A machine learning approach for stride speed estimation based on a head-mounted IMU

P. Tasca[1], F. Salis[2], S. Rosati[1], G. Balestra[1] and A. Cereatti[1]
[1] *Politecnico di Torino*
[2] *Università degli Studi di Sassari*

***Abstract*—Walking speed in real-life conditions is typically estimated through wearable inertial sensors mounted on waist, lower limbs, or wrists. Very recently, head-mounted inertial sensors are emerging for gait assessment. The present study explores the feasibility of measuring the stride speed with a head-mounted inertial sensor in both laboratory and real-world settings. The developed algorithm exploits a Temporal Convolutional Network for the detection of the gait events and a Gaussian Process Regression for the stride speed estimation. The experimental evaluation was carried out on healthy young participants during both standardised indoor and real-world walking trials. For indoor trials, errors were smaller than previous studies (0.05 m/s). As expected, errors increased at lower speed regimes due to a reduced signals amplitude. During 2.5-hours real-world evaluation, errors were slightly larger but acceptable (0.1 m/s). Reported results are encouraging and show the feasibility of estimating gait speed with a single head-worn inertial sensor.**

***Keywords*— machine learning, gait analysis, head IMU, gaussian process regression**

I. Introduction

GAIT is the main form of locomotion for human beings, and it is strictly related to the individual's quality of life [1]. For this reason, gait-related spatio-temporal parameters have been widely explored in both healthy and pathological cohorts [2]. Walking speed is particularly relevant in the assessment of mobility and it is an indicator of one's health status [3]. In the last decades, accelerometers and gyroscopes housed within a single Inertial Measurement Unit (IMU) have emerged as the election technology for the tracking of gait in the real-world, thanks to their wearability and affordability [1]. Multiple body-worn IMUs configurations for estimating walking speed have been extensively explored in literature [4]. Although multi-sensor configurations provide more accurate results, they are less likely to gain acceptance in customer-oriented solutions, as they are less easy to use and comfortable for the subject [5]. Therefore, research groups have started to focus on single-sensor configurations [5]. The most exploited body sites for the positioning of single sensors include wrist [6], foot [7], and waist [8]. However, current IMU-based solutions typically lack in aesthetics and ease of use, resulting in low user engagement [9]. Recently, the use of head-worn devices with IMUs started to spread due to the unique characteristics of this location [10] and its potential integration in smart-glasses for daily life use [9]. Some head- worn devices with inertial sensors have already been translated into commercial solutions and validated for walking speed estimation, such as the Recon Jet [11] smart glasses. Not surprisingly, the use of head-mounted IMUs (H-IMUs), and head-worn devices in general, is expected to grow in the healthcare sector [12]. Traditionally, IMU-based walking speed estimation methods include biomechanical models of gait, direct integration methods and machine learning techniques [2]. Human gait models and direct integration methods respectively suffer from lack of customisation and drift errors [2]; moreover, walking speed is derived indirectly by estimating time and distance. Instead, machine learning allows to directly infer the stride speed through a mapping between stride-specific features and the walking speed [11]. The purpose of the present study is the estimation of the stride speed from signals recorded by a single H-IMU. Foot-to-ground initial contacts (ICs) are estimated through a Temporal Convolutional Network (TCN) and used to segment the strides. Then, time and frequency-domain features are extracted from the stride acceleration norm and fed to a GPR model that outputs the stride speed. The models are trained using data recorded in indoor standardised conditions and tested on data not used for training recorded in both standardised and real-world conditions.

II. Materials and methods

A. Experimental setup

Each subject was equipped with the INertial module with DIstance sensors and Pressure insoles (INDIP) system [14] with the configuration shown in Fig. 1. For the purpose of this study, it included an H-IMU positioned on the left side of the head, in addition to the standard INDIP setup used as wearable reference system (three IMUs on right foot, left foot and lower back, two pressure insoles and two distance sensors). Each INDIP IMU contains a 3-D gyroscope (±2000 dps), a 3-D accelerometer (±16 g) and a 3-D magnetometer (±50 G, not used in this study).

