

Enabling Stroke Rehabilitation in Home and Community Settings: A Wearable Sensor-Based Approach for Upper-Limb Motor Training

Original

Enabling Stroke Rehabilitation in Home and Community Settings: A Wearable Sensor-Based Approach for Upper-Limb Motor Training / Lee, Sunghoon I; Adans-Dester, Catherine P; Grimaldi, Matteo; Dowling, Ariel V; Horak, Peter C; Black-Schaffer, Randie M; Bonato, Paolo; Gwin, Joseph T. - In: IEEE JOURNAL OF TRANSLATIONAL ENGINEERING IN HEALTH AND MEDICINE. - ISSN 2168-2372. - 6:(2018), pp. 2100411-11. [10.1109/JTEHM.2018.2829208]

Availability:

This version is available at: 11583/2712380 since: 2018-09-06T15:51:03Z

Publisher:

IEEE

Published

DOI:10.1109/JTEHM.2018.2829208

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

Received 24 September 2017; revised 16 January 2018; accepted 28 March 2018.
Date of publication 2 May 2018; date of current version 11 May 2018.

Digital Object Identifier 10.1109/JTEHM.2018.2829208

Enabling Stroke Rehabilitation in Home and Community Settings: A Wearable Sensor-Based Approach for Upper-Limb Motor Training

SUNGHOON I. LEE¹, (Member, IEEE), CATHERINE P. ADANS-DESTER^{2,3}, MATTEO GRIMALDI²,
ARIEL V. DOWLING⁴, (Member, IEEE), PETER C. HORAK⁵, RANDIE M. BLACK-SCHAFER²,
PAOLO BONATO², (Senior Member, IEEE), AND JOSEPH T. GWIN⁴, (Member, IEEE)

¹College of Information and Computer Sciences, University of Massachusetts, Amherst, MA 01003, USA

²Department of Physical Medicine and Rehabilitation, Harvard Medical School, Spaulding Rehabilitation Hospital, Charlestown, MA 02129, USA

³School of Health and Rehabilitation Sciences, MGH Institute of Health Professions, Charlestown, MA 02129, USA

⁴BioSensics, LLC, Watertown, MA 02472, USA

⁵Rensselaer Polytechnic Institute, Troy, NY 12180, USA

CORRESPONDING AUTHOR: S. I. LEE (silee@cs.umass.edu)

This work was supported by the National Center for Med Rehab Res, National Institute for Child Health and Human Development under Grant R44HD084035.

ABSTRACT High-dosage motor practice can significantly contribute to achieving functional recovery after a stroke. Performing rehabilitation exercises at home and using, or attempting to use, the stroke-affected upper limb during Activities of Daily Living (ADL) are effective ways to achieve high-dosage motor practice in stroke survivors. This paper presents a novel technological approach that enables 1) detecting goal-directed upper limb movements during the performance of ADL, so that timely feedback can be provided to encourage the use of the affected limb, and 2) assessing the quality of motor performance during in-home rehabilitation exercises so that appropriate feedback can be generated to promote high-quality exercise. The results herein presented show that it is possible to detect 1) goal-directed movements during the performance of ADL with a *c*-statistic of 87.0% and 2) poorly performed movements in selected rehabilitation exercises with an *F*-score of 84.3%, thus enabling the generation of appropriate feedback. In a survey to gather preliminary data concerning the clinical adequacy of the proposed approach, 91.7% of occupational therapists demonstrated willingness to use it in their practice, and 88.2% of stroke survivors indicated that they would use it if recommended by their therapist.

INDEX TERMS Machine learning, m-health, rehabilitation, remote health monitoring, stroke, wearable sensors, wearable technology.

I. INTRODUCTION

Stroke is a leading cause of severe long-term disability. In the US alone, nearly 800,000 people suffer a stroke each year [1]. The number of individuals who suffer a stroke each year is expected to rise in the coming years because the prevalence of stroke increases with age and the world population is aging [2]. Approximately 85% of individuals who have a stroke survive, but they often experience significant motor impairments. Upper-limb paresis is the most common impairment following a stroke. It affects 75% of stroke survivors and leads to limitations in the performance of Activities of Daily Living (ADL) [4].

Inability to use the stroke-affected upper limb for ADL often leads to a phenomenon that is referred to as *learned non-use* [5]. As patients rely more and more on the unaffected (or less impaired) upper limb [5] they progressively lose motor abilities of the stroke-affected upper limb that they may have recovered as a result of a rehabilitation intervention [6].

A high dosage of motor practice using the stroke-affected upper limb during the performance of ADL, despite considerable difficulty, stimulates neuroplasticity and motor function recovery [7]–[9]. Thus, it is clinically important to encourage stroke survivors to continue making appropriate use of the affected upper limb [10]–[13], in addition to engaging in

rehabilitation exercises that focus on range-of-motion and functional abilities [14]–[16].

The use of wearable sensors has recently emerged as an efficient way to monitor the amount of upper-limb use after a stroke [17]–[22]. However, despite growing evidence of the clinical potential of these devices [23], their widespread clinical deployment has been hindered by technical limitations. A shortcoming of currently available wrist-worn devices is that they cannot distinguish between Goal-Directed (GD) movements (i.e., movements performed for a specific purposeful task) and non-Goal-Directed (non-GD) movements (e.g., the arm swinging during gait). Instead, these sensors focus on recording the number and/or intensity of any type of arm movements [10]. Consequently, non-GD movements are reflected as part of the measurements with equal importance as GD movements. This results in an overestimation of the amount of actual arm use [24]. Furthermore, monitoring the aggregate number of stroke-affected upper limb movements is not sufficient for the purpose of providing timely feedback to encourage the use of the affected limb during the performance of ADL. To promote the use of the stroke-affected limb, it is critical that feedback reflects the relative use of the affected upper limb compared to the contralateral one.

Wrist-worn movement sensors have also been applied to monitoring rehabilitation exercises in the home setting [25]–[28]. However, existing systems primarily focus on quantifying the dosage/intensity of the exercises (e.g., the duration of the exercises and the number of movement repetitions) and do not monitor if the quality of the performed exercise is appropriate. Ensuring good quality of movement during the performance of rehabilitation exercises is critical for maximizing functional recovery after a stroke [29]. Moreover, providing customized feedback regarding the quality of exercise movements can increase motivation, promote long-term adherence to a prescribed exercise regimen, and ultimately maximize clinical outcomes [30]. One of the reasons for limited exercise participation by stroke survivors is the lack of access to resources to support exercise including performance feedback from rehabilitation specialists [31]. There are no technical solutions that provide feedback regarding the quality of exercise performance for upper-limb rehabilitation after stroke.

