Analysis of the Recent AI for Pedestrian Navigation With Wearable Inertial Sensors

Hanyuan Fu, Valerie Renaudin, Yacouba Kone, Ni Zhu

Analysis of the Recent AI for Pedestrian Navigation with Wearable Inertial Sensors

Hanyuan Fu, Valerie Renaudin, Yacouba Kone, Ni Zhu

***Abstract*— Wearable devices embedding inertial sensors enable autonomous, seamless, and low-cost pedestrian navigation. Appealing it is, the approach faces several challenges: measurement noises, different device-carrying modes, different user dynamics, and individual walking characteristics. Recent research applies Artificial Intelligence (AI) to improve inertial navigation's robustness and accuracy. Our analysis identifies 2 main categories of AI approaches depending on the inertial signals segmentation: either using human gait events (steps or strides) or fixed-length inertial data segments. A theoretical analysis of the fundamental assumptions is carried out for each category. Two state-of-the-art AI algorithms (SELDA, RoNIN), representative of each category, and a gait-driven non-AI method (SmartWalk) are evaluated in a 2,17 km long open access dataset, representative of the diversity of pedestrians' mobility surroundings (open-sky, indoors, forest, urban, parking lot). SELDA is an AI-based stride length estimation algorithm, RoNIN is an AI-based positioning method, and SmartWalk is a gait-driven non-AI positioning method. The experimental assessment shows the distinct features in each category and their limits with respect to the underlying hypotheses. On average, SELDA, RoNIN, and SmartWalk achieve 8 m, 22 m, and 17 m average positioning errors (RMSE) respectively, on six testing tracks recorded with two volunteers in various environments.**

***Index Terms*— Indoor positioning, inertial sensors, pedestrian navigation, dead reckoning, Machine Learning, Deep Learning**

I. Introduction

THE development of pedestrian navigation solutions has been an active field of research for almost two decades. The first technology employed is the Global Navigation Satellite System (GNSS) working in open-sky outdoor conditions. Indoors, radio beacon-based technology is deployed to locate pedestrians with ranging or mapping of signal footprints. These technologies are now widely commercialized but operate only in equipped infrastructure involving high installation and maintenance costs. Other approaches, aiming at fully autonomous navigation, are still being developed. They rely on image processing with SLAM, structure from motion or odometry methods, and inertial signals processing to infer the pedestrian's dynamics using wearable sensors. Inertial pedestrian navigation is very attractive because it doesn't require infrastructure, is operational with low-cost sensors that can be attached to several locations on the person's body (wrist, trousers pocket, etc.), and comply with the European General Data Protection Regulation recommendation promoting privacy by design technologies.

But pedestrian inertial navigation faces challenges. It accumulates positioning errors over time due to low-cost sensor noises. It lacks robustness when the pedestrian motion mode changes (slow/normal/fast walking, staircases, stationary, etc.) or when the wearable fixing point varies (handheld sensors, in the trouser/vest pocket, etc.). It also fails to adapt to individual walking gait characteristics (disability, injury). Artificial Intelligence (AI) is interesting for simultaneously addressing all these varying conditions and thus providing improved robustness and positioning accuracy. Consequently, it is increasingly applied to inertial pedestrian navigation research. Less explicit than traditional physics-based approaches, it raises design and hyper-parametrization issues.

This paper aims at analysing the recently proposed approaches in the field of pedestrian navigation with inertial wearable sensors to identify the key features that contribute to the success or limitations of robust and accurate positioning. It extends the previous conference proceeding [1], which classifies AI-based inertial pedestrian navigation methods into two main categories depending on the inertial signals segmentation. A detailed analysis of the fundamental hypotheses in each category, their likelihood, and their impact on positioning performance is conducted. A comparison with a non-AI pedestrian navigation approach is added. The experimental performance assessment is conducted with handheld inertial sensors on a larger open-access dataset in three different environments: urban, forest, and a shopping mall parking lot, including indoor and outdoor parts and staircases.

Section II presents the AI-based pedestrian navigation state-of-the-art methods: human gait and sampling frequency-driven AI methods, along with their underlying hypotheses. The three methods selected for evaluation are described in section III. Section IV is dedicated to the experimental evaluation and comparison of the three selected methods on pedestrian tracks. Section V concludes the article.

II. State-of-the-art on current AI methods for pedestrian inertial navigation

AI-based methods for pedestrian inertial navigation can be classified into two classes [1]: (A) human gait-driven and (B) sampling frequency-driven methods. The first category is inspired by the nature of human walking and the inertial

This work was supported by the French National Research Agency under Grant ANR-20-LCV1-0002

Hanyuan Fu, Yacouba Kone, Valerie Renaudin, Ni Zhu are with AME-GEOLOC, Gustave Eiffel University, F-44340 Bouguenais, France. {hanyuan.fu, yacouba.kone, valerie.renaudin, ni.zhu}@univ-eiffel.fr

signals are segmented with the user's gait (step or stride) events, whereas the second category segments inertial data into fixed-length sequences, usually overlapped.

