Use of Functional Data to Model the Trajectory of an Inertial Measurement Unit and Classify Levels of Motor Impairment for Stroke Patients

Yi-Ting Hwang, Wei-An Lu, and Bor-Shing Lin, *Senior Member, IEEE*

***Abstract*— Motor impairment evaluations are key rehabilitation-related assessments for patients with stroke. Currently, such evaluations are subjective; they are based on physicians' judgements regarding the actions performed by patients. This leads to inconsistent clinical results. Many inertial sensing elements for motion detection have been designed. However, to more easily and rapidly evaluate motor impairment, we require a system that can collect data effectively to predict the degree of motor impairment. Lin *et al.* used data gloves equipped with an inertial measurement unit (IMU) to collect movement trajectories for motor impairment evaluations in patients with stroke. The present study used functional data analysis to model data trajectories to reduce the influence of noise from IMU data and proposed using coefficients of function as features for classifying motor impairment. To verify the appropriateness of feature construction, five classification methods were used to evaluate the extracted features in terms of the overall and sensor-specific ability to classify levels of motor impairment. The results indicated that the features derived from cubic smoothing splines could effectively reflect key data characteristics, and a support vector machine yielded relatively high overall and sensor-specific accuracy for distinguishing between levels of motion impairment in patients with stroke. Future data glove systems can contain cubic smoothing splines to extract hand function features and then classify motion impairment for appropriate rehabilitation programs to be prescribed.**

***Index Terms*— Functional data, inertial measurement unit, motor impairment, spline, stroke.**

Manuscript received September 16, 2021; revised January 6, 2022 and February 27, 2022; accepted March 21, 2022. Date of publication March 25, 2022; date of current version April 18, 2022. This work was supported in part by the Ministry of Science and Technology in Taiwan under Grant MOST 109-2314-B-305-001 and Grant MOST 109-2221-E-305-001-MY2, in part by the University System of the Taipei Joint Research Program under Grant USTP-NTPU-TMU-109-03, in part by the Faculty Group Research Funding Sponsorship by the National Taipei University under Grant 2021-NTPU-ORDA-02, and in part by the "Academic Top-Notch and Features Field Project" funding sponsorship from the National Taipei University under Grant 109-NTPU_ORDA-F-005. *(Corresponding author: Bor-Shing Lin.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Chi-Mei Hospital under IRB No. 10102-019.

Yi-Ting Hwang and Wei-An Lu are with the Department of Statistics, National Taipei University, New Taipei City 23741, Taiwan (e-mail: hwangyt@gm.ntpu.edu.tw; vivialex96@gmail.com).

Bor-Shing Lin is with the Department of Computer Science and Information Engineering, National Taipei University, New Taipei City 23741, Taiwan (e-mail: bslin@mail.ntpu.edu.tw).

Digital Object Identifier 10.1109/TNSRE.2022.3162416

I. Introduction

CEREBROVASCULAR diseases (CVDs), the leading cause of death and disability worldwide, were responsible for an estimated 17.9 million deaths globally in 2019 [1]. According to data from the Taiwanese government and in 2020, CVDs were the fourth leading cause of death and approximately 50.1 of 100000 Taiwanese residents died of CVDs [2]. Nevertheless, the number of stroke survivors has been increasing owing to considerable improvements in acute medical care. As such, the need for rehabilitation to restore physical function has dramatically increased. Approximately two-thirds of survivors can walk independently after a stroke [3], [4], but less than half of them regain basic upper limb function within a year of rehabilitation [5]. Stroke limits their independence in activities of daily living and reduces their quality of life. Thus, upper limb rehabilitation is crucial for stroke survivors.

Appropriate rehabilitation can effectively aid patients in achieving favorable recovery that enables them to perform functional activities. A person's rehabilitation plan is related to their disability level. Clinical methods are available for evaluating functional recovery, including the Fugl-Meyer test, action research arm test, box and block test, Jebsen–Taylor hand function test, and Brunnstrom stage (BS) test. As demonstrated by [6], the outcomes of the Fugl-Meyer test, action research arm test, and box and block test are strongly correlated, and these tests have high test–retest reliability according to standardized guidelines. However, these clinical evaluation methods are subjective, and scores for the same patient can vary among clinicians. Furthermore, without detailed movement-related information, providing an individualized rehabilitation plan can be difficult [7].

To reduce the burden on clinicians and improve the quality of care, more accurate and efficient evaluations are required for quantifying disability levels. Several tools for evaluating upper limb and hand functions have been developed in recent years. [8] designed a triaxial accelerometer-based system to quantify the clinical features of Parkinson disease. [9], [10] have used magnetic and inertial sensors to collect data when patients performed the Movement Disorder Society–Unified Parkinson's Disease Rating Scale finger tapping (FT) task, and [11] used inertial sensors to collect data and determine the

features of parkinsonian tremors. Additionally, [12] designed and implemented a smartphone-based system for automated motor assessment by using low-cost off-the-shelf inertial sensors for measuring the movement of the joint angle of the upper body in patients with stroke. Furthermore, [13] designed an inexpensive and portable motion capture system by using a single reflective marker on the wrist to measure the kinematic movements of stroke survivors. For measuring hand kinematics, [14]–[17] have designed data gloves with inertial and magnetic sensors to detect joint movements in various directions.

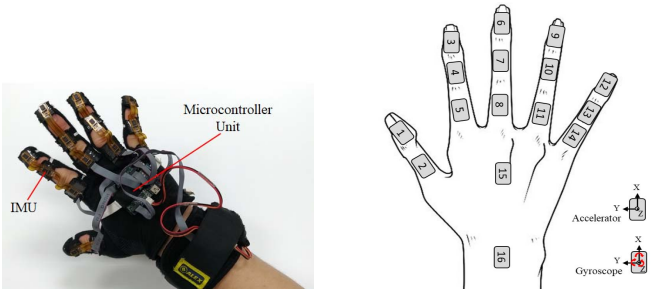
Sensors can be easily used to collect a large amount of data. Depending on the number of sensors and recording time, various functional trajectories are available for each participant. To conduct further analyses, these variable trajectories must be manipulated. [13] introduced five components, which were crudeness, two speed profile deviations, smoothness, and segmentation, to capture variable trajectories, and these components were measured using confidence scores, which required a reference trajectory and several threshold values. On the basis of physical characteristics, [16] developed only three features, namely average rotation speed, variations in movement completion time, and quality of movement, from raw data. [18] computed summary statistics, such as centrality and variablity, and the extremes of variable trajectories, which yielded over 3000 features. Related raw data were collected from signals transmitted from multiple devices, and data spikes, packet loss, and data gaps were noted when the server could not receive and process data instantly [19]. Furthermore, an inherent deficiency in related devices, such as clock stretching, can result in sporadic incorrect data points [20]. These outliers greatly affect the accuracy of summary statistics, especially extreme values.

Only some data trajectories have been examined in the literature. This study introduced a functional analysis that can capture entire data trajectories. Through a prespecified basis function and an objective function, a functional form is used to fit data trajectories. Instead of processing raw data, this functional form eliminates the noise induced by data collection [21]. This study investigated the feasibility of using functional analysis to analyze sensor data. Unlike summary statistics, which require the use of principle component analysis (PCA) to reduce the number of features [18], in functional analysis, the number of features depends on the number of bases and is determined objectively through crossvalidation. The feasibility of features derived from the functional approach was demonstrated by evaluating the accuracy in determining patients' motor impairment level using the data collected by [16]. Furthermore, because these features provide sufficient information on data trajectories, the importance of sensors was also examined, and the results can further reduce the cost of developing data gloves.

