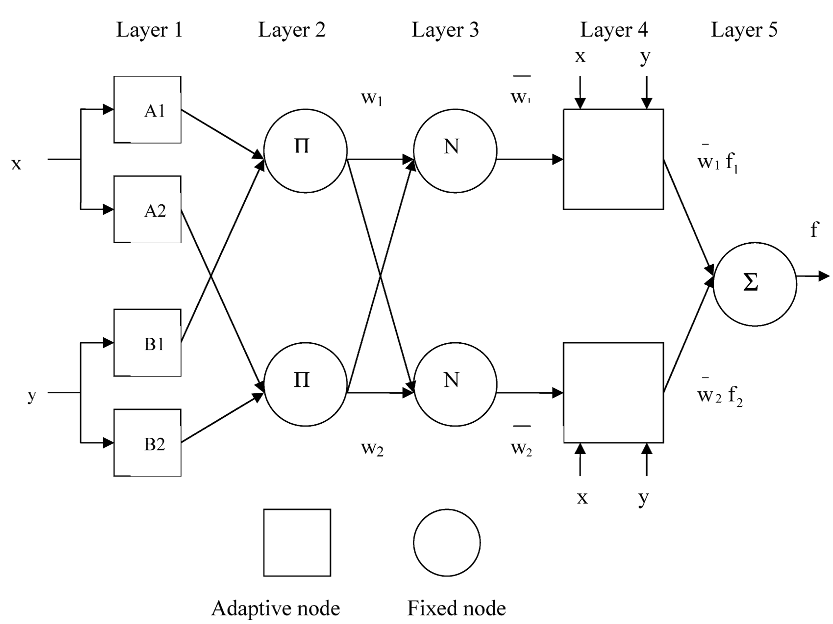

Fig. 4. The frameworks of ANFIS.

$$O_i^1 = \mu_{B_{i-2}}(y), \quad i = 3, 4, \tag{21}$$

where x (or y) is the input of node i, A_i (or B_{i-2}) is the linguistic variable related with this node. The membership function for A can be any appropriate parameterized membership function, such as the generalized bell function and generally adopts bell-shape with maximum and minimum equal to 1 and 0, respectively.

$$\mu_{A_i}(x) = \frac{1}{1 + \{((x - c_i)/a_i)^2\}^{b_i}} \tag{22}$$

where $\{a_i, b_i, c_i\}$ stands for the parameter set. If the values of this parameter set change, the bell-shape function will be changed in accordance. Meanwhile, the membership functions for a fuzzy set A, exhibit various forms. In this layer, the parameters are called as premise parameters.

Layer 2: Each node in this layer is a fixed node, indicated by circle node, whose output is the product of all incoming signals:

$$O_i^2 = w_i = \mu_{A_i}(x)\mu_{B_i}(y), \quad i = 1, 2, \tag{23}$$

The output signal w_i represents the firing strength of a fuzzy rule.

Layer 3: Each node in this layer is a fixed node, indicated by circle node. In order to find the normalized firing strength, the ratio of the ith rule's firing strength to the sum of all rules' firing strengths can be calculated as:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2, \tag{24}$$

Layer 4: Each node i in this layer is an adaptive node, indicated by square node, with a node function

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i), \quad i = 1, 2, \tag{25}$$

where $\bar{w}_i$ is the normalized firing strength from layer 3 and, $\{p_i, q_i, r_i\}$ is the parameter set in this layer which are referred to as the consequent parameters.

Layer 5: Each node in this layer is a fixed node, indicated by circle node, with node function to compute the overall output as the summation of all incoming signals

$$O_i^5 = \sum_{i=1}^{2} \bar{w}_i f_i = \frac{\left(\sum_{i=1}^{2} w_i f_i\right)}{w_1 + w_2}, \tag{26}$$

Hence we have built an adaptive network which is functionally equivalent to a Sugeno fuzzy model. Apparently, the assignment of node functions and the network configuration are arbitrary, as far as each node and each layer achieve significant and modular functionalities [52].

