OPEN ACCESS

International
Journal
Of Modern Engineering Research (IJMER)

Statistical, Energy Values And Peak Analysis (SEP) Approach For Detection of Neurodegenerative Diseases

A.Athisakthi[1],Dr.M.Pushpa Rani[2]

A.Athisakthi is with the Department of Computer Science,Mother Teresa Women's University, Kodaikanal, India e-mail: athisakthi.kumaresan@gmail.com.

Dr.M.Pushpa Rani is with Proff & Head, the Deapartment of Computer Science, Mother Teresa Women's University, Kodaikanal,India.e-mail: drpushpamtw@gmail.com

ABSTRACT: *In this paper, a technique of statistical, Energy values and peak analysis (SEP) approach is used for detection of neurodegenerative diseases from the signal of force sensitive resistors. In this work within the time series Left Stride Interval, Right Stride Interval, Left Swing Interval, Right Swing Interval, Left Stance Interval, Right Stance Interval and Double support interval are obtained and apply the SEP method. In statistical analysis, energy, standard deviation, mean, variance, co-variance are calculated. Two approximations and two details of energy values are extracted from wavelet decomposition. Average peak interval and peak histogram are calculated using peak analysis. Support Vector Machine (SVM) and Random Forest are used as a classifier. Data sets which include a healthy control (HC), various types of Neuro degenerative Diseases: Parkinson's Disease (PD), Huntington Disease (HD), Amyotrophic Lateral Sclerosis. For disease diagnostic Force Sensitive resistor signals are used for evaluation. The results show that the proposed technique can successfully detect the NDD pathologies. For NDD detection, the accuracy, the Sensitivity, the Specificity values are 97%, 97% and 97% using Random forest Classifier.*

Indexed Terms: *Neuro Degenerative Diseases (NDD), Parkinson's Disease(PD), Huntington Disease(HD), Amyotrophic Lateral Sclerosis(ALS), Support Vector Machine(SVM), Random Forest*

I. INTRODUCTION

Neurodegenerative disease is an umbrella term for a range of conditions which primarily affect the neurons in the human brain. An estimated number of people in the World currently suffer from some form neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, Amyotrophic lateral sclerosis, Huntington disease, and vascular cognitive impairment. The symptoms of these diseases vary, but they share a common and gradual decline in a person's cognitive abilities and memory, resulting from a progressive loss of brain cells or brain cell function. This affects the human Motor impairment. It is palpable that diagnosing different Neurodegenerative
diseases in a correct way and a correct time is highly important for the diseases. It would be valuable if it was done with a noninvasive approach. Moreover, identifying a correct disease in a abridge way will help to the physician to decrease the time and cost of the diagnosis process. Diagnosis time is an important factor in the treatment, especially for the progressive diseases.

Gait recognition has attracted attention as a method to identify people, however, this technology is not limited to this purpose, but it can be used in clinical field to classify diseases using gait patterns. Gait signal may be a good factor for discrimination the neurodegenerative disease that is caused by malfunctioning of some brain parts [3]. In this study a novel technique on multilevel feature analysis approach is proposed for detection of neurodegenerative diseases from Gait signals of healthy controls and patients with Parkinson Disease, Huntington Disease, and Amyotrophic Lateral Sclerosis.

Parkinson Disease (PD) is also a degenerative disorder of the central nervous system [7]. Parkinson's disease is a disorder of the brain that leads to shaking (tremors) and difficulty with walking, movement, and coordination. Nerve cells use a brain chemical called dopamine to help control muscle movement. Parkinson's disease occurs when the nerve cells in the brain that make dopamine are slowly destroyed. Without dopamine, the nerve cells in that part of the brain cannot properly send messages. This leads to the loss of muscle function. The damage gets worse with time. Exactly why these brain cells waste away is unknown. Parkinson's disease most often develops after age 50. It is one of the most common nervous system disorders of the elderly. Sometimes Parkinson's disease occurs in younger adults. It affects both men and women. The most common

symptoms are Constipation, Problems with balance and walking, Muscle aches and pains, Loss of small or fine hand movements, shaking called tremors [9]. According to the National Parkinson Foundation [26], it is estimated that between 4 and 6.5 million people worldwide, or about 1% of older adults are diagnosed with this disease

Huntington Disease (HD) is a genetic neurological disorder characterized by abnormal body movements called chorea and a lack of coordination [6]. Huntington's disease is a disorder passed down through families in which nerve cells in certain parts of the brain waste away, or degenerate. Huntington's disease is caused by a genetic defect on chromosome 4. The defect causes a part of DNA, called a CAG repeat, to occur many more times than it is supposed to. Normally, this section of DNA is repeated 10 to 28 times. But in persons with Huntington's disease, it is repeated 36 to 120 times. As the gene is passed down through families, the number of repeats tends to get larger, the number of repeats, the greater your chance of developing symptoms at an earlier age. Therefore, as the disease is passed along in families, symptoms develop at younger and younger ages. Common symptoms of Huntington's disease are slow uncontrolled movements, unsteady gait, Facial movements, including grimaces, loss of judgment, Head turning to shift eye position [9].

