2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) | 979-8-3503-7149-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/EMBC53108.2024.10781988

Integrating Wearable Sensor Technology and Machine Learning for Objective m-CTSIB Balance Score Estimation

Marjan Nassajpour[1], Mustafa Shuqair[1], Amie Rosenfeld, D.P.T.[2], Magdalena I. Tolea, Ph.D.[2], James E. Galvin, M.D, M.P.H.[2], and Behnaz Ghoraani, Ph.D.[1*]
[1]Department of Electrical Engineering & Computer Science, Florida Atlantic University, Boca Raton, FL, USA
[2]Comprehensive Center for Brain Health, Department of Neurology, University of Miami, Boca Raton, FL

***Abstract*—This study marks the first endeavor to utilize wearable technology combined with machine learning to objectively assess the Modified Clinical Test of Sensory Interaction on Balance (m-CTSIB). We focus on developing an affordable, easily accessible method for balance assessment, critical for adults at risk of falls and cognitive decline. Our novel approach uses a single inertial measurement unit sensor (APDM, INC.) to gather lumbar accelerometer and gyroscope data. This data is accompanied by ground truth scores obtained from m-CTSIB tests on a force plate (Falltrak II, MedTrak VNG, Inc.) from 34 participants aged 21 to 88. Using XGBOOST, we achieve a remarkable 0.94 correlation using accelerometer data and 0.90 with gyroscope data in the test dataset, demonstrating a strong correlation with actual scores in a subject-wise leave-one-out cross-validation. Offering objectivity, affordability, and potential for remote monitoring, our innovative approach holds promise for enhancing the diagnosis and management of balance disorders in adults, thereby improving their quality of life and independence.**

***Index Terms*—Balance Assessment, Wearable Sensors, Machine Learning, m-CTSIB (Modified Clinical Test of Sensory Interaction on Balance)**

I. Introduction

The term "balance" is frequently employed in the healthcare field across diverse clinical specialties. It is often used in conjunction with terms like stability and control of posture [1]. Around 33% to 50% of individuals aged 65 and older encounter challenges related to balance or walking [2]. Evaluating motor performance, including walking and stability, among older individuals can serve as a valuable clinical approach to anticipate various clinical consequences. These outcomes encompass the risk of falls, neurological conditions like Parkinson's disease, cognitive decline, and even mortality [3], [4].

Modified Clinical Test of Sensory Interaction on Balance (m-CTSIB) is among the commonly employed balance assessment methods, alongside others such as the Berg Balance Scale (BBS) and Timed Up and Go (TUG) test [5]–[7]. The conventional methods typically involve clinicians using stopwatches to measure the duration of independent standing under various sensory conditions [8]. However, this approach introduces variability and subjectivity in the obtained results. These limitations underscore the need for more objective and comprehensive assessment tools in balance evaluation.

This study was supported by the National Science Foundation (IIS 1942669) and Ed and Ethel Moore Alzheimer's Disease Research Program at the Florida Department of Health (AWD-001693).

On the other hand, force plate technology has emerged as a valuable means of objectively assessing balance and mobility. Utilizing distinct quantification systems, technologies like the NeuroCom Balance Master (NeuroCom™ International Inc.) [9] and Falltrak II (MedTrak VNG, Inc.) have been employed to measure m-CTSIB test scores. The NeuroCom Balance Master, for instance, evaluates sway velocity (measured in degrees per second) under each sensory condition, while the Falltrak II assesses deviations of the center of pressure (COP) from the center of mass (COM). However, the primary limitation of force plate technology lies in its dependence on expensive, specialized equipment, often rendering it inaccessible in various clinical settings. Additionally, it lacks the flexibility for remote assessments, limiting its application in broader contexts. These challenges highlight the imperative for developing accessible, objective, and affordable balance assessment methods, such as the one proposed in this study.

This study introduces a novel approach using wearable sensors and machine learning to estimate m-CTSIB scores. We utilize a single APDM wearable sensor to capture lumbar accelerometer and gyroscope data and apply the XGBOOST algorithm to estimate m-CTSIB scores. Notably, we utilize ground truth scores obtained concurrently with m-CTSIB assessments. This innovative approach holds the potential to provide an objective, accessible, and remote means of assessing balance, addressing the limitations associated with traditional methods and specialized equipment. This innovation can potentially enhance the diagnosis and management of balance disorders and consequential complications, improving their quality of life and independence.

