Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 12 (2): 180-185, 2012
ISSN 1812-5654 / DOI: 10.3923/jas.2012.180.185

Parkinson Disease Gait Classification based on Machine Learning Approach

Nooritawati Md. Tahir and Hany Hazfiza Manap
Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM),
Shah Alam Selangor, Malaysia

Abstract: This study discussed the ability of two machine classifiers namely Artificial Neural Network (ANN) and Support Vector Machine (SVM) in distinguishing gait pattern during self-selected speed walking due to the effect of motor Parkinson Disease (PD). There are three gait parameters that is utilized as features in classifying PD gait and normal subjects namely basic spatiotemporal, kinematic and kinetic. Firstly, the input features are pre-processed using two types of normalization technique specifically intra group as well as inter group normalization. Additionally, all the three features are classified solely followed by implementation of data fusion. Then, the effectiveness of the features vectors to identify PD patients or vice versa is evaluated based as inputs to both classifiers. Initial findings showed that basic spatiotemporal solely as feature vectors based on intra group normalization technique contributed perfect classification for both ANN and SVM as classifiers.

Key words: Parkinson disease, gait recognition, artificial neural network, support vector machine

INTRODUCTION

Parkinson Disease (PD) is the degenerative illness of the brain that affects patients' ability to perform a well-learned motor task such as walking, body balancing as well as writing appropriately (Morris and Lansek, 1996). Basically, there are two types of PD known as motor and non-motor. Four primary symptoms categorized in motor PD includes tremor or also known as hand shaking; a characteristic appearance of 'pill-rolling' movement between thumb and forefinger. Next is bradykinesia due to slowness of movements and loss of involuntary movement. Thirdly is rigidity or limbs stiffness and finally is postural instability that caused patients to fall easily (Bock and Nausieda, 1999). On the other hand, the symptoms of non-motor PD are neuropsychiatric disturbance that is a common symptom which includes depression, anxiety and apathy, sleep disturbance, autonomic disturbance such as constipation, dizziness, drooling of saliva, sexual dysfunction and bladder problem. Apart of it, the gastrointestinal tract problem also occur in non-motor PD patients that includes nausea, choking and sensation of incomplete voiding of bowels, difficulty swallowing, severe constipation and inability to taste as well as sensory symptom. Numerous researches have been reported related to motor PD. For instance, Morris *et al*. (1999) has studied the constraints on the kinetic, kinematic and spatiotemporal parameter of gait in a 71 year old woman with PD. Results showed that PD has greatly influence the subject's gait by reducing the speed, stride length, total range of movement during walking and power generation at push off in ankle power profile. Moreover, statistical test conducted by Sofuwa *et al*. (2005) and Roiz *et al*. (2010) also proven that PD did affect patients' gait pattern in terms of basic spatiotemporal, kinematic and kinetic parameter. Due to these findings by various researches that stated motor PD caused noticeable deviation of patients' gait, therefore this study will deem further the application of Artificial Neural Network (ANN) and Support Vector Machine (SVM) as classifiers to recognize gait pattern of PD patients.

Conversely, the use of ANN in human gait analysis is widely explored. Firstly, Gioftsos and Grieve (1995) had successfully recognized 98% of the walking speed in a given condition and 99% for the vice versa case. Lafunte *et al*. (1997) used ANN to discriminate between healthy and lower limb arthritis patients recorded 80% accuracy. Further, Begg and Kamruzzaman (2006) applied ANN for detecting and classifying walking pattern changes due to ageing obtained maximum of 83.3%. Holzreiter and Kohle (1993) implemented ANN in order to distinguish the gait pattern between 'normal' and 'pathological' with accuracy approximately 95%. Inspired by Holzreiter and Kohle (1993) and Barton and Lee (1997) used the approach of kinematic analysis to differentiate the gait pattern for three conditions, specifically normal gait, simulations of leg length difference along with simulation of leg weight difference. From this study,

Corresponding Author: Nooritawati Md. Tahir, Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM), Shah Alam Selangor, Malaysia

83.33% of unknown gait pattern were successfully assigned into the right category and proven that ANN is capable to be applied in clinical practice for automated diagnosis of gait disorder.

