33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Automatic Classification of Pathological Gait Patterns using Ground Reaction Forces and Machine Learning Algorithms

Murad Alaqtash, Thompson Sarkodie-Gyan, Huiying Yu, Olac Fuentes, Richard Brower, and Amr Abdelgawad

Abstract—**An automated gait classification method is developed in this study, which can be applied to analysis and to classify pathological gait patterns using 3D ground reaction force (GRFs) data. The study involved the discrimination of gait patterns of healthy, cerebral palsy (CP) and multiple sclerosis subjects. The acquired 3D GRFs data were categorized into three groups. Two different algorithms were used to extract the gait features; the GRFs parameters and the discrete wavelet transform (DWT), respectively. Nearest neighbor classifier (NNC) and artificial neural networks (ANN) were also investigated for the classification of gait features in this study. Furthermore, different feature sets were formed using a combination of the 3D GRFs components (mediolateral, anterioposterior, and vertical) and their various impacts on the acquired results were evaluated. The best leave-one-out (LOO) classification accuracy 85% was achieved. The results showed some improvement through the application of a features selection algorithm based on M-shaped value of vertical force and the statistical test ANOVA of mediolateral and anterioposterior forces. The optimal feature set of six features enhanced the accuracy to 95%. This work can provide an automated gait classification tool that may be useful to the clinician in the diagnosis and identification of pathological gait impairments.**

I. Introduction

GAIT analysis has become a crucial assessment tool in clinics and hospitals. It provides new insights to help understand various human movement patterns corresponding to different gait pathology. The traditional gait assessment methods are based on clinical observations. However, it has the drawbacks of being qualitative, whereby decisions were based on the subjective judgment of clinicians, and also on time-consuming processes. Current work in gait analysis intends to develop an automated gait assessment tool to manipulate the massive kinetic and kinematic gait data in a real-time and provide an objective decision. Hence, the gait classification tool may assist clinicians in diagnosis and in identification of pathological gait impairments.

Several studies have investigated different techniques for the representation and classification of data for the analysis of gait. Such techniques include: statistical analysis, mathematically transforms, and machine learning classifiers, e.g. support vector machines (SVM), nearest neighbor classifier (NNC), and artificial neural networks (ANN). The statistical methods are the most widely applied and understood for gait analysis. The most common adopted methods to represent kinetic data (dynamic data) in gait are the statistical methods [1] - [5]. This technique was performed in which the representative features were selected based on the peak levels of the GRFs and their corresponding time of occurrence. This approach has the benefit of selecting the most significant data points and hence reducing the massive data extracted from the gait kinetics (dynamics) in 3D. However, this approach is sensitive to noise and it is under the consideration that all subjects' data within the same class exhibit similar patterns and have similar features. Other studies investigated the mathematically transforms such as Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT) [6] - [8]. In these methods, the transforms compute approximation coefficients that are used as feature vectors to represent gait data. This leads to massive data reduction without the need to a selection process of specific data points. Several machine learning classifiers were investigated for the automatic recognition of gait patterns including: SVM [1], [9], and [2]; NNC [6] and [10]; ANN [7], and [11] - [13].

The existing techniques used for gait data have been evaluated in several reviews [14] - [18]. Preece *et al.* [18] investigated the different classification algorithms which have been used to classify normal activities to identify falls from body-worn sensor data. Although, a few investigated comparisons between different techniques suggested that ANN may give the highest classification accuracy, the differences are often small. SVM have shown promise in small pilot studies but have yet to be tested in large scale studies. Furthermore, recent work in combining different classifiers has been examined in the review. It has been found that the performance was improved by combining the output of different classifiers. ANN was found to be the most prevalent non-traditional methodology for gait analysis