We propose a system for aiding in functional recovery after a stroke that consists of two wearable sensors, one worn on the stroke-affected upper limb and the other on the contralateral upper limb [32] (Fig. 1). The proposed system can be used to provide timely feedback when ADL are performed. If the system detects that the patient consistently performs GD movements with the unaffected upper limb, and rarely uses the stroke-affected upper limb, then a visual or vibrotactile reminder can be triggered to encourage the patient to attempt GD movements with the stroke-affected limb. A benefit of this approach is that if a movement is critical (e.g., signing a check), patients can use the unaffected upper limb without receiving negative feedback as long as they have performed a sufficient number of movements with the

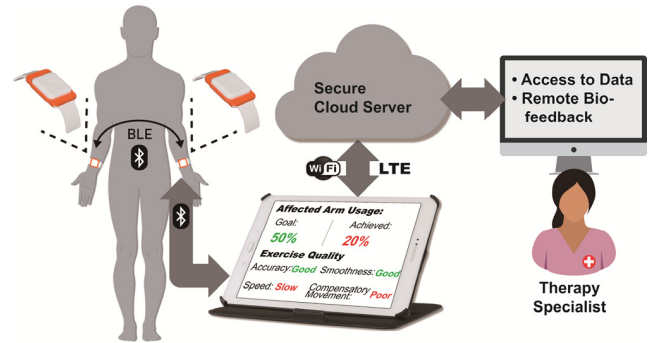


FIGURE 1. A conceptual representation of the wrist-worn sensor system for home-based upper-limb rehabilitation. The system consists of two wearable sensors, a tablet computer to be used by patients at home, and a backend web-service for clinician and caregiver data access.

affected upper limb throughout the day. Furthermore, the system promotes high-dosage motor practice with appropriate feedback to extend components of rehabilitation interventions into the home environment.

This paper is focused on the development of algorithms that form the foundation of the proposed system. These algorithms detect GD movements during ADL, and determine appropriate feedback during in-home rehabilitation exercise. This feedback is designed to be similar to how therapists would provide feedback to improve the quality of upper-limb movements during in-clinic exercise. The proposed algorithms extend previous work in which we showed that inertial data (e.g., acceleration and rotational velocity) recorded during the performance of a battery of functional motor tasks using sensors positioned on the upper body can be used to derive clinically-meaningful indicators of functional level and the severity of motor impairments [32]–[35]. In addition, this paper presents an evaluation of the appropriateness of the envisioned technological approach for the intended audience, including both stroke survivors and occupational therapists. This evaluation was carried out using an anonymous survey obtained from focus groups. The survey consisted of questions concerning whether clinicians would be interested in integrating the proposed system in their clinical practice and whether patients would be willing to adopt the proposed technology.

II. MATERIALS AND METHODS

A. DATA COLLECTION

We recruited 20 stroke survivors (54.4 ± 10.1 years old; average and standard deviation) and 10 aged-matched control subjects (53.8 ± 11.4 years old). Stroke survivors were recruited from Spaulding Rehabilitation Hospital (SRH) inpatient and outpatient units. Stroke survivors had chronicity of 4.6 ± 5.5 years and showed mild-to-moderate upper-limb motor impairments as evaluated using the upper-limb Fugl-Meyer Assessment (FMA) (patients' scores were 37.0 ± 8.0 out of 66 points) [36]. All subjects provided written informed consent. The SRH Institutional Review Board (IRB) approved the experimental procedures.

TABLE 1. Upper-limb motor tasks that resembled different types of activities of daily living.

ADL Type	Tasks
Unimanual (affected limb)	<ul style="list-style-type: none"> • Drinking from a can¹ • Turning a key in a lock¹ • Brushing hair²
Bimanual	<ul style="list-style-type: none"> • Pick up a pen from a tabletop, remove the cap, and place it back • Pick up a box on a table and bring it to the knees³ • Folding a hand towel¹
Stabilization (with affected limb)	<ul style="list-style-type: none"> • Stabilize a bottle while opening it⁴ • Hold a plastic container while removing the lid⁵ • Cutting a piece of meat in half using cutlery⁶
Passive	<ul style="list-style-type: none"> • Walking • Stand-to-sit transitions without using arm for bracing • Ascending and descending stairs

¹ Task taken from the Wolf Motor Function test [3]. ² Subjects were asked to pick up a comb and brush their hair from the front to the back of the head. ³ The box used in this exercise was 30x40x20 cm and was unloaded. ⁴ A 0.5L water bottle with a standard rotational cap was used. ⁵ A Rubbermaid food container was used. ⁶ Playdough was used to mimic a piece of meat, which was placed at the center of a plate. Subjects were asked to use a fork with the affected hand and a knife with the contralateral hand.

TABLE 2. Upper-limb rehabilitation exercises that are typical of home-based interventions.

Exercise Type	Tasks
Strength	<ul style="list-style-type: none"> • Arm raise in the sagittal and coronal planes • Elbow flexion/extension with weights • Forearm pro-supination with weights • Wrist flexion/extension with weights • Thera-band pulls
Range of Motion	<ul style="list-style-type: none"> • Bring the affected hand to the contralateral ear • Supine arm raises (towards ceiling) • Elbow flexion/extension • Forearm pro-supination • Wrist flexion/extension
Functional	<ul style="list-style-type: none"> • Flip a patty with a spatula • Reach for objects on a shelf¹ • Whisk material in a pot² • Scratching the back³ • Turn the pages of a magazine⁴

¹ Objects were placed above shoulder height. ² Subjects were asked to use a whisker in a medium-sized saucepan using the affected hand.

³ Subjects were asked to reach the lower back with the affected hand.

⁴ Subjects were asked to perform full-range pro-supination movements of the forearm.

When study participants arrived at the laboratory they were instrumented with six-axis inertial measurement units (three-axis accelerometer and gyroscope; Shimmer Research, Ireland) bilaterally on the wrist. Stroke survivors were first asked to perform motor tasks associated with the FMA for the evaluation of their upper-limb motor impairments. Then, both stroke survivors and control participants performed a battery of motor tasks resembling different types of ADL and rehabilitation exercises that could be performed independently at home, as summarized in Tables 1 and 2, respectively.

All motor tasks, except for the passive movements listed in Table 1, were performed while sitting in an armless chair in front of a table. Stroke survivors performed a subset of these motor tasks and exercises (i.e., a minimum of eight or more motor tasks and five exercises) that were carefully chosen by a therapist considering the subject's motor abilities. Age-matched healthy controls performed the complete set of motor tasks and exercises. Each motor task and exercise was repeated three to five times, depending on the patient's functional capability, in order to capture intra-subject variability in the movement patterns. All tasks were scripted and timed by a therapist. The sensor data were wirelessly streamed to a base-station (i.e., laptop) via Bluetooth at a sampling rate of 256 Hz throughout the entire experiment. A technician marked the beginning and the end of each repetition with a digital marker. The laboratory session was videotaped. Video recordings and sensor data were time-synchronized for offline analysis.

B. DATA ANNOTATION AND LABELING

1) GD VS. NON-GD MOVEMENTS

Data collected from the unaffected upper limb of stroke survivors and both limbs of control subjects were categorized as associated with either GD or non-GD movements by visual inspection of the time-synchronized video files. The performance of FMA tasks, non-passive ADL-related movements, and upper-limb exercises involving the target limbs were labeled as GD. All other passive movements, such as arm swing during gait, resting, and miscellaneous hand gestures during talking, were labeled as non-GD.

2) PERFORMANCE OF REHABILITATION EXERCISES

An experienced therapist reviewed the video recordings collected during the performance of rehabilitation exercises. The therapist provided feedback regarding the quality of the performed exercises with particular emphasis on accuracy and presence of compensatory movements. The inertial data collected during the performance of the rehabilitation exercises was labeled *Feedback* (if the quality of movement was not adequate) or *No Feedback* (if the movement quality was adequate) according to the therapist's opinion. In other words, the data was labeled according to how patients would have been provided with feedback based on the observed quality of movement in a regular rehabilitation session. The data was further annotated using binary labels reflecting the type of feedback generated by the therapist (i.e., *Accuracy* and/or *Compensatory Movement*). The exercise movements performed by the age-matched controls were also reviewed to make sure that they were performed appropriately. The sensor data collected from control subjects was labeled as *Control*.