A. Human gait-driven AI methods

The inertial data sequences are segmented using the gait events (step or stride instants) and processed to estimate the gait vector of each segment. Gait events are derived from the cyclic inertial signals patterns. For instance, [2] considers that a peak in the acceleration norm or a valley in the angular rate norm represents a step instant, regardless of the device's location. However, gait detection under irregular hand movements remains challenging.

Due to the different features needed to estimate the step/stride length and the walking direction, they are often treated separately.

1) Stride/step length estimation: Feature engineering is performed to select the most informative features for stride/step length estimation. According to [3], there are four categories of relevant features: statistical (mean, variance, amplitude, etc.), time-domain (number of peaks, zero-crossing ratio, etc.), frequency domain (dominant frequencies, spectrum energy, etc.), and higher-level features based on empirical models. Stride/step length can be regressed by different models: [4] uses a convolutional neural network, [5] uses an OS-ELM (Online Sequential Extreme Learning Machine) [6], [3] combines 6 regressors: Extreme Gradient Boost (XGBoost) [7], LightGBM [8], K-Nearest Neighbor (KNN) [9], Decision Tree (DT) [10], AdaBoost [11], and Support Vector Regression (SVR) [12].

An alternative to feature engineering is end-to-end regression with a sequence of inertial measurements over a gait interval. Feature extraction is usually performed by a deep network, [13] uses 2 stacked autoencoders, [14] uses a Deep Believe Network (DBN) built with multiple Gaussian Bernoulli Restricted Boltzmann Machines [15], [16] uses an LSTM [17], we will detail the model is section III. The regression task is done by one or several dense layers.

The main challenge of stride/step length estimation is the changeable device locations, user dynamics, and different individual walking characteristics. Classifying the device's location can improve robustness [3], [4]. Customizing the model for each individual is another way to improve accuracy [18].

2) Walking direction estimation: Estimating the user's walking direction with wearable inertial sensor measurements is a complex problem because the misalignment between the device's pointing direction and the user's walking direction is not necessarily constant. A common strategy is to express the inertial measurements in the navigation frame via device attitude tracking.

[6] uses the device's yaw angle and magnetometer measurements as features to infer the user's walking direction using an OS-ELM network. [19] doesn't need explicit device attitude tracking, instead, it performs data augmentation and adaptive alignment by a learnable Spatial Transformer Network (STN) to make the model invariant to the inertial data's rotation.

B. Sampling frequency-driven AI methods

This branch of AI methods considers pedestrian positioning as an end-to-end problem. Inertial measurement sequences are segmented into fixed-length segments, usually overlapped. Deep networks are trained to infer the user's average velocity or change in position over a segment.

The constant sampling frequency of the inertial measurements is a necessary condition for this branch of methods. If it is not the case, data interpolation and synchronization are needed.

RIDI [20], the "ancestor" of this branch, considers only strap-down configurations (leg pocket, in a bag, hand-held, body mounted). Raw inertial measurements are corrected thanks to a device attitude tracking algorithm and a neural network, before being integrated twice to obtain the user's position. The same team later proposed RoNIN [21], which can operate beyond strap-down configurations, that we selected for experimental evaluation (see Section III). IONet, another pioneer of this category, regresses the user's walking direction change and displacement every time it receives a 200-frame segment (200 Hz).

C. Hypotheses made for each AI category

The underlying hypotheses of each category are listed and analyzed.

1) Hypotheses for the Human Gait-Driven AI methods:

- Hypothesis GH1: Human walking is cyclic. A normal cycle of human locomotion is illustrated in Fig. 1. In [22], walking locomotion is described as a process in which the erect, moving body is supported by first one leg, and then the other. As the moving body passes over the supporting leg, the other leg swings forward to prepare for its next supporting phase.
- Hypothesis GH2: According to [23], The upper and the lower body movements are correlated. During normal walking, the head and the trunk move up and down as the center of gravity follows the lower limbs' periodic movements, and arms flex and extend reciprocally. This hypothesis makes gait event detection from inertial signals plausible.
- Hypothesis GH3: Human paces are regular and constrained. According to a statistical study presented by [3], within 10145 strides of gait measurements of different subjects and different walking dynamics, collected by a foot-mounted device, 99.5% of strides were within 1.55 m and no stride exceeds 1.75 m. The mean and standard deviation of stride lengths are 1.33 m and 0.18 m.
- Hypothesis GH4: Step/stride length and inertial signals collected from different body parts are correlated. Empirical models are developed based on hip or foot-mounted sensors. Weinberg [24] found a correlation between the hip's vertical acceleration amplitude and stride length. Ladetto [25] found a correlation between acceleration variance and stride length. The hypothesis is plausible considering the correlation between foot movements and those of the rest of the body (GH2).