In the ANFIS architecture, when the values of the premise parameters are fixed, the overall output can be expressed as a linear combination of the consequent parameters. Mathematically, the output of the ANFIS can be modelled as

$$f = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 \tag{27}$$

Substituting Eq. (5) into Eq. (9) yields

$$f = \bar{w}_1 f_1 + \bar{w}_2 f_2 \tag{28}$$

By substituting the fuzzy if–then rules into Eq. (10), we get

$$f = \bar{w}_1(p_1 x + q_1 y + r_1) + \bar{w}_2(p_2 x + q_2 y + r_2) \tag{29}$$

After rearrangement, the output can be written as

$$f = (\bar{w}_1 x)p_1 + (\bar{w}_1 y)q_1 + (\bar{w}_1)r_1 + (\bar{w}_2 x)p_2 + (\bar{w}_2 y)q_2 + (\bar{w}_2)r_2, \tag{30}$$

which is a linear combination of the adjustable resulting parameters p_1, q_1, r_1, p_2, q_2 and r_2. Therefore, a hybrid learning algorithm can be used for parameter estimation in this kind of models [29]. More specially, the hybrid learning algorithms consists of two phases:

- In the forward pass of the hybrid learning algorithm, node output values go forward until layer 4 and the resultant parameters are described by the least squares method.
- In the backward pass, the output errors are propagated backward and the premise parameters are updated by gradient descent algorithm.

It has been shown [29] that the resultant parameters identified in this way are optimal under the condition that the premise parameters are fixed. As a result, the hybrid approach converges much faster, because it decreases the search space dimensions of the original backpropagation method [52]. A detailed description of this algorithm can be found in [52,65–70].

In this work, a network of 20 nodes, 4 rules, 4 Gaussian membership functions, and 16 unknown parameters was used. It must be mentioned that the ANFIS methodology is been used here to obtain a first order Sugeno model. ANFIS can achieve a comparable error in only 20 epochs, which is a lot less than the 1000 epochs required by the networks presented before [70]. The ANFIS used 900 training data in 800 training periods and the step size for parameter adaptation had an initial value of 0.0045. In a real world domain all of the features used in the descriptions of instances may have different levels of relevancy. Therefore, in the present study changes of the final (after training) membership functions with respect to the initial (before training) membership functions of the input parameters were examined. Membership function of each input parameter was divided into three regions, namely, small, medium, and large. The examination of initial and final membership functions indicates that there are considerable changes in the final membership functions. After training, 300 testing data were used to validate the accuracy of the ANFIS model for EMG signal classification.

2.4. Development of neuro-fuzzy model

Neuro-fuzzy systems reduce the basic problem in fuzzy system design by using the learning capability of an ANN for automatic fuzzy if-then rule generation and parameter optimisation. As a result, those systems can utilize linguistic information from the human expert as well as measured data during modelling. Due to the nature of the EMG signal, it is practical to expect a large variation in the value of particular feature between individuals. An appropriate recognition method should accommodate the expected individual differences. First of all, the set of feature parameters is extracted from the sample EMG signals. Then neuro-fuzzy classifiers were trained so that it is expected to be more precise for EMG signal classification. The classifiers proposed for the classification of the EMG signals were implemented by using the MATLAB software package.

The purpose of the modelling phase in this application was to build up classifiers that are able to identify any input combination as belonging to either one of the three classes: NOR, MYO or NEU. For developing the neuro-fuzzy classifiers, the feature vectors are extracted by using AR, DWT and WPT from each EMG signal frame (1024 distinct samples) as described in Section 2.2. As a result we produced 1200 feature vectors for each feature extraction method. Then we create EMG signal patterns by using each of these feature vectors separately or by combining two of them. Then 900 EMG signal patterns were randomly taken from the 1200 distinct EMG signal patterns and used for training the neural networks. The remaining 300 signal patterns were kept aside and used for testing the

Table 1
Class distribution of the samples in the training and test data sets.

Class	Training set	Test set	Total
Normal	300 (5 subject)	100 (2 subject)	400 (7 subject)
Myopathic	300 (4 subject)	100 (3 subject)	400 (7 subject)
Neurogenic	300 (9 subject)	100 (4 subject)	400 (13 subject)
Total	900 (18 subject)	300 (9 subject)	1200 (27 subject)

validity of the developed models. The class distribution of the samples in the training and testing data set is summarized in Table 1.

The neuro-fuzzy models were designed with different features of EMG signal in the input layer; and the output layer consisted of three nodes representing normal, myopathic or neurogenic disorder. A value of [100] was used when the experimental study indicated a normal, [010] for myopathy and [001] for neurogenic.

