Low-Cost Tool for Human Gait Analysis Based on Computer Vision and Inertial Sensors

Eduardo Varela-Martínez[1], Misael Avila-López[1], Fernanda Nicole Gómez-Martínez[1], Sabi Serrato-Navarrete[2], Luz Jiménez-Terrazas[1], Alejandro Hernandez-Calixto[3], Rodrigo Hernández-Peña[4], and José Eduardo Chairez-Veloz[5]

[1] BSc Program on Biomedical Physics, Physics Department, Universidad Nacional Autónoma de México, Mexico City, Mexico

[2] BSc Program on Physics, Physics Department, Universidad Nacional Autónoma de México, Mexico City, Mexico

[3] BSc Program on Mechatronics Engineering, Universidad del Valle de México, Mexico City, Mexico

[4] Physics Department, Biomedical Physics, Universidad Nacional Autónoma de México, Mexico City, Mexico

jechairez@ciencias.unam.mx

Abstract. Accurate assessment of human locomotion is essential for early detection of movement disorders and effective rehabilitation, yet traditional methods are often imprecise and resource-intensive. We introduce Nemi, a system that integrates Inertial Measurement Units (IMUs) and computer vision for automated gait analysis. Using Bluetooth-IMUs and video footage of lower body movements, the system measures Range of Motion (RoM), velocity, step length, gait cycle and cadence. Convolutional neural networks extract lower body positions from video frames, enabling a secondary RoM estimation. A Support Vector Machine (SVM) model, trained on $\sim 2,300$ images of 8 healthy adults (20-30 years), classifies gait cycles stages with F1 scores ranging from 0.89 ± 0.01 to 0.96 ± 0.01 across all phases, and overall accuracy of 0.936 ± 0.001. Integrating both estimations refines RoMs values and minimizes uncertainty. Finally, results are displayed in a user-friendly graphical interface for clinical use.

Keywords: SVM model · computer vision · IMU · human gait

1 Introduction

Walking, though seemingly simple, requires precise coordination of multiple body systems. Variables such as speed, cadence, stride length, and joint mobility are fundamental for functional movement. Human Gait (HG) refers to the bipedal locomotion pattern involving coordinated and alternating lower limb motions.

The HG cycle is divided into support and swing phases. The support phase includes: *Initial Contact* (IC), when the heel makes first contact with the ground, *Load Response* (LR), as the foot progresses to full contact, *Mid-Stance* (MS), which begins when the body weight is propelled forward directly over the lower limb, *Terminal Stance* (TS), which involves lifting the heel off the ground and weight begins to transfer to the opposite leg, and *Pre-Swing* (PSw), when toes push off the floor. The swing phase comprises: *Initial Swing* (ISw), with ankle dorsiflexion and hip flexion, *Mid-Swing* (MSw), as the non-weight-bearing leg passes under the body, and finally *Terminal Swing* (TSw), when the leg prepares for the next heel strike.

Gait analysis, the study of this cycle, is essential for understanding locomotion and its neurological basis. Recent studies have linked gait irregularities to motor neuron damage in both developmental and degenerative disorders. For instance, Jarchi and collaborators in [6] highlighted accelerometry's clinical relevance, while Kikkert [7] found strong associations between gait and cognitive decline. Such findings confirm HG analysis as a valuable tool for diagnosis, prognosis and treatment planning [11]. Nevertheless, conventional gait assessment methods are often time-consuming, costly and lack precision [4]. These limitations can hinder early diagnosis and management of neurodegenerative diseases. For this reason, Bauchet et al. [1] underline the need for standardized and reproducible gait analysis to improve clinical outcomes.

Technologies such as Inertial Measurement Units (IMUs) and computer vision offer promising alternatives, each with distinct strengths and limitations. IMUs provide high-frequency kinematic data and portability, but are also susceptible to drift and noise. In contrast, computer vision systems excel in extracting spatial information under controlled environments but may lack portability. Recent studies propose fusing IMUs and vision data to combine their complementary advantages and thus mitigate individual shortcomings [5, 9, 12]. Despite these efforts, hybrid systems are still very scarce. In this direction, we introduce Nemi, an integrated system combining MediaPipe-based computer vision and IMUs for real-time HG analysis. We hypothesize that merging these technologies improves accuracy and reduces cost and complexity. Our goal is to develop and validate Nemi as an accessible tool for clinical HG assessment.

