

# Techniques of EMG signal analysis and classification of Neuromuscular diseases

V. Kehri<sup>1</sup>, R. Ingle<sup>1</sup>, R. Awale<sup>1</sup>, S. Oimbe<sup>2</sup>

<sup>1</sup>Department of Electrical Engineering, VJTI, Mumbai

<sup>2</sup>Cummins College Of Engineering for women, Pune

{vakehri@vjti.org.in; ringle@vjti.org.in ; mawale@vjti.org.in; aboimbe@gmail.com }

**Abstract.** Artificial intelligence techniques are being used effectively in medical diagnostic tools to increase the diagnostic accuracy and provide additional knowledge. Electromyography (EMG) signals are becoming increasingly important in clinical and biomedical applications. Detection, processing and classification of EMG signals are very desirable because it allows a more standardized evaluation to discriminate between different neuromuscular diseases. This paper reviews a brief explanation of the different features extraction and classification techniques for classifying EMG signals used in literatures. Wavelet Transform (WT), Principle Component Analysis (PCA), and Independent Component Analysis (ICA) are different feature extraction techniques. Literature presents different techniques to classify EMG data such as probabilistic neural network (PNN), Support Vector Machine (SVM), Artificial Neural Networks (ANN), etc. In this paper neuromuscular disease classification from electromyography (EMG) signals are proposed based on different combination of features extraction methods and types of classifiers. Combination of WT and SVM improved the classification accuracy than other combinations such as DWT with ANN, ICA with MLPN, PCA with ANN and DWT with PNN.

**Keywords:** *Electromyogram (EMG), Wavelet Transform (WT), Principle Component Analysis (PCA), Support Vector Machine (SVM), Multilayer Perceptron Neural Network (MLPNN), Probabilistic Neural Network (PNN).*

## 1 Introduction

The EMG signals are highly complex and non-linear signal. These signals are widely used in clinical trials for the diagnosis of neurological and neuromuscular problems [1]. Because of the complexity of EMG signals many times even experienced researchers are fail to provide enough information about these signals. EMG signals involve a great deal of information about the nervous system with anatomical and psychological properties of muscles. It is a record of electrical potentials generated by muscles cell [4]. The changes in the voltage difference between electrodes are sensed and amplified before it is transmitted to a computer program to display the tracing of the voltage potential recordings [2].

There are numerous neuromuscular disorders that influence the spinal cord, nerves or muscles. Early finding and diagnosis of these diseases by clinical examination is crucial for their management as well as their anticipation through prenatal diagnosis and genetic counseling. This information's are also valuable in research, which may lead to the understanding of the nature and eventual treatment of these diseases [4].

In the previous literatures Fast Fourier Transform (FFT) was used for analysis of EMG signals, but FFT suffers from large noise sensitivity [5]. Parametric power spectrum methods such as autoregressive (AR), reduces the spectral loss problems and gives better frequency resolution [5]. Since, EMG signals are non-stationary; the parametric methods are not suitable for frequency decomposition of these signals. Another method is Short Time Fourier Transform (STFT) which provides resolution in short window of time for all frequencies. FFT, AR, STFT do not have time and frequency resolution at same time [7]. To extract the features of EMG signals and to overcome the problems of STFT, AR, FFT, a power-full tool that is Wavelet Transform can be applied to extract the wavelet coefficients of discrete time signals. This procedure makes use of multi-rate signal-processing techniques [4,13]. Artificial Neural Networks (ANN) has been used in great number of medical diagnostic decision support system applications because of the belief that these have great predictive power.

Many authors have shown that combining the prediction of several models often results in a prediction accuracy that is higher than that of individual models [4, 10]. From the large number of neuromuscular diseases, we have selected neuropathy and myopathy for our study, as their consistency of clinical appearance is more than other EMG diseases.

The objective of this study is to discriminate between normal (NOR), myopathy (MYO) and neuropathic (NEURO) EMG signals. In this paper, we discuss various methods proposed for classification of EMG datasets. Rest of the paper is organized as follows: Section II describes EMG Data Acquisition. Section III includes, various feature Extraction Techniques proposed for EMG dataset. Section IV is about Classification Methods. Section V is Comparison of results; last section concludes this work.

