A knowledge-based intelligent framework for anterior cruciate ligament rehabilitation monitoring

Q1 S.M.N. Arosa Senanayke a, *, Owais Ahmed Malik a, Pg. Mohammad Iskandar b, Dansih Zaheer c

a Department of Computer Science, Faculty of Science, Universiti Brunei Darussalam (UBD), Gadong BE 1410, Brunei Darussalam
b Faculty of Integrated Technologies, Universiti Brunei Darussalam (UBD), Gadong BE 1410, Brunei Darussalam
c Sports Medicine & Research Center, Hassanal Bolkiah National Stadium, Berakas, Brunei Darussalam

1. Introduction

The monitoring of recovery progress after anterior cruciate ligament (ACL) injury/reconstruction is crucial for athletes for not only restoring their knee joint stability for dynamic activities and successful returning to sports but also minimizing the risks of re-injury, cartilage degeneration and early onset of osteoarthritis [1–6]. ACL injury causes deterioration in the sports performance or premature end to the career for athletes as ACL is one of the critical ligament for knee joint stability, maintaining normal gait patterns, preventing anterior tibial translation, and controlling knee axial rotation and varus movements [7]. It can be completely or partially ruptured as a result of deceleration due to sudden changes of direction, twisting and/or pivoting during sports, like soccer and basketball. The tearing or rupture of ACL affects gait patterns and variability, and causes changes in kinematics, kinetics and neuromuscular activities of athlete. The absence of ACL results in the loss of mechanical control of the knee, loss of proprioception due to unavailability of mechanoreceptors present in the ligament and neurophysiologic dysfunction, hindering the athletes to participate in sports activities. Efficient and effective rehabilitation programs are essential for athletes following ACL reconstruction as well as for those having ACL deficiency. The rehabilitation programs are designed to rebuilder muscle strength, re-establish joint and neuromuscular...
control and to enable the athletes to return to pre-injury activity level. The goal based criteria (e.g. muscle strength, functional and static knee stability, range of motion) are suggested to safe return to sports rather than relying on some specific time period for rehabilitation [7–9]. The conventional methods for post-ACL reconstruction recovery assessment test the static and dynamic knee stability of subjects. However, the dynamic knee stability analysis usually lacks quantifiable constructs, and some subjective or partially objective evaluations and knee rating scores are performed in order to evaluate the recovery progress of ACL reconstructed (ACL-R) subjects [10–15]. Moreover, physiological variations in subjects and coopers/non-coopers phenomenon indicate that the rehabilitation process may differ for an individual or a group of subjects [16]. Thus, in order to provide an effective cure and timely intervention during recovery process, a comprehensive rehabilitation monitoring system is essential for ACL reconstructed (ACL-R) subjects.

Intelligent techniques have been found effective for different medical applications including medical diagnosis, pathology classification, rehabilitation of lower limb movements after neurological disorders and implantation of artificial limbs [17–22]. These techniques assist clinicians in decision making process about subjects’ health condition. Although, there has been number of studies using intelligent mechanisms for assessment and classification of gait patterns for subjects having lower limb impairments due to stroke or cerebral palsy but few efforts have been made for gait and recovery analysis after ACL injury/reconstruction. Mostly the assessment of injured/reconstructed subjects is still based on the comparison of mean and peak values or some other statistics of kinematics, kinetics and neuromuscular parameters with controlled subjects [6,23–27]. Moreover, in most of the previous studies individual parameters are observed and few efforts have been made for developing an intelligent assistive tool for monitoring recovery of athletes after anterior cruciate ligament (ACL) injury and reconstruction [28,29].

In order to differentiate between ACL intact (ACL-I) and ACL deficient/injured, an intelligent classification system has been developed using adaptive neuro-fuzzy inference (ANFIS) system based on arthrometric data [30]. This system recognizes the normal and ACL injured subjects with high values of sensitivity and specificity but it does not consider the knee dynamics changes in ACL-R subjects. A classification model to test the gait normality of ACL-R subjects has been proposed in [28] based on principal component analysis and regression modeling techniques. This system used 3-D rotational kinematics and the classification results indicated the presence of abnormal gait patterns in frontal and transversal planes for ACL-R subjects still 1 year after surgery with an accuracy of 93.75%. Recently, an intelligent recovery classification model has been presented for ACL-R subjects to monitor their rehabilitation progress during different ambulation and balance testing activities [19]. This system integrates the kinematics and neuromuscular signals and the accuracy of the classification results is approximately 95% for different activities monitored. Although these studies used computational intelligent techniques for differentiating the normal and impaired human motion but post-surgical monitoring and prognosis of rehabilitation and sports performance of ACL-R subjects are fairly complex processes because of the involvement of multiple varying factors for each subject. The restoration of normal gait patterns does not ensure the complete recovery and return to high level sporting activities. In addition, keeping a record of pre- and post-surgery treatments and their effectiveness, adjustments in individual’s rehabilitation protocol and all the different experiences during convalescence is a complicated phenomenon for physiatrists and physiotherapists. These experiences can be used to learn and adapt the rehabilitation procedure for ACL-R subjects having similar parameters or patterns. In order to build such an exhaustive system, a knowledge base of athletes’ profiles can be developed to store the pre-injury, pre-surgery and post-surgery (during rehabilitation period) information about athletes’ knee dynamics and other relevant parameters. Based on this information and past experiences, a learning model can be developed for improving the recovery after ACL injury/reconstruction. An effective method to build this learning system is to use case based reasoning (CBR) paradigm which maintains a case repository of past experiences (old problem-solution pairs) and solves the new problems by using or adapting the existing solutions [31].

