Scientific
Research
Publishing

# Gait-Ground Reaction Force Sensors Selection Based on ROC Curve Evaluation

**Rami Alkhatib[1,2], Mohamad Diab[2], Bassam Moslem[2], Christophe Corbier[1], Mohamed El Badaoui[1]**

[1]Laboratoire LASPI, Université Jean-Monnet, Roanne, France
[2]Electrical and Computer Engineering Department, College of Engineering, Rafik Hariri University, Damour, Lebanon
Email: rami.alkhatib@universite-lyon.fr, khatibrh@rhu.edu.lb

## Abstract

Classification of normal gait from pathological gait as then can be used as indicator of falling among subjects requires the correct choice of sensor location in the insole. Such a flexi force- sensor can be used underneath foot to measure vertical ground reaction force. To start with, the most relevant information (parameters) that can characterize the recorded signals are extracted from the vertical ground reaction force signals. Then Receiver Operating Characteristic curve is used to evaluate the features upon 8 sensors underneath each foot located at different locations. To confirm results obtained, features are passed upon a chosen classifier, in this paper K-nearest neighbors algorithm is chosen. Results show that the sensor located at the inner arch of the sole of the foot (*i.e.* at the mid foot) holds the most relevant information needed for better classification compared to other sensors.

## Keywords

## 1. Introduction

Gait Ground reaction force (GRF) is widely used and modeled by many scientist [1] in order to extract valuable information. For example, it has been used for the purpose of differentiating between normal and Parkinson's subjects [2].

However, still many studies consider the sensor location to measure the GRF either at both toe and heel or usually by analyzing the total force from sensors underneath each foot [2]. Though, a need to investigate more

exploration of choosing the correct position of a sensor or set of sensors in the insole becomes crucial.

This is for the reason that some features are used for classification and better work when they are extracted from sensors data at a given location. In this paper, KNN classifier is used randomly as an example to verify the concept and not to test its power in classification. The selection of feature relayed on the features used in a number of papers for classification persistence such as skewness, kurtosis, median frequency and many more.

ROC curve is then used to evaluate each feature for all sensors. This could contribute to other research conducting in the goal of anticipating the risk of falling among people and especially among elderly [3].

In this paper the location of each sensor is studied independently in order to identify the sensor in the best location and which reveals the most relevant information needed for better classification compared to other sensor's data signals.

## 2. Gait Data—Vertical Ground Reaction Force (VGRF)

### 2.1. Database Description

VGRF database is obtained from PhysioNet [4]. A collection of signal measures VGRF (in Newton) as a function of time extracted from 8 sensors (Ultraflex Computer DuynoGraphy, Infotronic Inc.) underneath each of the right and left foot. In addition, two signals representing the summation of each of the 8 signals recorded from each foot are taken as a reference.

Sensor locations are shown in **Figure 1** inside the insole as lying approximately at the following (X, Y) coordinates measured as a person is comfortably standing with both legs parallel to each other. The origin (0, 0) is just between the legs and the person is facing towards the positive side of the Y axis.

Two groups of persons were recorded: the normal case, also named control and they count for 18 persons, and 29 patients with Parkinson disease. Each participant walked for two minutes at their own natural pace and with acquisition sampling rate of 100 Hz. **Figure 2** shows the force generated by eight sensors underneath right and left foot in addition to their total for both normal and Parkinson diseased person of age 72 years old.

### 2.2. Preprocessing

Before utilizing data for any purpose, preprocessing must be performed on the data to remove any undesirable characteristics that produced during acquisition. For instance, Filtering is used in cleaning and removing any unwanted disturbance in GRF data. That's because the presence of noise can totally mask the true information in data. In addition, it's significant to eliminate sources of variation on the measured VGRF like the influence of mediolateral and anterior-posterior variations. For illustration, Butterworth low pass filter of second order can be used for that purpose [5] [6]. Such a filter is used to filter the data with cut-off frequency of 30 Hz. This choose is due to studying the mean frequency of high oscillations denoted as Intrinsic mode function given by empirical mode decomposition of each signal.

