A Multi-Classifier Approach of EMG Signal Classification for Diagnosis of Neuromuscular Disorders

Muzaffar Khan
S.S.I. Research & Tech Sehore Bhopal (M.P) India

Jaikaran Singh
S.S.I. Research & Tech Sehore Bhopal (M.P) India

Mukesh Tiwari
S.S.I. Research & Tech Sehore Bhopal (M.P) India

ABSTRACT
Electromyographic (EMG) signal provides a significant source of information for diagnosis, treatment and management of neuromuscular disorders. This paper aims at introducing an effective multi-classifier approach to enhance classification accuracy. The proposed system employs both time domain and time-frequency domain features of motor unit action potentials (MUAPs) extracted from an EMG signal. Different classification strategies including single classifier and multiple classifiers with time domain and frequency domain features were investigated. Support Vector Machine (SVM) and K-nearest neighborhood (KNN) classifier used predict class label (Myopathic, Neuropathic, or Normal) for a given MUAP. Extensive analysis was performed on clinical EMG database for the classification of neuromuscular diseases and it is found that the proposed methods provide a very satisfactory performance in terms of overall classification accuracy.

Keywords
Support Vector Machine, EMG, Discrete wavelet Transform; K-nearest neighborhood (KNN)

1. INTRODUCTION
Electromyographic (EMG) signal analysis plays a major role in the diagnosis of neuromuscular diseases, such as amyotrophic lateral sclerosis (ALS) and myopathy. Neuromuscular diseases changes, the shape and characteristics of the motor unit action potentials (MUAPs) and firing patterns of the motor unit (MU) are affected. MUAPs detected from myopathic patients are characterized by high frequency contents, low peak-to-peak amplitude and MUAPs neuropathic patients are poly-phasic, low frequency, high peak-to-peak amplitude than the normal MUAPs. [1], [2]. The amplitude and time and frequency domain properties of the surface EMG signal are dependent on the timing and intensity of muscle contraction. When a patient maintains low level of muscle contraction, individual MUAPs can be easily recognized. As contraction intensity increases, more motor units are recruited. Different MUAPs will overlap, causing an interference pattern in which the neurophysiologist cannot detect individual MUAP shapes reliably [3]. The methods reported in [1], [11] used wavelet-domain features extracted through multi-level decomposition using a filter bank structure consisting of only the analysis bank with Daubechies 4 wavelet filters, and several time domain features are used, such as zero-crossing rate, turns-amplitude ratio, root-mean-square (RMS) value and autoregressive (AR) coefficients [13], [14]. Several classification methods such as fusion classifier, multi-classifier, an SVM that provides such probabilities for each class is reported in [1], [16]. Existing EMG signal decomposition methods can successfully decompose EMG signals extracting MUAPs by dominant MUAP selection method or thresholding active and non-active region [24], [26]. The motor unit potential trains (MUP) is assumed to have MUP shape validity, if motor unit MU discharges corresponding to a valid MUP occur at regular intervals and in general, have a Gaussian-shaped inter-discharge interval (IDI) histogram [16],[27],[28]. Empirical mode decomposition (EMD) is a kind of self-adapting signal processing method and it is very suitable for dealing with nonlinear and non-stationary signals, a new method is proposed which combines independent component analysis, empirical mode decomposition and AR model to extract and analyze the surface EMG. EMD plus AR model shows superior accuracy of 93% compared to AR model having accuracy 85% [3]. Sarbast Rasheed et al. [10], present design methodology for integrating heterogeneous classifiers ensembles by employing a diversity-based hybrid classifier fusion approach, whose aggregator module consists of two classifier combiners, to achieve an improved classification performance for motor unit potential classification during electromyographic (EMG) signal decomposition. The pool of base classifiers consists of different kinds of classifiers the adaptive certainty-based, the adaptive fuzzy k-NN, and the adaptive matched template filter classifiers. Miki Nikoli et al. [33], developed the system which decomposes signal decomposition algorithm consists of three stages, segmentation, clustering, and resolution of compound segments. Clustering of Segments is determined if two segments are similar in shape, a so called distance measure is used.

