Journal of Medical Systems (2019) 43: 245
https://doi.org/10.1007/s10916-019-1384-4

IMAGE & SIGNAL PROCESSING

# An efficient Gait Dynamics classification method for Neurodegenerative Diseases using Brain signals

**S. S. Sreeja Mole [1] · K. Sujatha [2]**

Received: 8 February 2019 / Accepted: 10 June 2019 / Published online: 25 June 2019


## Abstract

Neurons of the human brain are primarily affected by the Huntington's disease (HD), Amyotrophic Lateral Sclerosis (ALS), Parkinson's disease and so on. Classification of these neurodegenerative diseases (NDD) is clinically important to analyze the destruction of nerve cells. Early diagnosis of NDD'S helps in saving the human life. Based on the report of previous studies, motor impairment or human gait cycle is largely affected by the clinical symptoms of NDD. Accurate diagnosis of various neurodegenerative diseases in correct time is very important for early diagnosis of the disease. Diseases can be diagnosed earlier by means of characterizing the gait cycle. In this work, a gait dynamics classification method is proposed for determining the neurodegenerative diseases from the brain signals using multilevel feature extraction method. From force sensitive resistors, the left and right feet signals recorded in 60 one minute are included in the input database. It is obtained through fixing 16 healthy subjects, 13 ALS, 20 HD, and 15 PD. Using six levels of Discrete Wavelet Transform (DWT), the features are determined by means of decomposing the raw signal. Ultimately, the pathological gait signals are classified through exploiting three multilevel feature extraction techniques named as, (Detrended Fluctuation Analysis (DFA), Positive, Negative Peak Histogram Analysis (PNPHA) (proposed Method) and Statistical Temporal parameter Analysis (STA)). Experimental outcomes proved that the gait dynamics are successively distinguished between NDD and group of healthy controls using the proposed method.

**Keywords** Parkinson Diseases · Huntington Diseases · Amyotrophic Lateral Sclerosis · Neuro-Degenerative Diseases · Discrete Wavelet Transform · Positive negative Peak Histogram Analysis

## Introduction

In today's world, most of the peoples are suffering from crucial neurodegenerative diseases (NDDs) such as, amyotrophic lateral sclerosis (ALS), Parkinson's disease, Alzheimer's disease and so on. Memory and cognition of human brain are largely affected by a well-known degenerative disease called Alzheimer's disease. Mainly, memory and language nerve cells of brain are degenerated by the Alzheimer's disease.

This article is part of the Topical Collection on *Image & Signal Processing*

✉ S. S. Sreeja Mole
sreejaphd@gmail.com; sreebommy@gmail.com

[1] Department of ECE, Christhujyothi Institute of Technology and Science, Yeswanthapur, Jangaon, Telangana, India

[2] Department of Computer Science, Wenzhou-Kean University, Wenzhou, China

Alzheimer's disease is commonly seen in old aged people above 65 years. Conversely, this neurodegenerative disease belongs to the dementia [1] form. Other reasons for the cause of neurodegenerative diseases are alcohol intake, smoking, age and may be from genes (i.e. due to genetic factors) [2]. Alzheimer's disease can be recognized significantly from the symptoms such as complete memory loss, abnormal moods, verbal communication, misplacement of things, and poor decision making [3, 4]. Disease affected patients will enter into critical stage if not diagnosed or treated early the occurrence of disease. Three different stages used for diagnosing AD are, monitoring MRI scans, taking neuropsychological tests, and general physician consultation. Neurofibbrillary Tangles (NFT) and Mini-Mental state Examination (MMSE) are the currently tool for diagnosing the brain disease. But, these tools have no ability to diagnose the disease in early stage until the severity of disease wholly affects the brain [5]. In rare cases, dementias can show similar clinical symptoms; thereby, diagnosing this disease in the early stage is a tedious and complex task [6]. In present studies, for healthy adults, their gait interval parameters are measured using the computer-aided tools [7–20]. According to Aziz and Arif [7], different types of symbol sequence are produced through converting the series of stride time. Followed by this, the gait complexity is analyzed by means of applying a threshold dependent symbolic entropy method.

Human gait dynamics are simulated using the supercentral pattern generator (SCPG) of Scafetta et al. [8]. By considering ALS, HD, and PD, the fractal and stochastic properties are also discussed. With healthy controls and ALS subjects, the gait cadences are classified using statistical analysis method proposed by Wu and Shi [9]. Parzen-window method was utilized to estimate the gait cadence probability density functions. The accuracy rate achieved on classifying the control subjects and ALS stride patterns are 82.8%. A theory developed in the work [10] has proved that the swing interval turns count (SWITC) of healthy subjects is totally different from ALS patients. Depending on gait dynamics, the NDDs are diagnosed using an approach developed by Daliri [11]. For accurate diagnosis, different kernels included in support vector machines were analyzed. Uniformities of multi-resolution levels are analyzed to compare the series of right and left stance-interval. For stance time fluctuations, the multi-resolution entropy is analyzed to investigate gait asymmetries in NDDs (Liao et al. [13]). With HD, PD, and HLS, the similarity of gait is disturbed or collapsed. But, when compared to ALS subjects, the disturbance degree is less. High progression in the neurodegenerative impairment can increase the randomness of walking individuals [14–16].