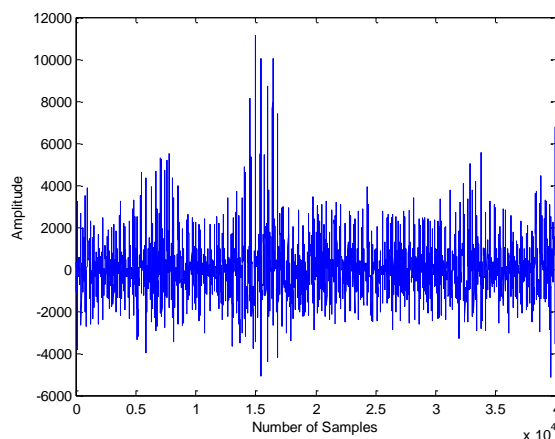
## 2 EMG Diseases

An EMG is a clinical test used to find function of muscles and the nerves that control them. EMG signals studies are used to help in the diagnosis of disorders such as the muscular dystrophies and neuropathies. Nerve conduction studies that measure how well and how fast the nerves conduct. Neuromuscular diseases are a group of disorders which contain motor nuclei of the cranial nerves, anterior cells of the spinal cord, nerve roots and spinal nerves cause muscular weakness [14].

Neuropathies describe damage to the peripheral nervous system which transmits information from the brain and spinal cord to every other part of the body. Patient may experience temporary numbness, tingling, pricking sensations and sensitivity to touch or muscle weakness. Some others symptoms are burning pain, muscle wasting, paralysis or organ or gland dysfunction [9]. Myopathy is a muscle disease where muscle fibers are not working properly. Patient may experience muscle stiffness, cramp and spasm [9]. Some others symptoms are difficulty in speaking and breathing. Early diagnosis may cure these diseases. Therefore, EMG signal analysis is carried out to diagnosis these types of diseases in early stage.

## 3 EMG Data Acquisition

For this work EMG signals are taken from publicly available clinical EMG database consisting of three different classes such as normal, neuropathy, and myopathy subjects. EMG signals were recorded from brachial biceps muscles under the usual conditions using standard needle electrode. The recordings were made at low voluntary and constant level of contraction. The high pass and low pass filters of the EMG amplifier were set at 1Hz and 12 kHz. The sampling rate of the EMG signal is 10 kHz [10].



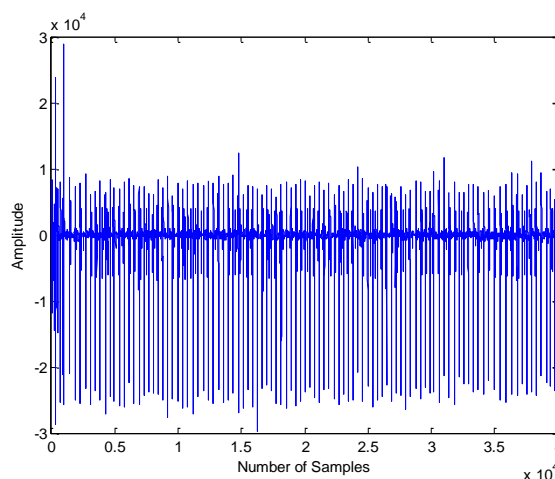
**Fig.1** EMG signal of normal subject.

## 4 Feature Extraction Techniques

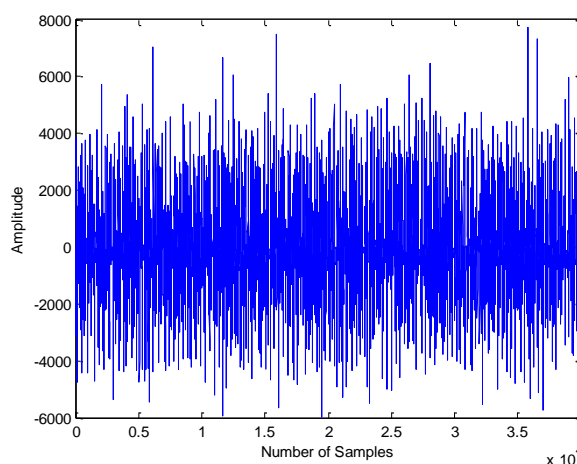
There are several features extraction techniques in the literature were used for analysis of EMG signals. Some of them which have been used in this work as follows:

### 4.1 Wavelet Transform:

The Wavelet Transform (WT) is designed to address the problem of non-stationary signals. It involves representing a time function in terms of simple, fixed building blocks, termed Wavelets [4].



