A Comprehensive Infrastructure and Methodology for Multi-Modal Data Acquisition to Empower AI-Based Rehabilitation

Katerina Tzatzimaki[1][a], Nick Portokallidis[2], George Drosatos[2][b], Eleni Kaldoudi[1, 2][c] and Stylianos Didaskalou[2][d]

[1]School of Medicine, Democritus University of Thrace, 68100 Alexandroupoli, Greece
[2]Institute for Language and Speech Processing, Athena Research Center, 67100 Xanthi, Greece
aikatzat@med.duth.gr, nikos.portokallidis@athenarc.gr, gdrosato@athenarc.gr, kaldoudi@athenarc.gr, stelios.didaskalou@athenarc.gr

Keywords: Inertial Measurement Unit, Clinical Rehabilitation, Physical Activity Recording, Human Activity Recognition, Artificial Intelligence.

Abstract: Accurate and synchronized motion tracking is essential for advancing quantitative assessment and personalized rehabilitation. This paper presents a comprehensive infrastructure and methodology for multi-modal data acquisition designed to power AI-based rehabilitation. The system integrates multiple inertial measurements units (IMUs) and dual-camera recordings within a unified software environment that ensures reliable connectivity, synchronization and calibration. A static calibration procedure corrects sensor-to-segment misalignments, while real-time visualization enables immediate assessment of signal quality during acquisition. Although video recordings will be used exclusively for model development, their combination with IMU data will enable the creation of multimodal datasets to train AI-models that rely solely on inertial data during clinical deployment. These models aim to enhance signal accuracy by compensating for noise, drift and alignment errors. The presented infrastructure and methodology established a robust foundation for the development of AI-based rehabilitation tools to empower unsupervised rehabilitation.

1 INTRODUCTION

Physical rehabilitation often requires continuous and accurate monitoring of patient movements to assess progress and guide personalized treatment plans. Traditional clinical assessments are often based on observational scoring or periodic evaluation sessions, lacking the temporal resolution and detailed kinematic insights, especially during unsupervised or home-based exercising.

This shortfall has pushed towards the development of digital tools that capture human body motion in detail, in real-world scenarios (Porciuncula et al., 2018).

In recent years, motion capture technologies have emerged as powerful instruments for assessing human movement during physical activity or rehabilitation (Adlou et al., 2025; Gu et al., 2023). Optical systems serve as the gold-standard, providing high spatial and temporal resolution but typically involve high cost, complex set up and restricted use outside the clinical environment. Meanwhile, wearable inertial measurement units (IMUs) have emerged as alternative, due to their portability and cost-effectiveness and more importantly because they allow motion capture in real-world settings (Gu et al., 2023). Their portability makes them attractive for both home and clinical-based rehabilitation scenarios. In addition, recent studies have demonstrated a strong agreement between IMU-based and optical systems in lower limb kinematics (Zhu et al., 2024). However, IMU-based systems do not come without limitations. Sensor drift, magnetometer disturbances, calibration inconsistencies and sensors-to-segment misalignments

[a] https://orcid.org/0009-0002-1393-1210
[b] https://orcid.org/0000-0002-8130-5775
[c] https://orcid.org/0000-0002-6054-4961
[d] https://orcid.org/0000-0001-5932-9626

Tzatzimaki, K., Portokallidis, N., Drosatos, G., Kaldoudi, E. and Didaskalou, S.
A Comprehensive Infrastructure and Methodology for Multi-Modal Data Acquisition to Empower AI-Based Rehabilitation.

In *Proceedings of the 19th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2026) - Volume 3: HEALTHINF*, pages 269-278
ISBN: 978-989-758-802-0; ISSN: 2184-4305

result in measurement degradation over time, potentially affecting the reliability of derived kinematics.

To overcome this limitation and to enable the development of robust, data-driven models for motion analysis, multi-modal data acquisition systems have gained increasing interest (García-de-Villa et al., 2022; Martínez-Zarzuela et al., 2023; Arias-Correa et al., 2024). By recording synchronized IMU and video data, it is possible to create rich data sets that facilitate the training and validation of artificial intelligence (AI) models for motion analysis, activity recognition and anomaly detection. Such models can subsequently be deployed using only wearable IMUs, without requiring video data in the clinical environment. Therefore, multimodal data collection serves as an essential intermediary step toward the creation of reliable, sensor-only AI models for self-guided rehabilitation systems.

In this context, the present work introduces a modular software platform designed to facilitate synchronized acquisition of IMU and video data with the aim to streamline dataset generation for AI model development, ensuring time-aligned, and high-quality multimodal recordings, while remain adaptable to various rehabilitation research needs.

This work is part of the ThrombUS+ project, cofunded by the European Union (EU) under Grant Agreement No. 101137227 (Kaldoudi et al., 2024). The ThrombUS+ project aims on developing a novel wearable device for early detection and prevention of Deep Vein Thrombosis (DVT). This infrastructure has been developed to facilitate the collection and preparation of activity data, required for the developing a self-supervised AI -based rehabilitation system which exploits extended reality for patient guidance and serious gaming for adherence and empowerment.