Past studies have evidently proved that the neurological disorders can be diagnosed easily by possibly extracting features included in the gait dynamics. The aim of this paper is to present a new approach for automatic diagnosis of neurodegenerative diseases in a computer-aided method. The rest of this manuscript has been organized in the following form: The next section describes the dataset description and proposed method. The results and discussion section explains about the implementation part and the conclusion describes in conclusion section.

# Method

## Dataset description

In this work, we have used a benchmark dataset from the public database Physionet [www.physionet.org]. The database includes 15 patients with Parkinson's disease, 20 patients with Huntington Disease, 13 patients with Amyotrophic Lateral Sclerosis and 16 healthy subjects. The force-sensitive resistors are used to assess the gait patterns. The result is approximately correlated to the subjects force under the foot. These unfiltered gait rhythms carry both left and right stride interval, swing interval, stance interval and double support interval. The frequency of the gait signal measured in seconds. For each subject, the clinical information like age, gender, height, weight, walking speed, disease severity are stated in the database. Figure 1 shows stride interval samples from different types of subjects.

## Proposed Approach

The outline of the proposed approach is illustrated in fig. 2. In this method, the intermediate differences ($ID = (x_n - x_{n-})$) are computed for each time series of the gait signal for feature extraction. This process helps to increase the accuracy rate, instead of using a raw gait signal. Figure 3 shows the intermediate difference of stance, stride and swing phases. For that predicament situation, three multilevel efficient feature extraction methods (Detrended Fluctuation Analysis (DFA), Positive, Negative Peak Histogram Analysis (PNPHA) and Statistical Temporal parameter Analysis (STA)) are exploited for classifying the pathological gait signals. The PNPHA is a proposed method to enhance the accuracy rate. The Random Forest machine learning algorithm is applied to classify the normal and morbid signal.

## Fluctuation Analysis

In complex time series analysis, quantifiers are playing a more important role in analyzing intricate and inconsistent signals. Detrended Fluctuation Analysis is one of the efficient quantifier, which quantifies auto correlation properties of the gait signal [21]. It is mandatory to fetch the attributes of the local oscillation at different time scales for the analysis process. In many clinical assessments, the DFA Quantifier can be used to discriminate healthy and diseased signal [22]. In the DFA algorithm, we calculate the internal difference of the time series. The total length of the signal is$N$.

$$y(k) = \sum_{i=1}^{k} \left[ S(i) - S_{avg} \right] \quad (1)$$

Where $S(i)$ is the $I^{th}$signal interval, $S_{avg}$ is the mean value of the signal. Then the length data signal divided into small segments. For each segment separately, we find the linear approximation $y_n$ using least square fit. It represents the trend in the given section. By subtracting the local trend $y_n(k)$ in each segment the time series is detrended. The average fluctuation $f(n)$ of this detrended time series is

$$zF(n) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} \left[ y(k) - y_n(k) \right]^2} \quad (2)$$

This data processing is repeated for all signal segments. Depending upon the size of the segment$n$, the $f(n)$ will increase. Next, we create a log-log plot of log$f(n)$versus the