II. Study design

A. Dataset

We conducted this study with a cohort of 34 participants, ranging in age from 21 to 88 years, consisting of 12 males and 23 females. The study adhered to ethical guidelines, obtained IRB approval, and followed the principles of the Helsinki Declaration. All participants provided informed consent through written consent forms. All subjects underwent the m-CTSIB clinical test, which assesses the balance performance of adults when one or more sensory systems (i.e., vision, somatosensory, and vestibular) are compromised. We employed the Falltrak II computerized balance testing system for this purpose. The m-CTSIB test was conducted with participants standing on

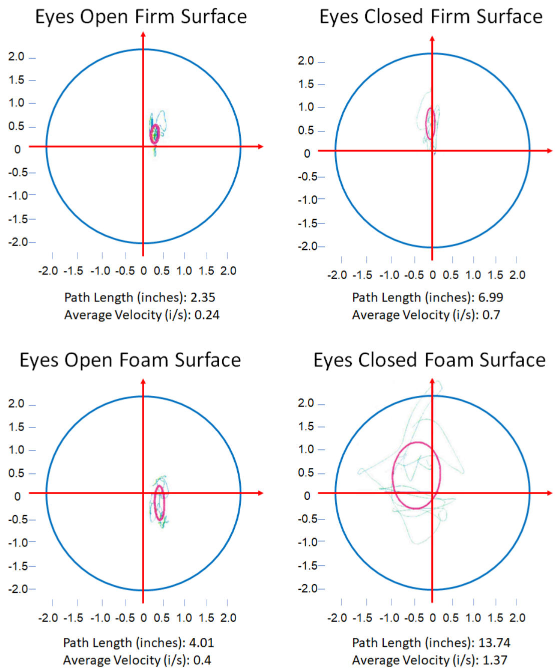

Fig. 1. The variations in path length (PL) and average velocity (AV) across four different conditions of one of our subjects. The horizontal and vertical axes represent the displacement of the subject along the x-axis and y-axis, respectively. The green lines represent the real-time tracing of the subject's center of pressure (COP) during the test, and the pink circles indicate the standard deviation of the subject's COP.

the Falltrak II plate, each trial lasting approximately 10 seconds. This test comprised four distinct conditions to assess balance under varying sensory inputs comprehensively. The first condition was Eyes Open, standing on a Stable Surface (EOSS), which served as a baseline measurement. The second condition, Eyes Closed, standing on a Stable Surface (ECSS), evaluated balance without visual cues. The third, Eyes Open, standing on a Foam Surface (EOFS), challenged the participants' balance on an unstable surface while retaining visual input. Finally, the most challenging condition was Eyes Closed, standing on a Foam Surface (ECFS), where participants had to maintain balance without visual cues and on an unstable surface.

The Falltrak II system measured the participants' balance using the path length (PL) and average velocity (AV). Stability measures how much the COP deviates from the COM during the test. PL is a measure of how much the COP moves during the test. A shorter path length indicates a better balance performance. AV is a measure of how fast the COP moves during the test. A lower average velocity indicates a better balance performance. Fig. 1 shows the Falltrak II report of a participant and how PL and AV vary through different conditions, with ECFS being the most challenging with the highest PL and AV scores.

Before initiating the m-CTSIB clinical test, we affixed an APDM wearable sensor to the lumbar region with the Z-axis oriented downward (see Fig. 2). This sensor recorded

TABLE I
SUMMARY OF MEASUREMENTS FROM DIFFERENT CONDITIONS

	PL (in)	**AV (in/s)**	**Sensor Data Duration (s)**
EOSS	3.3 ± 1.53	0.32 ± 0.15	12.00 ± 3.88
ECSS	6.34 ± 3.08	0.63 ± 0.31	11.62 ± 1.34
EOFS	7.05 ± 3.44	0.70 ± 0.34	11.51 ± 1.35
ECFS	19.84 ± 10.47	1.99 ± 1.05	11.74 ± 1.61

Measurements are in inches (in) for path length (PL) and average velocity (AV) and in seconds (s) for sensor data duration

accelerometer and gyroscope data in three dimensions (X, Y, and Z) at a sampling frequency of 128 Hz. Each participant generated four data files corresponding to the four test conditions. Initially, the sensor data was stored locally, and upon completing the data collection session, we utilized a Docking Station provided by APDM Wearable Technologies to transfer the data to a computer via USB cable.

We selected the lumbar region for sensor placement due to its proximity to the body's center of gravity, a critical factor in balance control. This location enables capturing the center of gravity movements to estimate the balance performance, as mentioned in previous studies [10], [11]. It also offers comfort and minimal interference with test conditions. Table I summarizes collected data, including average PL and AV scores and the duration of recorded wearable data.