On the other hand, SVM is another popular classifier too. For instance, Singla *et al.* (2011) concluded the superior performance of SVM over ANN classifer in detecting eye event from electroencephalographic (EEG) signal. Results obtained showed that SVM successfully classified 90.8% correct EEG pattern whilst ANN only gained 86.8% accuracy. Also, Wu and Liu (2011) implemented SVM and NN in finger-vein pattern identification and SVM showed good performance in recognizing finger-vein features. Another application of SVM in pattern classification was used by Roy and Bhattacharya (2005) for iris recognition and the performance of SVM as a classifier was proven far better than ANN. Thus, by referring to the performance of these two types of machine classifiers as reported by the previous researches, the classification ability of these classifiers will be evaluated to determine the most suitable classifier for PD gait pattern recognition.

MATERIALS AND METHODS

Gait data acquisition and gait features: Twelve PD patients and twenty healthy volunteers without any other diseases that could contribute to the additional effect on their body balance such as post-stroke or any additional muscular disease participated in this study. It is also confirmed that all subjects must also be able to walk freely without any cane or mechanical aid device during walking to eliminate the influence of aid device on the data acquired. Next, thirty seven reflective markers are adhered to the subjects' skin before the walking trials are preceded. This marker will be traced by infrared camera for kinematic analysis purpose. In addition, subjects are instructed to walk freely at their comfortable speed on two embedded force plate with sampling rate of 200 Hz. By walking on the force plate, Ground Reaction Forces (GRF) of the subjects are recorded. The three categories of GRF are lateral GRF (F_x), horizontal GRF (F_y) and vertical GRF (F_z). In this study, two types of significant GRF are chosen as the kinetic features specifically the horizontal and vertical GRF. The extraction of these two GRF based on the five significant walking events are as shown in Fig. 1.

As the subjects walk on the force plates, six infrared cameras are activated to record the walking event. These infrared cameras have the ability to trace the coordinate of reflective markers that is adhered to each subject's body. From the recording of walking event, all the relevant basic

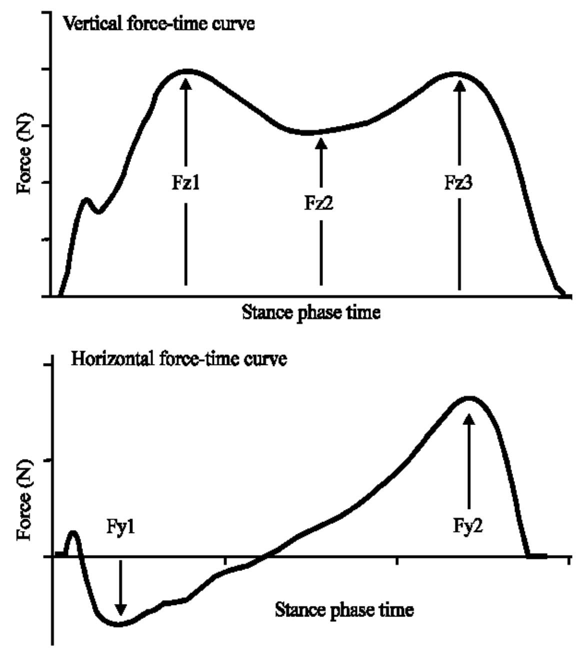

Fig. 1: The five kinetic features extracted from GRF reading

Table 1: Selected feature to be used as classifier inputs

Basic spatiotemporal	Kinematic	Kinetic
Stride time (sec)	Hip angle:	Maximum vertical heel contact force (Fz_1)
Cadence (steps min^{-1})	At heel strike	Vertical minimum mid-stance force (Fz_2)
Step length (m)	At toe off	Maximum vertical push off force (Fz_3)
Walking speed (m sec^{-1})	Maximum extension	Maximum horizontal heel strike force during braking phase (Fy_1)
	Maximum flexion	Maximum horizontal push-off force (Fy_2)
	Knee angle:	
	At heel strike	
	At toe off	
	Maximum extension	
	Maximum flexion	
	Ankle angle:	
	At heel strike	
	At toe off	
	Maximum extension	
	Maximum flexion	

