# Assessment of Gait Patterns of Chronic Low Back Pain Patients: A Smart Mobile Phone based Approach

Herman Chan
School of Computing and Mathematics, Ulster University
Jordanstown, Newtownabbey, Co.Antrim, BT37 0QB, UK
Chan-H@email.ulster.ac.uk

Huiru Zheng, Haiying Wang
Computer Science Research Institute, Ulster University
Jordanstown, Newtownabbey, Co.Antrim, BT37 0QB, UK
{hy.wang,h.zheng}@ulster.ac.uk

Dave Newell
Anglo-European College of Chiropractic
Bournemouth, UK
DNewell@aecc.ac.uk

***Abstract*—Chronic low back pain is a common and costly condition and has been shown to affect gait. This paper describes the use of gait analysis as measured by a smart phone in a group of chronic low back pain subjects. Reliability of features extracted from the smart phone sensors was investigated using a mutual information based minimum redundancy and maximum relevance feature selection method to identify a key feature set related to lower back pain. This analysis was carried out using a KStar classification model. Results indicate the feasibility of reducing gait features to 6 key components while still achieving very promising classification accuracy (92.50%). The results also demonstrated that it is feasible to use a smart mobile phone in gait tele-monitoring and tele-assessment suggesting potential as both a prognostic and potential treatment outcome. In addition, we show that predicting context such as age and gender using smart mobile phones is achievable, which has potential to provide personalised services and context-related monitoring and intervention.**

***Keywords*—Gait analysis, intraclass correlation coefficient feature selection, mutual information, smart phone, context awareness.**

## I. INTRODUCTION

Gait is a pattern that describes the movement of walking. Research shows that walking patterns can change as a result of certain health conditions, including serious neurodegenerative disease but also common functional conditions such as lower back pain [1], [2]. In addition the gait of an individual may start to deteriorate with age with elderly subjects above the age of 60 more likely to experience a decline in walking performance. Recent years have witnessed a growing interest in using gait analysis for rehabilitation, falls prevention and monitoring the progress of disease and chronic illness.

There are different fundamentals to analysing gait such as measuring temporal-spatial parameters, kinematics, centre of mass and whole body energetics, joint kinetics, electromyography (EMG) and joint power [3]. Existing technologies used in gait analysis are common and include measuring electric current produced by muscle contractions (EMG) and the use of video cameras, optoelectronic systems and force mats or plates [4]. However, these systems are laboratory based and may not be as suitable for the development of gait tele-monitoring and tele-assessment systems.

Emerging inertial sensor based systems have proven to be capable of obtaining the kinematics of a user's movement [5] and providing a low-cost alternative to laboratory-based motion capturing systems. In recent years, smart mobile phone manufacturers have adapted micro technologies such as micro electromechanical systems (MEMS) accelerometers and gyroscopes to determine device orientation. Potentially, the MEMS accelerometers in these devices can provide meaningful data concerning a user's movement and may be useful for gait analysis, assessment and monitoring in replacing the conventional piezo-electric crystal based accelerometers which are considered to be large and clumsy. Compared with standalone accelerometers, smart mobile phones are increasingly ubiquitous, low cost and in most cases, provide an intuitive touch screen interface increasing user accessibility. Yang et al. [6] explored the feasibility of using a smart mobile phone with built-in accelerometer in gait detection and analysis demonstrating the potential for using mobile phones in long term gait monitoring and assessment of the elderly.

The aim of this study was to explore the capabilities of the accelerometer within a smart mobile device, namely Apple's iPhone, for identification of gait patterns associated with chronic lower back pain (cLBP) patients. Arendt-Nielsen et al. [7] found that cLBP can change the motor performance during gait. Other authors have also investigated changes in gait in cLBP subjects and have suggested gait parameters could be used to monitor the progress of treatment effects in these subjects [8][9]. In this paper, we apply machine learning techniques to gait analysis for people with lower back pain (LBP) wearing a smart mobile phone. The following questions will be addressed: (1) Are features extracted from smart mobile phone accelerometers reliable for gait analysis?; (2) Can gait patterns associated with LBP be recognized using machine learning approaches? (3) How many features are needed in order to differentiate between LBP patients and normal subjects? and (4) Can context such as age and gender can be predicted using smart mobile? The remainder of this paper introduces the framework of a gait analysis system and the description of a methodology in Section II, followed by results and discussion in Section III. This paper concludes with a summary and suggestions for future work in Section IV.

## II. Methodology

The infrastructure used in the study for the gait tele-monitoring and assessment of cLBP patients is illustrated in Fig.1, which consists of the following three main components:

- Mobile-based Data collection
- Cloud System with data analysis
- Web user interface access for end-users, care-givers and doctors.