2 Methods

2.1 Gait analysis based on computer vision

Pose estimation algorithms offer a fast and markerless alternative for analyzing HG, compared to traditional methods such as goniometry or stereo vision. Comparative studies [3] have shown that computer vision algorithms like MediaPipe achieve results comparable to those gold-standard systems such as Kinect V2. These models can be evaluated using F1-scores, a metric that assesses the systems precision in identifying positive cases relative to negatives.

We developed an algorithm using MediaPipe Pose to compute body part segmentation and keypoint estimation from images, which were captured with a wide-angle 12 MP camera (f/1.8 aperture) placed 3.5 meters from the subject. Based on vector space geometry, three adjacent keypoints (e.g., hip, knee, ankle) defined two vectors, and the angle θ between them was computed as:

$$\theta = \arccos\left(\frac{\mathbf{v_1} \cdot \mathbf{v_2}}{\|\mathbf{v_1}\|\|\mathbf{v_2}\|}\right)$$

where $\mathbf{v_1}$ and $\mathbf{v_2}$ are the limb segments in vector form. This equation was applied to compute lower limb joint angles from the extracted coordinates.

To validate the estimated joint RoM, the algorithm-derived angles were compared to goniometer measurements during static poses, supervised by a trained specialist. For this purpose, a database of images from 8 university volunteers (aged 2030 years) performing four flexion exercises, with knee and hip angles simultaneously recorded by a goniometer as ground truth.

After validation, the algorithm analyzed 244 fps videos of people completing a walking cycle, extracting 2,300 frames classified into IC, LR, MS, TS, PSw, ISw, MSw and TSw stages. For each frame, the flexion angles and positions of the hip, knee, and ankle were recorded to train a Support Vector Machine (SVM) model for automatic gait phase detection. The model performance was evaluated using the F1-score, which combines the harmonic mean of precision and recall.

The MediaPipe Pose model was integrated to improve skeleton tracking visualization. In this implementation, body landmarks detected by the algorithm are assigned to lower limb joints, displaying their RoM in consistent colors.

2.2 Estimation of biomechanical parameters using IMUs

WT901BLE inertial sensors were used to estimate biomechanical parameters, including speed, stride length, and cadence. These devices integrate an accelerometer (±16g), gyroscope (±2000°/s), and three-axis magnetometer, providing linear acceleration, angular velocity, and orientation data with 0.05° accuracy on the XY axes. An Unscented Kalman Filter (UKF) optimized data fusion to enhance reliability during non-linear walking movements. The following describes the system architecture and procedures for data acquisition, sensor connectivity, calibration, and data processing.

The acquisition system uses the Bleak library, enabling Bluetooth 5.0 communication with WT901BLE sensors at 100 Hz sampling rate. Sensors were placed on the right and left Anterior Superior Iliac Spine (ASIS) (see Fig. 1) to enhance gait phase analysis. The DeviceModel class synchronizes data from individual sensors.

Each sensor is recognized by its MAC address and assigned to a DeviceModel instance based on ASIS placement. The UUID 0000ffe5-0000-1000-8000-00805f9a34fb GATT service enables data reception and calibration commands. Data capture can be also manually controlled via a parallel command.