CBR has been used as a decision supporting technique in variety of healthcare applications [32]. Based on CBR approach, a gait disorder analysis system has been proposed in [33] to help general practitioners. This system records lateral and forward movements of center of mass by using an accelerometer and retrieve similar cases based on the experts’ knowledge. In order to provide diagnosis and prognosis for stroke patients, CBR based system has been successfully developed using data collected through robotic tool [34]. The system retrieves a potential solution for stroke diagnosis by finding similar cases from the repository of stroke patients with explicit diagnosis and prognosis. Recently, different intelligent techniques have been combined with CBR approach to build hybrid/soft CBR systems for complex medical domains. A hybrid system combining Self Organizing Maps (SOM) and CBR has been developed in [35] for evaluating postural control of subjects based on trunk sway with a prediction accuracy of more than 90%. For diagnosis and treatment of stress, a reliable hybrid CBR system (88% accuracy) has been designed by using finger temperature [36]. The hybrid models based on integration of CBR and fuzzy decision tree provide accurate forecasting in the domains of breast cancer (98.4% accuracy) and liver disorders (81.6% accuracy) [37]. Another application of combining CBR and cluster analysis is the development of a health monitoring system for elderly subjects based on pulse rate and blood oxygen saturation data [38]. This system classifies the pulse rate and blood oxygen saturation with an accuracy of 93%.

This research aims at developing an intelligent framework using knowledge-based system and artificial intelligence (AI) techniques for monitoring and classification of recovery status of athletes, and evaluating their sports performance after ACL reconstruction. A knowledge base has been designed to store the subjects’ profiles and problem/solution pairs by integrating adaptive learning techniques and reasoning model to build an extensible recovery process monitoring system which provides efficient and effective solution based on the existing cases. Non-invasive body-mounted motion and electromyography (EMG) wireless sensors have been used to collect the kinematics and neuromuscular data from ACL reconstructed and healthy subjects during various ambulation and balance testing activities. Different statistical and time/frequency features have been extracted from kinematics and EMG data, and a feature reduction method has been applied to reduce the large number of features and generate a set of appropriate integrated kinematics and neuromuscular patterns. In order to facilitate the efficient selection and retrieval of cases and flexible knowledge base design, the fuzzy clustering technique has been used to form the groups of subjects based on their current recovery stage after ACL reconstruction. Different intelligent classification techniques (adaptive neuro-fuzzy inference system, fuzzy unordered rule induction algorithm and support vector machine) have been explored and compared for finding the similar case from the knowledge base. Once relevant cases are selected, adaptation is performed by adjusting the recovery protocols for individuals based on the previous known solutions and the performance evaluation can be done. This framework will facilitate the clinicians, physiotherapists, physiatrists and sports trainers in determining the recovery stage of ACL-R subjects based on the data collected during different rehabilitation activities and identifying the subjects lacking...
2. A knowledge-based intelligent framework for ACL recovery monitoring

The knowledge-based intelligent framework for monitoring rehabilitation progress and sports performance of ACL-R subjects is shown in Fig. 1. The components and sequential operations of the system are elaborated in the following sections.

2.1. System hardware and sensors’ placements

The knee joint dynamics are altered due to ACL injury/reconstruction so the major focus of the rehabilitation process is to restore the kinematics and neuromuscular control of ACL-R subjects back to the normal or pre-injury level. In order to measure these kinematics and neuromuscular signals, two subsystems have been used in this study. The kinematics data are recorded using wireless micro-electro-mechanical systems (MEMS) motion sensors units with command module and a radio for wireless data transmission (KinetiSense from ClevMed Inc.) and the neuromuscular signals are acquired through a wireless electromyography sensors unit (BioRadio) with electrodes and a USB receiver for wireless data transmission (BioCapture System from ClevMed. Inc.). The body-mounted MEMS motion sensors measure the 3-D motion using three orthogonal gyroscopes and three orthogonal accelerometers. The BioRadio, worn by the subjects, records the EMG signals through surface electrodes attached to the target muscles and then wirelessly transmits them to the computer using USB receiver. These small, light-weight and untethered sensors can be used without obstruction to monitor required signals from lower limbs during dynamic activities (e.g., walking, running, balance testing, one leg jumping) where the use of conventional equipment (e.g. arthrometer, goniometer, wired-EMG) may be cumbersome or not feasible.

The motion sensors were attached to each leg (one on thigh and one on shank as shown Fig. 1) of a subject using flexible bulk and Velcro straps to note the knee joint movements. For recording the EMG signals, foam snap electrodes were placed on four different knee extensor and flexor muscles including vastus medialis (VM), vastus lateralis (VL), semitendinosus (ST) and biceps femoris (BF) on both legs of the subjects during ambulation activities. For balance testing activity, EMG data were also collected from gastrocnemius medialis (GM) muscle in addition to above four muscles. SENIAM guidelines were followed to prepare the skin for better conductance of signals and placing the electrodes on identified position on muscles [39]. The command module and the BioRadio were worn by the subjects using a waist belt. These acquired kinematics and neuromuscular signals were the basis to derive raw pattern set for the formation of knowledge base as described in Section 2.2.

2.2. Data collection and processing

2.2.1. Data acquisition

Both sensing units (KinetiSense and BioRadio) and system software were initialized and setup for recording signals from the motion and EMG sensors. Each motion sensor recorded the 3-D angular rates and 3-D linear acceleration during the motion performed by the subjects. The motion sensor data were sampled at 128 Hz (12 bit analog/digital resolution within a frequency range 0–20 Hz) and digitized in the sensor unit. The recorded angular rates and linear acceleration from all four sensors were sent simultaneously in serial format to the Command Module which transferred the digital data to the computer through the USB Receiver connected to the computer’s USB port. The raw signals (angular rates and accelerations) from motion sensors were viewed and saved using KinetiSense software package for processing.

The surface EMG sensors were used to record the action potential of skeletal muscles indicating the force of muscles. The amplitude of the EMG measurements (usually in millivolts) was related to the amount of force generated from the muscle’s contraction while performing the ambulation or balancing activities. The sampling rate to collect EMG signals was set to 960 Hz at 12 bit Analog/Digital conversion and 2-D linear acceleration was also recorded from BioRadio unit to synchronize both KinetiSense and BioCapture devices. The EMG data recorded using BioRadio were transferred to the computer wirelessly through USB receiver...
connected to the computer’s USB port. These signals were viewed and stored using BioCapture software package for processing.

