Biomedical Signal Processing and Control 18 (2015) 138–144

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Comparison of decision tree algorithms for EMG signal classification using DWT

Ercan Gokgoz*, Abdulhamit Subasi

International Burch University, Faculty of Engineering and Information Technologies, 71000 Sarajevo, Bosnia and Herzegovina

ARTICLE INFO

Article history:
Received 16 September 2014
Received in revised form
16 December 2014
Accepted 19 December 2014
Available online 24 January 2015

Keywords:
Electromyography (EMG)
Motor unit action potentials (MUAPs)
Random forest
C4.5
CART
Multi-scale principle component analysis (MSPCA)
Discrete wavelet transform (DWT)

ABSTRACT

Decision tree algorithms are extensively used in machine learning field to classify biomedical signals. De-noising and feature extraction methods are also utilized to get higher classification accuracy. The goal of this study is to find an effective machine learning method for classifying ElectroMyoGram (EMG) signals by applying de-noising, feature extraction and classifier. This study presents a framework for classification of EMG signals using multiscale principal component analysis (MSPCA) for de-noising, discrete wavelet transform (DWT) for feature extraction and decision tree algorithms for classification. The presented framework automatically classifies the EMG signals as myopathic, ALS or normal, using CART, C4.5 and random forest decision tree algorithms. Results are compared by using numerous performance measures such as sensitivity, specificity, accuracy, *F*-measure and area under ROC curve (AUC). Combination of DWT and random forest achieved the best performance using *k*-fold cross-validation with 96.67% total classification accuracy. These results demonstrate that the proposed approach has the capability for the classification of EMG signals with a good accuracy. In addition, the proposed framework can be used to support clinicians for diagnosis of neuromuscular disorders.

1. Introduction

The electromyography (EMG) signal is a biomedical signal that consists of electrical currents generated in contraction and relaxation phase of muscles. The nervous system manages the muscle activity according to the structure. There are two types of structures which are named as anatomical and physiological. Specifications of muscles change according to the type of structure [1]. EMG signals are affected from environmental noise and signals while passing on different tissues. In addition, motor unit action potentials (MUAPs) can be affected by different signals during the signal acquisition. Besides, MUAPs provide essential information for diagnosis of neuromuscular disorders [2]. Thus, noise seriously distorts the EMG signal during the collection and recording process. These are the reasons that make EMG signal more complex and why they need to be de-noised. In our previous study, significantly better results were obtained by removing noise from the EMG signal waveform using multi-scale principal component analysis (MSPCA) technique [3].

Not only de-noising methods but also feature extraction methods are effective for obtaining higher classification accuracy. In this study discrete wavelet transform (DWT) is utilized as a feature extraction method by decomposing the signal into different frequency bands. New approaches and recent improvements in signal processing techniques have made the area more practical and reliable to develop advanced EMG signal analysis [4–6]. As a matter of fact, analysis of EMG signals is becoming a new trend in biomedical signal processing by using powerful and advanced signal processing techniques. EMG signal classification with a higher accuracy is important because outcomes are crucial in clinical diagnosis of neuromuscular disorders.

Classification of EMG signals by utilizing reliable and robust methods is a rising trend in biomedical engineering. Popularity of EMG signal analysis is tightly related to the usage field in real life clinical biomedical applications. Coherence between methods and techniques performs efficiently on diagnosis of neuromuscular disorders [1,7–10]. Higher classification accuracies can be obtained by readjusting used kernel parameters of the classifier [11,12]. Machine learning techniques such as artificial neural networks (ANN), dynamic recurrent neural networks (DRNN), support vector machines (SVM) and fuzzy logic systems are used for diagnosis of neuromuscular disorders [13]. Reasonable accuracy rates are obtained by applying spectrum matching and by extracting principle components [1]. Adaptive neuro-fuzzy inference system

* Corresponding author at: International Burch University, Faculty of Engineering and Information Technologies, Francuske Revolucije bb. Ilidza, Sarajevo 71000, Bosnia and Herzegovina. Tel.: +387 33 944 405; fax: +387 33 782 131.
E-mail addresses: egokgoz@ibu.edu.ba (E. Gokgoz), asubasi@ibu.edu.ba (A. Subasi).

http://dx.doi.org/10.1016/j.bspc.2014.12.005
1746-8094/

(ANFIS), multilayer perceptron neural networks (MLPNN), dynamic fuzzy neural network (DFNN) and combined feature extraction methods, such as Autoregressive, discrete wavelet transform and wavelet packed energy, presented due to their consistency in classification of EMG signals [14]. SVM classification technique optimized with particle swarm optimization (PSO) improves the accuracy of EMG signal classification [9]. Subasi et al. [15] developed a classification system with Autoregressive (AR) feature extraction method to produce inputs of feed forward error back propagation artificial neural networks (FEBANN) and wavelet neural networks (WNN) classifiers. The total accuracy for the WNN was 90.7% while 88% for FEBANN.

The contribution of this paper is to investigate the classification performance of different decision tree algorithms with discrete wavelet transform (DWT) coefficients using intramuscular EMG signals for diagnosis of neuromuscular disorders. EMG signal classification accuracy improved by utilizing MSPCA de-noising and DWT feature extraction methods. The effects of different signal processing and decision tree algorithms are explained compared and discussed using different performance measures.