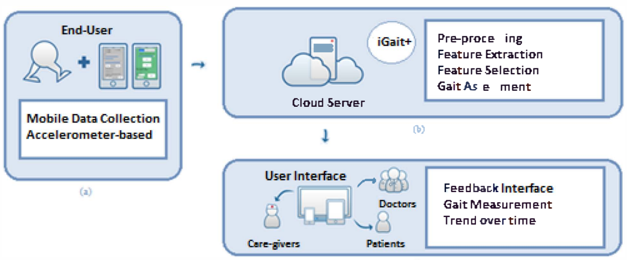


Figure 1. An illustration of the infrastructure used for the gait tele-monitoring and assessment of cLBP patients

This paper is focused on the first two elements of this system, i.e. data collection and analysis. Gait data associated with cLBP patients are collected by the accelerometer embedded in a smart mobile phone. The information processing is presented by a server-side computer where pre-processing of data, feature extraction and selection, and data mining is performed. The data are firstly preprocessed by an in-house gait analysis system, iGait+, which is an improved version of iGait [10]. Reliability is then tested for reproducibility of features derived from using smart mobile phones sensors to determine whether the features extracted from smart mobile phone accelerometers are valid compared to those from a tri-axial accelerometer. Those considered as reliable features are then ranked for feature selections. Finally the derived optimal set of features is used to assess gait patterns using classification techniques.

### A. *Mobile Application for Data Collection*

A mobile application for both iPhone and Android devices has been developed to collect user's gait information. An example of the user interface is illustrated in Fig. 2. The user presses the 'Start' button to begin recording the acceleration data and 'stop' when the recording is complete. The data captured is stored in a comma separated format to represent time, x, y and z-axis and can be emailed to the server. A cloud server was implemented to receive emails from the mobile application. The server interprets the CSV formatted data and stores the data in a secure database.

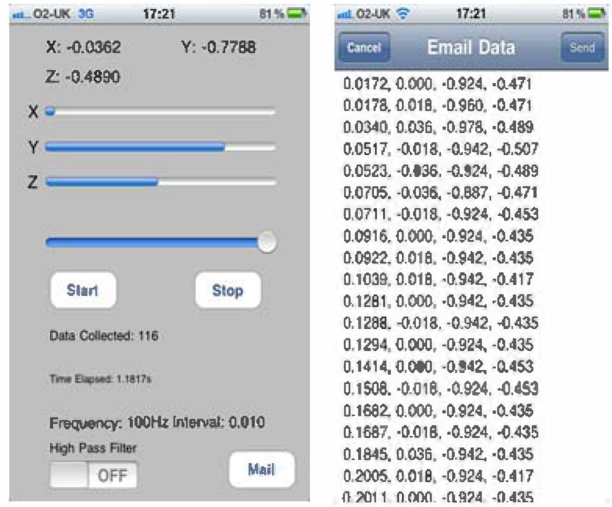


Figure 2 User Interface presented in the left screen and format of data captured on the right.

In order to validate the smart mobile phone data, a benchmark device was selected that records a different type of data as a means of validating the results of the iPhone acceleration data. The smart insoles provided by Tomorrow Options WalkinSense [11] were used as the benchmark device, which can detect gait events at the point of toe contact and heel contact during gait. Gait events recorded are used to validate data captured using the smart mobile phone accelerometer. A healthy male subject in his early 20's was asked to walk at a distance of 10 meters on a flat surface with an iPhone whilst also wearing trainers with smart insoles. In order to synchronize the two devices, a number of predefined motion tasks are performed before walking, such as a single step, jumping and standing stationary for a period of time. These motions will change accelerations in certain directions, and can be used as indicators for identifying the start of a task to assist synchronizing the acceleration data and pressure data.

#### 1) *Data Collection*

Patients with LBP attending a chiropractic teaching clinic in the UK (AECC) and triaged as having non-specific low back pain were recruited over the course of 6 weeks during late October to December of 2011. Twenty control subjects matched for age and gender with no existing LBP or walking difficulties were also recruited to perform the same tasks. In total, forty subjects participated and completed the study. Age and gender information is summarised in Table I. An iPhone 4 with custom software for collecting accelerometer data from the device was attached to a belt along with the Minimod device, which was used as a benchmark device in the feasibility study of using a smart mobile phone for extracting gait features. Subjects were asked to walk a short distance of 50 meters on a flat surface in a sports hall at their preferred walking speed [12].

TABLE I PARTICIPANTS AGE AND GENDER DISTRIBUTION

| Age | Patient (Male/Female) | Control (Male/Female) |
|---|---|---|
| **20-29** | 6 (2M/4F) | 8 (4M/4F) |
| **30-39** | 4 (2M/2F) | 4 (2M/2F) |
| **40-49** | 3 (2M/1F) | 3 (2M/1F) |
| **50-59** | 4 (3M/1F) | 4 (3M/1F) |
| **60-65** | 3 (2M/1F) | 1 (1M) |

This study follows the protocol given below in order to carry out the investigation of patients with LBP. The protocol was approved by the AECC Research Ethics committee in November 2011. There were two stages in the data collection protocol as shown below:

**LBP Data Collection Protocol**

1) **Preparation**
   a) Synchronise Devices using time.
   b) Measure length, clear walking path and make sure conditions are satisfactory for the subject to perform the trial.
   c) Demonstrate how to do the test.
2) **Data Collection**
   a) Place belt onto the subject without devices attached.
   b) Place three mobile phones onto belt and standalone accelerometer onto the subject's body.
   c) Start the data collection application on smart mobile phones.
   d) The subject will be asked to stand still in start position.
   e) The subject will then be asked to step once forward and then backwards, and then remain still for roughly 2 seconds.
   f) The subject will be asked to begin to walk along the path with their preferred speed.
   g) When the subject reaches the destination, the subject will be asked to stop and remain still for roughly 2 seconds.
   h) Stop all devices.
   i) Save data.