The data from both sensing units were simultaneously recorded and transmitted to the same computer. These stored signals (angular rates, accelerations and raw EMG data) were then exported to the files for further processing by custom-developed software.

2.2.2. Data processing

The raw signals from motion and EMG sensors were filtered and transformed into the required format in order to prepare the data for pattern set generation and extraction. For each motion sensor, measurements for zero-referencing were obtained prior to starting the experiment (actual motion) when the subjects were standing upright position and these measurements were subtracted from angular rate measurements of corresponding sensor during the experiment. The measurements obtained from the MEMS gyroscopes were low-pass filtered using 6th order Butterworth filter before computing the orientations in order to minimize the motion artifacts. After ACL injury/reconstruction, the kinematics of knee joint change and mainly the knee movements in sagittal plane are affected. In order to monitor these changes, the flexion/extension measurements were computed for each gait cycle using angular rates recorded through motion sensor units placed on the thigh and shank segments of both legs. The motion sensors were aligned to provide knee angle about the sagittal plane. Let \( \theta \) represents the orientation of a lower limb segment (thigh or shank) then the orientations of both lower extremities (\( \theta_{\text{T}h\text{igh}}, \theta_{\text{R}sh\text{ank}}, \theta_{\text{L}t\text{h}igh}, \text{and} \theta_{\text{L}sh\text{ank}} \)) can be estimated by applying trapezoidal integration method on respective angular rates (\( \dot{\theta}_{\text{T}h\text{igh}}, \dot{\theta}_{\text{R}sh\text{ank}}, \dot{\theta}_{\text{L}t\text{h}igh} \) and \( \dot{\theta}_{\text{L}sh\text{ank}} \)) of lower limbs. If \( \omega_{\theta}(t) \) represents the angular rate of either thigh or shank at time \( t \) and \( \Delta t \) is the sampling time, then the estimated orientation \( \theta_{\text{T}h\text{igh}}(\Delta t) \) of thigh and shank at time \( t \) is computed using (1).

\[
\theta_{\text{T}h\text{igh}}(\Delta t) = \frac{\omega_{\theta}(t) + \omega_{\theta}(t+1)}{2} \times \Delta t
\]

The knee angle was computed by subtracting \( \theta_{\text{L}sh\text{ank}} \) from \( \theta_{\text{T}h\text{igh}} \) for both legs. Knee flexion/extension and corresponding angular rates for thigh and shank are shown for a healthy (ACL intact) leg in Fig. 2.

Similarly, the raw EMG data from all muscles were band-pass filtered (10–450 Hz) using 4th order Butterworth filter and the mean of each signal was subtracted from it. The processed data are stored in the knowledge base and a preliminary profile of the subject was generated.

2.2.3. Data segmentation and feature extraction

Important features were extracted from the processed kinematics and EMG data in order to design the case and knowledge base, and perform the data analysis and classification tasks. The selection of right features minimizes the assessment error and helps in achieving the high retrieval accuracy. The feature set consisted of knee dynamics based on combined kinematics and EMG signals. Both signals were synchronized before feature set generation due to different sampling rates and recording delays and then based on the ambulation or balance testing activity, the features set generation and selection steps were performed. The data segmentation for ambulation activities was done by identifying the gait cycles during each trial for all recording sessions of a subject. The knee kinematics and relevant muscles’ strength vary in each phase of a gait cycle and these changes reflect the progress of the recovery after the ACL injury/reconstruction. The features were computed for different phases of a gait cycles for each subject (Fig. 3). The marking of gait phases and, segmentation of data from multiple EMG channels and motion sensors were based on the percentages defined in [40]. For balance testing activity, a window of four seconds was chosen (based on experiments) as a data segment for kinematics and EMG signals.

Three statistics values (root mean square, standard deviation and maximum value for knee flexion/extension) were computed for each segment (i.e. seven phases of a gait cycle for ambulation activities and a four seconds window for balance testing activity) for kinematics data. Thus, a derived actual pattern set generated consisted of a total of twenty-one kinematics features for each gait cycle for ambulation activities and three kinematics features for balance testing activity.

The EMG features represent the electrophysiological properties of the muscle fibers during contraction. In different applications, various time-domain, frequency-domain and time-frequency-domain features extracted from EMG signals have been used [41–43]. The wavelet transformation has been found more appropriate for analyzing such bio-signals as compared to only time/amplitude or frequency analysis due to stochastic and non-stationary nature of EMG signals. In this study, an EMG feature set has been generated based on multilevel discrete wavelet decomposition analysis. By employing discrete wavelet transform (DWT), EMG signal can be iteratively transformed into multi-resolution sub-levels of coefficients using suitable wavelet basis function. The time-domain EMG signal (\( EMG(t) \)) is passed through various high pass and low pass filters to obtain the approximation coefficient subsets (e.g. \( a_1, \ldots, a_3 \) as shown in Fig. 4) and the detail coefficient subsets (e.g. \( d_1, \ldots, d_3 \) as shown in Fig. 4) where the level of decomposition can be pre-defined (e.g. 3 in Fig. 4). Mathematically, after n level of decomposition, the original signal can be represented as follows (2):

\[
EMG(t) = c_{A0} + c_{D1} + c_{D2} + \ldots + c_{D_n} \tag{2}
\]

The choices of level and mother wavelet depend on the domain and applications, but for EMG analysis Daubechies 04/05 mother wavelet with four/five levels of decomposition has shown better performance results [43–45]. In this research, Daubechies 05 wavelet basis function with five levels of decomposition has been used to compute EMG features. An example of EMG signal analyzed by DWT with mother wavelet Daubechies 05 (db05) with 3-levels of decomposition is shown in Fig. 4.

