

Article

Not peer-reviewed version

---

# Improved Step Length Estimation with IMU Based Gait Analysis Using Sequential Neural Network

---

[Tuhin Ghosh](#) , [Aritra Paul](#) , [Pampa Sadhukhan](#) <sup>\*</sup> , Pradip K. Das <sup>†</sup> , Nirmalya Roy <sup>†</sup>

Posted Date: 29 December 2025

doi: 10.20944/preprints202512.2454.v1

Keywords: gait; step length; IMU; machine learning (ML); sequential neural network (SNN)





Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Improved Step Length Estimation with IMU Based Gait Analysis Using Sequential Neural Network

Tuhin Ghosh<sup>1</sup>, Aritra Paul<sup>1</sup>, Pampa Sadhukhan<sup>1,\*</sup> , Pradip K. Das<sup>2,†</sup>  and Nirmalya Roy<sup>3,†</sup> 

<sup>1</sup> School of Mobile Computing & Communication, Jadavpur University, India

<sup>2</sup> Dept. of Computer Sc. & Engineering, Jadavpur University, India

<sup>3</sup> Dept. of Information Systems, University of Maryland, Baltimore County

\* Correspondence: pampas.smcc@jadavpuruniversity.in; Tel.: +9133 2335 6122

† These authors contributed equally to this work.

## Abstract

Precise estimation of step length constitutes a fundamental requirement in contemporary gait analysis, particularly for applications in healthcare monitoring, rehabilitation, and intelligent wearable systems. In contrast to conventional approaches, which often rely on wired, foot-mounted inertial sensors and computationally intensive deep learning architectures, this study presents a wireless, thigh-mounted Inertial Measurement Unit (IMU)-based framework for step length estimation, employing a supervised learning paradigm to enhance accuracy, portability, and practicality in natural walking environments. Using an MPU-6050 IMU interfaced with an ESP32 module, the proposed SLE framework establishes a comprehensive data acquisition pipeline to enable seamless wireless transmission and real-time gait recording. Moreover, noise attenuation via Butterworth filtering and statistical normalization was also applied to refine motion signals. Additionally, fourteen engineered gait features extracted from segmented step events are employed to train four supervised learning algorithms, namely artificial neural networks (ANN), sequential neural networks (SNN), k-Nearest Neighbour (k-NN), and support vector machine (SVM), by this SLE framework. Evaluation of the proposed SLE model under both normal and fast walking conditions using a leave-one-out cross-validation scheme demonstrates SNN's superiority over the other considered supervised models, with an outstanding average accuracy of over 99.4% and a reasonably superior average accuracy of over 83.5% achieved in the wired and wireless environments, respectively, across diverse gaits under both walking modes. With a certain marginal performance degradation, the wireless configuration still underscores its robustness and exhibits its potential for real-time gait monitoring through resource-constrained devices.

**Keywords:** gait; step length; IMU; machine learning (ML); sequential neural network (SNN)

## 1. Introduction

Human gait analysis serves as a vital instrument for comprehending the mechanics of human movement and the intricate interaction between the musculoskeletal and nervous systems (Whittle, 2014) [1]. It facilitates a comprehensive assessment of walking patterns, providing valuable insights into balance, stability, and overall mobility (Baker, 2013; Lee et al., 2023) [2,3]. The study of gait involves a detailed examination of the cyclic motion of the lower limbs, primarily focusing on phases such as stance, swing, and double support, which together constitute the gait cycle (Everett & Kell, 2010) [4]. Each component of this cycle reflects distinct biomechanical and neuromuscular processes, rendering gait analysis an essential field of investigation across healthcare, sports science, and assistive robotics (Donoghue et al., 2021; Kelly et al., 2018; Phinyomark et al., 2015; Tong et al., 2024) [5–8].

Gait analysis plays an indispensable role in assessing human mobility and detecting deviations from normal walking behaviors. Abnormal gait patterns often serve as early indicators of various health disorders, including Parkinson's disease, cerebral palsy, arthritis, and post-stroke impairments (Hausdorff, 2009; Nguyen et al., 2024) [9,10]. Such conditions may alter stride length, cadence, or

limb coordination, thereby compromising balance and stability. In rehabilitation and clinical contexts, precise gait assessment assists in designing patient-specific therapeutic interventions, monitoring recovery, and evaluating the efficacy of assistive devices such as prosthetics and orthotics. Beyond healthcare, gait analysis also holds significant potential in optimizing sports performance and robotics. In athletics, quantitative gait data enable refined analysis of stride dynamics, assisting athletes in enhancing efficiency and reducing injury risk. In robotics and human–robot interaction, gait modeling provides foundational data for the development of humanoid robots capable of replicating natural human locomotion, thereby improving adaptability and control mechanisms in dynamic environments (Collins & Ruina, 2005) [11].

Over recent decades, gait analysis has evolved from optical and force-based systems to more compact and accessible technologies. Traditional laboratory-based methods, such as optical motion capture systems, force plates, and electromyography (EMG), may provide high-precision results. However, these systems are costly, require controlled environments, and often restrict natural movement, limiting their practicality for real-world applications (Cutti et al., 2010) [12]. In contrast, Inertial Measurement Units (IMUs), which integrate accelerometers, gyroscopes, and magnetometers to capture dynamic motion data efficiently, have emerged as a powerful and portable alternative. Their compactness, affordability, and wireless operation capabilities make them suitable for continuous, real-time gait monitoring outside laboratory conditions (Miezal et al., 2016) [13]. Nonetheless, due to the non-linear and multidimensional nature of gait signals, accurate interpretation of IMU data necessitates advanced computational techniques, especially those based on supervised machine learning and artificial intelligence.

In the context of human gait analysis, IMUs demonstrate exceptional potential. When strategically positioned on body segments such as the thigh, waist, or foot, they capture detailed kinematic information across multiple gait cycle phases (stance, swing, and double support). This enables the precise estimation of biomechanical parameters such as step length, step duration, cadence, and walking velocity (Díez et al., 2018) [14]. Compared to conventional laboratory-based motion analysis methods, IMU-based gait assessment offers numerous practical advantages, including cost-effectiveness, portability, ease of deployment, and the flexibility granted by wireless connectivity.