**Fig.2** EMG signal of patient with neuropathic diseases



**Fig.3** EMG signal of patient with myopathy diseases.

These building blocks are actually a family of functions which are derived from a single generating function called the mother wavelet by translation and dilation operations. The main advantage of WT is that it has a varying window size, being broad at low frequencies and narrow at high frequencies, thus leading to an optimal time frequency resolution in all frequency ranges [6]. One-dimensional WT is defined as,

$$W_f = \frac{1}{\sqrt{d}} \int_{-\infty}^{\infty} f(t) \Psi^* \left( \frac{t-m}{d} \right) dt \quad (1)$$

Where  $\psi^*(t)$  is the conjugate function of mother wavelet  $\psi(t)$  and  $d, m$  is called scale parameter and shift parameter respectively. Major advantage of WT for EMG signals is, it provides the better time frequency resolution at same time and compress data by preserving original information.

#### 4.2 Principle Component Analysis:

Principal component analysis (PCA) has been called one of the most valuable results from applied linear algebra. PCA is used abundantly in all forms of analysis from neuroscience to computer graphics - because it is a simple, non-parametric method of extracting relevant information from confusing data sets. PCA used for finding similarity in data pattern. For any data set  $Z = [z_1, z_2, z_3 \dots z_n]$  ( $n$ - dimension of data-point). Suppose  $Z$  includes  $m$  data-points, the new set of data-points is given by [11,12],

$$Y_i = z_i - \mu \quad i = 1, 2, 3, \dots, m \quad (2)$$

$\mu$  is a mean vector. We have to find the covariance matrix of  $Y_i$  can be expressed as,

$$C_{n \times n} = \text{cov}(z_i, z_j); i = 1, 2, 3 \dots n; j = 1, 2, 3 \dots n. \quad (3)$$

$$C = \sum_{j=1}^m (z_i - \mu)(z_i - \mu)^T \quad (4)$$

$m$  is total number of data-points. The Eigen value  $\lambda_i$  and Eigen Vectors  $\psi_i$  of covariance matrix  $C$  satisfy,

$$C\psi_i = \lambda_i\psi_i \quad (5)$$

$$E = [\psi_1, \psi_2, \psi_3, \dots, \psi_m] \quad (6)$$

Project the data set  $Y$  into the Eigen vector space, we get:

$$P^{m \times n} = E^T Y^T \quad (7)$$

Where,  $Y = [Y_1, Y_2, \dots, Y_n]$ ; We can easily mapped  $P$  back into the original coordinates, where

$$Y^T = EP \quad (8)$$

$E$  is an orthogonal matrix, and  $E^{-1} = E^T$

$$Y'^T = E'P' \quad (9)$$

If  $\lambda_i$  is arranged from the largest to the smallest and  $\lambda_i$  is arranged according to the value  $\lambda_i$ , we select the first  $k$  row of matrix  $P$  to form a matrix  $P'$ , which represents the principle component of the data set  $X$ .  $P'$  can construct  $m$  data points with  $k$  dimension. We can have mapped the principle component back into the original coordinates by,

### 4.3 Independent Component Analysis:

ICA is a signal processing technique in which observed random data is transformed into components that are statistically independent from each other. Linear ICA was used to separate neural activity between different types of EMG data. ICA would appear to be able to separate task-related potentials from other neural and artifactual [11]. EMG sources during different hand movement in form of independent components. Power spectra of the linear ICA transformations provided feature subsets with higher classification accuracy than the power spectra of the original EMG signals.

## 5 Classification Methods

Artificial neural network techniques are the recent trends for classification of 1-D and 2-D signals. In literature, classification of EMG diseases done by using various techniques such as feed forward network at different layers (FFNN), back propagation neural network (BPNN), principle component analysis (PCA), Support vector machine (SVM), Logistic Regression (LR), multilayer perceptron neural network (MLPNN), probabilistic neural network (PNN), etc.[8,12].