M. Alaqtash is with the Electrical and Computer Engineering Department, The University of Texas at El Paso, El Paso, TX 79968, USA (corresponding author to provide phone: +1-915-747-5541; fax: +1-915-747-7871; e-mail: msalaqtash@miners.utep.edu).
T. Sarkodie-Gyan is with the Electrical and Computer Engineering Department, The University of Texas at El Paso, El Paso, TX 79968, USA (e-mail: tsarkodi@utep.edu).
H. Yu is with the Electrical and Computer Engineering Department, The University of Texas at El Paso, El Paso, TX 79968, USA (e-mail: hyu@miners.utep.edu).
O. Fuentes is with the Computer Science Department, The University of Texas at El Paso, El Paso, TX 79968, USA (e-mail: ofuentes@utep.edu).
R. Brower is with the Neurology, Texas Tech University Health Science Center, El Paso, TX 79905, USA (e-mail: richard.brower@ttuhsc.edu).
A. Abdelgawad is with the Orthopedic Surgery & Rehabilitation, Texas Tech University Health Science Center, El Paso, TX 79905, USA (e-mail: amr.abdelgawad@ttuhsc.edu).

978-1-4244-4122-8/11/$26.00 ©2011 IEEE

recently [14], and [16]. However, the performance of ANN depends on internal learning parameters, weights and biases, which are difficult to interpret. Dobson *et al.* [17] presented a systematic review examining 18 studies for gait classification in children with cerebral palsy (CP). Half of the studies used qualitative pattern recognition method, whereas the remaining studies used quantitative method based on cluster analysis techniques. The study found that most classifications were constructed using only sagittal plan gait data, only one study used gait data in all three plans, which may limit classifications validity and restrict the applications.

Although, many techniques have been investigated for gait classification, the fact that different gait pathologies have different gait patterns imposes that more investigation and work are necessary in this field. The aim of this study is to develop an automated gait classification tool, which can be used in gait analysis application to classify pathological gait patterns, using 3D GRFs data and by utilizing different machine learning algorithms. Particularly, the proposed method was tested and evaluated using three groups of subjects: healthy, cerebral palsy (CP), and multiple sclerosis (MS). CP and MS are neurological diseases causing gait impairments.

II. Methodology

A. Participants

This study was approved by the Institutional Review Boards (IRB) of the University of Texas at El Paso (UTEP), and all subjects signed an informed consent form prior to participation. Twenty subjects were recruited for the purpose of the study: Twelve healthy adults (mean± std: age 27.1±5.9 years, weight 71.4±11.5 kg, height 171.6±8.3 cm, and body mass index (BMI) 24.2±2.8 kg/m^2), four spastic diplegic cerebral palsy patients (age 29.5±17.5 years, weight 67.8±17.9 kg, height 162.3±8.7 cm, and BMI 25.4±4.2 kg/m^2), and four relapsing remitting multiple sclerosis patients (age 50.3±11.5 years, weight 99.7±32.5 kg, height 167.8±14.5 cm, and BMI 34.5±5.8 kg/m^2).

Although the database is rather modest which may limit the generalization of extracted features and then affect the classifications accuracy, this work represents a preliminary study for investigation different machine learning algorithms that may be used for classification of pathological gait impairments.

B. Data Acquisition

An instrumented treadmill was used in this study (Bertec Corporation, USA). The instrumented treadmill is dual-belt type with two independent force plates mounted beneath the belts. Each force plate measures the GRFs in three directions (F_X: mediolateral, F_Y: anterioposterior, and F_Z: vertical) while subject walking at a self-selected natural speed for three minutes continuously. The GRFs data were acquired at frequency of 100 Hz and filtered using 20 Hz second order Butterworth low pass filter. The GRFs amplitudes were normalized based on body mass [1], [5], and [19]. The stride time was specified based on F_Z and time-normalization was accomplished by re-sampling and expressing each stride in percentage rather than time [20].

C. Feature Extraction

Two different feature extraction approaches have been evaluated in this study.