B. Experimental protocol

The experimental protocol included: i) a standardised indoor session comprising two motor tasks (12m-straight walk and double-ring walk, Fig. 2), each performed at different speeds (slow, comfortable, and fast) and repeated thrice, for a total of 12 trials; ii) a 2.5-hour unsupervised real-world acquisition in which the participants were asked to follow their daily life routine. All the INDIP sensors, sampling at 100 Hz, were synchronised before each acquisition by setting the timestamp of each unit. Data acquired during the standardised indoor session and the 2.5-hour real-world acquisitions will be referred to as the "standardised" and the "real-world" datasets, respectively.

C. Participants

Data were collected on fourteen young healthy adults (YHA) (6 females, age: 25 ± 3 years, height: 1.73 ± 0.08 m,

weight: 65 ± 10 kg). All participants provided written informed consent before taking part to the study. Of the 14 participants recruited, 11 participated to the standardised indoor session, while the remaining 3 performed the 2.5-hour real-world acquisition.

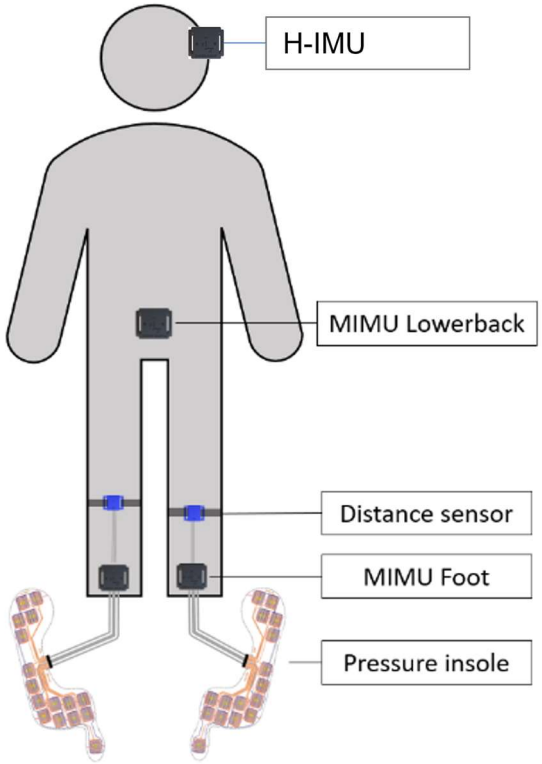

Fig. 1: The INDIP setup. IMUs and distance sensors are firmly attached to the body through Velcro bands to prevent any slippery o motion artifact.

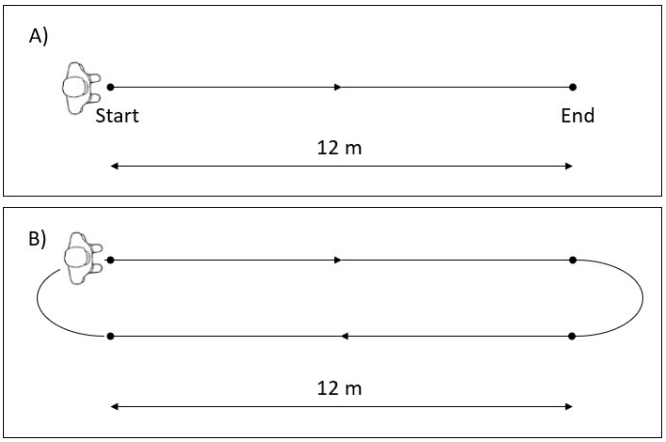

Fig. 2: Walking paths for the standardized acquisitions. A) Straight walk; B) "Double-ring" walk.