C. STAKEHOLDER SURVEY

A total of 17 chronic stroke survivors and 13 occupational therapists (4 practicing in the inpatient unit and 9 in the outpatient unit at SRH) volunteered to provide feedback regarding



FIGURE 2. Eight off-the-shelf wristbands presented to the stakeholders that provide different donning/doffing mechanisms, e.g., a regular watch bracelet, elastic bracelet, and slap bracelet. Participants were asked to try all wristbands and identify the ones that they could don/doff with minimal effort.

the clinical appropriateness of the proposed system and their willingness to use it for home-based rehabilitation. Stroke survivors were recruited via the SRH Stroke Support Group. Therapists were recruited by word of mouth. All sessions were held at SRH and lasted about 60 minutes.

The stroke survivors and the therapists participated in separate focus-group sessions. During each session, participants were provided with an overview of the proposed system via a PowerPoint presentation and demonstration of a *looks-like* prototype of the sensors. Research staff addressed questions that were raised during the demonstration. Then, participants were asked to fill out an anonymous questionnaire regarding the system's appropriateness for home-based rehabilitation.

Stroke survivors were asked questions about their willingness to use the system, if the system would be beneficial to them, if they would use the system when recommended by a therapist, if they liked the idea of receiving messages throughout the day to remind them to use their stroke-affected upper limb, and their preference for the reminder/feedback mechanism (visual, vibration, sound, or a combination of vibration and sound).

Therapists were asked if they would be willing to use the proposed system in their clinical practice, if they believed the proposed system would encourage patients to perform home-based exercises, if they would want to receive email alerts summarizing their patients' activity and stroke-affected upper limb use, and if they would want patients to receive reminders throughout the day to encourage the use of their stroke-affected upper limb.

Furthermore, both groups were presented with eight different types of wristband that provided different donning/doffing mechanisms for self-application of the sensor, which is particularly important for stroke survivors with limited upper-limb functional abilities. Stakeholders' feedback was gathered with the objective of identifying a wristband mechanism that would maximize the usability and long-term adherence to the technology. The tested wristbands included

1) the FitBit Flex, 2) a generic slap bracelet, 3) the Polar Loop, 4) the Moov Now, 5) the Jawbone UP, 6) the Apple Milanese Loop, 7) the Pebble Time Steel, and 8) a generic Velcro wrist band (Fig. 2). Participants were asked to try all wristbands and wear them tightly on their contralateral wrist, and identify the ones that they could don/doff with minimal effort.

D. DATA ANALYSIS

1) DETECTION OF GD MOVEMENTS

Fig. 3a shows a schematic representation of the data analysis pipeline that we designed to categorize movements as GD or non-GD, with the eventual goal of triggering feedback to motivate stroke survivors to use their affected limb.

Tri-axial acceleration time-series were preprocessed and data features relevant to detecting GD and non-GD movements were extracted. It is worth emphasizing that the detection of GD movements relied only on the accelerometer data. This choice was motivated by the observation that accelerometers use significantly less power than gyroscopes (microwatts vs. milliwatts) and are therefore better suited for long-term, continuous monitoring. Also, prior studies support the feasibility of accurate assessment of upper-limb movements using accelerometer data [19], [37], [38]

The orientation of the sensors was estimated by low-pass filtering the acceleration time-series with a cut-off frequency of 0.25 Hz. The orientation time-series was subtracted from the unfiltered signal to derive an approximation of the inertial component of the acceleration, which was again low-pass filtered with a cut-off frequency of 10 Hz to attenuate high-frequency noise. Velocity and displacement time-series were computed by trapezoid-integrating the acceleration data. After integration, the signals were high-pass filtered with a cut-off frequency of 0.25 Hz to attenuate the effect of integration drift. Next, we computed the magnitude (i.e., root mean square of the x, y, and z components) of the acceleration, velocity, displacement, and jerk vectors, as well as the inner products of each pair of acceleration

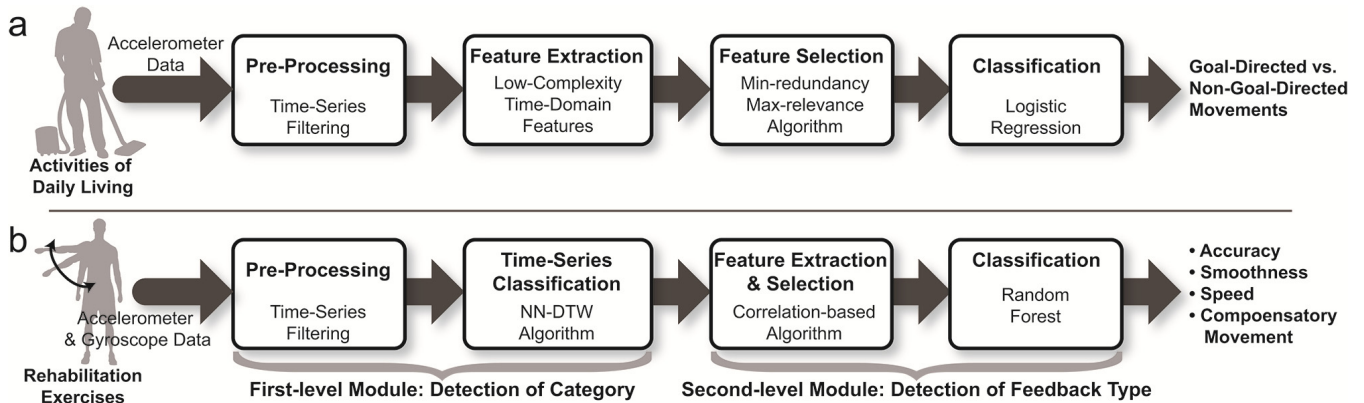


FIGURE 3. Data analysis pipeline for a) detecting GD vs. non-GD movements during the performance of ADL and b) generating appropriate feedback during the performance of rehabilitation exercises to be prescribed for home-based therapy.

signals (i.e., x-y, x-z, and y-z axes). We used the inner product as a lower computational cost surrogate for the Singular Value Decomposition (SVD) as previously proposed by Hong *et al.* [39] in gesture classification. This resulted in a total of 19 time-series (i.e., jerk, accelerometer, velocity, and displacement for the three axes and their magnitude, and the x-y, x-z, and y-z axes inner products of the acceleration) to be processed to extract data features. The signals were segmented using a 2.5 s sliding window without any overlap. Each segment was annotated as either GD or non-GD.

Data features that we anticipated to be relevant to the detection of GD movements were extracted from each of the above-described time-series. The data features included the minimum, maximum, range, mean, standard deviation, root-mean-square values, and the number of zero crossings of the time-series. These data features have been used in prior work to classify different activities [33], [35], [40]. We employed a feature selection algorithm – the minimal-redundancy maximal-relevance algorithm [41] – to identify the data features that were most relevant to classifying GD vs. non-GD movements. This algorithm ranks data features based on maximum correlation with the movement categories (relevance) and minimum correlation with each other (redundancy). The top 15 data features were fed to the classification algorithm described below.