2 RELATED WORK

Human body motion tracking can be achieved using motion capture systems, which can be generally categorized as either optical or non-optical, depending on whether they are using cameras to capture visual information. Optical systems can be further divided into marker-based (e.g. Vicon), which rely on reflective markers placed at specific position on the body, and markerless systems (e.g. Microsoft Kinect), which estimate motion directly from visual data. In contrast, non-optical system such as IMUs (e.g. XSens) use a combination of sensors, including accelerometers and gyroscopes, to measure linear acceleration and angular velocity of the body segments to which they are attached to. These measurements can further be used to estimate motion (Salisu et al., 2023). Beyond of these categories, various other sensing modalities, such as RGB cameras, radar, Wifi-based sensing, depth, skeletons and point clouds, have also been explored for human action recognition and motion analysis (Z. Sun et al., 2023).

Several studies have demonstrated the potential of motion capture technologies in rehabilitation. A recent review highlights the use of motion capture technology in rehabilitation to analyze motor disorders in Parkinson's disease and cerebral palsy, exploiting their distinctive movement patterns (Lam et al., 2023). Similarly, Cerfoglio et al., (Cerfoglio et al., 2023) validated Movella DOT IMUs in a rehabilitation context, reporting minimum angular errors, confirming the suitability for clinical grade motion analysis. Building upon these developments, IMUs have been extensively employed to monitor patients remotely and support movement assessments, a key step towards effective telerehabilitation. Caramia et al., (Caramia et al., 2025) proposed a low cost telerehabilitation system enabling remote exercise monitoring through inertial sensors, while Marsan et al. (Marsan et al., 2025) applied magnetometer free IMUs to assess gait in stroke patients, achieving clinically acceptable accuracy. Similarly, Sun et al., (Sun et al., 2025) utilized multi-IMU configurations for quantitative stroke motor evaluation, demonstrating strong correlations with established clinical scores.

Even though previous works have reported high levels of accuracy for IMU-based motion analysis, consistent calibration and alignment between sensors and body segments remain key challenges that can affect measurement stability. Several studies have proposed calibration or drift correction methods to enhance IMU measurement reliability. Even small misalignments between the sensor coordinate frame and the anatomical frame can lead to cumulative orientation errors during motion tracking. Hansen et al. applied functional calibration procedures based on accelerometer signals (Hansen et al., 2023). The results confirmed that the acceleration based functional calibration was effective in correcting magnetic fields disturbances, enabling reliable movement assessment even outside the laboratory environments. On the other hand, Niswander et al. (Niswander et al., 2020) investigated how the placement of IMUs affects the accuracy of the measurements during functionally relevant movements. Furthermore, Gonzalez -Alonso et al. (González-Alonso et al., 2021) proposed a simple, yet effective quaternion-based sensor to segment calibration method and developed a 2.4 GHz wireless protocol to improve data acquisition reliability in environments with potential inference.

Complementary to the IMU-focused research, a critical area of study involves the development of multimodal systems that synchronize and fuse inertial data with optical motion data (cameras), aiming to overcome the limitations of single modality systems. A critical contribution in this regard is the creation of publicly available datasets for robust validation. For instance, the PHYTMO database, includes data from various physical therapy exercises and gait variations, recorded using magneto-inertial sensors and an optical reference system (García-de-Villa et al., 2022). Similarly, the VIDIMU dataset includes 13 clinically relevant Daily Life Activities, recorded with synchronized video from a low-resolution camera and five IMUs (Martínez-Zarzuela et al., 2023). More recently, in their work, Arias-Correa et al proposed a data acquisition system for cyclists that records RGB camera and dual IMUs mounted on a bicycle and vehicle to record synchronized datasets of cyclist orientation and behavior in traffic (Arias-Correa et al., 2024).

Building upon this complementary approach, the present work introduces the design and development of o novel comprehensive infrastructure and methodology, for physical therapy exercise monitoring that leverages the combined strengths of multiple IMU sensors and a dual-camera setup to overcome the limitations of single-modality tracking, ensuring reliable motion measurement and synchronized multimodal data acquisition.

3 METHODS

3.1 System Description

The infrastructure is based on a comprehensive software application designed for the acquisition, visualization and storage of human motion data, utilizing six IMU sensors and two cameras (providing lateral and axial views). Developed in Python and built upon the Movella's Software Development Kit (SDK), the system is engineered to supports real time processing and synchronized recording of multimodal data streams.

3.2 Architecture

The system architecture is organized into four core subsystems as presented in Figure 1: (1) the IMU Manager which is responsible for detecting, connecting and retrieving data from up to six IMU sensors, ensuring continuous data flow and measurement integrity; (2) the Camera Manager, which manages the operation of two cameras, enabling simultaneous video streaming and recording for multi-angle motion capture; (3) the File Manager, which handles the structured and time synchronized storage of IMU data, video streams and system timestamps into various formats, including .csv, .txt and .mp4 files; and the Graphical User Interface (GUI) which serves as the central control layer, allowing the user to configure, initiate and monitor both sensor and camera operations in real time. This modular design separates, acquisition storage and visualization, enabling clear data flow and future extensibility.

