This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3110432, IEEE Transactions on Biomedical Engineering

Poststroke Grasp Ability Assessment using an Intelligent Data Glove based on Action Research Arm Test: Development, Algorithms, and Experiments

Debeshi Dutta, Srinivasan Aruchamy, *Member IEEE*, Soumen Mandal, Soumen Sen, *Member IEEE*

***Abstract*— Growing impact of poststroke upper extremity (UE) functional limitations entails newer dimensions in assessment methodologies. This has compelled researchers to think way beyond traditional stroke assessment scales during the out-patient rehabilitation phase. In concurrence with this, sensor-driven quantitative evaluation of poststroke UE functional limitations has become a fertile field of research. Here, we have emphasized an instrumented wearable for systematic monitoring of stroke patients with right-hemiparesis for evaluating their grasp abilities deploying intelligent algorithms. An instrumented glove housing 6 flex sensors, 3 force sensors, and a motion processing unit was developed to administer 19 activities of Action Research Arm Test (ARAT) while experimenting on 20 voluntarily participating subjects. After necessary signal conditioning, meaningful features were extracted, and subsequently the most appropriate ones were selected using the ReliefF algorithm. An optimally tuned support vector classifier was employed to classify patients with different degrees of disability and an accuracy of 92% was achieved supported by a high area under the receiver operating characteristic score. Furthermore, selected features could provide additional information that revealed the causes of grasp limitations. This would assist physicians in planning more effective poststroke rehabilitation strategies. Results of the one-way ANOVA test conducted on actual and predicted ARAT scores of the subjects indicated remarkable prospects of the proposed glove-based method in poststroke grasp ability assessment and rehabilitation.**

***Index Terms*— ARAT, data glove, poststroke rehabilitation, ROC, SVC**

Manuscript received April 12, 2021; revised July 19, 2021; accepted August 27, 2021. This work was supported by CSIR-Central Mechanical Engineering Research Institute and Department of Science and Technology, Govt. of India, under project 'Indo-Korean joint network center on robotics'-Node 3, Grant No. INT/KOREA/JNC/Robotics dated April 12, 2018.

D. Dutta is with the Academy of Scientific and Innovative Research (AcSIR), Ghaziabad-201002, India. S. Aruchamy is with the CSIR-Central Mechanical Engineering Research Institute, Durgapur, West Bengal, India and AcSIR. S. Mandal is with the CSIR-Central Mechanical Engineering Research Institute, Durgapur, West Bengal, India and AcSIR. S. Sen is with the Robotics and Automation Group, CSIR-Central Mechanical Engineering Research Institute, Durgapur-713209, West Bengal, India and with AcSIR (Correspondence e-mail: soumen_sen@cmeri.res.in).

I. Introduction

Upper extremity (UE) functional limitations are primary causes of long-term poststroke disability wherein grasping postures and manipulative movements are significantly affected [1]. Regular assessment of grasp abilities contributes noticeably towards systematic rehabilitation leading to faster recovery. However, considering the complex anatomy of the human hand, assessment of grasping skills is a daunting task requiring regular intervention of concerned therapists. Further, socioeconomic factors, especially dominant in developing countries, restrain regular assessment of patients thereby invoking less effective rehabilitation [2–5]. It is also seen that, though primary healthcare centers (PHCs) are abundant in developing countries, secondary and tertiary healthcare units that offer the most suitable stroke care by specialized experts and instruments are comparatively less [6]. Owing to the large number of stroke cases in developing countries [7], [8], [9], [10], many of the stroke survivors are deprived of appropriate treatments. As a result, they are left with residual impairments with no substantial rehabilitation gains. This necessitates a simple yet comprehensive stroke assessment method that would promote effective rehabilitation even in remote locations. Stroke assessment scales are age-old techniques for evaluating patients with disabilities and are quite popular among physicians and researchers [11]. Nevertheless, these techniques require domain expert's intervention, which is again difficult to be availed frequently. Additionally, a precise knowledge concerning patient-specific causes of poststroke functional limitations that helps in determining the most appropriate rehabilitation strategies is not addressed adequately by conventional stroke assessment methods. In effect, new vistas towards the quantification of stroke rehabilitation outcomes are proposed by researchers for decreasing the burden of the caregivers [12]. These techniques focus on a paradigm shift in approach from manually intervened clinical stroke assessment methods to minimally supervised intelligent wearable technologies for poststroke grasp ability assessment [1], [12–14]. However, not many researchers have contributed in this direction thereby leaving many prospective wearable sensors, assessment scales, and intelligent algorithms unexplored. Recent advancements in this domain are summarized in TABLE S1 of supplementary material. As evident from the table, while accelerometers [15–18] and IMUs [19],[20] appear in most of the poststroke UE or grasp evaluation techniques, use of flex sensors[12], [20–23] in this domain is comparatively low despite their proven efficiency. On the other hand, the use of force sensors is the rarest though the significance of these sensors has been identified by researchers [21]. In addition, the choice of assessment scales and/or learning algorithms required for quantitative grasp evaluation varied with the objective of the research [1], [11]. Comprehensive knowledge of parameters such as finger bending angles, finger-tip pressures, and hand acceleration profiles acquired from apposite sensors would further contribute towards a better assessment of the quality of grasp thereby assisting the planning of suitable rehabilitation strategies. The authors were hence motivated to fabricate an affordable instrumented glove with embedded flex sensors, force sensors, and accelerometer at appropriate dorsal and palmar locations for acquiring useful information about stroke-driven disability.

In addition to the selection of sensors, choosing an appropriate assessment scale is also crucial and is usually selected based on the degree and level of disability following stroke [11]. Although a few researchers attempted to quantify the functional ability of stroke patients based on general grasp and grip tasks, this could not be clinically correlated due to the non-involvement of any standard assessment scale [24]. To improve upon this shortcoming, validation of wearable technologies with suitable clinical assessment scales was found significant. The most frequently used scales for evaluating patients with UE disability are Fugl Meyer assessment (FMA) [25], action research arm test (ARAT) [26], Chedoke-McMaster stroke assessment (CMSA) [27], and box and block test (BBT) [28], of which, the last two techniques have been not so popular among the researchers. ARAT is a criteria-rated task-oriented performance measure that closely evaluates the functional abilities of stroke patients at all stages of recovery. The assessment involves a minimal number of items with reduced complexity and test duration. Also, on investigating the test-retest, intra-rater, and inter-rater reliability of different UE evaluation techniques, ARAT was found to outperform similar scales with higher intra-class coefficient, Pearson correlation, and Spearman rho coefficient [29].