As sensor and computational technologies continue to improve, IMU-based gait tracking is transitioning from controlled laboratory environments to portable, real-world, and wireless systems. This transition facilitates the development of intelligent, wearable solutions for continuous motion monitoring, early detection of abnormalities, and personalized gait rehabilitation. Nonetheless, current approaches face limitations that impede widespread practical implementation. Most existing systems rely on wired, foot-mounted IMUs, which restrict natural movement, diminish user comfort, and are susceptible to impact-induced noise during ground contact [15–18]. These constraints render them unsuitable for continuous or real-world gait monitoring. Additionally, many recent solutions depend on computationally intensive deep learning models requiring substantial processing resources and training time [19–22]. Such methods limit applicability in wearable and portable systems where real-time performance and energy efficiency are paramount.

Unlike prior approaches that utilized wired foot-mounted IMUs, this study introduces a wireless, thigh-mounted IMU system for step length estimation employing supervised learning algorithms. Mounting the sensor on the thigh offers several advantages: it reduces signal noise caused by foot impact, enhances comfort during extended use, and captures a broader range of motion data pertinent to the swing and stance phases of the gait cycle. Moreover, the wireless configuration eliminates physical restrictions, enabling natural movement and facilitating remote or continuous gait monitoring—an essential feature for contemporary wearable health systems.

Conversely, supervised learning methods effectively model complex relationships within IMU signals while maintaining lower computational demands compared to deep learning models. Their capacity to learn gait patterns efficiently makes them well-suited for real-time step length estimation in wearable environments, ensuring accuracy without excessive resource consumption (Zhou & Hu, 2008)

[23]. Consequently, this study aims to develop and evaluate a supervised learning-based step length estimation (SLE) model utilizing data from a wireless thigh-mounted IMU sensor. By leveraging four supervised learning models—k-nearest neighbors (k-NN), support vector machine (SVM), artificial neural network (ANN), and sequential neural network (SNN)—this research endeavors to achieve robust and precise step length estimation across various walking speeds. The proposed system seeks to bridge the gap between laboratory-grade gait analysis and practical, wearable solutions, providing a cost-effective, portable, and intelligent framework for continuous gait monitoring.

The organization of our article is as follows. The literature review in the current domain is carried out in Section 2. Section 3 presents the proposed methodology. The performance evaluation, as well as comparative analysis of the proposed SLE method, is conducted in Section 4. In Section 5, we draw our conclusions and then discuss the scope of further research in this arena.

## 2. Literature Review

Over the years, numerous research studies focusing on human gait analysis alongside step or stride length estimation have been proposed in the literature. At first, the work on gait analysis was elaborated, followed by a literature review on step or stride length estimation techniques.

### 2.1. Human Gait Analysis

Human gait is a highly coordinated and cyclic process involving complex neuromuscular interactions. The gait cycle, defined as the sequence of motions between two successive heel strikes of the same leg, serves as a fundamental basis for gait analysis and step length estimation. Various researchers have explored the segmentation of the gait cycle to enhance the precision of motion analysis and parameter estimation. According to Willemsen et al. (1990) [24], human bipedal locomotion can be divided into four distinct phases: stance, heel-off, swing, and heel-strike. During the stance phase, the leg remains in contact with the ground, generating minimal sensor variation due to stable movement. The heel-off phase initiates the transition toward the swing phase, in which the leg moves forward, followed by heel-strike, marking ground contact once again and completing the gait cycle. However, Everett and Kell (2010) [3] proposed a simplified two-phase model—comprising the stance (or stand) and swing (or sway) phases—offering a broader yet effective categorization of gait for practical applications. Such simplifications aid in developing computationally efficient gait detection algorithms while maintaining reasonable accuracy.

In sensor-based gait recognition, several threshold-based detection techniques have been employed to identify these gait phases. For instance, acceleration thresholds utilized [25] from foot-mounted IMUs to isolate stance phases due to their low fluctuation patterns. Other studies have applied angular velocity thresholds [26,27], acceleration signal thresholds [28], and magnetic field variations [29] to achieve more robust phase segmentation. These methods, while effective, can be sensitive to noise and require careful calibration depending on sensor placement and walking speed. In contrast to prior works employing foot-mounted sensors, recent studies have explored alternative placements such as the thigh or waist, which capture broader motion dynamics of the gait cycle. Thigh-mounted sensors offer more consistent data during both stance and swing phases due to reduced impact noise, thereby enhancing gait event detection in wireless and wearable systems.

### 2.2. Step or Stride Length Estimation

Step length and stride length are key gait parameters reflecting an individual's walking rhythm and stability. As defined by Godha et al. (2006) [30], stride length is the distance between two successive heel strikes of the same foot, whereas step length represents half that distance. Reliable estimation of these parameters is essential for applications in gait rehabilitation, sports performance analysis, and indoor navigation systems. Alongside machine learning methods and/or biomechanical models, traditional signal processing approaches were also employed for their estimation, as discussed below in classified form.

### 2.2.1. Integration-Based Methods

One of the most widely adopted methods for stride length estimation involves the double integration of acceleration data from IMU sensors. This approach computes velocity and displacement by integrating acceleration over time, thus determining step length. Studies by Bamberg et al. (2008) [31], Rampp et al. (2015) [32], and Truong et al. (2016) [33] demonstrated that numerical integration methods such as the trapezoidal rule or Simpson's method can yield accurate results when combined with proper error compensation. However, these techniques are often prone to drift errors caused by sensor noise and require sophisticated filtering techniques, such as the Zero Velocity Update (ZUPT) or Kalman filtering, to maintain accuracy.

### 2.2.2. Frequency and Acceleration-Based Estimation

Research revealed that stride length correlates strongly with walking frequency, speed, and body dynamics [34,35]. Building on this, it was also developed that a linear relationship exists between step frequency and stride length, simplifying the estimation process [35]. Although such models are computationally efficient, they often fail to capture individual gait variability influenced by leg length, terrain, and walking style [34]. Recent advancements in supervised learning methods now address this limitation by allowing data-driven models to learn these variations directly from sensor inputs [36].