*2) Gait Features*

A new and improved iGait system named iGait+ was developed for this study, which includes enhancement to user interface and logic implementation to support the extraction of 9 new features of gait. The iGait+ system introduces features of gait based on a combined direction of acceleration to provide a new direction of motion. The main rationale of introducing the new features was to minimise the impact of device placement, to provide additional gait-based features and also help describe patterns of gait. The combination of all three directions allows the subject's motion to be observed as a single entity path whereas the original iGait system only takes into account one dimensional features from vertical (VER), medio-lateral (ML) and anterior posterior (AP) directions [10]. Additional information may be present within the combined features that can enhance classification performances of machine learning algorithms. Rothney et al. [13] investigated predictions of activity energy expenditure by comparing a tri-axial accelerometer to uni-axial accelerometers (ActiGraph and Actical) and found a more accurate estimation using the tri-axial accelerometer and vector magnitude of the three measurements.

In total, 37 features (as shown in Table II) were extracted using iGait+, which include temporal-spatial features, root mean square (RMS) of the three axis, frequency domain features such as power spectral density of the signal (PSD), regularity and symmetry based features on vertical (VER), mediolateral (ML), anterior posterior (AP) and tri-axial directions. Feature 1 to Feature 28 is described in detail in a previous study [10], while the tri-axial direction used in Features 29 to 37 is defined in (1)

$$m_i = \sqrt{{x_i}^2 + {y_i}^2 + {z_i}^2} \quad (1)$$

where $x_i, y_i, z_i (i = 1,2, \dots, N)$ is the acceleration data of sample $i$ along three directions, i.e. vertical (VER), mediolateral (ML) and anterior posterior (AP) directions respectively.

TABLE II. 37 FEATURES EXTRACTED USING IGAIT+. IPSD: INTEGRAL OF POWER SPECTRAL DENSITY

| Feature No. | Description |
|---|---|
| **Feature 1** | Cadence (step/min) |
| **Feature 2** | Mean step length (m) |
| **Feature 3** | Velocity (m/s) |
| **Feature 4** | RMS VER |
| **Feature 5** | RMS ML |
| **Feature 6** | RMS AP |
| **Feature 7** | Intergral PSD VER |
| **Feature 8** | Frequency at 50% energy (Hz) VER |
| **Feature 9** | Frequency at 75% energy (Hz) VER |
| **Feature 10** | Frequency at 90% energy (Hz) VER |
| **Feature 11** | Frequency at 99% energy (Hz) VER |
| **Feature 12** | Intergral PSD ML |
| **Feature 13** | Frequency at 50% energy (Hz) ML |
| **Feature 14** | Frequency at 75% energy (Hz) ML |
| **Feature 15** | Frequency at 90% energy (Hz) ML |
| **Feature 16** | Frequency at 99% energy (Hz) ML |
| **Feature 17** | Intergral PSD AP |
| **Feature 18** | Frequency at 50% energy (Hz) AP |
| **Feature 19** | Frequency at 75% energy (Hz) AP |
| **Feature 20** | Frequency at 90% energy (Hz) AP |
| **Feature 21** | Frequency at 99% energy (Hz) AP |
| **Feature 22** | Symmetry in VER |
| **Feature 23** | Symmetry in AP |
| **Feature 24** | Stride Regularity in VER |
| **Feature 25** | Stride Regularity in ML |
| **Feature 26** | Stride Regularity in AP |
| **Feature 27** | Step Regularity in VER |
| **Feature 28** | Step Regularity in AP |
| **Feature 29** | RMS Tri |
| **Feature 30** | IPSD Tri |
| **Feature 31** | Frequency @ 50% Tri |
| **Feature 32** | Frequency @ 75% Tri |
| **Feature 33** | Frequency @ 90% Tri |
| **Feature 34** | Frequency @ 99% Tri |
| **Feature 35** | Step Regularity Tri |
| **Feature 36** | Stride Regularity Tri |
| **Feature 37** | Symmetry Tri |

*1) Reliability Testing*

In order to determine whether the features as collected by smart mobile phones are reliable and consistent with the features extracted from a stand-alone accelerometer, the statistical methodology, i.e. Intraclass Correlation Coefficient (ICC) [14], [15] is used. The ICC provides a scalar measure of agreement between the device measurements and can be calculated using the following formula:

$$ICC = \frac{\sigma^2(b)}{\sigma^2(b) + \sigma^2(w)} \tag{2}$$

where $\sigma^2(w)$ is the pooled variance within the extracted features from each device, and $\sigma^2(b)$ is the variance of the trait between the devices.

The coefficient in ICC represents agreement between two or more evaluation methods on the same set of subjects. In the case of this study, the evaluation methods are the different devices, i.e. iPhone 4 device and the stand-alone accelerometer device (Minimod).

There are benchmark standards to interpret the ICC output values. When the ICC is greater than 0.75 an excellent reliability is inferred, an ICC value between 0.4-0.75 indicates a fair-to-good reliability and a coefficient less than 0.4 reflects a poor reliability [16].