D. H-IMU algorithm

The estimation of the stride speed requires the previous segmentation of the stride. Therefore, the work was outlined as a two-tasks process:

i. Gait events detection
ii. Speed estimation

The gait events detection task was thought as a sequence classification task and was performed through a TCN [15], while the speed estimation task was organised as a regression task and was performed through a GPR model [11].

1) Pre-processing: H-IMU signals were virtually aligned to the direction of gravity, so that the x, y and z components of 3-D acceleration and angular rate were considered as parallel to the anteroposterior (AP), vertical (V) and mediolateral (ML) directions, respectively (Fig. 3).

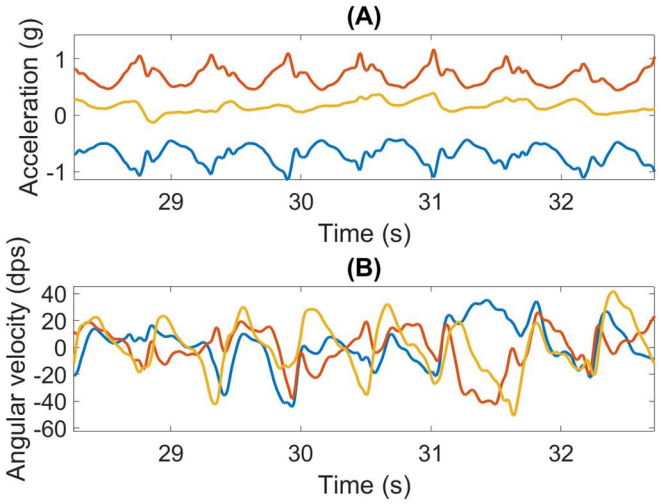

Fig. 3: A) Acceleration and B) angular velocity over four strides during a trial of standardized straight walking at comfortable speed. Acceleration is rotated with respect to gravity. Red: V component; Yellow: ML component; Blue: AP component.

2) *Gait events detection*: Norms of 3-D acceleration and angular rate were computed (Eq. 1).

$$a_n = \sqrt{{a_v}^2 + {a_{ml}}^2 + {a_{ap}}^2} \quad \omega_n = \sqrt{{\omega_v}^2 + {\omega_{ml}}^2 + {\omega_{ap}}^2} \qquad (1)$$

The six acceleration and angular rate components plus their norms were smoothed with a 5-points moving average filter and passed as predictors to a TCN. The network response was a binary signal of 0s and 1s with the same number of samples of the input sequence, with 0 denoting the double support phase (DS: both feet in contact with the ground, ~ 60% of the gait cycle) and 1 denoting the single support phase (SS: only right or left foot in contact with the ground, ~ 40% of the gait cycle). Portions of non-walking signals were ignored.

3) *Stride segmentation*: The ICs were detected at the falling edges of the network binary output i.e. at the end of the SS phases. Strides were determined as the time intervals between two ICs of the same foot. Since the TCN did not make a distinction between right and left SS phases, left and right ICs were not distinguished; however, the succession was preserved by picking adjacent ICs with a step of two.

4) *Stride speed estimation*: The raw acceleration norm of the segmented stride was filtered with a 4th-order Butterworth low pass filter with cut-off frequency of 5 Hz. A set of 134 features in the time and frequency domains were derived from the filtered norm and passed - together with height and weight of the subject - as input predictors to a GPR model (Table I). The decision regarding the number and typology of the input features was performed according to previous similar studies [11].