We trained a logistic regression classification model to distinguish GD from non-GD movements. The performance of the classifier was assessed using the leave-one-subject-out cross validation technique [42]. This technique evaluates the performance of an algorithm by selecting the data belonging to each subject (one at the time) as the test set and by training the algorithm using the data belonging to the remaining subjects. This approach avoids problems of overfitting and provides fair estimations of the expected classification accuracy. The overall classification performance was evaluated based on the Area Under the Curve (AUC) of the Receiver Operating Characteristics (ROC) curve, which

describes the tradeoff between sensitivity and specificity. The ROC AUC, which is also known as the *c*-statistic [43], reflects the predictive accuracy of a classifier [44]. The ROC curve was derived by varying from 0 to 1 the probability threshold for detecting the positive class (i.e., GD movements). The *c*-statistic ranges from 0.5 (unable to classify) to 1.0 (able to perfectly classify), where 0.8 represents a good classification ability.

2) DETERMINATION OF APPROPRIATE EXERCISE FEEDBACK

Fig. 3b shows a schematic representation of the data analysis pipeline that we designed to generate appropriate feedback regarding the quality of rehabilitation exercise performance. To assess feasibility, this analysis was performed on data from the exercise that was prescribed to the largest number of participating stroke survivors, which was the arm raise in the sagittal and coronal planes. Eleven stroke survivors performed 98 trials of this exercise. 17 of these trials were labeled as *No Feedback* and 81 as *Feedback*. All 11 stroke survivors required feedback during at least one trial. In addition, 9 age-matched healthy controls performed a total of 54 trials of this exercise.

The analysis used both accelerometer and gyroscope data. This choice was made based on previous findings that gyroscope data can significantly contribute to the assessment of movement quality [45]–[48]. In the context of evaluating exercise quality, the high power consumption of the gyroscopes would have a minimal impact on the overall power performance of the system because the home-based exercises are performed over a relatively short period of time (e.g., less than an hour). The magnitude of the accelerometer and gyroscope data (rather than its components) was used to minimize the dependence on the initial orientation of the sensing unit. The algorithm consisted of a cascade of two modules. The first module focused on classifying if a dataset belonged to the *Control*, *No Feedback*, or *Feedback* category. The second module detected the appropriate type of feedback within the *Feedback* category.

The detection of the category of each exercise was formalized as a time-series classification problem because the rehabilitation exercises considered in this work involve a pre-determined, simple movement that produced unique acceleration and gyroscope time-series patterns. This work employed the Nearest Neighbor (NN) algorithm with Dynamic Time Warping (DTW) for time-series classification. The NN-DTW algorithm with a sampling mechanism that effectively reduces the cardinality of the training set is known to allow accurate, fast, and computationally-efficient classification [49]. This is optimal for resource-constrained devices such as our wrist-worn sensors. NN-DTW has been used in gesture recognition and clinical event detection [37], [50], [51].

A preliminary analysis of the data showed that the accelerometer and gyroscope data in the *Control* cluster exhibited similar time-series patterns and formed a highly dense cluster; i.e., the mean μ_c and standard deviation σ_c of the time-warped distances among the *Control* time-series were extremely small. The *Feedback* and *No Feedback* clusters surrounded the *Control* cluster. A dataset with such a pattern is known as a Japanese flag dataset [49], [52]. Based on this observation, the algorithm was designed with two steps. First, it checked if a time-series belonged to the *Control* cluster. Then, if the time-series was classified as *Non-Control*, the algorithm performed the conventional K -NN algorithm between *No Feedback* and *Feedback* classes by comparing the number of NN instances belonging to the two classes.

Model training and performance evaluation was again performed using the leave-one-subject-out cross validation technique [42]. The algorithm classified a testing time-series as *Control* (or belongs to the *Control* cluster) by verifying that 1) the majority of NN of the testing time-series contained instances from the *Control* cluster and 2) the distances between the testing time-series and its NN instances are small enough to ensure that the testing time-series is indeed located within the boundary of the *Control* cluster. First, the algorithm performed a DTW between the testing dataset and all time-series in the training dataset, and selected K nearest neighbors for the accelerometer and gyroscope data separately, producing a total of $2 \times K$ neighbors. The K value was chosen as 10% of the cardinality of the training dataset. Then, the algorithm computed the number of *Control* instances (among the combined NN instances) that had DTW distances within the range of $\mu_c \pm 2\sigma_c$ (i.e., the boundary) of the training data. If such instances formed the majority of the NN instances, the testing time-series was classified as *Control*, otherwise as *Non-Control*.

If the testing time-series was classified as *Non-Control*, the algorithm then compared the number of NN instances belonging to the *No Feedback* and *Feedback* classes. Due to the imbalanced number of training instances in these two classes (i.e., the number of *Feedback* instances was three times greater than that of *No Feedback* instances in our dataset), the algorithm balanced the numbers of the *No Feedback* and *Feedback* instances in the NN instances using

weights ω_{NF} and ω_F :

$$\hat{n}_{NF}^{TR} = \omega_{NF} \cdot n_{NF}^{TR} \quad \text{and} \quad \hat{n}_F^{TR} = \omega_F \cdot n_F^{TR}.$$

where n_{NF}^{TR} and n_F^{TR} represent the numbers of *No Feedback* and *Feedback* instances in the training set, respectively. These weights were defined based on a *sigmoid* function that took the number of instances of the corresponding class in the training set as its input. The input to the sigmoid function was normalized such that the values of the weights were inversely proportional to the number of the instances per class. The weights were defined as:

$$\omega_{NF} = \frac{-1}{1 + \exp\left(\frac{n_{NF}^{TR}}{n_{NF}^{TR} + n_F^{TR}} - 0.5\right)} + 1$$

$$\omega_F = \frac{-1}{1 + \exp\left(\frac{n_F^{TR}}{n_{NF}^{TR} + n_F^{TR}} - 0.5\right)} + 1$$

The testing time-series was classified according to a majority rule. If the testing time-series was classified as *Feedback*, a data feature-based classification algorithm was implemented to identify the type of feedback (i.e., related to lack of accuracy and/or presence of compensatory movement). The method to identify feedback type consisted of a data pre-processing module, a data feature extraction module, a data feature selection module, and a Random Forest algorithm to determine the feedback type.

The data pre-processing method discussed in Section II-D.1 was employed to derive the magnitude of the linear jerk, linear acceleration, linear velocity, linear displacement, angular velocity, and angular displacement from the accelerometer and gyroscope time-series. Then we derived the following data features that reflected the speed, smoothness, and coordination of movement: 1) mean value, 2) root mean square value, 3) maximum and minimum amplitude, 4) skewness, 5) entropy, 6) kurtosis, 7) time it took to complete the movement and to reach the maximum amplitude, and 8) average DTW distance from the *Control* time-series in the training dataset. These features were previously found to be relevant to assessing movement quality [33]–[35]. A correlation-based feature selection algorithm was used to select data features that were relevant to detecting feedback type [53]. Then, a Random Forest with 100 trees was implemented as a binary classification model [42] to identify the exercise feedback type, and a leave-one-subject-out cross validation was performed. Classification performance was evaluated based on precision and recall (i.e., sensitivity). Precision is the percentage of Feedback events detected by the algorithm that were really Feedback events (i.e., the therapists gave feedback for that trial). In other words, this is the percentage of Feedback events reported by the algorithm that are correct. Recall is the percentage of real Feedback events that were detected by the algorithm. In other words, this is the percentage chance that a trial necessitating feedback will be detected

by the algorithm. Taken together, precision and recall are widely used measures to describe the overall performance of a classification algorithm. Overall performance was evaluated using the F -score, which is the harmonic mean of precision and recall [54]. Its value spans the range between 0 (no precision and recall) and 1 (perfect precision and recall).