*2) Feature selection methods*

In order to select a subset of features derived from iGait+ that correlate most strongly with the classification variable while being mutually as far away from each other as possible, we propose a mutual information based minimum redundancy maximum relevance (MI-mRMR) feature selection method [17] to improve the classification performance of gait analysis for people with LBP. The proposed method was compared to the correlation based feature ranking (CFR) method used in the previous study [19].

*a) CFR feature selection method*

The CFR method selects features based on the classification performance of each individual feature and the correlations among each feature. The algorithm is illustrated in Algorithm 1.

**Algorithm 1.** Correlation and feature ranking based feature selection methods

```
Rank all the features in terms of their predictive performance.
Repeat
    Calculate pair wise correlation among all feature pairs.
    Remove one feature with highest correlation against the highest
     ranking.
    Estimate prediction performance of the classifier using the
    remaining feature subsets.
Until there is only one feature left or significant deterioration of
   prediction performance.
```

*b) Mutual information based mRMR feature selection method*

Mutual information of two random variables is a quantity that measures the mutual dependence of the two variables. As a feature selection method, mRMR identifies subsets of data that are relevant to the existing parameters used and output of the information as Maximum Relevance [17].

For two discrete variable $X$ and $Y$, the mutual information denoted as $I(X;Y)$ is given by the following [18]:

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) log \frac{p(x,y)}{p(x)p(y)} \tag{4}$$

where $p(x,y)$ is the joint probability distribution function of $X$ and $Y$, and $p(x)$ and $p(y)$ are the corresponding marginal probability distribution functions. Mutual information is defined in terms of their probabilistic density function $p(x)$, $p(y)$, and $p(x,y)$. It can be used to characterize both the relevance and redundancy of two variables.

The aim of maximum relevance is to search in the feature set $S$ for all mutual information values between individual feature $f_i$ and class label $c$ to obtain which feature provides maximal relevance, i.e.

$$\max D\,(S,c), \qquad D(S,c) = \frac{1}{|S|} \sum_{f_i \in S} I(f_i; c) \tag{5}$$

where $D(S,c)$ measures the global relevance of the features s in $S$ with respect to the class label *c*.

The features selected by the max relevance could potentially provide high redundancy where dependencies among features exist. It has been suggested by Peng et al. [17] that if a feature was removed from a set of two that depends on each other, the class discriminative power would not change much. Then, the following minimal redundancy can be applied to select mutually exclusive features:

$$\min R\,(S), \qquad R(S) = \frac{1}{|S|^2} \sum_{f_i, f_j \in S} I(f_i; f_j) \tag{6}$$

The Equations (4) and (5) are combined to produce the minimum redundancy maximum relevance (mRMR) feature selection framework, by selecting the features maximizing relevance and taking away the redundancy. There are two common types of mRMR feature selection methods using mutual information, i.e. mutual information difference (MID) and mutual information quotient (MIQ), which can be computed using (6) and (7) respectively [20].

$$\mathrm{MID} = \max_{\mathrm{f_i} \in \Omega S} [\mathrm{I(f_i, c)} - \frac{1}{|S|} \sum_{f_i \in S} I(f_i; f_j)] \tag{7}$$

$$\mathrm{MIQ} = \max_{\mathrm{f_i} \in \Omega S} \{ \frac{\mathrm{I(f_i, c)}}{\left[\frac{1}{|S|} \sum_{f_i \in S} I(f_i; f_j)\right]} \} \tag{8}$$

*c) Classification models*

In this study, an instance-based classification model, i.e. *KStar* [21] is used to assess the predictive performance of the selected features for gait patterns of LBP. It has been shown that KStar can discriminate between the normal and abnormal gait pattern with high accuracy [6].

The model is implemented using the k-nearest neighbours (KNN) framework but uses entropy-based distance to measure the dissimilarity between input samples. The *KStar* function is defined as:

$$K^{*}(b|a) = -\log_2 P^{*}(b|a) \quad (3)$$

where $P^{*}(b|a)$is the probability of all paths from instance *a* to instance *b*. The distance between each instance is defined in *KStar* as the complexity of transforming one instance to another, whereby the complexity is calculated by mapping one instance to another through transformations. The reader is referred to [21] for a detailed description of the algorithm

The classification was implemented using WEKA [22]. Entropic distance measure was used; the global bending was set to 20 and if missing attributes existed, the average column entropy curve was used.

## III. RESULTS AND DISCUSSION

### A. *Validation with WalkinSense smart insoles*

As shown in Fig. 3, data collected from the iPhone accelerometer sensor are compared to the pressure data collected by the smart insoles. The heel section of left insole is used to clearly identify heel contact of the left foot. The peaks represent the highest pressure exerted by the foot. The distance between each peak represent the stride time. The average time between the six peaks from the smart insoles was found to be 1.15 seconds. The average time between the six troughs of the x-axis of the iPhone accelerometer was found to be 1.17 seconds. By using the trough of the acceleration data can provide a measurement of the stride time with +0.02s mean difference.