Amyotrophic Lateral Sclerosis (ALS) is a progressive neurodegenerative disease that affects nerve cells in the brain and the spinal cord. Motor neurons reach from the brain to the spinal cord and from the spinal cord to the muscles throughout the body. The progressive degeneration of the motor neurons in ALS eventually leads to their demise. When the motor neurons die, the ability of the brain to initiate and control muscle movement is lost. With voluntary muscle action progressively affected, people may lose the ability to speak, eat, move and breathe. The motor nerves that are affected when you have ALS are the motor neurons that provide voluntary movements and muscle control [7].

A. Normal and B. Abnormal Brain model

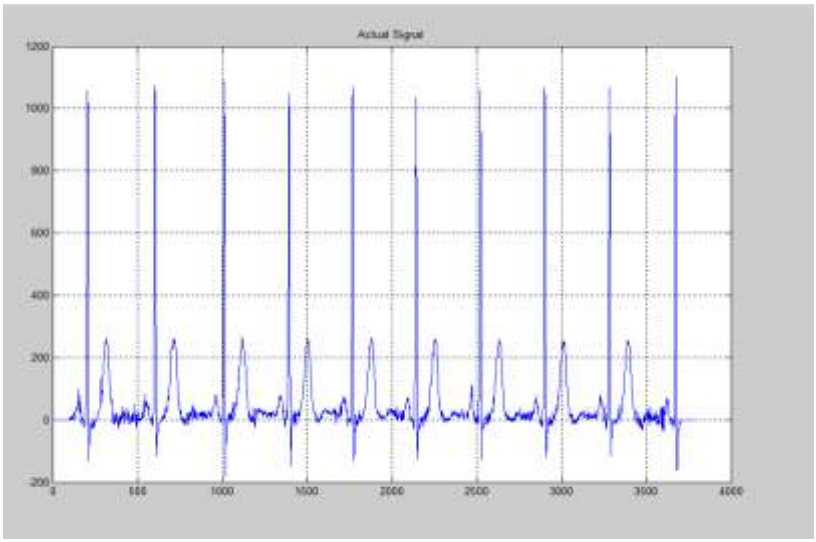

Fig1: Actual Signal of Parkinson Disease

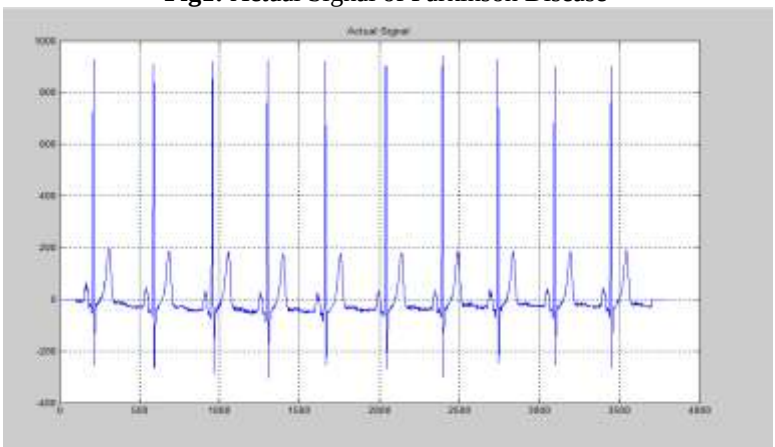

Fig2: Signal of Huntington Disease

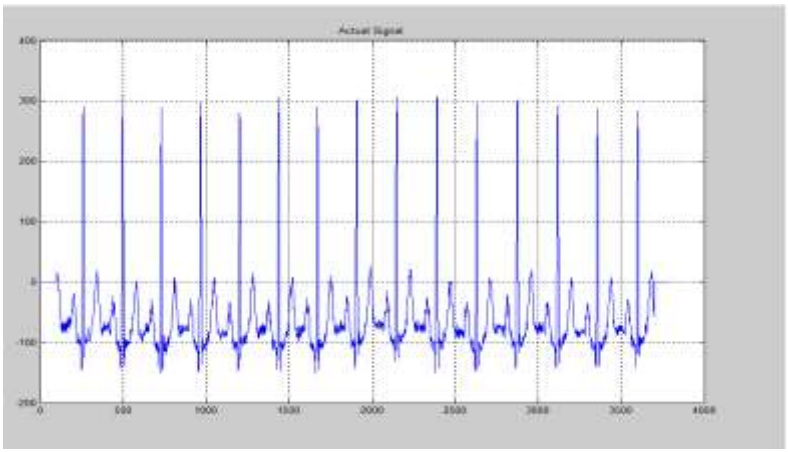

Fig3: Actual Signal of Amyotrophic Lateral Sclerosis

II. BACKGROUND STUDY

P.Murray et al. [5] proved gait is a recognizable pattern of cyclical movement in medical experiments, and did a preliminary analysis of the impact of height, age and other factors on identity. Simon. [2] Says Gait analysis is defined as the analysis of walking patterns of humans. A major application area is in the clinical decision-making and treatment processes for neuro musculoskeletal diseases, among others, such as security clearance systems and human identification. By extracting spatial, temporal parameters from human gait and posture, medical treatments would count with valuable additional information, allowing a better diagnosis and treatment assessment for diseases like Parkinson's [10]. J. M.Hausdorff [11] says Clinicians and researchers alike frequently require a gait analysis system that provides an accurate and reliable measure of temporal and spatial gait parameters for a broad range of clinical populations, where this system can be used in a wide range of settings, including outdoors. Of note importance is the ability to obtain accurate gait stride time statistics, including stride time standard deviations and stride time fractals, for a specified number of continuous walking strides.