B. Feature Extraction

As a preprocessing step, a 0.5 second interval was omitted from the beginning and end of the wearable sensor data before feature extraction. This adjustment was made to mitigate the potential influence of artifacts or disturbances when subjects transitioned between four conditions.

We independently extracted 42 features from the accelerometer and gyroscope data, considering each subject's performance in the X, Y, and Z dimensions. Consequently, there were 168 features available for each subject, separately for

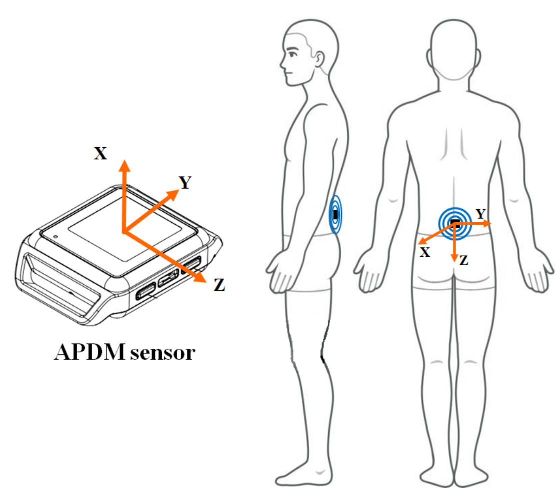

Fig. 2. The coordinate direction of APDM sensors and its location on the lumbar.

accelerometer and gyroscope data. The extracted features encompass a wide range of characteristics, including statistical measures (such as standard deviation, skewness, kurtosis, and sparsity), entropy-related metrics (including Shannon entropy, sample entropy, and frequency-domain entropy), frequency-domain attributes (main frequency, secondary frequency, power of the main frequency, and power of the secondary frequency), as well as time-domain features (such as difference sum, average jerk, and correlation between dimensions). These features collectively capture the nuances of the sensor signals and offer insights into the subjects' balance performance. These extracted features are comprehensively described in [12].

Notably, we opted not to perform feature selection in this study. The rationale behind this decision lies in utilizing the XGBOOST algorithm, a powerful and efficient machine-learning technique capable of handling high-dimensional and noisy data. XGBOOST can automatically discern the most pertinent features for the prediction task through its tree-based structure and regularization term. This approach is particularly advantageous for regression tasks, where traditional criteria for feature ranking might not be as straightforward. As a result, we retained all the extracted features as input for the XGBOOST model, allowing it to autonomously determine the most relevant features to predict m-CTSIB scores, enhancing the model's adaptability and performance.

C. Regression Model

This study employs the XGBOOST algorithm to estimate balance scores based on the lumbar sensor data. We utilized the XGBOOST package in Python to implement this regression model. We adopted a subject-wise One-Leave-Out cross-validation approach to ensure robustness and reliable model evaluation to partition the data into training and testing sets. Within the training set, we further divided the data into an 80% portion for model training and a 20% segment for model validation. Afterward, the data were normalized with the mean and standard deviation calculated from the training data. The grid search technique was applied to identify the optimal hyperparameters, specifically the number of trees, the depth of the decision tree, and the feature sampling rate. These three parameters are more important because beyond a certain point may lead to overfitting or longer training times. The correlation (r) and mean absolute error (MAE) were utilized to evaluate and assess the model's accuracy.

III. RESULTS AND DISCUSSION

The XGBOOST algorithm was independently applied to the extracted features obtained from the accelerometer and gyroscope to estimate AV and PL scores, resulting in four distinct models. The optimal models were determined based on minimizing MAE and maximizing r. Following cross-validation as described in section II-C, the selected XGBOOST model hyperparameters are reported in Table II.

Table III presents the outcomes, displaying the best results for PL and AV scores. Table III shows slight variations in the r values for PL score and AV score models in both the accelerometer and gyroscope. All r values for the training data are equal to one, and for the validation, it is approximately 0.97. In the case of the test data, the accelerometer-based models achieve the highest correlation coefficients, with values of 0.93 for AV and 0.94 for PL, respectively. Regarding MAE, models incorporating gyroscope data generally outperform those utilizing accelerometer data. Specifically, for the test data related to AV and PL scores, the MAE values are lower for gyroscope-based models, measuring at 0.26 and 2.4, respectively, compared to the accelerometer-based models with MAE values of 0.28 and 2.9.