In the other side, if one variable is to be normalized to another variable, it is important to understand the relation between them [3]. That's why, data normalized bynorm-2 as to make the data comparable to each other.
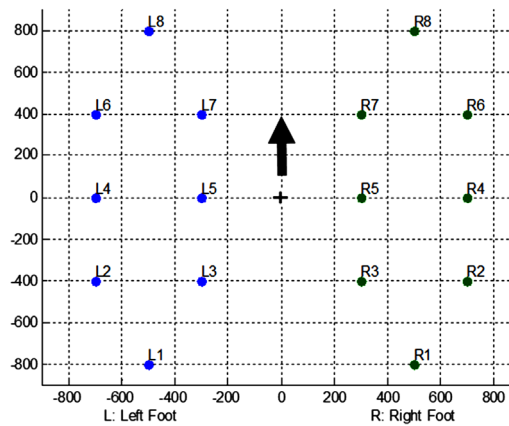


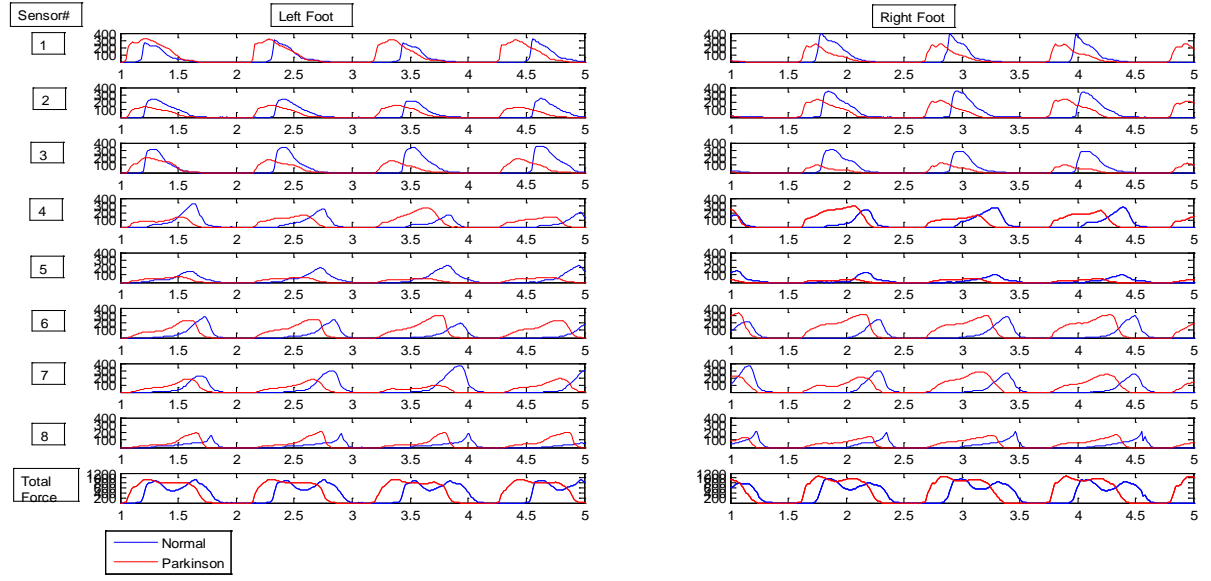**Figure 1.** Sensors location in both right and left insoles.

**Figure 2.** GRF signals from eight sensors underneath right and left foot for normal and Parkinson subject.

## 3. ROC Curve Overview

Plotting sensitivity or true positive rate (TPR) against fall-out (1-specificity) or false positive rate (FPR) exemplifies the performance of binary classifier. This is designated by ROC curve. Their calculation is given in equation one and two.

$$TPR = TP/(TP + FN) \tag{1}$$

$$TPR = FP/(FP + TN) \tag{2}$$

where *T* refers to true whenever the prediction matches the actual situation. Therefore, *TP* is true positive, *TN* is true negative, FP is false positive and *FN* is false negative. Thus, with 95% confidence interval, the area under the curve (AUC) will reflect the accuracy on how well a feature could well separate normal gait from Parkinson's gait. The following scores could be used to evaluate the accuracy:

0.9 - 1: Excellent
0.8 - 0.9: Good
0.7 - 0.8: Fair
0.6 - 0.7: Poor
0.5 - 0.6: Fail

As a result AUC = 1 refers to a perfect discrimination and has a ROC curve that passes through the upper left corner *i.e.* 100% sensitivity and 100% specificity with no overlap in the two distributions. **Figure 3** show a plot of the ROC curve for skewness feature tested over sensor 5 in the right foot among normal and Parkinson. Each point on the ROC curve represents the sensitivity and specificity pair corresponding to a particular decision threshold. The diagonal line dividing the ROC space is also called line of no-discrimination in which a point on this line corresponds to a completely guess. When points are above the diagonal, this indicates a good classification results and on the other side points below this line indicate a poor predictors. Therefore the distance from the random guess line is the best indicator of how much predictive power a method has.

