Infinite Science
Publishing

Cross-Dataset Evaluation of Sit-to-Stand Movement Classifiers for Post-Stroke Rehabilitation

L. Palumbi[1*], T. Kuhlgatz[1], M. Caruso[1,2], T. Seel[1] and L. Budde[1]

[1] Leibniz University Hannover, Institute of Mechatronic Systems, Hanover, Germany
[2] Center for Automation and Robotics, Spanish National Research Council (CSIC-UPM), Madrid, Spain
* Corresponding author, email: lucia.palumbi@imes.uni-hannover.de

Abstract: Automated assessment could support at-home post-stroke rehabilitation, yet ensuring cross-dataset generalizability is critical for real-world adoption. This study evaluates rule-based and Random Forest classifiers, trained on XSense IMU data, against the independent CeTI-Locomotion dataset. Zero-shot evaluation demonstrated the robustness of the rule-based approach (71.3% accuracy) compared to Random Forest (17.8%), which significantly improved to 58.2% with one-shot calibration. These findings indicate that generalizability is achievable through biomechanically grounded or adaptive strategies, marking a key step toward robust, clinically deployable rehabilitation systems.

I. Introduction

Stroke remains the third leading cause of disability worldwide [1]. To regain mobility and functional independence, post-stroke patients require high-dose, repetitive rehabilitation exercises. However, limited access to supervised physiotherapy and patient non-compliance between sessions significantly hinder recovery outcomes [2]. Wearable sensor-based systems for automated movement assessment offer a promising solution by enabling unsupervised monitoring during home-based rehabilitation.

Recent studies have demonstrated the feasibility of using inertial measurement units (IMUs) to assess rehabilitation exercises. E.g., Komaris et al. [3] showed that IMU-based metrics movement smoothness, intensity, and consistency – could effectively evaluate exercise performance at-home, while Ranganathan et al. [4] achieved 90% accuracy in detecting compensatory upper-limb movements during reaching and manipulation tasks. In our earlier work [5], we demonstrated the feasibility of distinguishing sit-to-stand (STS) movements from characteristic related post-stroke impairments, including asymmetrical weight-bearing, lateral trunk deviation, and compensatory stepping [6]. Using 6D IMUs, we evaluated both rule-based and learning-based classification models, which achieved average accuracies of 89.78% and 94.03%, respectively.

However, tthese approaches share a critical limitation: validation is performed exclusively on the same datasets used for model development, leaving cross-dataset generalizability untested. Algorithms trained and tested solely on small, controlled datasets often degrade significantly when applied to the heterogeneous movement patterns of real-life settings [7]. To eliminate the need for subject-specific calibration, reliable monitoring systems must be robust against these shifts in distribution.

In this work, we evaluate the cross-dataset transferability of our STS classifiers when applied out-of-the-box to the CeTI-Locomotion dataset [8], which contains IMU data from 50 healthy adults performing sit-to-stand. We assess both zero-shot and one-shot adaptation approaches, quantify the domain gap between training and target distributions and test the model's ability to distinguish natural kinematic variability from pathology, which is essential to ensure specificity and minimize false positives in clinical practice [2].

II. Material and methods

Datasets. We evaluated classifiers trained on the IMU dataset collected in [5] against the CeTI-Locomotion dataset [8]. The source domain data comprised 330 STS trials from 11 healthy participants mimicking post-stroke patterns under physiotherapist supervision: correct execution, asymmetrical movement simulating hemiparesis (plegie), uncontrolled descent or failure of completing the movement (fail), and compensatory stepping (step). The target dataset, CeTI-Locomotion, contains 499 STS repetitions from 50 healthy adults recorded with the IMU-based Rokoko Smartsuit Pro motion capture system, at 100 Hz. Since CeTI-Locomotion contains only correct executions, if our algorithms [5] generalize well, they should classify all trials in the CeTI_Locomotion dataset as correct, making misclassifications direct measures of false-positive compensation movements.

Preprocessing alignment. Both datasets provide joint angles (bilateral hip, knee and ankle and one trunk angle), enabling direct comparison despite different sensor modalities. CeTI data required several transformations:

sign correction for knee flexion, segmenting the trials from sit-to-stand-to-sit to STS, and resampling to 400 time steps. Finally, joint angles were normalized to movement cycle percentage. *Figure* 1 illustrates the kinematic comparison between the two datasets.

Classification approaches. We applied both classifiers from [5] with minimal modifications. The rule-based classifier uses biomechanical thresholds for asymmetry and range of motion; we adjusted the full-extension criterion from the original 20°≤ range, which was tailored to the to the dataset in [5], to the canonical value of 0°≤. The Random Forest classifier extracts kinematic features per joint and yields 7 predictions per trial, to then perform majority voting.