III. RESULTS

A. DETECTION OF GD MOVEMENTS

A total of 262 movements (220 GD and 42 non-GD) were collected from 10 control subjects, and 666 movements (367 GD and 299 non-GD) were collected from 20 stroke survivors. Segmenting the movements into 2.5 s windows produced 414 GD and 185 non-GD data points for the control subjects, and 2475 GD and 1325 non-GD data points for the stroke survivors.

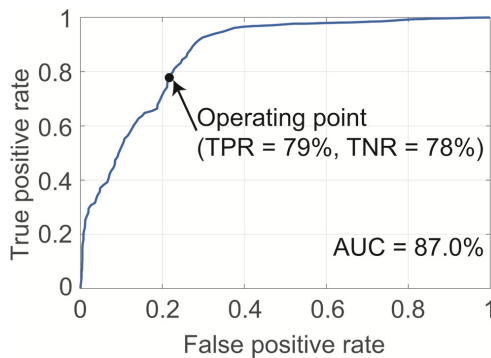


FIGURE 4. The ROC curve for the detection of GD movements. The area under the curve that indicates the predictive accuracy of the classifier was 87.0%. The highlighted operating point on the curve indicates a true positive rate of 79% and true negative rate of 78%.

Fig. 4 shows the ROC curve based on the classification results generated by the logistic regression classifier using the leave-one-subject-out cross validation technique. The c -statistic (ROC AUC) was 87.0%, thus indicating a strong classification ability. The operating point highlighted on the curve shown in Fig. 4 is associated with a true positive rate of 79% and true negative rate of 78%.

B. DETERMINATION OF EXERCISE FEEDBACK

The exercise that was prescribed to the largest number of participating stroke survivors was the “arm raise in the sagittal and coronal planes” (Table 2). Eleven stroke survivors performed 98 trials of this exercise. 17 of these trials were labeled as *No Feedback* and 81 as *Feedback*. All 11 stroke survivors required feedback on at least one trial whereas only 3 stroke survivors performed a trial that did not require feedback. In addition, 9 age-matched healthy controls performed a total of 54 trials of this exercise.

Fig. 5 shows confusion matrices for the first classification stage (i.e., detecting whether the performed exercise requires feedback). Fig. 5a shows the confusion matrix for classification of *Feedback*, where *No Feedback* and *Control* are grouped together. The precision and recall for detecting

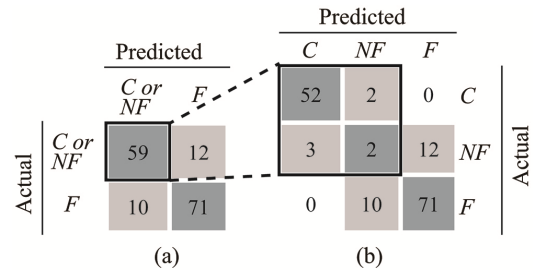


FIGURE 5. Confusion matrices for the first-level time-series classification when considering a) two classes (*Control* or *No Feedback* vs. *Feedback*) and b) three classes (*Control*, *No Feedback*, *Feedback*).

Feedback were 85.5% and 83.1%, respectively. This indicates that 85.5% of the movement instances that were classified as *Feedback* were true positives. Furthermore, 83.1% of the instances that were labelled as *Feedback* were successfully detected by the proposed analytic method. The computed F -score (the harmonic mean of precision and recall) was 84.3%. Fig. 5b shows the confusion matrix considering all three classes (i.e., *Control*, *No Feedback*, and *Feedback*). The average precision and recall were 64.8% and 65.2% with an F -score of 63.3%.

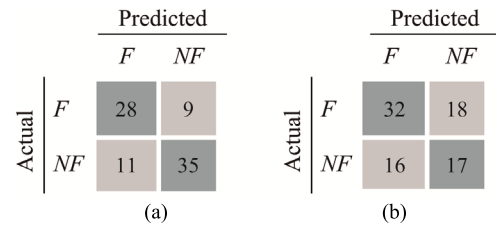


FIGURE 6. Confusion matrices for the second-level feature-based classifiers for detecting feedback regarding a) the accuracy of the movement and b) the use of compensatory movement. The F -scores for the accuracy and compensatory movement feedbacks were 73.7% and 65.3%, respectively.

Within the *Feedback* group, there were 83 data instances, of which 71 were true positives and 12 were false positives. Fig. 6 shows confusion matrices for the second classification stage (i.e., determining the appropriate feedback type within these 83 data instances). Fig. 6a shows the confusion matrix for detecting *Accuracy* feedback. The precision and recall were 71.8% and 75.7%, respectively, with an F -score of 73.7%. Fig. 6b shows the confusion matrix for detecting *Compensatory Movement* feedback. The precision and recall were 66.7% and 64.0%, respectively, with an F -score of 65.3%.

C. STAKEHOLDER QUESTIONERS

Fig. 7 shows a summary of the survey results obtained from stroke survivors regarding the use of the proposed wearable system. 76.5% of stroke survivors indicated their willingness to use the system and the reminders generated by the system during the day to encourage the use of the stroke-affected upper limb. 82.3% expressed confidence that the

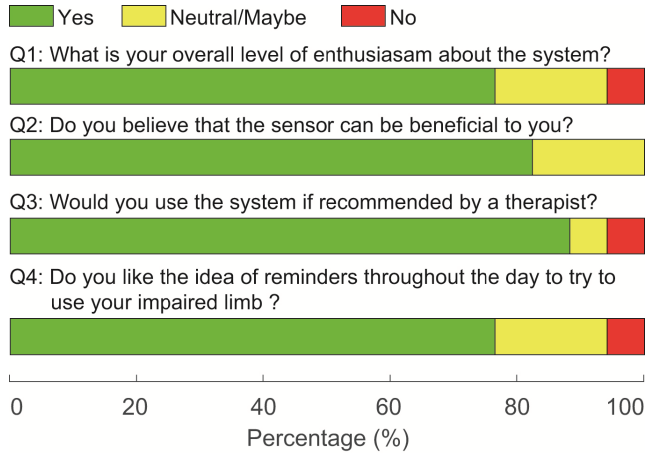


FIGURE 7. Survey results obtained from stroke survivors regarding the translational feasibility for at-home rehabilitation.

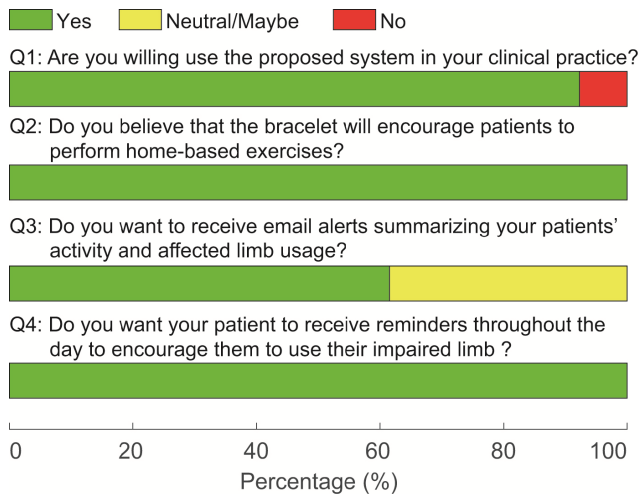


FIGURE 8. Survey results obtained from therapists regarding the translational feasibility for at-home rehabilitation.

system would help increase the use of their stroke-affected upper limb. 88.2% indicated that they would use the system if recommended by their therapist. Results regarding the preferred method of receiving feedback were equivocal; 29.4% preferred a visual message, 11.8% preferred a sound, 23.5% preferred a vibration, and 35.3% preferred a combination of vibration and sound.