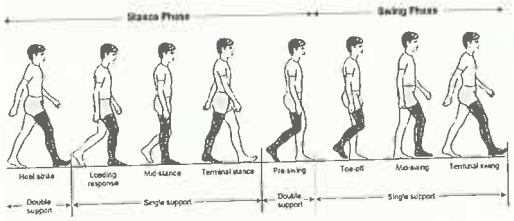


Fig. 1: Human gait cycle [26]

- Hypothesis GH5: A corollary of GH1, when the device location remains the same, the user's change in walking direction over two consecutive steps shall be approximately the same as the change in the device's pointing direction.

GH1 and GH3 are observed during natural walking, certain users such as senior citizens can easily break these assumptions. GH2 is a simplification of reality since users can move freely their arms and hands while walking. As for GH4, even if these correlations exist, they're likely to be significantly diversified depending on the device carrying mode, which explains the fact that most research works either consider a single carrying mode ([14], [16]) or perform carrying mode classification ([3], [4]). GH5 can be easily corrupted by noisy movements (swinging, device in pocket/bag).

2) Hypotheses for sampling frequency-Driven AI methods:

- Hypothesis FH1: The true kinematic of the user's center of mass is continuous and can be recovered from inertial measurements collected from different body parts.
- Hypothesis FH2: Each fixed-length segment is independent of the others. In other words, a segment contains sufficient information to recover the user's velocity or change in position over the same time window.
- Hypothesis FH3: The inertial signals are sampled at a constant sampling frequency.

This branch of methods is naturally suitable for using deep learning models which require fixed-length inputs. FH2 is approximately true if we consider regular and cyclic movements that the user's speed can be estimated with the signal's frequency and amplitude. However, the correlation between the user's speed and the signal's frequency or amplitude can vary from one individual to another. The same hypothesis also implies that a segment contains sufficient information to yield a walking direction inference. According to a survey [27] on traditional methods for walking direction estimation with an unconstrained device, existing methods such as Principal Component Analysis (PCA), [28], [29], Forward and Lateral Accelerations Modeling (FLAM) [30] and Frequency analysis of Inertial Signals (FIS) [31], assume that the walking direction is observable with handheld inertial sensors during one step/stride. As a result, we expect noisy inferences from this approach.

III. Selected Methods for the Experimental Assessment: SELDA, RONIN, and SmartWalk

We selected the method [16]: pedestrian Stride-length Estimation based on LSTM and Denoising Autoencoders (titled SELDA in the rest of the paper) among the gait-driven AI methods. The latter is representative of the category with sufficient implementation and data processing details, along with a benchmarking dataset. RoNIN [21] is selected among the sampling frequency-driven AI methods since the authors published their implementation, trained model weights, and a part of their dataset.

We would also like to evaluate a complete positioning non-AI gait-driven method for comparison, thus we selected SmartWalk. The method combines several techniques such as machine learning, statistical model, and an Extended Kalman Filter (EKF) to infer the user's trajectory, moreover, some parameters in the model are customized for each user.

A. SELDA: pedestrian Stride-length Estimation based on LSTM and Denoising Autoencoders

The stride length estimation AI model proposed in [16] takes as input stride segments of inertial measurements: 3-axis acceleration and angular rate, collected by a handheld smartphone, along with higher-level features from empirical models (Weinberg [24], Ladetto [25], Scarlett [32], etc.) computed with the acceleration segment.

SELDA requires the user to carry the device in "texting" mode, in such a way that the z-axis of the device points to the sky. Stride events for signal segmentation are detected by a foot-mounted device. This setup is constrained and impractical for deployment.

1) SELDA Dataset: publicly available [33], is collected by 5 volunteers of different gender and height, holding the smartphone horizontally in front of their chest.

The dataset covers both indoor and outdoor scenarios including staircases.

the inertial signal sampled at 100 Hz is segmented by stride instants provided by the foot-mounted reference tracker, which also provides stride length ground truth.

2) Adaptation for experimental assessment: We use only 4 higher-level features instead of 35 since only 4 definitions are available.

Among learning samples from the SELDA dataset, we only consider the stride length range between 0.3m and 1.8m. The original dataset is split into a 6319-sample training set and a 175-sample validation set.

Since SELDA only estimates stride length, for illustration's purpose, we pile up three modules namely stride detection, SELDA, and heading estimation to build a positioning system. Stride instants and walking directions are provided by our foot-mounted reference tracker ([34] and [35]).