3. Results and discussion

In this study, neuro-fuzzy classifiers were used to classify EMG signals when different methods (AR, DWT and WPE) were used as to extract feature vectors. These features were extracted from each signal frame to produce the total feature set used to characterize the EMG signal pattern. Then neuro-fuzzy classifiers were trained so that it is expected to be more precise for EMG signal classification. In order to calculate the performance of our approach, the whole EMG data is divided into training and test sets, and k-fold cross-validation is used subsequently. The training set is used to build a classification model and the test set is used to verify it. We choose k-fold cross-validation which is a well-known method for evaluation. k-Fold cross validation is used by numerous researchers to reduce the bias related with random sampling of the training and test sets. In k-fold cross validation, the whole data set is randomly split into k mutually exclusive subsets (folds) of approximately equivalent size. The classification model is trained and tested k times. In other words, it is trained on all but one of the folds and tested on the remaining single fold. The cross-validation accuracy (CVA) is the average of the k individual accuracy measures

$$CVA = \frac{1}{k}\sum_{j=1}^{k} A_i \tag{31}$$

where k (10 in this case) is the number of folds used, and A_i is the accuracy measure of each fold, $i = 1, \ldots, k$ [70]. In this work, the whole EMG data contain three classes: NOR, MYO and NEU. While stratifying the data by these three classes, each of the 10 folds contains approximately the same proportions of NOR, MYO and NEU cases as those in the whole data set.

The test performance of the neuro-fuzzy models was determined by the computation of the following statistical parameters:

Specificity: number of correct classified healthy subjects/number of total healthy subjects.
Sensitivity (*myopathy*): number of correct classified subjects suffering from myopathy/number of total subjects suffering from myopathy.
Sensitivity (neurogenic): number of correct classified subjects suffering from neurogenic disorder/number of total subjects suffering from neurogenic disorder.
Total classification accuracy: number of correct classified subjects/number of total subjects.

The values of statistical parameters for MLPNN are given in Table 2. As it can be seen from Table 2, the sensitivity for myopathy value of each feature extraction methods is almost the same. For WPE, the sensitivity of neurogenic disorder is the best (97%) of all three feature extraction methods, but specificity is bad (74%). The specificity value for AR is the best (97%). Hence, after combining these two feature extraction methods, better classification accuracy was taken. After trained with training data, testing data was used to validate the accuracy of DFNN model for EMG signal classification. The values of these statistical parameters for DFNN are given in Table 3. As it can be seen from Table 3, the sensitivity for myopathy value of each feature extraction methods is almost the same. For WPE, the sensitivity of neurogenic disorder is the best (97%) of all three feature extraction methods, but specificity is bad (77%). The specificity value for AR is the best (98%). Hence, after combining these two feature extraction methods, better classification accuracy was taken. The values of these statistical parameters for ANFIS are given in Table 4. As it can be seen from Table 4, the sensitivity for myopathy value for WPE and AR + DWT is the best (94%). For WPE, the sensitivity of neurogenic disorder is the best (97%) of all three feature extraction methods, but specificity is bad (77%). The specificity value of ANFIS with AR + DWT is the best (99%).

The performance demonstrated by the neuro-fuzzy knowledge model for EMG signal classification lies in following two special considerations: input variable choice and neuro-fuzzy method selection. The most valuable parameters which are derived from the EMG are dependent upon the signal processing technique used. The parameters, which are most excellent suited for EMG signal classification, should be used as the inputs of the model. For this reason, along with an AR, DWT and WPE are selected, because they are appropriate for classification of the nonlinear dynamics underlying muscle activity and allow insight into the development of complexity and regularity of the EMG. The advantages of ANFIS over other neuro-fuzzy methods make it as a tool to map the relationship between the parameters and the features. The combined use of

Table 2
Comparison of different feature extraction methods using MLPNN for EMG signal classification.

Statistical parameters	AR	DWT	WPT	AR + WPT	AR + DWT	DWT + WPT
Specificity (%)	97	96	74	96	97	96
Sensitivity (%) (myopathic)	90	89	90	90	90	88
Sensitivity (%) (neurogenic)	86	91	97	90	92	92
Total classification accuracy (% mean ± SD)	91 ± 9	92 ± 6	87 ± 5	92 ± 5	93 ± 6	92 ± 5

Table 3
Comparison of different feature extraction methods using DFNN for EMG signal classification.