II. Methods

A. Participants

This study enrolled 15 patients with stroke and 15 healthy age-matched individuals (H). The inclusion criteria for the study were 1) age of 20–80 years and 2) the ability to maintain a sit-up position for longer than 40 min. The exclusion criteria included 1) a diagnosis of hemispatial akinesia or visual attention deficit, 2) cognitive impairment, 3) joint defects in the upper extremities prior to a stroke. In this study, the BS test was used to evaluate motor impairment level. Of the 15 patients with stroke, 4 had BS4, 10 had BS5, and 1 had BS6. A therapist determined and provided patients' BS as the ground truth for classification. Because patients at BS6 with stroke typically have hand function that is clinically similar to that of healthy individuals, the patient with BS6 was treated as a healthy individual.



(a) Data glove (b) Positions

Fig. 1. (a) Photograph of the data glove. (b) Positions of the IMU sensors.

The average ages for healthy individuals and patients were 62.6 and 59.3 years, respectively. Nine patients and five healthy participants were men. This study was performed at Chi-Mei Hospital, Tainan, Taiwan, and was approved by the Ethics Committee of Chi-Mei Hospital (IRB No. 10102-019). All participants provided written informed consent.

B. Device and Task

Self-developed data gloves containing 16 inertial measurement units (IMUs; LSM330DLC; STMicroelectronics, Geneva, Switzerland) were used for data collection. A gyroscope and an accelerometer were included in each IMU to capture data on triaxial angular velocity and acceleration. The IMU data were collected by a microcontroller unit (MSP430; Texas Instruments, Dallas, TX, USA) and were wirelessly transmitted as encapsulated packets to a laptop through a Bluetooth interface. Fig. 1 (a) presents a photograph of the data glove, and Fig. 1 (b) displays the IMU sensor positions on the glove. The detailed mechanical design is provided in [16].

All participants were asked to wear the data glove to detect their hand motions when performing a thumb task (TT) and a grip task (GT). The tools for these tasks are illustrated in Fig. 2. The TT was designed to evaluate thumb dexterity. In this task, each participant was asked to hold a cylinder in their hand [Fig. 2 (a)] and then use their thumb to repeatedly push a button on the cylinder. A complete motion was defined as the completion of a press and release action. The GT was aimed at assessing the dexterity of the entire hand. For this task, each participant was asked to perform a grip and release motion using the tool displayed in Fig. 2 (b). The grip and release motions constituted a single complete motion. For both the GT and TT, each participant was asked to perform a complete motion within 4 s and repeat it ten times. For the TT and GT, a complete motion comprised ten cycles.

(a) (b)

Fig. 2. (a) TT and (b) GT tools.

TABLE I
DATA RECORDING

Task	Normal	BS5	BS4	Total
Thumb task (TT)	29	23	8	60
Grip task (GT)	33	22	8	63
Total	62	45	16	123

Depending on their physical condition, each participant was asked to repeat each task two or three times. In total, 60 and 63 sets of measurements were obtained for the TT and GT, respectively, and 62 and 61 sets of measurements were obtained from healthy individuals and patients with stroke, respectively (Table I). The data were collected from 16 IMU sensors. Furthermore, acceleration and angular velocity in three dimensions were extracted for each sensor.

C. Data Analysis

Because the sensors collected data every 23 ms, the data were manually divided into ten cycles to exclude the influence of between-cycle data. Although the participants were instructed to complete both the TT and GT in 4-s cycles, the actual cycle duration varied among them. To determine the acceleration and angular velocity trajectories, the time space was adjusted to [0, 1]. These trajectories were thus a function of time, and functional data analysis could be used to model the data trajectories.

Let y_i denote the outcome measured at time x_i, where $i = 1, \cdots, N$. The general model is expressed as follows:

$$y_i = f(x_i) + \epsilon_i, i = 1, \cdots, N, \tag{1}$$

where $f(\cdot)$ is a prespecified smoothing function, N is the number of sample points, and ϵ_i denotes independent zero-mean random variables. The smoothing function is often expressed in terms of basis functions as follows:

$$f(t) = \sum_{j=1}^{k} \phi_j(t) c_j, \tag{2}$$

where k is a prespecified number of bases, $\phi_j(\cdot)$ is a prespecified jth basis, and c_j is the corresponding jth coefficient. Although the task motions were performed periodically, the duration of each cycle varied among participants. In this study, two basis systems were thus considered: truncated power base splines (i.e., regression splines; commonly used for nonperiodic functional data) and Fourier bases (commonly used for periodic functional data).

1) Regression Splines: To define a polynomial spline, the range is first divided into T disjoint subintervals by t_l, where $l = 1, \ldots, T$ (called knots) and t_0 denotes the lower bound of the time interval. For a subinterval (t_l, t_{l+1}), a truncated power function of order d is used as follows:

$$\phi_j(t) = t^{j-1}, 1 \le j \le (d+1) \tag{3}$$

$$\phi_{d+j}(t) = (t - t_j)_+^d, 1 \le j \le T, \tag{4}$$

where $(t)_+ = t$ if $t > 0$; otherwise, $(t)_+ = 0$. The number of parameters is $(d+1) \times (T+1) - T \times d = d + T + 1$. According to [22], a cubic spline ($d = 3$) is sufficient to fit the data trajectory properly when a sufficient number of knots is considered. The bases are presented as follows: $\phi_1(t) = 1$, $\phi_2(t) = t$,

$$\phi_{j+2}(t) = d_j(t) - d_{T-1}(t), 1 \le j \le T-2, \tag{5}$$

where

$$d_j(t) = \frac{(t - t_j)_+^3 - (t - t_T)_+^3}{t_T - t_j}. \tag{6}$$

Each of these basis functions satisfies $d''(\cdot) = d'''(\cdot) = 0$.

To prevent the selection of a maximal set of knots, the penalized residual sum of squares was used as follows:

$$\mathrm{RSS}(f, \lambda) = \sum_{i=1}^{N} (y_i - f(x_i))^2 + \lambda \int (f''(t))^2 \, dt, \tag{7}$$

where $\lambda \in (0, \infty)$, denoting a fixed smoothing parameter, is used to identify the estimator of c_js. An explicit unique minimizer, a natural spline with knots at the unique values of x_i for $i = 1, 2, \ldots, N$ can be obtained as follows:

$$f(t) = \sum_{j=1}^{N} \phi_j(t) c_j, \tag{8}$$

where $\phi_j(t)$ is an N-dimensional set of basis functions for natural splines. The criterion defined in (7) can be reduced as follows:

$$RSS(\mathbf{c}, \lambda) = (\mathbf{y} - \Phi\mathbf{c})'(\mathbf{y} - \Phi\mathbf{c}) + \lambda \mathbf{c}' \mathbf{\Omega}_\phi \mathbf{c}, \tag{9}$$

where $\{\mathbf{\Phi}\}_{ij} = N_j(x_i)$ and $\{\mathbf{\Omega}\}_{jk} = \int \phi_j''(t)\phi_k''(t)dt$. The tuning parameter λ determines the smoothness of the fitted curve.

The multivariate adaptive regression spline procedure is then used to determine a suitable set of optimally positioned knots. The optimal value of λ is determined through generalized crossvalidation and is represented as

$$\mathrm{GCV}(\lambda) = \frac{\sum_{i=1}^{N} (y_i - \hat{f}_\lambda(x_i))^2}{(1 - M(\lambda)/N)^2}, \tag{10}$$

where $M(\lambda)$ is the effective number of parameters in the model, including λ, the number of knots, and the optimal positions of the knots.