The most recent and valuable technological advancement in this area lies in the systematic merging of wearable sensor data and transmission hardware with intelligent learning algorithms to generate unique disability-specific scores to the users thereby leading to a more effective assessment tool for patients, researchers, and medical practitioners [30], [31]. The choice of learning model exclusively depends upon the nature of data and the objective of the research. Still, prominently used learning algorithms include random forests (RF), extreme learning machine (ELM), support vector machine (SVM), k-nearest neighborhood (k-NN), Bayesian network (BN), convolutional neural network (CNN), and trivial clustering and regression methods [20], [24], [32–38]. A few limitations of these algorithms are presented in TABLE I. A comparative study by a few researchers [39], reported the highest overall accuracy and least sensitivity to training sample sizes for SVM followed by RF and k-NN. Additionally, recent literature establishes that SVMs perform exceptionally well in the classification of locomotion quality in humans using wearable sensor data [40–43]. Although SVMs come with the cost of problem-specific tuning of hyper-parameters, this might be addressed by an additional step involving a grid search algorithm that ends up estimating the right hyperparameters for the given problem. Again, SVM hyperplanes sometimes exhibit poor generalization, which can be avoided by supplying the model with a training dataset that adequately represents the overall data population [49].

TABLE I
DRAWBACKS OF THE LEARNING ALGORITHMS FOUND IN LITERATURE

Algorithm	Drawbacks
BN	Computationally expensive; Fails to portray correlated variables [44]
CNN	Requires huge dataset; Translational invariance; Pooling layers leads to loss of information [45], [46]
ELM	Slow evaluation [47]
RF	Suited for small dataset; Black box model [48]
k-NN	Not suited for high dimensional data; sensitive to noise [39]

As summarized from the literature, the most prominent gaps in automated poststroke grasp ability assessment include paucity in the number of sensors used, the need for performance-based stroke assessment scale in experimental procedures, and the requirement of a simpler and robust learning algorithm. Furthermore, causes of poststroke grasp disability might vary between patients and this necessitates better apprehension of motor functions for each patient to design the best possible rehabilitation strategies. Notwithstanding the significance of the same, this perspective has received limited attention in the available literature. Hence, we propose an instrumented data glove comprising force sensors on the fingertip in addition to flex sensors and accelerometer at suitable locations for the acquisition of pressure, bend angle, and acceleration data respectively wherein the influence of the force sensors could be significantly established. Furthermore, grasp ability assessment was conducted based on ARAT considering its suitability and simplicity compared to other commonly used scales [29]. In addition, an appropriately tuned SVC algorithm, known for good overall accuracy was used to generate quantitative grasp ability measures in the current research.

The novelty of the proposed methodology lies in the fact that it can effectuate systematic assessment of grasp ability of stroke patients just by using an intelligent data glove supported by ARAT protocols even in a home setting. The assessment can be conducted with the help of a minimally trained individual or a family member who would provide basic assistance to the patient. Moreover, additional knowledge related to patient-specific causes of functional limitations can be extracted, thereby assisting the concerned physician in deciding appropriate rehabilitation strategies. Therefore, the proposed intelligent data glove-based methodology would promote an approachable and comfortable rehabilitation process leading to faster recovery of stroke patients.

II. METHODOLOGY

With advancements in automated stroke assessment methodologies, researchers are indulged in multiple sensor-based intelligent wearables for quantitative evaluation of UE disability after stroke [50]. Focusing on the grasping behavior of stroke patients, in this research we have fabricated a right-handed data glove for the acquisition of uninterrupted mobility information through properly located sensors on the human hand. An experimental protocol was designed with the help of a group of stroke patients and a few healthy controls whose clinical assessment scores were noted at first. Thereafter, the subjects were instructed to don the glove to perform a list of pre-defined activities that examined their grasp abilities. On obtaining a set of reliable data from the experiments, an SVM-based classifier was used to predict the scores, and correlation results were gradually examined. In this paper, the SVM based classifier is termed support vector classifier (SVC). The proposed methodology and experimental strategies are summarized in Fig. 1. As evident from Fig. 1, data collection from the subjects was initiated after the ARAT assessment following which pre-processed labeled data with relevant features was trained using optimally tuned SVC to predict grasp ability scores without expert intervention. The data collection and interpretation procedure has also been illustrated explicitly in Fig. 1. The methodology has been discussed in the following sections.

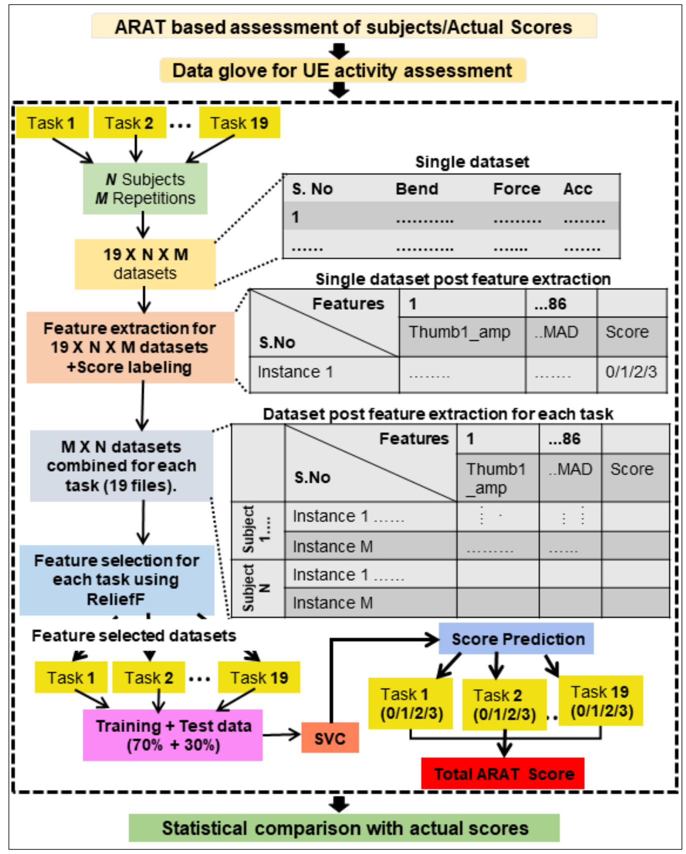

Fig. 1. Summary of the proposed methodology and the experimental strategies.

A. *Fabrication and calibration of the glove for UE mobility assessment*

An instrumented wearable set-up for the right hand was assembled on a polyamide/Lycra glove procured from B'TWIN®. The glove, made up of 0.38 mm thick stretchable tulle, ensured a reliable interface between the skin and the object dealt with. The palmar side of the glove had embedded grip pads which enhanced friction with the concerned object. Further, it was expected that a sufficiently thin fabric would allow better fitting and shape conformity with the human hand. Additionally, to check whether the glove posed any restriction to finger movements, trials were conducted on a few healthy individuals as well as stroke patients, and practically no complaint of discomfort was received. Based on the gaps identified in existing literature, the glove-based assessment platform was incorporated with multiple sensing yet holding on to simple construction. The sensor assembly constituted a suitably calibrated accelerometer, 6 bend sensors, and 3 force sensors for approximating real-life functional activities of stroke patients. Constructional details and validation of the fabricated glove can be found in the supplementary material and a pilot work [51]. Fig. 2 (i) shows the location of each sensor, details being described in TABLE II. It was further noted that the bend and force sensors were reset to zero at the initial rest condition on a flat surface. The fabricated instrumented glove developed for the assessment is shown in Fig. 2 (ii).