spatiotemporal gait features namely walking time, speed, cadence and step length can be calculated. In addition, the angle of significant walking event also can be obtained from the coordinate of the reflective markers. All the input features from three gait parameter specifically basic, kinematic and kinetic with description are as listed in Table 1. Moreover, all subjects participated are instructed to complete three gait trials and the basic, kinetic and kinematic parameters are calculated based on the mean of the three trials. As a precaution, if the observer found that any subject purposely extended or shortened their normal stride in order to ensure contact with the force plates as required or their foot overlapped the edge of the other foot, the trial is eliminated and need to be repeated.

Normalization method: The use of data normalization often provided great improvement in pattern recognition. In this paper, there are three normalization methods implemented prior to classification which are:

GRF data normalization: Normalization of GRF data is obtained with GRF data acquired as dividend and body weight as divisor. This is to eliminate the influence of subjects' body weight to GRF reading which can be differ for each person regardless their health status:

$$\text{Normalized GRF} = \frac{\text{GRF (N)}}{\text{Body weight (N)}} \tag{1}$$

Intra group data normalization: Here, all input data recorded during gait recording experiments are mathematically normalized by dividing all data acquired with the maximum value of the respective group of both normal and PD patients:

$$X_{intra} = \frac{X_i}{X_{max\ (Intra\ class)}} \tag{2}$$

Inter group data normalization: As for inter group normalization, input data are divided by the highest value amongst both normal and PD group:

$$X_{intra} = \frac{X_i}{X_{max\ (Inter\ class)}} \tag{3}$$

Artificial Neural Network (ANN): The architecture of the ANN used in this study is the multilayer feed-forward network with Back Propagation (BP) algorithm. BP algorithm is a generalization of the least mean squared algorithm that minimizes the mean squared error between the desired and actual output of the network by modifying its network weight (Mehrotra *et al.*, 1997). As BP algorithm used supervised learning, the network is trained using the known data of inputs and desired output. After the data is trained, the network weights are frozen to be used later for new input samples to compute an output values (Mehrotra *et al.*, 1997).

The proposed ANN configuration in this study consisted of an input layer with number of input neurons corresponding to the gait feature vectors that acted as inputs to the ANN. Additionally, one hidden layer with eleven neurons and one output layer with a single neuron indicate normal and abnormal gait pattern. Weight are adjusted by employing Lavenberg-Marquardt as the training algorithm and the learning rate are set at 0.3. Lavenberg-Marquardt is chosen as the training algorithm due to fast training and its ability to minimize the error function (Abraham, 2005). Then, tanh-sigmoidal activation function (tansig) is applied for both hidden and output layer.

Support Vector Machine (SVM): SVM which has been introduced by Vapnik is an excellent learning technique that is developed based on the framework of structural risk minimization (SRM) and the VC bounds theory cited by Kumar and Gopal (2011). In a binary classification task, SVM will map the input space to a high-dimensional feature space based on the selected kernel function. Then, an Optimal Separating Hyperplane (OSH) that separates the two classes by maximizing the margin between the closest point of the classes are constructed, as illustrated in Fig. 2. In this study, three type of kernel function are evaluated during classification namely linear, radial basis function and polynomial. The mathematical formula

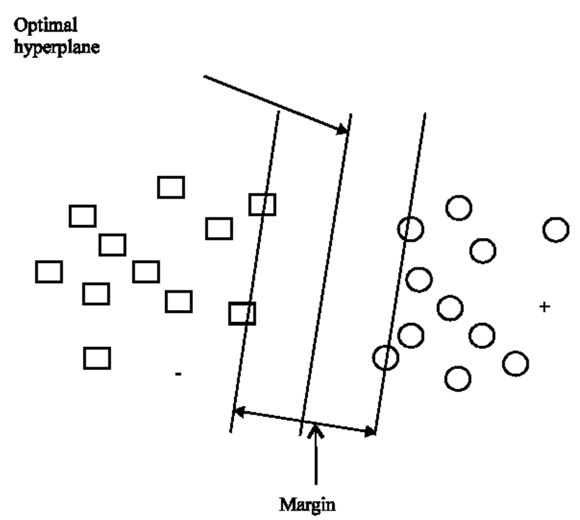

Fig. 2: Example of SVM binary classification with OSH and maximum margin

Table 2: Three SVM kernel functions used in classification task and its mathematical formula