TABLE I: FEATURES FOR STRIDE SPEED PREDICTION

Domain	*Parameter*	*Definition*
Time (6)	Median	[g]
	Mode	[g]
	Signal Magnitude Area (SMA)	$\sum_{i=1}^{N}\lvert a_i\rvert$ [g]
	Energy	$\sum_{i=1}^{N}(a_i)^2$ [g^2]
	Zero-crossing rate	-
	Mean Absolute Value (MAV)	$\frac{1}{N}\sum_{i=1}^{N}\lvert a_i\rvert$ [g]
Frequency (128)	FFT coefficients	$\left\lvert\sum_{i=1}^{N} a_i e^{\frac{2\pi i}{N}(i-1)(k-1)}\right\rvert$ [g]
Anthro (2)	Weight	[kg]
	Height	[cm]

Features used by the GPR model to predict the stride speed. *Frequency (128)*: magnitudes of the 128 coefficients of the Fast Fourier transform (FFT). *g*: norm of gravity acceleration (9.81 m/s^2); *a*: samples of the acceleration norm for the current stride; *N*: number of samples of a; *Anthro*: anthropometric.

E. Training and testing

Both the TCN and the GPR models were trained with data of the standardised dataset (11 out of 14 participants). Data from participants 1-10 (construction set) were aggregated and used to train the models, while data from participant 11 (standardised test set) were used only for testing the trained architecture at the end of training. The TCN architecture included four residual blocks, a dilation base of 2, 64 filters of size 5 and a drop-out factor of 0.005. The TCN layer was followed by a fully connected layer, a SoftMax layer and a classification layer. The TCN was trained with a hold-out cross-validation scheme. The construction set was randomly split into a training set (TRS) of 7 subjects (for network training) and a validation set (VS) of 3 subjects (to evaluate the network performance and prevent overfitting) for five times. The network was trained for each of the five TRS-VS combinations for 60 epochs with the ADAM learning rule. Eventually, the network with the best performance overall was selected. At training stage, a label "0" was associated to samples in the DS phase, while a label "1" was associated to samples in the SS phase. To determine the beginning and the end of the DS and SS phases at training stage, gait events estimated by the INDIP reference system were exploited. The performance of the best TCN was evaluated on data of subjects not used for training, *i.e.* participant 11 of the standardised dataset and on the real-world dataset (3 out of 14 participants). The GPR model was trained according to a 5-fold cross-validation scheme on the construction set. At training stage, the GPR model was fed with 136 stride-derived features (predictors) and the reference value of the stride speed (target response) estimated by the INDIP system. At training stage, subjects' anthropometric features were used as constant predictors for their respective strides. An exponential kernel was selected for the GPR model basing on previous studies [11]. Performance of the best GPR model was evaluated on the standardised test set and on the real-world dataset.

F. Data analysis

Stride speed values predicted by the H-IMU algorithm were compared to the ones estimated by the INDIP reference system, that leverages validated algorithms [13] to predict gait parameters and events from data recorded by the pressure insoles and the IMUs (except the H-IMU). The adopted metrics for the method evaluation included mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), coefficient of adjusted determination (adR2) and Pearson index of correlation (ρ). Predicted and target response values were divided into three groups according to speed regime with two thresholds at 1 m/s and 1.5 m/s. Absolute percentage error was computed for each observation; then, errors out of the range [25th-percentile - 1.5 IQR; 75th-percentile + 1.5 IQR] were classified as outliers and discarded. A Shapiro-Wilk test was used to test the normality of the distribution. Then, as the distribution resulted to be non-normal, Wilcoxon ranked sum test was performed against the alternative hypothesis that the population mean of the slow-speed regime error distribution is greater than the population mean of the medium-and-fast-speed regimes error distribution. The significance level of the statistical test was set to 0.05.

III. RESULTS

Gait events, used to segment the strides, were detected by the TCN with a MAE of 0.05 ± 0.04 s on the test set and 0.02 ± 0.02 s on the training set.

Table II shows the results obtained from the comparison of the H-IMU algorithm with the reference system on the standardised dataset. Fig. 4 and 5 show the distributions of the absolute percentage errors for three speed regimes in the standardized construction set and test set, respectively. In both cases, the alternative hypothesis (Section II-F) was accepted at the specified significance level (p-value < 0.001). Table III shows the results obtained by the algorithm on the real-world dataset.