2) Fourier Series: A Fourier series uses sin and cos functions as the basis functions. In other words, we have $\phi_{2r-1}(t) = \sin(r\omega t)$ and $\phi_{2r}(t) = \cos(r\omega t)$, where ω represents the period. The Fourier series is then defined as follows:

$$f(t) = c_0 + c_1 \sin \omega t + c_2 \cos \omega t + c_3 \sin 2\omega t + c_4 \cos 2\omega t + \cdots. \tag{11}$$

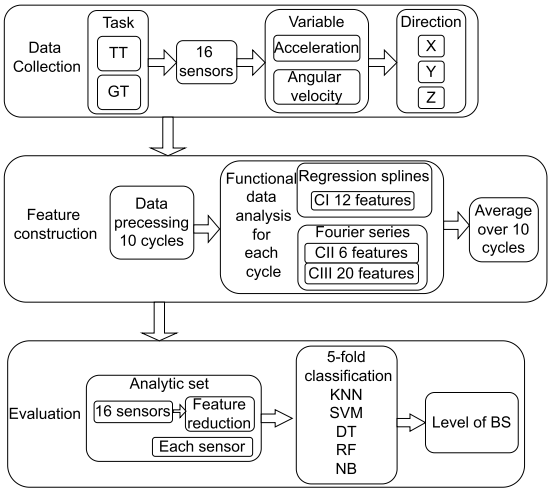
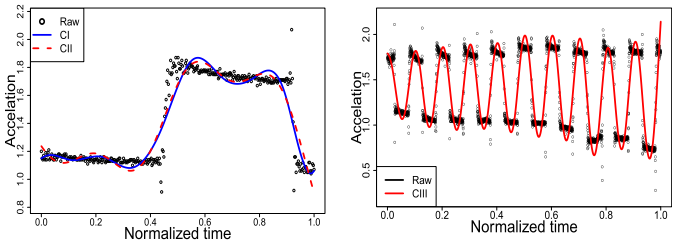


Fig. 3. Data analysis flowchart.



(a) (b)

Fig. 4. Plot showing raw data, (a) the fitted smoothing spline and Fourier series for one cycle, and (b) Fourier series for ten cycles.

D. Methodological Summary

As stated, in the GT and TT, a complete motion comprised ten cycles. A smoothing trajectory based on the regression spline was obtained for each of these cycles. For the Fourier series, a smoothing trajectory was obtained for each of these cycles and also for ten cycles. The feasibility of these feature constructions for classifying the level of stroke-related motor impairment was evaluated using the k-nearest neighbor (KNN), support vector machine (SVM), decision tree (DT), random forest (RF), and naive Bayes classifier (NB) methods. The results of two assessments were considered. The first was an overall evaluation of the entire system, and the second was used examine the importance of each of the 16 sensors. Owing to the limited numer of measurements taken, feature reductions were required for the first assessment. Fig. 3 illustrates the flow of the data analysis.

1) Feature Construction: The regression spline was applied using the *smoothing.spline* package in R, and the default hyperparameters were used, except for the maximum number of knots (nknots). In this study, 10 knots was the maximum. The resulting smoothing splines had 12 coefficient estimates. For an individual variable, ten sets of coefficient estimates and their corresponding averages were obtained. This was the first feature type and was termed CI.

The Fourier function in R was used for data fitting. When the smoothing trajectory was being derived for only one cycle, three Fourier bases provided a satisfactory fit for each cycle and resulted in six coefficient estimates. The average estimates of six coefficients corresponding to ten cycles were then obtained. This was the second feature type and was termed CII.

In accordance with the preliminary data exploration related to various bases, ten Fourier bases were selected for ten cycles and resulted in 20 coefficient estimates. This was the third feature type and was termed CIII. The fitted curves for acceleration are provided in Fig. 4 (a) and (b). These fitted curves could capture trajectories and were not influenced by outliers.

2) Feature Reductions: In each recording, each of the 16 sensors collected data on six variables. In other words, data on 96 variables were collected in each recording. Feature construction was performed for each variable. The total number of CI, CII, and CIII features in each recording was equal to 96 variables multiplied by the number of coefficients for each construction, (1152, 576, and 1920, respectively). As indicated in Table I, the number of recordings was relatively small. To avoid overfitting in the classification, one-way analysis of variance and a Wilcoxon signed-rank test were performed to identify the essential features. Heat maps were generated to present the importance of features on the basis of these two tests. The columns present the features, and the rows present the triaxial (X, Y, and Z) acceleration (A) and angular velocity (G); all these data were captured by the 16 sensors. Different colors represent different levels of significance. A feature was regarded as essential if and only if both p values derived from these two tests were significant (< 0.05).

TABLE II
NUMBER OF ESSENTIAL FEATURES DERIVED FOR DIFFERENT MEASURES AND THE DIRECTIONS OF EACH CONSTRUCTION FROM THE TT AND GT

		Measure													
		Acceleration Direction						Angular velocity Direction							
		X		Y		Z		X		Y		Z		Total	
Task		n	%	n	%	n	%	n	%	n	%	n	%	n	%
TT	CI	84	7.3	59	5.1	91	7.9	4	0.3	43	3.7	13	1.1	294	25.5
	CII	39	6.8	5	0.9	32	5.6	4	0.7	27	4.7	12	2.1	119	20.7
	CIII	53	2.8	30	1.6	64	3.3	9	0.5	108	5.6	65	3.4	329	17.1
GT	CI	78	6.8	71	6.2	119	10.3	24	2.1	154	13.4	26	2.3	472	41.0
	CII	63	10.9	18	3.1	65	11.3	13	2.3	60	10.4	25	4.3	244	42.4
	CIII	165	8.6	22	1.1	61	3.2			241	12.6	41	2.1	530	27.6

Table II summarizes the number of essential features derived using different measures and the directions of CI, CII, and CIII from the TT and GT. For the TT, 25.5%, 20.7%, and 17.1% of the features constructed from CI, CII, and CIII, respectively, were identified as essential features. CI yielded a slightly larger number of essential features than did CII and CIII. Among the essential features in CI, 20.3% were derived from acceleration and were equally distributed in terms of direction. Two-thirds of the essential features derived from CII were also derived from acceleration but were primarily distributed in the X and Z directions. The essential features derived from CIII were obtained in equal number from acceleration and angular velocity. Most essential features obtained from acceleration were derived from the X and Z directions, whereas the largest number of essential

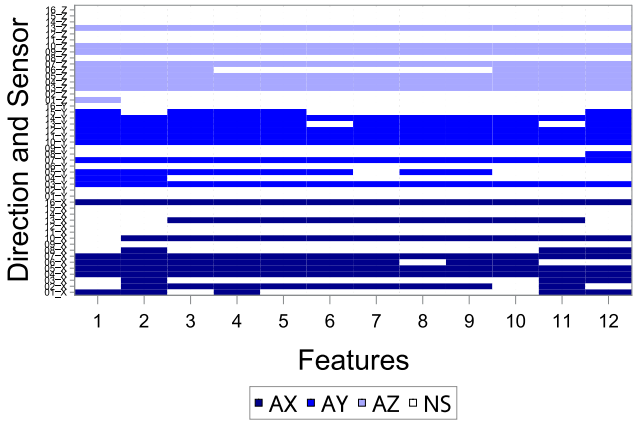


Fig. 5. Heat map of the *p* values of features constructed from CI on the basis of TT data, where colored areas represent essential features from different sensors and directions for A. NS denotes nonsignificance.

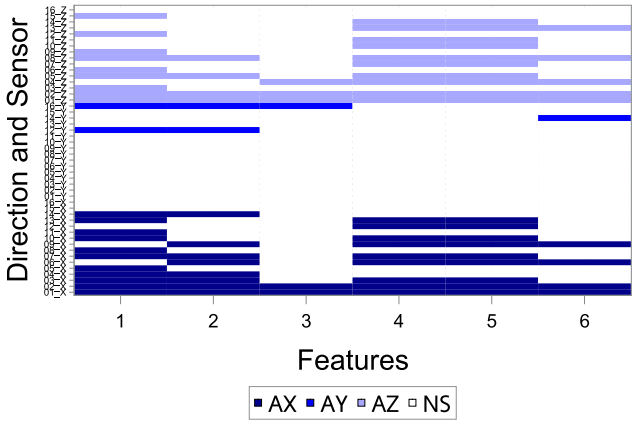


Fig. 6. Heat map of *p* values of features constructed from CII on the basis of TT data, where colored areas represent essential features from different sensors and directions for A. NS denotes nonsignificance.

features obtained from angular velocity were derived from the Y direction.