Statistical parameters	AR	DWT	WPT	AR + WPT	AR + DWT	DWT + WPT
Specificity (%)	98	97	77	97	98	96
Sensitivity (%) (myopathic)	91	89	93	91	92	90
Sensitivity (%) (neurogenic)	87	92	97	90	92	92
Total classification accuracy (% mean ± SD)	92 ± 8	92.6 ± 6	89 ± 4	92.6 ± 5	94 ± 6	92.6 ± 4

Table 4
Comparison of different feature extraction methods using ANFIS for EMG signal classification.

Statistical parameters	AR	DWT	WPT	AR + WPT	AR + DWT	DWT + WPT
Specificity (%)	98	98	79	97	99	96
Sensitivity (%) (myopathic)	92	90	94	92	94	91
Sensitivity (%) (neurogenic)	88	92	97	93	92	93
Total classification accuracy (% mean ± SD)	92.6 ± 6	93.3 ± 3	90 ± 2	94 ± 3	95 ± 4	93.3 ± 2

AR parameter estimation, DWT, WPE (for deriving input variables) and ANFIS (for classification) results in the better performance of the derived classification system.

Since the fuzzy logic component gives the algorithm a degree of robustness, and combined it with the learning abilities of the neural network, the use of the ANFIS is desirable. A robust detection system is attractive given the potential for minor variations in muscle activity. The learning abilities of the ANFIS allow for the generation of membership functions and rule bases without any predetermined notions or knowledge base of the data set. ANFIS creates an accurate detection scheme that could be used for EMG pattern classification. Furthermore, the neural network characteristics of the ANFIS let for multiple efforts of generating the rule base and membership functions. This provides a detection scheme that has a low error rate in predicting the output, with the fewest number of rules and membership functions to keep computation times to a minimum.

The developed ANFIS model classifies EMG signals using different features as the input with an accuracy of greater than 95%. Although it allows for variability to the input signal, ANFIS is still based on neural network techniques and the capability to build a detection scheme based on repeatable patterns. In order for this system to be satisfactory as a classification structure, it must meet or go beyond the criteria as set by currently used interfaces. The ANFIS algorithm presents an excellent means to derive a classification system that could be quickly generated and customized to the individual user. Nevertheless, one disadvantage to the use of the ANFIS alone is that there is no memory in this classification scheme, and as a result no way to keep track of previous classification result to decide if the new output value is in agreement with the progression of different events. The ANFIS as described in this study becomes comparable to other existing technologies in classification of EMG signals. It is verified that the ANFIS with AR + DWT is robust enough to variations in the MUP to classify EMG signals well beyond what it is trained for. It is also repeatable over time to be used on data collected after the initial classification system is built.

Based on the results of the present paper and familiarity in EMG signal classification problems, we would like to emphasize the following:

i. The first observation is that with using different EMG features; ANFIS achieves the best classification rate (95%), superior to those obtained by DFNN (94%) and MLPNN (93%). This result shows that the ANFIS is able to gain better generalization ability than the other two for testing EMG patterns.
ii. The second observation is that by combining different features to each classifier, it can be found that the improvement on classification accuracy for each of the three classifiers is significant. For example, AR + DWT perform much better than AR and DWT alone, and the improvements in classification rate are 2%. This shows the great advantage of combining different features in our system, hence the combination of AR and DWT features well represent the EMG signals, and by using these features a good distinction between three classes can be obtained.
iii. The average success rate for the ANFIS with AR + DWT was highest success rate (95%), and for the MLPNN with WPE technique is the lowest (87%). Examining the classification success rate for each class, the highest success rate was obtained for the normal group using AR + DWT and the lowest for the normal group using WPE. This was the case for three classification methods.