1) Approach I - GRFs Parameters: A common approach to represent gait data is based on the amplitudes and temporal parameters of data [5]. The maximum and minimum GRF amplitudes and their time of occurrence, in percentage, were computed for the three GRFs components. A total of 19 features were extracted to represent each subject as shown in Fig.1. Within the three minutes walking, the average value of each feature was calculated across all strides. Moreover, left and right values were transformed into a common mode value represent the average value for both sides, this done to avoid feature dependences on the laterality of the disease [5].

2) Approach II - Discrete Wavelet Transform (DWT): Wavelets are basis functions obtained from a prototype called mother function and it can be used to approximate data. The DWT compute the approximation coefficients using digital filtering to represent the feature vector for GRFs data. A mother wavelet function *Symlet3* can be used to calculate the feature vector for GRFs data [6]. In this work, a mother wavelet *Symlet3* and a feature vector of length 16 coefficients were determined experimentally.

D. Classification

1) Nearest Neighbor classifier (NNC):

The principle of nearest neighbor or *k*-nearest-neighbor algorithm is that the properties of any particular input point *x* are likely to be similar to those of points in the neighborhood of *x* [21]. The neighborhood is defined to include *k* points, and a distance metric, e.g. Euclidean distance, is used to identify the nearest neighbors of a query point. Nearest neighbor algorithm is very simple to implement and often performs quit well [21]. The typical value of *k* varies from 1 to a small percentage of the training data and is selected using trial and error [18].

2) Artificial Neural Network (ANN):

ANN is among the most effective learning methods currently known for learning to interpret complex real-world sensor data [22]. One of the most common ANNs is a multi-layer feedforward neural network [23]. A multi-layer feedforward neural network has been a standard method for various applications including gait analysis [16]. It consists of interconnected set of simple nodes, neurons, distributed in a hidden layer or layers. The inputs are mapped to the nodes through an input layer and the outputs are controlled by transfer function within each node and by adjusting the weights of links between nodes. The learning process in neural networks consists of adjusting the weights through training in which the actual output approximately matches the desired output. Although ANNs can have multiple hidden layers, research suggests that there should be one hidden layer with a number of nodes less than input nodes [14]. In this study, a multi-layer feedforward neural network of one hidden layer has been used.

3) Classification Evaluation and Performance Measures: Cross validation is a common method to evaluate the accuracy of classifiers [24]. The classifier is trained with training dataset which includes most of the subjects and then tested with testing dataset which has few subjects. In Leave-One-Out (LOO) cross validation, one subject is used for testing and the rest are used for training. The classification result is then computed and repeated until all subjects have participated in the testing dataset. The overall classification result is then computed as the average of all testing subjects.

A typical measure of classification algorithms performance is a confusion matrix [25]. An example of a confusion matrix with two classes, positive and negative, has four elements: *TP*, true positive, is the number of correctly classified positive instances; *FN*, false negative, is the number of positive instances incorrectly classified as negative; *TN*, true negative, is the number of correctly classified negative instances; and *FP*, false positive, is the number of negative instances incorrectly classified as positive. The classification accuracy is defined as the ratio of correctly classified instances to all instances.

$$Accuracy = (TP+TN) / (TP+FN+TN+FP)$$

Precision and recall are another typical classification performance measures defined as [26]:

$$Precision = TP / (TP+FP)$$

$$Recall = TP / (TP+FN)$$

Precision is the ratio of correctly classified positive instances to both correctly and incorrectly classified instances as positive. Recall is the ratio of correctly classified positive instances to positive instances.

III. Results And Discussion

A. Feature Sets

In order to evaluate the effect of each GRFs component on the classification accuracy, different feature sets were formed using a combination of the three GRFs components. Seven feature sets (F_Z; F_Y; F_X; F_XF_Y; F_YF_Z; F_XF_Z; $F_XF_YF_Z$) were examined.

A summary of LOO classification accuracy using different approaches and classifiers is presented in Table I. The highest achieved classification accuracy 85% corresponded to: $F_XF_YF_Z$ using NNC and approach I (GRFs parameters); the same set of features using ANN and approach II (DWT); F_XF_Y using ANN and approach I. The results on a single force component indicate that F_X is the most discriminatory and F_Z is the worst.