### 2.2.3. Machine Learning Based Estimation

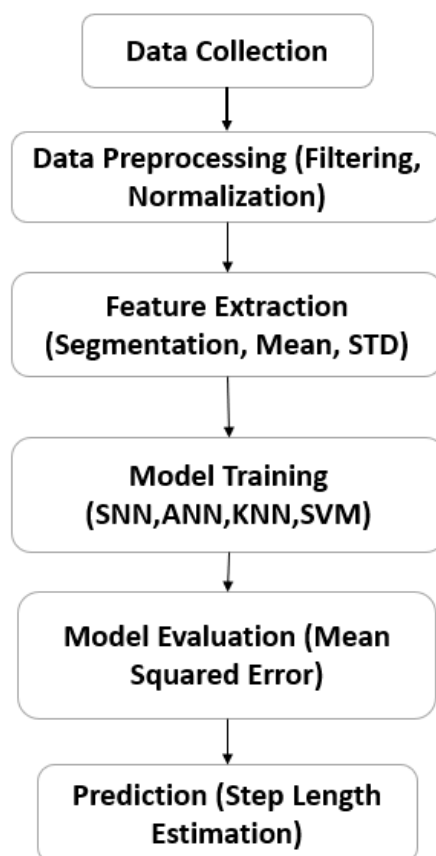
In recent years, several machine learning methods, including deep learning, have been actively explored for IMU-based step and stride length estimation in wearable gait analysis research. Zadka et al. proposed a single lower-back IMU system combined with machine-learning regression to estimate step length in both healthy individuals and patients with neurological disorders [36]. Their study demonstrated promising accuracy and strong clinical relevance, supported by a large participant cohort. However, the system was primarily evaluated in controlled walking environments, with limited consideration for naturalistic free-living scenarios, turning movements, or varied terrains. Furthermore, although a single-sensor configuration simplifies deployment, lower-back placement remains sensitive to posture changes and may not fully capture segment-level gait dynamics. In contrast, Sui and Chang (2022) [37] introduced a deep learning framework incorporating self-supervised pre-training followed by fine-tuning for stride length estimation during walking and running. While this approach effectively addressed the challenge of limited labeled dataset and achieved competitive accuracy, the computational demands of deep neural networks present challenges for real-time, low-power wearable devices. Additionally, the study primarily focused on healthy gait and controlled locomotion patterns, limiting its applicability to diverse gait conditions.

Complementing these efforts, [38] developed a lightweight step length estimation model using smartphone-based IMU signals and principal component analysis to extract relevant motion features. Their investigation across multiple sensor mounting locations, including the pelvis and thigh, highlighted the importance of sensor placement in IMU-based gait estimation accuracy. The method provided a computationally efficient alternative to deep learning models; however, the dataset size was limited to healthy individuals in structured environments, and the potential effects of loose sensor attachment and natural movement variability remained insufficiently addressed.

*While recent studies have demonstrated substantial progress in IMU-based step length estimation, existing work reveals persistent limitations related to wired or posture-sensitive sensor configurations, reliance on computationally intensive deep learning architectures, and restricted evaluation under controlled laboratory conditions. These gaps highlight the need for practical systems that support unconstrained movement, maintain low computational overhead, and employ strategically chosen sensor placement to improve signal robustness. In this context, a wireless, thigh-mounted IMU coupled with efficient supervised learning techniques presents a promising direction for achieving accurate and user-friendly step length estimation in real-world settings.*

### 3. Proposed Step Length Estimation Method

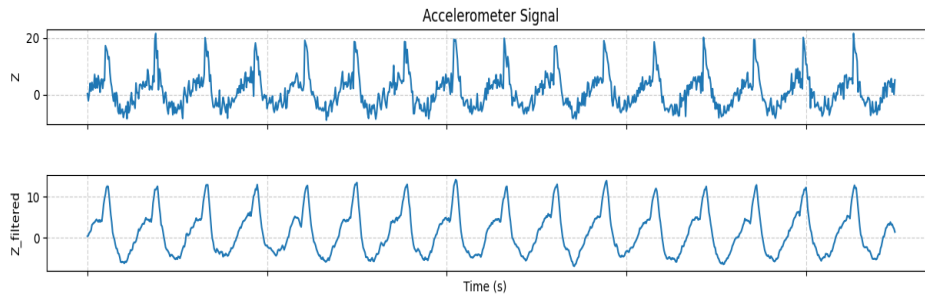
This section presents our proposed step length estimation approach, which leverages a multi-model machine learning framework to process sensor data collected from a single thigh-mounted Inertial Measurement Unit (IMU). The proposed approach incorporates four models—Artificial Neural Network (ANN), Sequential Neural Network (SNN), k-Nearest Neighbors (k-NN), and Support Vector Machine (SVM)—to compare performance and identify the most effective method. All models are trained in a supervised regression setting, aiming to predict step length from features extracted from the IMU data. A simplified flow chart of the methodology is illustrated in Figure 1.



**Figure 1.** Flow-chart of the Proposed Step Length Estimator.

#### 3.1. Data Collection

The IMU sensor module, containing a 3-axis accelerometer and a 3-axis gyroscope, is mounted on a single thigh. Data is recorded wirelessly during walking trials at a sampling frequency of 100 Hz. The coordinate axes are aligned in such a way that the X and Y axes represent movements along the left-to-right and forward-to-backward directions, respectively, whereas the Z-axis represents the vertical movement. Acceleration and angular velocity are recorded along each axis for the duration of the walk. The dataset used in this study contains multiple walking sequences with varying step lengths and cadences. In our proposed approach, the MPU6050 chip is interfaced with the ESP32-WROOM-32D for data recording. This chip combines a three-axis accelerometer and a three-axis gyroscope, allowing precise motion tracking with minimal power usage. The accelerometer detects linear acceleration, whereas the gyroscope measures angular velocity. The site used for data collection and experimentation, as illustrated in Figure 2, is a pedestrian pathway in our University Campus. Such a site is a suitable environment for testing and validating the proposed system, enabling performance evaluation under real-world conditions.