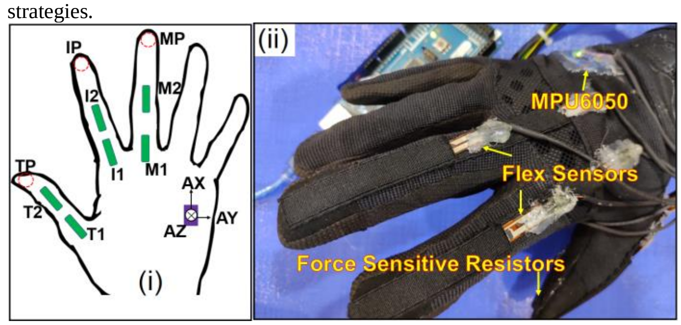

Fig. 2. (i) Sensor locations for the fabricated glove (ii) The fabricated glove used for experiments

TABLE II
EXPLANATION OF ABBREVIATION FOR EACH SENSOR

Abbreviation	Meaning
T, I, and M	Thumb, Index, and Middle finger
1, 2	Bend of MCP and PIP (or IP) joints of each finger
TP, IP, and MP	Tip pressures
AX, AY, and AZ	Acceleration in X, Y, and Z directions

B. *Data Collection*

The experiments were conducted in a comfortable laboratory/home set-up with 15 stroke patients with right hemiplegia and 5 healthy controls aged between 47–64 years. The stroke patients were on their 20th–82nd week following a stroke while the experiments were conducted. It was ensured that they were able to follow verbal instructions and were mobile enough to wear the glove. In concurrence with this, few subjects (in addition to the 20 considered) had to be opted out since they were unable to don the glove due to deformed and stiff fingers. Patients experiencing nociceptive pain and depending on passive exercises were also excluded. The selection criteria for the subjects along with general statistical information are presented in TABLE III. Based on the recommendations of a few researchers [26], [29], in this research, the ARAT based scoring technique was implemented to evaluate the participants. Each patient was seated

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3110432, IEEE Transactions on Biomedical Engineering

comfortably on an armless chair adjacent to a table of height ~25 inches, large enough to accommodate all the items of the ARAT kit. At this position, the patients were administered and scored based on their performance while using the affected (right) arm. The ARAT protocol allows a maximum time of 60 s for each task. The patients were rated on a scale of 0 (no movement within the maximum time limit), 1 (movement task is partially performed within the maximum time limit), 2 (movement task is completed but takes more than 5 seconds), and 3 (movement is performed normally within 5 seconds). The ARAT assessment and scoring strategies are briefly demonstrated in Fig. 3.

TABLE III
SELECTION CRITERIA AND STATISTICAL MEASURES FOR SUBJECTS

Eligibility criteria for inclusion of stroke patients					
Sl.	**Criteria**	**Eligibility**			
1.	Mobility status	Partial ability to perform ADL, no signs of spasticity			
2.	Upper limb paresis	Right			
3.	Cause of disability	Stroke			
4.	No. of stroke attacks	Not more than 1			
General information of the participating subjects					
		Stroke Patients (11 Males + 4 Females)		**Healthy Controls** (3 Males + 2 Females)	
Sl.	**Criteria**	**Mean**	**SD**	**Mean**	**SD**
1.	Age (in years)	56.92	5.42	57.5	2.64
2.	Duration after stroke (weeks)	50.33	15.63	NA	NA
3.	ARAT Score	32	12.25	57	NA

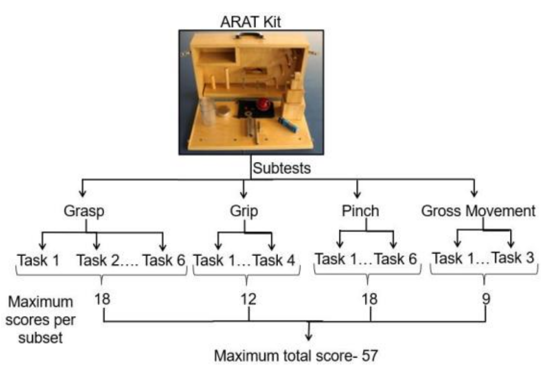

Fig. 3. Activities associated with ARAT assessment scale where the best possible scores for each subtest viz. grasp, grip, pinch, and gross movement totaling to a maximum score of 57 as reference

The tasks associated with each subtest of ARAT assessment scale are elaborated in TABLE S2 of supplementary material. The entire process was supervised by three therapists (raters) considering all compensatory movements [52]. It was ensured that the patients completed the tasks without any extra assistance and the other arm was kept at a fixed position with confirmed non-use. Due care was taken so that the participants were aware of the experimental protocol. 2–3 trials were allowed till the participants were confident enough for the main assessment. Scores assigned to the patients sometimes varied between raters and the same was decided based on the opinion of each rater and the maximum voted score was considered. This step was skipped for the healthy controls and each of them was assigned the highest score (=57). Keeping the experimental conditions unaltered, each one of them was instructed to don the glove and redo the tasks after a few successful trials. Considering the high flexibility of the glove, deformed/immobile fingers of patients could be comfortably inserted in the glove by pulling and adjusting it appropriately. Healthy subjects initiated the data collection procedure and the data collected for each task was set as reference and was considered ideal for the rest of the experiment. For every individual, each task was repeated 5–7 times and corresponding data was noted. Sufficient intervals were ensured between repetitions to avoid fatigue-related anomalies in the recorded data. The first few experiments were conducted in presence of the therapist who confirmed no considerable dissimilarity in performances with and without wearing the glove. The present research being particularly based on the design considerations of the developed glove, the rest of the data were collected under the supervision of a clinically experienced individual (typically a physiotherapy assistant- WHO ISCO code-3255) easily available at PHCs. Although as far as the deployment of the methodology is concerned, the protocol could be executed by a minimally trained individual and/or family member. This individual would be supplied with a step-wise instruction set as in TABLE S3, adhering to which would not require clinical expertise or familiarity with ARAT based scoring. Considering the average duration of the tasks, the sampling frequency of all the sensors was fixed at 10 Hz based on existing literature [20], [32], [33], [52], [54]. The glove housed a microcontroller that transmitted the data to a PC/ laptop via USB serial communication.

Time-stamped data was transmitted in the format readable by PLX DAQ as stated in the PARALLAX datasheet and was stored as .csv files in form of matrices, where, each row represented one instance while the columns indicated respective sensor readings. Fig. 1 can be referred to for the stepwise data gathering process. Every experimental session was video recorded for future reference. The experimental process was approved by the institutional ethical committee. Each subject and/or the accompanying person was requested to go through the details before signing the informed consent form at the onset of the experiment.