TABLE II: RESULTS ON STANDARDIZED DATASET

Set	*Number of strides*	*MAE (m/s)*	*RMSE (m/s)*	*MSE (m^2/s^2)*	ρ	*adR^2*
CS	3668	0.05 (4.6 %)	0.07	0.00	0.97	0.89
TS	354	0.07 (5.8 %)	0.08	0.01	0.97	0.70
CS+TS	4022	0.05 (4.7 %)	0.07	0.00	0.97	0.88

CS: Construction set; *TS*: Test set.

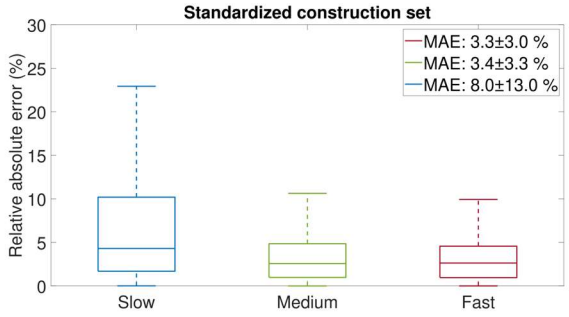

Fig. 4: Boxplot representation of the absolute percentage error distribution in the standardized construction set.

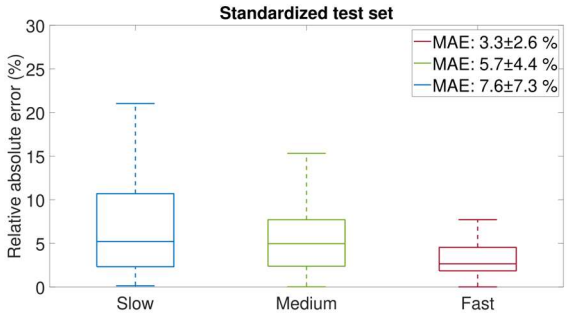

Fig. 5: Boxplot representation of the absolute percentage error distribution in the standardised test set.

TABLE III: RESULTS ON REAL-WORLD DATASET

Set	*Number of strides*	*MAE (m/s)*	*RMSE (m/s)*	*MSE (m^2/s^2)*	ρ	*adR^2*
RW	26498	0.1 (9.2 %)	0.12	0.13	0.80	0.37

RW: Real-world dataset.

IV. DISCUSSION

As shown in Table. II, a strong correlation exists between the predicted and the actual stride speed values for the standardised walking trials (ρ = 0.97). The fact that adR^2 is significantly lower than ρ suggests that some features have small contribution to the prediction output; therefore, dimensionality reduction methods shall be employed. The TCN-GPR-based algorithm has obtained a MAE of 0.05 m/s (4.7 %) on the standardised walking trials. Previous studies based on a single H-IMU for the estimation of the stride speed [11] have obtained higher errors in standardised conditions (6.5 %). As expected, significantly greater errors (p-value < 0.001) were found at lower speed regimes (MAE = 3.3±3.0% for medium and fast speed regime, MAE = 8.0±13.0% for slow speed regimes) due to smaller signals amplitude. This also suggests that the slow speed walking patterns of the participants are remarkably different than the medium and fast walking patterns. In real-world conditions, gait features show greater variability. As a result, correlation indices decrease, as shown in Table III (ρ = 0.80). The reported error (RMSE = 0.1 m/s) was equal to the one obtained by [6] through a single wrist-IMU (RMSE = 0.1 m/s). In the present study, the impact of gait events detection errors on the estimation of the stride speed was not investigated. However, MAE values obtained by the TCN were considerably low and in line with what achieved in other studies [15].