Haeri, Sarbaz, & Gharibzadeh, 2005[3] says Gait signal may be a good factor for discriminating movement disorders that is caused by malfunctioning of some brain parts. It can also be used for validation of models that are introduced for the diseases. Huntington's and Parkinson's diseases are both neurological diseases characterized by basal ganglia pathology which affects motor control regulation [12]. In subjects with Amyotrophic Lateral Sclerosis (ALS), the basal ganglia remain intact. However, motoneuron pathology develops in the motoneuron of the cerebral cortex, brain stem, and spinal cord which causes the gait cycle to become abnormal [13]. Stride interval correlation is subject to degeneration due to diminished nerve conduction velocity, loss of motoneurons, decreased reflexes, decreased muscle strength, and decreased central processing capabilities induced by neurodegenerative diseases like Huntington's disease, Parkinson's disease and Amyotrophic Lateral Sclerosis (ALS)

[12]. Hausdorff, et al, [12] proposed that aging and certain diseases affects the stride interval correlation. In their study, they found that the degree of stride to stride interval correlations in subjects with ALS and Huntington's disease were inversely associated with the degree of functional impairment [12]. Hausdorff, et al, [12,13] has shown that all three diseases alter gait rhythm, but it is unknown as to whether these diseases affect the right and the left stride intervals at an equal rate. L. N. Sharma,et al [14] says energy values detection is a good method for capturing relevant components from signals. Jonghee Han, et al [15] Peak analysis can be helpful in the understanding of the disease advancing state which is useful information for the treatments.

III. METHODOLOGY

These diseases were selected because they are known to have similar causes that may complicate diagnosis of them. In this study, we have used a public database from Physionet. The database includes data recorded from 15 patients suffering PD, 20 of HD, 13 of ALS and 16 healthy (control) subjects. The raw data were obtained using force-sensitive resistors, with the output roughly proportional to the force under the foot. Each record included two signals recorded from each foot of the Stride-to-stride measures of footfall contact times were derived from these signals. In this work within the time series Left Stride Interval, Right Stride Interval, Left Swing Interval, Right Swing Interval, Left Stance Interval, Right Stance Interval and Double support interval are obtained and apply (SEP) statistical, Energy values of wavelet decomposition and peak analysis techniques for feature extraction. The selected parameters are expected to be characteristic of the person's overall walking performance.