C. Data pre-processing, feature extraction, and feature selection

Denoising of a signal is inevitable especially when accelerometer data is used [20]. Raw sensor output was denoised using a Savitzky-Golay (SGolay) filter to maximize the precision of data without affecting the trend of the signal [55]. An SGolay filter of polynomial order 4 and frame size 11 was chosen for the present problem. Smoothening of data was also performed using a 5-point Moving Average (MA) filter the performance of both the filters was compared based on Percentage Root-mean-square Difference (PRD) and estimated Signal to Noise Ratio (SNR_E) of the original signal and its denoised counterpart [56]. In addition, the effect of filtering on the jerk feature was also evaluated. Selection and comparison of the filters were based on their widespread use in biomedical signal processing as evident in related literature [20], [57]. By investigating the features commonly used for time-domain signal analysis for physical activity monitoring available in the literature [58], the most prominent features were extracted from the dataset. The most relevant features extracted for our experiments included- (i) Mean- average of absolute values of sensor readings, (ii) Peak- maxima of absolute values of sensor readings), (iii) Maximum amplitude (amp)- the difference between mean and peak of sensor readings, (iv) Root-mean-square (RMS) value- square root of the arithmetic mean of the squares of sensor readings, (v) Jerk (JERK)- RMS value of the derivative of the acceleration time-series (or jerk time series) [59], [60], (vi) Approximate entropy (ApEn)- calculated using standard equations with embedding dimensions m=2 and r=0.15 times standard deviation [61], [62], (vii) Mean absolute deviation (MAD)- mean of the absolute deviations of the sensor readings around the mean of the sensor readings, and (viii) Pearson correlation coefficient (Corr) [63]. Features associated with flex and force sensors majorly represent grasping postures and forces. On the other hand, features related to the accelerometer sensor are indicative of dynamic components thereby responding to compensatory movements [64], [65]. The extracted features for each sensor along with detailed nomenclature with reference to the abbreviations used in Fig. 2(i), is presented in TABLE S4 of the supplementary material, where, VX, VY, and VZ are velocities computed for X, Y, and Z directions respectively by integrating respective accelerations with the initial condition at rest.

A feature selection algorithm was introduced to select the most appropriate features for each item of ARAT to deal with dimensionality issues and enhance the performance of the learning model. Inspired by the results obtained by various researchers, the ReliefF algorithm was incorporated to select the most relevant features for every task [66], [67]. ReliefF feature selection was executed in Weka 3.8.3 toolbox and a set of features ranked in order of importance were acquired. The number of features considered for the final assessment was selected based on experiments. Before considering the dataset for predictions the feature set was scaled to ensure the best results.

Unlike manual poststroke assessment through conventional methods, the outcomes of feature extraction facilitate a comprehensive understanding of hand movement distinctive for every patient. The extracted features like RMS, correlation, ApEn, and jerk reveal unique disability characteristics such as dynamic energy, coordination, randomness, and smoothness respectively [68]. The influence of these significant features on the task undertaken by the examinees was analyzed.

D. *Classification of subjects using C-SVC*

The subjects were classified using C-SVC based on which scores were predicted for unknown instances, mathematical formulations been elaborated in the supplementary material [69], [70]. Training and classification were conducted through a few noteworthy steps. At first, the acquired dataset with selected features was labeled with pre-defined ARAT scores and split into 70% and 30% for training and testing respectively. To adhere to the minimum generalization criteria of the SVM, the training data was nearly balanced among classes 0, 1, 2, and 3 for most of the tasks. The test data was verified for at least one data instance from each subject and the appearance of all classes in the dataset was ensured to visualize how each class was discriminated from the other. Next, a one-versus-one classification-based SVC algorithm with a polynomial kernel was adopted to train the samples and later predict scores on the test set in Python 3.8. Optimal hyper-parameters (C and γ) were chosen, where C stands for the cost of misclassification that must be as low as possible and γ is the inverse of radius of influence of samples selected by the model as support vectors that plays a significant role in kernel function. A grid search was conducted for tuning the hyper-parameters until the best plane for classification was obtained. To identify the best possible support vector hyperplane for the given problem and complying with recommendations by Hsu C *et al.* [71], an exponentially growing sequence of C and γ was chosen. A loose grid search was conducted at first to obtain C and γ, where the input sequence was set as $C = 2^{-10}, 2^{-9}\ldots,2^{10}$ and $\gamma = 2^{-15}, 2^{-14}\ldots,2^{3}$. On identifying the C and γ values after loose grid search, a finer grid search was conducted on their neighborhood to further fine-tune them. For instance, $C = 2^{-6}$, $\gamma = 2^{1}$ obtained after loose grid search was subjected to a finer grid search on a neighborhood specified by $C = 2^{-5}, 2^{-5.1}\ldots,2^{-7}$ and $\gamma = 2^{0}, 2^{0.1}\ldots,2^{2}$. Subsequently, the fine-tuned C and γ values were used to fit a 3rd degree polynomial kernel on the dataset to classify different stages of grasp disability (0, 1, 2 or 3) based on SVC. Prediction accuracy was determined based on the test dataset. Subject-specific scores were predicted for each task and the final score of each stroke patient was estimated on a 0–57 scale by summing up the individual scores over 19 different tasks. The performance of the classifier was evaluated using various statistical aspects. To scrutinize the performance of the model, Area Under the Receiver Operating Characteristics (AUROC) was computed. This metric was indicative of the model's capability of distinguishing between classes. The estimated ARAT scores and the predicted scores were subjected to a one-way ANOVA with a significance level of 0.05 to draw conclusions related to variability between the two.

III. Results And Discussions

A. *Characterization of the instrumented glove*

Experimental trials on the glove affirmed easy movement and no significant constraint in finger movement was observed, especially for stroke patients with functional limitations. Further, the glove was found to fit firmly on adult Indian subjects with the force sensors lying suitably on the fingertips. To ensure reliable sensory information, each sensor was carefully characterized before installation. The calibration results of the sensors are shown in Fig. S1 of supplementary material. The R-square and Root Mean Squared Error (RMSE) values found in TABLE S5 of supplementary material exhibited a good fit and hence the proposed equations could be incorporated into the functional code to ensure proper interpretation of targeted disability.

B. *Signal conditioning*

The SGolay and 5-point MA filtering techniques were implemented on raw datasets and a comparison between the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3110432, IEEE Transactions on Biomedical Engineering

two was carried out. The effect of denoising on sensor data for a healthy subject as well as a stroke patient considering the 3rd task from the grasp subtest (grasping a 5 cm^3 wooden block) is demonstrated in Fig. 4. The first task of each subtest was the toughest (for example, grasp a 10 cm^3 wooden block or grip and pour water from glass to glass), the second one was the easiest (for example, grasp a 2.5 cm^3 wooden block or grip a tube of diameter 2.25 cm), while the others were of intermediate complexity (for example, grasp a 5 cm^3 wooden block or grip a tube of diameter 1 cm). In this context, the characteristic of PIP joint bend angle and tip-pressure of index finger and acceleration along X-axis related to the 3^{rd} task of grasp subtest was considered for demonstration purposes. The PRD and SNR_E values for both the filters are illustrated in Fig. 5. The SGolay filter outperformed in all cases with the highest SNR_E and least PRD. The effect of using the SGolay filter on the jerk metric for a few subjects is shown in Table S6, which depicts a uniform change in jerk metric related to gross subtest post-filtering. This in turn ensures that the computation of jerk metric using filtered data would not affect the overall classification accuracy.