The rest of the paper is prepared as follows: in the next section, the subjects and data are explained and different methods such as MSPCA, DWT, random forest, CART and C4.5 methods are presented. In Section 3 complete experimental results in respect to different classification accuracy measurements such as area under ROC curve, *F*-measure and total classification accuracy are presented. In Section 4, discussion is given on the impact of used de-noising and classification methods. Finally, the conclusions are given in Section 5.

2. Materials and methods

2.1. EMG data

EMG signals were gathered at low voluntary and constant level of contraction which is just above threshold by using concentric needle electrode. The EMG signals were collected from five different places at deep, medium and low level of insertions. The high and low pass filters were set at 2 Hz and 10 kHz on the EMG amplifier which can be thought of as boundaries of data. The gathered signal contains control group, myopathy and ALS patients. The control group contains 10 normal subjects aged 21–37 years, 4 of them are females and rest are males. 6 of the 10 normal patients were in very good physical shape, and 3 of the 10 normal patients were in general good shape and the remaining one not in good shape. The myopathy group includes 7 patients. 2 of 7 are females and the rest are males who are between 19 and 63 years old. All of them had clinical and electrophysiological signs of myopathy. The ALS group includes 8 patients. 4 of 8 are females and the other 4 are males who are between 35 and 67 years old. In addition, no one in the control group had signs or history of neuromuscular disorders. The data were collected from brachial biceps and medial vastus since they were the most frequently investigated in the two patient groups [10,16]. Window length of the signal is 12,500 and sampling rate is 20,000.

2.2. MSPCA de-noising method

Multi-scale principal component analysis (MSPCA) combines the characteristics and ability of principal components analysis (PCA) to de-correlate the variables by procuring a stable interrelationship. Wavelet analysis was employed to get beneficial features and closely de-correlate the autocorrelated values. At each scale, MSPCA calculates the PCA of the wavelet coefficients and combines the results at defined scales. MSPCA is feasible for signal modeling which changes over time and frequency and it can be thought as an advantage of the multiscale approach [17].

PCA method is used in numerous fields of science and engineering [10,17–23]. PCA is used to analyze shorter signal segments to improve practicality of biomedical signals like EMG [24]. This approach is applied in different studies such as data reduction, beat detection, classification, signal segmentation and feature extraction [3,25]. The expected value of processed signals can be obtained by selecting principal components (PC) based on energy features of matrices. PCA uses these matrices for signal de-noising. PCA de-noising approach can be explained as follows: holding the principal components which have the highest variance to rebuilt the decomposed signal. Meanwhile, the noise corresponding to the low variance components can deliberately be omitted hence the noise component in the observed signal is reduced. The decision of multi-scale matrices and eigenvalues includes the expected energy [10,17,21]. PCA transforms the data matrix statistically using diagonalizing method via covariance matrix. The process extracts the correlation and interrelation between the variables in the data matrices. If any similarity is detected, and if the measured and evaluated variables are related in the first few entities, system determines the correlation between the variables. The PCA method is used one by one for the coefficients at each scale [17].

Multiscale PCA (MSPCA) includes the capacity, capability and characteristic of PCA to get the correlation between orthonormal wavelets and variables. The quantities of each variable or column are decomposed into its wavelet coefficients using the same orthonormal wavelet in order to combine the advantages of PCA and wavelets. Details are given in [3,17].

2.3. Feature extraction using discrete wavelet transform (DWT)

The wavelet transform (WT) can be used as another way to describe and extract features of a waveform which changes over time. To handle this situation the waveform is divided into segments of scale instead of sections of time [26]. In wavelet transform, the capability of the transformation depends strongly on mother wavelet $\psi(t)$ selection which is the certain function and can be represented by the following equation (1)

$$\psi(t) = \frac{1}{\sqrt{S}}\psi\left(\frac{t-u}{S}\right) \tag{1}$$

where s used as scale and u as a translation parameter. These parameters can be produced in time with various frequencies and midpoint localities, which are called as *baby wavelets* or *wavelet atoms*. Handling of correlations at different frequencies are gained by a form of the signal $x(t)$ with following wavelets as

$$W_x(u, S) = \frac{1}{\sqrt{S}}\int_{-\infty}^{+\infty} x(t)\psi*\left(\frac{t-u}{S}\right)dt \tag{2}$$

WT coefficients also can be used to define the frequency in the signal, thus information of the signal $x(t)$ can be detected in time and frequency

$$(s_j, u_k) = (2^j, k2^j : j, k \in \mathbb{Z}) \tag{3}$$

DWT utilizes two different functions which are scaling and wavelet functions. These functions are interrelated with low- and high-pass filters. Each step combines two filters and two down-samplers. These high- and low-pass filters observe details, D_1 and approximations, A_1 for the first one, by using the down-sampled outputs. There are set of processes for each approximation, from A_1 to A_6, which are decomposition, processing and reconstruction, orderly. Details are given in [14,27,28].