The evaluation of skewness shows the following:

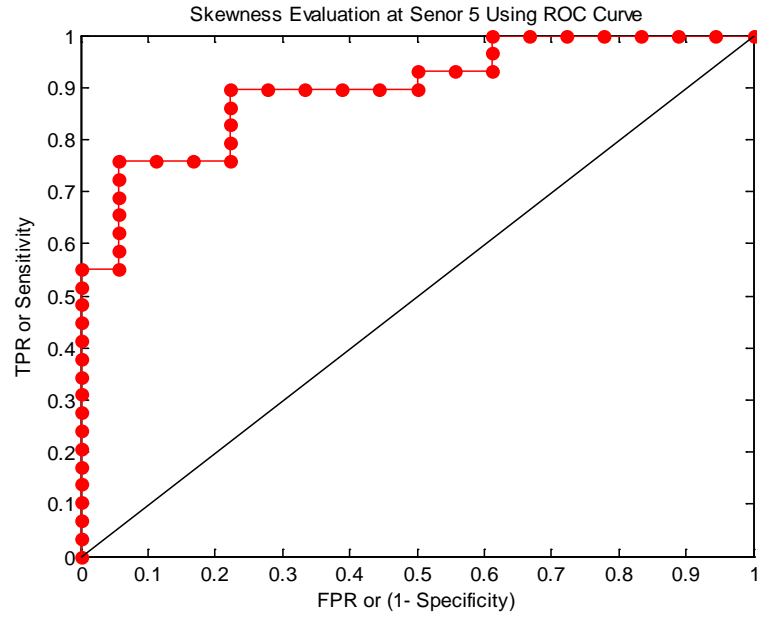| | |
|---|---|
| **ACC** | 85.1064 |
| **MCC** | 0.6822 |
| **Sensitivity** | 0.8966 |
| **Specificity** | 0.7778 |
| **Area Under Curve** | 0.902 |

**Figure 3.** Skewness ability in discrimination between Normal and Parkinson Gait using ROC curve.

Using the above listed data, 0.9 as area under the curve indicates an excellent performance of skewness in discriminating normal from Parkinson using VGRF from sensor 5.

## 4. Methodology

### 4.1. Feature Extraction

In this part, simple features were extracted after the data being pre-processed. Then their performance was tested. The features used here are used from literature, we found 23 features. In this paper the nine most relevant features were retained. A list of the commonly wide used features are shown below:

- Mean: Signal averaging
- Median: numerical value separating the higher half of a data from the lower half.
- standard deviation: measures the amount of variation or dispersion from the average
- range: the difference between the maximum and minimum values
- Interquartile range: robust estimate for the spread of the data being equal to the difference between the upper and lower quartiles IQR = $Q_3 - Q_1$.
- 95% percentiles of the distribution of the signal. A percentile (or a centile) is a measure used in statistics indicating the value below which a given percentage of observations in a group of observations fall.
- Skewness: measure of lack of symmetry
- Kurtosis: measure of whether the data are peaked or flat relative to a normal distribution.
- Power of the signal
- Mean power frequency
- Magnitude of peak frequency

### 4.2. Features Evaluation by ROC

As too many statistical features could be extracted and evaluated in time domain analysis and frequency domain, in this section one feature is used to demonstrate the evaluation. However this is done for all features among all sensors for the 47 subjects. Let's take for instance the skewness as in many studies related to vertical ground reaction force [2] shows its capability in distinguishing between normal and Parkinson's diseased person.

The results of ROC evaluation among all the 47 subjects are shown in **Table 1** for each sensor.