C. Feature selection

The feature selection step was a noteworthy inclusion following the extraction of relevant features that aimed at unleashing meaningful insights from raw sensory information. It was observed that inter-subtest features varied significantly i.e., features selected for tasks belonging to different subtests were largely dissimilar. On the other hand, intra-subtest feature variability was comparatively lower i.e., features selected for tasks belonging to the same subtest were most common. It was further observed that the tasks included in the gross movement focused on the features concerning acceleration whereas the pinch subtest majorly concentrated on the tip-force data. A thrust on features related to the index and middle fingers while executing task 2 and task 6 respectively of the pinch subtest further confirmed the relevance of the feature selection step. A few tasks were chosen to demonstrate this phenomenon as summarized in TABLE IV. The table further shows important insights into the use of force sensors since feature 86 i.e., MAD of index-tip pressure appears frequently in most of the listed tasks followed by the RMS value of index-tip pressure in grasp and grip tasks.

The information acquired from the feature extraction step was significant in discriminating patients based on mobility. To investigate this phenomenon, a few impactful selected features were chosen from four different tasks viz. RMS_IP from grip/task 2, Corr_I2T2 from pinch/task 2, ApEn_AX, and JERK_AX from gross/task 3. The said features were selected considering their multiple appearances in several tasks with reference to TABLE IV. 3 subjects each from classes 1 and 2 were selected to establish intra-class and inter-class comparisons. Fig. 6 maps magnitudes of feature-related reference information influencing motor performances of stroke patients for a selected set of tasks. It can be found in Fig. 6 that unlike the first two patients with a score of 2, the third one exhibited almost twice ApEn_AX (randomness), a higher Corr_I2T2 (co-ordination), and a lower JERK_AX (smoothness).

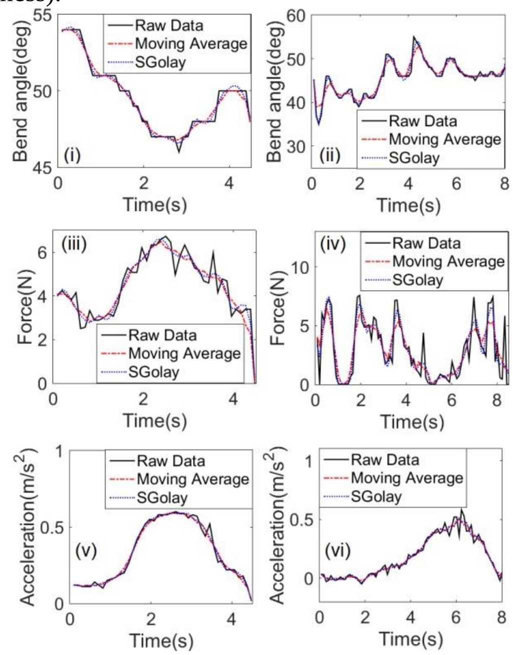

Fig. 4. Effect of denoising using S-Golay and MA filtering for a healthy subject (i, iii and v) and a stroke patient (ii, iv, vi) on grasp/task 3.

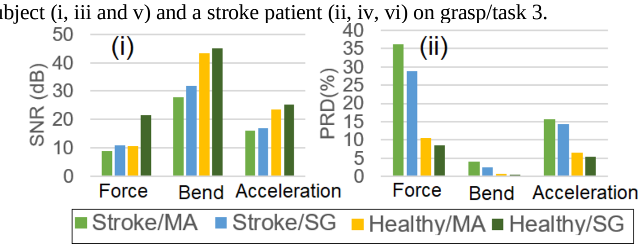

Fig. 5. Performance evaluation of moving average and SGolay filters using- (i) SNR_E, and (ii) PRD on grasp/task 3 for a healthy subject and a stroke patient.

TABLE IV
FEATURES SELECTED USING RELIEFF ALGORITHM DEMONSTRATING INTER AND INTRA CLASS VARIABILITY

Task	Selected features (Nomenclature as per TABLE S4)
Grasp/ Task 1	8, 86, 44, 28, 60, 16, 40, 17, 4, 56, 29, 5, 41, 20
Grasp/ Task 2	4, 44, 7, 43, 86, 85, 16, 38, 14, 40
Grip/ Task 2	32, 37, 20, 26, 44, 86, 1, 14, 13, 25, 57, 8, 38, 58, 79
Pinch/ Task 2	8, 20, 7, 19, 2, 4, 3, 86, 68, 65, 85
Pinch/ Task 6	21, 9, 45, 7, 5, 6, 2, 42, 87, 85, 30
Gross/ Task 3	49, 50, 72, 75, 34, 70, 62, 61, 46, 76, 22

Fig. 6. A comparative study of feature-related additional information influencing the motor performance of stroke patients

This indicated that the rehabilitation strategy for this patient must be designed such that special care is taken for randomness and smoothness in movement. Again, the last patient with a score of 1 demonstrated a JERK_AX nearly equal to that of patients with score 2 and indicated relatively

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3110432, IEEE Transactions on Biomedical Engineering

enhanced Corr_I2T2 and ApEn_AX when compared to other patients with the same score. This infers that rehabilitation exercises for this patient must be chosen such that the randomness in movement is reduced.

D. C-SVC

1) *Choosing the hyperparameters*

To obtain the best as well as most precise hyperparameters for the C-SVC a loose grid search was followed by a finer grid search. After a loose grid search, a hyperparameter set of ($C = 2^{-6}$, $\gamma = 2^{1}$) was obtained. After a finer grid search on the neighborhood of ($C = 2^{-6}$, $\gamma = 2^{1}$) the hyperparameter set was further updated to ($C = 2^{-6.2}$, $\gamma = 2^{1.4}$).

2) *Performance evaluation of C-SVC*

The obtained hyperparameter set was considered to fit a 3rd-degree polynomial kernel on the dataset. The corresponding separating plane plotted on test data after fine grid search for two highest-ranked features viz. feature number 8 (Mean_IP) and 86 (MAD_IP) for the first task of grasp subtest can be found in Fig. S2 of supplementary material.

It was observed that scores of 2 and 0 were misclassified as 1 for a few instances. After this stage, based on experiments it could be seen that any modification in the hyperparameters or the degree of the polynomial ended up with a poor accuracy or an over-fit. Subsequently, datasets for the rest of the tasks were subjected to classification based on the same SVC model and respective accuracies were computed. The classification accuracies pertaining to the number of features varying from 6 to 20 for each task were plotted and displayed in Fig. 7. While in most cases noticeable drifts in accuracies could be observed from 6 to 10 features, none indicated a significant change in accuracy beyond 15 features. Hence, the number of features considered for final predictions was limited to 15. The accuracies for all the tasks and respective confusion matrices are tabulated in TABLE S7 of supplementary material. For ease of explanation, the results obtained for task 1 of the grasp subtest are selected for discussion as follows.