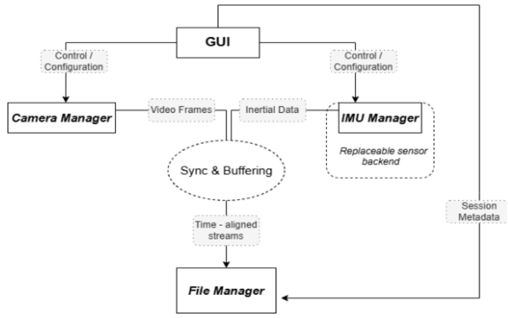


Figure 1: System architecture of the proposed recording infrastructure.

During acquisition, the GUI orchestrates the workflow by applying the selected configuration (e.g., sampling frequency, sensor assignment, recording folder and exercise label) and triggering start/stop actions. The IMU Manager streams synchronized inertial packets (orientation/quaternions, acceleration, angular velocity, magnetic field and sensor timestamps) from up to six sensors, while the Camera Manager captures frames from the two USB cameras.

Both data streams are handled concurrently and forwarded through thread-safe buffering mechanisms (e.g., events/queues) to ensure continuous flow without data loss. In parallel, the File Manager consumes the buffered streams and stores them in a structured manner (IMU data in .csv/.txt and video in .mp4), together with session metadata (date/time and protocol/exercise label), preserving temporal alignment across modalities.

The modular separation between acquisition (IMU/Camera Managers), control and visualization (GUI) and storage (File Manager) enables straightforward extensibility. First, the sensing backend can be adapted to other wearable inertial devices by replacing or extending the IMU Manager while keeping the same interface, without requiring changes to the GUI or File Manager. Second, the system can support different rehabilitation protocols by updating the configurable session parameters and exercise labelling scheme (e.g., menus and metadata) while reusing the same acquisition, synchronization and storage pipeline. This design allows the infrastructure to remain reusable across studies and protocols with minimal implementation effort.

3.3 Implementation Details

The system was implemented in Python using widely adopted open-source tools to support graphical user interaction, real-time data acquisition, visualization and efficient data handling. The implementation prioritizes modularity, robustness and real-time performance, rather than language-specific optimizations.

To support simultaneous acquisition from multiple IMU sensors and cameras, the architecture adopts concurrent execution and event-driven synchronization, ensuring non-blocking data flow and stable operation during recording sessions. Buffered data exchange mechanisms are employed to preserve data integrity and prevent packet loss under varying acquisition conditions.

Robustness is further handling enhanced through integrated error handling and logging mechanisms, enabling continuous monitoring of system status and automatic recovery from transient hardware or communication faults without interrupting ongoing recordings. This design allows the application to operate autonomously while providing real-time feedback to the user.

3.4 Sensor Placement and Connectivity

The IMU sensors are placed at specific anatomical locations on the lower limbs, as illustrated in 2. This configuration allows full lower limb motion tracking.

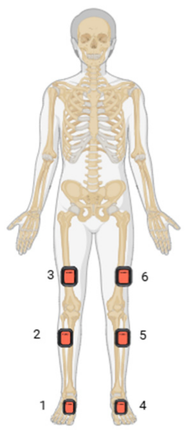


Figure 2: Sensor placement and numbering scheme. Six IMU sensors (1-6) are attached to the left and right anterior mid- thighs, anterior mid-shanks, and on the dorsal surface to capture lower-limb motion during data acquisition. The numbering corresponds to the data labelling convention used in the recording and analysis process.

As it shown in Figure 3 sensor connectivity is achieved via Bluetooth Low Energy (BLE), a protocol that ensures low power consumption and stable wireless communication. Each sensor transmits data including quaternions, acceleration, angular velocity and magnetic field, at a sampling rate defined by the user. Simultaneously, two cameras are connected via USB to record video from two different angles. The synchronized operation of sensors and cameras is managed through threading events, each responsible for a distinct system function. These events allow multiple threads to run concurrently without interference, ensuring smooth data flow and minimizing latency during acquisition.

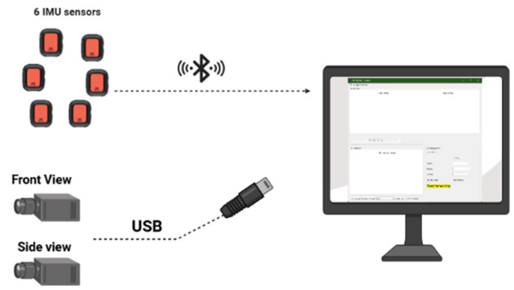


Figure 3: Overview of the infrastructure's setup, illustrating the connection of six IMU sensors via Bluetooth and dual camera via USB to the main recording and visualization interface.

