ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 5, Issue 6, June 2016

Classification of EMG Signals using SVM-kNN

PREETI MEENA, MALTI BANSAL

Abstract **–Electromyography (EMG) is a way to measure the electrical activity of the contraction of muscles and nerves.There are so many methods already developed, in the time and frequency domains, for EMG signals analysis. In this research, a method is proposed for the EMG signals analysis using their coiflet wavelet transform and with their features such as mean, energy and standard deviation. These features are then utilized on three different types of EMG signals like normal, myopathic and neuropathic signals to classify them. The method proposed can automatically classifies these signals into their respective classes. For this, various classifiers are used in this research but it is observed that SVM-kNN classifies the signals with highest accuracy in comparison to the other classifiers. SVM-kNN has come with 95% (approximately) accuracy which is also very compared to the other methods used earlier.**

Index terms **--- Electromyography, Wavelet transform, SVM, SVM-kNN.**

I. INTRODUCTION

Electromyography (EMG) is an analytic way to access the health of muscles and nerves. Motor neurons present in the body transmit electrical signals that cause muscles to contract. These changes in muscles contraction is recorded in the form of graph using EMG.

EMG signals classification is an important analysis and anorganisedstudy is required for their classification. For the same reason, number of computer-based quantitative EMG analysis algorithms have been developed.Among these algorithms, some of the already used feature extraction techniques includes Discrete wavelet transform (Daubechies-6) [1], AR modelling [2], autoregressive cepstral analysis [3], PSO [4], wavelet packet energy [5] etc. These methods were utilized by other researchers for the features extraction in the EMG signals that were used during their research. Then, these features are exploit in the classifiers to classify the signals used. Classifiers that are already exploitedincludes Fuzzy [5], Artificial neural networks (ANN) [6], Fuzzy-genetic [7], Neuro-fuzzy [8], Deep fuzzy neural network [1], SVM [9]etc. All these models give satisfactory and even good results. But, for better result performance in terms of accuracy,SVM-kNN is utilized here because SVM-kNN can also provide efficient results in the field of classification.

In this paper, a method is proposed for EMG signals classification. A brief overview of the proposed method is shown in the block diagram in Fig.1. As shown in the figure, three different types of EMG signals (Normal, Myopathy and Neuropathy) are taken. Detail coefficients of coiflet of order 5 of the EMG signals are computed. Then, for D4 coefficient, features like mean, energy and standard deviation are calculated. At last, using these calculated features, classifiers like SVM and SVM-kNN are exploited to classify the EMG signals and the best classifier from them is find out.

The organisation of different sections of the paper is as follows: Section 2 provides a brief overview of the proposed method in the form of a block diagram. Section 3 provides details about the dataset used, definitions and explanations of the methods used. Section 4 provides the process, explained using block diagram, of the proposed method. Section 5 contains the classification results followed by conclusion in Section 6.

II. BLOCK DIAGRAM

A brief outline of the work is shown in the block diagram in Figure. As shown in the figure, three different types of EMG signals (Normal, Myopathy and Neuropathy) are taken. Detail coefficients of coiflet of order 5 of the EMG signals are computed.

Preeti Meena*, Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India.*

Malti Bansal*, Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India.*

ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 5, Issue 6, June 2016

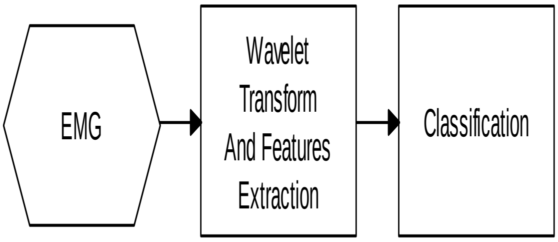

Fig.1.EMG Signals Classification

Then, for D4 coefficient, features like mean, energy and standard deviation are calculated. At last, using these calculated features, classifiers like SVM and SVM-kNN are exploited to classify the EMG signals and the best classifier from them is find out.

III. PROPOSED METHOD

A. Data Set Used

Data set of EMG signals is loaded from the MIT-BIH database.

From the data set, waveforms of three types of EMG signals are taken and shown in the Fig. 2. These waveforms represent the normal, myopathy and neuropathy signals.