**Figure 2.** Filtered acceleration data.

### 3.2. Data Preprocessing

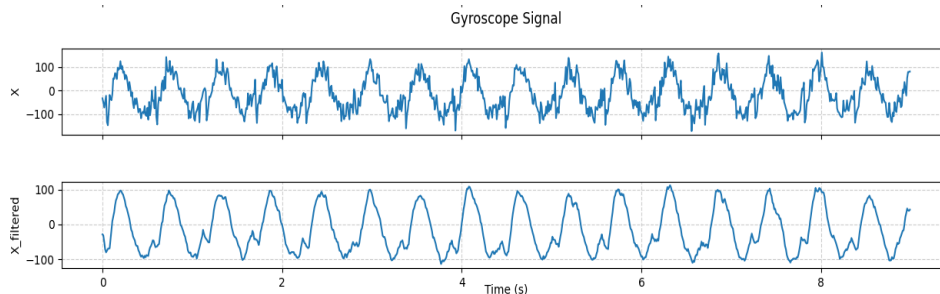
The raw data obtained from the inertial measurement unit (IMU) sensors required preprocessing to remove noise and scale the features appropriately before being used for model training. This crucial stage ensures that the subsequent feature extraction and machine learning models operate on a clean and standardized dataset. The preprocessing pipeline consisted of two main steps: low-pass filtering and normalization.

#### 3.2.1. Low-Pass Filtering

To remove high-frequency noise, a 5th-order Butterworth low-pass filter [39] with a cutoff frequency of 10 Hz is applied separately to each accelerometer and gyroscope axis. denoted as  $\omega_c$ , is computed as:

$$\omega_c = \frac{f_c}{0.5 \times f_s} \quad (1)$$

where  $f_c$  is the cutoff frequency and  $f_s$  is the sampling frequency. Filtering smooths the raw sensor data while retaining essential gait dynamics. A smoothing filter was applied to the raw Z-axis acceleration and X-axis gyroscope signals. The resulting filtered data are presented in Figures 2 and 3.



**Figure 3.** Filtered angular data.

#### 3.2.2. Normalization

After filtering, features are normalized to zero mean and unit variance using the standard scaler from scikit-learn [40]. The normalized yield, denoted as  $x'$  of a time series data  $x$ , is obtained as follows.

$$x' = \frac{x - \mu}{\sigma} \quad (2)$$

where  $\mu$  and  $\sigma$  are the mean and standard deviation of the data  $x$ . The normalization process ensures that all features are weighted equally during model training.

### 3.3. Feature Extraction

A peak detection algorithm was applied to detect the step boundaries, based on which the continuous data were segmented into a set of steps. The same set of fourteen features, which were effectively used in our proposed SNN-based prior SLE model based on the foot-mounted IMU data [18], was considered as a potential feature set to be extracted from each segmented step in this work.

The mean and standard deviation of both the acceleration and angular velocity components along their three axes,  $X$ ,  $Y$ , and  $Z$ , respectively, constitute 12 features, and the other two are step duration and step frequency. Their detailed definitions are provided in our earlier work [18]. Thus, the computed feature matrix  $F$ , which is a collection of all feature vectors, each one corresponding to a single step segment, is mathematically formulated as

$$F = \begin{bmatrix} f_1^1 & f_2^1 & \cdots & f_{13}^1 & f_{14}^1 \\ f_1^2 & f_2^2 & \cdots & f_{13}^2 & f_{14}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ f_1^k & f_2^k & \cdots & f_{13}^k & f_{14}^k \end{bmatrix} \quad (3)$$

where  $k$  is the number of step segments.

### 3.4. Model Training

Our proposed SLE model considers four distinct supervised learning algorithms, which are k-nearest neighbors (k-NN), support vector machine, artificial neural network, and the sequential neural networks (SNNs) to regress the step length. The working methodology alongside the training procedure of each considered supervised learning algorithm is provided in a separate subsection.

#### 3.4.1. k-Nearest Neighbors (k-NN)

The k-NN is a simple yet effective non-parametric machine learning method used for both classification and regression tasks. Unlike other ML models involving a dedicated training phase, k-NN operates on a lazy learning principle, i.e., not explicitly learning from the data during training [41]. Based on a distance metric, like Euclidean distance, it calculates the distance of all data points in the training set from the query point to identify the 'k' nearest or most similar data points to the query point, and their target values are used to predict the output. In regression problems, such as predicting step length, k-NN computes the average of the target values from these nearest neighbors to produce the final output. In our implementation, the number of nearest neighbors ( $k$ ), denoted by the parameter 'n\_neighbors', was set to 3. This approach is intuitive, easy to implement, and often performs well on datasets where similar input features correspond to similar outputs. However, its effectiveness can depend on the choice of 'k' and the scaling of the input features.

#### 3.4.2. Support Vector Machine

A Support Vector Machine (SVM) is not only a robust but also a flexible machine learning algorithm widely used for both classification and regression tasks. In the context of regression, often referred to as Support Vector Regression (SVR), the goal of the model is to find an optimal function that best fits the data while maintaining predictions within a specified margin of tolerance, defined by the epsilon ( $\epsilon$ ) parameter [42]. Unlike many other algorithms that rely on all training data points, SVM focuses only on a subset of critical data points, known as support vectors, which lie closest to the decision boundary (or regression line) and play a key role in defining its position and orientation. Depending on only these influential samples, SVM achieves a high level of efficiency and generalization, even with complex or high-dimensional data. In our implementation, the parameter kernel is set to 'rbf' (which means radial basis function), enabling the model to capture nonlinear relationships between input features and the target variable. This kernel projects the original input data into a higher-dimensional space, where the model can identify more complex patterns that a simple linear function might fail to do. Overall, SVM combines mathematical rigor with computational efficiency, making it a powerful choice for predicting step length and other continuous variables, especially when the underlying relationships in the data are nonlinear and multidimensional.