3.5 Coordinate System and Calibration Methods

Each IMU sensor operates within its own local axis system (Figure 4), which depends on the specific placement of the device on the body (D'Alcala' et al., 2021). Therefore, a calibration procedure is required prior to each use to ensure consistent orientation across all sensors.

Figure 4: Movella DOT sensor's coordinate system (D'Alcala' et al., 2021).

Initially, the reset heading method is applied to align all sensors to a shared orientation (common heading) (D'Alcala' et al., 2021). However, in addition to orientation discrepancies, errors may arise due to soft tissue artifacts, imperfect mounting and individual anatomical differences, as shown in Figure 5.

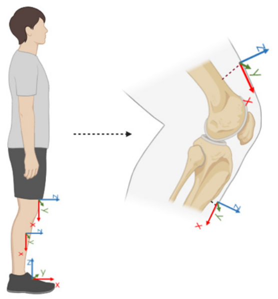


Figure 5: Schematic representation of IMU placement and sensor coordinate systems on the lower limb. Each sensor defines a local reference frame (X–Y–Z) used for computing segment orientation and joint kinematics during motion analysis.

To address these issues, a static calibration procedure is performed before each recording session, to ensure alignment between the sensor and the body segment reference frame. Briefly, data collection study participants are instructed to adopt a predefined neutral position (e.g. lying, or sitting, or standing), and quaternions are recorded. Subsequently, these quaternions are used to normalize the orientation data acquired during the exercises using equation (1).

$$q' = q_{calib}^{-1} * q \tag{1}$$

3.6 Recording Protocol

Before each recording session, all six IMU sensors are powered on and are connected via Bluetooth, followed by the synchronization process, to ensure a common internal clock across all devices. The software automatically verifies sensor connectivity and signal stability before allowing data acquisition. Finally, the heading of all sensors is aligned to a common reference point.

Data collection is carried out according to a standardized protocol with the aim of ensuring the reproducibility of measurements and the comparability of results (Didaskalou et al., 2025). The experimental setup utilizes six IMU sensors placed symmetrically on the participant's lower limbs via securing straps (Figure 2). Also, a dual-camera setup is also employed to capture both axial and lateral views of the participant's motion, as mentioned before. The cameras are positioned and configured to capture the full range of lower limb exercise movements while ensuring that no facial data are recorded.

Following camera positioning, system initialization, and sensor placement, a static calibration step is performed by requesting participants to adopt a predefined pose for 3 seconds before each recording.

Subsequently, the recording session begins, and all sensor and camera streams are synchronized. During recording, the interface displays real-time visualization of IMU measurements through quaternion-based plots, synchronized with live camera feeds. Upon manual termination by the operator, the system automatically stores all data streams in the folder selected by the operator and labelled with date, time, and the exercise variation, chosen from a predefined list prior to recording. IMU data are saved in .csv and .txt format, and video streams in .mp4 format.

3.7 Timestamp Stability Analysis

To assess the reliability and accuracy of the recorded timestamps, a dedicated analysis was performed. The experimental measurements were conducted using a single healthy volunteer who was instrumented with six IMU sensors, positioned according to the predefined anatomical configuration described in the recording protocol. The participant performed one exercise from the proposed physiotherapy routine, designed to simulate a typical rehabilitation movement. Data acquisition followed the standardized protocol, ensuring consistent sensor attachment, proper calibration, and controlled environmental conditions. A detailed evaluation of timestamp consistency is reported in **Section 4.2**.

3.8 Sensor Placement Analysis

To quantify the impact of sensor misplacement on the accuracy of recorded acceleration signals, a controlled series of trials was conducted. In these, the sensor on the left thigh was positioned at various anatomical locations other than the correct one. The acceleration signals from each incorrect placement were systematically compared to those from the correct one, in order to calculate key statistical error metrics, providing a quantitative assessment of the signal deviation.

Measurements were performed in a standing posture, where the gravitational component dominates the X-axis of the Movella coordinate system, while Y and Z axes reflect dynamic factors such postural sway. The results of this analysis are presented in **Section 4.4.**

4 RESULTS

4.1 Overview of the GUI

The GUI of the infrastructure includes three main modules: the Main tab for system control, the Settings tab for configuration and sensor assignment and the Visualization tab for synchronized data visualization and post-recording analysis. The application is open-source and available on GitHub at https://github.com/thrombusplus/recording-studio under MIT license.

The Main tab provides real-time control of sensors, allowing data streaming, data recording and data visualization from all sensors in real time (Figure 6).

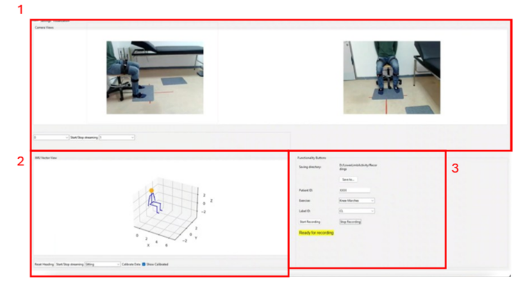


Figure 6: Main tab of the application interface. It provides real-time control of streaming and recording, with live camera previews (front and side views) at the top (1), a 3D IMU Vector View displaying sensor orientations and control buttons (2). Buttons and menus for starting and stopping recordings are placed on the bottom right (3).