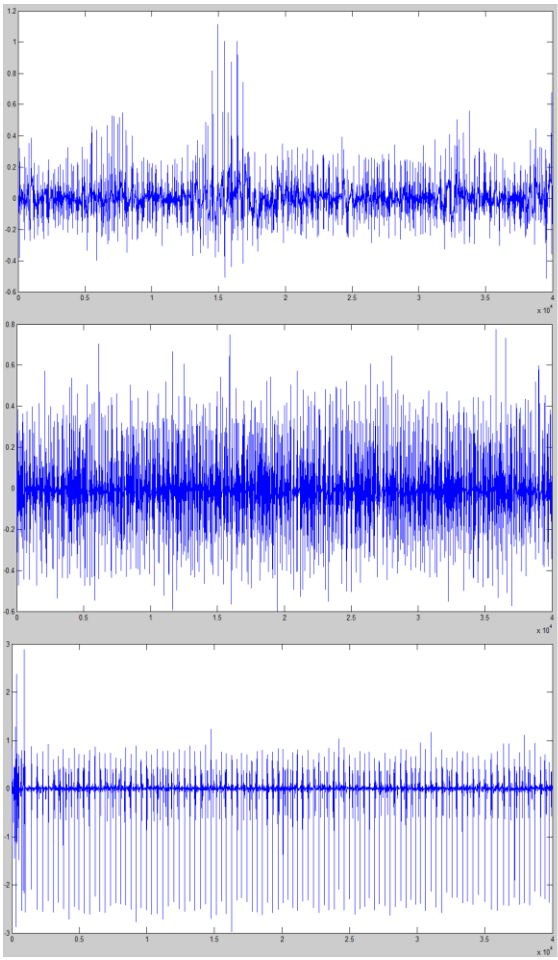

Fig.2. EMG Signals – Normal, Myopathy and Neuropathy

A description of the data set is shown in the Table I.

Table I
Description of dataset used

EMG Signals	Total
Normal	500
Myopathy	500
Neuropathy	500

B. WAVELET TRANSFORM

Wavelet transformis the representation of any signal in the time–frequency domain[10] simultaneously. It can provide time and frequency information at the same point of time, thus giving a time–frequency representation of the signal. It is an already known fact that the higher frequency elements are resolved, better, in time domain while the lower frequency components are resolved, better, in frequency domain. Hence, the wavelet transform can be defined as:

$$C(\alpha,\beta)=\int_{-\infty}^{\infty}x(t)\frac{1}{\sqrt{\alpha}}\varphi\left(\frac{t-\beta}{\alpha}\right)dt \quad (1)$$

It can be seen from the above equation that the resultant signal obtained is a function of two variables, α and β. Here, φ is the transforming function called as the mother wavelet, α is the translation parameter and β is the scale parameter.

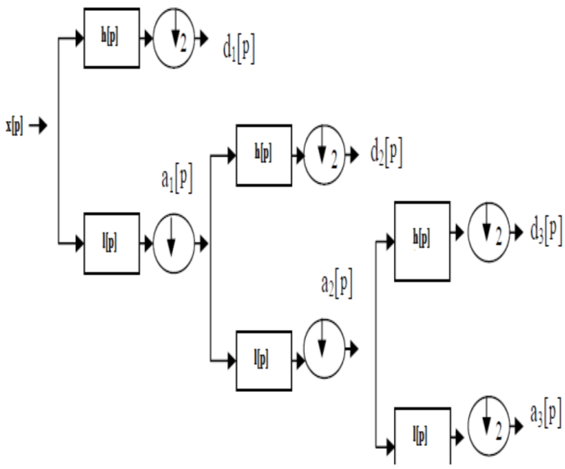

Fig. 3. Sub-band decomposition of DWT implementation; h[p] is the high pass filter, l[p] the low pass filter.

For this, the signal is passed through various high- and low-pass filters. This process is repeated for the either portion or both of the filter's output.