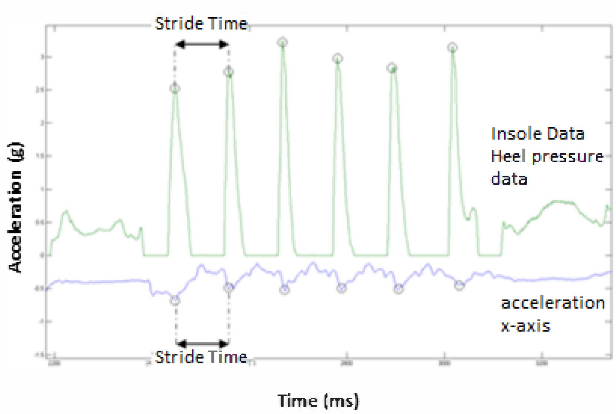


Figure 3. Insole data synchronised with iPhone acceleration data by time

### B. *Reliability testing*

We first compared the complete dataset which includes 40 cLBP patients and 40 control subjects. The result shows that there are 11 features with ICC values greater than 0.75, 17 features between 0.75 and 0.4, and finally 9 features having an ICC value less than 0.4.

In order to remove any features that may be inconsistent and unreliable which were not captured in the dataset as a whole, ICC analysis is instigated on the patient and the control group dataset to individually validate which features are consistently reliable across the two separate datasets. A total of 22 features as depicted in Table III are considered reliable as these features showed reliability (with ICC greater than 0.4) consistently across the two groups, i.e. patients and control data) for both groups.

TABLE III. FEATURES ARE RELIABLE CONSISTENTLY ACROSS THE TWO GROUPS, I.E. PATIENTS AND CONTROL DATA.

| Feature | Description |
|---|---|
| Feature 1 | Cadence (step/min) |
| Feature 2 | Mean step length (m) |
| Feature 3 | Velocity (m/s) |
| Feature 4 | RMS VER |
| Feature 5 | RMS ML |
| Feature 6 | RMS AP |
| Feature 7 | Intergral PSD VER |
| Feature 8 | Frequency at 50% energy (Hz) VER |
| Feature 11 | Frequency at 100% energy (Hz) VER |
| Feature 13 | Frequency at 50% energy (Hz) ML |
| Feature 14 | Frequency at 75% energy (Hz) ML |
| Feature 15 | Frequency at 90% energy (Hz) ML |
| Feature 21 | Frequency at 100% energy (Hz) AP |
| Feature 24 | Stride Regularity in VER |
| Feature 25 | Stride Regularity in ML |
| Feature 26 | Stride Regularity in AP |
| Feature 27 | Step Regularity in VER |
| Feature 29 | RMS Tri |
| Feature 31 | Frequency at 50% energy (Hz) Tri |
| Feature 34 | Frequency at 100% energy (Hz) Tri |
| Feature 35 | Step Regularity Tri |
| Feature 37 | Symmetry Tri |

In observing the results, features showing poor reliability include: Features 9 and 10 (frequency domain features of the vertical acceleration axis, of 75% and 90% of the cumulated integral power spectral density (IPSD)). Features 17, 18 and19 (frequency domain features of the anterior posterior (AP) axis, the IPSD, 50% and 75%. Feature 30 (IPSD of all the axis combined). A reasonable cause for this lack of reliability is the sway of the device from moving forwards. This is a limitation of using smart mobile devices as device placement may become loose. Feature 32 is the 75% IPSD of the cumulated combined axis, as this feature is considered poor reliability for both VER and AP. This can have an impact on the combined frequency domain feature along with Feature 33 which was shown to fall under poor reliability.

### C. *Classification*

In order to determine whether the reliable gait features from smart mobile phones can be used to differentiate between patient and control groups in the case of LBP, the KStar based classification model was employed where a 10-fold cross validation is used to evaluate the model with accuracy (in percentage) and AUC as prediction measures.

When using all reliable features, KStar achieved 83.75% accuracy and the AUC value is 0.952, which demonstrated that people with LBP exhibit different gait patterns from people without LBP. As shown in Fig. 4, the KStar correctly classed 35 instances of cLBP patientS with 5 misclassified instances with a true positive rate of 0.875.

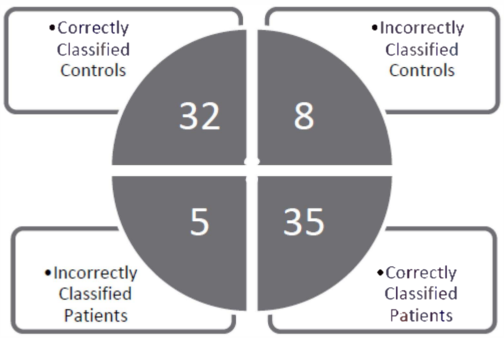


Figure 4. Confusion Matrix to describe KStar classification breakdown

### D. *Correlation based Feature Selection(CFS)*

Table IV shows the classification results using the subsets of features selected by CFS. It performed at best when using 11 features of gait outputting 88.75% classification accuracy.