4. Conclusion

The results of this study make obvious the feasibility of automated EMG signal classification using intramuscular EMG signals. This was accomplished using neuro-fuzzy computing techniques with three different feature extraction methods; namely AR, DWT and WPE. The neuro-fuzzy computing techniques were capable of classifying the EMG signals with a high degree of accuracy and repeatability. Also some conclusions related to the impacts of features on the classification of EMG signals were obtained through analysis of the neuro-fuzzy classifiers. So, the classification error rates showed that the ANFIS classification method is superior to the other methods for EMG signal classification. The classification accuracy for ANFIS with AR + DWT is 95%. Therefore we have concluded that the diagnostic decision support systems can become helpful when the physician's judgment is dependent on some other invasive and expensive tests; hence these techniques reduce the need for those assessments.

Acknowledgements

The author thanks to Dr. Mustafa Yilmaz for providing the EMG data utilized in this research.

References

[1] A. Fuglsang-Frederiksen, The utility of interference pattern analysis, Muscle and Nerve January (2000) 18–36.
[2] C.J. DeLuca, Towards Understanding the EMG Signal, 4th ed., Williams & Wilkinson, Baltimore, 1978.
[3] C. Krarup, Pitfalls in electrodiagnosis, Journal of Neurophysiology 81 (1999) 1115–1126.
[4] K.C. McGill, Optimal resolution of superimposed action potentials, IEEE Transactions on Bio-medical Engineering 49 (2002) 640–650.
[5] J. Kimura, Electrodiagnosis in Diseases of Nerve and Muscle: Principle and Practice, 3rd ed., Oxford University Press, New York, 2001.
[6] A. Fuglsang-Frederiksen, Power spectrum of the needle EMG in normal and diseased muscles, Methods in Clinical Neurophysiology 2 (1980) 1–8.
[7] H. Christensen, A. Fuglsang-Frederiksen, Power spectrum and turns analysis of the EMG at different voluntary efforts in normal subjects, Electroencephalography and Clinical Neurophysiology 64 (1986) 528–535.
[8] E.W. Abel, H. Meng, A. Forster, D. Holder, Singularity characteristics of needle EMG IP signals, IEEE Transactions on Bio-medical Engineering 53 (2006) 219–225.
[9] E.K. Richfield, B.A. Cohen, J.W. Albers, Review of quantitative and automated needle electromyographic analyses, IEEE Transactions on Bio-medical Engineering 28 (7) (1981) 506–514.
[10] C.D. Katsis, Y. Goletsis, A. Likas, D.I. Fotiadis, I. Sarmas, A novel method for automated EMG decomposition and MUP classification, Artificial Intelligence in Medicine 37 (2006) 55–64.
[11] C.I. Christodoulou, C.S. Pattichis, Unsupervided pattern recognition for the classification of EMG signals, IEEE Transactions on Bio-medical Engineering 46 (2) (1999) 169–178.
[12] G.H. Loudon, N.B. Jones, A.S. Sehmi, New signal processing techniques for the decomposition of EMG signals, Medical & Biological Engineering & Computing November (1992) 591–599.
[13] C.S. Pattichis, M.S. Pattichis, Time-scale analysis of motor unit action potentials, IEEE Transactions on Bio-medical Engineering 46 (11) (1999) 1320–1329.
[14] C.S. Pattichis, A.G. Elia, Autoregressive cepstral analyses of motor unit action potentials, Medical Engineering & Physics 21 (1999) 405–419.