This process is known by the name decomposition. This constitutes one level of decomposition and can be expressed as:

$$y_{hi}(m) = \sum_{p} x(p).h(2m-p) \quad (2)$$

$$y_{lo}(m) = \sum_{p} x(p).l(2m-p) \quad (3)$$

Here, yhi[k] is the output from the high-pass filter and ylo[k] is the outputfrom the low-pass filter after sub-sampling by 2. It is shown in the Fig.3. Decomposition of any signal halves its time resolution, that is, only half the number of samples is required to describe the entire signal. However, this will also result in double the frequency resolution. The process shown in Fig.3 is known as the sub-band coding. This can be further repeated for more decomposition levels. In Fig.3 shows is the process of level decomposition. In the figure, x[n] represents the original signal that need to be decomposed.l[p] and h[p] are low- and high-pass filters, respectively. Thisdecomposition of the x[p] will give the detail and the approximation coefficients. The approximation coefficient may further be decomposed as the process is shown in Fig.4.For the EMG signal, these approximation and detail coefficients are calculated for the Coiflet 5 (coif5) wavelet transform.

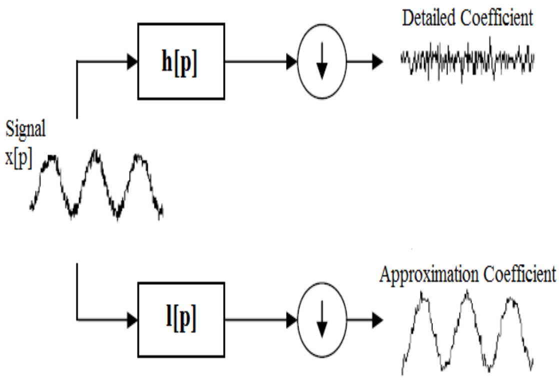

Fig. 4. EMG signal and its wavelet decomposition into Approximation and Detailed coefficients.

Wavelet coefficients extracted as detailed and approximation ones gives the representation of the distribution of EMG signal in time and frequency domain simultaneously.

Now, detail coefficients computed for the EMG signals are selected and the coefficient which highly resembles its original signal is taken out. Here, this coefficient is 4th one i.e. D4 coefficient. It is shown in the Fig. 5 that D4 coefficient resembles its original signal. Feature extraction of the EMG signals is now begins. Features that are going to be calculated for these signals D4 coefficient are mean, energy and standard deviation [11]. These are briefly described as follows:

(1) Mean of the absolute values of the coefficient 4 in each sub-band.

$$Mean = \frac{\sum_{w} z_w}{w} \quad (4)$$

(2) Energy of the wavelet coefficient 4 in each sub-band.Energy of the sub-signal xwk (t) is calculated by

$$Energy = \sum_{w}\sum_{j} \left|D_j^w\right|^2 \quad (5)$$

(3) Standard deviation of the coefficient 4 in each sub-band.

$$\text{Standard Deviation} = \sqrt{\frac{\sum_{w}(z_w - mean)^2}{w}} \quad (6)$$

The features extracted as mean, energy and standard deviation are now utilized to train and test on the classifiers which are going to be explained in the next sub-sections.

C. SVM

The Classifiers are now applied using the features computed in the previous sub-section that includes coiflet wavelet transform of the EMG signals and features like mean, energy and standard deviation for the D4 coefficient of the transform. These features are now utilized to train the SVM model and then test on the same trained SVM for the classification.

SVM (Support Vector Machine) is a classifier which requires training of the model for the testing of the samples and such technique in which training is given is called as supervised learning technique. In this, model is firstly trained to learn according to the features of the classes which need to be classified. Then that trained model is exploited in the classification process. Learning given to the classifier provides better results compared to the other classifiers in which unsupervised learning is used. This classifier is simple and easy to understand as it constructs a hyperplane between the different classes which need to be classified. The classifier used may be linear or non-linear. In linear SVM, training samples of the classes are linearly separable. But it is very difficult in practical situations that a straight line is sufficient to classify each and every sample. For such cases, non-linear classifier is exploited.

In these, kernel functions may also be utilized to increase the dimensionality of the mapping. If the samples are not distinguishable in lower dimensional space, then kernel functions are used. In this, a non-linear operator maps the inputs to the classifier into a higher dimensional space so that samples can be classified easily.
If a linear function is given by the equation

$$g(z) = az + b \tag{7}$$

Then its dimensionality can be increased by using the equation as

$$g(z) = a.\phi(z) + b \tag{8}$$

In this, after plotting the samples in a space, a hyperplane is drawn according to the condition that margin between the support vectors and the hyperplane need to be maximized. It can be seen in [9].