B. RoNIN: Robust Neural Inertial Navigation

RoNIN expresses acceleration and angular rate measurements in the navigation frame, via the device attitude provided by android's Game Rotation Vector (GRV) and a spatial alignment procedure. An AI model (RoNIN ResNet) is trained to infer the user horizontal velocity (V_x, V_y), given a fix-length segment (200 frames) of acceleration and angular rate expressed in the navigation frame. Inferred velocities are integrated to obtain the user's trajectory. Both RoNIN dataset,

model implementation, and trained model weights are publicly available [36].

1) RoNIN dataset: is collected mainly in indoor environments by 100 volunteers and 3 android devices, covering usual scenarios such as a smartphone in a bag, in the pocket, handheld, walking, sitting, etc.

The ground truth trajectories are provided by visual-inertial SLAM performed by a tango phone attached to the volunteer's chest.

All measurements are synchronized and sampled at 200 Hz.

2) Importance of the Game Rotation Vector: Game Rotation vector is a quaternion provided by android API, indicating the device's orientation w.r.t some gravity-aligned reference frame [37]. To better understand it, we compare it to a transparent EKF device attitude tracking algorithm MAGYQ [38]. We attach rigidly our "home-made" navigation device (see section IV) ULISS and a smartphone on an aluminum plate to align their z axes (Fig. 2). We recorded 2 tracks.

- Track 1: slow and steady rotations.
- Track 2: random rotations during walking.

In Fig. 3a and Fig. 3b, We plot the Euler angles of the smartphone given by GRV (top figure) and those of the ULISS given by MAGYQ (middle figure). The roll angle of ULISS evolves in the same way as the pitch angle of the android device and the pitch angle of ULISS evolves in the same way as the android device's roll angle. ULISS yaw angle is opposite to the yaw angle of the android device. In the bottom figure, we plot the variation of their angular offsets given by Eq. 1. The under-script "a" stands for android and "u" stands for ULISS.

$$\Delta_{roll} = roll_u - roll_u(t=0) - (pitch_a - pitch_a(t=0)) \quad (1)$$

$$\Delta_{pitch} = pitch_u - pitch_u(t=0) - (roll_a - roll_a(t=0)) \quad (2)$$

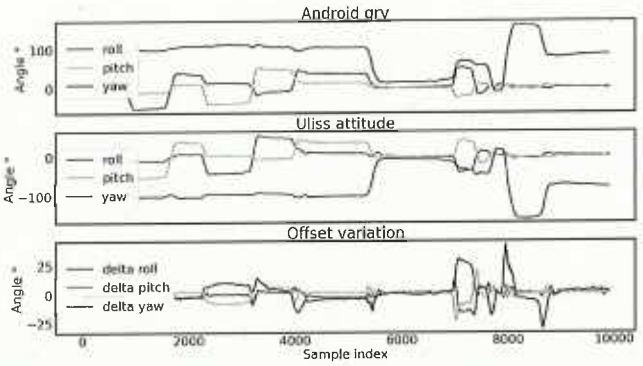
$$\Delta_{yaw} = yaw_u - yaw_u(t=0) + (yaw_a - yaw_a(t=0)) \quad (3)$$

Fig. 2: Experimental setup for studying the game rotation vector

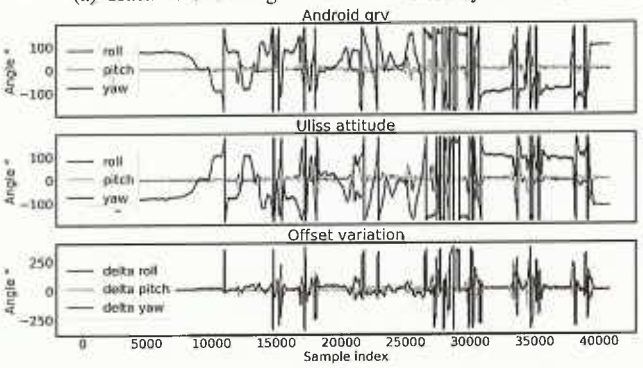
The figures show that the game rotation vector is good at estimating roll and pitch (related to gravity). There is no large difference between the game rotation vector and the MAGYQ result. The offset between game rotation yaw and ULISS yaw is almost constant during several minutes of recordings.

We observe punctual peaks in the bottom plots, which are due to are due to slight synchronization lags or the difference in response time of the two devices.

We can conclude that the offset between the game rotation vector and the device's orientation, w.r.t to North-East-down frame, is approximately constant for several minutes. For this reason, we replace the game rotation vector with MAGYQ in our experiments to make the RoNIN more transparent.