In this paper, DWT based feature extraction schemes are proposed for the classification of normal, ALS and myopathy subjects. First an MUAP based scheme is proposed where the MUAPs are first extracted from the EMG data by using a decomposition technique. A dominant MUAP selection criterion is introduced to extract features only from selected MUAPs. Statistical features are obtained from the DWT of dominant MUAPs. Next design of multi-classifier majority voting using SVM as base classifier and K-nearest neighborhood (KNN) classifier is employed. Finally, experimental results with comparative analysis are presented.

2. MUAP EXTRACTION BY USING EMG DECOMPOSITION
The first step is the filtering part, in which the EMG signal is band pass filtered (10 Hz to 3 kHz). Now EMG signal contain so-called inactive segments with low activity and active segments containing MUAPs. Window function is used to extract MUAPs around the peak this low activity segments can affect time domain feature. So it removed in beginning before applying window function. To remove inactive segment threshold parameter (±λ) is set around baseline if the signal sample lays between ±λ and ~λ for more than 0.4 ms is discarded. Segmentation of EMG signals carried by finding the peaks of the MUAPs, then a window of 180 sampling points is centered at the identified peak, size of window depends on sampling rate [23]. The selection criteria
for the MUAP extracted from EMG signal is based on dominant MUAP based on temporal energy. In case of myopathy MUAPs become low in amplitude and short in duration, while for the neurogenic disorders, MUAPs exhibit higher amplitude and longer duration than normal. Hence, the energy content of MUAPs provides significant information about the EMG signal and idea about pathology. ALS group is the highest followed by the normal group and the myopathic dominant MUAP has the lowest energy. Once the dominant MUAPs for different datasets are obtained, these are then used for the feature extraction [19].

3. TIME AND TIME-FREQUENCY FEATURE EXTRACTION AND SELECTION NORMAL

3.1 Time Domain Features Extraction

Time domain features are morphological features of the MUAPs which are used for visual assessment. MUAPs myopathic patients are characterized by high frequency contents, low peak-to-peak amplitude. Neuropathic patients are polyphasic, low frequency, high peak-to-peak amplitude than the normal MUAPs. The following morphological features were employed to represent each MUAP [1], [20], [21].

1. Rise Time: The time between the initial positive to the next negative peak within the main spike.
2. Ratio of Peak to Peak magnitude to RMS value
3. Spike Duration: The time between the first to the last positive peak.
4. Ratio of ascending slope to descending slope positive spike of MUAP.
5. Ratio Area of positive to Area of negative spike MUAP
6. Phases: The number of baseline crossings where amplitude exceeds ±25 µV, plus one.
7. Thickness: The ratio of the area to the peak-to-peak amplitude.
8. Peak-to-Peak Samples Number: Total number of samples between the minimum positive and the maximum negative peak.

3.2 DWT Based Feature Extraction Scheme

The DWT is a multi-resolution technique that offers localization both in time and frequency. Hence, the DWT is chosen to extract features from the EMG signal

The DWT of a signal S(n) can be represented as

\[ W_{d}(j, K) = \frac{1}{\sqrt{M}} \sum_{n} s(n) \psi_{j,k}(n) \]  
\[ W_{o}(j, K) = \frac{1}{\sqrt{N}} \sum_{n} s(n) \phi_{j,k}(n) \]

where, \( j \geq j_{0i} \) and \( s(n) \), \( \psi_{j,k}(n) \) and \( \phi_{j,k}(n) \) are functions of discrete variables \( n = 0,1, \ldots, M-1 \) select M to be a power of 2 \((M = 2^k)\). \( K \) and \( j \) are selected by the optimal procedure. The approximation coefficients and equation (2) computes the detail coefficients

The original signal being filtered via high pass \( W_{o}(j, K) \) and a low-pass \( W_{d}(j, K) \) filter produces output expressed as

\[ W_{o}(j, K) = \frac{1}{\sqrt{M}} \sum_{n} s(n) (m - 2k)\sqrt{2} \phi_{2^j+1, n-m} \]  
\[ W_{d}(j, K) = \frac{1}{\sqrt{N}} \sum_{n} s(n) \psi_{j,k}(n) \sqrt{2} \phi_{2^j+1, n-m} \]  
\[ W_{o}(j, K) = \sum_{n} h_{j}\psi_{j}(m + 2k) W_{o}(j+1, k) \]

DWT coefficients at adjacent scales. Both \( W_{o}(j, K) \) and \( W_{d}(j, K) \) are obtained by convolving the scale \((j+1)\) approximation coefficient \( W_{o}(j + 1, k) \) with \( h_{j}(n) \) and \( h_{j}(n) \) respectively and then subsampling the convolved output by a factor of 2.