However, classification skill of SVM is better as compared to the other methods, but, some complications still persists in its applications which includesclassifying accuracy as low in the complex applications. It is also difficultto choose the kernel function parameters. In an attempt to find solution for these problems, SVM is combined with k-nearest neighbour classifier (kNN). It will be seen in the results section that doing this is a good approach and better results are obtained using the SVM-kNN. This explained in the next sub-section.

D. SVM-kNN

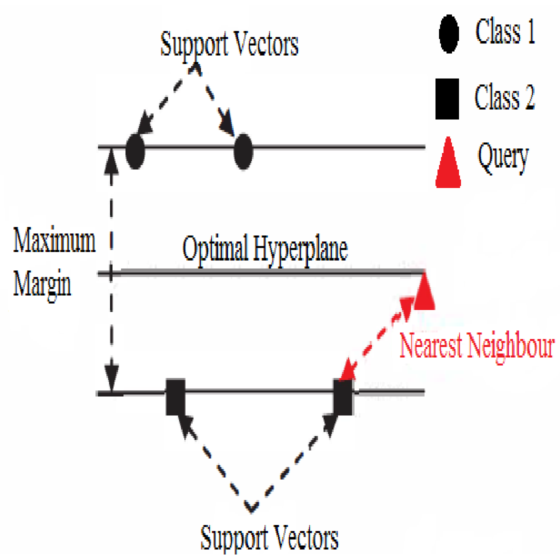

Fig.5. SVM-kNN Classification

SVM-kNN classifier is a hybrid of Support Vector Machine and k-Nearest Neighbour classifiers. In the previous section studied, SVM is a 1NN classifier [9]. That is, according to the nearest neighbour approach, SVM uses only a single representative point as support vectors for each class. But on combining the SVM with kNN, more than one points are chosen from the sample points i.e. say k-points are selected and accordingly, the class is decided for the tested sample.

It can be seen from the Fig.5 that this hybrid algorithm (SVM-KNN) considers more than one support vectors as the representative points of a class. This is much better than the SVM in which samples present nearest to the hyperplane represents the support vectors. In that, only one representative point is chosen for support vectors in each class and this representative point only represents the whole class which becomes quite complicated in the complex situations. In SVM-kNN, all support vectors are considered as representative points for the class and hence, maximum of information of a class is exploited in the classification process. In this classifier, radial basis kernel function is used with SVM to train the classifier. And, during the testing of the signals, nearest neighbour (that is, support vector) is computed as the query point using k-nearest neighbour algorithm. In the next section, the experimental point of view is shown.

IV. Experimental View

EMG Dataset: Data set of EMG signals is loaded from the MIT-BIH database. It is described in the table shown in Table II.

Table II
Description of dataset used

EMG Signals	Training	Testing	Total
Normal	350	150	500
Myopathy	350	150	500
Neuropathy	350	150	500

Computing the wavelet transform and their features: Discrete wavelet transform (DWT) of the EMG signals are computed using the coiflet wavelet transform of the order of 5. Then, Features such as Energy, Mean and Standard Deviation for the D4 coefficient are computed for each and every sample of dataset of the different types of EMG signals used during the research. These calculated features in the form of mean, energy and standard deviation serve as an input to train and test the various classifiers.

Classification: During the classification, computed features of the EMG signals in the second stage are exploited by the classifiers like SVM and SVM-kNN to determine the corresponding class of the samples. This feature set consists of mean, energy and standard deviation of the D4 coefficient of the