IMU calibration involved sequentially sending three byte arrays ([0xFF, 0xAA, 0x01, 0x00], [0xFF, 0xAA, 0x02, 0x00], [0xFF, 0xAA, 0x03, 0x00]), with 1-second intervals between commands. This process adjusted the zero points of the accelerometer, gyroscope, and magnetometer axes to correct drift and ensure measurement accuracy for linear acceleration, angular velocity and orientation. To

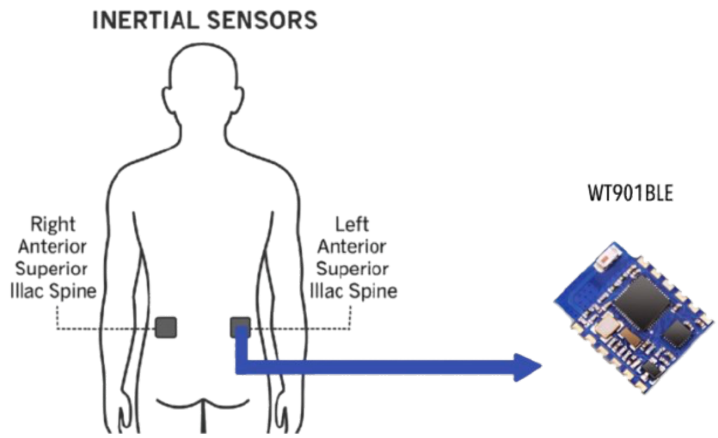

Fig. 1. Placement of inertial sensors WT901BLE on the right and left ASIS. The sensors were secured to the waist using a custom-designed 3D printed module with adjustable straps to ensure proper alignment and stability.

verify data reliability, a Bland-Altman analysis compared 600 data points on the X, Y, and Z axes between the developed calibration and the manufacturer's (WitMotion) calibration. This confirmed agreement between both methods and consistency with our approach.

Raw data, transmitted in 11-byte packets at 100 Hz, was processed in real time using the `on_data_received` method, which accumulates bytes and extracts subpackets for analysis. According to the WitMotion, each packet includes a header byte (0x55), frame identifiers (0x51 for acceleration, 0x52 for angular velocity and 0x53 for Euler angles), sensor data and checksums. Data were extracted as follows: accelerations (a_x, a_y, a_z) from bytes 27 in 0x51 frames, angular velocities (g_x, g_y, g_z) from 0x52 frames, and Euler angles (ang_x, ang_y, ang_z) from 0x53 frames. Values were converted to physical units using scale factors (16g, 2000°/s, 180°).

UKF fused measurements to enhance reliability during non-linear walking movements. It used sigma points ($\alpha = 0.1, \beta = 2.0, \kappa = 0.0$) to estimate a 7-dimensional state (linear velocities $[v_x, v_y, v_z]$ and quaternions $[q_x, q_y, q_z, q_w]$), with process and measurement noise covariance $\mathbf{Q} = 0.0\mathbf{I}$ and $\mathbf{R} = 100\mathbf{I}$. The transition function integrated linear accelerations into the global frame (via a rotation matrix derived from quaternions) and updated quaternions with angular velocities, applying normalization to ensure unitarity. The observation function projected gravitational accelerations onto the sensor frame to reduce noise and drift.

After data processing, step detection was performed by monitoring ang_z, which reflects pelvic rotation in the transverse plane, a biomechanical indicator of step initiation during alternating leg movements. A threshold of $\text{ang}_z > 20°$ and a minimum interval of 0.3 seconds between steps were applied to avoid false positives, increment a step counter, and record timestamps. The Z-axis was chosen because its rotation in the transverse plane is less affected by external movements than the X or Y axes, ensuring robust step detection.

Cadence was calculated as:

$$\text{Cadence} = \frac{\text{steps}}{\text{total time}} \cdot 60$$

Speed was estimated as the average magnitude of the UKF-derived linear velocities:

$$\text{Speed} = \sqrt{v_x^2 + v_y^2 + v_z^2}$$

Step length was computed as:

$$\text{Step Length} = \frac{\text{Speed}}{\text{Step Frequency}}$$

refined by UKF predictions. Processed data were stored in a CSV file containing biomechanical parameters.

2.3 API architecture and interface development

Integrating computer vision and IMUs for gait analysis requires a robust software architecture capable of synchronizing multiple data sources in real-time and presenting information clearly to the user. To achieve this, we designed an Application Programming Interface (API), which enables communication between hardware and software modules, as well as an interactive graphical interface in Unity to support accurate acquisition, processing, fusion and visualization of kinematic data (see Fig. 2).