[15] C.S. Pattichis, C.N. Schizas, L.T. Mittleton, Neural network models in EMG diagnosis, IEEE Transactions on Bio-medical Engineering 42 (1995) 486–496.
[16] M.H. Hassoun, C. Wang, A.R. Spitzer, NNERVE. Neural network extraction of repetitive vectors for electromyography—Part I: algorithm, IEEE Transactions on Bio-medical Engineering 41 (1994) 1039–1052.
[17] M.H. Hassoun, C. Wang, A.R. Spitzer, NNERVE. Neural network extraction of repetitive vectors for electromyography—Part II: performance analysis, IEEE Transactions on Bio-medical Engineering 41 (1994) 1053–1061.
[18] R.T. Lauer, B.T. Smith, R.R. Betz, Application of a neuro-fuzzy network for gait event detection using electromyography in the child with cerebral palsy, IEEE Transactions on Bio-medical Engineering 52 (9) (2005) 1532–1540.
[19] S.E. Hussein, M.H. Granat, Intention detection using a neuro-fuzzy EMG classifier, IEEE Engineering in Medicine and Biology November/December (2002) 123–129.
[20] K. Englehart, B. Hudgins, P.A. Parker, A wavelet-based continuous classification scheme for multifunction myoelectric control, IEEE Transactions on Bio-medical Engineering 48 (3) (2001) 302–311.
[21] A. Subasi, M. Yilmaz, H.R. Ozcalik, Classification of EMG signals using wavelet neural network, Journal of Neuroscience Methods 156 (2006) 360–367.
[22] F.H.Y. Chan, Y.-S. Yang, F.K. Lam, Y.-T. Zhang, P.A. Parker, Fuzzy EMG classification for prosthesis control, IEEE Transactions on Rehabilitation Engineering 8 (3) (2000) 305–311.
[23] R.F. Weir, A.B. Ajiboye, A multifunction prosthesis controller based on fuzzy-logic techniques, EMBS, 2003, in: Proceedings of the 25th Annual International Conference of the IEEE, vol. 2, 2003, pp. 1678–1681.
[24] L.A. Zadeh, Fuzzy sets, Information and Control 8 (3) (1965) 338–353.
[25] M. Sugeno, T. Yasukawa, A fuzzy-logic based approach to qualitative modeling, IEEE Transactions on Fuzzy Systems 1 (1) (1993) 7–31.
[26] C.A. Pena-Reyes, M. Siper, A fuzzy-genetic approach to breast cancer diagnosis, Artificial Intelligence in Medicine 17 (1999) 131–155.
[27] D. Nauck, R. Kruse, Obtaining interpretable fuzzy classification rules from medical data, Artificial Intelligence in Medicine 16 (1999) 149–169.
[28] A. Subasi, Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using feature extraction, Computers in Biology and Medicine 37 (2007) 227–244.
[29] J.S.R. Jang, ANFIS: adaptive network based fuzzy inference system, IEEE Transactions on Systems, Man and Cybernetics 23 (3) (1993) 665–683.
[30] S. Karlsson, J. Yu, M. Akay, Time–frequency analysis of myoelectric signals during dynamic contractions: a comparative study, IEEE Transactions on Bio-medical Engineering 47 (2000) 228–238.
[31] S. Karlsson, J. Yu, M. Akay, Enhancement of spectral analysis of myoelectric signals during static contractions using wavelet methods, IEEE Transactions on Bio-medical Engineering 46 (1999) 670–684.
[32] J.L. Semmlow, Biosignal and Biomedical Image Processing MATLAB-Based Applications, Marcel Dekker Inc, New York, NY 10016, 2004.
[33] J. Proakis, D. Manolakis, Digital Signal Processing Principles, Algorithms, and Applications, Prentice Hall, NJ, 1996.
[34] S.Y. Tseng, R.C. Chen, F.C. Chong, T.S. Kuo, Evaluation of parametric methods in EEG signal analysis, Medical Engineering & Physics 17 (1995) 71–78.
[35] S.M. Kay, S.L. Marple, Spectrum analysis—a modern perspective, Proceedings of the IEEE 69 (1981) 1380–1419.
[36] A. Isaksson, A. Wennberg, L.H. Zetterberg, Computer analysis of EEG signals with 4 parametric models, Proceedings of the IEEE 69 (4) (1981) 451–461.
[37] H. Akaike, A new look at the statistical model identification, IEEE Transactions on Automatic Control 19 (1974) 716–723.
[38] R. Polikar, The Wavelet Tutorial, http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html (accessed at 28.08.11).
[39] I. Daubechies, The wavelet transform, time–frequency localization and signal analysis, IEEE Transactions on Information Theory 36 (5) (1990) 961–1005.
[40] I. Daubechies, S. Mallat, A.S. Willsky, Introduction to the special issue on wavelet transforms and multiresolution signal analysis, IEEE Transactions on Information Theory 38 (1992) 529–532.