Kernel function	Mathematical formula
Linear	$K(x_i, x_j) = \langle x_i, x_j \rangle$
Radial basic function	$K(x_i, x_j) = \exp(-\lvert x-z\rvert^2 / 2\sigma^2)$, σ is the with of the function
Polynomial	$K(x_i, x_j) = (\langle x_i, x_j \rangle + 1)^d$, d is the degree of polynomial

for the entire kernel function are as given in Table 2 (Kumar and Gopal, 2011).

Cross-validation and classification method: Cross-validation is a statistical method of dividing data into two segments, which are training and testing set. Since the PD subjects in this study is relatively small, cross-validation method are employed to overcome the constraint, inspired from technique used by Barton and Lee (1997). Data set obtained from both category that generated total of 32 subjects are divided into 4 subsets. Initially, one subset will be used as training set, while the remainder are assigned as unseen or testing data. The process repeated until all subsets are utilized as testing data.

Conversely, input data from each gait category, namely basic spatiotemporal, kinematic and kinetic are classified solely and the classifier performance are recorded. Then, gait features from two parameters are fused or combined to evaluate the effectiveness of these features in classifying gait of PD. This classification task are repeated for both ANN and SVM classifier based on intra group and inter group normalization method too. By performing these two classification mode, potential of each gait parameter can be evaluated in identifying the most significant gait features that contributed for recognition of PD gait.

RESULT AND DISCUSSION

The overall performances of both classifiers based on the two technique of data normalization are as tabulated in Table 3 and 4. Additionally, three statistical indices are computed specifically accuracy, sensitivity and specificity using the computed results from True Acceptance Rate (TAR), False Rejection Rate (FRR), True Rejection Rate (TRR) and False Acceptance Rate (FAR).

The detail of TAR, FRR, TRR and FAR are as below:

- TAR: Classifiers identify PD as PD
- FRR: Classifiers identify PD as normal
- TRR: Classifiers identify normal as normal
- FAR: Classifiers identify normal as PD
- Sensitivity: Classifiers ability to identify PD as PD
- Specificity: Classifiers ability to identify normal as normal

$$\text{Accuracy} = \frac{\text{TAR} + \text{TRR}}{\text{TAR} + \text{TRR} + \text{FRR} + \text{FAR}} \quad (4)$$

$$\text{Sensitivety} = \frac{\text{TAR}}{\text{TAR} + \text{FRR}} \quad (5)$$

$$\text{Sensitivety} = \frac{\text{TRR}}{\text{TRR} + \text{FAR}} \quad (6)$$

As observed in Table 3, based on intra group normalization technique, both ANN as well as SVM attained perfect accuracy rate based on four features that is extracted from the basic spatiotemporal category namely stride time, cadence, step length and walking speed. Moreover, basic spatiotemporal data also recorded perfect sensitivity and specificity rate that indicated the classifiers ability to identify PD subjects accurately. On the other hand, it is realized that kinematic parameter that comprised of twelve feature vectors extracted from hip angle, knee angle as well as ankle angle during heel strike, toe off, along with maximum extension and maximum flexion contributed worst recognition rate based on both ANN and SVM as classifiers. It is also observed that the sensitiveness of both classifiers deteriorated to correctly detect PD gait features. In addition, results attained also verified that kinematic features either solely as inputs or based on data fusion are unsuitable as classification inputs due to insignificant or less deviation of joint angle parameters during walking between the PD subjects and normal group. Apart from basic spatiotemporal and kinematic features, kinetic gait parameter is also vital in gait recognition. ANN classifier obtained classification rate of 90.6% for kinetic parameter as feature inputs as in Table 3. On the whole, recognition accuracy as well as sensitivity rate with kinetic parameters as feature inputs is higher as compared to kinematic parameters for both ANN and SVM classifiers. Hence, it is proven that basic spatiotemporal solely acted as the most suitable gait features for classification of PD gait. From the findings of this study, it is also suggested for future researches to only consider basic and kinetic gait parameter as input features for recognition of PD gait.

