Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Smart Lower-Limb Spasticity Assessment via Embeddable Pressure Insoles and Rotational Clustering Granular Logistic Regression

Wenting Qin
Tongji University
Qiangqiang Chen
Fudan University
Linjie He
Sichuan University
Xiaoyu Chen
Fudan University
Taiyang Liu
Shanghai Sunshine Rehabilitation Center
Lingyu Liu
Tongji University
Xingchen Dong
Fudan University
Hongyu Chen
Fudan University
Lingjing Jin
lingjingjin@tongji.edu.cn

Tongji University

Research Article

Keywords: Automated assessment, lower limb spasticity, machine learning, plantar pressure, stroke rehabilitation, wearable devices

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11605613/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Smart Lower-Limb Spasticity Assessment via Embeddable Pressure Insoles and Rotational Clustering Granular Logistic Regression

Wenting Qin[1†], Qiangqiang Chen[2†], Linjie He[3], Xiaoyu Chen[2], Taiyang Liu[1], Lingyu Liu[1], Xingchen Dong[2], Hongyu Chen[4*], Lingjing Jin[1*]

[1]Department of Neurology and Neurological Rehabilitation, Shanghai Disabled Persons' Federation Key Laboratory of Intelligent Rehabilitation Assistive Devices and Technologies, Yangzhi Rehabilitation Hospital (Shanghai Sunshine Rehabilitation Center), School of Medicine, Tongji University, Shanghai, China.
[2]College of Biomedical Engineering, Fudan University, Shanghai, 200433, China.
[3]College of Computer Science, Sichuan University, Chengdu, Sichuan, 610065, China.
[4]Greater Bay Area Institute of Precision Medicine (Guangzhou), Fudan University, Shanghai, 200433, China.

*Corresponding author(s). E-mail(s): chenhongyu@ipm-gba.org.cn; lingjingjin@tongji.edu.cn;
Contributing authors: 2280271@tongji.edu.cn; 23110720133@m.fudan.edu.cn; linjiehe_jason@foxmail.com; xiaoyuchen23@m.fudan.edu.cn; supermanjakun@163.com; Happyneurologist@163.com; 24110720130@m.fudan.edu.cn;
†These authors contributed equally to this work.

Abstract

Lower-limb spasticity is a motor dysfunction after stroke. Accurate spasticity assessment is crucial for its treatment and rehabilitation. However, the complex variability in the activity and activation patterns of spastic muscles challenges

automated, precise evaluation of lower-limb spasticity. We develop flexible, non-intrusive plantar pressure (PP) sensing insoles (48 units) that integrate naturally into regular shoes without external hardware modules. PP data are collected from 49 patients with lower limb spasticity. The high-dimensional 297 PP features are extracted to enrich the characterization of spastic movement patterns. A novel rotational clustering granulation logistic regression (RCGLogR) assessment algorithm was proposed to improve diagnostic performance for patients with lower-limb plantar flexor spasticity (PFS) and patients with lower-limb combined plantar flexor-knee extensor spasticity (CPFKES). Using leave-one-subject-out cross-validation, RCGLogR achieved sample-level accuracies of 84.6% and 78.0% and individual patient-level accuracies of 93.9% and 87.8% for the PFS and CPFKES tasks, respectively. At the individual patient level, RCGLogR achieved the highest accuracy among all evaluated classifiers. It exceeded the best-performing classical classifiers by 8.2 and 10.2 percentage points and the best-performing granular classifiers by 6.2 percentage points for PFS and CPFKES, respectively. To our knowledge, this is the first study to explore the feasibility of automated precision assessment for lower limb spasticity. The developed automated assessment system, with its dual advantages of convenience and accuracy, promises to address current evaluation limitations that rely entirely on manual manipulation and rehabilitation physicians' clinical experience.

Keywords: Automated assessment, lower limb spasticity, machine learning, plantar pressure, stroke rehabilitation, wearable devices

1 Introduction

Lower limb muscle spasticity is a common motor dysfunction after stroke. It mainly manifests as muscle contracture, pain, increased muscle tone, and abnormal movement patterns, severely affecting stroke patients' ability to live independently [1]. Accurate identification and assessment of lower limb spasticity severity are essential for developing effective treatment and rehabilitation programs [2]. The Modified Ashworth Scale (MAS) is a standard clinical tool for assessing lower limb muscle spasticity. Based on the resistance encountered during passive stretching of the lower limb in supine or lateral decubitus positions, muscle tone is classified into six grades (MAS 0, 1, 1+, 2, 3, 4) to reflect spasticity severity [3]. However, the MAS scale assessment relies solely on manual manipulation and subjective judgment by clinicians [4]. This leads to a heavy reliance on clinical experience, a complex, time- and labor-intensive evaluation process, significant strain on medical resources, and low inter-rater reliability due to its inherent subjectivity. Therefore, an urgent clinical need remains for an automated assessment tool that can objectively and accurately evaluate lower limb spasticity.

Recent studies have focused on analyzing spastic muscle movement and activation characteristics [5–11]. Santis et al. used wearable inertial measurement units (IMUs) to quantify knee joint motion and time-related properties during the pendulum test in patients with spastic knee extensors [8]. They found a significant correlation between the first swing angle and the MAS scale. Yeh et al. employed phase-amplitude coupling

metrics to correlate surface electromyography (sEMG) of knee extensors with joint motion during knee swing tests in patients with hemiplegic stroke, revealing abnormal coupling patterns between movement and muscle activation during spasticity [9]. To assess knee flexion capacity during spasticity, Choi et al. developed a wearable IMUs system that quantifies the range of motion during passive stretching of muscles, providing a portable clinical assessment tool [10]. Gurluk et al. placed a single IMU on the dorsal aspect of the foot and deployed sEMG sensors on the plantarflexor and its antagonist [11]. Through active dorsiflexion and passive stretching tests, they precisely calibrated the timing of muscle activation and joint restriction, revealing the relationship between neural activation patterns and joint movement limitations during spasticity. However, these studies primarily aimed to estimate muscle activation characteristics or kinematic parameters in patients with spasticity. They have not yet achieved an effective clinical diagnostic assessment of spasticity. A systematic review of 93 relevant studies published between 2009 and 2024 revealed a lack of automated, objective, and accurate clinical assessments of lower limb spasticity severity [5–11].

Moreover, IMU and/or sEMG sensors, which mainly constituted the technical schemes adopted in current spasticity-related studies [5–11], have numerous limitations in clinical settings. The IMU sensors face challenges from placement inconsistencies and interference from soft tissue movements [12, 13], while sEMG signals suffer from instability caused by varying skin conditions, imprecise electrode placement, and environmental electromagnetic disturbances [14]. Both IMUs and sEMG technologies require considerable effort to ensure proper skin attachment, typically needing adhesive materials or securing straps, which complicates practical deployment in clinical settings. In contrast, plantar pressure (PP) sensors, which measure the forces generated during foot-ground contact, demonstrate inherent resistance to electromagnetic interference, movement-related artefacts, and placement variations [15–17]. It can be seamlessly integrated into standard shoe insoles, significantly simplifying implementation and enhancing their suitability for routine clinical use [16]. In addition, PP signals have been shown to demonstrate remarkable diagnostic value in the domain of neurologically related movement disorders [17]. Shalin et al. utilized PP data during walking to detect freezing of gait in Parkinson's disease patients, achieving 82.1% sensitivity and 89.5% specificity [18]. Chen et al. achieved 91.89% accuracy using an ensemble learning algorithm and PP data from 37 stroke patients to assess the severity of lower-limb motor dysfunction [16]. Despite these benefits, few studies have explored the diagnostic value of PP sensors for assessing lower-limb spasticity. Accordingly, we aim to explore the feasibility of assessing lower-limb spasticity using PP signals, with a particular focus on plantar flexor spasticity (PFS), which is closely related to distal motor control impairments [19], and combined plantar flexor–knee extensor spasticity (CPFKES), disproportionately affecting overall movement patterns and functional outcomes [20].

Prior studies have shown that PP signals, as biomechanical outputs at the distal end of the lower limb kinetic chain, struggle to directly reflect neuromuscular pathology [20]. This indirect biomechanical transmission may challenge the effective identification of spastic motor features. To address this, we introduce granular computing theory into PP-signal-based lower-limb spasticity assessments to improve evaluation performance.

As a cognition-inspired computing paradigm, granular computing constructs multi-level information granules to reveal latent structures [21]. It enables multi-granular feature representation and cross-granular fusion to better handle fuzziness and uncertainty, improving knowledge extraction and decision accuracy [22]. Fuzzy granular models have shown considerable success in biomedical applications [23–25]. Samuel et al. reported 91.10% accuracy for heart failure risk prediction using an ANN with multi-granularity feature fusion, outperforming a conventional ANN by 4.40% [23]. Diao et al. achieved a 3.5% improvement in gesture classification with a granular fuzzy decision fusion algorithm [24]. Chen et al. substantially reduced RMSE for antigenic drift prediction on influenza data using a fuzzy granular sparse learning model [25]. By introducing fuzzy granular computing theory, we aim to present a high-precision algorithm for lower-limb spasticity classification using PP signals.