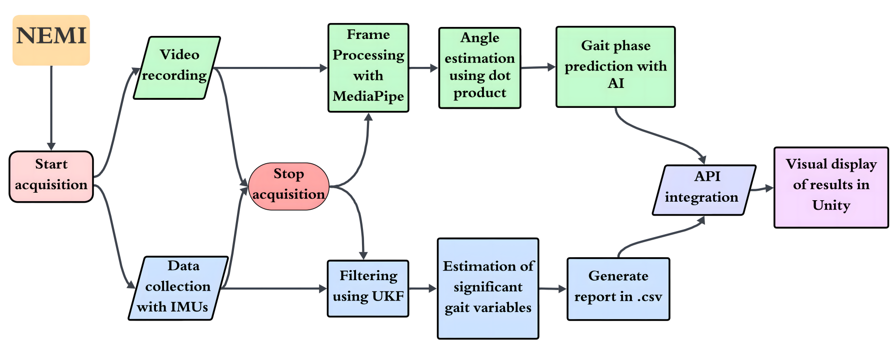

Fig. 2. Workflow of Nemi integrating video and IMU data for gait analysis, AI classification, and visualization in Unity.

The architecture was implemented using the FastAPI framework, enabling real-time handling of multiple asynchronous requests across modules. In the computer vision module, the MediaPipe Pose algorithm was adapted to structure the extracted data in JSON format and send it to the API via POST requests with timestamps for synchronization. The IMU module incorporated custom Python code for calibration, filtering, and data processing. The API exposes endpoints to start, stop, store data, and retrieve computed gait variables. Acting as a back-end, it integrates with a Unity-based graphical interface that provides real-time visualization of computed parameters and skeleton tracking, ensuring seamless interaction between the computer vision and IMU modules.

3 Results

3.1 Validation of the computer vision algorithm

The images were validated against measurements obtained using a manual goniometer willing an absolute error of 6.44%. Furthermore, as shown in Fig. 3, the SVM model achieved consistent F1 scores across all phases, with a minimum score of 0.89 in MSw and a maximum score of 0.96 in PSw, and an overall accuracy of 0.94 in the other phases.

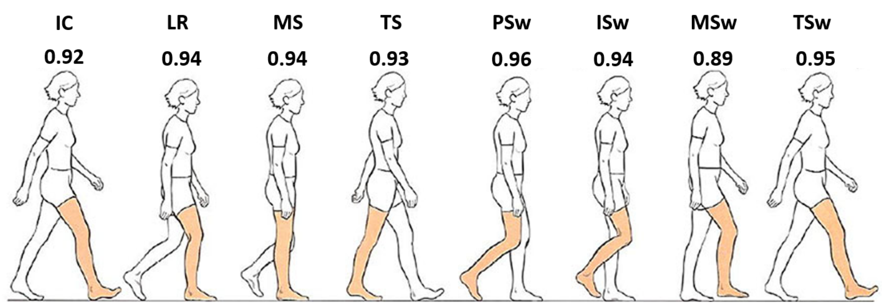

Fig. 3. Performance of the gait phase classifier, showing F1-scores across the eight phases of the walking cycle.

Additionally, Fig. 4 shows the visual output of the gait analysis algorithm during the IC phase, showing the lower limb landmarks and the RoM values. Futhermore, the current gait phase detected by the SVM model is displayed in the upper left corner of the screen.