For various ambulation and balance testing activities, wavelet coefficients (\( c_{D1} – c_{D3} \) and \( c_{A3} \)) were computed which represent the energy distribution of the EMG signals from four/five identified muscles. From these coefficients, following statistical features (maximum, minimum, mean absolute value, standard deviation and average power of the six coefficients) were calculated for each phase of gait cycle and balance testing segment. These features were chosen based upon previous studies for bio-signal classification [43,46] and the experiments performed in this study for recovery and performance assessment. Thus, a derived actual pattern set consisted of a total of 840 EMG features (6 coefficients × 5 EMG features × 4 muscles × 7 phases) for each gait cycle for ambulation activities and 150 EMG features (6 coefficients × 5 EMG features × 5 muscles) for balance testing activity.

2.2.4. Feature selection/projection

The lengths of a feature vector (combined kinematics and EMG parameters) for each ambulation and balance activity were 861 (6 × 5 EMG features × 4 muscles × 7 phases + 3 kinematics feature × 7 phases) and 153 (6 × 5 EMG features × 5 muscles + 3 kinematics feature), respectively. In order to reduce the length of these feature vectors, different feature selection and reduction algorithms were investigated. Instead of using the feature selection algorithms which select only a subset of the features to represent the model, feature projection/reduction method was preferred due to its efficiency and effectiveness. The selection algorithms (sequential forward selection/sequential backward selection) were

---

Fig. 2. (a) Knee flexion/extension (degrees), (b) Angular rate for thigh (rad/s), (c) Angular rate for shank (rad/s) for multiple gait cycles of an ACL intact leg.
found to be slow to test each subset for the prediction error for the large feature number of features and were unable to re-assess the feature’s importance as once a feature was added to the subset it could not be removed. In this study, one of the widely used feature reduction/projection technique namely principal component analysis (PCA) was applied to determine the best combination of the original features to form a new and smaller feature set. PCA has been successfully used reduce the dimension of bio-signals data set by removing the redundancy in the data and replacing the group of variables with a single variable while still not rejecting some of the features completely from the data set [47, 48]. PCA transformed the original feature set of kinematics and EMG parameters \( f_i \in F \subseteq \mathbb{R}^{M} \) into a new feature set of variables \( v \in V \subseteq \mathbb{R}^{N} \) of reduced dimension by minimizing the mean-square error (MSE) between the original set \( F \) and projected set \( V \) [49]. For a given set of kinematics and EMG input vectors, \( f_i \) (\( i = 1, \ldots, n \)), where each \( f_i \) is of dimension \( N \) (\( N=861 \) for ambulation and \( 153 \) for balance testing activity), PCA linearly transformed each vector \( f_i \) into new vector \( v_i \) by using (3).

\[
v_i = A^T f_i
\]

where \( A \) is the \( N \times N \) orthogonal matrix whose \( i \)th column \( a_i \) is the \( i \)th eigenvector of the sample covariance matrix in (4).

\[
\text{Cov} = \frac{1}{n} \sum_{i=1}^{n} f_i f_i^T
\]  

(4)

These new variables, called principal components, were the linear combination of original kinematics and EMG features, which form an orthogonal basis for the space of data. The feature projection step can only be performed after collecting and extracting features for multiple trials from a group of healthy and ACL-R subjects so that PCs covers variations in kinematics and EMG parameters. Hence, a final pattern set was generated based on this method which was used as the input to develop the knowledge based system.

2.3. Fuzzy clustering

In order to make the case selection and retrieval process more efficient and to avoid searching the whole case repository, the knowledge base was designed using clustered based structure. The

![Fig. 3. Variations in knee flexion/extension during different gait phases.](image)

![Fig. 4. EMG signal for vastus lateralis of a healthy (ACL intact) subject walking at 5 km/h analyzed by DWT with Daubechies 05 mother wavelet at 3-levels of decomposition.](image)

clustering was performed by grouping the subjects according to similarities in their kinematics and EMG parameters. There could be various groups based on the health condition or stage of recovery of subjects so unsupervised fuzzy clustering has been adopted to assign the subjects to a group due to the imprecise nature of motion and neuromuscular parameters. In the domains of recovery classification or gait analysis, variations in data are more common and the clusters of cases may be overlapping such that one subject may belong to more than one group with different degree of memberships. Fuzzy clustering has been applied for gait analysis and identifying the effect of temporal patterns on the walking speed [50, 51].

Fuzzy clustering partitions the sample space and organizes the data into approximate clusters [52, 53]. In this study, fuzzy C-means (FCM) algorithm was applied to the transformed feature set ‘V’ of kinematics and neuromuscular data collected during various ambulation and balance testing activities. The fcm function in Fuzzy Logic Toolbox from MATLAB starts with an initial guess for the cluster centers for marking the mean location of each cluster. Each feature vector as data point is then assigned a membership grade for each cluster. fcm follows an iterative process to minimize the objective function (5) and then decides the right cluster centers (6) and membership grade (7) for each feature vector.

\[ \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m ||v_i - c_j||^2 \]  

(5)

\[ c_j = \frac{\sum_{i=1}^{N} u_{ij}^m v_i}{\sum_{i=1}^{N} u_{ij}^m} \]  

(6)

\[ u_{ij} = \frac{1}{\sum_{k=1}^{C} (||v_i - c_k||/||v_i - c_j||)^{2/(m-1)}} \]  

(7)

where \( v_i \) is the \( i \)th of \( M \)-dimensional data from PCA, \( c_j \) is the \( M \)-dimension center of the cluster, \( u_{ij} \) is the degree of membership of \( v_i \) in the cluster \( j \), \( || \cdot || \) represents the similarity (Euclidian distance) between any measured data and the center, and \( m \) is the number of clusters. In the initial stage of populating the case library, the variation in number of clusters indices is possible which is reduced as more cases are added to the repository. In order to identify the number of initial clusters, a modified cluster validity measure has been used [54] which is mathematically illustrated as follows (8).

\[ \text{Validity} = \frac{\sum_{i=1}^{K} \sum_{v \in V} u_{ix} \times ||v - c_i||^2}{K N} \]  

(8)

where \( c_i \) and \( c_j \) are the cluster centers, \( N \) is the number of records in the data set, \( k \) is the maximum number of clusters, \( v \) is a subset (selected features) of \( V \) and \( ||v - c_i||^2 \) and \( ||c_i - c_j||^2 \) represent the intra and inter cluster distances, respectively.