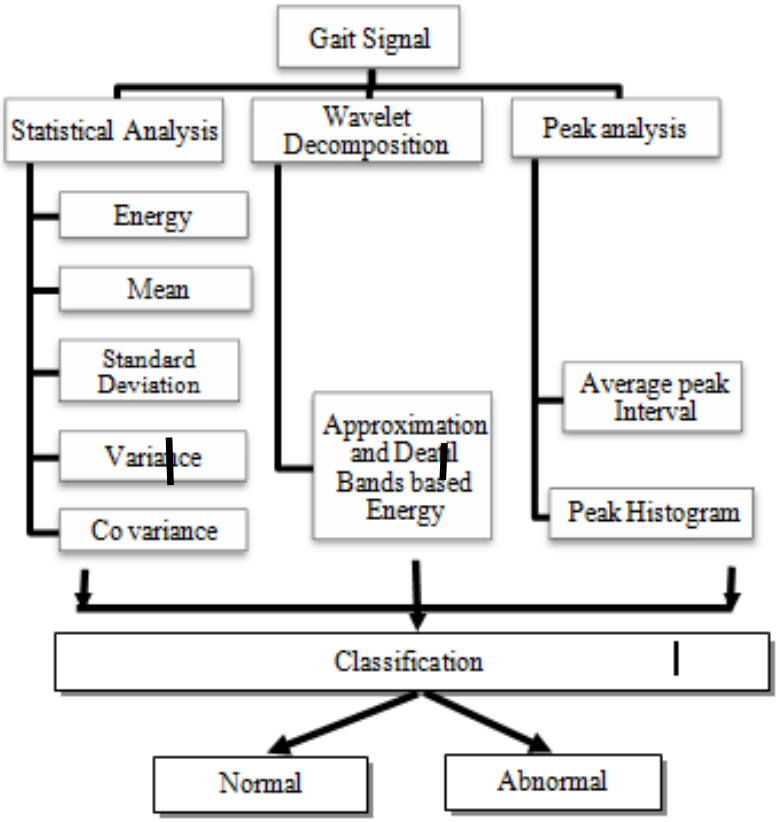

Fig. 4 Detection of Neuro Degenerative Disease from Gait Signal

A gait is a periodic movement and can be decomposed to a single gait cycle. Each gait cycle consists of the stance and swing phase. The stance phase is about 60% of the gait cycle and can be subdivided into the double-leg and single-leg stance [16]. The stance is the term used to designate the entire period during which the foot is on the ground. Stance begins with initial contact. The word swing applies to the time the foot is in the air for limb advancement. Swing begins at the foot is lifted from the floor.

Right initial contact occurs while the left foot is still on the ground and there is a period of double support between initial contact on the right and toe off on the left. During the swing phase on the left side, only the right foot is on the ground, giving a period of right single support, which ends with initial contact with the left foot. There is then another period of double support, until toe off on the right side. Left single support corresponds to the right swing phase and the cycle ends with the next initial contact on the right [17].

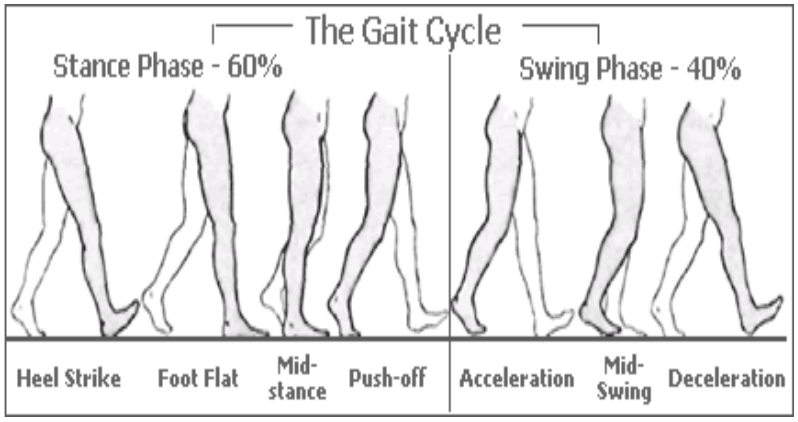

Fig.5. Stance and Swing phase in a Gait Cycle [19]

The stride length is the distance between two successive placements of the same foot. It consists of two step lengths, left and right, each of which is the distance by which the named foot moves forward in front of the other one[18]. In each gait cycle, there are thus two periods of double support and two periods of single support. The stance phase usually lasts about 60% of the cycle, the swing phase about 40% and each period of double support about 10%[17].