### 3.4.3. Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) is a computational system inspired by the information processing strategy of the human brain, consisting of a network of interconnected units, i.e., neurons, grouped into layers. These layers work together to learn patterns and relationships within data through mathematical operations and iterative adjustments [43]. Such a network receives information through its input layer, where each neuron corresponds to a specific feature out of the fourteen gait features considered by our proposed SLE framework. At the hidden layer, the data undergoes several transformations as each neuron applies a nonlinear activation function after combining inputs from the previous layer through weighted connections before passing it to the next layer. Through this process, the hidden layers enable the network “learn” abstract representations of the gait features that are useful for predicting the target variable, i.e., step length in this case, at the output layer. The network’s architecture, which includes layers such as Dense(64, ···) and Dense(32, ···), classifies it as a deep neural network because it contains multiple hidden layers. These additional layers allow the model to capture more sophisticated relationships between the input features and the output. During training, the model makes predictions and then computes an error or loss by comparing them to the actual values to tune the internal parameters known as weights and biases, in small steps over time for error reduction. This optimization process is typically guided by algorithms like Adam or stochastic gradient descent (SGD). By repeatedly analyzing examples, the ANN learns to approximate complex relationships within the data, acting as a powerful pattern recognition system.

### 3.4.4. Sequential Neural Network

A Sequential Neural Network (SNN), implemented using Keras’ Sequential API [44], is a linear stack of layers where data flows from the input layer through multiple hidden layers to the output layer. This structure makes it ideal for regression tasks such as predicting step length from gait features. In this model, input data with 14 gait features passes through three hidden layers with 128, 64, and 32 neurons, each one leveraging the ReLU activation function to introduce nonlinearity and capture complex relationships in the data. The output layer, consisting of a single neuron without an activation function, produces a continuous prediction of step length. The detailed architecture of this model was provided in our earlier work [18].

The target variable is scaled using the same scaler applied to the training data to maintain consistency. The model is compiled with the Adam optimizer and trained using the Mean Squared Error (MSE) loss function over 50 epochs with a batch size of 16. A 10% validation split helps monitor the model’s performance and reduce overfitting. After training, the predictions are inverse-transformed to their original scale. The model’s accuracy is evaluated using standard regression metrics such as MSE and root mean square error (RMSE), along with a step-length-based accuracy measure. In summary, this SNN effectively learns and models the nonlinear relationships between gait features and step length, providing accurate and reliable predictions of the latter.

### 3.4.5. Model Evaluation

All the considered models are evaluated using mean squared error (MSE), which is one of the most common ways for measuring the performance of a regression model. It shows, on average, how far the considered model’s predictions are from the actual values and makes larger errors have a higher impact on the result by squaring the differences. Thus, a smaller MSE value indicates better accuracy in the model’s predictions. Mathematically, it is formulated as follows.

$$MSE = \frac{\sqrt{\sum_{i=1}^n (y_i^o - y_i^p)^2}}{n}, \quad (4)$$

where,  $n$  is the total number of data points,  $y_i^o$  and  $y_i^p$  denote the actual (ground truth) step length value and its predicted value from the considered model.

### 3.4.6. Step Length Prediction

Once the trained model is validated, it is then used to estimate the step length of a user based on a new input feature set.

$$\overline{SL}_{pred} = \frac{\sum_{i=1}^n y_i^p}{n}, \quad (5)$$

where  $n$  is the total number of predictions and  $y_i^p$  is the predicted value for  $i^{th}$  observation. The formula computes the arithmetic mean of all predicted values, which is simply the total of all predictions divided by the number of predictions.

## 4. Results and Discussions

The performance of the proposed SLE technique, developed using both the traditional supervised learning models (k-NN and SVM) and deep learning models (ANN and SNN), was rigorously evaluated in terms of accuracy. The results were then compared with our previously proposed wired SLE method, which uses the Peak-Valley detection algorithm and SNN model. The accuracy of the previously and newly proposed SLE methods under each considered ML model is computed as follows.

$$Accuracy(\%) = \left(1 - \frac{|sl^p - sl^o|}{sl^o}\right) \times 100, \quad (6)$$

where  $sl^p$  represents the model-predicted step length, and  $sl^o$  represents the ground truth (i.e., observed) step length measured during the experiment. The formulae for accuracy measure the closeness of the predicted step length to its true value. The smaller the difference between the estimated step length and its ground truth counterpart, the higher the accuracy. An accuracy of 100% means the predicted step length perfectly matches the observed one, while lower percentages indicate greater error. In simple terms, accuracy evaluates how perfectly the model can estimate a person's step length compared to the ground truth value.

### 4.1. Experimental Evaluation of Wired Setup

For experimental validation, a 60 – meter straight walking track with a device mounted on the foot was selected, and data were collected from ten participants aged between 24 and 50 years, including both males and females, with heights from 160 to 180 cm and weights from 40 to 80 kg. To ensure robust performance evaluation and minimize bias, a leave-one-out cross-validation (LOOCV) strategy was adopted for each participant, in which the model was trained with all participants' data except the one used for the test case. Furthermore, the proposed SLE technique was examined under two distinct walking conditions, normal and fast gait speeds, for all participants. This dual-speed assessment was conducted to verify the adaptability and consistency of the proposed method across varying locomotion patterns. Table 1 shows the achieved accuracy of all four ML models considered by our proposed SLE framework, alongside our previously proposed peak-valley detection-based SLE method under normal walking conditions. The same results obtained under the fast-walking condition by our newly proposed ML model-based SLE framework and the previous SLE method are presented in Table 2.

**Table 1.** Accuracy comparison of four ML models and peak-valley detection-based SLE method under normal walking in a wired setup.

Participant No.	Peak-Valley Detection (%)	SNN (%)	ANN (%)	k-NN (%)	SVM (%)
1	92.03	99.61	96.87	96.40	96.17
2	92.27	99.52	90.92	90.34	90.53
3	98.29	99.76	87.89	87.13	88.48
4	97.44	99.55	91.30	93.85	93.72
5	96.54	99.78	98.18	96.19	95.18
6	93.83	99.68	96.49	96.40	97.35
7	98.31	99.46	98.70	98.54	98.71
8	88.35	99.78	92.42	93.45	91.64
9	93.68	99.52	91.60	89.83	91.48
10	98.29	99.67	90.84	90.90	89.93

**Table 2.** Accuracy comparison of four ML models and peak-valley detection-based SLE method under fast walking in a wired setup.