Fig. 8 shows a summary of the survey results obtained from the therapists. 91.7% of the therapists reported that they would be willing to use the system in their clinical practice. All therapists (100%) thought that the device would help encourage patients to perform home-based exercises. 61.5% reported that they would want to receive periodic email alerts about their patient's performance, and the rest (i.e., 38.5%) were undecided. All therapists (100%) expressed interest in a system that can provide patients with reminders throughout the day to encourage them to use their stroke-affected upper limb.

TABLE 3. A summary of stakeholder preferences for wristband mechanisms in terms of its ease-of-use under limited upper-limb functionality. The numbers represent the percentage of responders (i.e., patients and therapists) who preferred the associated donning/doffing mechanism for the presented technology.

	Fitbit Flex	Slap Bracelet	Polar Loop	Moov Now
Patients	5.9%	52.9%	11.8%	0%
Therapists	0%	61.5%	23.1%	0%
	Jawbone Up	Apple Milanese Loop	Pebble Time Steel	Velcro
Patients	47.1%	11.8%	23.5%	11.8%
Therapists	92.3%	38.5%	0%	7.7%

Table 3 provides a summary of stakeholder preferences for different wristbands regarding their suitability for self-application to the wrist. The slap bracelet and Jawbone UP were preferred by both patients and therapists, whereas the FitBit Flex and Moov Now were not deemed appropriate.

IV. DISCUSSION AND CONCLUSIONS

We have presented a novel technological approach that utilizes two wearable sensors (i.e., one worn on the stroke-affected upper limb and the other on the contralateral upper limb) for detecting GD movements during ADL and for determining appropriate feedback during in-home rehabilitation exercise (based on the sensor on the stroke-affected limb). Furthermore, we have demonstrated the feasibility of two algorithms that form the foundation of the system, with the eventual goal of providing timely reminders to encourage the use of the stroke-affected limb during ADL and appropriate feedback to promote high-quality home-based rehabilitation exercise. The presented results show that it is possible to detect GD movements during the performance of ADL with a c -statistic of 87.0%, and incorrectly-performed home-based rehabilitation (arm raise) exercises with an F -score of 84.3%. These results indicate that movements can be successfully classified into either GD or non-GD movements with a predictive accuracy of 87.0%, and the incorrectly-performed exercise movements with an accuracy of 84.3%, thus enabling the generation of appropriate feedback.

Reminder-based approaches to ADL modification should provide timely feedback. In our case, this means providing a reminder to use (or attempt to use) the stroke-affected limb at the time of a movement with the contralateral limb. Such approaches must also provide relevant feedback and avoid alarm fatigue. In our case, this means not providing feedback every time a contralateral limb movement is detected. Instead, we propose to trigger feedback at the time a GD movement is detected based on the relative usage of the affected compared to the contralateral limb. For example, if the patient consistently performs GD movements with the contralateral upper limb, and rarely uses the stroke-affected upper limb, then a reminder is triggered. On the other hand, if the patient has been making an effort to use the stroke-affected limb

throughout the day then they will not receive a reminder when a unilateral GD movement is performed with the contralateral limb.

Detecting GD movements is, therefore, an important aspect of the presented approach. We focused on detecting GD movements of the non-affected upper limb in stroke survivors. To accurately monitor the relative usage of the stroke-affected and contralateral upper limbs it is also necessary to detect GD movements of the stroke-affected limb; otherwise patients could simply game the system by performing non-GD (e.g., passive arm swing) movements with stroke-affected limb. Future studies should examine GD movements performed by the stroke-affected upper limb and develop an algorithm that can determine the optimal ratio between the relative usage of the two limbs. We hypothesize that different algorithms may be needed for patients with different functional abilities. To our knowledge, this is the first time that an algorithm has been developed to detect GD movements. Future applications of this technique will expand upon prior work that has explored the use of the stroke-affected and contralateral upper limbs in stroke survivors [13], [55]–[57].

The presented algorithm for detecting the type of feedback appropriate for a given exercise extends our prior work using wearable sensor data to derive clinically-meaningful indicators of functional level and the severity of motor impairments in stroke survivors [33]–[35]. In the present study, and our prior work, a wide range of data features were selected from the acceleration, velocity, displacement, and angular velocity time-series to achieve reliable detection. This highlights the importance of using both accelerometers and gyroscopes when monitoring movement quality.

There were 17 *No Feedback* instances, twelve of which were misclassified as *Feedback* (i.e., the false positives in Fig. 5a). Interestingly, these 17 instances were associated with only three stroke survivors, and two of these three stroke survivors performed movements requiring both *Feedback* and *No Feedback* during a total of 10 repetitions of the exercise. This suggests that their movements were positioned on the “therapist’s decision boundary” for determining *Feedback* and *No Feedback*, making the classification especially challenging. On the other hand, 8 out of the 10 false negatives were generated by a single subject, which suggests that the movement patterns of that specific subject were similar to a particular set of *No Feedback* data in the training set. We argue that the primary cause of these misclassifications was the relatively small dataset available for analyses. To lessen the potential bias caused by the small data size, we performed the same analysis using a 10-fold cross validation instead of a leave-one-subject-out cross validation. This resulted in an *F*-score of 95.1% with only four false positives and four false negatives. These results suggest that a larger sample size would likely address the shortcomings of the classifier developed in this preliminary study.

There are several limitations of the study that deserve consideration. We did not consider movements that were both GD and non-GD in nature, nor did we handle cases

where a GD movement and a non-GD movement occurred in the same 2.5s window of time. This important analysis is left for future work. The sample size was relatively small (i.e., 20 stroke survivors and 10 healthy controls) and thus the reported results may not be generalized to the target patient population of chronic stroke survivors with mild-to-moderate upper-limb motor impairments. The limited sample size also restricted the number of rehabilitation exercises that were prescribed by a therapist to the stroke survivors who participated in the study. Consequently, data analyses were performed on the exercise that was prescribed to the largest number of stroke survivors, which was the “arm raise in the sagittal and coronal planes”. Therefore, the reported results for detecting feedback regarding the quality of rehabilitation exercise performance may not be applicable to other types of exercise movements. In an ongoing clinical study, we are collecting data from a larger patient population performing a variety of rehabilitation exercises. This dataset will allow us to assess the applicability of the proposed algorithms to other exercise movements. Future work to translate the research findings herein reported into clinical practice will require the deployment of the proposed system in the home setting and a rigorous evaluation of its robustness, usability and clinical effectiveness.

The envisioned technological approach could have a profound impact on stroke rehabilitation by extending rehabilitation interventions to the home. The results of the focus groups presented here indicate that both stroke survivors and therapists are amenable to the usage of such a technology. The proposed system would be especially useful in cases where high-dosage therapy is not feasible due to limited access to therapy programs. Patients with mild-to-moderate upper-limb impairments would be the ideal candidates for such a technology because they are able to engage in a wide range of ADL and rehabilitation exercises in the home environment.