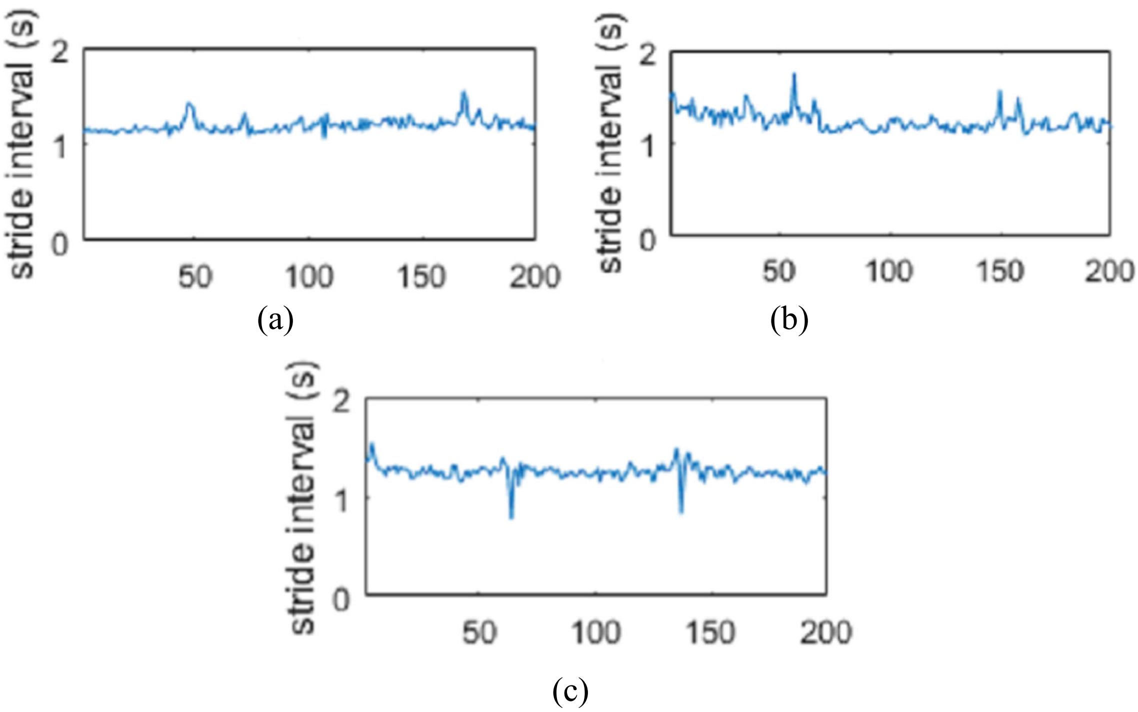
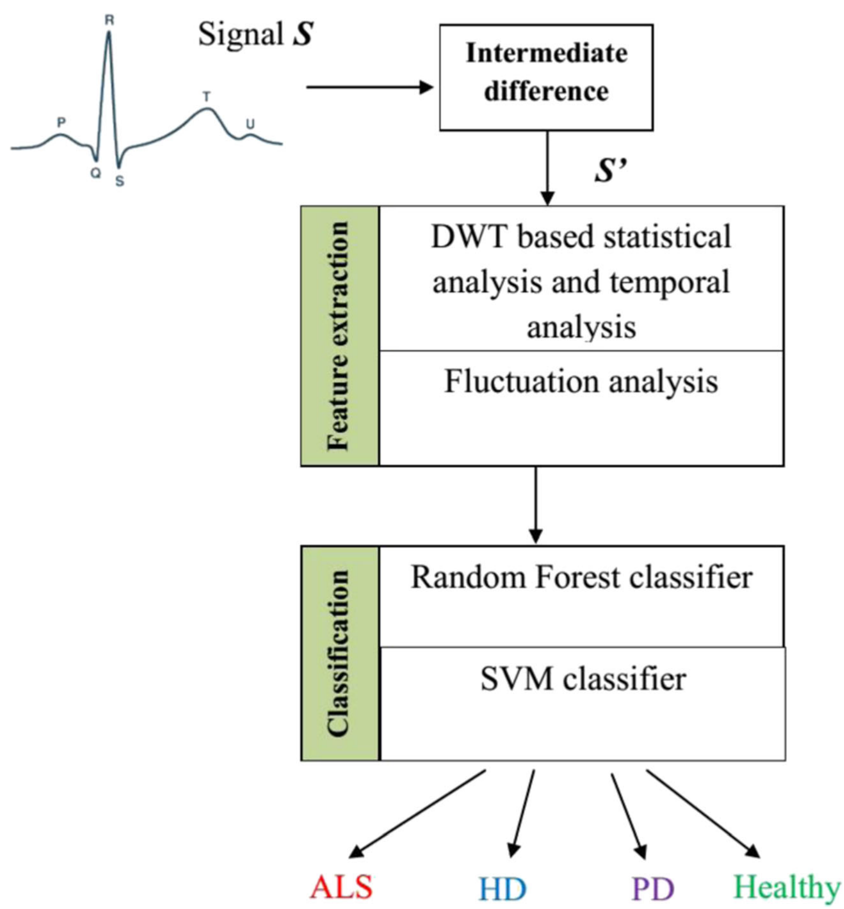