Fig. 3 visually represents the Mean Absolute Percentage Error (MAPE) observed between the actual and predicted values across the EOSS, ECSS, EOFS, and ECFS test scenarios for the four models. MAPE is chosen as the evaluation metric instead of MAE due to the varying ranges of balance scores among these four conditions, allowing for more effective result comparisons, as detailed in Table I. ECFS, EOFS, and ECSS conditions exhibit notably low MAPE values, with ECFS achieving the lowest MAPE in the range of 0.17 to 0.19. This indicates that the models perform exceptionally well in predicting balance scores under these challenging conditions, reflecting their proficiency in scenarios where maintaining balance is considerably more difficult.

Our results underscore the superior performance of our approach, especially in terms of the correlation coefficient, signifying its efficiency and reliability in balance measurement compared to previous studies. In contrast to prior research that frequently relies on the duration of tests or costly specialized force plate systems like the Falltrak II, our method, utilizing a single wearable APDM sensor placed on the lumbar region, delivers a more dependable measurement within approxi-

TABLE II
THE HYPERPARAMETERS OF THE SELECTED XGBOOST MODELS

	Balance Score	**Hyperparameters**		
		No. Trees	Max Depth	FS Rate
Accel	AV	150	5	0.4
	PL	130	3	0.4
Gyro	AV	90	8	0.3
	PL	70	12	0.2

Accel: Accelerometer **Gyro**: Gyroscope **No**: Number of **Max**: Maximum **FS**: Feature Subsampling

TABLE III
ESTIMATION RESULTS FOR M-CTSIB SCORES USING XGBOOST

	Balance Score	**Train**		**Validation**		**Test**	
		r	MAE	r	MAE	r	MAE
Accel	AV	1	0.02	0.97	0.23	0.93	0.28
	PL	1	0.11	0.97	2.2	0.94	2.9
Gyro	AV	1	0.02	0.98	0.19	0.90	0.26
	PL	1	0.18	0.97	2.1	0.89	2.4

Accel: Accelerometer **Gyro**: Gyroscope

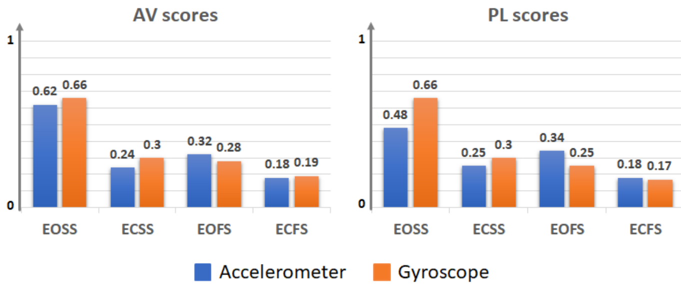

Fig. 3. Comparative analysis of minimum absolute error percentage (MAPE) between actual and predicted AV and PL scores in different m-CTSIB sensory assessments of EOSS, ECSS, EOFS, and ECFS.

mately 10 seconds per test under four different conditions. Our results indicate that XGBOOST models, using either accelerometer or gyroscope features from APDM wearable sensors on the lumbar region, accurately estimate m-CTSIB scores. With lower power consumption and greater affordability and availability than gyroscopes, accelerometer sensors emerge as a more practical choice. This enhances the overall feasibility of utilizing accelerometer sensors for m-CTSIB score estimation.

As our work is the first to estimate m-CTSIB scores, direct comparisons are challenging. Nevertheless, our evaluation metrics, including MAE, exhibited similarities to those reported by Choi *et al.* in predicting TUG scores [13]. Remarkably, our correlation coefficient ($r = 0.94$) outperformed the correlation coefficient reported by Tang *et al.* ($r = 0.37$) for predicting BBS scores, despite assessing a different type of balance score [14]. These findings underscore the effectiveness and potential of our approach in the field of balance assessment.

Our subject size of 34 participants aligns with previous studies using sensor technology and machine learning to measure balance. For instance, Tang *et al.* collected data from 30 older adults to predict BBS scores, while Choi *et al.* gathered data from 37 older adults to predict TUG scores. This subject size provides reliable and valid results for estimating m-CTSIB scores using wearable sensors and the XGBOOST algorithm. However, we recognize that a larger and more diverse subject population would enhance the generalizability and robustness of our approach. Consequently, our future work will involve recruiting a broader range of participants with varying characteristics and conditions.