In the Settings tab, the user can select the required sampling frequency of the IMU sensors, connect to available IMUs and available web-cameras and assign each connected IMU to a specific leg segment for proper activity tracking (Figure 7).

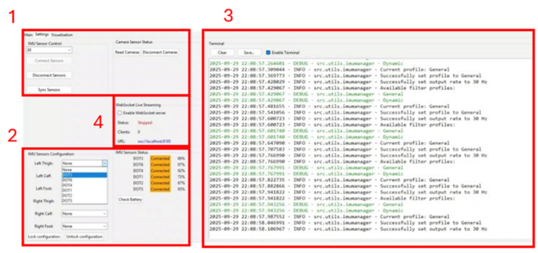


Figure 7: The Settings tab contains all configuration options necessary before recording. It allows the user to set the sampling frequency and connect to available sensors and cameras and synchronize IMUs (1), assign connected IMU sensors to leg segments or check connection and battery status (2). Also, it integrates a logging panel for printing all errors and system activity (3). Additionally, WebSocket activation buttons allow the streaming via a websocket to allow communication with other applications (4).

Finally, the Visualization tab allows users to load and review all visual data, including 3D skeletal poses over time along their respective synchronized video frames as well as IMU data plots (Figure 8).

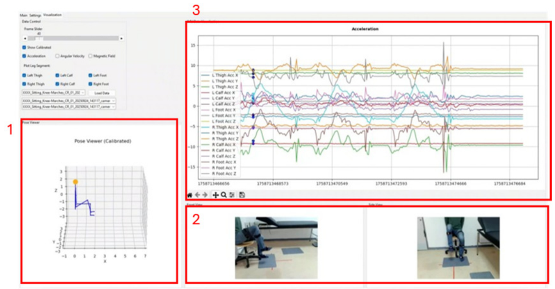


Figure 8: The Visualization tab combines all visualization modules. It displays quaternion-based motion plot utilizing a 3D skeletal model (Pose Viewer) (1), along the respective synchronized front and side camera feeds (2), as well as raw data recordings over time (3), allowing for comprehensive motion analysis.

4.2 Timestamp Stability Analysis

This section presents the results of the timestamp analysis, focusing on the evaluation of sampling rate consistency and synchronization accuracy among the IMU sensors.

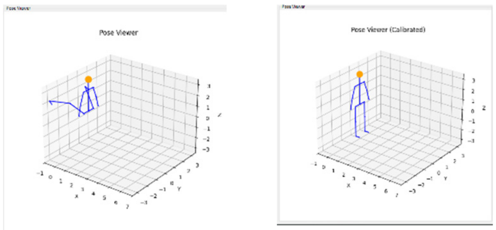
The timestamp analysis demonstrated high stability and synchronization consistency across all IMU sensors, as presented in Table 1 .The mean inter-sample interval was 33.4 ±0.5 ms (mean ± SD, n=1437 samples per sensor), corresponding to an effective sampling frequency of 29.9 ±0.2 Hz, in close agreement with the nominal rate of 30 Hz. The standard deviation ranged between 1.75 and 2.22 ms, confirming stable sampling performance. Occasional outlier (0 ms or 100 ms) was observed but occurred infrequently and has negligible influence on the overall data quality. The dropout rate remained exceptionally low (<0.004%) for all sensors, confirming the robustness of the acquisition pipeline and the temporal precision of the synchronized recordings.

4.3 Effect of Static Calibration on 3D Human Pose Representation

The 3D human model is employed to provide a visual illustration of the calibration process. Before calibration, when the sensors were placed on the corresponding anatomical sites of the volunteer, the initial visualization (Figure 9a) showed noticeable deviations in the alignment of the leg segments, indicating orientation offsets between sensors. After the calibration procedure, the updated visualization (Figure 9b) demonstrated a significantly improved alignment with the volunteer's actual pose and smoother motion during activity. These results highlight the effectiveness and necessity of the calibration step prior to data recording and visualization.

Table 1: Summary of timestamp analysis metrics for all IMU sensors, including mean and standard deviation of inter-sample intervals, effective sampling frequency, and minimum–maximum values recorded during acquisition.

IMU ID	Mean Δt (ms)	Standard Deviation (ms)	Effective Frequency (Hz)	Minimum value	Maximum Value	Dropouts (%)
Movella DOT 0	33.42638581	1.752692581	29.9164859	33.3	66.7	0.002771619
Movella DOT 1	33.40787140	2.218178302	29.9330654	0	66.7	0.003325942
Movella DOT 2	33.42638581	2.074847260	29.9164859	33.3	100	0.002217295
Movella DOT 3	33.42638581	1.752692581	29.9164859	33.3	66.7	0.002771619
Movella DOT 4	33.42638581	1.753745443	29.9164859	33.3	66.7	0.002771619
Movella DOT 5	33.44490022	1.920672214	29.8999247	33.3	66.7	0.003325942



(a) (b)

Figure 9: The effect of the calibration procedure on the IMU quaternion signals. The uncalibrated data (a) show noticeable offsets and orientation drift among sensors, while the calibrated data (b) exhibit improved alignment and temporal consistency across all segments.