Fig. 1 Sample of real data (a) ALS, (b) HD, (c) PD

Fig. 2 Flow diagram of proposed Gait signal classification



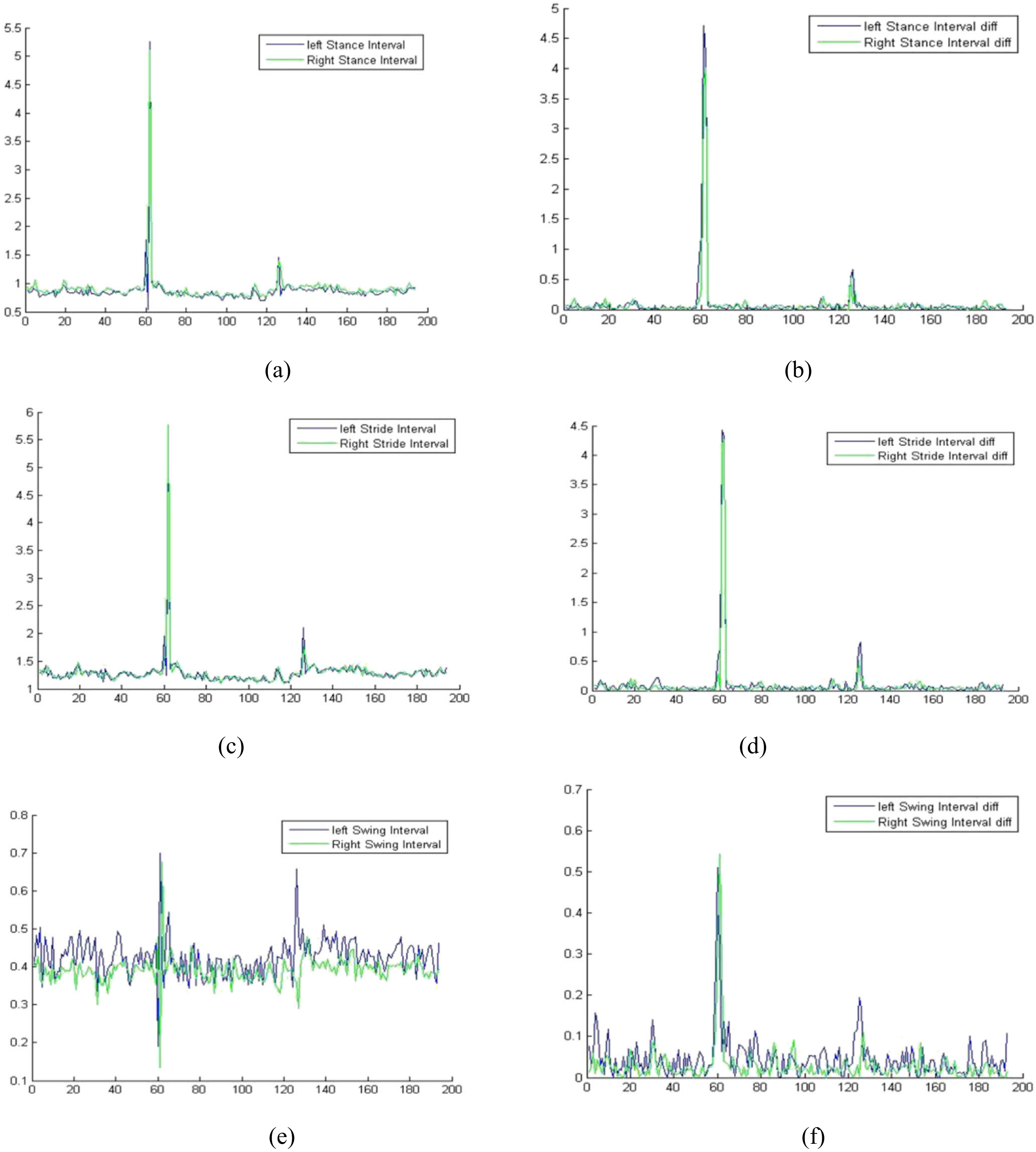



**Fig. 3** Intermediate Differences of (a) (b) stance, (c) (d) stride and (e) (f) swing phases

segment size$n$. The linear dependence indicates the self-fluctuation presence and the slope line determines $f(n) = n^{\alpha}$the self-similarity parameter [21, 22]. The autocorrelation properties of the signal represented by the$\alpha$. This DFA computation is proposed for a long time series of gait signals, which fetch from the force sensitive resistors.

## Statistical and Temporal Analysis

Before computing the Statistical and Temporal parameters we decompose the gait signal using Discrete Wavelet Transform. The wavelets are discretely sampled is known as Discrete Wavelet Transform [23]. It gives a better result in analyzing

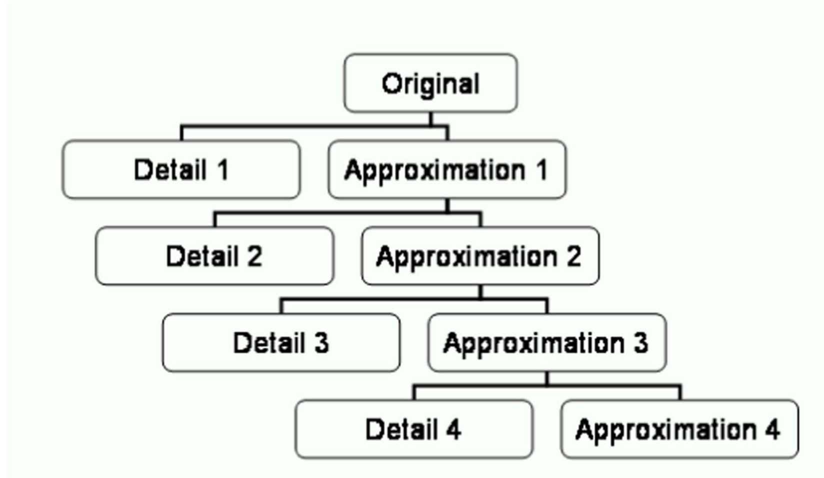



**Fig. 4** Wavelet Decomposition of Time series Signal

unstable time series gait signal. In the DWT Function (shown in fig. 4), the signals are decomposed using a Low pass filter and High pass filter simultaneously. The result of DWT, approximation, represents the low pass coefficient sequence and detail represents the high pass coefficient sequence. The high-frequency fluctuations are removed by low pass filters and the slow trends are removed by High pass filters.

$$x_{low}[n] = \sum_{k=-\infty}^{\infty} s[k]l[n-k] \tag{3}$$

$$x_{high}[n] = \sum_{k=-\infty}^{\infty} s[k]h[n-k] \tag{4}$$

**Statistical Analysis** In this study, we consider Energy, Standard Deviation, Mean, Variance and Co-Variance for each time series of a single subject are collected during the same walk.

(a) *Energy*

The Energy Discrete time complex signal is defined as

$$E_s = (x(n), x(n)) = \sum_{n=-\infty}^{\infty} |x(n)|^2 \tag{5}$$

(b) *Mean*

The Mean is the 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 $N$

$$\mu = \frac{1}{N} \sum_{I=0}^{N-1} X_I \tag{6}$$

(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 a, $\mu$ is the mean, $N$ is the number of samples, and $\alpha$ is the standard deviation.