For the GT, a larger number of essential features was provided; 41%, 42.4%, and 27.6% of such features were constructed from CI, CII and CIII, respectively. Different from the TT, for the GT, the number of essential features derived from acceleration and angular velocity were similar for all constructions. However, differences existed in their extraction in terms of direction. The essential features derived from acceleration were primarily collected from the X and Z directions, whereas those from G were from the Y direction. No essential feature was derived from the X direction for CIII.

The heat maps of the features related to the TT and GT are presented in Fig. 5; the columns contain information on the features and the rows contain information on measurements, directions, and sensors. Different colors are used to differentiate between the information presented in rows. Blue and red indicate essential features for A and G, respectively, and a gradient of colors from dark to light shades distinguishes between the X, Y, and Z directions.

The blue column in Fig. 5 represents approximately 80% of the essential features selected from A. The distribution of color depth was relatively uniform, and the colored column areas were more concentrated for sensors 3, 4, 5, 7, 10, and 13. The blue columns have had an feature distribution, but the red columns had a higher concentration of features 5 and 10 (data not shown).

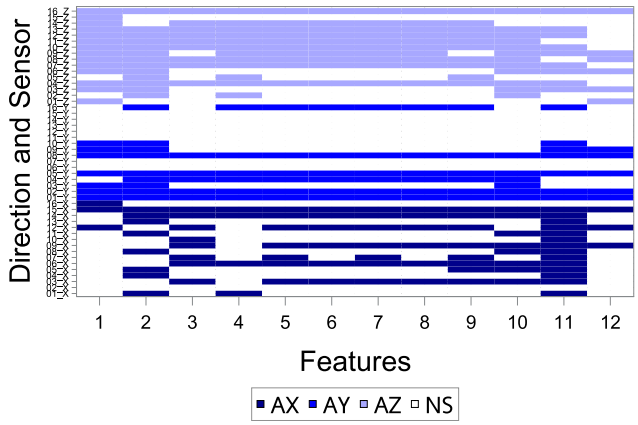
Most of the essential features for CII were features 1, 4, and 5 derived from the X and Z directions of A, as indicated in Fig. 6. For G, almost all six features were obtained from the Y direction of sensors 1, 2 and 3, and features 1–5 were obtained from the Z direction of sensors 3 and 4 (data not shown). For CIII, most of the essential features derived from A were features 11–20, and those from G were features 1–10 (data not shown).

Fig. 7 (a) and (b) display the essential features for A and G, respectively. For A, all 12 features were essential features. They were derived from sensor 15 from the X direction; sensors 1, 2, 5, and 8 from the Y direction; and sensor 16 from the Z direction. In addition, 15 features from sensors 4, 7, 8, 10, 12, and 13 from the Z direction were essential features, and Feature 11 from sensors 1 and 3–15 from the X direction was an essential feature. As indicated in Table II, most of the essential features for G were derived from the Y direction [see Fig. 7 (b)]. Feature 6 from the Y direction of all sensors was determined to be essential. Except for feature 12, most of the other features from sensors 1–14 were determined to be essential.

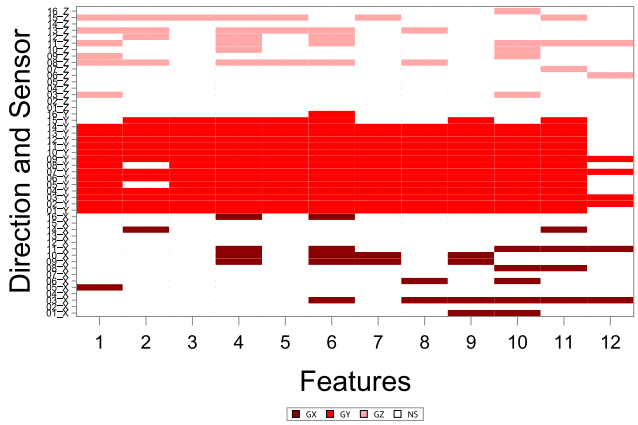
For CII, all features from the X direction of sensors 4, 5, 8, 10, 11, 13, and 14 and the Z direction of sensors 3, 4, 6, 7, 9, 10, and 12 for A were essential features, as shown in Fig. 8. For G, all features from the Y direction of sensors 2, 5, and 8–14 were essential features, and all features from the Z direction of sensor 9 were essential features. A similar pattern was also observed for CIII (data not shown).

3) Evaluations: Classification procedures were performed separately for the TT and GT. The results from each classification are summarized in a confusion matrix. Let n_{ij} represent the number of observations for the ith observed value and the jth predicted value, where 1, 2, and 3 represent individuals with a healthy status, BS5, and BS4, respectively [23]. The overall accuracy was determined by dividing the number of correct predictions by the number of observations ($\sum_{i=1}^{3} n_{ii} / \sum_{ij} n_{ij}$). Furthermore, the accuracy of classifications in identifying individuals without stroke, patients with BS5, and patients with BS4 was also computed. The accuracy in identifying individuals without stroke was determined by the number of correct predictions of individuals without stroke divided by the number of observations for that group. The other definitions were similar to the aforementioned one. Finally, the average overall accuracy was determined, as was the accuracy over 30 task repetitions by individuals without stroke, patients with BS5, and patients with BS4.

The classification was performed using the train function in R's *Caret* package. The training control is based on a fivefold crossvalidation scheme. For the KNN method, *tuneGrid* in the *Caret* package was applied such that the main parameters were decided automatically, and the value of k was set from 1 to 30. In accordance with a suggestion in [24], the radial basis function kernel was used for the SVM classification, and the

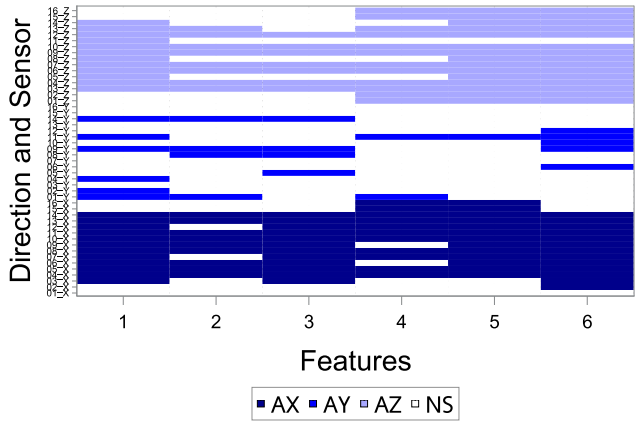


(a)

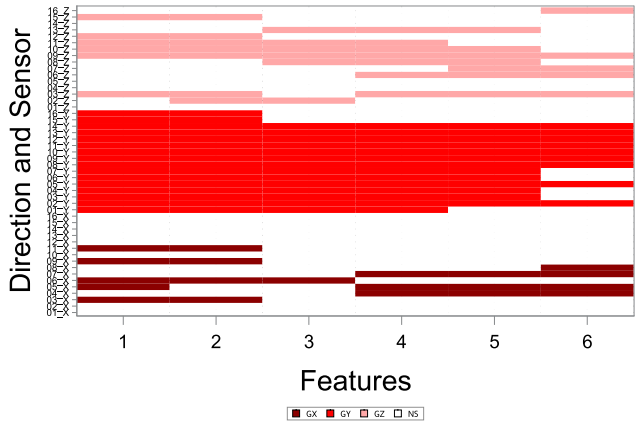


(b)

Fig. 7. Heat map of p values of features constructed from CI on the basis of GT data, where colored areas represent essential features from different sensors and directions for (a) A and (b) G. NS denotes nonsignificance.