We envision that the proposed system would be used first in an inpatient setting as an adjunct to standard therapy hours, which would allow patients to familiarize with the system and clinicians to choose optimal system settings on an individual basis (e.g., the reminder/feedback modality, the minimum time duration between feedback instances, and the threshold value to be applied to the relative usage of the two limbs to determine when feedback should be delivered). It would then be used in the outpatient setting to enable rehabilitation in home and community settings. The proposed technology could also be leveraged to assess stroke survivors’ responses to prescribed rehabilitation interventions. An accurate measure of limb use in the home environment would provide valuable information about a patient’s ability to translate the motor skills they acquired during therapy sessions to real life situations. This would allow therapists to closely monitor the motor recovery process and develop individualized therapy plans to maximize a patient’s ability to perform ADL and live independently, which is the ultimate goal of rehabilitation.

ACKNOWLEDGMENT

The authors would like to thank Edoardo Bonizzoni for his help in regard to the technical aspects of the data collections, A. O'Brien and S. Rajan for their help recruiting subjects and carrying out the data collections, and all the study participants (stroke survivors, control subjects, and occupational therapists) who volunteered their time to provide us with valuable feedback about the proposed system. S. I. Lee and M. Grimaldi were with the Department of Physical Medicine and Rehabilitation, Harvard Medical School, Charlestown, MA 02129, USA. A. V. Dowling and P. C. Horak were with BioSensics LLC, Watertown, MA 02472, USA.

REFERENCES

- [1] E. J. Benjamin *et al.*, "Heart disease and stroke statistics-2017 update: A report from the American heart association," *Circulation*, vol. 135, no. 10, pp. e146–e603, Mar. 2017.
- [2] J. S. Elkins and S. C. Johnston, "Thirty-year projections for deaths from ischemic stroke in the United States," *Stroke*, vol. 34, no. 9, pp. 2109–2112, 2003.
- [3] S. L. Wolf, P. A. Catlin, M. Ellis, A. L. Archer, B. Morgan, and A. Piacentino, "Assessing Wolf motor function test as outcome measure for research in patients after stroke," *Stroke*, vol. 32, no. 7, pp. 1635–1639, Jul. 2001.
- [4] J. M. Veerbeek, G. Kwakkel, E. E. H. van Wegen, J. C. Ket, and M. W. Heymans, "Early prediction of outcome of activities of daily living after stroke: A systematic review," *Stroke*, vol. 42, no. 5, pp. 1482–1488, May 2011.
- [5] A. Kunkel *et al.*, "Constraint-induced movement therapy for motor recovery in chronic stroke patients," *Arch. Phys. Med. Rehabil.*, vol. 80, no. 6, pp. 624–628, Jun. 1999.
- [6] E. Taub, G. Uswatte, V. W. Mark, and D. M. Morris, "The learned nonuse phenomenon: Implications for rehabilitation," *Eura Medicophys.*, vol. 42, no. 3, pp. 241–256, Sep. 2006.
- [7] E. Taub and D. M. Morris, "Constraint-induced movement therapy to enhance recovery after stroke," *Current Atherosclerosis Rep.*, vol. 3, no. 4, pp. 279–286, Jul. 2001.
- [8] S. L. Wolf *et al.*, "Effect of constraint-induced movement therapy on upper extremity function 3 to 9 months after stroke: The EXCITE randomized clinical trial," *JAMA*, vol. 296, no. 17, pp. 2095–2104, Nov. 01 2006.
- [9] J. A. Kleim and T. A. Jones, "Principles of experience-dependent neural plasticity: Implications for rehabilitation after brain damage," *J. Speech Lang. Hear. Res.*, vol. 51, pp. S225–S239, Feb. 2008.
- [10] K. S. Hayward, J. J. Eng, L. A. Boyd, B. Lakhani, J. Bernhardt, and C. E. Lang, "Exploring the role of accelerometers in the measurement of real world upper-limb use after stroke," *Brain Impairment*, vol. 17, no. 1, pp. 16–33, 2015.
- [11] M. Noorkoiv, H. Rodgers, and C. I. Price, "Accelerometer measurement of upper extremity movement after stroke: A systematic review of clinical studies," *J. Neuroeng. Rehabil.*, vol. 11, p. 144, Oct. 2014.
- [12] E. Taub, G. Uswatte, and R. Pidikit, "Constraint-induced movement therapy: A new family of techniques with broad application to physical rehabilitation—A clinical review," *J. Rehabil. Res. Develop.*, vol. 36, no. 3, pp. 237–251, Jul. 1999.
- [13] M. A. Urbin, K. J. Waddell, and C. E. Lang, "Acceleration metrics are responsive to change in upper extremity function of stroke survivors," *Arch. Phys. Med. Rehabil.*, vol. 96, no. 5, pp. 854–861, May 2015.
- [14] S. Barreca, S. L. Wolf, S. Fasoli, and R. Bohannon, "Treatment interventions for the paretic upper limb of stroke survivors: A critical review," *Neurorehabil. Neural Repair*, vol. 17, no. 4, pp. 220–226, Dec. 2003.
- [15] P. Duncan *et al.*, "A randomized, controlled pilot study of a home-based exercise program for individuals with mild and moderate stroke," *Stroke*, vol. 29, pp. 2055–2060, Oct. 1998.
- [16] A. Turton and C. Fraser, "The use of home therapy programmes for improving recovery of the upper limb following stroke," *Brit. J. Occupational Therapy*, vol. 53, no. 11, pp. 457–462, 1990.
- [17] R. R. Bailey, J. W. Klaesner, and C. E. Lang, "Quantifying real-world upper-limb activity in nondisabled adults and adults with chronic stroke," *Neurorehabil. Neural Repair*, vol. 29, no. 10, pp. 969–978, Nov. 2015.
- [18] G. Uswatte, C. Giuliani, C. Winstein, A. Zeringue, L. Hobbs, and S. L. Wolf, "Validity of accelerometry for monitoring real-world arm activity in patients with subacute stroke: Evidence from the extremity constraint-induced therapy evaluation trial," *Arch. Phys. Med. Rehabil.*, vol. 87, no. 10, pp. 1340–1345, 2006.
- [19] M. de Niet, J. B. Bussmann, G. M. Ribbers, and H. J. Stam, "The stroke upper-limb activity monitor: Its sensitivity to measure hemiplegic upper-limb activity during daily life," *Arch. Phys. Med. Rehabil.*, vol. 88, no. 9, pp. 1121–1126, 2007.
- [20] M. E. Michielsen, R. W. Selles, H. J. Stam, G. M. Ribbers, and J. B. Bussmann, "Quantifying nonuse in chronic stroke patients: A study into paretic, nonparetic, and bimanual upper-limb use in daily life," *Arch. Phys. Med. Rehabil.*, vol. 93, no. 11, pp. 1975–1981, 2012.
- [21] S. C. van der Pas, J. A. Verbunt, D. E. Breukelaar, R. van Woerden, and H. A. Seelen, "Assessment of arm activity using triaxial accelerometry in patients with a stroke," *Arch. Phys. Med. Rehabil.*, vol. 92, no. 9, pp. 1437–1442, 2011.
- [22] N. Gebruers, S. Truijen, S. Engelborghs, G. Nagels, R. Brouns, and P. P. De Deyn, "Actigraphic measurement of motor deficits in acute ischemic stroke," *Cerebrovascular Diseases*, vol. 26, no. 5, pp. 533–540, 2008.
- [23] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," *J. Neuroeng. Rehabil.*, vol. 9, no. 1, p. 21, 2012.
- [24] K. Leuenberger, R. Gonzenbach, S. Wachter, A. Luft, and R. Gassert, "A method to qualitatively assess arm use in stroke survivors in the home environment," *Med. Biol. Eng. Comput.*, vol. 55, no. 1, pp. 141–150, 2016.
- [25] L. A. Connell, N. E. McMahon, L. A. Simpson, C. L. Watkins, and J. J. Eng, "Investigating measures of intensity during a structured upper limb exercise program in stroke rehabilitation: An exploratory study," *Arch. Phys. Med. Rehabil.*, vol. 95, no. 12, pp. 2410–2419, Dec. 2014.
- [26] C. E. Lang, J. R. MacDonald, and C. Gnip, "Counting repetitions: An observational study of outpatient therapy for people with hemiparesis post-stroke," *J. Neurol. Phys. Therapy*, vol. 31, no. 1, pp. 3–10, Mar. 2007.
- [27] C. E. Lang, J. M. Wagner, D. F. Edwards, and A. W. Dromerick, "Upper extremity use in people with hemiparesis in the first few weeks after stroke," *J. Neurol. Phys. Therapy*, vol. 31, no. 2, pp. 56–63, Jun. 2007.
- [28] P. Saini *et al.*, "Philips stroke rehabilitation exerciser: A usability test," presented at the IASTED Int. Conf. Telehealth/Assistive Technol., Baltimore, MD, USA, 2008.
- [29] R. W. Teasell and L. Kalra, "What's new in stroke rehabilitation," *Stroke*, vol. 35, pp. 383–385, Feb. 2004.
- [30] S. A. Billinger *et al.*, "Physical activity and exercise recommendations for stroke survivors: A statement for healthcare professionals from the American heart association/American stroke association," *Stroke*, vol. 45, pp. 2532–2553, Aug. 2014.
- [31] G. Banks, J. Bernhardt, L. Churilov, and T. B. Cumming, "Exercise preferences are different after stroke," *Stroke Res. Treatment*, May 2012, Art. no. 890946.
- [32] J. T. Gwin, "Systems and methods for sensorimotor rehabilitation," U.S. Patent 9311 789, Apr. 12, 2016.
- [33] S. Del Din, S. Patel, C. Cobelli, and P. Bonato, "Estimating Fugl-Meyer clinical scores in stroke survivors using wearable sensors," in *Proc. Conf. IEEE Eng. Med. Biol. Soc.*, Aug./Sep. 2011, pp. 5839–5842.
- [34] S. Patel *et al.*, "A novel approach to monitor rehabilitation outcomes in stroke survivors using wearable technology," *Proc. IEEE*, vol. 98, no. 3, pp. 450–461, Mar. 2010.
- [35] S. I. Lee *et al.*, "Using wearable motion sensors to estimate longitudinal changes in movement quality in stroke and traumatic brain injury survivors undergoing rehabilitation," *Arch. Phys. Med. Rehabil.*, vol. 97, no. 10, p. e117, 2016.
- [36] J. Sanford, J. Moreland, L. R. Swanson, P. W. Stratford, and C. Gowland, "Reliability of the Fugl-Meyer assessment for testing motor performance in patients following stroke," *Phys. Therapy*, vol. 73, no. 7, pp. 447–454, Jul. 1993.
- [37] R. Muscollo, M. Schmid, S. Conforto, and T. D'Allesio, "Early recognition of upper limb motor tasks through accelerometers: Real-time implementation of a DTW-based algorithm," *Comput. Biol. Med.*, vol. 41, no. 3, pp. 164–172, 2011.
- [38] A. Vega-Gonzalez, B. J. Bain, P. M. Dall, and M. H. Granat, "Continuous monitoring of upper-limb activity in a free-living environment: A validation study," *Med. Biol. Eng. Comput.*, vol. 45, no. 10, pp. 947–956, Oct. 2007.