To address the lack of effective clinical diagnostic assessment of spasticity in current research, this study developed a novel objective, automated spasticity assessment system to enable convenient and accurate clinical evaluation of the severity of lower-limb PFS and CPFKES. The system consists of a pair of flexible, unobtrusive smart PP insoles and the proposed rotational clustering granular logistic regression (RCGLogR) classification algorithm. The primary contributions of this work are as follows: (i) a pair of flexible, unobtrusive smart insoles with 48*2 PP units was developed, enabling "zero-burden" clinical wear by naturally embedding into daily shoes without any external modules;(ii) a novel RCGLogR algorithm, integrated with granular computing, was proposed and shown to surpass eight widely used classifiers for the classification of patients with spasticity; (iii) we developed a pioneering automated system for assessing lower-limb spasticity based on smart PP insoles and RCGLogR. It achieved classification accuracies of 93.9% for PFS assessment and 87.6% for CPFKES evaluation, confirming its clinical feasibility and superiority; (iv) the impact of muscle spasticity on gait was revealed, and further, the diagnostic utility of PP features in spasticity assessment was discovered for the first time. This study's successful implementation is expected to provide an accurate, convenient, and automated tool for assessing lower-limb spasticity, with significant implications for early screening, treatment monitoring, and prognostic evaluation.

2 Experiment Protocol

2.1 Subject Information

The study enrolled 49 patients with post-stroke lower-limb spasticity. The inclusion criteria were: (i) chronic stroke (at least 6 months after stroke onset) with hemiplegia; (ii) ages between 18 and 80 years, regardless of gender; (iii) sufficient hearing, comprehension, and cognitive abilities to adhere to the experimental protocol; (iv) presence of PFS (MAS 0-3) and/or knee extensor spasticity (MAS 0-3) in the affected lower limb; (v) the ability to walk independently for a minimum of 14 meters without assistance. The exclusion criteria were: (i) a history of Parkinson's disease or other neurological disorders; (ii) recent botulinum toxin injections within the past 4 months; (iii) peripheral musculoskeletal disorders affecting walking. Table 1 presents patient information. The Medical Ethics Committee at Shanghai Yangzhi Rehabilitation Hospital,

Table 1 Subject demographic characteristics

Characteristics	Plantar flexor	Knee extensor
Number	49	49
Age (years)*	56.3±6.5	56.3±6.5
Weight (kg)*	73.2±4.1	73.2±4.1

*Data are presented as mean ± standard deviation.

China, approved this study (No. YZ 2024-023). All procedures adhered to applicable guidelines, and informed consent was obtained from all participants.

2.2 Experimental Equipment

We collected PP data during walking using a wearable smart insole device. The system comprised two main components: a pair of flexible smart insoles and a smartphone with data collection software. The insoles recorded PP signals at 20 Hz and wirelessly sent them via Bluetooth® to the smartphone for display and storage. As shown in Fig. 1, each insole consisted of a bottom layer of EBA foam and a top silicone protective layer, with a 48-unit sensing matrix in between. All electronic components, including the Bluetooth module, processor, and battery, were fully embedded within the flexible PP insole. This self-contained design lets the device fit directly into regular footwear without external modules. Our previous work offers a more detailed description of the smart insole [26].

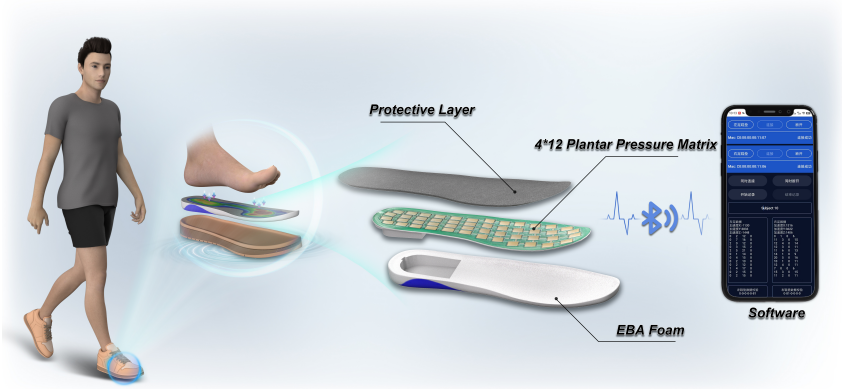


Fig. 1 Schematic diagram of the wearable smart insoles device and related components.

2.3 Data Acquisition

Before data collection, each participant was fitted with smart insoles under the supervision of medical staff, and a 2-minute acclimatisation period was allowed. Participants were then instructed to perform three walking trials along a 14-meter walkway at a

self-selected, normal speed, with rest periods of 1 to 2 minutes between trials. To analyse steady-state gait, we captured data exclusively from the central 10-meter segment of the walkway. Participants were advised to walk with their natural gait pattern and to pause the trial only if they experienced physical discomfort or unforeseen external events. The PP data collection process is shown in Fig. 2.



Fig. 2 Experimental setup for PP data collection: (left) experimental environment; (right) PP data collection during walking under medical supervision.

3 Method

3.1 Data Preprocessing

A zero-phase 6th-order Butterworth low-pass filter with a 5 Hz cutoff was applied to remove noise from the PP signals sampled at 20 Hz. According to previous research, the gait pattern of the unaffected side in patients with spasticity is closer to normal than that of the affected side [27]. Based on the variations in the total PP on the unaffected side and the clinical gait definition, the filtered PP data of each patient were divided into 30 continuous gait cycles. Each gait cycle was further subdivided into four sub-segments: first double support phase (DSP_1); unaffected foot (UF) single support phase sub-segment (SSP_U); second double support phase (DSP_2); and affected foot (AF) single support phase sub-segment (SSP_A). The PP data from each gait cycle corresponds to one sample. The dataset comprises 1470 samples from 49 individuals, with each contributing 30 gait cycles (30 samples).

3.2 Feature Extraction

A total of 45 PP feature extraction strategies were identified and summarized from related studies on the analysis of lower limb motor function based on PP signals [15, 28–30]. The specific definition of each strategy was provided in Table 2. These feature extraction strategies included 24 balance-related strategies for assessing gait balance, 12 stability-related strategies for evaluating single-foot stability, and nine time-related strategies for measuring gait temporal characteristics. Then three feature

Table 2 Feature extraction strategies based on PP data

Time-related Feature	Stability-related Feature		Balance-related Feature
Gait cycle (GC)	Center of pressure (COP) medial-lateral (ML) range	PP of midfoot region (MR)	Bilateral corresponding stability-related features ratio
Duration of DSP_1	COP ML SD	Ratio of MR to total PP	
Duration of SSP_U	COP path length	Mean pressure (MP) of forefoot region	Bilateral corresponding stability-related absolute differences
Duration of DSP_2	COP step length SD	MP of HR	
Duration of SSP_A	COP anterior-posterior (AP) range	Total pressure of MR	
Duration of DSP_1 to GC ratio	COP AP SD	Total sum pressure of all regions	
Duration of SSP_U to GC ratio	95% confidence circle area	MR ratio to total region sum	
Duration of DSP_2 to GC ratio	Covariance of COP projection in AP and ML	PP of FR	
Duration of SSP_A to GC ratio	Peak pressure of heel region	Ratio of FR PP to total PP	
	Ratio of heel region (HR) PP to total PP	MP of MR	
	Total pressure of fore region (FR)	Total pressure of HR	
	FR ratio to total region sum	HR ratio to total region sum	

extraction windows were created, as follows: (1) UF data in DSP_1 paired with AF data in DSP_2, (2) AF data in DSP_1 paired with UF data in DSP_2, and (3) UF data in SSP_U paired with AF data in SSP_A. Note that both the UF in SSP_A and the AF in SSP_U are in the swing phase. Since there isn't much reliable motion information in the pair of PP signals, they weren't taken into account while PP features were being extracted. For each feature extraction window, the stability-related and balance-related features were thus computed once. There were 297-dimensional features for a whole gait cycle.

3.3 RCGLogR Classification Algorithm

This work proposes a new granulation method, namely rotational clustering granulation, and subsequently designs the RCGLogR algorithm to improve the classification accuracy for patients with PFS and those with CPFKES. It is formulated by incorporating rotational clustering granular computing theory into the LogR framework. It begins by performing multi-granularity partitioning on the optimised PP features. Next, the model uses the defined granule calculation to input granule information for classification decision-making. Finally, the sample labels are predicted. The overall classification process of RCGLogR is illustrated in Fig. 3. The details of the RCGLogR algorithm are described as follows:

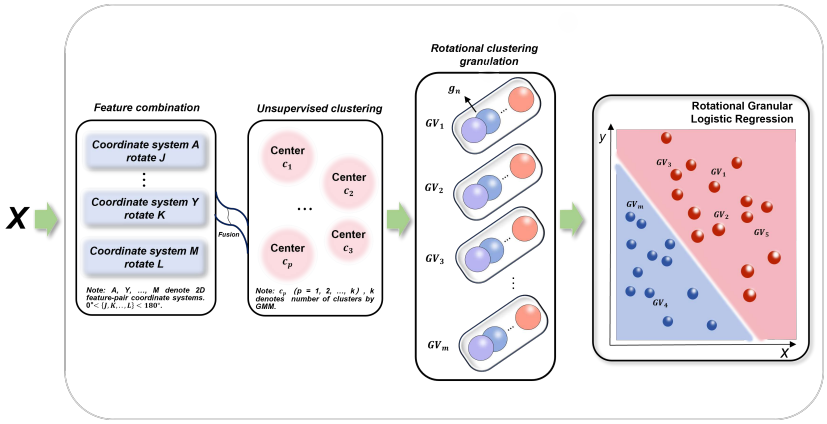


Fig. 3 Schematic diagram of the classification process of the proposed RCGLogR algorithm. X and Y denote the PP datasets and the predicted results of the samples, respectively. g_n denotes a granular component. GV_m denotes the granular vector formed by concatenating multiple g components.