(a)



(b)

Fig. 8. Heat map of the p values of features constructed from CII on the basis of GT data, where colored areas represent essential features from different sensor and directions for (a) A and (b) G. NS denotes nonsignificance.

total number of unique combinations was set to 10 for the grid search. The cp value in *tuneGrid* under the DT method was set to apply searches from 0 to 0.005 in steps of 0.0005. The mtry value was set to 1:15 for *tuneGrid* under the RF method. Under the NB method, the rf parameter was set to search from 0 to 5 in steps of 1, the usekernel parameteer was set as TRUE, and the adjusted parameter was set to search from 0 to 5 in steps of 1. All these classification processes were repeated 30 times to determine the final accuracy ($\bar{A}$). Furthermore, the sample standard deviation (SD) of the accuracy level after 30 repetitions was obtained to derive the error interval for average accuracy, where the lower and upper bounds were $\bar{A}-1.96\times\text{SD}$ and $\bar{A}+1.96\times\text{SD}$, respectively. The lower bound was set to 0 when the computed lower bound was negative. When the computed upper bound exceeded 1, the upper bound was set to 1.

III. Results

A. Data Classification Involving Three Groups

The essential features obtained from three constructions were used as classification inputs. Analyses were performed for each task as follows:

1) TT: The average overall accuracy from 30 repetitions of the TT was computed, as indicated in Fig. 9 (a). The highest overall classification accuracy was that for CI, as determined by the KNN method, followed by that for CI and CII, as determined by the SVM method. Regardless of the classification method used, the essential features in CI had the best classification performance of all features. The classification ability from the use of essential CII and CIII features varied based on classification method. Under the KNN, DT, and NB methods, CI yielded a slightly higher accuracy than did CII and CIII. The error interval for CI under the KNN method was slightly shorter than those for the other methods, whereas the error intervals under the DT and NB methods were slight longer than those for the other methods. The error intervals derived from CI, CII, and CIII for the SVM and RF methods mostly overlapped. Under the KNN method, the error interval derived from CI was significantly higher that derived from CIII. Three error intervals for the DT method were siginificantly lower.

Fig. 9 (b)–(d) presents the average accuracy in identifying individuals without stroke, patients with BS5, and patients with BS4. The highest average accuracy for individuals without

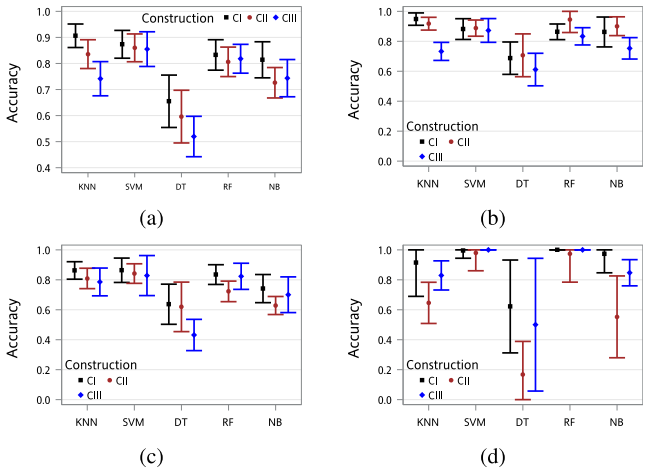


Fig. 9. Average accuracy on the basis of TT data: (a) overall, (b) for individuals without stroke, (c) for patients with BS5, and (d) for patients with BS4.

stroke was achieved using the RF method and essential CII features. The second highest accuracy was achieved using the KNN method and essential CI features. The performance of classification methods varied based on the construction of essential features. Except for when the the KNN and SVM methods were applied, the highest classification accuracy was observed when essential CII features were used. The essential CIII features had the poorest classification ability among all features. The error intervals for CI, CII, and CIII overlapped when the SVM method was applied, and the error interval was slightly shorter for CII than it was for CI and CIII. In particular, the lower bound of the error interval was very close to 0.8. The error interval derived from CII was slightly higher than those derived from CI and CIII when the RF method was used. The error interval of the average accuracy derived from CI and CII under the KNN method was significantly longer than that derived from CIII. The error interval was longer when the DT and NB methods were applied. The DT method provided significantly lower error intervals than did the other methods.

The highest average accuracy with which the patients with BS5 were identified slightly exceeded 0.85, as indicated in Fig. 9 (c). Except for the DT and NB methods, the classification methods through which the essential CI features were input performed favorably. Classification method performance varied when the essential CII and CIII features were input. Under the RF and NB methods, the average accuracy for CIII was slightly higher than that for CII. Compared with individuals without stroke, the error interval was slightly longer for patients with BS5. Under the SVM and KNN methods, the error intervals for CI, CII, and CIII overlapped, but the error intervals for CII obtained from the SVM and KNN methods were shorter than those for CI and CII. The error intervals for CI and CIII under the RF method were slightly longer than the relevant error interval for CII. Under the DT and NB methods, the error intervals for CI, CII, and CIII overlapped.

Only eight recordings were included for patients with BS4. Regardless of which construction was used, excellent classification ability was noted for this group under the SVM and RF methods. Except for when the DT method was applied, the average accuracy for CI was >0.91. For CIII, the average accuracy was >0.81. When the essential CII features were input, classification was only successfully performed under the SVM and RF methods. Overall, the DT method had the poorest classification ability. The longest error interval was observed for patients with BS4. The upper bound for the KNN, SVM, and RF methods was set to 1, and the lower bound for the DT method was set to 0. The length of the error interval derived from CI and CIII under the RF method was close to 0. That is, overfitting fitting was noted under the RF method for patients with BS4. Under the SVM method, the length of the error interval for CIII was close to 0, and the error intervals for CI and CII were slightly wider than that for CII. The KNN, DT, and NB methods provided a similar pattern of error intervals for CI, CII, and CIII, where the interval lengths derived from CIII for the KNN and NB methods and from CII for the DT method were slightly shorter. The performance of the DT method varied substantially across 30 repetitions.

2) GT: Fig. 10 (a)–(d) presents the average accuracy of 30 repetitions of the GT. The essential CI features under the SVM method had the highest overall average accuracy, followed by the essential CII features under the KNN method. Among constructions, the SVM method had the best classification ability. Except for those under the DT method, essential CI features had an average classification accuracy of >0.78. When the essential CII features were input, the average accuracy under the KNN and SVM methods was >0.84. Except for when the SVM method was used, the essential CIII features did not provide sufficient information to enable classification. The error interval patterns under the SVM, DT, and RF methods were similar. The error interval for CI was significantly higher than that for CIII. When the KNN method was used, the accuracy derived from three essential constructions differed significantly. The error interval derived from CII was higher than that from CI, and that from CIII was significantly lower. Most of the error intervals derived from CI, CII, and CIII overlapped under the NB method and the DT method.

Fig. 10 (b) presents the average accuracy for classifying individuals without stroke. When the essential CI features were used as inputs under the SVM method, the average accuracy was closer to 1. Except under the DT method, the average classification accuracy was >0.85 when the essential CI features were input, whereas that when the essential CII features were input was >0.80. The essential CIII features had the poorest classification ability of all features. The error intervals derived from CI under the SVM and RF methods were slightly longer than those from CII and CIII. The lower bound of the error interval for the SVM method exceeded 0.8. Under the KNN method, the lower bound of the error interval derived from CI and CII exceeded 0.8, whereas the upper bound under the error interval derived from CIII was close to 0.8. The patterns of error intervals under the DT and NB methods were similar.