TABLE IV. THE PERFORMANCE OF KSTAR WHEN USING FEATURESUBSETS GENERATED BY THE CFS METHOD. THE HIGHEST ACCURACY % IS HIGHLIGHTED IN BOLD

| Feature set | ACC (%) | AUC |
|---|---|---|
| **1 Feature** | 56.25 | 0.536 |
| **2 Features** | 70.00 | 0.682 |
| **3 Features** | 70.00 | 0.703 |
| **4 Features** | 66.25 | 0.751 |
| **5 Features** | 66.25 | 0.743 |
| **6 Features** | 65.00 | 0.784 |
| **7 Features** | 72.50 | 0.821 |
| **8 Features** | 77.50 | 0.884 |
| **9 Features** | 83.75 | 0.891 |
| **10 Features** | 85.00 | 0.931 |
| **11 Features** | **88.75** | 0.921 |
| **12 Features** | 83.75 | 0.899 |
| **13 Features** | 82.50 | 0.924 |
| **14 Features** | 80.00 | 0.900 |
| **15 Features** | 81.25 | 0.923 |
| **16 Features** | 87.50 | 0.843 |
| **17 Features** | 83.75 | 0.932 |
| **18 Features** | 86.25 | 0.939 |
| **19 Features** | 86.25 | 0.948 |
| **20 Features** | 86.25 | 0.949 |
| **21 Features** | 83.75 | 0.950 |
| **22 Features** | 83.75 | 0.952 |

### E. *mRMR feature selection*

The CFS-based feature selection method shows that there are possibilities of reducing features whilst increasing the performance of classification. In this section, we examined the behavior of the proposed mRMR. The subsets of features selected by the proposed method is shown in Table V. The prediction performance of the KStar by using these subsets as input is listed in Table VI. The best performance (92.50%) detailed in Fig. 5 was achieved by using only 6 features, i.e. Features 1, 2, 4, 27, 5 and 6. The features within this subset are made up of mainly temporal spatial features (cadence, mean step length) and statistical measures such as the root mean square of AP, VER and ML direction with the addition of Feature 27, the step regularity in the vertical direction, suggesting that cLBP can affect people's walking patterns in all three directions. Also, it is not surprising that cadence and speed are found in this optimal feature set, as both have been widely used in physical activity monitoring and assessment, and this is consistent with clinical findings [23].

TABLE V. FEATURE SUBSET ORDERED BY MUTUAL INFORMATION. THE FIRST SUBSET NAMELY 1 FEATURE ONLY INCLUDES THE TOP FEATURE, I.E. FEATURE 1 FOR BOTH MID AND MIQ, THE SECOND SUBSET NAMELY 2 FEATURES INCLUDE THE TOP 2 FEATURES, I.E. FEATURES 1 AND 2 FOR MID AND FEATURES 1 AND 27 FOR MIQ, AND SO ON

| mRMR | Feature Subset ordered by mutual information |
|---|---|
| **MID** | 1, 2, 4, 27, 5, 6, 29, 37, 35, 15, 24, 3, 26, 25, 11, 7, 13, 31, 14, 8, 34, 21 |
| **MIQ** | 1, 27, 37, 15, 3, 11, 24, 14, 21, 13, 25, 8, 26, 34, 31, 7, 2, 4, 5, 6, 29, 35 |

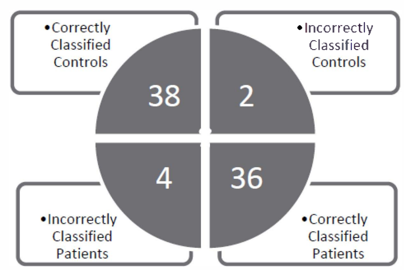


Figure 5. Confusion Matrix to describe KStar classification breakdown when using the optimal feature subset, i.e. Features 1, 2, 4, 27, 5 and 6.

### F. *Contextualising age and gender in gait analysis*

In computing, the use of context is significant for interactive applications to provide relevant information or services to an end-user [24]. Contextualising age and gender can provide personalised services and context-related intervention [25], allowing the device to operate in context specific settings. In this section, we investigate whether the computational models proposed in this study can be applied to detect the age or gender of cLBP patients.

The breakdown of the patient and control groups by age and gender is shown in Fig. 6. The gender and age are used as a class label for classification. The 22 reliable features determined from the reliability study are used. There are two age groups, 20-39 for younger adults and 40-65 for older adults. The two groups for gender are male and female. A 10-fold cross validation is used to assess the classification performance.

TABLE VI. THE PERFORMANCE OF KSTAR WHEN USING FEATURE SUBSETS GENERATED BY THE MID AND MIQ METHODS. THE HIGHEST ACCURACY % IS HIGHLIGHTED IN BOLD

| Feature set | MID (ACC%) | MID (AUC) | MIQ (ACC%) | MIQ (AUC) |
|---|---|---|---|---|
| **1 Feature** | 62.50 | 0.697 | 62.50 | 0.679 |
| **2 Features** | 66.25 | 0.791 | 61.25 | 0.691 |
| **3 Features** | 80.00 | 0.844 | 65.00 | 0.706 |
| **4 Features** | 78.75 | 0.844 | 68.75 | 0.713 |
| **5 Features** | 86.25 | 0.924 | 72.50 | 0.845 |
| **6 Features** | 92.50 | 0.973 | 77.50 | 0.881 |
| **7 Features** | 90.00 | 0.975 | 75.00 | 0.844 |
| **8 Features** | 87.50 | 0.959 | 85.00 | 0.858 |
| **9 Features** | 87.50 | 0.973 | 83.75 | 0.919 |
| **10 Features** | 90.00 | 0.961 | 86.25 | 0.918 |
| **11 Features** | 85.00 | 0.938 | 78.75 | 0.889 |
| **12 Features** | 87.50 | 0.962 | 83.75 | 0.908 |
| **13 Features** | 82.50 | 0.923 | 81.25 | 0.880 |
| **14 Features** | 81.25 | 0.906 | 80.00 | 0.892 |
| **15 Features** | 83.75 | 0.953 | 81.25 | 0.904 |
| **16 Features** | 88.75 | 0.957 | 83.75 | 0.916 |
| **17 Features** | 87.50 | 0.958 | 81.25 | 0.918 |
| **18 Features** | 88.75 | 0.958 | 85.00 | 0.928 |
| **19 Features** | 88.75 | 0.963 | 87.50 | 0.943 |
| **20 Features** | 83.75 | 0.936 | 87.50 | 0.910 |
| **21 Features** | 85.00 | 0.949 | 88.75 | 0.963 |
| **22 Features** | 83.75 | 0.952 | 83.75 | 0.952 |