### 5.1 Support Vector Machine (SVM):

SVM is a supervised learning machine based on Statistical theory. It trains a classifier by finding an optimal separating hyperplane which maximizes the margin between two classes of data in the kernel induced feature space [8]. A training sample set  $\{x_i, y_i\}; i=1-N$  is considered, where  $N$  is total number of samples. The hyperplane  $f(x) = 0$  that separates the given data can be obtained as a solution to the following optimization problem

Minimize:

$$\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i \quad (10)$$

Subject to,

$$\begin{cases} y_i(\omega^T x_i + b) \geq 1 - \xi_i \\ \xi \geq 0, \quad i = 1, 2, 3, \dots, n \end{cases} \quad (11)$$

Where constant C representing error penalty. Rewriting the above optimization problem in terms of Lagrange multiplier ( $\lambda$ ), leads to the following problem [7],

Maximize,

$$W(\lambda) = \sum_{i=1}^N \lambda_i - \frac{1}{2} \sum_{i,j=1}^N y_i y_j x_i x_j \lambda_i \lambda_j \quad (12)$$

Subject to,

$$0 \leq \lambda_i \leq C \quad (13)$$

$$\sum_{i=1}^N \lambda_i y_i = 0, \quad i = 1, 2, 3, \dots, N \quad (14)$$

Holding hyper plane vectors are termed as support vectors. In literature SVM has been used to classify multiple datasets of EMG signals.

## 5.2 Probabilistic Neural Network (PNN):

A single PNN is capable of handling multiclass problems. This is opposite to the so-called one-against-the rest or one-per-class approach taken by some classifiers, such as the SVM, which decompose a multiclass classification problem into dichotomies and each classifier, has to separate a single class from all others [10]. The PNN architecture is composed of many interconnected processing units or neurons organized in successive layers. The input layer unit does not perform any computation and simply distributes the input to the neurons in the pattern layer. On receiving a pattern  $x$  from the input layer, the neuron  $x_{ij}$  of the pattern layer computes its output is given by [11],

$$\phi_{ij} = \frac{1}{(2\pi)^{\frac{d}{2}} \sigma^d} \exp \left[ \frac{-(x - x_{ij})^T (x - x_{ij})}{2\sigma^2} \right] \quad (15)$$

Where  $d$  denotes the dimension of the pattern vector  $x$ , is the smoothing parameter, and  $x_{ij}$  is the neuron vector.

## 6 EMG Diseases Classification And Comparative Discussion

The brachial biceps muscles were used in this work because they were the frequently investigated in the different patient groups. For the purpose of detecting the EMG diseases from the given EMG data, 10 EMG datasets of 5 normal subjects and 5 abnormal subjects are used. Each EMG dataset has total number of 242,124 samples at a rate of 22,438 samples per second. Thus, each of these single channel datasets has total time duration of

10.184 sec. In order to consider the middle steady regions, from each dataset 20 frames (from 25<sup>th</sup> frame to 45<sup>th</sup> frame) are selected out of 64 frames of both the normal persons and abnormal patients for feature extraction. In this paper two types of experiments have done. In the first experiment, EMG raw data of each subject containing 4096 samples used directly for feature extraction using discrete wavelet transform. From each frame, arranging the DWT coefficients in a descending order, the first six higher valued coefficients are used for features extraction. The results presented are computed using the 'db2' type mother wavelet. In the second experiment a rectangular window which is formed by 256 discrete data was selected. After down-sampling, for the 100 series set, a total of 1600 vectors are obtained from each set for features extraction. Figure 4 represent the block diagram of proposed work for EMG disease classification.

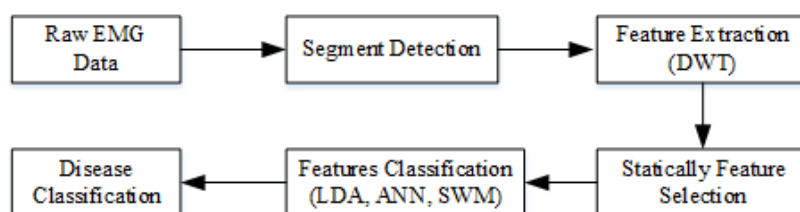


Fig.4. Functional block diagram of EMG analysis for diseases classification

The classification performance was evaluated for EMG datasets (normal, myopathy and neuropathy subject). In the experiment different features are extracted for EMG datasets and classified using (combination of features extraction techniques and types of classifiers) DWT+ANN, PCA+ANN, DWT+PNN and DWT+SVM classification technique. The total-classification accuracies obtained in the first experiment (training and testing sets consisted of raw data for all three subjects) were 90.4%, 92.4%, 93.05% and 94.28%, respectively, as shown in table I.