The average accuracy for classifying patients with BS5 was similar to that for classifying individuals without stroke. Under

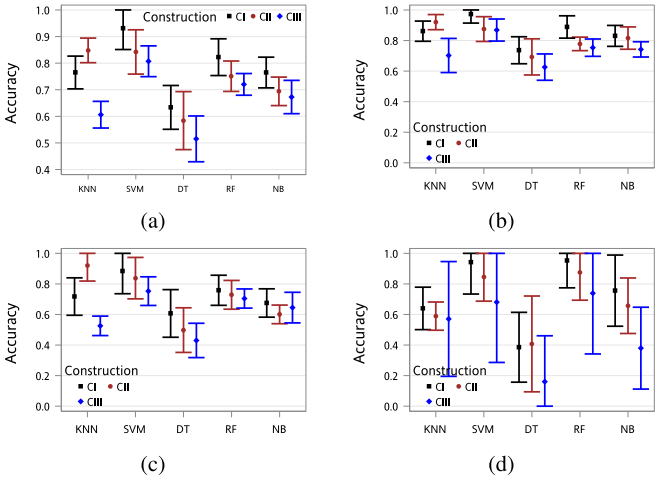


Fig. 10. Accuracy on the basis of the GT data: (a) overall, (b) for the individuals without stroke, (c) for the patients with BS5, and (d) for patients with BS4.

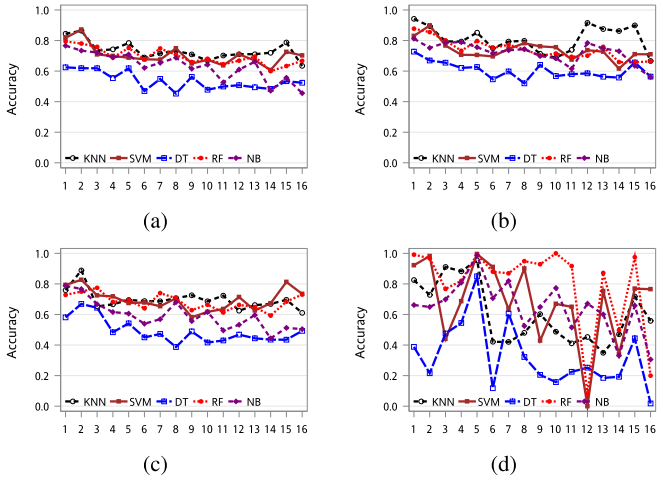


Fig. 11. Average accuracy of each sensor on TT in CI (a) all individuals, (b) individuals without stroke, (c) patients with BS5, and (d) patients with BS4.

the SVM method, when essential CI features were used, the highest average accuracy of 0.88 was obtained, as indicated in Fig. 10 (c). When the essential CII features were used, the second highest average accuracy of 0.85 was obtained under the SVM method, and the third highest average accuracy was obtained under the KNN method. Regardless of construction, the RF and NB methods had average accuracy levels slightly exceeding 0.7 and 0.6, respectively. Three error intervals overlapped under the SVM, DT, RF, and NB methods, where the error interval derived from CI was slightly higher than those derived from CII and CIII. The error interval patterns derived CI, CIII and CIII under the KNN method were similar to those for the individuals without stroke.

Under the RF and SVM methods, when the essential CI features were used, the average accuracies obtained were approximately 1 and 0.94, respectively [Fig. 10 (d)]. Regardless of essential feature construction, the average accuracy under the RF method was >0.9. Classification was not possible under the DT method. The patterns of error intervals under the SVM and RF methods were similar. The interval derived from CI was significantly shorter that that derived from CIII. Under the KNN method, the the error intervals derived from CI and CII were significantly shorter than the interval derived from CIII. All the error intervals under the DT and NB methods were wide.

B. Sensor-Specific Classification Among Three Groups

The essential information derived from the sensor data varied with task design. The classification performance based on sensor data is discussed as follows. All of the features derived from CI were input.

1) TT: Fig. 11 (a)–(b) presents the average classification accuracy for each sensor worn during the TT for the CI feature type.

When the essential features derived from CI were used, the classification accuracy for most of the sensors was comparable under the KNN, SVM, and RF methods [Fig. 11 (a)]. The features derived from the data collected by sensors 1 and 2 provided valuable information for differentiation between participants. The average accuracy obtained under the KNN method was higher than those obtained under other methods for most sensor data. When sensors 6, 10, and 16 were excluded, the average accuracy was >0.7.

Fig. 11 (b) presents the average accuracy obtained using the CI features for individuals without stroke. The highest average accuracy was again derived from the features extracted from data collected by sensors 1 and 2. The KNN method used features from sensor 1 data, and the SVM method used features from sensor 2 data. As for the features obtained from the data of most sensors, the KNN method had higher classification power than other methods. The performance of the SVM, RF, and NB methods was comparable for data collected by sensors other than sensors 1 and 2.

The results related to average accuracy in classifying patients with BS5 are detailed in Fig. 11 (c). Data from sensor 2 provided more information than those from other sensors for identifying individuals in this group under the KNN and SVM methods. In contrast to the results for individuals without stroke, under the SVM method, features were derived from sensor 15 data. For the remaining sensor data, the performance of the KNN, SVM, and RF methods was comparable.

The performance of the methods varied considerably in classifying patients with BS4. For sensors 1, 2, 5, 8, 10, and 15, the RF method outperformed the other methods, and the average accuracy exceeded 0.95. For sensors 1, 2, 5, 6, and 8, the SVM method also yielded an average accuracy of >0.9. However, when sensor 12 data were used, differentiation between participants was not possible under the SVM and RF methods. When data from sensors 1, 3, 4, and 5 were used, the average accuracy obtained under the KNN method was >0.85. Finally, the NB method yielded a slightly higher average accuracy when data from sensors 4, 5, and 7 were used.

The average accuracy for each group for CII and CIII had similar patterns but was lower than the overall average accuracy (data not shown).

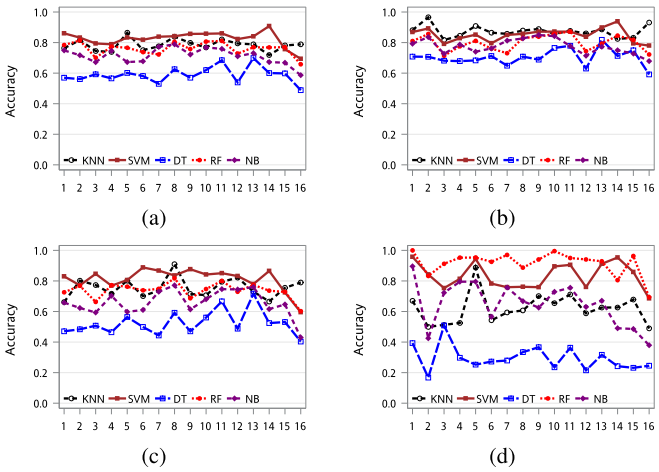


Fig. 12. Average accuracy of each sensor for the GT in CI for (a) all participants, (b) individuals without stroke, (c) patients with BS5, and (d) patients with BS4.

2) GT: Fig. 12 (a) presents the average accuracy for CI on the GT. Most of the sensor data provided sufficient information for the SVM method to differentiate between the participants. Almost all the sensor data had an average accuracy of >0.8. The performance of the KNN and RF methods was comparable, and the average accuracy slightly exceeded 0.75.

The results related to the average accuracy in classifying individuals without stroke, patients with BS5, and patients with BS4 (using the CI features) are presented in Fig. 12 (b), (c), and (d), respectively.