3.2 Validation of IMU

The Bland-Altman analysis revealed a systematic bias between the two calibration methods, however, it remained stable across all axes with low variability. For the X-axis, the mean bias was −0.6207 with a standard deviation of 0.0018 and limits of concordance between −0.6296 and −0.6118. For the Y-axis, the

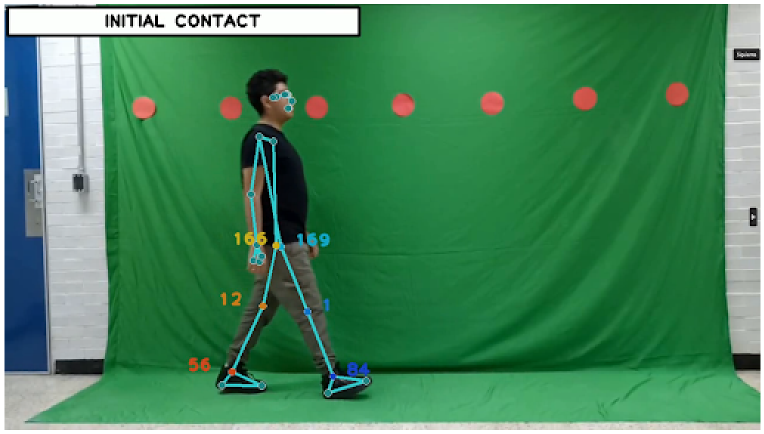

Fig. 4. MediaPipe Pose-based visualization of the IC phase. Lower limb joints are highlighted with colored markers, and real-time RoM values are shown for each segment. The SVM model prediction of the current gait phase appears in the upper-left corner.

bias was 0.6427, with a standard deviation of 0.0044 and limits between 0.6209 and 0.6645. Finally, for the Z-axis, the bias was −2.4017, with a standard deviation of 0.0018 and limits between −2.4106 and −2.3929, confirming consistence between both calibration approaches

The Z-axis angle data from the IMUs after were filtered using an UKF to reduce noise and improve signal stability. Figure 5 shows trajectories for the left and right ASIS during gait, illustrating the alternating pelvic rotation pattern of human walking. This filtered signal enabled the algorithm to compute key gait parameters, later integrated with computer vision data to support phase classification and thus provide a comprehensive analysis of gait dynamics.

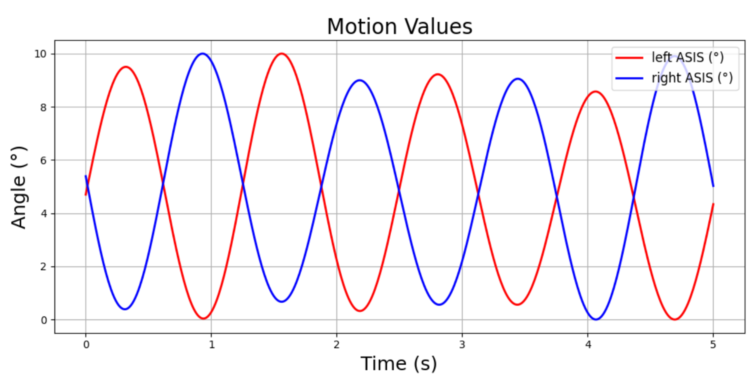

Fig. 5. Z-axis angle trajectories of the left and right anterior superior iliac spines (ASIS) after UKF filtering, illustrating pelvic rotation during walking.

3.3 Integration and visualization

The Nemi system includes a Unity-based graphical interface designed for usability and intuitive visualization. The home screen (Fig. 6a) offers two analysis modes: computer vision and IMU sensors. Each mode supports real-time recording, video upload, and simulation. In the computer vision module (Fig. 6b), the original video and the analyzed video with MediaPipe Pose landmarks are displayed side by side for direct comparison. The sensor module (Fig. 6c) allows users to control data acquisition and view gait parameterssuch as steps, cadence, speed, and stride length, in a dynamic table.

Fig. 6. Unity-based interface of the Nemi gait analysis system. a) Home screen showing analysis mode options, b) Computer vision module with side-by-side video comparison, and c) Sensor module displaying real-time gait metrics.

4 Discussion

MediaPipe Poses validation compares hip and knee angle estimations in static poses with goniometer measurements, yielding a 6.4% mean absolute error consistent with the findings of Latreche in [8] and Dill et al. [4], who reported mean absolute errors ranging from 5% to 10% in physical exercises, confirming clinical suitability. The SVM model, trained on 2,300 frames labeled across eight gait phases, achieved an overall accuracy of 0.94, with F1 scores ranging from 0.89 to 0.96. This approach was chosen for efficiency and robustness with small datasets, outperforming complex neural networks, though kernel sensitivity may limit scalability to pathological gaits.