(a) Track 1: Euler angles of slow and steady rotations



(b) Track 2: Euler angles of random rotations during walking

Fig. 3: Comparison of attitude angles estimated by android game rotation vector and MAGYQ for two scenarios

3) Adaptation for experimental assessment: The article proposes 3 variants based on different deep learning models: ResNet [39], LSTM [17], and Temporal Convolutional Network (TCN) [40] to estimate the user's position. We only assess RoNIN ResNet, since it yields the best results. We use the published model implementation and trained model weights for experimental assessment. We replace the game rotation vector and the spatial alignment procedure with MAGYQ.

C. SmartWalk

It appeared interesting to make a comparison with an older solution based on classical signal processing techniques instead of recent AI tools: SmartWalk [41]. It is a pedestrian positioning algorithm fusing data from a tri-axis accelerometer, a tri-axis gyroscope, a tri-axis magnetometer, a barometer, and a GNSS receiver. In this paper, GNSS raw data are not included in the positioning algorithm to ensure a fair comparison with the other AI-based solutions. Smartwalk contains several modules:

1) Step detection, step length estimation and carrying mode classification: This module processes the wearable raw acceleration and angular rate readings at 200 Hz. The original step detection module was replaced with SmartStep [2], [42]. Step length s is estimated by a linear model [43]

$$s = h \times (a \times f + b) + c \quad (4)$$

where h is the user's height, f is the frequency of the acceleration magnitude and a, b and c are universal parameters learned on a group of users. A personalized calibration is as well possible [44]. The device's carrying mode is classified into either texting or swinging mode. Irregular and static phases are also detected [43].

2) GMM for walking direction inference: This module adopts a Gaussian Mixture Model (GMM) to model the misalignment between the device's pointing direction and the user's walking direction [45]. First, the device's accelerations are transformed into the local North-East-Down frame using the Extended Kalman Filter-based device attitude tracking algorithm: MAGYQ [38]. Then, the distribution of the horizontal accelerations, when the user is walking toward the North (0° heading), is modeled by a weighted sum of 2-D Gaussian distributions:

$$facc(x) = \sum_{k=1}^{q} \tau_k N(x, m_k, P_k) \tag{5}$$

Where τ_k is the weight of Gaussian component k, parameterized by mean m_k and variance P_k. A GMM model is learned for each individual for a device handling mode, by expectation maximization. Finally, the walking direction of a step is inferred by rotating the learned GMM by angle θ to maximize the log-likelihood of the horizontal accelerations cloud over one step. θ is the inferred user heading for this step.

3) Corrections: An EKF completes SmartWalk with the following corrections:

- Identify the stairs with a barometer and use a fixed step length (30cm) when the user is on stairways.
- Use a fixed step length (50cm) in the propagation model and the estimated step lengths as observations.
- Fuse the GMM's inference with the device's pointing direction given by MAGYQ according to hypothesis GH5. High confidence is given to the GMM at the beginning of the trajectory to match the initial heading with the GMM prediction.

Table I summarizes the inputs and outputs of the three selected methods.

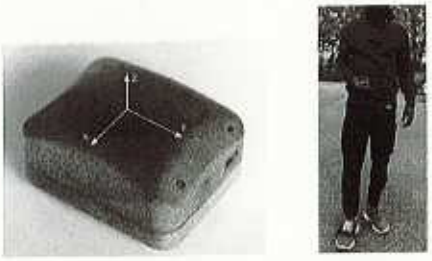
	SELDA	RoNIN	SmartWalk
Input	stride instants, acceleration, angular rate	acceleration, angular rate, attitude of the device (magnetic field)	acceleration, angular rate, magnetic field, the user's height, pressure
Output	stride length	user's position	user's position

TABLE I: Summary of the inputs and output of each selected method

IV. Experimental assessment

A. Experimental setup

1) Hardware: The device ULISS, shown in Fig. 4(a), was developed by the GEOLOC laboratory at University Gustave Eiffel and is used for the experiments. It is a state-of-the-art



(a) (b)

Fig. 4: (a) ULISS sensor; (b) Experimental setup: one ULISS sensor is held in the user's right hand and the other on the user's right foot

Inertial Navigation System containing an Xsens Mit-7 IMU-Mag sensor, a barometer, and a GNSS receiver, providing acceleration, angular rate, magnetic field, and atmospheric pressure readings at 200 Hz, GNSS reading at 5 Hz, using GPS timestamps. It is used for the experimental assessment instead of a smartphone because the signal acquisition is controlled and the sensors are calibrated. One ULISS is mounted on the foot and is the reference solution, i.e. ground truth, with a 0.3% positioning error over the traveled distance [46]. It is the winning solution of the 3-year French national competition (MALIN) for the non-collaborative positioning of soldiers in challenging indoor environments [35].