$$\sigma = \frac{1}{N-1} \sum_{i-0}^{N-1} (a_i-\mu)^2 \tag{7}$$

(d) *Variance*

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

$$Var(X) = E\left[(X-\mu)^2\right] \tag{8}$$

(e) *Co-Variance*

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

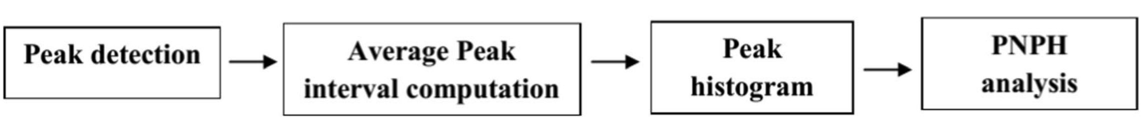



**Fig. 5** Temporal Analysis

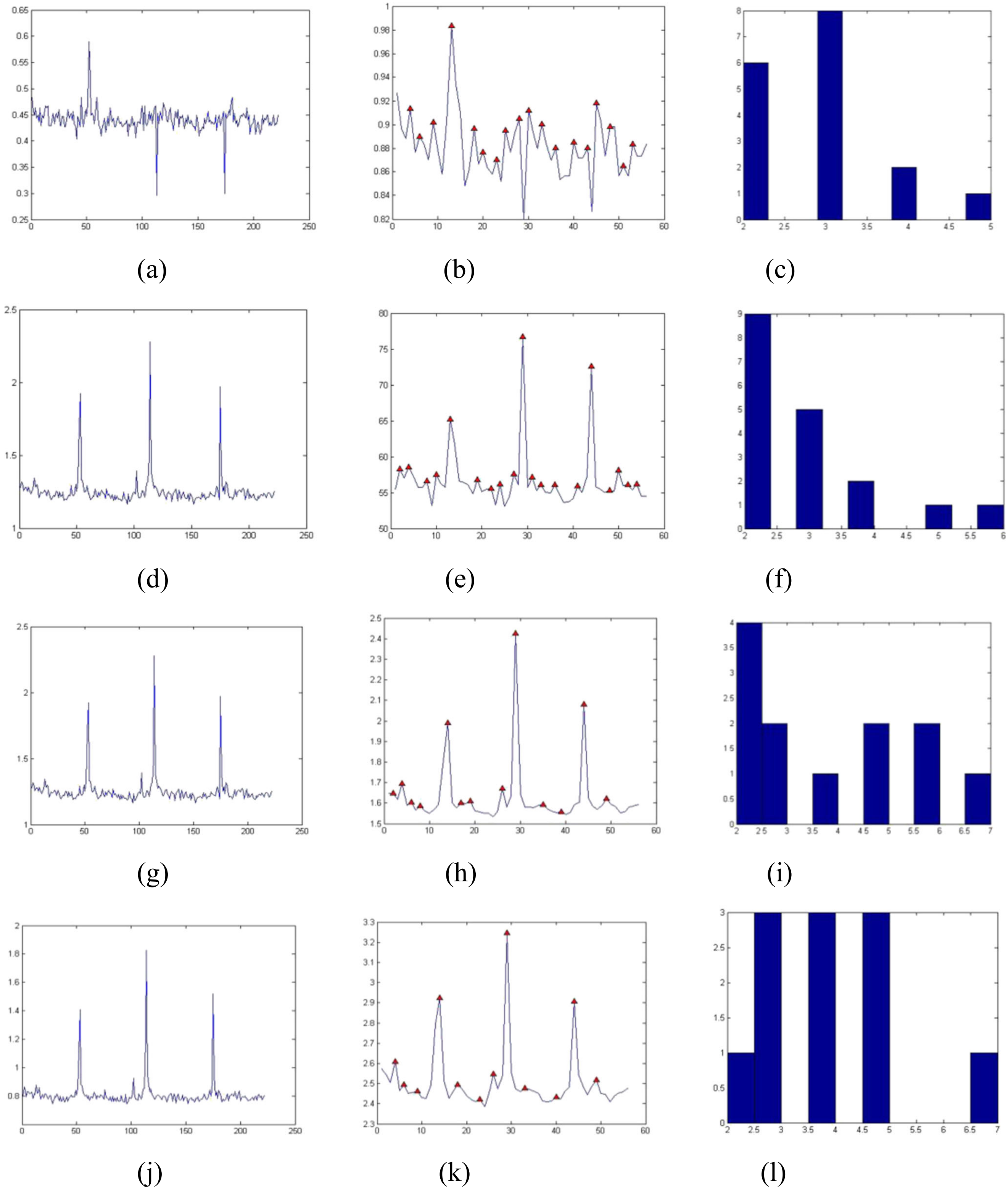



**Fig. 6** (a) (b) (c) Average Peak Interval, (d) (e) (f) Peak Detection and Peak Histogram of Swing phase, (g) (h) (i) Support phase, (j) (k) (l) Stride Phase and stance phase