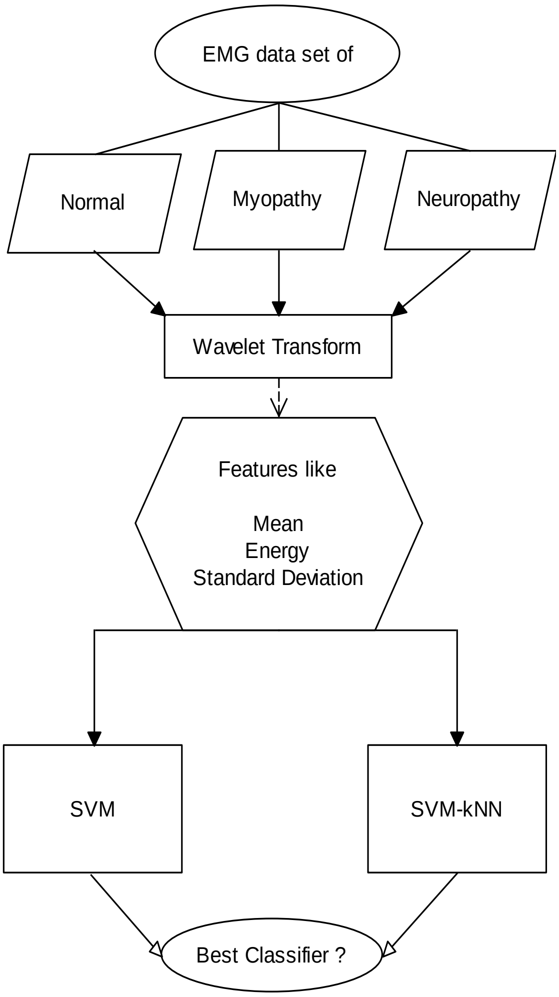

Fig.6. Block Diagram of EMG Signals Classification

coiflet wavelet transform that should efficiently characterize the variations in the input signals for accurate detection and classification of the EMG signals. The calculated features will be applied to the classifiers like SVM and SVM-kNN classifiers as training and testing data to classify the EMG signals in their corresponding classes.

V. Results

In this study, SVM and SVM-kNN classifiers are utilized for the classification of three different types of EMG signals (Normal, Myopathy and Neuropathy). As the features required to train the classifiers, D4 coefficient of the coiflet wavelettransform is used. Then, as the features, mean, energy and standard deviation is used.

Now, data set utilized for this research is shown in the Table I. For this data of the EMG signals, firstly, wavelet transform is applied. In the wavelet transform, coiflet family of order 5 is used. The results for each type of data set used is shown in the figures 7, 8 and 9.