Calibration of WT901BLE inertial sensors at the ASIS ensured reliable joint measurements. The comparison under controlled conditions showed high clinical agreement, with a mean bias $< 1°$ and narrow limits of agreement of $\pm 3°$ in orientation. These results are consistent with Benedetti et al. in [2], who emphasized the clinical appropriateness of such calibration thresholds. Calibration errors ranged from 5% to 10% for manual inertial sensor calibration, with limits of agreement around $\pm 4°$ which are acceptable for biomechanical applications. Note that the UKF provided stable velocity estimates, mitigating drift and supporting sensor calibration reliability. The current implementation employs standard webcams, which significantly reduces the cost of the system and enhances its potential for widespread adoption in resource-limited settings. Although this study was limited to healthy young adults, a further validation in patients with gait impairments is required to confirm clinical applicability.

5 Conclusion

Nemi offers an objective, efficient, and low-cost approach to human gait analysis, achieving 94% accuracy. Its intuitive interface supports clinical diagnostics and rehabilitation monitoring. Additionally, integrating this information into a Unity API provides an intuitive and accessible interface with clear visualizations for both specialists and patients. This facilitates feedback and monitoring of therapeutic progress. Future work will focus on evaluating its effectiveness in patients with specific pathologies or human gait disorders in order to validate its diagnostic capabilities in collaboration with specialists in clinical settings.

Acknowledgments

Authors acknowledge the support from UNAM-DGAPA PAPIIT IN115825.

References

1. Beauchet, O., Allali, G., et al. Guidelines for Assessment of Gait and Reference Values for Spatiotemporal Gait Parameters in Older Adults: The Biomathics and Canadian Gait Consortiums Initiative. Frontiers In Human Neuroscience. 11, (2017).
2. Benedetti, M.G., Beghi, E., et al. SIAMOC position paper on gait analysis in clinical practice: General requirements, methods and appropriateness. Results of an Italian consensus conference. Gait & Posture. 58, 252-260 (2017).
3. De Gusmão Lafayette, T.B., De Lima Kunst, V.H., et al. Validation of Angle Estimation Based on Body Tracking Data from RGB-D and RGB Cameras for Biomechanical Assessment. Sensors. 23, 3 (2022).
4. Dill, S., Rösch, A., et al. Accuracy Evaluation of 3D Pose Estimation with MediaPipe Pose for Physical Exercises. Current Directions In Biomedical Engineering. 9, 563-566 (2023).
5. Hellstén, T., Arokoski, J., et al. Reliability and validity of computer vision based markerless human pose estimation for measuring hip and knee range of motion. Healthcare Technology Letters. 12, (2025).

6. Jarchi, D., Pope, J., et al. A Review on Accelerometry-Based Gait Analysis and Emerging Clinical Applications. IEEE Reviews In Biomedical Engineering. 11, 177-194 (2018).
7. Kikkert, L.H.J., Vuillerme, N., et al. Walking ability to predict future cognitive decline in old adults: A scoping review. Ageing Research Reviews. 27, 1-14 (2016).
8. Latreche, A., Kelaiaia, R., Chemori, A., Kerboua, A.: Reliability and validity analysis of MediaPipe-based measurement system for some human rehabilitation motions. Measurement. 214, 112826 (2023).
9. Li, T., Yu, H. Visual Inertial Fusion-Based Human Pose Estimation: A review. IEEE Transactions On Instrumentation And Measurement. 72, 1-16 (2023).
10. Patterson, K.K., Nadkarni, N.K., et al.Gait symmetry and velocity differ in their relationship to age. Gait and Posture. 35, 590-594 (2012)
11. Pearl, O., Shin, S., et al. Fusion of video and inertial sensing data via dynamic optimization of a biomechanical model. Journal of Biomechanics, 155, 111617 (2023).