Features of the EMG signal can be extracted by means of discrete wavelet coefficients in time and frequency using some mathematical techniques. Signals can be characterized by their statistical information to make them more applicable. Those statistical features reduce dimensionality of the signal. Better performance is obtained by reducing dimension of EMG signal in numerous studies [14,23,28]. In this study, EMG signals are represented by using following features of coefficients c with following formulas [29]:

(1) Mean of the coefficients for each sub-band.

$$\text{Mean} = \frac{\sum_{i=1}^{n} C_i}{n} \tag{4}$$

(2) Average power of the wavelet coefficients in each sub-band.

$$\text{Average} = \frac{1}{N}\sum_{i=0}^{N}(c_i)^2 \tag{5}$$

(3) Standard deviation of the coefficients in each sub-band

$$\text{Standard deviation} = \sqrt{\frac{\sum_{i=1}^{n}(c_i-\mu)^2}{n}} \tag{6}$$

(4) Ratio of mean values of neighbor sub-bands

$$\text{Ratio} = \frac{\sum_{i=1}^{n} c_i/n}{\sum_{j=1}^{n} c_j/n} \tag{7}$$

1st and 2nd features are extracted for evaluation of the frequency distribution of the signal. 3rd and 4th features are extracted for the evaluation of the changes in the frequency distribution. Seven different features are extracted from (1), (2) and (3); six different features are extracted from (4). Hence, totally 27 features are extracted which consist of mean, average power of the wavelet coefficients in each sub-band, standard deviation and ratio of mean values of neighbour sub-bands. These extracted features are used on wavelet coefficients to make it more applicable and feasible [23,30–33].

These features are calculated for D1–D6 and A6 which are frequency bands and used as input to classifiers [34]. Higher classification accuracy and lower computation cost are obtained with "db4" wavelet filter.

2.4. Decision tree algorithms

2.4.1. Classification and regression trees (CART)

The classification and regression tree (CART) is a decision tree algorithm developed by Breiman et al. [35]. CART uses a recursive partitioning technique which has splitting criteria to create nodes. Tree is built by using created and split nodes related to splitting criteria and function. Before applying the split criteria, we need to know what the best split point is. The quality of the splitting criteria is measured by a function which is obtained by processing variance function. Generated function is applied to each split point to calculate the best point for splitting [36]. Different criteria can be used to define the splits f_i such as Gini which is defined as [35]:

$$\text{Gini}(t) = 1 - \sum_i f_i^2 \tag{8}$$

CART has four process steps. First one is to build a tree by using recursive splitting of nodes where splitting criteria reached. Second one is to stop tree building process after the learning dataset fitted and formed according to the attributes. Third one is tree pruning, through cutting important nodes off to produce simpler trees. The last one is the selection of optimal tree from the sequence of pruned trees, which are not over fitting with the information when fitting with the learning dataset [37]. Efficient results are obtained by assigning minimum number terminal nodes as 2 and number of folds in cross-validation are 5 for pruning.

2.4.2. C4.5

C4.5 is a popular decision tree algorithm which is an upgraded version of ID3 decision tree algorithm that is developed by Quinlan, too [38]. ID3 is upgraded with some features such as categorization of continuous attributes, handling missing values, pruning of decision trees and rule derivation.

C4.5 algorithm starts by selection of attribute to test root of the tree. Each attribute is calculated statistically to check feasibility by comparing training examples. The most proper attribute is used to test root node. Branch nodes of the root node are created after examining each potential attribute values in respect to training examples. This process is repeated for each branch node to select the most proper attribute by testing associated training examples. Essence of C 4.5 algorithm is to choose the correct attribute according to the information gain to test nodes [39,40]. Details are given in [41,42]. The confidence factor for pruning is 0.6. The minimum number of values in a leaf is 2. To reduce the error of pruning, data on the tree folded 3 times.

2.4.3. Random forest

In the area of bioinformatics and biomedical engineering, the random forest (RF) classifier, which is a set of decision trees, became a popular option in the machine learning processes. Increased use of random forest can be seen in computational biology field, owing to its advantages in dealing with limited sample size, complex data structures and multidimensional feature space [43].

Random forest uses numerous independent decision trees which are created by randomly chosen variables. The independent trees are built by an algorithm. After trees are created they vote to find out the most popular class [44]. The algorithm guarantees that all trees in the forest are different. Randomness is applied in two steps. First step is to use different bootstrap sample data to build each tree. Second step is to choose a subset of predictors randomly then splitting each node of trees with the best subset instead of all predictors [45]. There are two reasons to have bootstrap step. First one is classification accuracy increment when random features are used and the second one is to reduce generalization error [44]. Random selection of splitting does better than bagging in terms of generalization error [46]. The strength of the individual tree classifiers is important on classification. Sometimes, this algorithm works better than some other classifiers such as support vector machines, neural networks and discriminant analysis [47]. Even though random forest was presented as decision trees, it can work with other classifiers. In this study, the optimum number of trees is defined as 30 for efficient results.