In general, the feature extraction approach I, GRFs parameters, depicts better accuracy than DWT associated with NNC, whereas the result varies in the case of ANN. Both NNC and ANN have the same range of classification accuracy among all feature sets.

B. Features Selection

In view of previous results, further investigation has been made to enhance the results by applying effective features selection techniques. In order to select the most significant features, GRFs parameters, it is important to investigate the contribution of each feature set to classification for each class of gait patterns: healthy, CP, and MS. Fig.2 shows the LOO classification accuracy for all classes using NNC.

Alaqtash *et al.* [27] recorded that F_Z of MS patients shows a significant longer stance phase. Moreover, the contact force was relatively flat with a single peak force as shown in the same force curve, whereas the same curve of healthy subjects exhibits the M shape with two peaks to reflect the weight transfer from the heel to the mid-foot and to the ball of the foot for push-off.

Although F_Z provides full discrimination accuracy between healthy and pathological gait pattern, F_Z leads to lower classification accuracy for both pathological gaits CP and MS as shown in Fig.2. Therefore, another representation of F_Z has been investigated based on the M shape [3]. F_Z force was represented as either *M-shaped* or *non M-shaped.* *M-shaped* was defined as both ratios F_{Z2} / F_{Z1} and F_{Z2} / F_{Z3} are less than 0.9; others were defined as *non M-shaped*. The number 0.9 was determined experimentally.

As a result, using an input feature set consists of the complete F_X and F_Y features sets and one feature represents F_Z (1: *M-shaped*, 0: *non M-shaped*), the classification outcomes are recorded in Table II. The result shows 100% for precision and recall measures of healthy subjects. This result demonstrated that full discrimination between healthy and pathology gait can be achieved using one feature, *M-shaped*, represents F_Z in addition to the complete F_X and F_Y features sets. Moreover, the confusion matrix depicts that all misclassified instances are located within the pathological gait classes. Therefore, further enhancement can be made if a certain features selection technique applied for F_X and F_Y sets.

The statistical significant difference test, analysis of variance (*ANOVA*), was employed in this study to evaluate the effect of each feature on the difference between CP and MS classes. The *p-value* depicts the significant difference between CP and MS classes corresponding to a certain feature. The *p-value* was calculated for each feature in F_X and F_Y sets and therefore features were sorted based on *p-value*. Fig.3 demonstrates the features selection order and the corresponding classification accuracy. It can be observed that the best LOO classification accuracy 95% was achieved using the optimal feature set of six features; one feature from F_Z : *M-shaped*; three from F_X : F_{X1}, T_{X3}, and F_{X3}; two from F_Y: T_{Y3} and T_{Y2}. The classification performance measures: confusion matrix, precision, and recall, using the optimal feature set are shown in Table III. The results clearly indicate that the classification performance significantly increased using the optimal feature for the three classes. The overall performance measures are 96% and 95% for precision and recall, respectively.

IV. CONCLUSION

In this study, an automated gait classification tool was developed. A classification of healthy, CP, and MS gait patterns, was performed using the three components of GRFs data (F_X, F_Y, and F_Z) and by utilizing different machine learning algorithms. The proposed method was tested and evaluated using three groups of subjects: healthy, CP, and MS. Two feature extraction algorithms (GRFs parameters and DWT) and two classifiers (NNC and ANN) have been investigated and the result was evaluated. Further enhancement on classification accuracy has been made applying a features selection algorithm based on *M-shaped* value of F_Z and the statistical test ANOVA of F_X and F_Y. The best LOO classification accuracy 85% was achieved without features selection. An improvement of the result was achieved using features selection. The accomplished classification accuracy was 95% corresponding to the optimal set of six features.