In the proposed framework, fuzzy clustering step was an iterative process and as more cases would be generated and added to the repository, the number of clusters would be refined. The clusters were generated based on the data collected from two trials of each activity from all subjects so that the variations in the gait/balance patterns could be recorded and used to generate the classes. These clusters were used to train the intelligent classifier(s) for efficient retrieval of cases.

2.4. Case generation

The generation of cases was based on two scenarios; initially when the case repository was empty and subsequently when case base was already populated with some cases. In the first situation when there was no case in the repository, data from a group of healthy and ACL-R subjects at different recovery stages were collected and clustered using fuzzy clustering technique to form the groups based on their current health condition and recovery stage. The labels for these groups were identified by finding the distances of their centers from the center of healthy subjects’ cluster and were manually verified. Then new cases were generated based on a semi-automatic process where both extracted features and recommendations from the clinicians were stored using the proposed case representation (explained in Section 2.6.2). In the later scenario, a new case was added based on the outcome of CBR cycle (retrieve, reuse, revise and retain) and re-grouping/re-training of classifier was performed if required. Each solved case, stored in the repository, included the problem description, average value of transformed attribute/value pairs for multiple trials, recovery stage, class, recommendations and protocols followed during rehabilitation.

2.5. Intelligent classification mechanisms

The retrieval of a case from the case repository essentially means matching a current input pattern with one or more stored patterns or cases. Intelligent classification techniques have been proved very efficient for performing pattern matching task. The variations in kinematics signals, and non-stationary nature of EMG data lead the recovery classification task challenging. The adaptive intelligent techniques can be more useful for building models for such inputs. These techniques can effectively identify the stochastic changes in bio-signals, and can also deal with the imprecision in measurements and variations due to subjects’ physiological conditions [55, 56]. More specifically, fuzzy logic based techniques are very useful when the data are incomplete, noisy, or imprecise. Three different intelligent classification techniques have been explored and compared in this study in order to implement the case retrieval task for recovery and performance analysis.

2.5.1. Adaptive neuro-fuzzy inference system (ANFIS)

ANFIS has been applied for addressing the problem of selecting cases from overlapping cluster regions. The ANFIS is a fuzzy Sugeno model that adapts the membership function parameters using neural network and learns from the given data set [57, 58]. ANFIS combines the learning capabilities of neural networks with fuzzy set theory which handles the uncertainty arising due to overlapping cases. The system adjusts the membership function (\( \mu \)) parameters based on the given data, and the number of rules and output of fuzzy rules is minimized. The overall output of an \( n \)-input system is given in (9) where \( i \) is the previous layer’s output and \( w_i \) is called the firing strengths of the rules (10).

\[ y = \frac{\sum_{i=1}^{n} w_{i} \mu_i}{\sum_{i=1}^{n} w_{i}} \]  

(9)

\[ w_i = \prod_{j=1}^{n} \mu_j(x_j) \]  

(10)

The subtractive-clustering method was used to partition the data due to large number of inputs, no requirement of setting the number of clusters in advance and noise robustness. It is a one-pass algorithm for estimating number of clusters by finding the high density data point regions in feature space. The cluster center is the point with the highest number of neighbors. The learning parameters of membership functions (premise parameters) and output (consequent parameters) were tuned using the hybrid learning algorithm. This algorithm combines the least square method and gradient descent method that make the convergence faster in the large search space. The forward pass (least square method) and
backward pass (gradient descent method) are used to optimize the consequent and premise parameters, respectively. After determining the consequent parameters, the output of ANFIS is calculated and the premise parameters are adjusted based on output error by using the back-propagation algorithm [57].

2.5.2. Fuzzy unordered rule induction algorithm (FURIA)

FURIA is a recent classification algorithm based on the RIPPER algorithm with certain modifications and extensions [58]. It learns the fuzzy rules having soft boundaries as compared to conventional strict rules and provides an unordered rule set using one vs. rest strategy such that there is no default rule. A rule stretching mechanism is used in case when a new problem is not covered by any of the existing rules. A fuzzy rule in FURIA is obtained through replacing the RIPPER’s intervals with fuzzy intervals \( F_i \) of the form of trapezoidal membership functions specified with four parameters as follows:

\[
F_i = \begin{cases} 
1 & b \leq v \leq c \\
\frac{v-a}{b-a} & a < v < b \\
\frac{d-v}{d-c} & c < v < d \\
0 & \text{else}
\end{cases}
\]  

(11)

where \( b \) and \( c \) are the lower and upper bound of the core of the fuzzy set, respectively, and \( a \) and \( d \) are the lower and upper bound of the support, respectively. If a fuzzy selector \( A_i \in F_i \) covers an instance \( x = (x_1, \ldots, x_n) \) to the degree \( F_i(x) \) then a fuzzy rule \( r^i \) involving \( k \) selectors \( A_i \in F_i \), \( i = 1, \ldots, k \), covers \( x \) to the degree as follows:

\[
\mu_{r^i}(x) = \prod_{i=1}^{n} F_i(x_i) 
\]  

(12)

The fuzzification of a \( A_i \in F_i \) is done by considering only the relevant training data \( D^p_i \) and \( D^n_i \) is partitioned into positive and negative instances. For a new query instance \( x \), the support of this class is defined by

\[
s_j(x) = \sum_{i=1}^{n} \mu_{r_i}(x_i) \cdot CF(r_i) 
\]  