Analyzing **Table 1** indicates thatunlike other studies similarto [2] that consider total summation of force sig-

nals from all sensors as the most important, however, its clearly shown that sensor 5 is the most important sensor (AUC = 0.9) to consider in building acquisition system to acquire data for analyses. **Figure 4** shows the ROC curve of the total ground reaction force from sensors of the right foot. The sensitivity is recorded to be 0.6552 while the specificity is 0.5556. This yield AUC = 0.5460 which refers a fail level of accuracy in classification.

This conclusion is generalized as the same procedure is applied over the rest of the features chosen in this study. If expanding data is needed, then adding sensor 7 and 6 corresponds as main sensors also to be considered for classification as shown in **Figure 5**.

Not to add, the average of the strides of the 2 classes corresponding to the 120 seconds of walking is also considered. As a result, each series of strides is represented by its average, that is, one stride. Next, 3 features were extracted: the amplitude of the first peak, that is, the peak that corresponds to the heel contact (in case of total force), time to the first peak, stride time. The ROC evaluation also infers a better accuracy for sensor 5 compared to other sensors.

**Table 1.** ROC evaluation of skewness among all sensors.

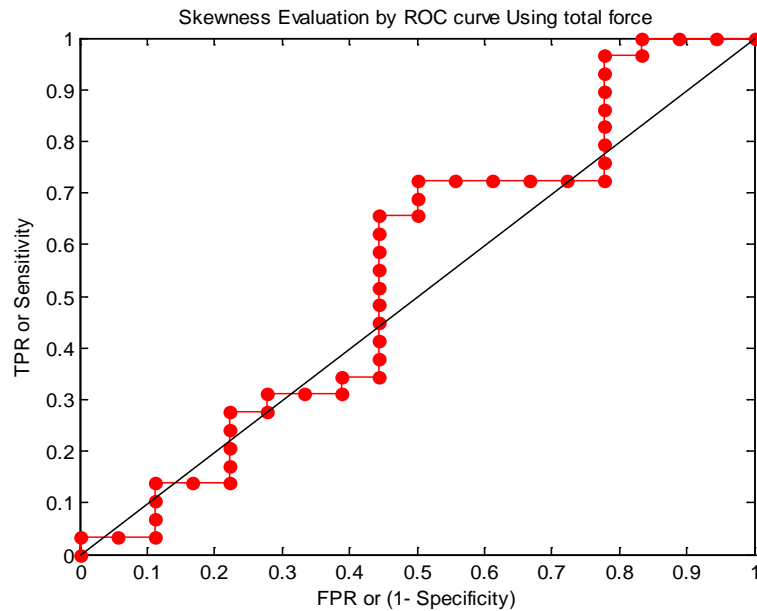| Sensor # in Foot | ROC-AUC | |
| --- | --- | --- |
| | Left Foot | Right Foot |
| 1 | 0.645594 | 0.655172 |
| 2 | 0.609195 | 0.637931 |
| 3 | 0.609195 | 0.613027 |
| 4 | 0.701149 | 0.676245 |
| 5 | 0.904215 | 0.842912 |
| 6 | 0.787356 | 0.764368 |
| 7 | 0.850575 | 0.808429 |
| 8 | 0.62069 | 0.703065 |
| VGRF corresponds to summation of the 8 sensors | 0.605364 | 0.611111 |



**Figure 4.** Skewness is extracted from the total force from sensors located under the right foot and its performance in binary classification is evaluated using ROC curve.
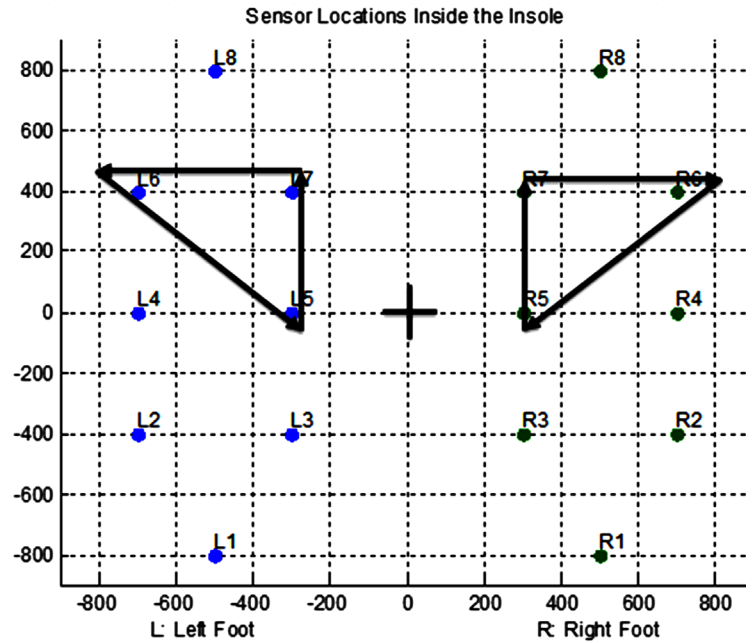
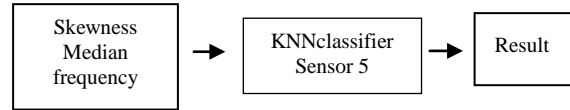**Figure 5.** The most important locations to acquire data for Gait analysis.