2) Scenarios: As shown in Fig. 4(b), the test person holds one ULISS horizontally (the z-axis points to the sky) and walks naturally, as requested by SELDA. All scenarios started outdoors for the initialization of MAGYQ and the foot-mounted reference solution. Both algorithms need a magnetometer calibration without strong artificial magnetic fields for the initialization. For the sake of fairness, the ground truth initial walking direction is given to all three implemented positioning methods.

Four different surroundings, representing the diversity of common pedestrian navigation contexts, were chosen for the experiments. Fig. 5 shows the diversity of these environments. Tests 1 to 3 were recorded on the campus of the university by a healthy man (volunteer 1, 1.66 m height), and tests 4 to 6 in various environments (forest, city, and parking lot), by another healthy man (volunteer 2, 1.80 m height). The six recorded tracks are provided in open source [47]. They include raw inertial signals, calibration parameters, and ground truth trajectories. Table II summarizes the main characteristics of these six evaluation tracks.

B. Performance evaluation

Three metrics are selected to evaluate the horizontal trajectories estimated by each selected method: the Scale Factor (SF), the EndPoint error Rate (EPR), and the Root Mean Square Error (RMSE).

1) The scale factor: is the ratio of the total length of the estimated trajectory l_{es} to the total length of the ground truth trajectory l_{gt}. The ratio is expected to be close to 1.

$$SF = \frac{l_{es}}{l_{gt}} \tag{6}$$

(a) (b) (c) (d) (e)

Fig. 5: Environments for experiments: (a) campus; (b) office building; (c) woods; (d) city; (e) parking lot in a shopping mall

Test	volunteer	location	length (m)	outdoor	indoor	stairways	surface texture	magnetic perturbation
1	1	campus	251	✓	✓	✓	concrete	✓
2	1	campus	230	✓			concrete	✓
3	1	campus	288	✓	✓	✓	concrete	✓
4	2	woods	374	✓			soil/grass	
5	2	city	563	✓			concrete	✓
6	2	parking lot in a shopping mall	464	✓			concrete	✓

TABLE II: Characteristics of the six tests in open access [47].

2) The endpoint error rate: is the ratio of endpoint error (EPE) to the ground truth trajectory's total length l_{gt}.

$$EPE = \sqrt{[(x_{end} - \hat{x}_{end})^2 + (y_{end} - \hat{y}_{end})^2]} \quad (7)$$

$$EPR = \frac{EPE}{l_{gt}} \quad (8)$$

3) The Root Mean Square Error: measures the standard deviation on the horizontal positioning accuracy.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2]} \quad (9)$$

where n is the number of points in the trajectory, (x_i, y_i) is the user's ground truth position at time step i, and $(\hat{x}_i, \hat{y}_i)$ is the predicted one.

The experimental results are reported in Table III. Estimated and ground truth trajectories are shown in Fig. 6 for the on-campus datasets (volunteer 1) and in Fig. 7 for the forest/urban/parking dataset (volunteer 2).

C. Analysis of the inertial pedestrian positioning estimates

1) Walking distance: The scale factor (SF) evaluates the quality of the estimated walking distances. Table III shows that RoNIN always underestimates the walking distance. But the standard variation of RoNIN's scale factor (0.067 for volunteer 1 and 0.019 for volunteer 2) is smaller than the one of SELDA (0.079 for volunteer 1 and 0.065 for volunteer 2), which over or underestimates the walking distance. Globally, RoNIN is able to better follow the pedestrian's dynamics changes as compared to SELDA. But important drifts are observed in the RoNIN trajectories for Tests 2 and 3. These observations are further detailed in Fig. 8 to Fig. 11.

Fig. 8 and Fig. 9 show the stride lengths predicted by SELDA (blue) against the ground truth (orange) for both volunteers. To complete the analysis, stride lengths estimated by the state-of-the-art Weinberg model (green) are plotted over.

The "nearly flat" line of the SELDA's predictions shows that it fails to capture the variations in the user's movement, especially when the user is taking stairs (smaller strides). Weinberg estimates, sharing one higher-level feature with SELDA, are much better at tracking variations in walking dynamics. The RoNIN's velocity plots (Fig. 10 and Fig. 11) show better performance in the tracking of walking changes: start, stop, and taking stairs. But, as foreseen by the theoretical analyses (FH2), the velocity estimates are noisy. The integration of the inferred velocities smooths this noise.

Globally, the two categories of methods show completely opposite behaviors. The failure of SELDA and the robustness of RoNIN can be explained by their training datasets. All stride length labels from SELDA's training set and their distribution are shown in Fig. 12. The mean and standard deviation of SELDA stride length labels are 1.36 m and 0.078 m. Most of the labels are close to the mean value, which is the best guess that the model can achieve. On the other hand, RoNIN regresses the two components of the velocity, whose variations are more important due to the infinity possibility of the walking direction.