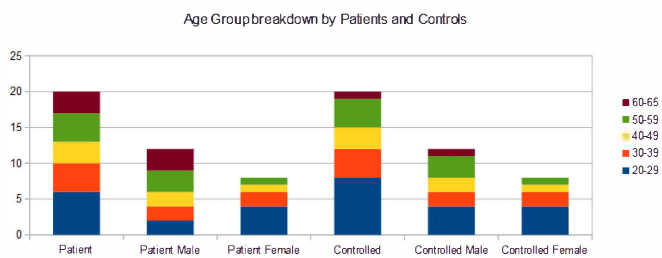


Figure 6. Age and gender groups breakdown by 20 patients and 20 controls

Using 22 features, KStar can achieve the performance of 78.75% for the classification between the younger and older adult age groups and 85.00% for differentiating between gender groups (Table VII). In an attempt to derive the optimal feature set for classification, mRMR MID is applied. An observation was made of the subsets in Table 5, the features 1, 2, 4, 5, 6, 29 and 35 (cadence, mean step length, RMS vertical axis, RMS medio lateral axis, RMS anterior posterior axis, RMS of the new direction and Step regularity of the new direction) were in the same order for both subsets. Mutual information values are high in these values to generate the subset, indicating that the first 7 features of the subset hold higher mutual dependency to both age and gender class variables.

The results in Table VII show mRMR MID feature selection produced an increase of performance of 83.75% accuracy for age group classification using 15 features [1, 2, 4, 5, 6, 29, 35, 37, 21, 26, 7, 24, 25, 27 and 14]. A reduction of features required for classification is also achieved using this method of feature selection. The results in Table 7 show no significant increase (T test, $p > 0.05$) in performance of classification for the gender classification. However, a reduction in feature subset is achievable, 12 features of gait [1, 2, 4, 5, 6, 29, 35, 14, 37, 26, 25, 7 and 13] can produce 85% performance of accuracy in classification.

TABLE VII MRMR MID CLASSIFICATION RESULTS FOR AGE AND GENDER WITH KSTAR

| | AGE GROUP ACCURACY (%) | GENDER GROUP ACCURACY (%) |
|---|---|---|
| **1 Feature** | 51.25 | 60.00 |
| **2 Features** | 60.00 | 61.25 |
| **3 Features** | 65.00 | 57.50 |
| **4 Features** | 71.25 | 80.00 |
| **5 Features** | 76.25 | 80.00 |
| **6 Features** | 77.50 | 75.00 |
| **7 Features** | 71.25 | 75.00 |
| **8 Features** | 72.50 | 81.25 |
| **9 Features** | 76.25 | 83.75 |
| **10 Features** | 73.75 | 80.00 |
| **11 Features** | 76.25 | 81.25 |
| **12 Features** | 73.75 | **85.00** |
| **13 Features** | 73.75 | 82.50 |
| **14 Features** | 77.50 | 81.25 |
| **15 Features** | **83.75** | 80.00 |
| **16 Features** | 80.00 | 78.75 |
| **17 Features** | 73.75 | 81.25 |
| **18 Features** | 78.75 | 80.00 |
| **19 Features** | 76.25 | 85.00 |
| **20 Features** | 78.75 | 82.50 |
| **21 Features** | 76.25 | 83.75 |
| **22 Features** | 78.75 | 85.00 |

*Highest accuracy % is highlighted in bold.

## IV. SUMMARY AND FUTURE WORK

In this paper, we propose two mutual information based feature selection methods to select a subset of gait features to differentiate gait pattern of cLBP from control subjects. Results showed that using machine learning methods, we were going able to achieve 92.50% accuracy in the classification by using only 6 features. In addition, it has been demonstrated that predicting context such as age and gender using smart mobile phones is achievable, which has potential to provide personalised services and context-related intervention. This suggests that the use of a smart mobile phone for monitoring and assessing gait patterns may hold promise for tracking the progress of the gait related diseases and conditions associated with gait changes including, the assessment of treatment outcomes. Following the work presented in this paper, a tele-gait monitoring system, namely iterGait, has been developed by the University of Ulster authors.

In this study, the mobile phone was attached to participants' lower back, which can be inconvenient for daily monitoring. The impact of the placement positions of the mobile deserves further study.