Definition 1 Let the plantar-pressure spasticity assessment system be $U = (X, M, \{d\})$, where the sample set is $X = \{x_1, x_2, ..., x_n\}$, the feature set is $M = \{m_1, m_2, ..., m_k\}$, and $\{d\}$ denote spasticity labels. For a given sample $x \in X$ and a feature $m \in M$, let $t(x, m)$ denote the normalized value of sample x on feature m.

Definition 2 Let the k-dimensional feature set be $M = \{m_1, m_2, ..., m_k\}$. Select any pair of features $\{m_i, m_j\} \in C$ with $i < j$ to form a two-dimensional feature coordinate system $\langle m_i, m_j \rangle$. For a given sample $x \in X$, all pairwise feature combinations yield $h = k * (k-1)/2$ planar feature coordinate systems. For any point $t = (t_i, t_j)^T$ in such a plane, a counterclockwise rotation by angle θ produces a new rotated feature coordinate:

$$t' = R_\theta \cdot t, \tag{1}$$

$$where\, R_\theta = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}. \tag{2}$$

Definition 3 Given a sample set $X = \{x_1, x_2, ..., x_n\}$ and feature set $M = \{m_1, m_2, ..., m_k\}$, apply the Gaussian Mixture Model (GMM) to obtain p cluster centers $C = \{c_1, c_2, ..., c_p\}$. For any center $c \in C$ and any feature plane $\langle m_i, m_j \rangle$ from Definition 2, define the center's reference coordinate in that plane as

$$c_t = (c_{ti}, c_{tj})^T \tag{3}$$

Definition 4 For a sample $x \in X$, its rotated planar feature is $t' = (t'_i, t'_j)^T$. For a cluster center $c \in C$, its reference plane is $c_t = (c_{ti}, c_{tj})^T$. Define the distance function between the rotated feature plane $(t'_i, t'_j)^T$ and the cluster reference plane $c_t = (c_{ti}, c_{tj})^T$ as

$$D((t'_i, t'_j)^T, (c_{ti}, c_{tj})^T) \tag{4}$$

Definition 5 Given $t' = (t'_i, t'_j)^T$ and $c_t = (c_{ti}, c_{tj})^T$, the distance can be instantiated in three ways:

- Manhattan distance

$$D_m = |t'_i - c_{ti}| + |t'_j - c_{tj}| \tag{5}$$

- Euclidean distance

$$D_o = \sqrt{(t'_i - c_{ti})^2 + (t'_j - c_{tj})^2} \tag{6}$$

- Squared distance

$$D_s = (t'_i - c_{ti})^2 + (t'_j - c_{tj})^2 \tag{7}$$

Definition 6 Let $U = (X, M, \{d\})$ be the plantar-pressure dataset, with cluster centers $C = \{c_1, c_2, ..., c_p\}$. For any sample $x \in X$ and any feature plane $\langle m_i, m_j \rangle$ with $t = (t_i, t_j)^T$, define the rotational reference granulation of feature plane $\langle m_i, m_j \rangle$ with respect to C as

$$\begin{aligned} g(x_{i,j}) = \{r_1, \ldots, r_u\} = \bigcup_{v=0}^{p} \{ & D_m((t'_i, t'_j)^T, (c_{ti}, c_{tj})^T), \\ & D_o((t'_i, t'_j)^T, (c_{ti}, c_{tj})^T), \\ & D_s((t'_i, t'_j)^T, (c_{ti}, c_{tj})^T)\} \end{aligned} \tag{8}$$

where $\bigcup$ denotes the element-aggregation operation in granular computing. $g(x_{i,j})$ is termed a rotational reference granule.

Definition 7 For $U = (X, M, \{d\})$ and any sample $x \in X$, let $M = \{m_1, m_2, ..., m_k\}$. By Definition 2, x induces h feature planes. The granular vector of x with respect to C, denoted $GV(x)$, is defined as the rotational granular vector:

$$GV(x) = (g_{x_{v0}}, g_{x_{v1}}, \ldots, g_{x_{vh}}) \tag{9}$$

The granulation of data into granules and granular vectors necessitates the definition of relevant mathematical operations. A set of these operations for granules and granule vectors has been defined in previous work by He [31] and Chen et al [22].

Definition 8 Let $g_1(x)$ and $g_2(x)$ be two granules of sample x with respect to different reference cluster centers. The element-wise arithmetic on granules (addition, subtraction, multiplication, and division) is defined as follows:

$$g_1(x) + g_2(x) = \{g_1(x)_s + g_2(x)_s\}_{s=1}^{u} \tag{10}$$

$$g_1(x) - g_2(x) = \{g_1(x)_s - g_2(x)_s\}_{s=1}^{u} \tag{11}$$

$$g_1(x) * g_2(x) = \{g_1(x)_s * g_2(x)_s\}_{s=1}^{u} \tag{12}$$

$$g_1(x)/g_2(x) = \{g_1(x)_s / g_2(x)_s\}_{s=1}^{u} \tag{13}$$

The result of any of these arithmetic operations is also a granule.

Definition 9 Let the granular vectors be $GV(x_a) = (g_1(x_a), g_2(x_a), g_3(x_a), \ldots, g_h(x_a))^T$ and $GV(x_o) = (g_1(x_o), g_2(x_o), g_3(x_o), \ldots, g_h(x_o))^T$. Their dot product is defined by

$$\begin{aligned} GV(x_a) \cdot GV(x_o) = {} & g_1(x_a) * g_1(x_o) + g_2(x_a) * g_2(x_o) + \\ & g_3(x_a) * g_3(x_o) + \cdots + g_h(x_o) * g_h(x_o) \end{aligned} \tag{14}$$

By definition, the dot product of two granular vectors is a granule.

Definition 10 For the granule $g(x) = \{r_s\}_{s=1}^{u}$, its size is

$$\text{Size}(g(x)) = \sum_{s=1}^{u} r_s \tag{15}$$

Based on the aforementioned granule definitions and granular operations, we subsequently introduce the RCGLogR classification algorithm.

Definition 11 In the spasticity assessment system $U = (X, C, \{d\})$, per definition, granularize $x \in X$ to $GV(x) = (g_1(x), g_2(x), g_3(x), \ldots, g_h(x))$. Let the weight granular vector be $W = (w_1, w_2, w_3, \ldots, w_h, b)$. The rotational clustering–granule logistic regression is defined as

$$\text{GLogitreg}(x) = \frac{1}{1 + e^{-W \cdot GV(x)}} \tag{16}$$

where the term $W \cdot G(x)$ represents the dot product, calculated as:

$$W \cdot GV(x) = w_1 * g_1(x) + w_2 * g_2(x) + \cdots + w_h * g_h(x) + 1 * b.$$

Given that the set of granules is closed under the relevant arithmetic operations (Definition 8,9), the dot product $W \cdot G(x)$ necessarily yields a granule. The decision granule is then obtained by applying the sigmoid function element-wise.

Definition 12 In the spasticity assessment system $U = (X, C, D)$, there are the granule vector $GV(x) = (g_1(x), g_2(x), g_3(x), \ldots, g_h(x), 1)$, a weight-granule vector $W = (w_1, w_2, w_3, \ldots, w_h, b)$, and the logistic-regression label granule g_y. The granule loss function is:

$$L(g_{y_i}, \text{GLogitreg}(x_i)) = -\frac{1}{n} \sum_{i=1}^{n} [g_{y_i} \log(\text{GLogitreg}(x_i)) + (1 - g_{y_i}) \log(1 - \text{GLogitreg}(x_i))] \tag{17}$$

To ensure the rotational clustering–granule logistic regression model accurately captures patterns in the training data, we initiate iterative refinement. We use gradient descent to systematically adjust the model's parameters. In each iteration, the algorithm calculates the gradient of the loss function and updates the parameters to incrementally reduce the overall loss.

3.4 Model Training and Metrics

The classification performance of RCGLogR and eight other algorithms was evaluated on two datasets: one for PFS assessment and the other for CPFKES evaluation. These eight algorithms include six classical classifiers: k-nearest neighbours (kNN) [32], support vector machine (SVM) [33], random forest (RF) [34], AdaBoost [35], linear discriminant analysis (LDA) [36], and logistic regression (LogR) [37], as well

Algorithm 1 Rotational clustering–granule logistic regression learning algorithm

Input: Training set $U = (X, Y)$, where $x_i \in X \subseteq \mathbb{R}^m$ is an m-dimensional feature vector and $y_i \in Y$ is the label $i = 1, 2, ..., n$; learning rate η.
Output: Granular weight matrix W.