IV. CONCLUSIONS

In this paper, we introduced a novel approach to estimating m-CTSIB balance test scores using wearable sensors and machine learning. Our dataset consisted of 34 participants who performed the m-CTSIB test under four sensory conditions using the Falltrak II system, with a single APDM sensor attached to their lumbar region. We extracted features from the accelerometer and gyroscope signals and trained XGBOOST models to predict AV and PL scores for each condition. Our results demonstrated high accuracy and correlation with actual scores, with correlation coefficients (r) ranging from 0.89 to 0.94. Our approach offers significant advantages over existing methods for measuring balance performance. It provides an objective and reliable measurement that eliminates the need for human observation or judgment. Moreover, it utilizes a single wearable sensor, ensuring ease of use and reducing system complexity and cost. Additionally, the test is quick and straightforward, taking less than a minute to complete and minimizing participant inconvenience and fatigue. Based on our findings, we recommend using the accelerometer due to its comparable performance to the gyroscope and lower power consumption, making it a more power-efficient choice. Our proposed approach could enhance the assessment and management of balance performance and consequential complications, particularly for older adults and at-risk individuals. Its objectivity, simplicity, and efficiency make it a valuable contribution to the field of balance evaluation, paving the way for improved clinical applications and research in this domain.

REFERENCES

[1] A. S. Pollock, B. R. Durward, P. J. Rowe, and J. P. Paul, "What is balance?," *Clinical rehabilitation*, vol. 14, no. 4, pp. 402–406, 2000.

[2] M. Mancini and F. B. Horak, "The relevance of clinical balance assessment tools to differentiate balance deficits," *European journal of physical and rehabilitation medicine*, vol. 46, no. 2, p. 239, 2010.

[3] S. Fritz and M. Lusardi, "Walking velocity: The sixth vital sign," *Journal of Geriatric Physical Therapy*, vol. 32, no. 2, pp. 5–5, 2009.

[4] B. Schoneburg, M. Mancini, F. Horak, and J. G. Nutt, "Framework for understanding balance dysfunction in parkinson's disease," *Movement disorders*, vol. 28, no. 11, pp. 1474–1482, 2013.

[5] A. C. Alonso, N. M. Luna, F. N. Dionísio, D. S. Speciali, L. E. G. Leme, and J. M. D. Greve, "Functional balance assessment," *Medicalexpress*, vol. 1, pp. 298–301, 2014.

[6] L. B. Horn, T. Rice, J. L. Stoskus, K. H. Lambert, E. Dannenbaum, and M. R. Scherer, "Measurement characteristics and clinical utility of the clinical test of sensory interaction on balance (ctsib) and modified ctsib in individuals with vestibular dysfunction," *Archives of physical medicine and rehabilitation*, vol. 96, no. 9, pp. 1747–1748, 2015.

[7] M. Seifallahi, J. E. Galvin, and B. Ghoraani, "Detection of mild cognitive impairment from quantitative analysis of timed up and go (tug)," in *2022 IEEE ICDMW*, pp. 248–253, IEEE, 2022.

[8] D. L. Freeman, G. Gera, F. B. Horak, M. T. Blackinton, M. Besch, and L. King, "The instrumented test of sensory integration for balance: A validation study," *Journal of geriatric physical therapy (2001)*, vol. 41, no. 2, p. 77, 2018.

[9] P. Suttanon, K. D. Hill, K. J. Dodd, and C. M. Said, "Retest reliability of balance and mobility measurements in people with mild to moderate alzheimer's disease," *International Psychogeriatrics*, vol. 23, no. 7, pp. 1152–1159, 2011.

[10] W.-H. Wang, Y.-L. Hsu, P.-C. Chung, and M.-C. Pai, "Predictive models for evaluating cognitive ability in dementia diagnosis applications based on inertia-and gait-related parameters," *IEEE Sensors Journal*, vol. 18, no. 8, pp. 3338–3350, 2018.

[11] R. Mc Ardle, S. Pratt, C. Buckley, S. Del Din, B. Galna, A. Thomas, L. Rochester, and L. Alcock, "Balance impairments as differential markers of dementia disease subtype," *Frontiers in Bioengineering and Biotechnology*, vol. 9, p. 639337, 2021.

[12] M. D. Hssayeni, M. A. Burack, J. Jimenez-Shahed, and B. Ghoraani, "Assessment of response to medication in individuals with parkinson's disease," *Medical engineering & physics*, vol. 67, pp. 33–43, 2019.

[13] J. Choi, S. M. Parker, B. A. Knarr, Y. Gwon, and J.-H. Youn, "Wearable sensor-based prediction model of timed up and go test in older adults," *Sensors*, vol. 21, no. 20, p. 6831, 2021.

[14] W. Tang, G. Fulk, S. Zeigler, T. Zhang, and E. Sazonov, "Estimating berg balance scale and mini balance evaluation system test scores by using wearable shoe sensors," in *2019 IEEE BHI*, pp. 1–4, IEEE, 2019.