A. Statistical Analysis

The purpose of statistical gait analysis is to describe gait functionally, analyzing several tens or hundreds of consecutive steps and it is intended to evaluate the patient during a "functional" walk, typical of the daily life. In this study for each time series of a single subject, we consider Energy, Standard Deviation, Mean, Variance and Co Variance are collected during the same walk. This allows adopting a "statistical gait analysis" approach, ensuring repeatable and accurate results

a) Energy

The energy of a continuous-time complex signal x(t) is defined as

$$Ex = \int_{-\infty}^{\infty} |X(t)|^2 \, dt$$

b) Mean

The Mean is an average value of the signal. To compute the mean value, sum the values in the signal, X_i by letting the index I, run from 0 to N-1. Then finish the calculation by dividing the sum by N

$$\mu = \frac{1}{N} \sum_{I=0}^{N-1} X_I$$

c) Standard Deviation

The standard deviation is similar to the average deviation, except the averaging is done with power instead of amplitude. The signal is stored in x, μ is the mean, N is the number of samples, and α is the standard deviation.

$$\sigma = \frac{1}{N-1} \sum_{i-0}^{N-1} (X_i - \mu)^2$$

d) Variance

The formula for computing the variance of the signal is the mean of its squares minus the square of its mean

$$Var(X) = E[(X - \mu)^2]$$

e) Co Variance

The formula for computing the covariance of the variable X and Y is

$$COV = \sum_{i-1}^{n} (X_i - \bar{x})(Y_i - \bar{y})$$

With $\bar{x}$ and $\bar{y}$ are the means of X and Y respectively

B. Energy Calculation

a). Wavelet Decomposition

Wavelet analysis is one of the better tools for analyzing signals in signal processing. The main properties of wavelets for the time domain are (i) a wavelet is a finite function and it's admissible, i.e. in spite of oscillation, it has a zero average; (ii) a wavelet is a regular function, with the derivative properties of smoothness and continuity; (iii) a wavelet is a function with compact support, that means it is located in the space. It is used to break a signal down into its constituent parts for analysis. The diagnostic information of the original signal is distributed in different wavelet sub bands based upon their bandwidth or frequency content. It has been reported that the lower frequency sub bands contain most of the diagnostically significant information of the original signal [21].In this work two level wavelet Decomposition is used for extracting features. Totally one approximation and three details were obtained from the decomposed signal.

Basically, a wavelet is a function $\varphi \in L^2(R)$ with a zero $\int_{-\infty}^{\infty} \varphi(t)dt = 0$.

The continuous Wavelet Transformation (CWT) of a signal x(t) is then defined as:

$$CWT_{\varphi}x(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t)\varphi^* \left[\frac{t-b}{a}\right] \square dt$$

Where y (t) is called the mother wavelet, the asterisk denotes complex conjugate, while a and b (a, b∈ R) are scaling parameters respectively. The scale parameter a determines the oscillatory frequency and the length of the wavelet, and the translation parameter b determines its shifting position. The scaling function is

associated with the low-pass filters (LPF), and the wavelet function is associated with the high-pass filters (HPF). The decomposition procedure starts by passing a signal through these filters. The approximations are the low-frequency components of the time series and the details are the high-frequency components. The signal is passed through a HPF and a LPF. Then, the outputs from both filters are decimated by 2 to obtain the detail coefficients and the approximation coefficients at level 1 (A1 and D1). The approximation coefficients are then sent to the second stage to repeat the procedure. Finally, the signal is decomposed at the expected level [20].

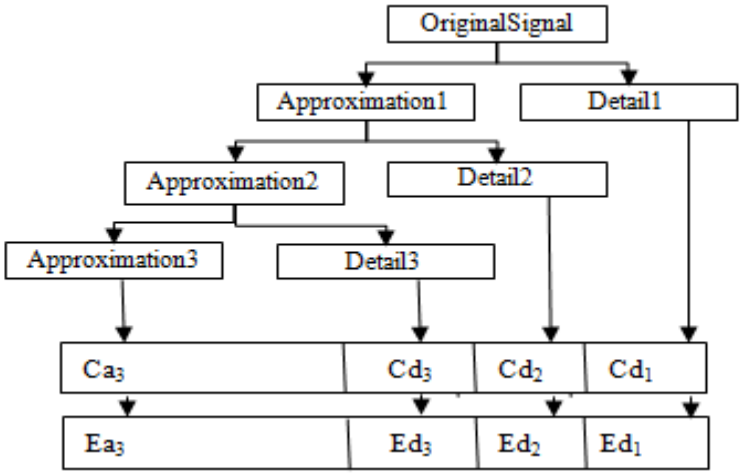

Fig.5 Energy extraction using wavelet decomposition

The energy of a vague signal can be separated at different resolution levels. Mathematically, this can be presented at

$$ED_{i=}\sum_{j=1}^{N}|D_{ij}|^2, i = 1, \ldots . l$$

$$EA_{i=}\sum_{j=1}^{N}|A_{ij}|^2$$

Where i =1,.....,l is the wavelet decomposition level from level 1 to level l. N is the number of the coefficients of detail or approximate at each decomposition level. ED_i is the energy of the detail at decomposition level i and EA_l is the energy of the approximate at decomposition level l.