3.3 Mother Wavelet Selection

In this work, the best mother wavelet was determined experimentally using cross validation technique. The choice of mother wavelet can be based on it can be selected based on correlation \( \gamma \) between the signal of interest and the wavelet denoised signal. It determines estimation of the original signal, but also affect the frequency spectrum of the de-noised signal

\[ \gamma = \sum \frac{(x - \bar{x})(y - \bar{y})}{(x - \bar{x})^2} \]  

Where \( \bar{x} \) and \( \bar{y} \) are mean value of set X and Y, respectively. The family of five mother wavelets consisting of Symlet, Daubechies, Morlet, Coiflet and Haar were studied. Symlet4 and Daubechies4 provided the most discriminative frequency band for three groups (myopathic, neuropathic, and normal). The DWT is a multi-resolution technique that offers localization both in time and frequency [17].

3.4 DWT Features Reduction.

Once the best discriminative frequency band was determined, the following statistics were estimated and used to represent the time-frequency distribution of the isolated MUAPs and reduce the dimension of DWT features [19].

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. CLASSIFICATION STRATEGIES

In this paper, Multi-Classifiers Majority Voting (MCMV) classification strategies were evaluated. Multi classifier as shown in figure 1, consist three group in parallel. Each group consist of four SVM classifier as base classifier, two schemes is employed for class discrimination one against one (OAO) and one against all (OAA) given in table 1, [29], [30]. The selected SVM has Gaussian radial basis function (RBF) kernel, which is stated as follows

\[ K(x, x') = e^{-\gamma / \|x-x'\|^2} \]  

Where x the input feature vector to the SVM, \( x' \) is the center of the support vector and \( \gamma \) is the width of the kernel [1]. The multi-classifier scheme base classifier C1 to C12 are grouped into three groups, the first group consist base classifier from
C1 to C4 (myopathic class label), second group consist base classifier from C5 to C68 (Neuropathic group) and third group consist of base classifier from C9 to C12 (Normal class label). SVM was first trained as a standard SVM and then a sigmoid function was trained which maps the SVM outputs to the posterior probabilities. The conditional probabilities of the two classes for given input vector x is given by

\[ P1(x) = \frac{1}{1 + \exp(-\beta_1 f(x) + \beta_2)} \] ... 9

\[ P2(x) = 1 - P1(x) \] ... 10

\( f(x) \) is the output standard SVM, where \( \beta_1 \) and \( \beta_2 \) are parameter of sigmoid function, these parameters are derived from maximum likelihood estimation during training phase.

### Table 1. List of Base Classifiers with Time Frequency Feature Employed

<table>
<thead>
<tr>
<th>Base classifier</th>
<th>Priority Ranking</th>
<th>Class Discrimination</th>
<th>Group/Class label</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>P1</td>
<td>Myopathic Vs. Others (Normal &amp; Neuropathic)</td>
<td>Group 1 Myopathic</td>
</tr>
<tr>
<td>C2</td>
<td>P2</td>
<td>Myopathic Vs. Myopathic</td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>P3</td>
<td>Myopathic Vs. Neuropathic (inverted votes)</td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>P3</td>
<td>Myopathic Vs. Normal (inverted votes)</td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>P1</td>
<td>Neuropathic Vs. Others (Normal &amp; Myopathic)</td>
<td>Group 2 Neuropathic</td>
</tr>
<tr>
<td>C6</td>
<td>P2</td>
<td>Neuropathic Vs. Neuropathic</td>
<td></td>
</tr>
<tr>
<td>C7</td>
<td>P3</td>
<td>Neuropathic Vs. Myopathic (inverted votes)</td>
<td></td>
</tr>
<tr>
<td>C8</td>
<td>P3</td>
<td>Neuropathic Vs. Normal (inverted votes)</td>
<td></td>
</tr>
<tr>
<td>C9</td>
<td>P1</td>
<td>Normal Vs. Others (Myopathic &amp; Neuropathic)</td>
<td>Group 3 Normal</td>
</tr>
<tr>
<td>C10</td>
<td>P2</td>
<td>Normal Vs Normal</td>
<td></td>
</tr>
<tr>
<td>C11</td>
<td>P3</td>
<td>Normal Vs. Neuropathic (inverted votes)</td>
<td></td>
</tr>
<tr>
<td>C12</td>
<td>P3</td>
<td>Normal Vs. Myopathic (inverted votes)</td>
<td></td>
</tr>
</tbody>
</table>