Evaluation protocol. Zero-shot evaluation applied trained models directly to CeTI-Locomotion data, expecting "correct" classification for all 499 trials. One-shot adaptation used a single reference sample per subject to calibrate prediction thresholds before classifying remaining trials.
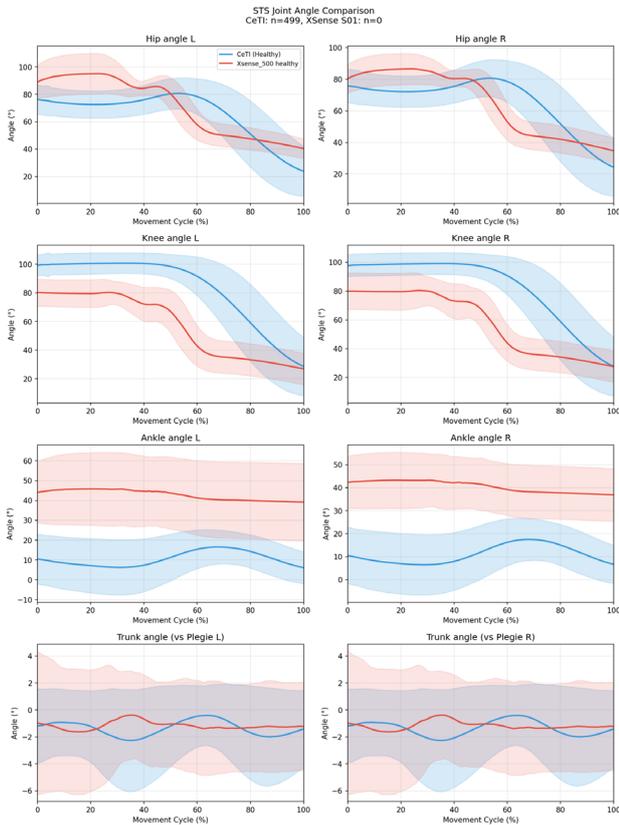

Figure 1: Mean ± SD joint angles for CeTi_locomotion (blue) and our dataset (red) [5] during a normalized Sit-to-Stand task.

III. Results and discussion

Table 1 summarizes the classification results on CeTI-Locomotion data. The rule-based approach achieves 71.3% correct classification in zero-shot evaluation, outperforming the Random Forest (17.8%). One-shot calibration improves the learning-based model to 58.2%.

The observed offset in ankle kinematics from *Figure 1* can be attributed to discrepancies in the definition of the neutral reference frame.

Table 1. Classification results on CeTI-Locomotion dataset (healthy subjects; expected class: Correct).

Method	Condition	Correct	Plegie	Fail	Step
Rule-based	Zero-shot	**71.3%**	24.0%	1.8%	2.8%
Random Forest	Zero-shot	17.8%	47.4%	34.4%	0.4%
Random Forest	One-shot	**58.2%**	17.7%	24.1%	0.0%

IV. Conclusions

This study demonstrates that cross-dataset generalizability for STS assessment is an achievable and critical objective for clinical adoption. The rule-based model's zero-shot performance (71.3%) validates the robustness of biomechanical principles, while the success of one-shot calibration highlights the potential of adaptive strategies. However, since validation was limited to healthy subjects, a critical next step is assessing transferability to pathological conditions.

Future work should prioritize transfer learning and domain-invariant representations to ensure automated rehabilitation systems can reliably scale to clinical populations.

REFERENCES
[1] V. L. Feigin *et al.*, 'World Stroke Organization: Global Stroke Fact Sheet 2025', *International Journal of Stroke*, vol. 20, no. 2, pp. 132–144, Feb. 2025, doi: 10.1177/17474930241308142.
[2] R. Campbell, M. Evans, M. Tucker, B. Quilty, P. Dieppe, and J. L. Donovan, 'Why don't patients do their exercises? Understanding non-compliance with physiotherapy in patients with osteoarthritis of the knee', *J Epidemiol Community Health*, vol. 55, no. 2, pp. 132–138, Feb. 2001, doi: 10.1136/jech.55.2.132.
[3] D.-S. Komaris, G. Tarfali, B. O'Flynn, and S. Tedesco, 'Unsupervised IMU-based evaluation of at-home exercise programmes: a feasibility study', *BMC Sports Science, Medicine and Rehabilitation*, vol. 14, no. 1, p. 28, Feb. 2022, doi: 10.1186/s13102-022-00417-1.
[4] R. Ranganathan, R. Wang, B. Dong, and S. Biswas, 'Identifying compensatory movement patterns in the upper extremity using a wearable sensor system', *Physiol. Meas.*, vol. 38, no. 12, p. 2222, Nov. 2017, doi: 10.1088/1361-6579/aa9835.
[5] T. Kuhlgatz, M. A. Caruso, S. Ferrante, and T. Seel, 'Assessing sit-to-stand motion execution using 6D IMUs only - A proof of concept', *Current Directions in Biomedical Engineering*, vol. 11, no. 1, pp. 429–432, Sept. 2025, doi: 10.1515/cdbme-2025-0209.
[6] A. Boukadida, F. Piotte, P. Dehail, and S. Nadeau, 'Determinants of sit-to-stand tasks in individuals with hemiparesis post stroke: A review', *Ann Phys Rehabil Med*, vol. 58, no. 3, pp. 167–172, June 2015, doi: 10.1016/j.rehab.2015.04.007.
[7] G. Nicora *et al.*, 'Systematic review of AI/ML applications in multi-domain robotic rehabilitation: trends, gaps, and future directions', *J NeuroEngineering Rehabil*, vol. 22, no. 1, p. 79, Apr. 2025, doi: 10.1186/s12984-025-01605-z.
[8] S. Hanisch, L. Pogrzeba, E. Muschter, S.-C. Li, and T. Strufe, 'A kinematic dataset of locomotion with gait and sit-to-stand movements of young adults', *Scientific Data*, vol. 11, no. 1, p. 1209, 2024, doi: 10.1038/s41597-024-04020-6.