3. Experimental results

EMG signal decomposition into MUAPs and classification are frequently studied as a machine learning problem [15]. It is hard to decompose an EMG signal into MUAP since the waveform includes all electrical activities of muscles. In order to decompose the signal we used DWT method whose reliability is defined in [33]. The main goal of the analysis of EMG signals is to get better classification accuracy and to find efficient feature extraction techniques. Our EMG data are collected from 25 patients. The control group contains 10, myopathy group contains 7 and ALS group contains 8 patients [16].

In this study, DWT feature extraction and decision tree algorithms have been applied for EMG signal classification (Fig. 1). DWT feature extraction technique is applied after MSPCA de-noising to observe the effect on both the relation between methods and the estimation of the classification error. The EMG data is separated into training data set and test data set, since both sets have to be

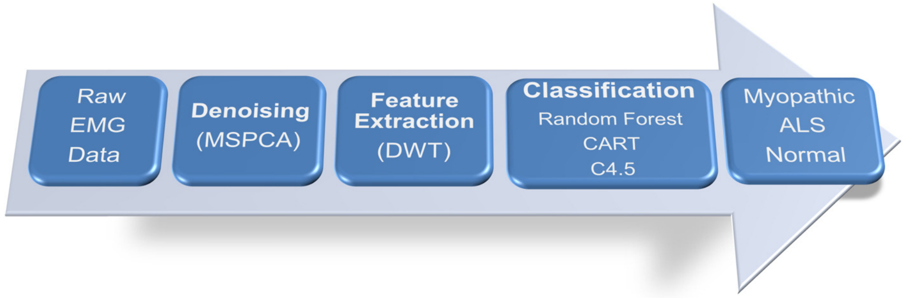

Fig. 1. Block diagram of proposed system.

chosen independently. The classification model is built via training set. The test data is used for validation.

K-fold cross-validation method is a well-known and reliable method for predicting the error rate of a classification technique. It is used by many researchers to decrease the bias which is related to the random sampling of the produced data sets [48–50]. *K*-fold cross-validation arbitrarily divide the data into a determined number of subsets which are called as folds. Cross-validation accuracy is formulated as follows

$$\mathrm{CVA} = \frac{1}{k}\sum_{j=1}^{k} A_i \tag{9}$$

where k is the number of employed folds, and A_i is the accuracy of each fold [51].

The number of true negatives (TN), true positives (TP), false positives (FP) and false negatives (FN) are used as a measurement to evaluate the performance of classifiers. Different definitions are used to explain the results on different domains. The specificity, sensitivity and accuracy are widely used in diagnostic and detection tests and defined as follows

$$\text{Sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \times 100\% \tag{10}$$

$$\text{Specificity} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}} \times 100\% \tag{11}$$

$$\text{Accuracy} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TN}+\mathrm{TP}+\mathrm{FN}+\mathrm{FP}} \times 100\% \tag{12}$$

F-measure is another measurement which is used to evaluate performance [9] and formulated as follows:

$$F\text{-measure} = \frac{2\mathrm{TP}}{2\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} \tag{13}$$

Three different classification algorithms are applied namely CART, C4.5 and random forest, with and without MSPCA de-noising method. Total classification accuracy is the number of correctly classified data which is divided by the number of all data. Results are presented in Table 1 and illustrated in Fig. 2.

Effect of MSPCA de-noising can be seen easily by comparing results in Tables 1 and 2 and Fig. 2. Random forests give the best classification accuracy with 96.67%, C4.5 is the second best with 91.11% and CART is the last one with 90.88% accuracy which is

Table 1
Comparison of performance results of decision tree based classifiers.

Statistical parameters	CART	C4.5	Random forest
Specificity (%)	88.7	87.7	94.75
Sensitivity (Myopathic) (%)	90	90.7	95.66
Sensitivity (ALS) (%)	94	95	99.58
Total classification accuracy (%)	90.88	91.11	96.67

slightly lower than C4.5. Besides, random forest is better than the others for EMG pattern classification with MSPCA de-noising and DWT decomposition method. Advantage of using DWT and random forest methods is to have capability of modeling high dimensional feature space. Numerous studies have proven that random forest and the other decision tree algorithms are reliable and accurate since they are widely used as a robust classification algorithm [47,52,53]; due to the fact that features of EMG data need to be extracted and de-noised to obtain better performance from classifiers [54]. Hence, more accurate data are provided to the classifiers by using DWT decomposition and MSPCA de-noising. The total accuracy for random forest increased from 72.33% to 96.67% after de-noising with MSPCA. 23.34% improvement in accuracy proves that the better performance is achieved by using MSPCA de-noising.

The performance of classifiers is also measured by receiver operating characteristic (ROC) curve. The curve is generated by plotting true positives as percentage of all positives and negative ones in the sample. The number of true and false positives in the test set is counted in each fold of cross-validation in order to plot the results on ROC curve [55]. Thus, classification performance can be evaluated by the area, which is under the ROC curve (AUC). The mean of AUC is another way of performance evaluation which shows reliability of the results by using input data [56–61]. Since AUC performance result is a measure of total accuracy of independent thresholds, it is accepted as the index of performance [48,51,56,57,62–66]. AUC of random forest is the best again and followed by C4.5 which is slightly better than CART. Even though, total accuracy of C4.5 is better than CART, CART classify ALS better than C4.5. As a result, it can be said that CART is more convenient than C4.5 just for ALS data. Sensitivity of ALS is higher than Normal and Myopathic for all classifiers.