Nevertheless, it is observed too that intra group normalization technique contributed higher accuracy rate. For instance, SVM with polynomial kernel contributed as the most significant kernel amongst the three kernels as well as ANN as classifier with 95.8% accuracy rate for kinetic parameters based on intra group normalization and the accuracy rate reduced for inter group normalization category. As in Table 4, the classification rates of both classifiers worsened and similar achievement of recognition rate is observed too irrespective of the feature inputs. Therefore, this finding confirmed that inter group normalization technique is unsuitable for PD gait pattern recognition conducted in this research.

Table 3: Results based on intra group normalization data

	ANN			SVM (Linear)			SVM (RBF)			SVM (Polynomial)		
Gait features	Acc	Sens	Spec	Acc	Sens	Spec	Acc	Sens	Spec	Acc	Sens	Spec
Basic	100	100	100	100	100	100	100	100	100	100	100	100
Kinematic	78.1	66.7	85	68.3	33.3	89.2	68.5	18.1	98.8	65.9	9.03	100
Kinetic	90.6	86.7	95	88.3	86.1	89.6	90.1	82.5	94.6	95.8	88.9	100
Fusion of basic and kinematic	90.6	83.3	95	90.9	95.1	88.3	91.7	79.9	98.8	84.3	58.3	100
Fusion of basic and Kinetic	96.9	91.7	100	82.8	76.4	86.7	98.2	95.1	100	84.5	58.3	100
Fusion of kinematic and kinetic	81.3	66.7	90	80.2	61.1	91.7	85.2	62.5	98.8	77.9	40.9	100
Fusion of all three features	84.4	75	90	93.2	95.8	91.7	93.5	84.7	93.8	93.2	95.8	90

Table 4: Results based on inter group normalization data

	ANN			SVM (Linear)			SVM (RBF)			SVM (Polynomial)		
Gait features	Acc	Sens	Spec	Acc	Sens	Spec	Acc	Sens	Spec	Acc	Sens	Spec
Basic	78.1	79.2	80.8	68.3	42.4	84.2	75.6	77.1	76.3	82.3	54.9	98.8
Kinematic	68.6	52.1	76.3	66.4	47.9	77.5	62.5	35.4	93.3	70.1	31.3	93.3
Kinetic	71.9	77.1	69.6	68.5	20.1	97.5	75.5	70.8	78.3	59.1	22.9	80.8
Fusion of basic and kinematic	71.9	60.8	78.3	73.2	72.2	73.8	81.5	52.8	98.8	80.7	48.6	100
Fusion of basic and kinetic	71.9	72.5	73.3	74.2	40.1	94.2	80.2	52.1	97.1	78.9	48.6	97.1
Fusion of kinematic and kinetic	71.9	69.7	77.5	58.3	26.4	77.5	72.4	39.6	92.1	72.7	27.1	100
Fusion of all three features	76.5	65	77.5	70.1	66.7	72.8	81.8	53.5	98.8	79.2	44.4	100

SVM with RBF as kernel performed well in distinguishing gait pattern based on data fusion of two gait parameters as showed in Table 3 and 4 with fusion of basic spatiotemporal and kinetic gait parameter achieved highest accuracy that is 98.2%. The increased in accuracy rate for SVM upon fusion of input features indicated the incapability of SVM as classifier based on small number of input features. This confirmed the lower accuracy of SVM using basic, kinematic or kinetic as inputs solely. Overall, for data fusion, the classifiers performances are enhanced based on fusion of three parameters as compared to data fusion of two parameters. Finally, SVM with polynomial kernel attained highest specificity that confirmed the classifer ability not to generate false detection (normal healthy gait pattern) rather than the ability to accurately recognize the pathological gait pattern.