(1)Perform rotational clustering–based granulation of X with respect to C to obtain $GV(X)$.
(2)Construct the rotational clustering granular logistic regression model and randomly initialise the granular weights W.
(3)Feed granular vectors into the rotational granular logistic regression to obtain decision granules.
(4)Evaluate the loss L between the decision and label granules, and update the granule weight vector via backpropagation along the gradient direction.
(5)Repeat steps 3–4 until convergence or the maximum number of iterations is reached.
(6)Output the learned weight granules $W = (w_1, w_2, w_3, \ldots, w_h, b)$.

Table 3 The distribution of patients with different spasticity severity levels

Spasticity type	MAS-0 group	Mild group	Moderate group
PFS†	17	12	20
CPFKES†	17	19	13

†Values indicate the number of patients.

as two granular classifiers: granular sigmoid extreme learning machine (GSELM) [38] and granular eXtreme gradient boosting (GXGBoost) [39]. To reduce class imbalance and obtain three study-defined severity categories, we grouped participants separately for the PFS and CPFKES classification tasks. For PFS, we assigned participants with plantar-flexor MAS grade 0 to the MAS-0 group, indicating no increased resistance detected by the MAS in the assessed plantar flexors. MAS grades 1 and 1+ were combined into the mild (MIL) group, whereas MAS grades 2 and 3 were combined into the moderate (MOD) group. For CPFKES, the combined severity category was determined according to the concurrent involvement of the plantar flexors and knee extensors. Participants with MAS grade 0 in both muscle groups were assigned to the MAS-0 group. Participants with MAS grade 2 or 3 in both muscle groups were assigned to the MOD group. All remaining participants with an MAS grade of at least 1 in one or both muscle groups were assigned to the MIL group [40–42]. The participant distribution for each category is presented in Table 3. To reduce feature dimensionality without introducing information leakage, we incorporated feature selection, normalization, and hyperparameter optimization into a nested, subject-wise cross-validation framework [43, 44]. We performed the complete procedure separately for the PFS and CPFKES classification tasks and repeated it independently for each classifier. The outer loop used leave-one-subject-out cross-validation (LOSO-CV) to estimate classification performance. In each of the 49 outer iterations, all 30 gait-cycle samples from

one patient were reserved as the outer test set. The remaining 48 patients formed the outer training set. We did not involve the held-out patient in normalization, feature selection, or hyperparameter optimization. Within each outer training set, the 48 patients were divided into five inner folds. We stratified the folds at the patient level to preserve the distribution of the three spasticity categories as closely as possible. All gait-cycle samples from the same patient were assigned to the same inner fold. During each inner iteration, four folds were used for training, and the remaining fold was used for validation. We performed normalization separately in each inner iteration. The mean and standard deviation of each feature were calculated using only the patients in the four inner-training folds. These parameters were then used to normalize both the inner-training data and the corresponding inner-validation data. Sequential forward selection (SFS) started with an empty feature set. At each selection step, every feature not yet selected was temporarily added to the current feature set. Each resulting candidate feature set was evaluated using the same five inner folds. For each candidate feature set, all predefined hyperparameter combinations of the corresponding classifier were evaluated by grid search. For every combination, the classifier was trained on four inner folds and evaluated on the remaining fold. This procedure was repeated five times so that every inner fold served once as the validation set. The balanced accuracies from the five validation folds were then averaged. The candidate feature and hyperparameter configuration producing the highest mean inner-validation balanced accuracy were retained at that selection step. The selected feature was permanently added to the current feature set, and the remaining features were evaluated in the next step. Feature addition stopped when none of the remaining candidate features improved the mean inner-validation balanced accuracy. The SFS procedure produced a sequence of increasingly larger feature subsets. To avoid retaining unnecessary features, the smallest subset whose mean inner-validation performance was within one standard error of the highest performance was selected. The hyperparameter configuration associated with this subset was retained for the corresponding outer iteration. After the inner selection procedure was completed, a new Z-score normalizer was fitted using all 48 outer-training patients. The selected features and hyperparameters were then used to refit the classifier on the complete outer training set. Finally, we applied the fitted normalizer and classifier once to the held-out patient. The outer test data were not used to modify any part of the model. This nested procedure was repeated until every patient had served once as the outer test case. Feature selection was therefore conducted independently in every outer iteration. Consequently, the number and identity of the selected features were allowed to vary across the 49 outer folds. We did not use any feature subset obtained from the complete dataset for performance evaluation. For the subsequent feature analysis, selection stability was evaluated using the SFS results from the proposed RCGLogR classifier. We performed the analysis separately for the PFS and CPFKES tasks. For each task, the selection frequency of a feature was defined as the number of outer folds in which that feature was retained by SFS. Features were ranked from the highest to the lowest selection frequency. When two or more features had the same frequency, the feature with the earlier median entry position in the SFS sequence was ranked higher. The 10 highest-ranked features in each task were defined as the top 10 stability-ranked features for

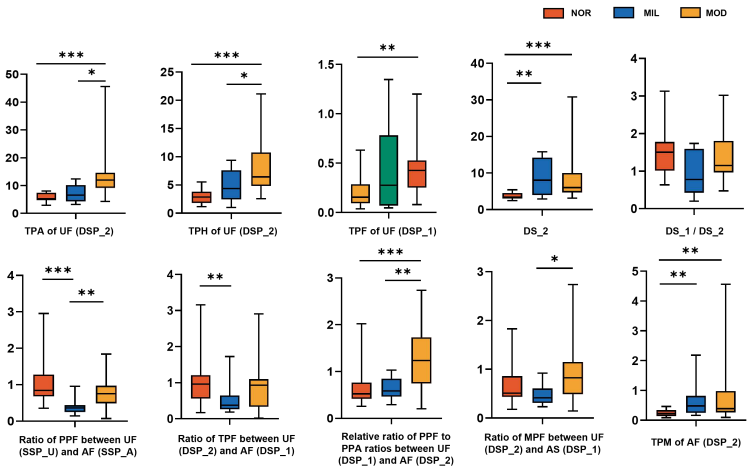


Fig. 4 Box-plot distributions of the 10 highest stability-ranked PP features across the PFS severity groups. TPA: total plantar pressure across the whole foot. TPH: total pressure in the heel region. DS_1: duration of the first double-support phase. DS_2: duration of the second double-support phase. PPF: peak pressure in the forefoot region. PPA: peak plantar pressure across the whole foot. TPF: total pressure in the forefoot region. MPF: mean pressure in the forefoot region. TPM: total pressure in the midfoot region. The label "NOR" in the embedded plot denotes the MAS-0 group. Significance levels: $*P<0.05$, $**P<0.01$, $***P<0.001$.

subsequent visualization and clinical interpretation. This ranking was performed only after all outer-fold predictions had been completed and did not affect model training or performance evaluation. Sample-level (SL) performance was determined from the predictions of the gait-cycle samples in the outer test sets. Each outer iteration produced predictions for the 30 gait cycles of the held-out patient. After completion of the 49 outer iterations, the SL classification accuracy was calculated from all outer-test gait-cycle predictions. Because every patient contributed 30 gait cycles, this result was equivalent to averaging the gait-cycle classification accuracy across the 49 patients. Individual patient-level (IPL) performance was derived from the same outer-test predictions. For each held-out patient, the predicted labels of the 30 gait cycles were aggregated by majority voting to generate one patient-level prediction. If two or more classes received the same number of votes, the class with the highest mean predicted probability across the 30 gait cycles was assigned. IPL classification accuracy was calculated by comparing the 49 patient-level predictions with the corresponding clinical labels.

4 Results

4.1 Feature Analysis

To explore the clinical characteristics of PP features, we analysed the 10 highest stability-ranked features from each assessment task. Box plots showing their distributions across the study-defined groups are presented in Fig. 4 for the PFS assessment and Fig. 5 for the CPFKES assessment. The results showed that TPH of UF (DSP_2),

TPA of UF (DSP_2), and TPM of AF (DSP_2) exhibited clear upward trends with increasing PFS severity, while the ratio of PPF between UF (SSP_U) and AF (SSP_A) showed a pattern of initial decrease followed by recovery. The MPF of AF (DSP_2) and the ratio of PPH to PPA in UF (DSP_2) increased significantly with CPFKES severity, and the PPH of UF (SSP_U) was notably elevated in the MOD group for CPFKES. Non-parametric analysis using Kruskal-Wallis test and post-hoc Dunn test revealed that ratio of MPF between UF (DSP_2) and AS (DSP_1) (H = 5.04, P = 0.08), TPM in UF (DSP_2) (H = 3.65, P = 0.16), MPH of AF (DSP_2) (H = 1.04, P = 0.59), and DS_1/DS_2 (H = 5.36, P = 0.06) showed no significant differences, while the remaining features demonstrated important inter-group differences ($P < 0.05$). In the CPFKES assessment, PPH of UF (SSP_U) showed substantial differences between the MAS-0 group and other groups ($P < 0.01$), demonstrating strong discriminative ability. The ratio of PPH to PPA in UF (DSP_2) differed significantly between the MAS-0 group and other groups ($P < 0.05$), and the ratio of TPF between UF (DSP_2) and AF (DSP_1) showed significant differences between the MIL and MOD groups ($P < 0.05$). For PFS evaluation, DS_2 and TPM of AF (DSP_2) showed significant differences between the MAS-0 group and other groups ($P < 0.01$). The ratio of PPF between UF (SSP_U) and AF (SSP_A) differed significantly between the MIL and MOD groups ($P < 0.01$). Additionally, TPA of UF (DSP_2) and TPH of UF (DSP_2) exhibited similar inter-group difference patterns.