F-measure is another performance evaluation method which is coincident with our classification accuracy results. As it can be seen

Table 2
Confusion matrix for decision tree based classifiers.

Actual[a]/Predicted	CART			C4.5			Random forest		
	ALS	Myopathic	Normal	ALS	Myopathic	Normal	ALS	Myopathic	Normal
ALS	282	5	12	285	2	13	1195	4	16
Myopathic	3	270	22	2	272	25	1	1148	47
Normal	15	25	266	13	20	263	4	48	1137

[a] According to the medical expert.

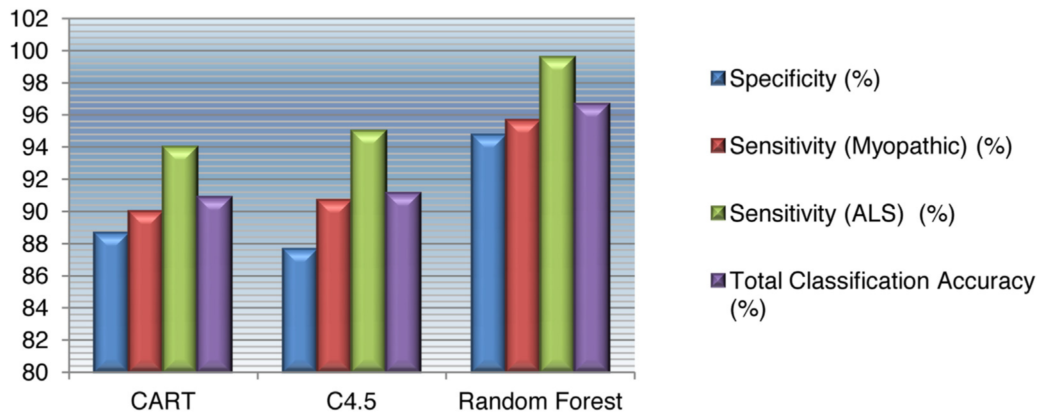

Fig. 2. Graphical representation of decision tree based classifiers' performances.

Table 3
AUC results for decision tree based classifiers.

	CART	C4.5	Random forest
Normal	0.917	0.921	0.995
Myopathic	0.937	0.947	0.996
ALS	0.967	0.96	1
Total	0.94	0.942	0.997

from Tables 1 and 3, total accuracy and AUC results are very close to each other for random forest algorithm which confirms the reliability of results. Hence obtained AUC result is 0.997 which is coincident with total accuracy (96.67%).

4. Discussion

Different EMG signal classification studies can be found in literature [4,8,22,54,67,68]. Data collection method is important to compare EMG classification results precisely. While some studies use surface EMG data, some others use intramuscular EMG data. Besides, in one side, simulated EMG data are used; on the other side, real EMG data are used. Since data collection methods are not the same, it is not easy to compare classification accuracies of these methods altogether (Table 4).

Random forest is used to classify biomedical data since it has advantages in dealing with limited sample size, complex data structures and multidimensional feature space [43]. Furthermore, multi-scale principal component analysis is utilized for de-noising, meanwhile number of trees have been increased to raise the dimensionality and number of variations. Contribution of this study is to design an efficient classification tool by applying discrete wavelet transform and MSPCA approach for diagnosis of neuromuscular disorder. Random forest classifier achieved the best classification accuracy rate with 96.67%; C4.5 is following Random forest with 91.11% classification accuracy. CART is the last one in these three decision tree algorithms with 90.88% accuracy. This result shows that the random forest is better than others for EMG signal classification to diagnose neuromuscular disorder. The improvement on total classification accuracy for all three classifiers is significant with MSPCA de-noising method as seen from the tables. The classification accuracy of random forest without MSPCA de-noising is 73.33%, with MSPCA de-noising 96.67%.

Table 4
F-measure of decision tree based classifiers.

	CART	C4.5	Random forest
Normal	0.878	0.875	0.952
Myopathic	0.908	0.916	0.958
ALS	0.942	0.942	0.99
Total	0.909	0.911	0.967

Table 5
Comparison of results and methods with other studies which used same dataset.

Studies	Approach	Accuracy (%)
Parsaei and Stashuk [70]	Adaptive gap-based Duda and Hart (AGDH) method used for component analysis on motor unit potential trains	91.3%
Filho et al. [71]	Multiscale multidimensional parser applied and results evaluated percent root mean for encoding EMG signal	91.1% of the false positives, 95% correctly identified
Parsaei and Stashuk [72]	Gap statistics and motor unit potential trains utilized together for decomposition of EMG signal	93.8%
Yejin et al. [73]	Montreal and fuzzy expert decomposition techniques applied	95.7% (at %20 contraction level)
Marateb et al. [74]	A novel decomposition method applied on motor unit potential trains	>90%
Nikolic and Krarup [75]	A novel EMGtools decomposition method applied on constituent MUAPs and firing patterns	95% correctly identified
Our proposal	Features of denoised EMG signal extracted by using discrete wavelet transform then classified with random forest	96.67%