(13)

where \( CF \) (certainty factor of rule \( r_{i}^{j} \)) is defined as follows:

\[
CF(r_{i}^{j}) = \frac{(2D_p^j/D_T) + \sum_{x \in D_p^j} \mu_{r_i}(x)}{\sum_{x \in D_p^j} \mu_{r_i}(x)} 
\]  

(14)

2.5.3. Support vector machines (SVMs)

SVMs combine the statistical learning theory with optimization techniques and kernel mapping to solve the classification problem. In case of binary support vector classification, the feature vectors are separated based on optimal hyperplanes [59]. For linear separable data, the hyperplane separating the data into classes can be described using following equation:

\[
x^T w + w_0 = 0 
\]  

(15)

where \( x \) represents the feature vector set and \( w \) is normal to the hyperplane. An optimal separating hyperplane maximizes the sum of distances from the hyperplane to the closest data points on each side. If \( y_i \) represents the possible class labels (+1, -1) then the optimal hyperplane can be found by minimizing \( ||w^2|| \) subject to

\[
y_i(x_i^T w + w_0) \geq 1 
\]  

(16)

Different kernel mapping (quadratic, Gaussian radial basis function, polynomial etc.) can be utilized other than the linear function to improve the classifier’s performance based on the specific domain. This study has compared the effect of three types of kernel functions (linear, quadratic and radial basis function) on the classification accuracy for ACL recovery analysis. In order to deal with the multi-class problem, the binary SVM has been modified using one vs. all strategy.

2.6. Hybrid case based reasoning

A hybrid intelligent framework was developed by combining case-based reasoning (CBR) approach and adaptive intelligent mechanisms for recovery stage diagnosis and prognosis of ACL-R subjects. The framework utilizes the concept of solving new problems by using/modifying the similar previous experiences (problem–solution pairs). CBR problem-solving cycle consists of four steps [31]:

- Retrieve: Finding similar case(s) from the knowledge base whose problem description best matches with the given problem.
- Reuse: Reusing the solution of most similar case to solve the new problem.
- Revise: Adapting/modifying the chosen solution according to the differences in new problem.
- Retain: Storing the new problem–solution pair as a case once it has been solved.

The details about the different components of the hybrid intelligent knowledge-based system are described in next sections.

2.6.1. Knowledge base (KB)

The overall structure of knowledge base is depicted in Fig. 5. The knowledge base contains different types of information including case library, raw and processed data, domain knowledge, historical data available for subjects (pre-injury, post-injury) and session data during convalescence. In order to manage the knowledge base repository, a relational database was used to reduce the storage redundancy and provide flexibility. The knowledge base evolves with the time-period when new problems are presented and new cases are added to the system. This evolution process makes it more useful for domains where subject’s specific monitoring and prognosis mechanisms are required.

2.6.2. Case representation

Each case in the knowledge base is composed of two components: activity set and overall set. The activity set consists of subsets of activity-based (walking at different speeds and balance testing) problem–solution pair. The problem part of each subset is represented as attribute/value pairs for the selected kinematics and EMG features and the corresponding solution part is made up of the recovery classification and the treatments suggested for the next stage. The overall set contains the performance evaluation for the athletes, treatments given and followed at current stage of recovery, link to the previous sessions’ outcomes (if any) and case description. The case description contains the subject’s biographical, sports and other relevant clinical information. This information may provide assistance in refining the retrieval and adaption of similar cases.

2.6.3. Retrieval of similar cases

In order to retrieve the most similar cases for the given scenario, a two step process is performed. In the first step, the best matching class is identified for the new problem by using trained classifier as described in Section 2.5. In the second step, the most similar
cases are retrieved from the identified cluster by using case density function given in (11) [60].

\[
Case\ Density(e, C) = \frac{\sum_{e' \in C} SM_{e'}}{|C| - 1}
\]

where \(SM_{e'}\) represents the similarity between case \(e\) and \(e'\) (computed by cosine/Euclidian measure) and \(|C|\) is the number of cases in class \(C\). The retrieved cases are arranged in the descending order of case densities and first \(k\) cases are selected where \(k\) is the user’s choice.

2.6.4. Reuse and repair of cases

After retrieving the most similar case(s), the next step is to use and adapt the solution of this case to improve the recovery process of the athlete for the next stage. This semi-automatic process requires the involvement of the clinicians/physiotherapist to decide any changes in the rehabilitation protocol based on the recommendations and indication of performance level from the retrieved case(s). Additionally, modifications may also be done in the recommendation section of the previous stages of the new problem.

2.6.5. Retaining the learned case

The new case may be retained in the library after formulating a solution based on the adjustments of parameters for individuals. The attribute/value pairs for the selected features are generated by taking an average of their values for all sets. Whenever a new case is added, the fuzzy clustering algorithm generates new clusters (if required) and assigns different membership grades to the related cases.

![Fig. 5. Overall structure of knowledge base containing different types of data.](image-url)
4. Results and discussion

In order to generate the feature set and populate the case repository, data were collected from healthy and ACL-R subjects during three ambulation and one balance testing activities. A large number of kinematics and EMG features were extracted from the human motion during these activities to evaluate their effect on the recovery analysis after ACL reconstruction. However, the large feature vector/set represented the data redundancy and made the classification process inefficient so PCA was applied to reduce the feature sets for all activities monitored during this study. Before applying PCA, the total data set was partitioned into two groups. The first group consisted of data from 14 subjects (4 healthy and 10 ACL-R) and the second group contained data from 5 subjects (2 healthy and 3 ACL-R). The data from first group were used to generate coefficient matrix from PCA, form clusters, train and test classifiers, and generate cases. For ambulation activities, the size of the pattern set for first group was 280 × 861 (i.e. 240 records from 14 subjects for 20 gait cycles with 861 kinematic and EMG features). For balance testing activity, the size of the pattern set for the same group was 56 × 153 (i.e. 56 records from 14 subjects for 4 segments with 153 kinematics and EMG features). The latter group was used to test the performance of the framework. PCA was applied on the data for first group and the transformed features with respective coefficient matrix were generated for each activity. The variations explained by principal components (PCs) for these activities are shown in Fig. 6. The number of PCs required for a cumulative distribution of variance of 90% or more vary among activities and were found as follows: 42 PCs for walking at 4 km/h, 39 PCs for walking at 5 km/h, 37 PCs for walking at 6 km/h and 6 PCs for balance testing. The coefficient matrix for each activity was stored to apply to the data for new subjects. Fig. 7 shows the original data projected on first three PCs for walking at 4 km/h. Another advantage of PCA is that it can also be useful in identifying the subjects who have very different values of the input parameters (marked inside circles in Fig. 7). These data points can be treated as outliers/error and removed from the data set before clustering/classification step or further steps can be taken to investigate the subjects for whom these data points belong to.