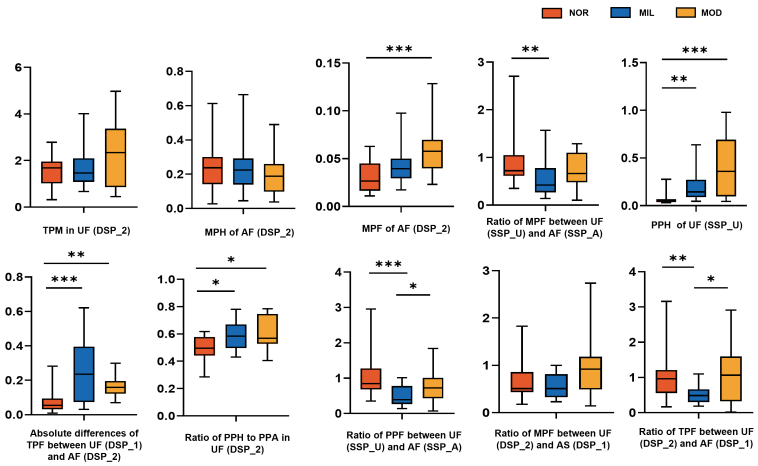


Fig. 5 Box-plot distributions of the 10 highest stability-ranked PP features across the CPFKES severity groups. PPF: peak pressure in the forefoot region. PPH: peak pressure in the heel region. PPA: peak plantar pressure across the whole foot. TPF: total pressure in the forefoot region. MPH: mean pressure in the heel region. MPF: mean pressure in the forefoot region. TPM: total pressure in the midfoot region. The label "NOR" in the embedded plot denotes the MAS-0 group. Significance levels: $*P<0.05$, $**P<0.01$, $***P<0.001$.

4.2 Classification Results of RCGLogR and LogR in CPFKES and PFS assessment

The SL classification results generated by RCGLogR and LogR algorithms in CPFKES and PFS evaluation are shown in Fig. 6. In the MAS-0 group, RCGLogR achieved higher SL classification accuracy than LogR for 10 of the 17 participants in the PFS assessment. Notably, subject four was misclassified by LogR but correctly classified by RCGLogR at IPL. Although subjects eight and six showed lower classification accuracy with RCGLogR compared to LogR, this did not affect their correct IPL predictions. In the MIL group, RCGLogR achieved higher SL classification accuracy than LogR for subjects 18, 21, 24, 25, 27, and 29. In the MOD group, the classification accuracy of LogR was below 60% for most patients. In contrast, 16 out of 20 patients in this group achieved classification accuracy exceeding 70% with RCGLogR. Importantly, subjects 32 and 44 were misclassified by LogR but correctly classified by RCGLogR at IPL. For the CPFKES assessment, RCGLogR achieved higher SL classification accu-

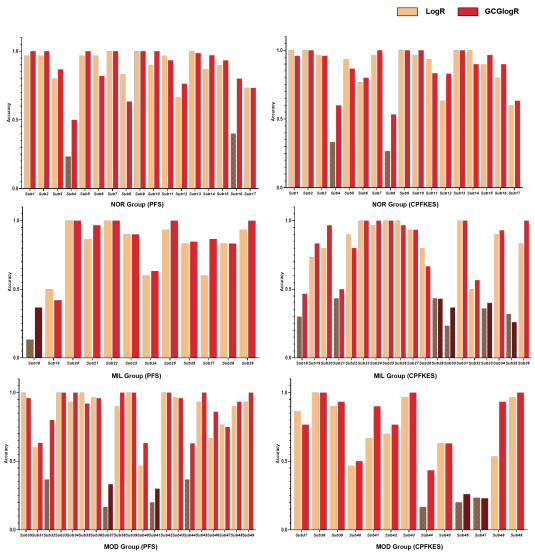


Fig. 6 The SL classification results of RCGLogR and LogR in CPFKES and PFS assessment for each of the 49 patients; individually misclassified patients are shown as densely cross-hatched bars. The label "NOR" in the embedded plot denotes the MAS-0 group.

racy than LogR for more than half of the participants in the MAS-0 group, with an average improvement of 11.5 percentage points. RCGLogR showed notably lower classification performance than LogR for subjects five, 11, and 14 in the MAS-0 group, but this did not affect their correct IPL classification. Furthermore, RCGLogR successfully corrected the classifications for subjects four and eight at IPL, both of whom were misclassified by LogR. Both classifiers achieved near 100% accuracy for subjects 23, 25, and 31 in the MIL group. Compared with LogR, RCGLogR achieved higher SL classification accuracy for most participants in the MIL group, including subjects

18, 20, and 30. In particular, RCGLogR corrected LogR's misclassifications at the IPL for Subjects 18 and 21 in the MIL group. In the MOD group, RCGLogR improved the classification accuracy for patients with lower accuracy by LogR (subjects 41, 44, and 48). Overall, compared to LogR, RCGLogR substantially improved SL classification accuracy in more than half of the patients in both the CPFKES and PFS assessments. The IPL classification results for 49 patients by RCGLogR and LogR in CPFKES and PFS assessments are shown in Fig. 7. Compared with LogR in the PFS assessment, RCGLogR improved average IPL classification accuracy by 11.8% in the MAS-0 group and 10% in the MOD group. For the CPFKES assessment, RCGLogR increased the IPL classification accuracy by 11.8, 10.5, and 7.7 percentage points in the MAS-0, MIL, and MOD groups, respectively. RCGLogR misclassified three patients in the PFS assessment and six patients in the CPFKES assessment, whereas LogR misclassified seven and 11 patients, respectively. Misclassifications primarily occurred between adjacent groups.

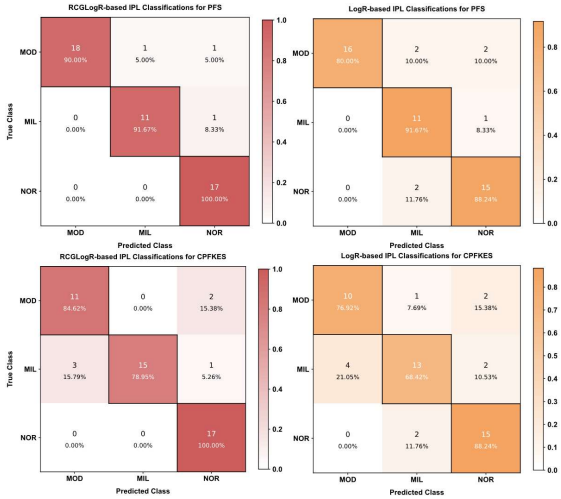


Fig. 7 The IPL classification results of RCGLogR and LogR for CPFKES and PFS assessment. Red heatmaps represent RCGLogR, and yellow heatmaps represent LogR. The label "NOR" in the embedded plot denotes the MAS-0 group.

4.3 Performance Comparison of RCGLogR and Other Classifiers

The SL and IPL classification performance of the proposed RCGLogR was compared not only with LogR but also with five classical classification algorithms (kNN, SVM, RF, AdaBoost and LDA) and two recently popular granular classifiers (GSELM and GXGBoost). In the PFS assessment, the classification results for each classifier are reported in Table 4. For the CPFKES evaluation, the corresponding results are presented in Table 5, with the highest accuracy in each group highlighted in bold.

Table 4 Performance Comparison of Different Classifiers for SL and IPL Classification in PFS Assessment.

Algorithm	SL (%)				IPL (%)			
	MAS-0 group	Mild group	Moderate group	Overall	MAS-0 group	Mild group	Moderate group	Overall
kNN [32]	81.6 ± 4.68	68.6 ± 8.43	69.7 ± 6.94	73.5 ± 3.96	94.1 ± 5.93	75.0 ± 13.05	75.0 ± 9.92	81.6 ± 5.67
SVM [33]	82.4 ± 5.30	71.4 ± 9.33	74.5 ± 6.34	76.5 ± 3.94	88.2 ± 8.06	75.0 ± 13.05	85.0 ± 8.24	83.7 ± 5.32
RF [34]	79.4 ± 5.91	45.0 ± 10.57	71.8 ± 8.33	67.9 ± 5.02	88.2 ± 8.06	58.3 ± 14.93	75.0 ± 9.92	75.5 ± 6.27
LDA [36]	84.7 ± 4.94	72.2 ± 9.08	58.5 ± 8.60	71.0 ± 4.71	94.1 ± 5.93	75.0 ± 13.05	65.0 ± 10.94	77.6 ± 6.03
AdaBoost [35]	79.8 ± 5.92	60.0 ± 8.79	71.7 ± 6.16	71.6 ± 3.94	82.4 ± 9.52	75.0 ± 13.05	70.0 ± 10.53	75.5 ± 6.27
GSELM [38]	80.2 ± 5.54	70.3 ± 9.07	80.0 ± 5.82	77.7 ± 3.73	82.4 ± 9.52	83.3 ± 11.24	85.0 ± 8.24	83.7 ± 5.32
GXGBoost [39]	86.4 ± 7.56	78.3 ± 7.92	76.3 ± 4.38	80.3 ± 5.41	94.1 ± 5.93	75.0 ± 13.05	**90.0 ± 6.93**	87.7 ± 4.72
LogR [37]	83.3 ± 5.36	76.1 ± 7.45	75.7 ± 6.58	78.4 ± 3.76	88.2 ± 8.06	**91.7 ± 8.36**	80.0 ± 9.28	85.7 ± 5.14
RCGLogR	**87.9 ± 3.78**	**81.9 ± 6.52**	**83.4 ± 4.93**	**84.6 ± 2.85**	**100.0 ± 0.00**	**91.7 ± 8.36**	**90.0 ± 6.93**	**93.9 ± 3.52**

*Data are presented as mean ± standard error.