4.4 Impact of Sensor Placement

Table 2, Table 3 and Table 4 present the results of the analysis, as the quantified error for every misplacement configuration examined.

For example, when the sensor on the left thigh was displaced 5 cm downward, the resulting differences remained negligible along the X-axis (static gravitational component) and Y-axis (lateral motion, with RMSE =0.045 ± 0.025g and 0.039 ± 0.020g, respectively (n=748). In contrast, a higher deviation was observed along the Z-axis, where RMSE increased to 0.01 ±0.064g, indicating amplification of local soft-tissue motion near the lower placement.

In the inward lateral displacement (5cm) case, larger deviations were observed across all axes. The Z-axis again showed the greatest sensitivity (RMSE = 0.243±0.307g), while the X-axes and Y-axes exhibited moderate distortions (RMSE=0.104±0.083g and 0.0055±0.090g, respectively). The corresponding Peak Amplitude Error (PAE) reached 0.938g on the Z-axis, confirming significant alternation of the motion amplitude.

Table 2: Statistical error metrics (Bias, RMSE, MAE, STDEVP, PAE) for the configuration in which the sensor on the left thigh was displaced 5 cm downward from its correct anatomical position.

Axes	X	Y	Z
Bias	$2.11e^{-15}$	$2.13e^{-16}$	$-2.30e^{-15}$
RMSE	0.045032865	0.039051966	0.101372191
MAE	0.042132865	0.030248034	0.120372191
STDVEP	0.024534761	0.034593523	0.064299617
PAE	0.0541	0.0259	0.2608

Table 3: Statistical error metrics (Bias, RMSE, MAE, SD, PAE) for the configuration in which the sensor on the left thigh was displaced 5 cm inward lateral from its correct anatomical position.

Axes	X	Y	Z
Bias	$-5.3e^{-16}$	$-1.6e^{-16}$	$-2.60e^{-16}$
RMSE	0.104249854	0.005541691	0.242949563
MAE	0.067649854	0.066358309	0.198149563
STDVEP	0.083136731	0.089622743	0.30691868
PAE	0.3788	0.5336	0.938

For outward lateral displacement (5cm), the deviations remained substantial, particularly for the Y-axis (RMSE=0.079±0.148g, PAE=1.436), suggesting strong misalignment between the sensor and the anatomical axes. The Z-axis presented RMSE=0.131±0.140g and PAE = 0.505g, indicating persistent vertical instability.

Table 4: Statistical error metrics (Bias, RMSE, MAE, SD, PAE) for the configuration in which the sensor on the left thigh was displaced 5 cm ward lateral from its correct anatomical position.

Axes	X	Y	Z
Bias	$2.11e^{-15}$	$2.13e^{-16}$	$-2.30e^{-15}$
RMSE	0.073503207	0.078750146	0.131481633
MAE	0.057803207	0.050549854	0.181681633
STDVEP	0.057367601	0.147739335	0.140133462
PAE	0.4574	1.4364	0.505

5 DISCUSSION

The experimental validation demonstrates that the proposed infrastructure and methodology provide a reliable approach for synchronized multimodal motion data acquisition. The timestamp analysis confirmed stable communication with all six IMU sensors, maintaining a consistent sampling frequency of approximately 30 Hz (33.4 ± 0.5 ms), with an insignificant dropout rate (<0.004%). These findings verify the system's ability to maintain temporal stability and synchronization, which are essential for accurate sensor fusion and motion reconstruction.

Regarding the acceleration analysis, the controlled sensor misplacement experiments revealed the strong influence of positional errors on the recorded signals. Downward displacement of the thigh sensor produced minor deviation along the X and Y- axes (RMSE< 0.05g), but significantly increased error along the Z-axis (RMSE = 0.10 ± 0.06g), where vertical motion components dominate. Lateral displacements (inward and outward) resulted in even higher deviations, with PAE exceeding 1g, confirming that small variations in sensor positioning can substantially affect measurement integrity. These results are in agreement with those of Niswander et al.,(Niswander et al., 2020) and Hansen et al.,(Hansen et al., 2023), who reported that IMU-base accuracy strongly depends on proper anatomical alignment and calibration procedures. From a rehabilitation perspective, these findings highlight the importance of ensuring precise and repeatable sensor placement during patient monitoring and exercise evaluation. The system's synchronized multimodal setup, combining IMU and camera data, ensures that researchers can obtain synchronized quantitative kinematic metrics and visual confirmation of patient movements to build comprehensive data sets. This dual perspective enhances clinical interpretability and allows for explainable and robust AI models.

Despite its promising performance, the proposed system presents certain limitations. The accuracy of IMU data remains sensitive to sensor's placement errors and environmental magnetic disturbances. In addition, the current implementation does not include automated calibration or motion segmentation algorithms.