Total classification accuracy result of our approach is the best when compared to other studies which used the same EMG data set, because, MSPCA de-noising approach is not utilized in other studies. According to comparison results it can be seen easily that appropriate and coherent decomposition and de-noising methods perform better on EMG signal classification. On the other hand simulated datasets are used in some studies since it is hard to find compact and consistent EMG dataset. As a result of the inflexibility, different studies used simulated EMG data and higher classification accuracies have been obtained by using simulated EMG data [73]. But it can be said that, since simulated data is not a real one, results of these studies cannot be reliable for clinical usage or cannot be used as a decision support tool. Even if they cannot be reliable enough, they can help to understand structure and tendency of EMG signal [69]. Comparison of results and methods with other studies which use the same EMG dataset are presented in Table 5.

5. Conclusion

In this study, we have found out considerable coherence between classifiers, de-noising and decomposition methods, which are proved by the results. This was revealed using classifiers such as random forest, CART, C4.5 and DWT decomposition method with MSPCA de-noising. The contribution of this study is to develop an efficient classification algorithm for intramuscular EMG signals and to find out an appropriate decision tree algorithm. This study achieved significantly better performance by using MSPCA de-noising. Random forest classifier with DWT feature extraction method accomplished better performance with MSPCA de-noising over the three EMG signal patterns: normal, myopathic and ALS. While examining coherence between methods, we realized how higher accuracy can be achieved by rearranging the parameters of the classifiers. Generally it is not efficient to use default parameters of classifiers since they have been set for common purpose. The classification accuracy with MSPCA improved to 96.67% which was 73.33% without MSPCA using DWT and random forest combination. Besides, classification accuracy of other classifiers also increases by using MSPCA de-noising. DWT feature extraction and random forest classifier with MSPCA de-noising can be helpful to the clinician for diagnosis of neuromuscular disorders.