### 4.1 Majority Voting

The group with more votes is selected as the ultimate decision. The votes of base classifier trying classify other than its group label are inverted for majority voting method to be used. However, in the case of equal number of votes between two groups, then decision is based top two priority classifier within the group. Classifier with priority P1 is highest and P3 is lowest.

### 4.2 Distance weighted k-Nearest Neighborhood (D WKNN)

DWKNN is treated as a benchmark classifier employed for the purpose of comparison. K-nearest neighbor has an identical weight in decision making and neglects that closer neighbor contribute more to classification. Dudani proposed the weight the distance weighted k-Nearest Neighbor (KNN) rule (WKNN) in which votes of the different members of the one of the K neighbors set are computed by the function of their distance to the query [31]. In this scheme, the i-th weight of the corresponding nearest neighbor is given as

\[ W_i = \begin{cases} \frac{(d_{iNN} - d_{1NN})}{(d_{KNN} - d_{1NN})} \times \left( \frac{d_{KNN} + d_{1NN}}{2} \right) & \text{if } d_{KNN} \neq d_{1NN} \\ 1 & \text{if } d_{KNN} = d_{1NN} \end{cases} \] ... 11

Where \( d_{iNN} \) is the distance to the query of the i-th nearest neighbor \( d_{1NN} \) is the distance the nearest neighbor and \( d_{KNN} \) is the distance of the K-furthest neighbor. Then, the query is assigned to the majority weighted voting class label \( y_{j\max} \) using the following rule

\[ y_{j\max} = \arg\max_{y_i} \sum_{(x, y) \in T} W_i \times I(y = y_{iNN}^N) \] ... 12

Algorithm for DWKNN can be state as

1. Compute the distances of nearest neighbors of the query \( \bar{x} \).
2. Sort the distances in an ascending order.
3. Calculate the dual weights of k nearest neighbors, \( \bar{W} = \{ \bar{W}_1, ..., \bar{W}_k \} \) from equation 11.
4. Assign a majority weighted voting class label \( y_{j\max} \) to the query \( \bar{x} \).

### 4.3 Evaluation Methodology

The performance of classifiers is evaluated using indices. Classification accuracy indices is defined for this purpose: accuracy for classifier (Ac), indices is given by

\[ Ac = \frac{\text{Number of correctly Classified EMG}}{\text{Total number EMG sample}} \] .... 13

Misclassification of classifier is given by (Mc) is given by

\[ Mc = \frac{\text{Number of Misclassified EMG Class label}}{\text{Total number EMG sample / class Label}} \] .... 14
5. RESULTS AND DISCUSSION

As it seen from graph shown in figure 2 and Table 2 classification accuracy is high within the same class. Whereas classification strategy one against all class label gives second highest accuracy for base classifier in all groups. The proposed multi-classifier model provides average accuracy for base classifier in all groups. The models were tested on data of 150 EMG signal, 50 sample of each class. EMG is carried by remove inactive region around base line and use of window function around peak gives simple approach for MAUPS extraction. The dominant MAUPs selected for Time and Time-frequency domain feature extraction. The base classifier used in the multi-classifier model is trainable, a sigmoid function was trained which maps the SVM outputs to the posterior probabilities. All posterior probabilities are sorted in descending order, then from posterior values related to misclassified vector, threshold value for misclassified bound selected. Time-frequency domain features are selected since time domain feature fail to map spectrum behavior and complexity of EMG signal. The only differences are that on the one hand, DWKNN offers a large variety of possible kernel functions in order to produce different weighting schemes. The main purpose of this extended method is to gain a technique that up to a certain degree is independent of a bad choice for k resulting in a high misclassification error. Now this number of nearest neighbors is implicitly hidden in the weights. DWKNN classifiers aggregated by a (weighted) majority vote and this aggregated result is used as prediction.
7. ACKNOWLEDGMENTS

The authors would like to gratefully thank Dr. G.R. Selokar director S.S.S.I.T for sharing research knowledge, guidance and giving access to laboratory facility for the research work.