The KNN and SVM methods identified individuals without stroke with high accuracy, as indicated in Fig 12 (b). The average classification accuracy under the KNN method was slightly higher, and features derived from sensor 2 data had the highest average classification accuracy. The highest average accuracy under the SVM method was achieved using features derived from sensor 14 data. The average accuracy under the RF method was 5% lower than that under the SVM method.

Patients with BS5 were identified under the SVM method at an average accuracy of >0.85 [Fig. 12 (c)]. The features derived from data from sensors 9 and 14 provided sufficient information for classification. When the features derived from sensors 2, 5, 8, 11, and 12 were used, the KNN method had an average accuracy slightly exceeding 0.8. The average accuracy under the RF method slightly exceeded 0.7. The average accuracy under the DT method ranged from 0.4 to 0.6.

The classification of patients with BS4 differed from that of the individuals without stroke, as indicated in Fig. 12 (d). When most of the sensor data were used, the highest average accuracy was achieved under the RF method; the average accuracy achieved was 1 when data from sensors 1 and 10 were used. The second highest average accuracy was obtained under the SVM method; the average accuracy was $>$ 0.9 when features derived from sensors 1, 5, 11, and 13–15 were used. When features derived from sensor 1 data were used, the NB method had an average accuracy of 0.9. Under the DT method, the average accuracy was <0.5.

IV. Discussion

The approach of smoothing bases through regression (smoothing) splines and Fourier bases (series) was used to extract features from accelerometer and gyroscope data. The hand function of patients with stroke was evaluated using the TT and GT. According to the results, to differentiate between individuals without stroke, patients with BS4, and patients with BS5, CI coefficient estimates can be applied to either the TT or GT. For the TT, the coefficient estimates related to smoothing splines derived from sensors 1 and 2 were fairly accurate; for the GT, the corresponding estimates derived from sensor 14 data had the highest accuracy, followed by those derived from data from sensors 1, 5, and 9–11.

Model performance was estimated using a fivefold cross-validation scheme and customized tuning grids. Furthermore, all processes were repeated 30 times to ensure the reliability of the classification result. For the TT, the KNN and SVM methods provided good classification accuracy when the essential features from CI were input. The KNN method had a slighter higher accuracy in identifying individuals without stroke, whereas the SVM method had higher accuracy in identifying patients With BS4. For the GT, the SVM method outperformed the other methods in terms of classification when essential CI features were used. In addition, although only eight recordings of patients with BS4 were included in this study, classification accuracy remained satisfactory under the SVM and RF methods. However, the DT method was not appropriate for use with imbalanced data.

Although participants were asked to perform the task repeatedly, some time gaps remained between repetitions. The cyclical characteristics of sample data remained somewhat unclear. Consequently, the lowest classification accuracy was achieved when CIII features were used. Thus, for similar experiments, the extraction of features for each cycle is advised. Furthermore, the CI and CII features were derived by computing the average coefficient estimates for ten cycles. Since the data trajectories of patients with BS4 or BS5 was not as regular as that of individuals without stroke, the variability levels in coefficient estimates for ten cycles can be treated as potential features for classification in addition to averages for summarizing features.

The number of degrees of freedom selected for functional form was restricted because the number of variables already considerably exceeded the number of participants. If a larger sample size is employed, the model accuracy can be improved using a more complex functional form. The functional form of our accelerometer and gyroscope data was nonperiodic. Because a Fourier series was designed to model the cyclical data, a wavelet series may also be applicable for these data.

Table III details the methods used in [16], [18], [25], [26] and the current method for comparison. In [16], the researchers directly extracted features from the movement characteristics of each participant. By contrast, [18] computed summary statistics in terms of centrality and variability in each cycle to extract raw data from each task. To reduce the number of features, an independent-samples t-test and PCA were performed, and the features derived from the PCA were

TABLE III
DETAILS OF PREVIOUS METHODS AND THE PROPOSED METHOD FOR COMPARISON

	[16]	[18]	[25]	[26]	Proposed method
Tasks	TT, GT and CTT	TT, GT and CTT	Random movement motor task	Finger-to-nose task	TT and GT
Feature extraction	Processing of kinematic parameters	Kinematic parameters	Kinematic parameters	Kinematic parameters	Functional form
Feature selection	None	PCA	Statistical extraction	Statistical extraction	One-way ANOVA and Wilcoxon test
Important features	Indirect	Indirect	Indirect	Indirect	Direct
Important sensor position	No	Indirect	No	No	Direct
Important task	No	Yes	No	No	Yes
Mapped clinical scale	Brunnstrom stage	Brunnstrom stage	Fugl-Meyer assessment scores	FMA-UE, ARAT and MBI	Brunnstrom stage
Results	Accuracy = 70.22%	No	$r^2 = 0.70$	55%, 51% and 32% of variance in FMA-UE, ARAT and MBI, respectively	Accuracy $> 90\%$

CTT: card-turning task; FMA-UE: Fugl-Meyer assessment for upper extremities; ARAT: action research arm test; MBI: modified Barthel index

TABLE IV
ACCURACY UNDER VARIOUS CLASSIFICATION METHODS ON TT AND THE GT

		Sensors															
Task	Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
TT	H	2,1	3,0	0,1		1,1			1,0				1,0	1,0	1,0	1,0	
	BS5		1,1													0,1	
	BS4	2,1	2,0	1,0	1,1	4,1	1,1	1,0	2,0	1,0	1,0	1,0		1,0		1,0	
	Total	4,2	6,1	1,1	1,1	5,2	1,1	1,0	3,0	1,0	1,0	1,0	1,0	2,0	1,0	2,1	
GT	H	2,1		0,1	0,2	2,1	1,1	2,1	2,2	3,1	3,1	3,0	1,1	2,1	2,1	0,2	1,0
	BS5	0,1	0,1	1,0		0,2	1,0	1,0	1,2	1,0	0,1	1,1	0,2		1,0		
	BS4	3,0	0,2	1,0	1,2	3,1	1,0	1,0	1,0	1,0	2,0	2,0	1,0	2,0	1,1	2,0	
	Total	5,2	0,3	2,1	1,4	5,4	3,1	4,1	4,4	5,1	5,2	6,1	2,3	4,1	4,2	2,2	1,0

Note : x and y denote the number of classification methods under which the average accuracy was higher than 0.85 and between 0.8 and 0.85, respectively.

linear combinations of summary statistics of centrality and variability. Notably, the summary statistics, especially those for variability, were highly sensitive to outliers. [25] and [26] have extracted motion periodic motion parameters as features and then used statistical methods, such as the Kruskal–Wallis test, for feature selection. Under these methods, features with repeated effects may be selected, which in turn affects the prediction effect of the classifier. However, in the proposed method, a functional form was fitted to the trajectory of each cycle in each task. The features, which are coefficient estimates, represent the direct characteristics of functional forms. Furthermore, generalized crossvalidation was conducted to determine the number of parameters that in turn determined the smoothness of data trajectories. Unlike the summary statistics computed from all sample data, the proposed method used the coefficients of the smoothed curve and hence the outliers had a negligible influence on the features.