**Figure 6.** Skewness and Median Frequency are extracted from signals of sensor 5 in mid-foot.

## 4.3. Verification

In order to verify results, the features are passed through a chosen classifier, k-nearest neighbors (KNN) in this case. As mentioned, KNN is not used to test its power in classification, but to test the power of sensor 5 in a given classifier. Ten subjects from each class are chosen as training and the rest are tested by the classifier.

In this study, this is done in two ways. First, select one sensor among all subjects and then choose two features randomly and iterate between them. The feature chosen will have a high score from ROC evaluation and then feed them to KNN-classifier one example is shown in **Figure 6**.

In a second case, fix the feature and iterate a number of sensors among the KNN classifier. The results of KNN classifier indicate an accuracy of around 83% on average in most cases where sensor five exist. Other sensor shows a relatively smaller value. While the total force when used shows an accuracy of around 15% smaller than sensor five.

## 5. Conclusion and Future Research

More attention to which sensor is chosen must be made. This is especially recommended when building an acquisition system. This study shows the sensor located at the inner arch of the sole of the foot (*i.e.* at the mid foot) near the axis of the center of body holds the most important information given certain features for classification. This could help more in using such sensor location to model walking by 3D-link dynamics as one foot is in contact with ground while the other is in swing phase. The classification of the VGRF between normal and Parkinson patient using classical feature is something to pay attention as features are chosen. In addition, more advanced parameters (features) should be extracted from the signals .Conventionally, a classifier, such as a neural network with rbf or an SVM will be used in later studies where the input of the classifier will be more than one feature prior to their use. Therefore, higher classification accuracy is expected to be obtained. Classification of the 8 signals individually from one foot then using a fusion technique may increase the classification accuracy.

## References

[1]  Sabri, K., El Badaoui, M., Guillet, F., Belli, A., Millet, G. and Morin, J. (2010) Cyclostationary Modeling of Ground Reaction Force Signals. *Signal Processing*, **90**, 1146-1152.

[2]  Alkhatib, R., Diab, M., Moslem, B., Corbier, C. and ElBadaoui, M. (2014) Nonstationary Analysis of Gait-Ground Reaction Force Signals. 20*th Anual LAAS Conference*, AlHadath, Lebanon.

[3]  Beauchet, O., Allali, G., Berrut, G., Hommet, C., Dubost, V. and Assal, F. (2008) Gait Analysis in Demented Subjects: Interests and Perspectives. *Neuropsychiatric Disease and Treatment*, **4**, 155-160. http://dx.doi.org/10.2147/NDT.S2070

[4]  Jeffry, M. and Research, F. (2005) Gait in Parkinson's Disease. National Institutes of Health, National Parkinson Foundation, and the Parkinson's Disease Foundation. http://www.physionet.org/

[5]  Worobets, J.T. and Stefanyshyn, D.J. (2005) Normalizing Vertical Ground Reaction Force Peals to Body Weight in Heel-Toe Running. Human Performance Laboratory, Faculty of Kinesiology, University of Calgary.

[6]  Minamisawa, T., Sawahata, H., Takakura, K. and Yamaguchi, T. (2012) Characteristics of Temporal Fluctuation of the Vertical Ground Reaction Force during Quiet Stance in Parkinson's Disease. *Gait & Posture*, **35**, 308-311. http://dx.doi.org/10.1016/j.gaitpost.2011.09.106