Thanks to the sampling frequency-driven data segmentation, RoNIN is more data-intensive than SELDA for the same amount of measurements. For example, if a normal walking sequence lasts 12 s and is sampled at 200 Hz, only about 10 strides segments can be extracted from the track for SELDA. An average gait cycle lasts about 1.2 s. On the other hand, $((12 \times 200) - 200)/5 + 1 = 441$ segments can be extracted for RoNIN, i.e. 200 measurement points in one segment, with a stride of 5 frames.

For volunteer 1, SmartWalk tracks the best the user's dynamics (with a 0.038 standard variation), whereas it is RoNIN (with a 0.019 standard variation) for volunteer 2. To understand the observation, stride lengths estimated by the SmartWalk's step model without (orange) and with (green) EKF correction are illustrated against the ground truth (blue) in Fig.13 and Fig. 14. Step instants detected by processing the handheld inertial sensors data are projected on the ground truth to observe the derived stride (2 steps) length. Strides larger than 2 m can be observed in the blue dots, illustrating the under-detection of gait events in the SmartWalk approach. Similar to SELDA, SmartWalk's estimations form a relatively flat (less than SELDA) line when there is no staircase. When the staircase detection function is operational with the barome-

	SELDA			RoNIN			SmartWalk		
Volunteer 1	SF (unitless)	EPR (%)	RMSE (m)	SF (unitless)	EPR (%)	RMSE (m)	SF (unitless)	EPR (%)	RMSE (m)
test 1	1.08	9	12.15	0.98	0.23	6.38	0.90	4.78	7.59
test 2	0.96	0.66	0.41	0.82	14.5	13.42	0.84	6.51	8.55
test 3	1.16	6.29	10.7	0.87	6.65	16.53	0.93	2.27	3.20
average	1.03	5.33	7.76	0.89	7.11	12.11	0.89	4.52	6.45
Standard deviation	0.079	3.49	5.23	0.067	5.82	4.25	0.038	1.74	2.33
	SELDA			RoNIN			SmartWalk		
Volunteer 2	SF (unitless)	EPR (%)	RMSE (m)	SF (unitless)	EPR (%)	RMSE (m)	SF (unitless)	EPR (%)	RMSE (m)
test 4	0.9	6	3.09	0.68	7.45	29.47	0.77	2.81	21.15
test 5	0.79	3.64	20.17	0.68	1.65	41.66	0.72	16.36	57.23
test 6	0.95	0.3	1.99	0.64	3.80	22.71	0.86	3.31	6.94
average	0.88	1.51	8.42	0.67	4.30	31.28	0.78	7.49	28.44
Standard deviation	0.065	1.51	8.32	0.019	2.39	7.84	0.058	6.27	21.17

TABLE III: Performance evaluation of SELDA-based PDR, RoNIN, and SmartWalk

	SELDA			RoNIN			SmartWalk		
	SF (unitless)	EPR (%)	RMSE (m)	SF (unitless)	EPR (%)	RMSE (m)	SF (unitless)	EPR (%)	RMSE (m)
Overall average	0.96	3.42	8.09	0.78	5.71	21.7	0.84	6.01	17.45

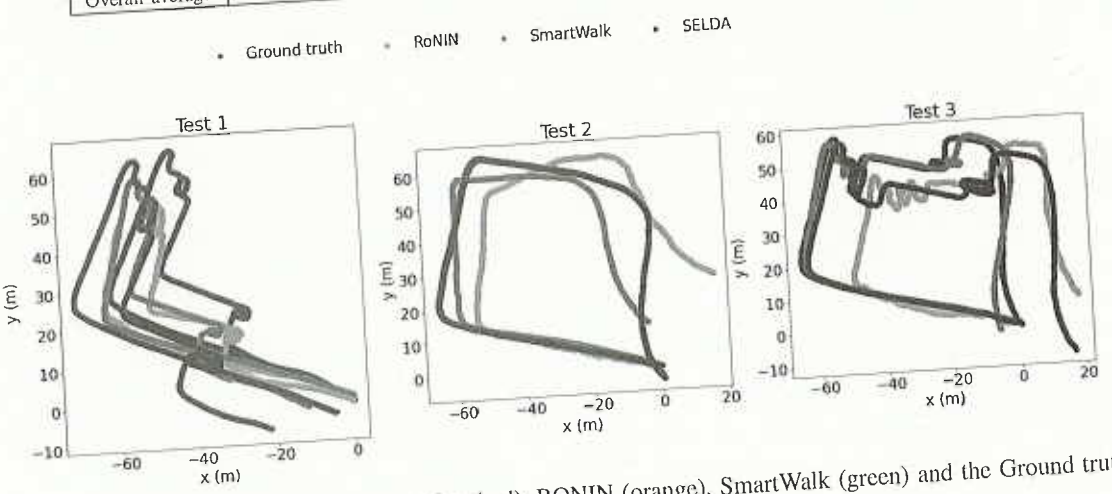


Fig. 6: Volunteer 1: Estimated trajectories by SELDA (red), RONIN (orange), SmartWalk (green) and the Ground truth (blue)