V. CONCLUSION

In this study, feasibility of walking speed estimation based on a head-mounted IMU using a TCN-GPR-based algorithm is investigated. TCN and GPR models were trained with the data acquired from 10 young healthy adults during standardised walking trials. Performance of the model was evaluated for both standardised and real-world walking trials. Results showed that the algorithm can estimate the stride speed in standardised conditions with lower errors with respect to other IMU-based methods. Nevertheless, accuracy of the method decreases at lower speed regimes. The algorithm also showed good generalisation skill on data acquired in real-world conditions. However, training of the algorithm with data acquired in real-world settings is suggested to increase accuracy. An approach based on a single H-IMU may be successful for many applications of fitness and health monitoring; however, it also comes with some limitations, as the upper the sensor location, the harder the extraction of gait features [16]. In the future, further aspects related to the implementation of models for the estimation of the stride speed should be addressed, such as dimensionality reduction, hyperparameters tuning, dataset completeness and eventually validation on pathological patients, where a machine learning-based approach may be effective for seizing the underlying features of abnormal walking patterns.

REFERENCES

[1] A. Cereatti, D. Trojaniello, U. Della Croce, Accurately measuring human movement using magneto-inertial sensors: techniques and challenges, *International Symposium on Inertial Sensors and Systems (ISISS),* 2015.

[2] S. Yang and Q. Li, Inertial sensor-based methods in walking speed estimation: a systematic review, *Sensors*, 2012.

[3] S. Fritz, White Paper: "walking Speed: the sixth vital sign", *J Geriatr Phys Ther,* 2009.

[4] W. Yeoh, I. Pek, Y. Yong, X. Chen, A.B. Waluyo, Ambulatory monitoring of human posture and walking speed using wearable accelerometer sensors, *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008.

[5] R. J. Mobbs, J. Perring, S. M. Raj, M. Maharaj, N. K. M. Yoong, et al., Gait metrics analysis utilizing single-point inertial measurement units: a systematic review, *mHealth*, 2022.

[6] A. Soltani, H. Dejnabadi, M. Savary and K. Aminian, Real-world gait speed estimation using wrist sensor: a personalized approach, *IEEE Journal of Biomedical and Health Informatics*, 2020.

[7] B. Mariani, C. Hoskovec, S. Rochat, C. Bula, J. Penders, et al., 3D gait assessment in young and elderly subjects using foot-worn inertial sensors, *Journal of Biomechanics*, 2010.

[8] S. Byun, H.J. Lee, J.W. Han, J.S. Kim, E. Choi, et al., Walking-speed estimation using a single inertial measurement unit for the older adults, *PLoS One*, 2019.

[9] A. Cristiano, A. Sanna and D. Trojaniello, Validity of a smart-glasses-based step-count measure during simulated free-living conditions, *Information*, 2020.

[10] C. Buckley, B. Galna, L. Rochester, C. Mazza, Attenuation of upper body accelerations during gait: piloting an innovative assessment tool for Parkinson's disease, *Biomed Res Int*, 2015.

[11] S. Zihajehzadeh and E. J. Park, A Gaussian process regression model for walking speed estimation using a head-worn IMU, *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* 2017.

[12] N. Hunn, The market for hearable devices 2016–2020, *Technical report November, WiFore Consulting, London*, UK, 2016.

[13] F. Salis, S. Bertuletti, K. Scott, M. Caruso, T. Bonci, et al., A wearable multi-sensor system for real world gait analysis, *Annu Int Conf IEEE Eng Med Biol Soc.* 2021, 2021.

[14] F. Salis, T. Bonci, S. Bertuletti, M. Caruso, K. Scott, et al., Performance of a multi-sensor wearable system for validating gait assessment: preliminary results on patients and healthy, *Gait & Posture*, 2022.

[15] R. Romijnders, E. Warmerdam, C. Hansen, G. Schmidt and W. Maetzler, A deep learning approach for gait event detection from a single shankworn IMU: validation in healthy and neurological cohorts, *Sensors*, 2022.

[16] J. J. Kavanagh, S. Morrison, R. S. Barrett, Coordination of head and trunk accelerations during walking, *Eur J Appl Physiol*, 2005.