C. Peak Analysis

Peaks are most noticeable and useful features for analyzing signals. In this study within the time series signal, the peaks are detected; the average peak intervals and histogram are calculated from Left Stride Interval, Right Stride Interval, Left Swing Interval, Right Swing Interval, Left Stance Interval, Right Stance Interval and Double support interval signals.

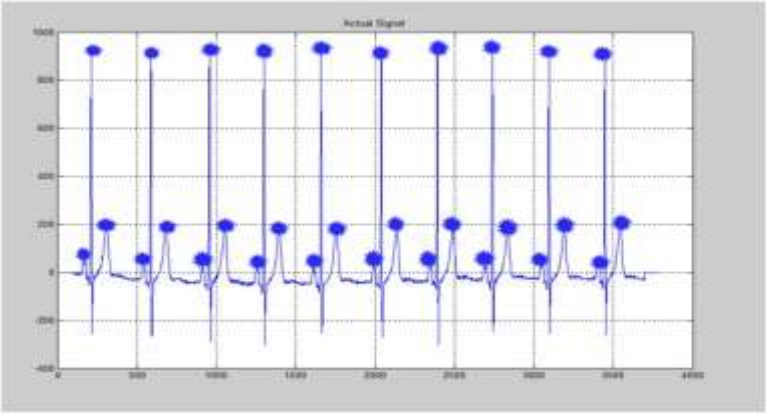

Fig7. Peak detection

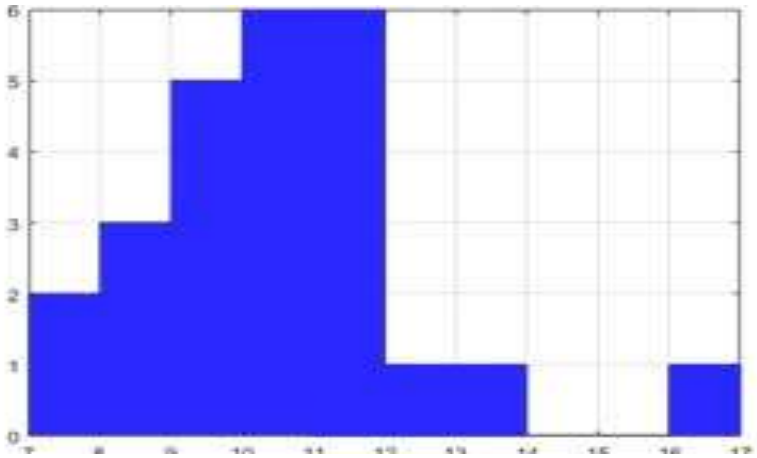

Fig8. Histogram of Peak Interval

D. Classification

In this work, two classifiers as support vector machine (SVM) and Random Forest (RF) are used for detection and classification. The RF and the SVM are supervised learning models which are used for identifying the class levels of the test feature vectors [8]. The SVM is a two class classifier. It works with an objective to maximize the hyper-plane [23] corresponding to the decision boundary. The multiclass classification using SVM is accomplished through various multiclass coding techniques like 'one VS one', 'one VS all' etc. [24]. The SVM classifies function uses results from SVM train to classify vectors x according to the following equation:

$$c = \sum_{i} \alpha_{ik(s_{i,x})+b}$$

Where S_i is the support vector, α_i is the weight, b is the bias, and k is a kernel function. In the case of a linear kernel, k is the dot product. If c >= 0, then x is classified as a member of group 1, otherwise it is classified as a member group 2.

The other classifier random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution of all trees in the forest. In definition Leo Breiman [10] says A random forest is a classifier consisting of a collection of tree structured classifiers {h(x,Θk), k=1, ...} where the {Θk} are independent identically distributed random vectors and each tree casts a unit vote for the most popular class at input x.

The algorithm for random forests applies the general technique of bagging, to tree learners. Given a training set $X = x_{1...},x_n$ with responses $Y= y_1... y_n$, bagging repeatedly (B times) selects a random sample with replacement of the training set and fits trees to these samples for b = 1,.., B, Sample, with replacement, n training examples from X, Y; call these X_b, Y_b. Train a decision or regression tree f_b on X_b, Y_b. After training, predictions for unseen samples x' can be made by averaging the predictions from all the individual regression trees on x':

$$\hat{f} = \frac{1}{B}\sum_{b=1}^{B} \hat{f}_b(x')$$

IV. PERFORMANCE MEASURES

The performances of the classifiers are evaluated in terms of sensitivity, specificity and accuracy. These parameters are estimated by comparing the actual test output and the predicted output. A confusion matrix visualizes the number of true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN) for a classifier. The sensitivity [25] relates to the ability of trained model to identify positive results of Neuro degenerative diseases. The sensitivity (SE) is evaluated as [25]

$$SE = \frac{TP}{TP + FN}$$

The specificity (SP) [25] is related to the ability of identifying the negative outcomes (healthy control or non-infracted). The specificity (SP) is evaluated as [25].