The results show that RCGLogR achieved the highest SL classification accuracy in each group for PFS assessment. Furthermore, for IPL classification, RCGLogR performed best in the MAS-0 and MOD groups, while showing comparable performance to LogR in the MIL group. RCGLogR achieved the highest SL and IPL accuracies across all patients in the PFS assessment, at 84.6% and 93.9%, respectively. In the CPFKES evaluation, GXGBoost achieved the highest SL accuracy across all patients at 80.2%. RCGLogR showed the highest IPL accuracy in the MIL group, while performing similarly to GXGBoost in the MAS-0 and MOD groups. However, RCGLogR achieved the best IPL classification performance across all patients at 87.6% in the CPFKES assessment.

Table 5 Performance Comparison of Different Classifiers for SL and IPL Classification in CPFKES Assessment.

Algorithm	SL (%)				IPL (%)			
	MAS-0 group	Mild group	Moderate group	Overall	MAS-0 group	Mild group	Moderate group	Overall
kNN [32]	76.5 ± 4.83	57.9 ± 6.92	42.6 ± 9.74	60.3 ± 4.47	88.2 ± 8.14	68.4 ± 11.04	46.2 ± 14.46	69.4 ± 6.75
SVM [33]	78.8 ± 6.04	62.6 ± 8.19	46.4 ± 9.21	63.9 ± 4.86	82.4 ± 9.55	63.2 ± 11.45	46.2 ± 14.46	65.3 ± 6.95
RF [34]	71.8 ± 7.52	47.7 ± 7.46	30.0 ± 8.27	51.4 ± 5.09	70.6 ± 11.41	57.9 ± 11.67	30.8 ± 13.35	55.1 ± 7.28
LDA [36]	79.8 ± 6.23	67.4 ± 7.91	57.4 ± 9.25	69.0 ± 4.54	88.2 ± 8.14	73.7 ± 10.48	69.2 ± 13.39	77.6 ± 6.04
AdaBoost [35]	58.0 ± 7.64	56.8 ± 7.38	47.7 ± 7.76	54.8 ± 4.31	58.8 ± 12.36	52.6 ± 11.82	53.8 ± 14.43	55.1 ± 7.28
GSELM [38]	76.3 ± 4.65	53.7 ± 8.24	35.1 ± 9.28	56.6 ± 4.89	82.4 ± 9.55	63.2 ± 11.45	61.5 ± 14.06	69.4 ± 6.75
GXGBoost [39]	**90.0 ± 3.17**	**75.0 ± 7.43**	**75.0 ± 7.85**	**80.2 ± 2.36**	**100.0 ± 0.00**	63.2 ± 11.45	**84.6 ± 10.48**	81.6 ± 5.61
LogR [37]	82.7 ± 5.48	70.8 ± 8.02	63.8 ± 8.13	73.1 ± 4.39	88.2 ± 8.14	68.4 ± 11.04	76.9 ± 12.27	77.6 ± 6.04
RCGLogR	87.0 ± 4.15	74.1 ± 6.26	71.9 ± 7.68	78.0 ± 3.72	**100.0 ± 0.00**	**78.9 ± 8.07**	**84.6 ± 10.48**	**87.6 ± 4.05**

*Data are presented as mean ± standard error.

5 Discussion

Addressing the effective clinical diagnostic assessment of spasticity that current research has yet to achieve, this study developed flexible wearable smart PP insoles combined with the proposed RCGLogR algorithm to establish an automated lower-limb spasticity assessment system. The system offers dual advantages of precision and convenience for the PFS and CPFKES patient. Key findings include: (1) the smart insoles, nearly identical in appearance and usage to conventional insoles, seamlessly integrate into daily footwear for stable PP acquisition without external devices, substantially reducing clinical complexity; (2) RCGLogR outperformed eight other classification algorithms, whether for CPFKES or PFS patient classification; (3)

the system achieved clinical evaluation accuracies of 93.9% and 87.6% for PFS and CPFKES patient respectively, confirming clinical feasibility; (4) the study revealed how MAS-assessed spasticity severity impacts gait, with PP distribution during walking showing regular variations across different spasticity levels.

5.1 Clinical Reliability and Usability of Smart Insoles

The self-developed sensing matrix performance has been validated in prior studies [26]. The matrix exhibits rapid response characteristics and a minimal relaxation duration, enabling the robust acquisition of routine gait signals. Material stability assessments conducted over 12-hour periods have established adequate durability for extended monitoring applications. During walking trials, the smart insole achieved zero step frequency deviation when benchmarked against the commercial QH-JBE-Y gait analysis platform [45]. Deviations for support and swing phase measurements remained below $\pm 3\%$ relative to the commercial reference system, with temporal parameters including step duration, support phase duration, and swing phase duration exhibiting discrepancies under ± 0.02s. These findings substantiate the reliability of plantar pressure signals obtained via the developed smart insole during ambulation. The system achieves "zero burden" wearability via two principal design elements: (1) seamless integration within conventional footwear, eliminating requirements for anatomical alignment or external components, and (2) demonstrated user autonomy with complete success in independent device application across all participants (100% self-donning rate).

5.2 Analysis of the Clinical Significance of PP Features

According to the PP feature analysis results, PP distribution exhibited significant differences between the PFS and CPFKES patient groups, as well as across severity levels within each of the PFS and CPFKES groups. It provided clinical evidence for effective algorithmic classification and may explain the favorable assessment outcomes achieved in this study. A key finding showed that with increasing spasticity severity, significant increases were occurred in total forefoot and hindfoot plantar pressure during DSP_2 of UF in PFS patients, as well as in peak forefoot and hindfoot pressures during SSP_U of UF in CPFKES patients, compared to the corresponding affected sides, as indicated by changes in specific features (PFS groups: TPH of UF (DSP_2), TPM of AF (DSP_2), TPA of UF (DSP_2); CPFKES groups: ratio of PPH to PPA in UF, MPH of AF (DSP_2), PPH of UF in SSP_U). This clinical alteration in PP distribution pattern corresponds closely to the clinical definition of motor compensation. It represents the first investigation to demonstrate biomechanical motor compensatory patterns in spastic patients.

The significant inter-group differences in PP features across spasticity levels suggest their potential for diagnosing lower limb spasticity. To explore the diagnostic value, four representative features (DS_2, ratio of PPF between UF (SSP_U) and AF (SSP_A), PPH of UF (SSP_U), and ratio of TPF between UF (DSP_2) and AF (DSP_1)) were selected from the 20 most discriminative features identified through SFS, due to their salient distributions and variations across different groups. For PFS assessment, DS_2 demonstrated excellent classification performance in distinguishing

the MAS-0 group from other groups, achieving an AUC of 0.86 and an accuracy of 0.77. After excluding individuals from the MAS-0 group, the ratio of PPF between UF (SSP_U) and AF (SSP_A) showed superior diagnostic ability in identifying MOD group patients from MIL group patients, with an AUC of 0.81 and an accuracy of 0.81. Similarly, for CPFKES evaluation, PPH of UF (SSP_U) demonstrated excellent classification capability in distinguishing the MAS-0 group from other groups, achieving an AUC of 0.83 and an accuracy of 0.80. After excluding patients from the MAS-0 group, the ratio of TPF between UF (DSP_2) and AF (DSP_1) demonstrated strong performance in differentiating MOD group patients from MIL group patients, with an AUC of 0.72 and an accuracy of 0.81.

These observations led to the development of a multi-feature, combined diagnostic strategy. For PFS assessment, DS_2 first identifies participants in the MAS-0 group, followed by the ratio of PPF between UF (SSP_U) and AF (SSP_A) to distinguish between the MOD and MIL groups. For CPFKES evaluation, PPH of UF (SSP_U) initially identifies participants in the MAS-0 group, then the ratio of TPF between UF (DSP_2) and AF (DSP_1) differentiates between the MOD and MIL groups. This strategy demonstrated robust diagnostic performance, achieving an AUC of 0.74 and classification accuracy of 0.69 in PFS assessment, and an AUC of 0.76 and classification accuracy of 0.74 in CPFKES assessment. ROC curves were plotted to illustrate the diagnostic performance of the four individual features and the two combined diagnostic strategies, as shown in Fig. 8. These results emphasize a crucial point: PP features, which are readily obtainable through smart wearable insoles, demonstrate robust spasticity diagnostic capabilities. It is worth noting that this work initiates the revelation of the diagnostic value of PP features for clinical evaluation of lower limb spasticity, facilitating the creation of biomechanical indicators for assessing lower limb spasticity.