Table 1. Classification Accuracy (%) for different classification techniques

CLASSIFICATION TECHNIQUES	CLASSIFICATION ACCURACY (%)
DWT + ANN	90.40
PCA + ANN	92.40
DWT + PNN	93.05
DWT + SVM	94.28

## 7 Conclusion And Future Work

Muscle is a vital organ of the body responsible for movements. Study on EMG is broad starting from the design of electrodes to recording methods, analysis techniques and application for various used. Many papers have been used in an attempt to study different muscle disorders using time and frequency domain analysis. Classification of EMG diseases is a challenging task for researchers. This paper described few feature extraction and classification techniques which have been used for classification of EMG datasets. DWT reduced data and provided better features as compared with PCA and ICA methods. Experimental results show that combination of DWT + SVM classifier can be effectively trained for classification of EMG datasets. We have managed to obtain an accuracy of 94.28%. Performance of classifier further improved by using combination of different machine learning and artificial neural techniques with number of hidden layers.

## References

- [1] Ricamato, A.L.; Absher, R.G.; Moffroid, M.T.; Tranowski, J.P. A Time-Frequency Approach to Evaluate Electromyography Recordings. In Proceedings of the Fifth Annual IEEE Symposium on Computer-Based Medical Systems, Durham, NC, USA, 14–17 June 1992; pp. 520–527.
- [2] Englehart, K.; Hudgins, B.; Parker, P.A.; Stevenson, M. Classification of the myoelectric signal using time-frequency based representations. *Med. Eng. Phys.* 1999, *21*, 431–438.
- [3] Ismail Babaoglu, Oğuz Findik, Mehmet Bayrak, "Effects of principle component analysis on assessment of coronary artery diseases using support vector machine," *Expert Systems with Applications*, vol.37, pp.2182–2185, 2010.
- [4] A. B. M. S. U. Doulah, Student Member, IEEE, S. A. Fattah, Member, IEEE, W.-P. Zhu, Senior Member, IEEE, and M. O. Ahmad, Fellow, IEEE, "Wavelet Domain Feature Extraction Scheme Based on Dominant Motor Unit Action Potential of EMG Signal for Neuromuscular Disease Classification", *IEEE Trans. Biomed. Eng.*, vol. 8, no. 2, APRIL 2014.
- [5] F. G. Nihal and K. Sabri, "Classification of EMG signals using PCA and FFT," *J Med. Syst.*, vol. 29, no. 3, pp. 15–25, 2005.
- [6] EEG signal Classification using Principle Component Analysis and Wavelet Transform with Neural Network" International conference on communication and signal processing, IEEE conference, vol.1, pp.456-459, April 3-5, 2014.
- [7] Alkan, A.; Günay, M. Identification of EMG signals using discriminant analysis and SVM classifier. *Expert Syst. Appl.* 2012, *39*, 44–47.
- [8] N S. A. Fattah, A. B. M. S. U. Doulah, M. A. Iqbal, C. Shahnaz, W.-P. Zhu, and M. O. Ahmad, "Identification of motor neuron disease using wavelet domain features extracted from EMG signal," in Proc. IEEE Int. Conf. Circuits and Systems, 2013, pp. 1308–1311.
- [9] Prof. Dr. Onsy Abdul Alim, Dr. Mohamed Moselhy, and Eng. Fatima Mroueh 'EMG Signal Processing and Diagnostic of Muscle Diseases" 2012 2nd international conference ACTEA.
- [10] D. F. Specht, "Probabilistic neural networks," *Neural Netw.*, vol. 3, no. 1, pp. 109–118, 1990.
- [11] Kavita Mahajan, M. R. Vargantwar, Sangita M. Rajput, "Classification of EEG using PCA, ICA and Neural Network", *International Journal of Engineering and Advanced Technology (IJEAT)* ISSN: 2249 - 8958, vol.1, Oct .2011.
- [12] S.S.Lekshmi; V.Selvam; M.Pallikonda Rajasekaran, "EEG signal Classification using Principle Component Analysis and Wavelet Transform with Neural Network "International conference on communication and signal processing, IEEE conference, vol.1, pp.456-459, April 3-5, 2014.
- [13] Englehart, K.; Hudgin, B.; Parker, P. A wavelet based continuous classification scheme for multi-function myoelectric control. *IEEE Trans. Biomed. Eng.* 2001, *48*, 302–311.
- [14] Shravanti Kalwa, Prof.H.T.Patil, Neuromuscular Disease Classification based on Discrete Wavelet Transform of Dominant Motor Unit Action Potential of EMG Signal.