References

[1] R. Begg, D.T.H. Lai, M. Palaniswami, Computational Intelligence in Biomedical Engineering, CRC Press, Boca Raton, 2008.
[2] K.C. McGill, Optimal resolution of superimposed action potentials, IEEE Trans. Biomed. Eng. 49 (7) (2002) 640–650.
[3] E. Gokgoz, A. Subasi, Effect of multiscale PCA de-noising on EMG signal classification for diagnosis of neuromuscular disorders, J. Med. Syst. 38 (4) (2014).
[4] C.D. Katsis, Y. Goletis, A. Likas, D.I. Fodiatis, I. Sarmas, A novel method for automated EMG decomposition and MUAP classification, Artif. Intell. Med. 37 (1) (2006) 55–64.
[5] G.H. Loudon, N.B. Jones, A.S. Sehmi, New signal processing techniques for the decomposition of EMG signals, Med. Biol. Eng. Comput. 30 (6) (1992) 591–599.
[6] C.S. Pattichis, M.S. Pattichis, Time-scale analysis of motor unit action potentials, IEEE Trans. Biomed. Eng. 46 (11) (1999) 1320–1329.
[7] G. Bienvenu, L. Kopp, Optimality of high resolution array processing using the eigensystem approach, IEEE Trans. Acoust. Speech Signal Process. 31 (5) (1983) 1235–1248.
[8] A. Subasi, Classification of EMG signals using combined features and soft computing, Appl. Soft Comput. 12 (8) (2012) 2188–2198.
[9] A. Subasi, Classification of EMG signals using PSO optimized SVM for diagnosis, Comput. Biol. Med. 43 (5) (2013) 576–586.
[10] M. Misra, H.H. Yue, S.J. Qin, C. Ling, Multivariate process monitoring and fault diagnosis by multi-scale PCA, Comput. Chem. Eng. 26 (9) (2002) 1281–1293.
[11] C.L. Huang, C.J. Wang, A GA-based attribute selection and parameter optimization for support vector machine, Expert Syst. Appl. 31 (2) (2006) 231–240.
[12] S.Y. Ho, L.S. Shu, J.H. Chen, Intelligent evolutionary algorithms for large parameter optimization problems, IEEE Trans. Evol. Comput. 8 (6) (2004) 522–541.
[13] M.B.I. Raez, M.S. Hussain, F. Mohd-Yasin, Techniques of EMG signal analysis: detection, processing, classification and applications, Biol. Proced. Online 8 (2006) 11–35.
[14] A. Subasi, Medical decision support system for diagnosis of neuromuscular disorders using DWT and fuzzy support vector machines, Comput. Biol. Med. 42 (8) (2012) 806–815.
[15] A. Subasi, M. Yilmaz, H.R. Ozcalik, Classification of EMG signals using wavelet neural network, J. Neurosci. Methods 156 (1–2) (2006) 360–367.
[16] M. Nikolic, Findings and Firing Pattern Analysis in Controls and Patients with Myopathy and Amytrophic Lateral Sclerosis, Faculty of Health Science, University of Copenhagen, Copenhagen, 2001.
[17] B.R. Bakshi, Multiscale PCA with application to MSPC monitoring, AIChE J. 44 (1998) 1596–1610.
[18] M.P.S. Chawla, A comparative analysis of principal component and independent component techniques for electrocardiograms, Neural Comput. Appl. 18 (2009) 539–556.
[19] H. Hotelling, Analysis of a complex of statistical variables into principal components, J. Edu. Psychol. 24 (6) (1933) 498–520.
[20] I.T. Joliffe, Principal Component Analysis, 2nd ed., Springer, New York, NY, 2002.
[21] B.R. Bakshi, Multiscale analysis and modeling using wavelets, J. Chemom. 13 (1999) 415–434.
[22] E.N. Kamayuako, D. Farina, K. Yoshida, W. Jensen, Relationship between grasping force and features of single channel intramuscular EMG signals, J. Neurosci. Methods 185 (1) (2009) 143–150.
[23] A. Kandaswamy, C.S. Kumar, R.P. Ramanathan, S. Jayaraman, N. Malmurugan, Neural classification of lung sounds using wavelet coefficients, Comput. Biol. Med. 34 (6) (2004) 523–537.
[24] M. Kotas, Application of projection pursuit based robust principal component analysis to ECG enhancement, Biomed. Signal Process. Control 1 (4) (2006) 289–298.
[25] F. Castells, P. Laguna, L. Sornmo, A. Bollmann, J. Roig, Principal component analysis in ECG signal processing, EURASIP J. Adv. Signal Process. (2007) 21p, Article ID: 74580.
[26] J.L. Semmlow, Biosignal and Biomedical Image Processing, Marcel Dekker, New York, 2004.
[27] M. Akay, Wavelet applications in medicine, IEEE Spectr. 34 (1997) 50–56.
[28] A. Subasi, EEG signal classification using wavelet feature extraction and a mixture of expert model, Expert Syst. Appl. 32 (2007) 1084–1093.
[29] D.G. Manolakis, V.K. Ingle, S.M. Kogon, Statistical and Adaptive Signal Processing, McGraw-Hill, MA, 2005.
[30] J.T. McClave, P.G. Benson, T. Sincich, A First Course in Business Statistics, Pearson College Division, 2008.
[31] X. Peng, A.K. Chan, Fast and robust neural network based wheel bearing fault detection with optimal wavelet features Proceedings of the 2002 International Joint Conference on Neural Networks, vol. 3, 2002, pp. 2076–2080.
[32] F. Morchen, Time series feature extraction for data mining using DWT and DFT International Conference on Machine Learning and Applications, vol. 33, 2003, pp. 735–739.
[33] A. Phinyomark, C. Limsakul, P. Phukpattaranont, Application of wavelet analysis in EMG feature extraction for pattern classification, Meas. Sci. Rev. 11 (2) (2011) 45–52.
[34] X. Qi, Z. Hui, W. Yongji, H. Jian, Fuzzy support vector machine for classification of EEG signals using wavelet-based features, Med. Eng. Phys. 31 (2009) 858–865.
[35] L. Breiman, J.H. Friedman, R.A. Ohlsen, C.I. Stone, Classification and Regreesion Tree, Wadsworth International Group, Belmont, CA, 1984, pp. 43–49.
[36] H. Shien-Ping, C. Chih-Chou, Process monitoring with ICA-based signal extraction technique and CART approach, Qual. Reliab. Eng. Int. 25 (2009) 631–643.
[37] R.J. Lewis, An introduction to classification and regression tree (CART) analysis, in: Annual Meeting of the Society for Academic Emergency Medicine, San Francisco, CA, 2000.
[38] R.J. Quinlan, C4. 5: Programs for Machine Learning, Morgan Kaufmann, USA, 1993.
[39] T.M. Mitchell, Machine Learning, McGraw-Hill, 1997.
[40] J.R. Quinlan, Improved use of continuous attributes in C4.5, J. Artif. Intell. Res. 4 (1996) 77–90.
[41] A.B.M.S. Ali, S.A. Wasimi, Data Mining: Methods and Techniques, Thomson, Victoria, Australia, 2007.
[42] M.M. Mazid, A.B.M. Shawkat Ali, K. Stickle, Improved C4.5 algorithm for rule based classification, in: AIKED'10 Proceedings of the 9th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases, Wisconsin, 2010.
[43] Y. Qi, Random forest for bioinformatics, in: Z. Cua, Y. Ma (Eds.), Ensemble Machine Learning Methods and Applications, Springer, NJ, 2012, pp. 307–323.
[44] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.
[45] Z. Yi, J. Pan, Application of random forest to stellar spectral classification, in: 3rd International Congress on Image and Signal Processing, Yantai, 2010.
[46] T.G. Dietterich, An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization, Mach. Learn. 40 (2) (2000) 1–22.
[47] A. Liaw, M. Wiener, Classification and regression by random forest, R News 2 (2002) 18–22.
[48] C. Ambroise, G.J. McLachlan, Selection bias in gene extraction on the basis of microarray gene-expression data, Proc. Natl. Acad. Sci. U. S. A. (2002) 6562–6566.
[49] A. Subasi, Classification of EMG signals using combined features and soft computing, Appl. Soft Comput. 12 (8) (2012) 2188–2198.
[50] J.J. Sandvig, B. Mobasher, R. Burke, A survey of collaborative recommendation and the robustness of model-based algorithms, IEEE Comput. Soc. Tech. Comm. Data Eng. 31 (2) (2008) 3–13.
[51] R. Zhang, G. McAllister, B. Scotney, S. McClean, Combining wavelet analysis and bayesian networks for the classification of auditory brainstem response, IEEE Trans. Inf. Technol. Biomed. 10 (2006) 458–467.
[52] R. Díaz-Uriarte, S. Alvarez de Andrés, Gene selection and classification of microarray data using random forest, BMC Bioinform. 7 (3) (2006) 142–146.
[53] V. Svetnik, A. Liaw, C. Tong, J.C. Culbertson, R.P. Sheridan, B.P. Feuston, Random forest: a classification and regression tool for compound classification and QSAR modeling, J. Chem. Inf. Model. 43 (6) (2003) 1947–1958.
[54] A.P. Dobrowolski, M. Wierzbowski, K. Tomczykiewicz, Multiresolution MUAPs decomposition and SVM-based analysis in the classification of neuromuscular disorders, Comput. Methods Programs Biomed. 107 (3) (2011) 393–403.
[55] I. Witten, E. Frank, Data Mining, Practical Machine Learning Tools and Techniques, Morgan Kaufmann Publishers/Elsevier, San Francisco, CA, 2005.
[56] J.A. Handley, B.J. McNeil, The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology 143 (1982) 29–36.
[57] N.A. Obuchowski, Receiver operating characteristic curves and their use in radiology, Radiology 229 (2003) 3–8.
[58] J.A. Swets, ROC analysis applied to the evaluation of medical imaging techniques, Invest. Radiol. 14 (1979) 109–121.
[59] J.M. Deleo, Receiver operating characteristic laboratory (ROCLAB): software for developing decision strategies that account for uncertainty, in: Proceedings of the Second International Symposium on Uncertainity Modelling and Analysis, 1993, pp. 318–325.
[60] A. Devos, L. Lukas, J.A. Suykens, L. Vanhamme, A.R. Tate, F.A. Howe, C. Majos, A. Moreno-Torres, M. Graaf, S. Arus, S. Huffel, Classification of brain tumours using short echo time ^{1}H NMR spectra, J. Magn. Reson. 170 (2004) 164–175.
[61] M.H. Zweig, G. Campbell, Receiver–operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine, Clin. Chem. 39 (8) (1993) 561–577.