Furthermore, a limitation of the present study is that the synchronization accuracy between the IMU sensors and the video streams was not quantitively assessed. Although the system enforces simultaneous start and end events for all streams through a unified software workflow and common control mechanisms, a standard multimodal synchronization analysis (such as frame-level alignment between video timestamps and IMU data) was not included within the scope of this experimental validation. This decision stems from the primary objective of the study, which was the technical validation of the reliability, temporal stability and robustness of the inertial infrastructure, the core sensing modality for the system's intended clinical application. Video recordings were integrated as a supplementary information source, aimed at supporting data labeling and model development, rather than serving as a high-precision reference for direct kinematic measurement.

However, it should be noted that while quantitative verification is pending, the recorded data can be observed in a synchronized manner through the system's visualization tab, providing qualitative assurance of temporal alignment during acquisition.

6 FUTURE WORK

Recent advances in AI and ML have shown great potential in automating motion analysis and rehabilitation assessment. Building upon the technical validation of the acquisition infrastructure presented in this study, specifically regarding synchronization stability and robustness, our future work will focus on expanding the software's capabilities to bridge the gap between data acquisition and automated clinical insight.

A key direction will be the quantitative validation of cross-modal synchronization between IMU sensors and camera streams. Future developments will include frame-level temporal alignment analysis, controlled synchronization experiments and systematic estimation of inter-stream latency, aiming to further strengthen the temporal coherence of the multimodal datasets generated by the system.

In parallel with these technical refinements, we plan to integrate advanced data labelling and annotation features within the Studio platform. By enhancing the software infrastructure to support the creation of high-quality, structured datasets, we aim to facilitate the robust training of ML models. Such enhancements will enable the automatic recognition of specific rehabilitation exercises and movement patterns. Ultimately, these capabilities will serve as the backbone for developing intelligent, adaptive feedback systems designed to provide personalized therapy.

Figure 10 presents an AI-based rehabilitation pipeline, where the developed infrastructure serves as the foundation. The process begins with multimodal data acquisition (IMU and video), which provides the high quality, synchronized datasets necessary for AI model development. These models will be integrated into clinical and home-based deployment, such as serious games and extended reality environments, creating a feedback driven system that supports personalized therapy and continuous model refinement.

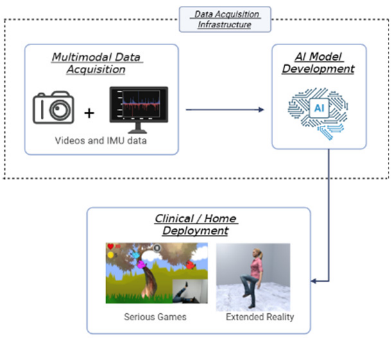


Figure 10: Conceptual overview of an AI-based rehabilitation system, illustrating the role of the proposed infrastructure within the overall pipeline from data acquisition to clinical deployment and feedback-driven model refinement.

Alongside these technical improvements, a large-scale clinical data collection is planned. The software platform will be deployed in clinical settings to record rehabilitation exercises performed by patients according to standard protocols. This growing database will allow for the detailed study of differences between subjects. These datasets will provide the necessary foundation to build, train, and test AI models, with the final goal of improving rehabilitation support and assisting in clinical decision-making. Parallel to these technical developments, a large-scale clinical data collection is currently underway. The proposed software platform is being deployed in clinical settings to record rehabilitation exercises performed by patients under standardized protocols. This growing repository of clinical data will be instrumental for future research, enabling comprehensive inter-subject variability analysis. These datasets will provide the necessary foundation for the development, training and rigorous evaluation of AI-based models aimed at enhancing rehabilitation support and clinical decision-making.

7 CONCLUSIONS

The infrastructure and methodology presented here ensures reliable synchronization, calibration and real-time visualization, enabling the collection of high-quality datasets suitable for both clinical research and model training. While this work establishes a robust foundation for synchronized multimodal motion capture, future extensions will focus on the full quantitative validation of camera-to-IMU synchronization to further enhance dataset reliability. Future work will focus on incorporating annotation and labeling tools within the platform to streamline dataset preparation and facilitate the development of self-supervised and adaptive learning models for intelligent, personalized rehabilitation systems.

ACKNOWLEDGEMENTS

This work is co-funded by the European Union, under the Horizon Europe Innovation Action ThrombUS+ (Grant Agreement No. 101137227). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or HADEA as the granting authority. Neither the European Union nor the granting authority HADEA can be held responsible for them. Also, his work was carried out in partial fulfilment of the requirements for the MSc degree of K.T. in the MSc Course Biomedical Informatics, which is co-organized by the School of Medicine, Democritus University of Thrace, and the ATHENA Research Center, Greece.

REFERENCES

Adlou, B., Wilburn, C., & Weimar, W. (2025). Motion Capture Technologies for Athletic Performance Enhancement and Injury Risk Assessment: A Review for Multi-Sport Organizations. *Sensors*, *25*(14), 4384.