Participant No.	Peak-Valley Detection (%)	SNN (%)	ANN (%)	k-NN (%)	SVM (%)
1	96.62	99.52	84.53	85.38	83.73
2	96.4	99.55	98.39	98.32	98.27
3	93.43	99.46	82.03	82.44	82.91
4	98.37	99.61	87.69	87.83	87.27
5	97.6	99.52	93.66	92.56	93.13
6	92.77	99.56	87.02	84.68	84.21
7	97.79	99.67	92.10	91.75	92.90
8	93.94	99.61	97.34	98.40	97.92
9	96.3	99.46	94.67	92.96	93.09
10	94.78	99.52	93.70	93.41	93.66

The accuracy results reported in Tables 1 and 2 demonstrate that the SNN-based SLE approach outperforms the other three ML models as well as the previously proposed peak-valley detection-based SLE approach under both walking conditions. With a very low average error rate of only 0.5%, the proposed step length estimator employing SNN has an amazing average accuracy of over 99.4%. The findings also show that although the SNN-based method is the most accurate, the combination of ANN, K-NN, and SVM in the SLE framework provides a great trade-off between accuracy and computational economy. Moreover, the evaluation across a set of ten participants with sufficient variability in their ages, heights, and genders reveals the SNN model's strong robustness and resilience across diverse gait patterns, indicating its enormous strength for possible use in actual gait monitoring and step length estimation tasks.

Furthermore, the inconsistency in the achieved accuracy for the ML models k-NN, SVM and ANN across the diverse gait patterns as well as their accuracy drop under the fast walking condition reveals that these three models are affected by both the gait pattern and gate speed, where the performances of the SNN model remain almost uninfluenced by such factors as it achieves a consistent average accuracy above 99.4% across the diverse gait patterns and speed. Another noticeable point is that, like the k-NN, SVM, and ANN, the accuracy of previously proposed peak-valley detection-based SLE methods falls somewhat under the fast-walking condition, which can be attributed to the reduced signal fluctuation and smoother sensor communication at moderate walking speeds of the pedestrians. This phenomenon can be attributed to the fact that with an increase in gait speed, the sensor readings are contaminated with more noise.

#### 4.2. Experimental Evaluation of Wireless Setup

The performances of the four machine learning models – k-NN, SVM, ANN, and SNN -in terms of achieved accuracy were also thoroughly assessed under the wireless setup using a thigh-mounted IMU. For experimental validation, data were collected from five individuals aged 24 to 50 years, including both males and females, with heights from 166 to 180 cm and weights from 53 to 64 kg, during their walk along a 60-meter straight path with an IMU mounted on their thigh. To ensure reliable evaluation and reduce potential bias, a leave-one-out cross-validation (LOOCV) method was also implemented in this case, as performed under the wired setup. In addition, the performance of the proposed SLE framework was tested under two different gait conditions, normal and fast walking speeds, for every participant in this case as well. This comparative evaluation aimed to demonstrate the model's adaptability, accuracy, and stability across varying walking patterns. Table 3 presents the achieved accuracy of all four ML models considered by our proposed SLE framework, evaluated in the wireless setup, under normal walking conditions. The same results obtained under the fast-walking condition by all four ML models in the wireless setup are reported in Table 4.

**Table 3.** Accuracy comparison of the newly proposed four ML models-based SLE method under normal walking in a wireless setup.

Participant No.	SNN (%)	ANN (%)	k-NN (%)	SVM (%)
1	98.13	94.10	97.18	96.74
2	99.09	96.80	98.27	97.66
3	96.64	95.60	94.2	94.23
4	94.92	94.34	93.49	93.03
5	93.66	92.85	92.77	93.28

**Table 4.** Accuracy comparison of the newly proposed four ML models-based SLE method under fast walking in a wireless setup.

Participant No.	SNN (%)	ANN (%)	k-NN (%)	SVM (%)
1	92.54	91.32	91.91	91.39
2	93.54	88.10	89.40	90.42
3	93.78	90.81	91.13	89.71
4	83.58	82.96	82.87	80.89
5	90.73	86.34	87.49	88.28

As shown in Tables 3 and 4, the Sequential Neural Network (SNN) model achieved the highest overall accuracy in both scenarios, averaging above 90% for fast walking and exceeding 95% for normal walking. This demonstrates the robustness and stability of the SNN in handling gait data transmitted wirelessly. The ANN, K-NN, and SVM models also produced commendable results, indicating that wireless data transmission had minimal impact on model precision. Notably, all models performed slightly better during normal walking compared to fast walking, as also observed in the wired setup.

However, the performance of all four models, including the SNN, in terms of achieved accuracy, somewhat degrades in the wireless setup when experimented with a thigh-mounted IMU compared to the same carried under a wired environment with a foot-mounted IMU. Such a phenomenon can be attributed to two reasons. Firstly, noise contamination in wirelessly transmitted sensor readings is more common than in wired settings. Perhaps the second one is that an IMU mounted on the foot probably has greater capability in recognizing gait phases and events than the same affixed to the thigh.

## 5. Conclusions and Future Scope of Research

Based on two traditional supervised learning models (k-NN and SVM) and two deep learning models (ANN and SNN), a step length estimation (SLE) framework is proposed in this study. The

proposed SLE framework is evaluated across diverse gait patterns and two distinct walking speeds (normal and fast) in two different contexts: a wired setup equipped with a foot-mounted IMU and a wireless setup with a thigh-mounted one. The evaluations in both contexts demonstrate their high accuracy, stability, and adaptability across different walking speeds and connectivity modes. In the wired setup, the models achieved strong overall performance, with an average accuracy of 99.56% for the SNN, 90.16% for ANN, 91.37% for k-NN, and 90.03% for SVM during fast walking, while the normal walking trials recorded 99.63% (SNN), 93.03% (ANN), 93.03% (k-NN), and 93.12% (SVM). These results indicate the excellent precision of the proposed models under stable, wired data transmission conditions. Similarly, during wireless connectivity testing, the system maintained competitive accuracy, averaging 92.32% (SNN), 87.91% (ANN), 88.56% (K-NN), and 88.14% (SVM) for fast walking. Under normal walking conditions, the accuracy further improved to 96.49% (SNN), 94.34% (ANN), 95.20% (K-NN), and 94.59% (SVM). The slight reduction in performance during wireless operation can be attributed to minor transmission delays or signal interference, yet the overall accuracy remained above 87% for all models, confirming strong reliability. In summary, the findings validate that the proposed SLE system maintains high accuracy and robustness across both wired and wireless configurations. The marginal accuracy difference is less than 5–7% between the two setups, establishing the framework's capability for real-time, portable gait analysis applications without significant loss of precision.