CONCLUSION

In conclusion, SVM as classifier performed better as compared to ANN specifically for data fusion of gait parameters. However, both classifiers attained higher classification rate based on intra- group normalization method instead of inter group normalization irrespective of data fusion or otherwise. Finally, initial results proven that basic spatiotemporal contributed as the best feature inputs based on perfect accuracy, specificity and sensitivity rate attained and the feature vectors can be utilized to classify pathological gait of PD.

ACKNOWLEDGMENT

This study was supported by Ministry of Science, Technology and Innovative (MOSTI) Malaysia under the E-Science Grant No: 06-01-01-SF0314.

REFERENCES

Abraham, A., 2005. Artificial Neural Network. In: Handbook of Measuring System Design, Sydenham, P.H. and R. Thorn (Eds.). John Wiley and Sons, Ltd, Unites States, pp: 901-908.

Barton, J.G. and A. Lee, 1997. An application of neural networks for distinguishing gait patterns on the basis of hip-knee joint angle diagrams. Gait Posture, 5: 28-33.

Begg, R. and J. Kamruzzaman, 2006. Neural networks for detection and classification walking pattern change due to ageing. Aust. Phys. Eng. Sci. Med., 29: 188-195.

Bock, G. and P. Nausieda, 1999. Parkinson's Disease: What You and Your Family Should Know. The National Parkinson Foundation Inc., USA.

Gioftsos, G. and D.W. Grieve, 1995. The use of neural networks to recognize pattern of human movement. J. Clin. Biomech., 10: 179-183.

Holzreiter, S.H. and M.E. Kohle, 1993. Assessment of gait patterns using neural networks. J. Biomech., 26: 645-651.

Kumar, M.A. and M. Gopal, 2011. Reduced one-against-all method for multiclass SVM classification. Expert Syst. Appli., 38: 14238-14248.

Lafunte, R., J.M. Belda, J. Sanchez-Lacuesta, C. Soler and J. Prat, 1997. Design and test of neural networks and statistical classifiers in computer-aided movement analysis: A case study on gait analysis. Clin. Biomech., 13: 216-229.

Mehrotra, K., C.K. Mohan and S. Ranka, 1997. Elements of Artificial Neural Networks. The MIT Press, Boston, pp: 344.

Morris, M.E. and R. Iansek, 1996. Characteristics of motor distrubance in Parkinson's disease and strategies for movement rehabilitation. Hum. Movement Sci., 15: 649-669.

Morris, M.E., J. McGinle, F. Huxham, J. Collier and R. Iansek, 1999. Constraints on the kinetic, kinematic and spatiotemporal parameters of gait in Parkinson's disease. Hum. Movement Sci., 18: 461-483.

Roiz, R.M., E.W. Chacho, M.M. Pazinatto, J.G. Reis, A. Jr.Cliquet and E.M. Barasnevicius-Quagliato, 2010. Gait analysis comparing Parkinson's disease with healthy elderly subjects. Arq Neuropsiquiatr, 68: 81-86.

Roy, K. and P. Bhattacharya, 2005. Iris Recognition with Support Vector Machines. Springer-Verlag Berlin Heidelberg, Germany, pp: 486-492.

Singla, R., B. Chambayil, A. Khosla and J. Santosh, 2011. Comparison of SVM and ANN for classification of eye events in EEG. J. Biomed. Sci. Eng., 4: 62-69.

Sofuwa, O., A. Nieuboer, K. Desloovere, A. Willems, F. Chavret and I. Jonkers, 2005. Quantitative gait analysis in Parkinson's disease: Comparison with a healthy control group. J. Arch Phys. Med. Rehabil., 86: 1007-1013.

Wu, J.D. and C.T. Liu, 2011. Finger-vein pattern identification using SVM and neural network technique. Expert Syst. Appl., 38: 14284-14289.