Arias-Correa, M., Robledo, S., Londoño, M., Bañol, J., Madrigal-González, C., Ballesteros, J. R., & Branch-

Bedoya, J. W. (2024). CYCLOPS: A cyclists' orientation data acquisition system using RGB camera and inertial measurement units (IMU). *HardwareX*, *18*, e00534.

Caramia, F., D'Angelantonio, E., Lucangeli, L., & Camomilla, V. (2025). Validation of Low-Cost IMUs for Telerehabilitation Exercises. *Sensors (Basel, Switzerland)*, *25*(10), 3129.

Cerfoglio, S., Capodaglio, P., Rossi, P., Conforti, I., D'Angeli, V., Milani, E., Galli, M., & Cimolin, V. (2023). Evaluation of Upper Body and Lower Limbs Kinematics through an IMU-Based Medical System: A Comparative Study with the Optoelectronic System. *Sensors (Basel, Switzerland)*, *23*(13), 6156.

D'Alcala', E. R., Voerman, J. A., Konrath, J. M., & Vydhyanathan, A. (2021). *Xsens DOT Wearable Sensor Platform White Paper*.

Didaskalou, S., Papageorgiou, P., Tzatzimaki, K., Liapi, M., Portokallidis, N., Drosatos, G., & Kaldoudi, E. (2025). *ThrombUS+ Study B2 Lower Limb Activity Data Collection Manual*.

García-de-Villa, S., Jiménez-Martín, A., & García-Domínguez, J. J. (2022). A database of physical therapy exercises with variability of execution collected by wearable sensors. *Scientific Data*, *9*(1), 266.

González-Alonso, J., Oviedo-Pastor, D., Aguado, H. J., Díaz-Pernas, F. J., González-Ortega, D., & Martínez-Zarzuela, M. (2021). Custom IMU-Based Wearable System for Robust 2.4 GHz Wireless Human Body Parts Orientation Tracking and 3D Movement Visualization on an Avatar. *Sensors*, *21*(19), 6642.

Gu, C., Lin, W., He, X., Zhang, L., & Zhang, M. (2023). IMU-based motion capture system for rehabilitation applications: A systematic review. *Biomimetic Intelligence and Robotics*, *3*(2), 100097.

Hansen, C., Chebil, B., Cockroft, J., Bianchini, E., Romijnders, R., & Maetzler, W. (2023). Changes in Coordination and Its Variability with an Increase in Functional Performance of the Lower Extremities. *Biosensors*, *13*(2), 156.

Lam, W. W. T., Tang, Y. M., & Fong, K. N. K. (2023). A systematic review of the applications of markerless motion capture (MMC) technology for clinical measurement in rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, *20*, 57.

Marsan, T., Clauzade, S., Zhang, X., Grandin, N., Urman, T., Linton, E., Elsayed-Aly, I., Ricciardi, C. E., & Temporelli, R. (2025). REEV SENSE IMUs for Gait Analysis in Stroke: A Clinical Study on Lower Limb Kinematics. *Sensors (Basel, Switzerland)*, *25*(16), 5123.

Martínez-Zarzuela, M., González-Alonso, J., Antón-Rodríguez, M., Díaz-Pernas, F. J., Müller, H., & Simón-Martínez, C. (2023). Multimodal video and IMU kinematic dataset on daily life activities using affordable devices. *Scientific Data*, *10*(1), 648.

Niswander, W., Wang, W., & Kontson, K. (2020). Optimization of IMU Sensor Placement for the Measurement of Lower Limb Joint Kinematics. *Sensors (Basel, Switzerland)*, *20*(21), 5993.

Porciuncula, F., Roto, A. V., Kumar, D., Davis, I., Roy, S., Walsh, C. J., & Awad, L. N. (2018). Wearable Movement Sensors for Rehabilitation: A Focused Review of Technological and Clinical Advances. *PM & R: The Journal of Injury, Function, and Rehabilitation*, *10*(9 Suppl 2), S220–S232. https://doi.org/10.1016/j.pmrj.2018.06.013

Salisu, S., Ruhaiyem, N. I. R., Eisa, T. A. E., Nasser, M., Saeed, F., & Younis, H. A. (2023). Motion Capture Technologies for Ergonomics: A Systematic Literature Review. *Diagnostics*, *13*(15), 2593.

Sun, Y., Song, Z., Mo, L., Li, B., Liang, F., Yin, M., & Wang, D. (2025). IMU-Based quantitative assessment of stroke from gait. *Scientific Reports*, *15*(1), 9541.

Sun, Z., Ke, Q., Rahmani, H., Bennamoun, M., Wang, G., & Liu, J. (2023). Human Action Recognition From Various Data Modalities: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(3), 3200–3225.

Zhu, Y., Li, H., Wu, X., & Chen, N. (2024). Accuracy Validation of a Sensor-Based Inertial Measurement Unit and Motion Capture System for Assessment of Lower Limb Muscle Strength in Older Adults—A Novel and Convenient Measurement Approach. *Sensors*, *24*(18), 6040.