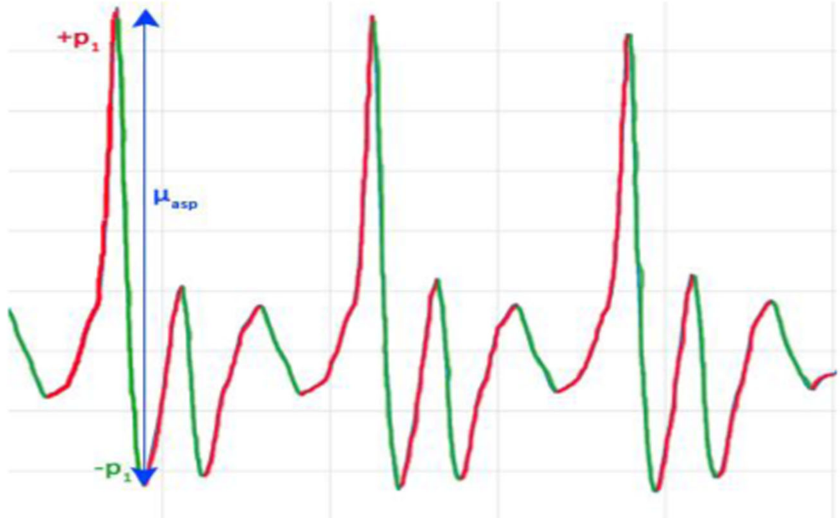



**Fig. 7** Positive negative Peak Analysis

$$COV = \sum_{i-1}^{n}\left(X_i - \overline{x}\right)\left(Y_i - \overline{y}\right) \quad (9)$$

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

## Temporal Analysis

**Peak Analysis** With the help of three significant features, the temporal analysis or peak analysis has been done in this study. 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. Figure 5 denotes computation of the temporal analysis. Figure 6 denotes the average peak interval, Peak detection, and peak histogram of swing phase Support phase, Stride Phase and stance phase time series signal.

**PNPH Analysis** General Histogram method merely represented by the following formula.

$$n = \sum_{i=1}^{k} m_i \quad (10)$$

In this function, $m_i$that counts the number of observations fall into each of the bins. Whereas $n$ is the total number of observations and $i$ is the total number of bins. The cumulative histogram $M_i$ of a histogram $m_i$ defined as

$$M_i = \sum_{j=1}^{i} m_i \quad (11)$$

Positive Negative Peak Histogram Analysis is a proposed method to analyze the multifaceted biomedical signal. After the decomposition, for $n^{th}$ the level of approximation coefficient, we identify the$\mu_{(asp)}$. Let $\mu_{(asp)}$ is the average value of the approximation coefficient between two successive positive and negative peaks are shown in the fig. 7.

With the help of$\mu_{(asp)}$a histogram was constructed with 9 bin feature vector size.

$$PNPH_{bin}\left(\mu_{asp}+1\right) = PNPH_{bin}\left(\mu_{asp}+1\right)+1 \quad (12)$$

## Classification stage

Several machine learning algorithms [24, 25] have been developed for classification purpose in recent years. Once the features are extracted for all the time series of the

**Table 1** Summary of demographics details for patients

| Group | Age (years) (mean ± SD) | Height (meters) (mean ± SD) | Weight (kg) (mean ± SD) | Gait speed (m/s) (mean ± SD) |
|---|---|---|---|---|
| Parkinson's disease | 64.1 ± 15.10 (Range,44–80) | 1.85 ± 0.13 | 78.4 ± 14.52 | 1.12 ± 0.150 |
| Huntington's disease | 41.6 ± 7.35 (Range,35–54) | 1.85 ± 0.08 | 80.5 ± 15.5 | 1.42 ± 0.310 |
| ALS patients | 51.9 ± 14.58 (Range,36–66) | 1.78 ± 0.06 | 86.55 ± 12.04 | 1.15 ± 0.2 |

**Table 2** Performance of Proposed approach for Control Vs ALS

| Methods | Sensitivity (%) | Specificity (%) | Accuracy (%) | Youden index (%) |
|---|---|---|---|---|
| Work [17] | 81.33 | 68.8 | 82.4 | 90.27 |
| Work [18] | 90.33 | 81.25 | 85.88 | 88.54 |
| Proposed | 93.33 | 86.67 | 90.95 | 92.34 |

**Table 3** Performance of Proposed approach for Control Vs PD

| Methods | Sensitivity (%) | Specificity (%) | Accuracy (%) | Youden index (%) |
|---|---|---|---|---|
| Work [17] | 81.33 | 81.32 | 92.54 | 88.10 |
| Work [18] | 88.11 | 85.56 | 86.88 | 89.22 |
| Proposed | 98.32 | 97.25 | 98.25 | 98.25 |

corresponding gait signal, and then fed to the classifier for classification. In our work, Support Vector Machine (SVM) and Random Forest (RF) classifiers are used. To prove the efficiency of the proposed method, we compare two recent methods [17, 18]. For the evaluation, four performance measures like as Sensitivity, Specificity, Accuracy and Youden index are used. These measures are defined as follows:

$$Sensitivity = \frac{TP}{TP + FN} \tag{13}$$

$$Specificity = \frac{TN}{TN + FP} \tag{14}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{15}$$

$$Youden \quad Index = Sensitivity + Specificity - 1 \tag{16}$$

Where $TP$ is the number of true positives, $FN$ is the number of false negatives, $TN$ is the number of true negatives and$FP$ is the number of false positives.

## Result and Discussion

Experiments are implemented using Matlab software and tested on an Intel Corei5 3.5 GHz computer with 4 GB RAM. Spatial and temporal along with fluctuation features are estimated for all healthy and NDDs subjects. The obtained features were evaluated using the means of 10-trials for 5-fold cross-validation with a Random Forest classifier.

The original data were first divided into five equal subsets, and one subset was tested using the classifier trained on the remaining nine subsets. This procedure was repeated until every subset had been used once for testing. The overall discrimination accuracy of the classifier is based on the average performance over the 10 runs of classifications. Here we evaluate our proposed method for the classification of NDDs using gait dynamics. Several experiments are conducted to test the ability of the proposed method. First, we evaluate each group of NDDs subjects against healthy control subjects. Then we evaluate all groups of NDDs against the healthy control subjects. Table 1 demonstrated the summary of demographics of various groups.