$$SP = \frac{TN}{TN + FP}$$

The classification accuracy (Acc) of a measurement system is the degree of closeness of measurements of a quantity to that of its actual (true) value and it is defined as

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

V. RESULTS AND DISCUSSION

In this work the signals are obtained from the public database Physionet. The database includes data recorded from 15 patients suffering Parkinson Disease, 20 of Huntington Disease, 13 of Amyotrophic Lateral Sclerosis and 16 healthy (control) subjects. The raw data were obtained using force-sensitive resistors.

In this study NDD detection is evaluated using the proposed method (SEP) that is Statistical, Energy Space and Peak analysis. In statistical analysis, Energy, Standard deviation, Mean, Variance and Co Variance are evaluated for from Left Stride Interval, Right Stride Interval, Left Swing Interval, Right Swing Interval, Left Stance Interval, Right Stance Interval and Double support interval signals. In energy calculation, the energy space is calculated from wavelet decomposed signals. In peak analysis Average peak interval and Histogram was calculated. These three analysis methods are combined for performance evaluation. Table-I shows the performance of SVM and RF classifiers classification rate. Table-II shows the Sensitivity Rate and Table – III Shows the Specificity Rate

TABLE I
Classification Rate of SVM and RF Classifiers

Classification Rate	Stride	Swing	Stance	Support
RF	93.75	93.75	93.75	96.875
SVM	81.25	84.375	84.375	81.25

TABLE II
Sensitivity Rate of SVM and RF Classifiers

Sensitivity Rate	Stride	Swing	Stance	Support
RF	91.6667	87.5	87.5	93.75
SVM	81.25	70.833	85.416	83.33

TABLE III
Specificity Rate of SVM and RF Classifiers

Sensitivity Rate	Stride	Swing	Stance	Support
RF	91.6667	96.153	96.153	98
SVM	70.7407	79.4372	78.9855	76.666

Table IV and Fig 9 Shows Total Accuracy, Sensitivity and specificity rate of Support Vector Machine is 96.875%, 85.4167%, 88.8571% respectively. And total Accuracy, Sensitivity and specificity rate of Random forest Classifier is 96.875%, 96.755%, 96.725%.

TABLE IV

Total Accuracy, Sensitivity and specificity rate of SVM and RF Classifiers

Performance Analysis	RF	SVM
Accuracy	96.875	90.625
Sensitivity	96.755	85.4167
Specificity	96.725	88.8571

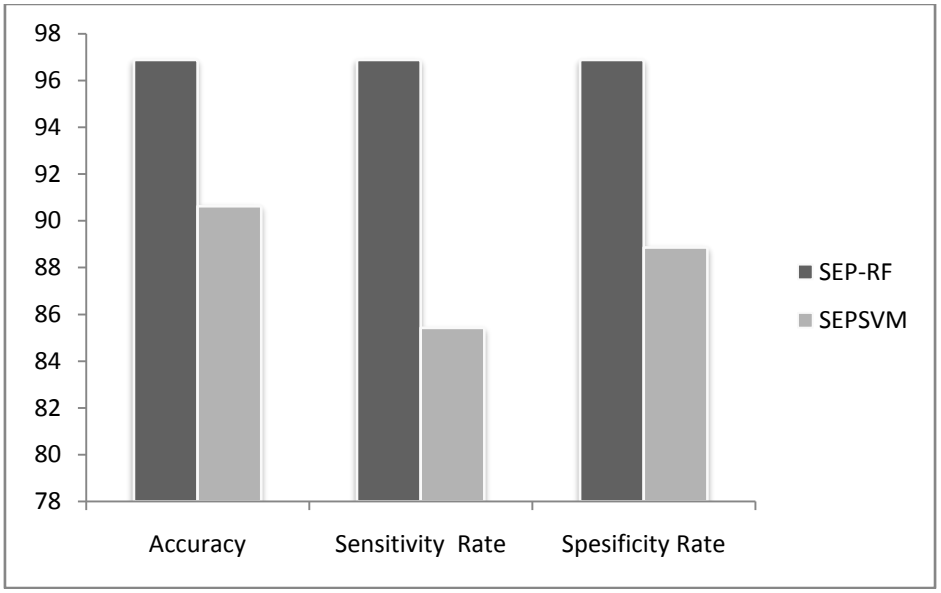

Fig 9. Classification Rate Diagram

VI. CONCLUSION

In this work, SEP approach is proposed for detection of Neuro degenerative diseases from force sensitive resistors signals. Most of the present works are based on single method analysis. The proposed NDD detection is combined three set of analysis. This increases the accuracy percentage of analysis. The SEP features show an accuracy of 96.875% for NDD detection using Random Forest Classifier. The proposed method is simpler compared to earlier method for NDD detection. This method does not require preprocessing of Signal.