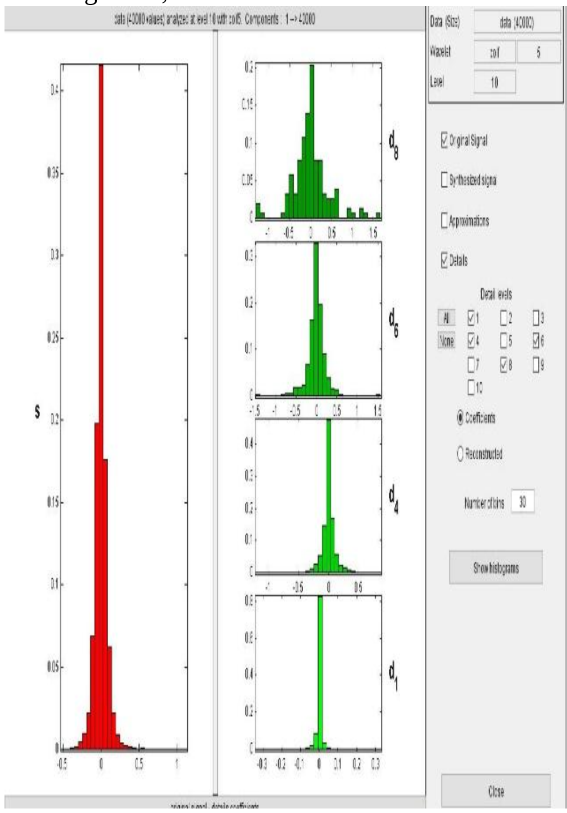

Fig.7. Histogram of wavelet transform of normal person EMG

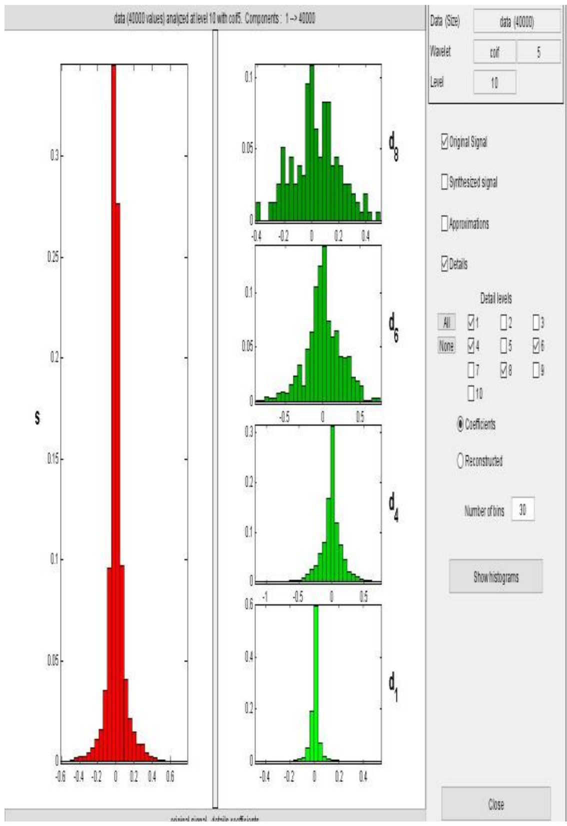

Fig.8. Histogram of wavelet transform of myopathy person EMG

ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 5, Issue 6, June 2016

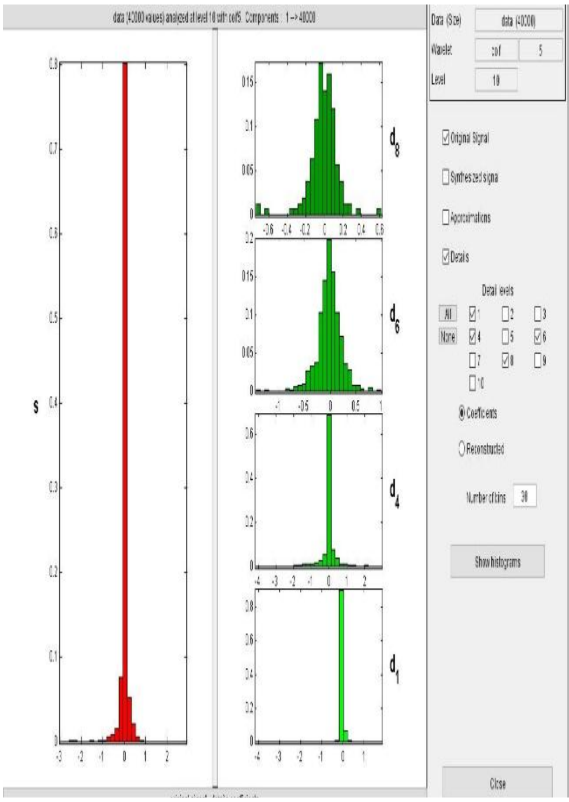

Fig.9. Histogram of wavelet transform of myopathy person EMG

These figures shows the EMG signals wavelet transform in the form of histograms. And it can be observed from the figures that histogram of D4 coefficient maximally resembles the original signals histogram. This implies that maximum information lies in that coefficient and hence, for further analysis, D4 coefficient is sufficient to give the maximum features alone. Therefore, mean, energy and standard deviation is computed for D4 coefficient only. Sample values for these features are shown in the Table III.

Table III. Sample values of features for EMG signals

EMG Signal	Mean	Energy	Standard Deviation
Normal	7.43	56.58	1.173
Myopathy	5.26	28.46	0.890
Neuropathy	9.48	85.34	1.212

In the Table III shown is the sample values for the features computed for the D4 coefficients of the coiflet of order 5. These features include mean, energy and standard deviation values of the signal. These features are calculated for each and every sample of data set of the three types of EMG signals used.

Now, from these features computed, features of 350 samples from each class are selected and used to train the SVM. After that, remaining 450 samples (Normal-150, Myopathy-150 and Neuropathy-150) are tested on the same trained SVM. Its classification results are in the Table IV in the form of a confusion matrix.