8. REFERENCES


[7]. Aaron J. Young, Lauren H and Smith, Elliott . “Classification of Simultaneous Movements Using Surface EMG Pattern Recognition” IEEE Transactions On Biomedical Engineering, Vol. 60, No. 5, May 2013


Table 2: Class discrimination and percentage accuracy of classifier

<table>
<thead>
<tr>
<th>Base classifier</th>
<th>Class Discrimination</th>
<th>Percentage Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Myopathic Vs Others</td>
<td>93.21%</td>
</tr>
<tr>
<td>C2</td>
<td>Myopathic Vs Myopathic</td>
<td>100%</td>
</tr>
<tr>
<td>C3</td>
<td>Myopathic Vs Neuropathic</td>
<td>90%</td>
</tr>
<tr>
<td>C4</td>
<td>Myopathic Vs Normal</td>
<td>77.21%</td>
</tr>
<tr>
<td>C5</td>
<td>Neuropathic Vs Others</td>
<td>96.2%</td>
</tr>
<tr>
<td>C6</td>
<td>Neuropathic Vs Neuropathic</td>
<td>100%</td>
</tr>
<tr>
<td>C7</td>
<td>Neuropathic Vs Myopathic</td>
<td>78.2%</td>
</tr>
<tr>
<td>C8</td>
<td>Neuropathic Vs Normal</td>
<td>75.21%</td>
</tr>
<tr>
<td>C9</td>
<td>Normal Vs Others</td>
<td>90.38%</td>
</tr>
<tr>
<td>C10</td>
<td>Normal Vs Normal</td>
<td>100%</td>
</tr>
<tr>
<td>C11</td>
<td>Normal Vs Neuropathic</td>
<td>73.21%</td>
</tr>
<tr>
<td>C12</td>
<td>Normal Vs Myopathic</td>
<td>79.2%</td>
</tr>
<tr>
<td>Multi-classifier Model</td>
<td>One Vs All</td>
<td>97%</td>
</tr>
<tr>
<td>DWKNN Classifier</td>
<td>One Vs All</td>
<td>95%</td>
</tr>
</tbody>
</table>

Table 3: Classification accuracy of DWKNN classifier

<table>
<thead>
<tr>
<th>S.No</th>
<th>Class Label</th>
<th>Feature type</th>
<th>Percentage accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Neuropathic Vs All</td>
<td>Time domain &amp; Frequency domain</td>
<td>97.00%</td>
</tr>
<tr>
<td>2</td>
<td>Myopathic Vs All</td>
<td>Time domain &amp; Frequency domain</td>
<td>92.00%</td>
</tr>
<tr>
<td>3</td>
<td>Normal Vs All</td>
<td>Time domain &amp; Frequency domain</td>
<td>96.00%</td>
</tr>
<tr>
<td>Average Accuracy</td>
<td></td>
<td></td>
<td>95.00</td>
</tr>
</tbody>
</table>

6. CONCLUSION

This paper focuses on evaluating two classification strategies to classify the MUAPs into the following classes, normal, myopathic and neuropathic. The proposed classification strategies consist of several base classifiers which take different MUAPs features such as time domain features, time-frequency features (wavelet coefficients). These classification strategies can be employed in other pattern recognition applications because they segment a big decision into several detailed decisions where the input of each decision node can be separately optimized. Multi-classifier overcomes limitation of single stage classifier but with a cost of complexity and processing time. Although the result of time-frequency features is superior to the time domain ones, selecting both types of feature result in promising results (97%) for the three classes. These classification strategies can be employed in other pattern recognition applications. Through our experiments, the proposed method always outperforms the DWKNN classifiers among a large range of k and its effectiveness was demonstrated with good performance. For extending this research is to investigate influence of the recording conditions on the classification accuracy.