Before generating the cases, the transformed features were clustered using FCM to form the groups of subjects who were healthy or at similar stage of recovery after ACL reconstruction. This step is part of the initial grouping and case generation process and can be applied if re-grouping is required. Three clusters were initially identified for each activity using proposed validity index (8) which were manually verified and found to be appropriate. Fig. 8 and Fig. 9 show 3-D scatter plot for the first three PCs and the cluster centers identified by FCM for walking at 5 km/h and 6 km/h respectively, where clusters 1, 2, and 3 represent subjects less than 6 months after surgery, subjects between 6 months and 1 year of surgery and subjects without ACL injury (i.e. healthy subjects), respectively. Some of the data points lie on the boundary of second and third clusters which depicts the overlapping of groups such that some of the subjects belong to both clusters with certain membership grades and cannot be completely categorized into one group. These are the subjects who have recovered to a level where they are similar to the subjects in healthy group. This is natural as even after following the same rehabilitation protocol, the recovery may depend on individuals’ other physical parameters. After forming the clusters, the case generation step was performed by storing problem–solution pairs, corresponding treatments given/followed by the subjects and the recommendations from the physiatrists.

In order to build an efficient case retrieval model, three different classifiers (ANFIS, FURIA and SVM) were trained and tested for each activity based on the respective clustered data for healthy and ACL-R subjects, and their performance measures were compared. The classifiers models for ANFIS and SVM were developed using MATLAB 7.0 functions and Weka (3.7.9) was used for generating classifier models for FURIA. The cross validations of all models were done by partitioning the data into two groups: training data (70% of the total data) and test data (30% of the total data). The partitions for ANFIS and SVM classifiers were made by using cross validation function from MATLAB which randomly partitions the observations into a training set and a test set with stratification such that both training and test sets contain approximately the same class proportions as in the clusters’ groups (original groups). For FURIA, the partitioning task was implemented using Weka Experimentert where the randomized data split option was used for training and testing the model. The training/testing phase for all classifiers was repeated 100 times and the average values of different performance measures were computed.

The ANFIS networks were trained for 1000 epochs and the step size was set to an initial value of 0.01. The networks were generated using *genfis2* with *gaussmf* input membership functions. The parameters for initial membership functions were tuned using the training/testing of the networks and the examinations of final (after training) membership functions indicated that there were significant differences in the parameters for these membership functions for all activities. Fig. 10 shows the initial (before training) and final (after training) membership functions of two inputs (PCs) for walking at 5 km/h speed. Similar trends were found for other activities also. Both ANFIS and FURIA based classifiers generated the rules after training phase and these rules (models) were saved for further validation and retrieval of cases for new problems. However, it was found that the number of rules generated by FURIA based classifiers were, in general, much less than the rules generated by ANFIS models for different activities which indicates the efficient rule learning/pruning mechanism in FURIA. The number of rules generated by ANFIS and FURIA models for four different activities is shown in Table 1. Multi-class SVM models were also developed for different activities using three different kernel functions namely linear, quadratic and Radial Basis Function (RBF). The overall classification accuracies of all three types of models were more than 95% but linear-SVM model was found more accurate in classifying the health status of subjects as compared for quadratic- and RBF-based SVM models (Fig. 11).

The experiments showed that only first few PCs were enough to achieve high classification accuracy for all three types of classifiers. In general, it was found that first 3–7 PCs produced that maximum classification accuracy for all techniques for different activities. Fig. 11 shows a comparison of overall classification accuracy of three models for different activities. ANFIS based classifiers performed better than other two models for two walking activities (4 km/h and 5 km/h) but it poorly classified the data for balance testing activity. For walking at 6 km/h and balance testing activities, FURIA and SVM classifiers were found more effective and these models were able to classify the data for these two activities.

<table>
<thead>
<tr>
<th>Activity/technique</th>
<th>ANFIS</th>
<th>FURIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking at 4 km/h</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Walking at 5 km/h</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Walking at 6 km/h</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Single leg balance test</td>
<td>15</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 1: Number of rules generated by ANFIS and FURIA classifiers for different activities.
without any error (100% accuracy). The sensitivity, specificity and F-measure computed for each classification model for three groups (class 1 – subjects after less than 6 months of surgery, class 2 – subjects between 6 months and 1 year of surgery, class 3 – healthy subjects without ACL injury) are shown in Tables 2–4. The high values of these performance measures indicate that all three types of classifiers are suitable for classifying the recovery state/health condition of subjects after ACL reconstruction (Fig. 12).

The trained models and cases for each activity were stored and the second data set (data from 5 new subjects with known classes) was used for validation of the retrieval process and the framework. PCA coefficient matrices were applied on the given

Fig. 6. Percentage of variance explained by first 50 principal components for ambulation and balance testing activities.

Fig. 7. Original data projected on first three principal components.