From TABLE V it could be observed that unlike classes 0, 2, and 3, class 1 demonstrated a reduced precision of 0.57 indicating the classifier's inaccuracy towards positive predictions. Class 2 had the lowest recall followed by class 0 inferring the reduced ability of the classifier in identifying all the positive instances that belonged to these classes. This phenomenon can also be recognized from the confusion matrix heatmap presented in Fig. S3 of the supplementary material. It is observed from the results that the proneness of a higher score being misclassified to lower is more abundant and therefore the cost of misclassification can be considered mild since a person with better functional abilities was scored inadequately which is undoubtedly less vital compared to overscoring. The latter might misguide a therapist/patient assuring an improvement in functional abilities that did not eventuate. It is also to be noted that throughout the experiments the class with the highest score (=3) was neither misclassified to other classes nor any other class was misclassified into it. This ensures that the classifier could sufficiently distinguish between healthy subjects and stroke patients. In addition, misclassification took place only within consecutive classes, for example, class 2 was never misclassified as class 0. This characteristic is an indication of good estimation accuracy of total scores and minimal deviation.

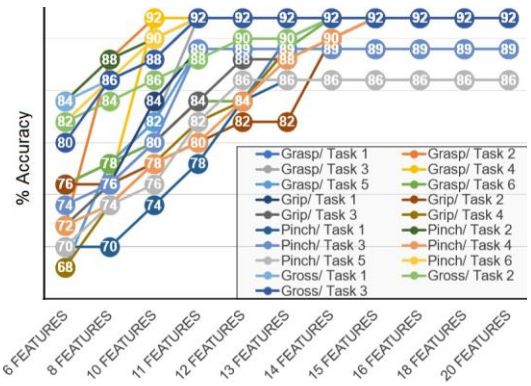

Fig. 7. Influence of number of features selected on classification accuracy of SVC

TABLE V
CLASSIFICATION REPORT FOR C-SVC FOR ALL SUBJECTS CONSIDERING GRASP/ TASK 1

Classes/Scores	Precision	Recall	F1-score	Support
0	1.00	0.88	0.93	8
1	0.57	1.00	0.73	4
2	1.00	0.83	0.91	12
3	1.00	1.00	1.00	13
Accuracy			0.92	

An accuracy of 92% confirmed a satisfactory classification model when compared to the existing literature furnished in Table S1, where the accuracies obtained ranged between 82% and 90%. Fig. 8 displays the ROC curves for each class as well as micro and macro-average of all classes binarized which again shows evidence of class 0 and 2 being misclassified to 1, whereas class 3 meets the 'ideal' point. The weighted macro-average AUROC score of 0.99 for the one-vs-one scheme of classification indicated satisfactory model performance.

Scores predicted for each task were added up over 19 different tasks to obtain the final predicted score for each patient after which a single data instance from each patient was chosen for the prediction. A chart depicting the comparison between actual and predicted scores is presented in Fig. 9, where predicted scores deviate insignificantly from the actual ones with a mean error of 1.15.

The results of ANOVA, as presented in TABLE VI, depict a much lower sum of squares (SS) and mean-size-squared (MS) for between-group variance ($SS_{between}$=13.22, $MS_{between}$=13.22) when compared to within-group variance (SS_{within}=8943.75, MS_{within}=235.36). This infers that the individual means of the expected and predicted scores do not differ much and hence their differences with the grand mean won't vary significantly either.

Here, SS and MS for the total sample indicates the sum of the squared deviations from the grand mean and the variance for the total sample respectively. An F-ratio as low as 0.05 and a significantly high F-critical value of 4.09 signifies major statistical similarity between the expected and predicted scores. Furthermore, a P-value (=0.81) much greater than the significance level (=0.05) confirms the null hypothesis i.e., the two groups viz. actual scores and predicted scores have statistically insignificant differences. Based on these

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2021.3110432, IEEE Transactions on Biomedical Engineering

inferences, it is evident that the presented automated assessment procedure can adequately quantify post-stroke grasp ability complying ARAT scale to assist the concerned physician to correlate grasp ability status clinically.

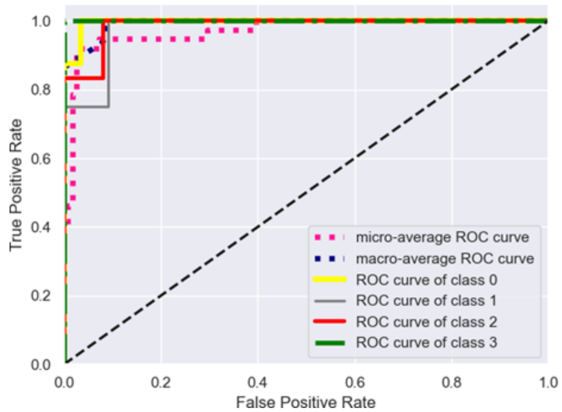

Fig. 8. ROC curves for each class as well as micro and macro-average of all classes considering grasp/task 1

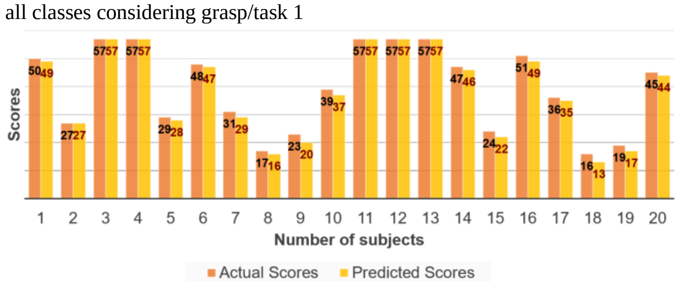

Fig. 9. Comparison of actual and predicted total scores for the subjects

TABLE VI
RESULTS OF ONE-WAY ANOVA ON ACTUAL AND PREDICTED SCORES

Source of Variation	**SS**	**df**	**MS**	**F-ratio**	**P-value**	**F-critical**
Between Groups	13.22	1	13.22	0.05	0.81	4.09
Within Groups	8943.75	38	235.36	-	-	-
Total	8956.97	39	248.58	-	-	-

IV. CONCLUSIONS

In this research, an improved assessment methodology is presented towards the quantitative evaluation of the grasping ability of stroke patients. The new findings include- i) an instrumented glove aided with SVM algorithm that could quantify grasp ability of stroke patients in an automated manner; ii) assisting physicians in designing suitable rehabilitation strategies by providing additional mobility information from extracted features; iii) glove housing force sensors on finger-tips establishing the significance of finger-tip pressures towards stroke assessment not found earlier; iv) a minimally supervised poststroke grasp ability assessment method using comprehensive and simple ARAT scale that outperformed similar scales. The research hence opens new vistas towards assessment of poststroke upper extremity functional limitations involving ARAT scoring technique. Nevertheless, SVC performed exceptionally well with an accuracy of 92% supported by a weighted macro-average AUROC score of 0.99.