[41] M. Vetterli, C. Herley, Wavelets and filter banks: theory and design, IEEE Transactions on Signal Processing 40 (1992) 2207–2232.
[42] A. Subasi, Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients, Expert Systems with Applications 28 (2005) 701–711.
[43] A. Subasi, EEG signal classification using wavelet feature extraction and a mixture of expert model, Expert Systems with Applications 32 (2007) 1084–1093.
[44] M. Akay, Wavelet applications in medicine, IEEE Spectrum 34 (5) (1997) 50–56.
[45] A. Subasi, Epileptic seizure detection using dynamic wavelet network, Expert Systems with Applications 29 (2005) 343–355.
[46] A. Kandaswamy, C.S. Kumar, R.P. Ramanathan, S. Jayaraman, N. Malmurugan, Neural classification of lung sounds using wavelet coefficients, Computers in Biology and Medicine 34 (6) (2004) 523–537.
[47] R.R. Coifman, M.V. Wickerhauser, Entropy-based algorithms for best basis selection, IEEE Transactions on Information Theory 38 (1992) 713–718.
[48] S. Blanco, A. Figliola, R.Q. Quiroga, O.A. Rosso, E. Serrano, Time–frequency analysis of electroencephalogram series. III. Wavelet packets and information cost function, Physical Review E 57 (1998) 932–940.
[49] X. Hu, Z. Wang, X. Ren, Classification of surface EMG signal using relative wavelet packet energy, Computer Methods and Programs in Biomedicine 79 (2005) 189–195.
[50] F. Rosenblatt, Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms, Spartan, 1962.
[51] M. Minsky, S. Papert, Preceptrons, MIT Press, 1969.
[52] P. Melin, O. Castillo, Hybrid Intelligent Systems for Pattern Recognition Using Soft Computing, Springer-Verlag, Berlin, Heidelberg, 2005.
[53] A.H. Fielding, Cluster and Classification Techniques for the Biosciences, Cambridge University Press, The Edinburgh Building, Cambridge, UK, 2007.
[54] S. Haykin, Neural Networks: A Comprehensive Foundation, Macmillan, New York, 1994.
[55] M.J. Er, S. Wu, A fast learning algorithm for parsimonious fuzzy neural systems, Fuzzy Sets and Systems 126 (2002) 337–351.
[56] S. Wu, M.J. Er, Dynamic fuzzy neural networks—a novel approach to function approximation, IEEE Transactions on Systems, Man, and Cybernetics B 30 (2) (2000) 358–364.
[57] K.-L. Du, M.N.S. Swamy, Neural Networks in a Softcomputing Framework, Springer-Verlag London Limited, 2006.
[58] Y. Becerikli, Y. Oysal, A.F. Konar, Trajectory priming with dynamic fuzzy networks in nonlinear optimal control, IEEE Transactions on Neural Networks 15 (2) (2004) 383–394.
[59] Y. Becerikli, On three intelligent systems: dynamic neural, fuzzy and wavelet networks for training trajectory, Neural Computing & Applications 13 (4) (2004) 339–351.
[60] Y. Oysal, A comparative study of adaptive load frequency controller designs in a power system with dynamic neural network models, Energy Conversion and Management 46 (2005) 2656–2668.
[61] Y. Becerikli, A.F. Konar, T. Samad, Intelligent optimal control with dynamic neural networks, Neural Networks 16 (2003) 251–259.
[62] M.J. Er, T.P. Tan, S.Y. Loh, Control of a mobile robot using generalized dynamic fuzzy neural networks, Microprocessors and Microsystems 28 (2004) 491–498.
[63] Y. Oysal, Y. Becerikli, A.F. Konar, Generalized modeling principles of a nonlinear system with a dynamic fuzzy network, Computers and Chemical Engineering 27 (2003) 1657–1664.
[64] N.K. Kasabov, Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering, A Bradford Book, The MIT Press, Cambridge, MA/London, England, 1998.
[65] J.S.R. Jang, C.T. Sun, Neuro-fuzzy modeling and control, Proceedings of the IEEE 83 (3) (1995) 378–406.
[66] B. Kosko, Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence, Prentice-Hall, 1992.
[67] A. Kandel, Fuzzy Expert Systems, CRC Press Inc, 1992.
[68] M. Sugeno, G.T. Kang, Structure identification of fuzzy model, Fuzzy Sets and Systems 28 (1988) 15–33.
[69] J.-S.R. Jang, C.-T. Sun, E. Mizutani, Neurofuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence, Prentice-Hall, 1997.
[70] R. Zhang, G. McAllister, B. Scotney, S. McClean, G. Houston, Combining wavelet analysis and Bayesian networks for the classification of auditory brainstem response, IEEE Transactions on Information Technology in Biomedicine 10 (3) (2006) 458–467.