Fig. 8. Clusters’ centers identified by FCM for walking at 5 km/h – cluster 1 (×) cluster 2 (o) cluster 3 (+).
data, and three classification models were used to identify the classes and retrieve the cases from the case repository. The classification performances of three models were compared (Fig. 13) and it was found that FURIA based classification model performed better than the other two models for data from new subjects. After classifying the recovery stage, the relevant cases were selected using case density function (17). The performance of each subject during each activity was compared with the most similar

<table>
<thead>
<tr>
<th>Activity</th>
<th>Class</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>F-measure (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking 4 km/h</td>
<td>1</td>
<td>99.66</td>
<td>99.75</td>
<td>99.54</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>98.98</td>
<td>99.75</td>
<td>99.19</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Walking 5 km/h</td>
<td>1</td>
<td>99.54</td>
<td>97.31</td>
<td>98.85</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>95.99</td>
<td>99.62</td>
<td>97.42</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Walking 6 km/h</td>
<td>1</td>
<td>99.32</td>
<td>100.00</td>
<td>99.65</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>100.00</td>
<td>99.70</td>
<td>99.64</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100.00</td>
<td>99.90</td>
<td>99.88</td>
</tr>
<tr>
<td>Single leg balance</td>
<td>1</td>
<td>79.83</td>
<td>94.17</td>
<td>82.81</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>87.00</td>
<td>95.29</td>
<td>87.96</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100.00</td>
<td>93.94</td>
<td>94.54</td>
</tr>
</tbody>
</table>

**Table 2**

ANFIS classification performance for all activities.
Table 3
FURIA classification performance for all activities.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Class</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>F-measure (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking 4 km/h</td>
<td>1</td>
<td>96.50</td>
<td>97.80</td>
<td>96.50</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>94.10</td>
<td>98.50</td>
<td>94.40</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>98.60</td>
<td>100.00</td>
<td>99.30</td>
</tr>
<tr>
<td>Walking 5 km/h</td>
<td>1</td>
<td>98.50</td>
<td>97.80</td>
<td>98.50</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>96.60</td>
<td>98.80</td>
<td>96.60</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Walking 6 km/h</td>
<td>1</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Single leg balance test</td>
<td>1</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Table 4
SVM classification performance for all activities.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Class</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>F-measure (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking 4 km/h</td>
<td>1</td>
<td>98.22</td>
<td>100.00</td>
<td>99.08</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>100.00</td>
<td>98.91</td>
<td>99.51</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Walking 5 km/h</td>
<td>1</td>
<td>89.79</td>
<td>100.00</td>
<td>94.44</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>100.00</td>
<td>93.09</td>
<td>97.69</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Walking 6 km/h</td>
<td>1</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Single leg balance test</td>
<td>1</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Fig. 12. Overall classification accuracy using ANFIS, FURIA and SVM (linear) for different activities monitored during recovery analysis.

Fig. 13. Overall classification accuracy for case retrieval using trained classifiers (ANFIS, FURIA and SVM) for new subjects to test the framework.

5. Conclusions and future work

A knowledge-based framework has been implemented using hybrid intelligent techniques which will help in diagnosis and prognosis of ACL-R subjects during recuperation. A case repository of healthy and various ACL-R subjects at different stages of recovery has been constructed using fuzzy clustering of integrated kinematics and EMG data for implementing CBR model. The framework was validated for a group of healthy and ACL-R subjects with very high retrieval accuracy. Three intelligent classifiers were trained and tested, and during training/testing phase FURIA and SVM models classified all instances of the data for ambulation at a speed of 6 km/h and balance testing activities with 100% accuracy. While for other two activities (walking at speeds of 4 km/h and 5 km/h), ANFIS classifiers performed slightly better than the other models. The trained classifiers were then applied to a group of subjects not included in the training and testing phase, and the results indicated that FURIA based classification model performed well for all activities. This framework can provide a valuable feedback to clinicians and physiotherapist in order to evaluate the rehabilitation status of the subjects after ACL injury/surgery and determine the potential solutions to the knee joint and neuromuscular problems during rehabilitation. Moreover, maintaining the athletes’ pre-injury, post-injury and post surgery profiles can provide a comprehensive platform for ACL injury prevention and sports performance evaluation. The system will be expanded in future by including data retrieved cases by using their recommendations and next stage results.

This study demonstrates the successful development of a knowledge-based framework for rehabilitation and performance monitoring of ACL-R subjects using hybrid intelligent techniques (CBR and intelligent clustering and classification algorithms). Anterior cruciate ligament injury alters the human locomotion and results in various short- and long-term effects including loss of proprioception, cartilage degeneration and early onset of osteoarthritis. Although, through ACL reconstruction certain improvements in the knee dynamics are achieved but the changes in the kinematics and neuromuscular parameters have been reported to persist in anterior cruciate ligament reconstructed (ACL-R) subjects for several months after surgery. A comprehensive system based on CBR methodology and hybrid intelligent techniques can be used to identify the need for timely intervention in the current rehabilitation protocol of new subjects and making adjustments as per the requirements, thus minimizing the risk of re-injury. The knowledge-based system can store the pre-injury, post-injury and recovery data in order to provide a detailed analysis about the recovery of subjects and a comparison of current and past performance in any specific activity. The collection of relevant existing experiences for rehabilitation can help in designing a personalized rehabilitation protocol as well as improving the generalized rehabilitation process.

Three different types of classifiers (ANFIS, FURIA and SVM) were investigated and the results show that for retrieving similar cases (i.e. differentiating the health/recovery stage of different subjects), the performance of FURIA based model is slightly better than the other two models (Fig. 13). Various statistical features were extracted from the kinematics and electromyography data for ambulation and balance testing activities and these features provided useful information for developing accurate classifiers. The integration of kinematics and electromyography data has been proved quite effective in identification of health condition of different subjects. The data processing of stochastic and non-stationary EMG signals was done by using DWT analysis which provided reliable features in order to form data clusters and designing accurate classifiers for case selection and retrieval.
Acknowledgments

This work is supported by the University Research Council (URC) grant scheme at the Universiti Brunei Darussalam under the grant no. UBD/PNC2/2/RG/1(195) with the title “Integrated Motion Analysis System (IMAS).” The authors would like to acknowledge the support provided by Ministry of Sports and Ministry of Defense, Brunei Darussalam for providing national athletes as test subjects for this study.