[62] A.H. Fielding, Cluster and Classification Techniques for the Biosciences, Cambrigde University Press, Cambridge, 2007.
[63] C.-F. Lin, S.-D. Wang, Fuzzy support vector machines, IEEE Trans. Neural Netw. 13 (2002) 464–471.
[64] C.-F. Lin, S.-D. Wang, Fuzzy support vector machines with automatic membership setting, Support Vector Mach. Theory Appl. 177 (2005) 233–254.
[65] W. Tai-Yue, C. Huei-Min, Fuzzy support vector machine for multi-class text categorization, Inf. Process. Manag. 43 (2007) 914–929.
[66] E. Karageorgiou, The logic of exploratory and confirmatory data analysis, Cent. Cognit. Sci. (2011) 35–48.
[67] L.J. Pino, D.W. Stashuk, S.G. Boe, T.J. Doherty, Motor unit potential characterization using pattern discovery, Med. Eng. Phys. 30 (5) (2008) 563–573.
[68] Z. Yan, Z. Wang, H. Xie, The application of mutual information-based feature selection and fuzzy LS-SVM-based classifier in motion classification, Comput. Methods Programs Biomed. 90 (3) (2008) 275–284.
[69] E. Gokgoz, Evaluating Effects of Denoising and Feature Extraction Methods on Classification on EMG Signals, International Burch University, Sarajevo, 2014.
[70] H. Parsaei, D.W. Stashuk, Adaptive motor unit potential train validation using MUP shape information, Med. Eng. Phys. 33 (2011) 581–589.
[71] E.B. Filho, E.A. da Silva, M.B. de Carvalho, On EMG signal compression with recurrent patterns, IEEE Trans. Biomed. Eng. 55 (7) (2008) 1920–1923.
[72] H. Parsaei, D.W. Stashuk, Evaluating a motor unit potential train using cluster validation methods, in: Canadian Student Conference on Biomedical Computing and Engineering, 2010, pp. 31–35.
[73] L. Yejin, D. Chenyun, E.A. Clancy, Cross-comparison between two multi-channel EMG decomposition algorithms assessed with experimental and simulated data, Northeast Bioeng. Conf. 39 (2013) 191–192.
[74] H. Marateb, S. Muceli, K. McGill, R. Merletti, D. Farina, Robust decomposition of single-channel intramuscular EMG signals at low force levels, J. Neural Eng. 8 (6) (2011) 192.
[75] M. Nikolic, C. Krarup, EMGTools, an adaptive and versatile tool, IEEE Trans. Biomed. Eng. 58 (10) (2011) 2707–2718.