REFERENCES

[1]. Simon, S.R., 2004. Quantification of human motion: Gait analysis – benefits and limitations to its application to clinical problems. J. Biomech. 37, 1869–1880.
[2]. Haeri, M., Sarbaz, Y., & Gharibzadeh, S. (2005). Modeling the Parkinson's tremor and its treatments. Journal of Theoretical Biology, 236(3), 311–322.
[3]. C. A. Ross and W. W. Smith, "Gene-environment interactions in
[4]. Parkinson's disease," Parkinsonism Relat. Disord., vol. 13, no. 3, pp. S309–S315, 2007.
[5]. Murray, M.P.: Gait as a total pattern of movement. American Journal of Physical Medicine 46(1), 290–333 (1967)
[6]. Banaie, M., Sarbaz, Y., Gharibzadeh, S., & Towhidkhah, F. (2008). Huntington's disease: Modeling the gait disorder and proposing novel treatments. Journal of Theoretical Biology, 254(2), 361–367. 21 September 2008.

[7]. Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). Principles of neural science (4th ed.). New York: McGraw-Hill.

[8]. Huang, Chenn-Jung and et al, "Application of wrapper approach and composite classifier to the stock trend prediction," Expert Systems with Applications, vol. 34, no. 4, pp. 2870–2878, 2008.
[9]. www.ncbi.nlm.nih.gov
[10]. RANDOM FORESTS Leo Breiman Statistics Department University of California Berkeley, CA 94720 September 1999.
[11]. J. M. Hausdorff, "Gait dynamics in Parkinson's disease: Common and distinct behavior among stride length, gait variability, and fractal-like scaling," Chaos, vol. 19, no. 2, p. 26113, Jun. 2009.
[12]. J.M. Hausdorff, "Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis," in Journal of Applied Physiology, vol. 88, issue 6, , 2000, pp. 2045-2053.
[13]. J.M Hausdorff,, "Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease," in Journal of Applied Physiology, vol. 82, no. 1, 1997, pp. 262-269.
[14]. L. N. Sharma, R. K. Tripathy and S. Dandapat, "Multiscale energy and eigenspace approach to Detection and Localization of Myocardial Infarction",in IEEE Transactions on Biomedical Engineering. 2015.
[15]. Marion Trew, "Human Movement", CHURCHILL LIVINGSTONE.
[16]. G. Rau, "Movement biomechanics goes upwards: from the leg to the arm", Journal of Biomechanics, vol. 33, Issue 10, pp. 1207-1216, Mar. 2000.
[17]. Ashutosh Kharb, Vipin Saini, Y.K Jain, Surender Dhiman," A review of gait cycle and its parameters", IJCEM International Journal of Computational Engineering & Management, Vol. 13, July 2011.
[18]. Christopher Kirtl,"Clinical Gait Analysis: Theory and Practice" December 1, 2005.
[19]. http://fasthewoodlands.com.
[20]. Omerhodzic, S. Avdakovic, A. Nuhanovic, K. Dizdarevic, "Energy Distribution of EEG Signals: EEG Signal Wavelet-Neural Network Classifier"
[21]. Sharma, L. N. and et al, "ECG signal denoising using higher orderstatistics in wavelet subbands," Biomedical Signal Processing and Control, vol. 5, no. 3, pp. 214 – 222, 2010.
[22]. KAISER, G. A friendly guide to wavelets. Boston: Birkhauser, 1994. 325 p.
[23]. Corinna Cortes and Vladimir Vapnik, "Support vector machine," Machine learning, vol. 20(3), pp. 273–297, 1995.
[24]. Suykens, Johan AK and et al, Least squares support vector machines.World Scientific, 2002, vol. 4
[25]. D. G. Altman and J. M. Bland, "Diagnostic tests 1: Sensitivity and specificity." Medical Statistics Laboratory, Imperial Cancer Research Fund, London., vol. 308, p. 1552, June 1994.
[26]. National Parkinson Foundation, http://www.parkinson.org