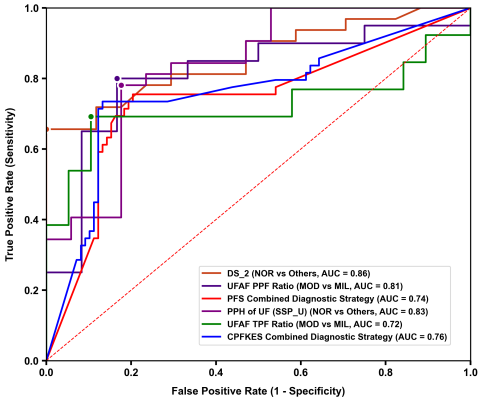


Fig. 8 Comparison of individual feature and multi-feature combined clinical diagnosis of PFS and CPFKES patients. The optimal classification performance under the maximum Youden index is marked as colored dots with white outlines on each ROC curve, with the dot color matching its respective curve. UFAF TPF Ratio: the ratio of TPF between UF (DSP_2) and AF (DSP_1). UFAF PPF Ratio: the ratio of PPF between UF (SSP_U) and AF (SSP_A).

5.3 Enhanced Classification Performance of RCGLogR

Overall, compared to LogR, RCGLogR significantly improved the SL classification accuracy in over half of the patients. It enhanced both mean SL and IPL classification accuracy in each group, with the average improvement across all groups exceeding 5.2%, whether in CPFKES or PFS assessment. The enhanced classification performance is primarily attributed to the incorporation of granular computing theory, which employs a hierarchical granulation mechanism to abstract PP data into multi-granularity information granules. This enables both global discriminative feature extraction and local detail preservation, thus more effectively mining spasticity-related features and improving spastic state discrimination. To our knowledge, this is the first work to apply granular computing theory creatively in PP signal analysis. With the help of granular computing theory and the rich spastic motion features contained in PP data, RCGLogR achieved the best performance compared to six popular traditional classifiers and two current granular classifiers for both PFS and CPFKES assessment. The only exception was in CPFKES classification for CPFKES assessment, where RCGLogR showed lower classification accuracy than GXGBoost. The lower classification accuracy can be attributed to differences in their underlying classification frameworks and approaches to granular vector construction, which result in distinct characteristics of data structure extraction. However, RCGLogR demonstrated superior clinical applicability with better IPL classification performance. Significant inter-individual variability exists in the gait manifestations of spastic patients, potentially arising from differences in disease onset time, severity, and ambulatory patterns. Robust algorithms are therefore crucial for ensuring reliable clinical evaluations in the presence of such heterogeneity. RCGLogR achieved notable reductions in standard error of 24.3% and 14.0% in PFS and CPFKES patient classification, respectively, compared to LogR. This indicates that RCGLogR possesses better capability in handling inter-patient data distribution variations, reflecting the inherent adaptive capacity of granular structures to data distribution shifts.

As shown in Fig. 7, RCGLogR completely separates the MAS-0 group from the MIL and MOD groups in CPFKES and PFS assessment, relative to LogR. Although RCGLogR yields a lower misclassification rate between adjacent groups than LogR, misclassifications remain. This is likely due to the high similarity of gait patterns among patients with neighboring spasticity grades. Notably, misclassification of MOD patients into the MAS-0 group occurs in both LogR and RCGLogR, albeit less frequently with RCGLogR. Post-hoc video review and PP feature analysis collectively revealed that pronounced compensatory strategies occurred in patients misclassified from MOD to the MAS-0 group, which may mask the underlying motor patterns and present as pseudo-normal gait patterns. The design of more sensitive test tasks should be considered in future work to better reveal patients' actual motor capacity. In conclusion, RCGLogR exhibits superior classification performance and robustness, whether in CPFKES or PFS assessment, demonstrating significant clinical application value.

5.4 Comparison with Existing Related Studies

A systematic review of 93 studies (published between 2009 and 2024) revealed that current research primarily focuses on the objective monitoring of abnormal lower-limb spastic movements and muscle activation patterns, with less attention paid to clinically objective automated assessment of muscle spasticity severity [5–11]. A few studies have proposed metrics to quantify spasticity severity and demonstrated correlation with clinical scales [7, 9]. It is important to note that correlations indicate potential clinical utility rather than clinical diagnostic effectiveness.

This study constructed a novel automated diagnostic system for lower-limb spasticity by deeply exploring the characteristics of dynamic PP collected via the smart insoles and proposing an advanced RCGLogR evaluation algorithm, addressing effective clinical diagnostic assessment of spasticity that current research has yet to achieve. The system demonstrated advantages in both convenience and accuracy of evaluation. The flexible, unobtrusive smart insole can be naturally embedded into regular footwear without requiring any external devices, enabling clinical "zero burden" deployment. Moreover, patients only need to perform several normal walking steps to complete the assessment of lower-limb PFS and CPFKES, without any additional or specific movements. The constructed system exhibited superior clinical diagnostic performance with classification accuracy of 93.9% and 87.6% for PFS and CPFKES patients, respectively, demonstrating a significant advancement.

5.5 Limitations and Future Research Directions

Despite the promising results, several limitations should be noted in this study. The validation cohort was relatively small and homogeneous in terms of patient characteristics, while the cross-sectional study design limited the dynamic assessment of spasticity recovery. Additionally, the current system may face challenges in adapting to special populations, including patients with varying foot sizes (both significantly larger or smaller than average), which could affect measurement accuracy and system reliability. Future work should focus on: (1) validating system performance in larger, multi-centre cohorts with heterogeneous patient populations, (2) improving system adaptability to different foot dimensions through advanced plantar region auto-detection algorithms, (3) conducting longitudinal studies to track rehabilitation progress, and (4) integrating additional clinical gait features to optimise the diagnostic model. These improvements will facilitate the translation of this technology into routine clinical practice, ultimately enhancing patient assessment and treatment outcomes for spasticity.

6 Conclusion

A pioneering automated assessment system for lower limb spasticity was constructed, effectively addressing the technical limitations of existing studies that had prevented the automated diagnosis of spasticity severity. Based on the developed flexible smart insoles and RCGLogR assessment algorithm, the system achieved dual advantages in clinical convenience and diagnostic accuracy. The flexible smart insoles can be seamlessly integrated into regular footwear without the need for external devices, achieving

a "zero-burden" clinical implementation. The system with the proposed RCGLogR algorithm achieved assessment accuracy of 93.9% and 87.6% for PFS and CPFKES patients, respectively, demonstrating excellent diagnostic performance. This work provides a convenient and objective automated assessment tool while establishing a novel technical pathway for biomechanics-based evaluation and monitoring of spasticity.

Acknowledgements. The authors would like to thank all patients who participated in this study. We also thank the medical staff at Shanghai Yangzhi Rehabilitation Hospital for their assistance in collecting and evaluating the clinical data.

Declarations

Ethics approval and consent to participate. The Ethics Committee of Shanghai Yangzhi Rehabilitation Hospital, Shanghai, China, approved all ethical and experimental procedures and protocols under Approval No. YZ 2024-023. All procedures adhered to applicable guidelines, and informed consent was obtained from all participants.

Consent for publication. Not applicable

Availability of data and materials. The de-identified datasets generated and/or analysed during the current study are not publicly available due to participant privacy and ethical restrictions but are available from the corresponding author on reasonable request, subject to approval by the relevant institution and ethics committee.

Competing interests. The authors declare that they have no competing interests.

Funding. This work was supported in part by the National Natural Science Foundation of China (Grant No. 82402979), in part by the National Key Clinical Specialty Discipline Construction Project of China (Z155080000004), and in part by the Shanghai Hospital Development Center Foundation (SHDC22026304).

Authors' contributions. WQ contributed to study conceptualization, clinical investigation, participant recruitment, data collection and curation, formal analysis, visualization, and preparation of the original manuscript. QC contributed to methodology development, algorithm design and implementation, software development, validation, formal analysis, visualization, and preparation of the original manuscript. LH contributed to methodology development, software implementation, and algorithm validation. XC contributed to smart-insole development, experimental implementation, data preprocessing, and validation. TL and LL contributed to participant recruitment, clinical assessment, data collection, and clinical interpretation. XD contributed to device implementation, data management, and experimental validation. HC contributed to study conceptualization, methodology, technical supervision, project administration, and manuscript review and editing. LJ conceived and supervised the clinical study, provided clinical resources, coordinated the project, acquired funding, and reviewed and edited the manuscript. WQ and QC contributed equally and share first authorship. All authors read and approved the final manuscript and agreed to be accountable for all aspects of the work.

Acknowledgements. The authors thank all participants and their families, and the clinical and laboratory staff of the Neurorehabilitation Centre and Human Movement Laboratory at Yangzhi Rehabilitation Hospital for their support with recruitment, assessment, and data collection.