This study has demonstrated some significant points: 1) a comparison between two different feature extraction approaches depicts that GRFs parameters approach associated with NNC has higher classification accuracy than DWT; 2) both NNC and ANN have the same range of classification accuracy among all feature sets; 3) full discrimination between healthy and pathology gait can be achieved using one feature, *M-shaped*, represents F_Z in addition to the complete F_X and F_Y features sets; and 4) a features selection method can be used to enhance the classification accuracy and the statistical test *ANOVA* is a possible technique for that.

This work can provide an automated gait classification tool that may be useful to the clinician in the diagnosis and identification of pathological gait impairments.

ACKNOWLEDGMENT

The authors wish to thank the Stern Foundation for providing the funds for this research work. The authors also would like to thank all the subjects who participated in this study.

REFERENCES

[1] D. Lai, P. Levinger, R. Begg, W. Gilleard, and M. Palaniswami, "Automatic recognition of gait patterns exhibiting patellofemoral pain syndrome using a support vector machine approach", *IEEE Trans. Information Technology in biomedicine,* vol. 13, issue 5, pp. 810-817, 2009.

[2] R. Begg, and J. Kamruzzaman, "A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data", *Journal of Biomechanics,* vol. 38, pp. 401-408, 2005.

[3] T. Takahashi, K. Ishida, D. Hirose, Y. Nagano, K. Okumiya, M. Nishinaga, Y. Doi, and H. Yamamoto, "Vertical ground reaction force shape is associated with gait parameters, timed up and go, and functional reach in elderly females", *J. Rehabil. Med*., vol. 36, pp. 42-45, 2004.

[4] H. Gok, S. Ergin, and G. Yavuzer, "Kinetic and kinematic characteristics of gait in patients with medial knee arthrosis," *Acta Orthopaedica,* vol. 73, issue 6, pp. 647-652, 2002.

[5] R. Lafuente, J.M. Belda, J. Sanches-Lacuesta, C. Soler, and J. Prat, "Design and test of neural networks and statistical classifiers in computer aided movement analysis: a case study on gait analysis", *Clinical Biomechincs,* vol. 13, issue 3, pp. 216-229, Nov. 1997.

[6] N. Mezghani, S. Husse, K. Boivin, K. Turcot, R. Aissaoui, N. Hagemeister, and J. de Guise, "Automatic classification of asymptomatic and osteoarthritis knee gait patterns using kinematic data features and the nearest neighbor classifier", *IEEE Trans. Biomedical Engineering*, vol. 55, issue 3, pp.1230-1232, March 2008.

[7] M. Kohle, and D. Merkl, "Analyzing Human Gait Patterns for Malfunction Detection". *In Proc. 2000 ACM symposium on Applied computing* 1, ACM, New York, NY, 2000, pp. 41-45.

[8] A. R. Ismail, and S. S. Asfour, "Discrete wavelet transform: a tool in smoothing kinematic data", *Journal of Biomechanics,* vol. 32, pp. 317-321, 1999.

[9] R. K. Begg, M. Palaniswami, and B. Owen, "Support vector machines for automated gait classification", *IEEE Trans. Biomedical Engineering*, vol. 52, issue 5, pp. 828-838, May 2005.

[10] L. Wang, T. Tan, W. Hu, and H. Ning, "Automatic gait recognition based on statistical shape analysis", *IEEE Trans. Image Processing,* vol. 12, issue 9, pp. 1120-1131, Sept. 2003.

[11] M. A. Hanson, H. C. Powell, A. T. Barth, J. Lach, and M. Brandt-Pearce, "Neural network gait classification for on-body inertial sensors". *Sixth International Workshop on Wearable and Implantable Body Sensor Networks,* Berkeley, CA, June 2009, pp. 181-186.

[12] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, "Activity classification using realistic data from wearable sensors", *IEEE Trans. Information Technology in Biomedicine*, vol. 10, issue 1, pp. 119-128, 2006.