- [39] F. Hong, S. You, M. Wei, Y. Zhang, and Z. Guo, "MGRA: Motion gesture recognition via accelerometer," *Sensors*, vol. 16, no. 4, p. 530, 2016.
- [40] S. Arneja and J. Leith, "Review article: Validity of the KT-1000 knee ligament arthrometer," *J. Orthopaedic Surg.*, vol. 17, no. 1, pp. 77–79, Apr. 2009.
- [41] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.
- [42] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer, 2009.
- [43] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer, 2009.
- [44] A. Linden, "Measuring diagnostic and predictive accuracy in disease management: An introduction to receiver operating characteristic (ROC) analysis," *J. Eval. Clin. Pract.*, vol. 12, no. 2, pp. 132–139, Apr. 2006.
- [45] F. B. van Meulen *et al.*, "Objective evaluation of the quality of movement in daily life after stroke," *Frontiers Bioeng. Biotechnol.*, vol. 3, p. 210, Jan. 2015.
- [46] S. Chaeibakhsh, E. Phillips, A. Buchanan, and E. Wade, "Upper extremity post-stroke motion quality estimation with decision trees and bagging forests," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 4585–4588.
- [47] O. Giggins, D. Kelly, and B. Caulfield, "Evaluating rehabilitation exercise performance using a single inertial measurement unit," in *Proc. 7th Int. Conf. Pervasive Comput. Technol. Healthcare*, Venice, Italy, 2013, pp. 49–56.
- [48] O. M. Giggins, K. T. Sweeney, and B. Caulfield, "Rehabilitation exercise assessment using inertial sensors: A cross-sectional analytical study," *J. NeuroEng. Rehabil.*, vol. 11, p. 158, Nov. 2014.
- [49] F. Petitjean, G. Forestier, G. I. Webb, A. E. Nicholson, Y. Chen, and E. Keogh, "Dynamic time warping averaging of time series allows faster and more accurate classification," in *Proc. IEEE Int. Conf. Data Mining*, Dec. 2014, pp. 470–479.
- [50] J. Margarito, R. Helaoui, A. M. Bianchi, F. Sartor, and A. G. Bonomi, "User-independent recognition of sports activities from a single wrist-worn accelerometer: A template-matching-based approach," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 788–796, Apr. 2016.
- [51] A. Dalton *et al.*, "Development of a body sensor network to detect motor patterns of epileptic seizures," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 11, pp. 3204–3211, Nov. 2012.
- [52] X. Xi, K. Ueno, E. Keogh, and D.-J. Lee, "Converting non-parametric distance-based classification to anytime algorithms," *Pattern Anal. Appl.*, vol. 11, nos. 3–4, pp. 321–336, 2008.
- [53] M. A. Hall, "Correlation-based feature subset selection for machine learning," Ph.D. dissertation, Dept. Comput. Sci., Univ. Waikato, Hamilton, New Zealand, 1998.
- [54] D. M. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, Feb. 2011.
- [55] R. R. Bailey, J. W. Klaesner, and C. E. Lang, "An accelerometry-based methodology for assessment of real-world bilateral upper extremity activity," *PLoS ONE*, vol. 9, no. 7, p. e103135, 2014.
- [56] M. A. Urbin, R. R. Bailey, and C. E. Lang, "Validity of body-worn sensor acceleration metrics to index upper extremity function in hemiparetic stroke," *J. Neurol. Phys. Ther.*, vol. 39, no. 2, pp. 111–118, Apr. 2015.
- [57] R. R. Bailey, R. L. Birkenmeier, and C. E. Lang, "Real-world affected upper limb activity in chronic stroke: An examination of potential modifying factors," *Top Stroke Rehabil.*, vol. 22, no. 1, pp. 26–33, Feb. 2015.