The developed instrumented glove aided with a suitable learning algorithm can find non-clinical applications as well, such as in identifying the shape, size, and softness of unknown objects. General hand assessment that correlates with ARAT based assessment will be conducted as future work. Implementation of cloud interface and integration of appropriate Internet of Things (IoT) modules would encourage more effectual monitoring of motor functions and remote assistance to stroke patients, thus facilitating m-Health based stroke assessment.

ACKNOWLEDGEMENT

The authors thank CSIR-CMERI, the participants of the experiments, and the therapists.

REFERENCES

[1] P. Aqueveque et al., “After stroke movement impairments: A review of current technologies for rehabilitation,” in Physical Disabilities-Therapeutic Implications, ch. 7, 2017.

[2] A. Rauch et al., “How to apply the international classification of functioning, disability and health (ICF) for rehabilitation management in clinical practice,” Eur J Phys Rehabil Med, vol. 44, no. 3, pp. 329–342, 2008.

[3] E. Lang et al., “Assessment of upper extremity impairment, function, and activity after stroke: Foundations for clinical decision making,” J Hand Ther, vol. 26, no. 2, pp. 104–115, 2013.

[4] P. Langhorne, “Organised inpatient (stroke unit) care for stroke,” Cochrane Database Syst Rev, vol. 4, no. 4, 2013.

[5] H. K. Kristensen et al., “Research-based evidence in stroke rehabilitation: an investigation of its implementation by physiotherapists and occupational therapists,” Disabil. Rehabil., vol. 38, no. 26, pp. 2564–2574, 2016.

[6] J.D. Pandian and P. Sudhan, Stroke Epidemiology and Stroke Care Services in India, J Stroke, vol. 15, no. 3, pp. 128–134, 2013.

[7] T. K. Banerjee and S. K. Das, “Fifty years of stroke researches in India,” Ann Indian Acad Neurol, vol. 19, no. 1, pp. 1–8, 2016.

[8] R. V. Krishnamurthi et al., “Global and regional burden of first-ever ischaemic and haemorrhagic stroke during 1990–2010: findings from the global burden of disease study 2010,” Lancet Glob Heal, vol. 1, no. 5, pp. E259–281, 2013.

[9] V. L. Feigin et al., “Worldwide stroke incidence and early case fatality reported in 56 population-based studies: a systematic review,” Lancet Neurol, vol. 8, no. 4, pp. 355–369, 2009.

[10] H. Lee et al., “Development-assistance strategies for stroke in low-and middle-income countries,” J Korean Med Sci, vol. 30, no. 2, pp. S139–S142, 2015.

[11] J.K. Harrison et al., “Assessment scales in stroke: clinimetric and clinical considerations,” Clin Interv Aging, vol. 8, pp. 201–11, 2013.

[12] P. MacEira-Elvira, et al., “Wearable technology in stroke rehabilitation: towards improved diagnosis and treatment of upper-limb motor impairment,” J Neuroeng Rehabil, vol. 16, 2019, Art. no. 142.

[13] J. A. Franck et al., “Changes in actual arm-hand use in stroke patients during and after clinical rehabilitation involving a well-defined arm-hand rehabilitation program: A prospective cohort study,” PLoS One, vol. 14, no. 4, 2019, Art. no. e0214651.

[14] J. Kenry et al., “Emerging flexible and wearable physical sensing platforms for healthcare and biomedical applications,” Microsyst Nanoeng. 2016, vol. 2, Art. no. 16043.

[15] M. Noorkõiv et al., “Accelerometer measurement of upper extremity movement after stroke: a systematic review of clinical studies,” J Neuroeng Rehabil, vol. 11, 2014, Art. no. 144.

[16] S. I. Lee et al., "Predicting and monitoring upper-limb rehabilitation outcomes using clinical and wearable sensor data in brain injury survivors," IEEE Trans Biomed Eng, vol. 68, no. 6, pp. 1871–1881, 2021.

[17] R. J. M. Lemmens et al., “Accelerometry measuring the outcome of robot-supported upper limb training in chronic stroke: a randomized controlled trial,” PLoS One, vol. 9, no. 5, 2014, Art. no. e96414.

[18] E. Narai et al., “Accelerometer-based monitoring of upper limb movement in older adults with acute and subacute stroke,” J Geriatr Phys Ther, vol. 39, no. 4, pp. 171–177, 2016.

[19] M. El-Gohary and J. McNames, "Shoulder and elbow joint angle tracking with inertial sensors," IEEE Trans Biomed Eng, vol. 59, no. 9, pp. 2635–2641, 2012.