[13] W. L. Wu, F. C. Su, Y. M. Cheng, and Y. L. Chou, "Potential of the genetic algorithm neural network in the assessment of gait patterns in ankle arthrodesis", *Ann. Biomed. Eng.*, vol. 29, issue 1, pp. 83-91, Jan. 2001.

[14] S. R. Simon, "Quantification of human motion: gait analysis--benefits and limitations to its application to clinical problems", *Journal of Biomechanics,* vol. 37, issue 12, pp. 1869-1880, Dec. 2004.

[15] T. Chau, "A review of analytical techniques for gait data. Part 1: fuzzy, statistical and fractal methods", *Gait & Posture,* vol. 13, pp. 49-66, 2001.

[16] T. Chau, "A review of analytical techniques for gait data. Part 2: neural network and wavelet methods", *Gait & Posture,* vol. 13, pp. 102-120, 2001.

[17] F. Dobson, M. Morris, R. Baker, and H. Graham, "Gait classification in children with cerebral palsy: A systematic review", *Gait & Posture,* vol. 25, pp. 140-152, 2007.

[18] S. J. Preece, J. Y. Goulermas, L. Kenney, D. Howard, K. Meijer, and R. Crompton, "Activity identification using body-mounted sensors—a review of classification techniques", *Physiol. Meas.*, vol. 30, pp. 1-33, 2009.

[19] L. Jonesa, M. Beynonb, C. Holta, and S. Royc, "An application of the Dempster–Shafer theory of evidence to the classification of knee function and detection of improvement due to total knee replacement surgery", *Journal of Biomechanics,* vol. 39, pp. 2512–2520, 2006.

[20] D. A. Winter, *Biomechanics and Motor control of Human Movement*. 4th ed., Hoboken, New Jersey: John Wiley & sons, 2009.

[21] S. Russell, and P. Norvig, *Artificial Intelligence a modern approach*. 2nd ed., New Jersey: Prentice-Hall, 2003.

[22] T. M. Mitchell, *Machine Learning.* 1st ed., McGraw-Hill, 1997.

[23] S. Haykin, *Neural Networks; a Comprehensive Foundation*. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[24] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. 2nd ed., New York: Wiley, 2001.

[25] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique", *Journal of Artificial Intelligence Research,* vol. 16, pp. 321-357, 2002.

[26] D. L. Olson, and D. Delen, *Advanced Data Mining Techniques*. Springer, 2008.

[27] M. Alaqtash, H. Yu, R. Brower, A. Abdelgawad, and T. Sarkodie-Gyan, "Application of wearable sensors for human gait analysis using fuzzy computational algorithm," *Engineering Applications of Artificial Intelligence*, In Press, Available online May 2011.

TABLE I
SUMMARY OF CLASSIFICATION ACCURACY USING THE DIFFERENT APPROACHES AND CLASSIFIERS

Classifier	Features	F_Z	F_Y	F_X	F_X,F_Y	F_Y,F_Z	F_X,F_Z	F_X,F_Y,F_Z
NNC	Approach I [a]	65%	70%	80%	75%	75%	75%	85%
	Approach II [b]	65%	65%	70%	60%	65%	75%	60%
ANN	Approach I	70%	70%	75%	85%	75%	70%	80%
	Approach II	60%	75%	80%	80%	70%	70%	85%

[a] GRFs parameters [b] DWT

TABLE II
THE CLASSIFICATION PERFORMANCE MEASURES USING *M-SHAPED*, F_X AND F_Y FEATURES

(a) CONFUSION MATRIX

	CP	MS	Healthy
CP	**2**	2	0
MS	1	**3**	0
Healthy	0	0	**12**

(b) PRECISION AND RECALL

	Precision	Recall
CP	75%	50 %
MS	60%	75%
Healthy	1	1

TABLE III
THE CLASSIFICATION PERFORMANCE MEASURES USING THE OPTIMAL FEATURE SET

(a) CONFUSION MATRIX

	CP	MS	Healthy
CP	**3**	1	0
MS	0	**4**	0
Healthy	0	0	**12**

(b) PRECISION AND RECALL

	Precision	Recall
CP	100%	75%
MS	80%	100%
Healthy	100%	100%
Weighted Average [a]	96%	95%

[a] Average of 12 healthy, 4 CP, and 4 MS.