Future developments of the proposed Step Length Estimation (SLE) system can focus on incorporating more advanced deep learning techniques, such as convolutional and recurrent neural networks, to enhance prediction accuracy and adaptability across diverse walking patterns. Expanding the study to include a wider range of participants—varying in age, height, and mobility levels—will help improve the model's generalization. Additionally, refining the wireless communication framework can further reduce latency and ensure smoother real-time data transmission. The system also holds strong potential for clinical and rehabilitation use, where it could be integrated with wearable IoT devices or mobile platforms to provide continuous, personalized gait monitoring and analysis.

**Author Contributions:** Conceptualization, T.G, A.P. and P.S.; methodology, T.G and A.P.; software, T.G and A.P.; validation, T.G and A.P.; formal analysis, T.G, A.P., P.S., P.D. and N.R.; investigation, T.G, A.P. and P.S.; resources, T.G and A.P.; data curation, T.G and A.P.; writing—original draft preparation, T.G. and P.S.; writing—review and editing, P.D. and N.R.; visualization, T.G. and A.P.; supervision, P.S.; project administration, P.S.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Acknowledgments:** The authors gratefully acknowledge the facilities and support provided by the Director and all other staff members of the School of Mobile Computing and Communication, Jadavpur University, a Center of Excellence set up under the "University with potential for Excellence" Scheme of the UGC.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Whittle, M. W. (2014). *Gait Analysis: An Introduction* (5th ed.). Butterworth-Heinemann.
2. Baker, R. (2013). *Measuring Walking: A Handbook of Clinical Gait Analysis*. Mac Keith Press.
3. H. J. Lee et al., A normative study of the gait features measured by a wearable inertial sensor in a healthy old population, *Gait & Posture*, Volume 103, 2023, Pages 32-36, ISSN 0966-6362, <https://doi.org/10.1016/j.gaitpost.2023.04.006>.
4. Everett, T. & Kell, C. (2010). *Human movement: an introductory text*. Elsevier Health Sciences.
5. Donoghue, O. A., Leahy, S. & Kenny, R. A. Longitudinal associations between gait, falls, and disability in community-dwelling older adults with type II diabetes mellitus: findings from the Irish Longitudinal Study on Ageing (TILDA). *J. Gerontol. A Biol. Sci. Med. Sci.* 76, 906–913 (2021).

6. Kelly A. Hawkins, Emily J. Fox, Janis J. Daly, Dorian K. Rose, Evangelos A. Christou, Theresa E. McGuirk, Dana M. Otzel, Katie A. Butera, Sudeshna A. Chatterjee, David J. Clark, Prefrontal over-activation during walking in people with mobility deficits: Interpretation and functional implications, *Human Movement Science*, Volume 59, 2018, Pages 46-55, ISSN 0167-9457, <https://doi.org/10.1016/j.humov.2018.03.010>.
7. Angkoon Phinyomark, Sean Osis, Blayne A. Hettinga, Reed Ferber, Kinematic gait patterns in healthy runners: A hierarchical cluster analysis, *Journal of Biomechanics*, Volume 48, Issue 14, 2015, Pages 3897-3904, ISSN 0021-9290, <https://doi.org/10.1016/j.jbiomech.2015.09.025>.
8. Y. Tong, H. Liu and Z. Zhang, "Advancements in Humanoid Robots: A Comprehensive Review and Future Prospects," in *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 2, pp. 301-328, February 2024, doi: 10.1109/JAS.2023.124140.
9. Hausdorff, J. M. (2009). Gait dynamics in Parkinson's disease: Common and distinct behavior among stride length, stride time, and gait variability. *Human Movement Science*, 28(2), 149–156.
10. Nguyen, T. X. D., Mai, P. T., Chang, Y.-J. & Hsieh, T.-H. Effects of transcranial direct current stimulation alone and in combination with rehabilitation therapies on gait and balance among individuals with Parkinson's disease: a systematic review and meta-analysis. *J. Neuroeng. Rehabil.* 21, 27 (2024).
11. S. H. Collins and A. Ruina, "A Bipedal Walking Robot with Efficient and Human-Like Gait," *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, Barcelona, Spain, 2005, pp. 1983-1988, doi: 10.1109/ROBOT.2005.1570404.
12. Cutti, A. G., Ferrari, A., Garofalo, P., Raggi, M., Cappello, A., & Ferrari, A. (2010). Outwalk: a protocol for clinical gait analysis based on inertial and magnetic sensors. *Medical & Biological Engineering & Computing*, 48(1), 17–25.
13. Miezal, M., Taetz, B., & Bleser, G. (2016). On inertial body tracking and the estimation of rigid body orientation. *Sensors*, 16(1), 20.
14. L. E. Díez, A. Bahillo, J. Otegui and T. Otím, "Step Length Estimation Methods Based on Inertial Sensors: A Review," in *IEEE Sensors Journal*, vol. 18, no. 17, pp. 6908-6926, 1 Sept.1, 2018.
15. Q. Li, M. Young, V. Naing, and J. M. Donelan, "Walking speed estimation using a shank-mounted inertial measurement unit," *J. Biomech.*, vol. 43, no. 8, pp. 1640–1643, 2010.
16. Z. Zhu, S. Wang. A Novel Step Length Estimator Based on Foot-Mounted MEMS Sensors. *Sensors (Basel)*. 2018 Dec 15;18(12):4447. doi: 10.3390/s18124447. PMID: 30558332; PMCID: PMC6308666.
17. Sadhukhan P, Gayen B, Chowdhury C, Mukherjee N, Wang X, Das PK. Human Gait Modeling with Step Length Estimation based on Single Foot Mounted Inertial Sensors. Available at SSRN 4830580. 2024 May 16.
18. Paul, Aritra, Pampa Sadhukhan, Pradip K. Das, and Nirmalya Roy. "Sequential Neural Network-Based Gait Analysis and Step Length Estimation with a Foot-Mounted IMU." (2024).
19. Guimarães, Vânia, Inês Sousa, and Miguel Velhote Correia. 2021. "A Deep Learning Approach for Foot Trajectory Estimation in Gait Analysis Using Inertial Sensors" *Sensors* 21, no. 22: 7517. <https://doi.org/10.3390/s21227517>.
20. J. Park, J. Hong Lee and C. Gook Park, "Pedestrian Stride Length Estimation Based on Bidirectional LSTM and CNN Architecture," in *IEEE Access*, vol. 12, pp. 124718-124728, 2024, doi: 10.1109/ACCESS.2024.3454049.
21. H. Jin, I. Kang, G. Choi, D. D. Molinaro and A. J. Young, "Wearable Sensor-Based Step Length Estimation During Overground Locomotion Using a Deep Convolutional Neural Network," 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Mexico, 2021, pp. 4897-4900, doi: 10.1109/EMBC46164.2021.9630060.
22. Z. Ping, M. Zhidong, W. Pengyu and D. Zhihong, "Pedestrian Stride-Length Estimation Based on Bidirectional LSTM Network," 2020 Chinese Automation Congress (CAC), Shanghai, China, 2020, pp. 3358-3363, doi: 10.1109/CAC51589.2020.9327734.
23. Zhou, H., & Hu, H. (2008). Human motion tracking for rehabilitation—A survey. *Biomedical Signal Processing and Control*, 3(1), 1–18.
24. Willemsen, A.T.M., van Alsté, J.A., & Boom, H.B.K. (1990). Real-time gait assessment utilizing a new way of accelerometry. *Journal of Biomechanics*, 23(8), 859-863.
25. Castaneda, N., & Lamy-Perbal, S. (2010). An improved shoe-mounted inertial navigation system. *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Zurich, Switzerland, 15–17 September 2010.
26. Borenstein J, Ojeda L, Kwanmuang S. Heuristic Reduction of Gyro Drift for Personnel Tracking Systems. *Journal of Navigation*. 2009;62(1):41-58.