As discussed in [18] and as determined through variable loading in the PCA, the most critical sensor position was indirect, and only features obtained from the TT could distinguish between patients with BS5 and BS4. However, the proposed method of constructing features was effective at classifying patients' BS and identifying the important position of sensors on the data glove. Table IV presents a summary, for each sensor in each task, of the number of classification methods under which the average group accuracy was higher than 0.85 and between 0.8 and 0.85 when CI features were used. For the TT, the information provided by sensor 2 was the most useful for classification, followed by the information provided by sensors 5, 1, 8, and 15. When the thumb was pressing a button, sensors 1 and 2 captured larger movements, sensor 15 captured muscle movements, and sensor 5 generated a signal upon impact or contact with the thumb. Determined according to physical principles, the order of inference effects is as follows: sensor 1 $>$ sensor 2 $>$ sensor 15 $>$ sensor 5. Sensor 1 was positioned at the fingertips and was therefore easily affected by the sliding of a glove and by the reaction force generated from pressing a button. Moreover, the amount of noise was relatively large, meaning that the related effect may have been reduced. In the GT, differentiation between groups was possible under most of classification methods when features derived from sensors 1, 5, 8, 10, and 11 were used. Sensor 16 was designed as a reference for computing another measurement and was demonstrated to be unimportant. Per physical principles, inference effects were expected to be strongest under sensors 1, 3, 6, 9, and 12, which were positioned on the fingertips. However, a similar problem as that encountered with the TT occurred: the fingertip sensors were susceptible to sliding of gloves and the reaction force generated by gripping. Furthermore, the amount of noise was relatively large, meaning that the effect may have been reduced. Sensor 1 captured a larger range of motion during the GT, meaning that it was the most effective sensor overall, followed by sensors 5, 8, 11, and 14, which were positioned at the base of the fingers. The position in sensor 1 in the rankings was pushed back once more because the movement of the little finger was too small.

Owing to the design of the data grove, the current study can only be used to evaluate patients with minor motor impairment, such as patients with BS4 to BS6. For patients with severe motor impairment, the priority of rehabilitation should be focused on movements of the entire upper limbs. Other device designs might be required. Furthermore, because of the inconsistent number of recordings in the TT and GT, the data for the TT and GT could not be merged. In future

studies, the number of recordings for each task should be set to be consistent such that multitask integration can be performed.

V. Conclusion

In accordance with the results of this study, functional analysis can be used to model the trajectory of data colllected by IMU sensors, and the corresponding coefficients can be treated as features for evaluating hand function. The GT task provided more sophisticated features to distinguish the extent of the hand function determined by BS. When the GT was used, the features had average accuracies of over 0.9 under the SVM method. When only data from sensors 1, 5, 8, 10, and 11 were used, the average accuracy was also over 0.91. Thus, in future data glove systems, only six sensors can be used, and a functional analysis model can be used to extract the features for evaluating hand function and formulating rehabilitation plans.

References

[1] World Health Organization. *WHO Reveals Leading Causes of Death and Disability Worldwide 2000–2019*. Accessed: Dec. 9, 2020. [Online]. Available: https://www.who.int/news/item/09-12-2020-who-reveals-leading-causes-of-death-and-disability-worldwide-2000-2019

[2] Ministry of Health and Welfare. *Taiwan's Leading Causes of Death in 2020*. Accessed: Jun. 18, 2020. [Online]. Available: https://www.mohw.gov.tw/cp-5017-61533-1.html

[3] Y. John and F. Anne, "Review of stroke rehabilitation," *BMJ*, vol. 334, no. 7584, pp. 86–90, 2007.

[4] Z. Yue, X. Zhang, and J. Wang, "Hand rehabilitation robotics on poststroke motor recovery," *Behav. Neurol.*, vol. 2017, Nov. 2017, Art. no. 3908135.

[5] R. Stockley, R. Peel, K. Jarvis, and L. Connell, "Current therapy for the upper limb after stroke: A cross-sectional survey of U.K. Therapists," *BMJ Open*, vol. 9, no. 9, 2018, Art. no. e030262.

[6] T. Platz, C. Pinkowski, F. van Wijck, I.-H. Kim, P. di Bella, and G. Johnson, "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, Jun. 2005.

[7] L. Yu, J. P. Wang, Q. Fang, and Y. Wang, "Brunnstrom stage automatic evaluation for stroke patients using extreme learning machine," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Hsinchu, Taiwan, Nov. 2012, pp. 380–383.

[8] J. Stamatakis *et al.*, "Finger tapping clinimetric score prediction in Parkinson's disease using low-cost accelerometers," *Comput. Intell. Neurosci.*, vol. 2013, Art. no. 717853.

[9] M. Djuriċ-Joviĉiċ *et al.*, "Finger tapping analysis in patients with Parkinson's disease and atypical parkinsonism," *J. Clin. Neurosci.*, vol. 30, pp. 49–55, Aug. 2016.

[10] Y. Sano *et al.*, "Quantifying Parkinson's disease finger-tapping severity by extracting and synthesizing finger motion properties," *Med. Biol. Eng. Comput.*, vol. 54, no. 6, pp. 953–965, Jun. 2016.

[11] H. Dai, P. Zhang, and T. C. Lueth, "Quantitative assessment of parkinsonian tremor based on an inertial measurement unit," *Sensors*, vol. 15, no. 10, pp. 25055–25071, 2015.

[12] W. W. Lee *et al.*, "A smartphone-centric system for the range of motion assessment in stroke patients," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 6, pp. 1839–1847, Nov. 2014.

[13] V. Venkataraman *et al.*, "Component-level tuning of kinematic features from composite therapist impressions of movement quality," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 1, pp. 143–152, Jan. 2016.

[14] P. Xia, T. Ding, Y. Peng, Q. Yang, and J. Li, "A multi-information data glove for hand function evaluation of stroke patients," *Investigación Clínica*, vol. 61, no. 1, pp. 328–339, Mar. 2020.

[15] H. G. Kortier, V. I. Sluiter, D. Roetenberg, and P. H. Veltink, "Assessment of hand kinematics using inertial and magnetic sensors," *J. Neuroeng. Rehabil.*, vol. 11, no. 1, pp. 1–15, Dec. 2014.

[16] B.-S. Lin, P.-C. Hsiao, S.-Y. Yang, C.-S. Su, and I.-J. Lee, "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, Nov. 2017.

[17] B.-S. Lin, I.-J. Lee, P.-Y. Chiang, S.-Y. Huang, and C.-W. Peng, "A modular data glove system for finger and hand motion capture based on inertial sensors," *J. Med. Biol. Eng.*, vol. 39, no. 4, pp. 532–540, Aug. 2019.

[18] B.-S. Lin, I.-J. Lee, P.-C. Hsiao, and Y.-T. Hwang, "An assessment system for post-stroke manual dexterity using principal component analysis and logistic regression," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 8, pp. 1626–1634, Aug. 2019.

[19] B. S. Lin, Y. S. Lin, I. J. Lee, and B. S. Lin, "A real-time missing data recovery method using recurrent neural network for multiple transmissions," in *Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing*, J. S. Pan, A. Ito, P. W. Tsai, and L. C. Jain, Eds. Cham, Switzerland: Springer, 2019, pp. 99–107.

[20] A. Hua *et al.*, "Evaluation of machine learning models for classifying upper extremity exercises using inertial measurement unit-based kinematic data," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 9, pp. 2452–2460, Sep. 2020.

[21] J. O. Ramsay and B. W. Silverman, *Applied Functional Data Analysis: Methods and Case Studies*. New York, NY, USA: Springer, 2002.

[22] R. L. Eubank, *Nonparametric Regression Spline Smoothing*, 2nd ed. New York, NY, USA: Marcel Dekker, 1999.

[23] A. Tharwat, "Classification assessment methods," *Appl. Comput. Inf.*, vol. 17, no. 1, pp. 168–192, 2021.

[24] G. L. Prajapati and A. Patle, "On performing classification using SVM with radial basis and polynomial kernel functions," in *Proc. 3rd Int. Conf. Emerg. Trends Eng. Technol.*, Nov. 2010, pp. 512–515.

[25] 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, Mar. 2020.

[26] Z.-J. Chen, C. He, M.-H. Gu, J. Xu, and X.-L. Huang, "Kinematic evaluation via inertial measurement unit associated with upper extremity motor function in subacute stroke: A cross-sectional study," *J. Healthcare Eng.*, vol. 2021, pp. 1–7, Aug. 2021.