### Classification of ALS

In our dataset, there are13 subjects with ALS disease (numbered 'als1', 'als2',. . ., 'als13') and 16 healthy control subjects (numbered 'control1', 'control2',. .., 'control16') are taken as the first configuration for evaluation. The following Table 2 shows the performance of the proposed approach in terms of sensitivity, specificity, accuracy and Youden index.

### Classification of PD

Based on the subjects of Parkinson's disease, the data collected are considered in this study. From 16 healthy controls and

**Table 4** Performance of Proposed approach for Control Vs HD

| Methods | Sensitivity (%) | Specificity (%) | Accuracy (%) | Youden index (%) |
|---|---|---|---|---|
| Work [17] | 80.30 | 81.12 | 92.32 | 85.22 |
| Work [18] | 82.18 | 84.22 | 83.80 | 86.10 |
| Proposed | 100 | 85.24 | 91.25 | 96.10 |

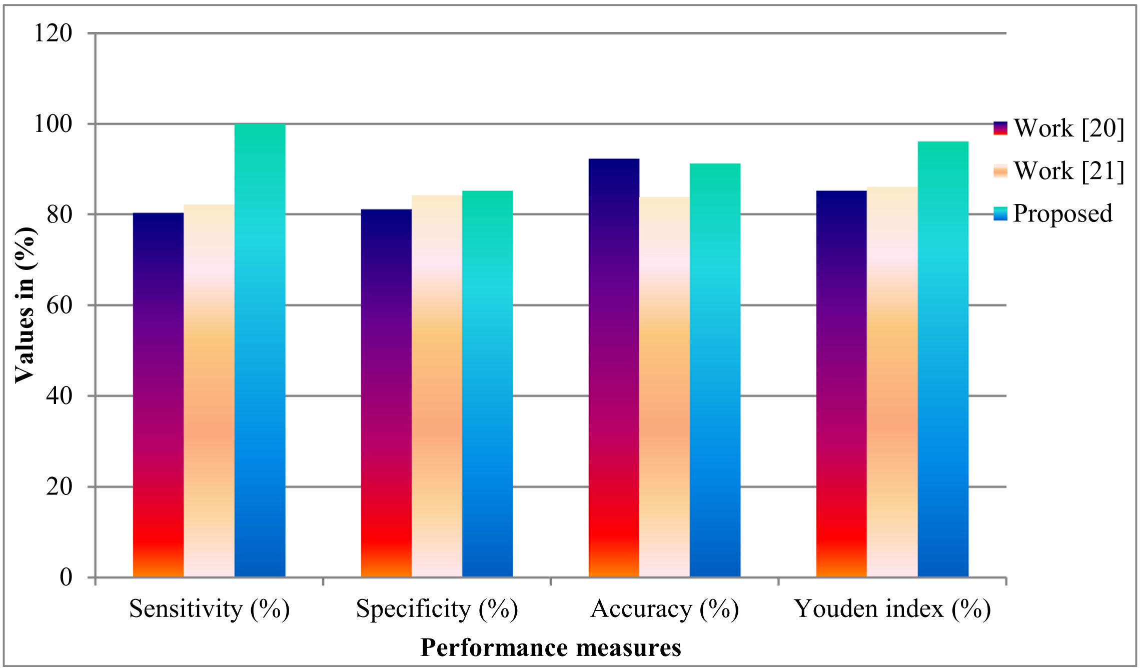



**Fig. 8** Classification performance of NDDs in terms of sensitivity, specificity, accuracy and youden index

15 PD subjects, the data are collected. The sensitivity, specificity, accuracy and Youden index for classification also measured in this case. The results have been shown in Table 3.

### Classification of HD

As explained above, the data are collected with 20 HD subjects. In this context, extracted gait dynamics from the subjects are used. In our database, evaluation is then performed against the healthy control subjects (16 cases) data. The results have been shown in Table 4.

## Classification of NDDs

In the last experiment, we put together the data from all three groups of NDDs including HD, PD and ALS. We evaluate the accuracy, sensitivity, specificity and youden index of the classification using the data from the healthy control subjects and all the NDDs. The results have been shown in Fig. 8. We obtained the accuracy of 91.25% and Youden index as 96.1% for the proposed method.

# Conclusion

This research proposes a new technique based on DWT in gait analysis to acquire triple features for the evaluation of Neuro-Degenerative Diseases distinction. The triple features are extracted from Force Sensitive Resistors (FSR) to provide essential benefaction with other implications for discrimination of NDD's. The obtained features are may use to develop different classification algorithms for a stable diagnosis of NDD's. Apart from that, the work presents a greater accuracy rate 98.99% for the discrimination of NDD's.

As a conclusion, the research provides a result to variations in the actual signal parameters due to conditions as exhaustion or complexity during long-term record obtained from the patients. In the future, by the usage of advanced artificial intelligence algorithms, the method can be developed to become a commercial device which is faster, highly accurate and more stable.

## Compliance with ethical standards

**Conflict of interest** The Authors and Co-Authors have no conflicts of Interests. The Paper is not submitted to any other Journals.