Monitoring gait has the potential to provide insights into both the prognosis and potential impact of clinical interventions in a number of conditions including low back pain. However, remote monitoring using tele assessment is restricted because of the traditional lab based approach to this endeavour.

We present a potential solution to such barriers and show that features unique to LBP patients a can be reliably detected using a mobile smart phones. Such metrics hold the potential to provide treatment monitoring and prognostic information for a range of conditions associated with gait changes.

## REFERNCES


[1] Nutt, J. G., Marsden, C. D., and Thompson,P. D. 1993, Human walking and higher-level gait disorders, particularly in the elderly. *Neurology*, 43(2), 268-268.

[2] Koller, W. C. and Trimble, J. 1985. The gait abnormality of Huntington's disease. *Neurology*, 35(10), 1450-1450, 1985.

[3] Kirtley, C. 2006. *Clinical gait analysis: theory and practice*.: Elsevier Health Sciences.

[4] Gao,T., Greenspan, D., Welsh, M., Juang, R. and Alm, A. 2005. Vital signs monitoring and patient tracking over a wireless network. *Engineering in Medicine and Biology Society*, 102-105.

[5] Morris, R. G. and Lawson, S. E. M. 2009. *A review and evaluation of available gait analysis technologies, and their potential for the measurement of impact transmission*. Newcastle University, UK.

[6] Yang, M., Zheng, H., Wang, H., McClean, S. and Harris, N. 2012. Assessing the utility of smart mobile phones in gait pattern analysis. *Health and Technology*, 2(1), 81-88.

[7] Arendt-Nielsen, L., Graven-Nielsen, T., Svarrer, H. and Svensson, P. 1996. The Influence of Low back pain on muscle activity and coordination during gait: a clinical and experiemental study. *Pain*, 64, 2, 231-240.

[8] Roberts, D., Hamill-Ruth, R., Parker, B., Maximous, S., Clark, B., Nelson, K. 2004. A new method for evaluating efficacy of low back pain therapy. *Pain,*5(3):S116.

[9] M van der Laan and Newell D. 2010. Measures of complexity during walking in chronic non-specific low back pain patients. *Clinical Chiropractic*. 13(1), 8–14.

[10] Yang, M., Zheng, H., Wang, H., McClean, S. and Newell, D. 2012. iGAIT: An interactive accelerometer based gait analysis system. *Computer methods and programs in biomedicine*. 108(2), 715-723.

[11] Tomorrow Options. (2010) WalkinSense. [Online]. http://www.tomorrow-options.com/en/walkinsense.

[12] Chan, H., Zheng, H., Wang, H., Sterritt, R. and Newell, D. 2013. Tele-monitoring and assessment of Gait using smart mobile phone for patients with lower back pain. *In Proceedings of* 5th Annual Translational Medicine Conference, Derry/Londonderry(May 2013).

[13] Rothney, M. P., Schaefer, E. V., Neumann, M. M., Choi, L. and Chen, K. Y. 2008. Validity of physical activity intensity predictions by ActiGraph, Actical, and RT3 accelerometers. *Obesity,* 16(8), 1946-1952.

[14] Maynard,V., Bakheit,A. M. O., Oldham, J. and Freeman,J. 2003. Intra-rater and inter-rater reliability of gait measurements with CODA mpx30 motion analysis system. *Gait & posture*, 17(1), 59-67.

[15] Shrout, P. E. and Fleiss, J. L. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*, 86(2), 420-428.

[16] Reiman, M. P. and Manske, R. C. 2009. *Functional Testing in Human Performance*. Functional Testing in Human Performance, ch. 1.1, pp. 5-6.

[17] Peng, H., Long, F. and Ding, C. 2005. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *Trans Pattern Analysis Machine Intelligence*. 27(8), 1226-1238.

[18] Shannon, C.E. and Weaver, W. 1949. *The mathematical theory of communication*. University of Illinois Press, Urbana, Illinois.

[19] Chan, H. Yang, M.J., Wang, H.Y., Zheng, H., McClean,S., Sterritt, R. and Mayagoitia, R.E. 2013. Assessing Gait Patterns of Healthy Adults Climbing Stairs: a Machine Learning and Statistical Approach. *International Journal of Intelligent Systems*, 28(3), 257-270.

[20] Gokhan, Z. C. and Gulgezen, L. Y. 2009. Stable and accurate feature selection. *Machine Learning and Knowledge Discovery in Databases*, 455-468.

[21] Cleary, J. G. and Trigg, L. E. 1995. K*: An Instance-based Learner Using an Entropic Distance Measure. *In Proceedings of the Twelfth International Conference on Machine Learning* (Tahoe City, 1995). 108-114.

[22] Hall, M. et al. (2009) The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, 11(1), 10-18.

[23] Thomas, L. K., Hislop, H. J. and Waters,R. L. 2008. Physiological Work Performance in Chronic Low Back Disability Effects of a Progressive Activity Program. *Physical therapy*, 60(4), 407-411.

[24] Abowd, G. D. et al. 1999. Towards a better understanding of context and context-awareness. *Handheld and ubiquitous computing*, 304-307.

[25] Bricon-Souf, N. and Newman, C. R. 2007. Context awareness in health care: A review. *International Journal of Medical Informatics*, 76(1), 2-12.