References

[1] Arene N, Hidler J. Understanding motor impairment in the paretic lower limb after a stroke: a review of the literature. Topics in stroke rehabilitation. 2009;16(5):346–356.

[2] Hao M, Fang Q, Wu B, Liu L, Tang H, Tian F, et al. Rehabilitation effect of intelligent rehabilitation training system on hemiplegic limb spasms after stroke. Open Life Sciences. 2023;18(1):20220724.

[3] Aamodt G, Kjendahl A, Aamodt G, Kjendahl A, Jahnsen R. Dimensionality and scalability of the Motor Assessment Scale (MAS). Disability and Rehabilitation. 2006;28(16):1007–1013.

[4] Lee IJ, Hu YH, Hsiao PC, Yang SY, Lin HT, Chen YC, et al. AI-based automatic system for assessing upper-limb spasticity of patients with stroke through voluntary movement. IEEE Journal of Biomedical and Health Informatics. 2022;28(2):742–752.

[5] Guo X, Wallace R, Tan Y, Oetomo D, Klaic M, Crocher V. Technology-assisted assessment of spasticity: a systematic review. Journal of neuroengineering and rehabilitation. 2022;19(1):138.

[6] OuYang Z, Shen C, Wang Y. Motion analysis for the evaluation of dynamic spasticity during walking: A systematic scoping review. Multiple Sclerosis and Related Disorders. 2025;94.

[7] He J, Luo A, Yu J, Qian C, Liu D, Hou M, et al. Quantitative assessment of spasticity: a narrative review of novel approaches and technologies. Frontiers in neurology. 2023;14:1121323.

[8] De Santis D, Perez MA. A portable system to measure knee extensor spasticity after spinal cord injury. Journal of neuroengineering and rehabilitation. 2024;21(1):50.

[9] Yeh CH, Young HWV, Wang CY, Wang YH, Lee PL, Kang JH, et al. Quantifying spasticity with limited swinging cycles using pendulum test based on phase amplitude coupling. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2016;24(10):1081–1088.

[10] Choi S, Shin YB, Kim SY, Kim J. A novel sensor-based assessment of lower limb spasticity in children with cerebral palsy. Journal of neuroengineering and

rehabilitation. 2018;15(1):45.

[11] Gurluk A, Bilal HM, Hocaoglu E. Quantitative Analysis of Ankle Spasticity Through Electromyography Signal. In: 2024 Medical Technologies Congress (TIPTEKNO). IEEE; 2024. p. 1–4.

[12] Wang G, Jin L, Zhang J, Duan X, Yi J, Zhang M, et al. Recurrent neural network enabled continuous motion estimation of lower limb joints from incomplete semg signals. IEEE Transactions on neural systems and rehabilitation engineering. 2024;.

[13] Mobbs RJ, Perring J, Raj SM, Maharaj M, Yoong NKM, Sy LW, et al. Gait metrics analysis utilizing single-point inertial measurement units: A systematic review. Mhealth. 2022;8:9.

[14] Cheng L, Li J, Guo A, Zhang J. Recent advances in flexible noninvasive electrodes for surface electromyography acquisition. npj Flexible Electronics. 2023;7(1):39.

[15] Chen X, Hu D, Zhang R, Pan Z, Chen Y, Xie L, et al. Interpretable evaluation for the Brunnstrom recovery stage of the lower limb based on wearable sensors. Frontiers in neuroinformatics. 2022;16:1006494.

[16] Chen Q, Chen X, Chen C, Qin W, Liu T, Yin B, et al. Evaluation of Stroke Rehabilitation Using Smart Pressure Insoles and Ensemble Learning Algorithm. In: 2024 IEEE 20th International Conference on Body Sensor Networks (BSN). IEEE; 2024. p. 1–4.

[17] Chen JL, Dai YN, Grimaldi NS, Lin JJ, Hu BY, Wu YF, et al. Plantar pressure-based insole gait monitoring techniques for diseases monitoring and analysis: a review. Advanced Materials Technologies. 2022;7(1):2100566.

[18] Shalin G, Pardoel S, Lemaire ED, Nantel J, Kofman J. Prediction and detection of freezing of gait in Parkinson's disease from plantar pressure data using long short-term memory neural-networks. Journal of neuroengineering and rehabilitation. 2021;18(1):167.

[19] Zhou Z, Sun Y, Wang N, Gao F, Wei K, Wang Q. Robot-assisted rehabilitation of ankle plantar flexors spasticity: a 3-month study with proprioceptive neuromuscular facilitation. Frontiers in neurorobotics. 2016;10:16.

[20] Li S. Stiff knee gait disorders as neuromechanical consequences of spastic hemiplegia in chronic stroke. Toxins. 2023;15(3):204.

[21] He L, Chen Y, Wu K. Fuzzy granular deep convolutional network with residual structures. Knowl Based Syst. 2022;258:109941.

[22] Chen Q, He L, Diao Y, Zhang K, Zhao G, Chen Y. A novel neighborhood granular meanshift clustering algorithm. Mathematics. 2022;11(1):207.

[23] Samuel OW, Asogbon GM, Sangaiah AK, Fang P, Li G. An integrated decision support system based on ANN and Fuzzy_AHP for heart failure risk prediction. Expert Syst Appl. 2017;68:163–172.

[24] Diao Y, Chen Q, Liu Y, He L, Sun Y, Li X, et al. A fuzzy granular logistic regression algorithm for sEMG-based cross-individual prosthetic hand gesture classification. J Neural Eng. 2023;20(2):026029.

[25] Chen Y, Cai Z, Shi L, Li W. A fuzzy granular sparse learning model for identifying antigenic variants of influenza viruses. Applied Soft Computing. 2021;109:107573.

[26] Chen H, Wang Z, Meng L, Qin W, Wu J, Liu J, et al. Non disturbance gait signal acquisition insole for daily monitoring. In: 2023 IEEE 19th International Conference on Body Sensor Networks (BSN). IEEE; 2023. p. 1–4.

[27] Li J, Zhang N, Xu Y, Wang J, Kang X, Ji R, et al. Dynamical network-based evaluation for neuromuscular dysfunction in stroke-induced hemiplegia during standing. J Neuroeng Rehabil. 2024;21(1):190.

[28] Hébert-Losier K, Murray L. Reliability of centre of pressure, plantar pressure, and plantar-flexion isometric strength measures: A systematic review. Gait & posture. 2020;75:46–62.

[29] Pan Z, Gao H, Chen Y, Xie Z, Xie L. Evaluation of hemiplegic gait based on plantar pressure and inertial sensors. IEEE Sens J. 2023;23(11):12008–12017.

[30] Song Z, Ou J, Shu L, Hu G, Wu S, Xu X, et al. Fall risk assessment for the elderly based on weak foot features of wearable plantar pressure. IEEE Trans Neural Syst Rehabil Eng. 2022;30:1060–1070.

[31] He L, Chen Y, Zhong C, Wu K. Granular elastic network regression with stochastic gradient descent. Mathematics. 2022;10(15):2628.

[32] Peterson LE. K-nearest neighbor. Scholarpedia. 2009;4(2):1883.

[33] Suthaharan S. Support vector machine. In: Machine learning models and algorithms for big data classification: thinking with examples for effective learning. Springer; 2016. p. 207–235.

[34] Rigatti SJ. Random forest. Journal of insurance medicine. 2017;47(1):31–39.

[35] Schapire RE. Explaining adaboost. In: Empirical inference: festschrift in honor of vladimir N. Vapnik. Springer; 2013. p. 37–52.

[36] Izenman AJ. Linear discriminant analysis. In: Modern multivariate statistical techniques: regression, classification, and manifold learning. Springer; 2013. p. 237–280.

[37] LaValley MP. Logistic regression. Circulation. 2008;117(18):2395–2399.

[38] Jiang H, Chen Y, Jiang H, Ni Y, Su H. A granular sigmoid extreme learning machine and its application in a weather forecast. Applied Soft Computing. 2023;147:110799.

[39] Lan B, Chen Y, Wu K. A granular XGBoost classification algorithm. Applied Intelligence. 2025;55(13):1–22.

[40] Li S, Ghuman J, Gonzalez-Buonomo J, Huang X, Malik A, Yozbatiran N, et al. Does spasticity correlate with motor impairment in the upper and lower limbs in ambulatory chronic stroke survivors? American journal of physical medicine & rehabilitation. 2023;102(10):907.

[41] Norbye AD, Midgard R, Thrane G. Spasticity, gait, and balance in patients with multiple sclerosis: A cross-sectional study. Physiotherapy Research International. 2020;25(1):e1799.

[42] Krewer C, Luther M, Koenig E, Müller F. Tilt table therapies for patients with severe disorders of consciousness: a randomized, controlled trial. PLoS One. 2015;10(12):e0143180.

[43] Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics. 2006;7:91. https://doi.org/10.1186/1471-2105-7-91.

[44] Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. PLOS ONE. 2019;14(11):e0224365. https://doi.org/10.1371/journal.pone.0224365.

[45] Shanghai Hanfei Medical Equipment Co , Ltd . Ji Bu'en Gait Detection, Evaluation and Training System QH-JBE-Y [database on the Internet]. 2023 [cited Jul. 10, 2023]. Available from: https://www.hanfeiylqx.com/Products-37124528.html.