[20] A. L. van Ommeren et al., "Detection of the intention to grasp during reaching in stroke using inertial sensing," IEEE Trans Neural Syst Rehabil Eng, vol. 27, no. 10, pp. 2128–2134, 2019.
[21] Yu et al., "A remote quantitative Fugl-Meyer assessment framework for stroke patients based on wearable sensor networks," Comput. Methods Programs Biomed, vol. 128, pp. 100–110, 2016.
[22] S. Lee et al., "Automated evaluation of upper-limb motor function impairment using Fugl-Meyer assessment," IEEE Trans Neural Syst Rehabil Eng, vol. 26, no. 1, pp. 125–134, 2018.
[23] G. B. Prange-Lasonder et al., "Applying a soft-robotic glove as assistive device and training tool with games to support hand function after stroke: preliminary results on feasibility and potential clinical impact," IEEE Int Conf Rehabil Robot, 2017, pp. 1401–1406.
[24] D. Dutta et al., "Bayesian network aided grasp and grip efficiency estimation using a smart data glove for post-stroke diagnosis," Biocybern Biomed Eng, vol. 37, no. 1, pp. 44–58, 2017.
[25] A. R. Fugl Meyer et al., "The post stroke hemiplegic patient. I. A method for evaluation of physical performance," Scand J Rehabil Med, vol. 7, no. 1, pp. 13–31, 1975.
[26] R. C. Lyle, "A performance test for assessment of upper limb function in physical rehabilitation treatment and research," Int J Rehabil Res, vol. 4, no. 4, pp. 483–492, 1981.
[27] A. Gowland et al., "Measuring physical impairment and disability with the Chedoke-Mcmaster stroke assessment," Stroke, vol. 24, no. 1, pp. 58–63, 1993.
[28] V. Mathiowetz et al., "Adult norms for the Box and Block Test of manual dexterity.," Am J Occup Ther, vol. 39, no. 6, pp. 386–91, 1985.
[29] T. Platz et al., "Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer test, Action Research Arm Test and Box and Block Test: a multicentre study," Clin Rehabil, vol. 19, no. 4, pp. 404–411, 2005.
[30] G. Kwakkel et al., "Effects of augmented exercise therapy time after stroke: a meta-analysis," Stroke, vol. 35, no. 11, pp. 2529–2539, 2004.
[31] M. Morrison, "Inertial Measurement Unit," US Patent- US4711125A, 1985.
[32] S. Patel et al., "Tracking motor recovery in stroke survivors undergoing rehabilitation using wearable technology," Annu Int Conf IEEE Eng Med Biol Soc, 2010, pp. 6858–6861, 2010.
[33] S. Del Din et al., "Estimating Fugl-Meyer clinical scores in stroke survivors using wearable sensors," Annu Int Conf IEEE Eng Med Biol Soc, 2011, pp. 5839–5842, 2011.
[34] A. Wittmann et al., "Self-directed arm therapy at home after stroke with a sensor-based virtual reality training system," J Neuroeng Rehabil, vol. 13, no. 1, 2016, Art.no. 75.
[35] L. Yu et al., "A compressed sensing-based wearable sensor network for quantitative assessment of stroke patients," Sensors (Basel), vol. 16, no. 2, 2016, Art. no. 202.
[36] Z. Zhang et al., "Objective assessment of upper-limb mobility for poststroke rehabilitation," IEEE Trans Biomed Eng, vol. 63, no. 4, pp. 859–868, 2016.
[37] M. Panwar et al., "Rehab-Net: deep learning framework for arm movement classification using wearable sensors for stroke rehabilitation," IEEE Trans Biomed Eng, vol. 66, no. 11, pp. 3026–37, 2019.
[38] B. Oubre et al., "Estimating upper-limb impairment level in stroke survivors using wearable inertial sensors and a minimally-burdensome motor task," IEEE Trans Neural Syst Rehabil Eng, vol. 28, no. 3, pp. 601–611, 2020.
[39] P. Thanh Noi and M. Kappas, "Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for land cover classification using Sentinel-2 imagery," Sensors (Basel), vol. 18, no. 1, 2017, Art. no. 18.
[40] R.C. King, et al., "Application of data fusion techniques and technologies for wearable health monitoring" Med Eng Phy, vol. 42, pp. 1–12, 2017.
[41] H. Qiu et al., "Application of wearable inertial sensors and a new test battery for distinguishing retrospective fallers from non-fallers among community-dwelling older people," Sci Rep, 2017, vol. 8, no. 1, Art. no. 16349.
[42] I. C. Gyllensten and A. G. Bonomi, "Identifying types of physical activity with a single accelerometer: evaluating laboratory-trained algorithms in daily life," IEEE Trans Biomed Eng, vol. 58, no. 9, pp. 2656–2663, 2011.
[43] E. Preatoni et al., "Supervised machine learning applied to wearable sensor data can accurately classify functional fitness exercises within a continuous workout," Front Bioeng Biotechnol, vol. 8, 2020, Art. no. 664.
[44] L. Uusitalo, "Advantages and challenges of Bayesian networks in environmental modelling," Ecol Modell, vol. 203, no. 3–4, pp. 312–318, 2007.
[45] S. B. Kotsiantis, et al., "Machine learning: a review of classification and combining techniques," Artif Intell Rev, vol. 26, pp. 159–190, 2006.
[46] R. Yamashita et al., "Convolutional neural networks: an overview and application in radiology," Insights Imaging, vol. 9, no. 4, pp. 611–629, 2018.
[47] S. Lin et al., "Is extreme learning machine feasible? A theoretical assessment (Part II)," IEEE Trans Neural Netw Learn Syst, vol. 26, no. 1, pp. 21–34, 2015.
[48] A. Statnikov et al., "A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification," BMC Bioinformatics, vol. 9, 2008, Art. no. 319.
[49] M. Awad and R. Khanna, "Support Vector Machines for classification," Efficient Learning Machines Apress Berkeley, 2015, ch. 3, pp. 39–66.
[50] S. Patel et al., "A review of wearable sensors and systems with application in rehabilitation," J Neuroeng Rehabil, vol. 9, 2012, Art. no. 21.
[51] D. Dutta et al., "Development of a smart glove for affordable diagnosis of stroke-driven upper extremity paresis," Int Conf Comp Elect Comm Eng, 2020, pp. 1–7.
[52] N. Yozbatiran et al., "A standardized approach to performing the action research arm test," Neurorehabil Neural Repair, vol. 22, no. 1, pp. 78–90, 2008.
[53] 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, 2010.
[54] J. Fridolfsson et al, "Effects of frequency filtering on intensity and noise in accelerometer-based physical activity measurements," Sensors, vol. 19, no. 9, 2019, Art. no. 2186.
[55] W. H. Press and S. A. Teukolsky, "Savitzky-Golay smoothing filters," Comput Phys, vol. 4, 1990, Art. no. 669.
[56] M. AlMahamdy and H. B. Riley, "Performance study of different denoising methods for ECG signals," Proc Comp Sci, vol. 37, pp. 325–332 2014.
[57] M. U. A. Bromba and H. Ziegler, "Application hints for Savitzky-Golay digital smoothing filters," Anal Chem, vol. 53, no. 11, pp. 1583–86, 1981.
[58] T. Hester et al., "Using wearable sensors to measure motor abilities following stroke," Int Workshop Wearable Implant Body Sensor Netw, 2006, pp. 4–8.
[59] N. Hogan and D. Sternad, "Sensitivity of smoothness measures to movement duration, amplitude, and arrests," J Mot Behav, vol. 41, no. 6, pp. 529–534, 2009.
[60] R. Young and R. Marteniuk, "Acquisition of a multi-articular kicking task: Jerk analysis demonstrates movements do not become smoother with learning," Human Mov Sci, vol. 16, pp. 677–701, 1997.
[61] S. Pincus and W. Huang, "Approximate entropy: statistical properties and applications," Comm Stat, vol. 21, no. 11, pp. 3061–3077, 1992.
[62] P. Zhu et. al., "Characterization of the stroke-induced changes in the variability and complexity of handgrip force," Entropy (Basel), vol. 20, no. 5, pp. 550–560, 2018.
[63] P. Ahlgren et al., "Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient," J Am Soc Inf Sci, vol. 54, pp. 550–560.
[64] J. Antonio et al. "Low-cost wearable data acquisition for stroke rehabilitation: A proof-of-concept study on accelerometry for functional task assessment," Topics Stroke Rehabil, vol. 21, no. 1, pp, 12–22, 2014.
[65] M. Airaksinen et al. "Automatic posture and movement tracking of infants with wearable movement sensors," Sci Rep, vol. 10, no. 1, 2020, Art. no. 169.
[66] Kononenko et al., "Overcoming the myopia of inductive learning algorithms with RELIEFF," Appl Intell, vol. 7, pp. 39–55, 1997.
[67] R. J. Urbanowicz et al., "Relief-based feature selection: introduction and review," J Biomed Informatics, vol. 85, pp. 189–203, 2018.
[68] B. Lin et al., "Data glove system embedded with inertial measurement units for hand function evaluation in stroke patients," IEEE Trans Neural Syst Rehabil Eng, vol. 25, no. 11, pp. 2204–2213, 2017.
[69] Y. Shihong, et al., "SVM classification: its contents and challenges," Appl Math, vol. 18, pp. 332–342, 2003.
[70] J. A. Matthews, "Support vector machine (SVM)," in Encyclopedia of Environmental Change, 2014.
[71] Chih-Wei Hsu et al., "A practical guide to support vector classification," BJU Int, vol. 101, no. 1, pp. 1396–1400, 2008.