Table IV. SVM Classification

Targets→ Outputs↓	Normal	Myopathy	Neuropathy	Accuracy
Normal	141/150	3/150	6/150	94%
Myopathy	11/150	134/150	5/150	89.33%
Neuropathy	7/150	4/150	139/150	92.67%

The confusion matrix shown in the Table IV shows a good result in terms of classifying the EMG signals into their respective classes. SVM gives 92% of accuracy in classifying these signals. This is quite good but SVM-kNN is utilized further to increase this percent of accuracy in classification of EMG signals.

Next, SVM-kNN classifier is used for the EMG signals classification. In this also, features of 1050 samples (Normal-350, Myopathy-350 and Neuropathy-350) are utilized to train the SVM-kNN. Then, the remaining samples are used to test on the same trained SVM-kNN.

Table V. SVM-kNN Classification

Targets→ Outputs↓	Normal	Myopathy	Neuropathy	Accuracy
Normal	147/150	2/150	1/150	98%
Myopathy	8/150	137/150	5/150	91.33 %
Neuropathy	3/150	5/150	142/150	94.67 %

And, it can can be clearly seen from the Table V that increase in the accuracy of classification is observed. Table V shows the confusion matrix of SVM-kNN classification and from this classifier an accuracy of approximately 95% is observed.

VI. CONCLUSION

In the research, a method is presented for the classification of three different types of EMG signals (i.e. Normal, Myopathy and Neuropathy) using the classifier SVM-kNN. This method is based on the extraction of features like mean, energy and standard deviation for the coiflet family of wavelet transform of the order of 5. Results achieved from the classification show that the proposed method can beconsidered as an alternative approach, compared to the other methods, for extracting relevant features and classification for EMG signals.

This method exploits only three features of the EMG signals compared to other existing methods. And, hence, this also allows an efficient classification results with 95% accuracy with only a small number of features.Therefore, now,it can be concluded that,with some more efforts in achieving more accuracy in classification, the expensive tests of diagnosing the EMG can be replaced by this automatic techniques of classifying EMG signals.

REFERENCES

1. Subasi, A., "Classification of EMG signals using combined features and soft computing techniques," Applied Soft Computing, vol. 12, pp 2188- 2198, 2012.
2. Kaur, G. Shatru-Aurora, A., Kumar-Jain, Vijendra., "EMG Diagnosis via AR Modeling and Binary Support Vector Machine Classification," Int Journal Of Eng. Science and Tech., vol. 2, pp. 1767-1772. 2010
3. C.S. Pattichis, A.G. Elia, Autoregressive cepstral analyses of motor unit action potentials, Medical Engineering & Physics 21 (1999) 405–419.
4. Subasi, A., "Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders," Computers in Biology and Medicine, vol. 43, pp 576-586, 2013.
5. Khushaba RN, Al-Jumaily A, Al-Ani, "A. Novel feature extraction method based on fuzzy entropy and wavelet packet transform for myoelectric Control," 2007 Int Symp Commun Inf Technol., pp. 352– 357, 2007.
6. Y. Becerikli, On three intelligent systems: dynamic neural, fuzzy and wavelet networks for training trajectory, Neural Computing & Applications 13 (4) (2004) 339–351.
7. C.A. Pena-Reyes, M. Siper, A fuzzy-genetic approach to breast cancer diagnosis, Artificial Intelligence in Medicine 17 (1999) 131–155.
8. 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.
9. Subasi, A., "Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders," Computers in Biology and Medicine, vol. 43, pp 576-586, 2013.
10. C. Parameswariah, S. Member, M. Cox, and S. Member, "Frequency Characteristics of Wavelets," vol. 17, no. 3, pp. 800–804, 2002
11. Sepulveda-Cano LM, Acosta-Medina CD, Castellanos-Dominguez G "Relevance Analysis of Stochastic Biosignals for Identification of Pathologies," 2011 EURASIP J. Adv. Sig. Proc. 2011.
12. Phinyomark, a. Quaine, F., Charbonnier, S., Serviere, C., Tarpin- Bernard, F., Laurillau, Y., "EMG feature evaluation for improving myoelectric pattern recognition robustness", Expert systems with applications, vol. 40, pp. 4832-4840, 2013.