**Ethical approval (involving human participants and/or animals)** This article does not contain any studies with human participants or animals performed by any of the authors.

**Informed consent** All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

# References

1. Rabeh, A. B., et al. Diagnosis of Alzheimer Diseases in Early Step Using SVM (Support Vector Machine). Computer Graphics, Imaging and Visualization (CGiV) 364–367, 2016.
2. Bhagyashree, S. R., and Sheshadri, H. S., An approach in the diagnosis of Alzheimer disease—a survey. Int J Eng Trends Technol (IJETT) 7(1):41–43, 2014.

3. Shree, S. B., Sheshadri, H. S., Diagnosis of Alzheimer's Disease using Rule based Approach. Indian Journal of Science and Technology 9(13), 2016.
4. Shree, S. B., Sheshadri, H. S., An initial investigation in the diagnosis of Alzheimer's disease using various classification techniques. In Computational Intelligence and Computing Research (ICCIC) 1–5, 2014.
5. Diker, A., Avci, D., Avci, E., and Gedikpinar, M., A new technique for ECG signal classification genetic algorithm Wavelet Kernel extreme learning machine. Optik 180:46–55, 2019.
6. Joseph, S., et al. A Statistical and Biological Approach for identifying misdiagnosis of incipient Alzheimer patients Using Gene expression Data. In Engineering in Medicine and Biology Society 5854–5857, 2006.
7. Aziz, W., and Arif, M., Complexity analysis of stride interval time series by threshold dependent symbolic entropy. Eur. J. Appl. Physiol. 98(1):30–40, 2006.
8. Scafetta, N., Marchi, D., and West, B. J., Understanding the complexity of human gait dynamics. Chaos 19(2):026108, 2009.
9. Wu, Y. F., and Shi, L., Analysis of altered gait cycle duration in amyotrophic lateral sclerosis based on nonparametric probability density function estimation. Med. Eng. Phys. 33(3):347–355, 2011.
10. Wu, Y. F., and Krishnan, S., Computer-aided analysis of gait rhythm fluctuations in amyotrophic lateral sclerosis. Med. Biol. Eng. Comput. 47(11):1165–1171, 2009.
11. Daliri, M. R., Automatic diagnosis of neuro-degenerative diseases using gait dynamics. Measurement 45(7):1729–1734, 2012.
12. Sundararaj, V., An efficient threshold prediction scheme for wavelet based ECG signal noise reduction using variable step size firefly algorithm. *International Journal of Intelligent Engineering and Systems* 9(3):117–126, 2016.
13. Liao, F. Y., Wang, J., and He, P., Multi-resolution entropy analysis of gait symmetry in neurological degenerative diseases and amyotrophic lateral sclerosis. Med. Eng. Phys. 30(3):299–310, 2008.
14. Scafetta, N., Moon, R., and West, B. J., Fractal response of physiological signals to stress conditions, environmental changes and neurodegenerative diseases. Complexity 12(5):12–17, 2007.
15. Hausdorff, J. M., Mitchell, S. L., Firtion, R., Peng, C. K., Cudkowicz, M. E., Wei, J. Y., and Goldberger, A. L., Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease. J. Appl. Physiol. 82(1):262–269, 1997.
16. Hou, F. Z., Zhuang, J. J., Ning, X. B., Yang, X. D., and Huo, C. Y., Decrease in Hurst exponent of human gait with aging and neurodegenerative diseases. Chinese Phys. B 17(3):852–856, 2008.
17. Ren, P., Zhao, W., Zhao, Z., Bringas, M. L., Valdes-Sosa, P. A., Kendrick, K. M., Analysis of Gait Rhythm Fluctuations for Neurodegenerative Diseases by Phase Synchronization and Conditional Entropy, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 1534–4320, 2015.
18. Zeng, W., and Wang, C., Classification of neurodegenerative diseases using gait dynamics via deterministic learning. Information Sciences 317:246–258, 2015.
19. Balakrishnan, S., Janet, J., Sujatha, K., and Sheeba Rani, S., An efficient and complete automatic system for detecting of lung module. *Indian J. Sci. Technol.* 11(26):1–5, 2018.
20. Sujatha, K., and Shalini Punithavathani, D., Optimized ensemble decision-based multi-focus imagefusion using binary genetic Grey-Wolf optimizer in camera sensor networks. *Multimed. Tools Appl.* 77(2):1735–1759, 2018.
21. Agnieszka Kitlas, G., Detrended Fluctuation Analysis (DFA) in biomedical signal processing selected examples. Logic, Grammar and Rhetoric 29(42):2012, 2012.
22. Rodriguez, E., Echeverria, J. C., and Alvarez Ramirez, J., Detrended fluctuation analysis of heart interbeat dynamics. Physica A 384:429–438, 2007.
23. Madan, R., Kr. Singh, S., Jain, N., Signal Filtering Using Discrete WaveletTransform. International Journal of Recent Trends in Engineering, Vol 2, No. 3, 2009.
24. Sundararaj, V., Muthukumar, S., and Kumar, R. S., An optimal cluster formation based energy efficient dynamic scheduling hybrid MAC protocol for heavy traffic load in wireless sensor networks. *Comput. Secur.* 77:277–288, 2018.
25. Sundararaj, V., Optimal task assignment in mobile cloud computing by queue based ant-bee algorithm. Wirel. Pers. Commun. 104(1): 173–197, 2019.