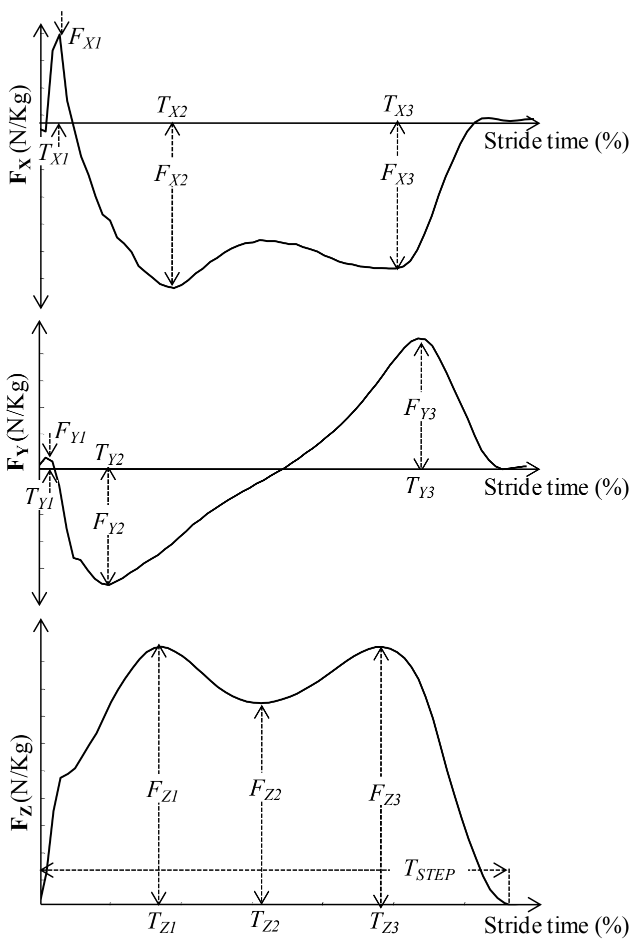

Fig. 1. GRFs parameters.

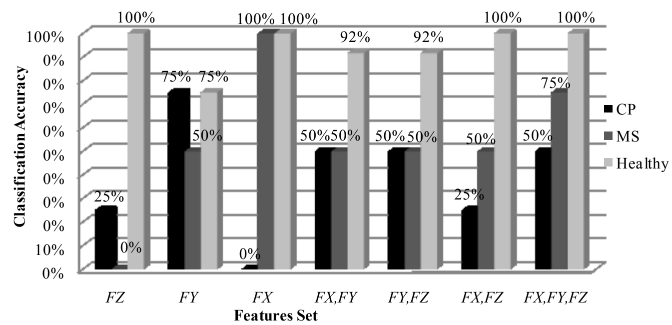

Fig. 2. Classification accuracy for all classes using NNC.

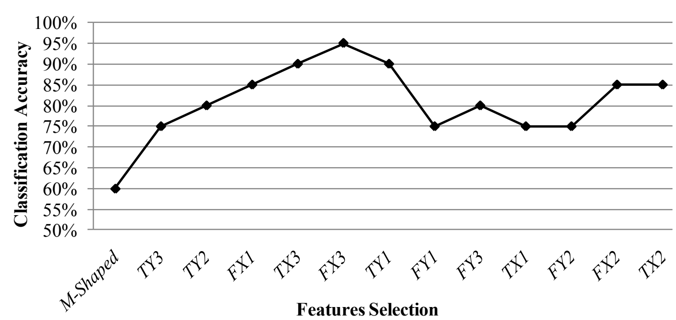

Fig. 3. Features selection order and the corresponding classification accuracy.