27. Woodman, O. J., & Harle, R. (2008). Pedestrian localisation for indoor environments. Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp'08), Seoul, Korea, 21–24 September 2008, 1–10.
28. Foxlin, E. (2005). Pedestrian tracking with shoe-mounted inertial sensors. *IEEE Computer Graphics & Applications*, 25(6), 38–46.
29. Jiménez, A. R., Seco, F., Prieto, J. C., & Guevara, J. (2009). A comparison of Pedestrian Dead-Reckoning algorithms using a low-cost MEMS IMU.
30. Kwakkel, S. & Godha, Saurabh & Lachapelle, Gérard. (2007). Foot and Ankle Kinematics During Gait Using Foot Mounted Inertial System. ION NTM San Diego CA.
31. Bamberg SJ, Benbasat AY, Scarborough DM, Krebs DE, Paradiso JA. Gait analysis using a shoe-integrated wireless sensor system. *IEEE Trans Inf Technol Biomed*. 2008 Jul;12(4):413-23.
32. Rampp, Alexander & Barth, Jens & Schüle, Sam & Gassmann, Karl & Klucken, Jochen & Eskofier, Bjoern. (2014). Inertial Sensor-Based Stride Parameter Calculation From Gait Sequences in Geriatric Patients. *IEEE transactions on bio-medical engineering*. 10.1109/TBME.2014.2368211.
33. Truong, P. H., et al. (2016). Stride Counting in Human Walking and Walking Distance Estimation. *Sensors*, 16(6), 823.
34. Ngueteu, A.-M., Otis, M. J.-D., & Batcho, C. S. (2025). Validation of novel stride length model-based approaches to estimate distance covered based on acceleration and pressure data during walking. *Sensors*, 25(19), 6217.
35. Vežočník, M., Kamnik, R., & Juric, M. B. (2022). Adaptive Inertial Sensor-Based Step Length Estimation Model. *Sensors*, 22(23), 9452.
36. Zadka, A., Rabin, N., Gazit, E. et al. A wearable sensor and machine learning estimate step length in older adults and patients with neurological disorders. *npj Digit. Med.* 7, 142 (2024). <https://doi.org/10.1038/s41746-024-01136-2>.
37. Sui, X., & Chang, Y.-C. (2022). IMU-based deep stride length estimation with self-supervised learning. arXiv preprint arXiv:2205.02977.
38. Vežočník, M., Kamnik, R., & Juric, M. B. (2021). Inertial Sensor-Based Step Length Estimation Model by Means of Principal Component Analysis. *Sensors*, 21(10), 3527.
39. S. Butterworth, "On the Theory of Filter Amplifiers," *Wireless Engineer*, vol. 7, no. 10, pp. 536-541, 1930.
40. Standard Scaler: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html>.
41. Sun, J., Du, W., & Shi, N. (2018). A Survey of kNN Algorithm. *Information Engineering and Applied Computing*, 9(2), Article 348305327.
42. Rivas-Perea, P., Cota-Ruiz, J., Chaparro, D., Venzor, J., Carreón, A., & Rosiles, J. (2013). Support Vector Machines for Regression: A Succinct Review of Large-Scale and Linear Programming Formulations. *International Journal of Intelligence Science*, 3(1), 5-1.
43. Qamar, R., & Baqar, H. (2023). Artificial Neural Networks (ANNs): An Overview. *Mesopotamian Journal of Computer Science*, 2023, 130–139.
44. Brownlee, J. (2022, August 5). Regression Tutorial with the Keras Deep Learning Library in Python. *Machine Learning Mastery*.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.