ter readings, SmartWalk outperforms both RoNIN and SELDA in scenarios with staircases (Tests 1 and 3). Without a stairway, RoNIN gives the best user dynamics tracking. On the other hand, SmartWalk can be sensitive to barometer noises and over-detect staircases. In Test 4, several tiny strides of 60cm are inferred but no stairway is included in the track.

Despite SELDA's failure of tracking the user's dynamics by giving almost constant stride length inferences, its scale factors are closer to 1 than the other two methods, when stairs are not included in the track (Tests 2, 4, 5, and 6). Indeed, the constant stride length estimated by SELDA corresponds to an average learned from its training set. Because human walking is regular and constrained, the deviation of a stride length around the average is bounded.

2) Walking direction: Only RoNIN and SmartWalk are compared for the walking direction estimation task since SELDA does not predict the latter. Noisy velocity inferences result in noisy walking directions estimated by RoNIN. SmartWalk is better than RoNIN since the mean RMSE of SmartWalk is smaller in both cases of volunteers 1 and 2. The EKF correction module is efficient here.

The benefit of SmartWalk's EKF correction is illustrated in Fig. 15 for Volunteer 2. Similar to RoNIN, GMM standalone (orange) yields noisy walking direction estimations. On the other hand, the trajectories estimated with MAGYQ's yaw angle (blue) have almost the same shape as the ground truth. Under the hypothesis that when the carrying mode remains steady, the change in the user's walking direction over two consecutive steps shall be the same as the change in the device's pointing direction (Hypothesis GH5). The device's yaw angle can be utilized to correct GMM results since MAGYQ results are more accurate and smooth. The fusion improves significantly the walking direction estimation, especially for turnings (green).

V. Conclusion

This article analyses the features of existing AI-based pedestrian inertial positioning techniques with wearable sensors both at the theoretical and experimental performance levels. A 2-category classification, based on the inertial segmentation strategy is presented: either using the human gait analysis or the inertial signals sampling frequency.

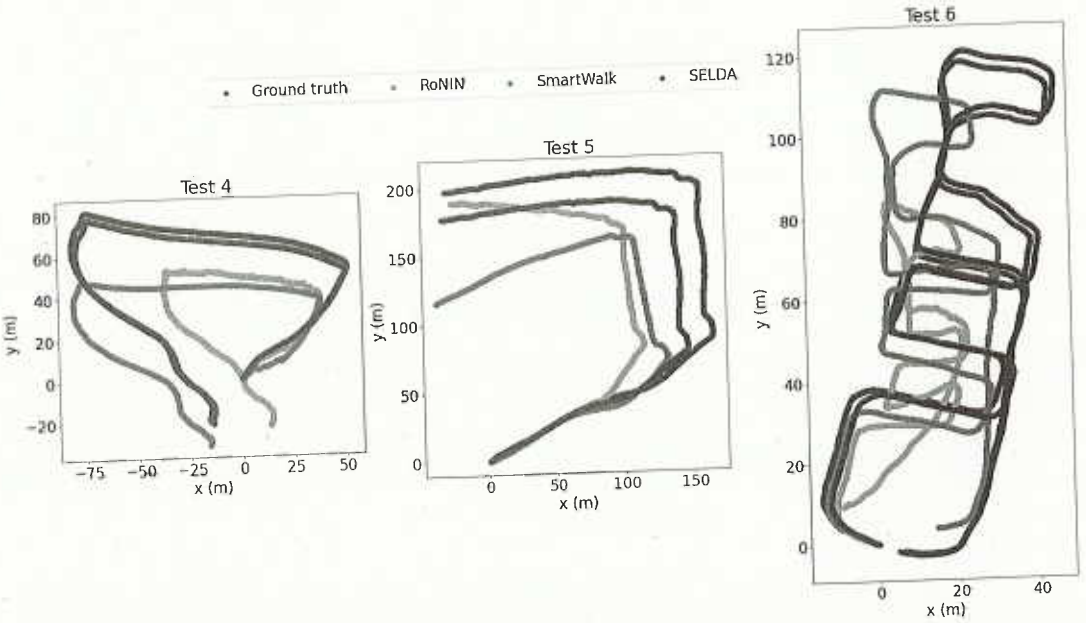


Fig. 7: Volunteer 2: Estimated trajectories by SELDA (red), RONIN (orange), SmartWalk (green) and the Ground truth (blue)

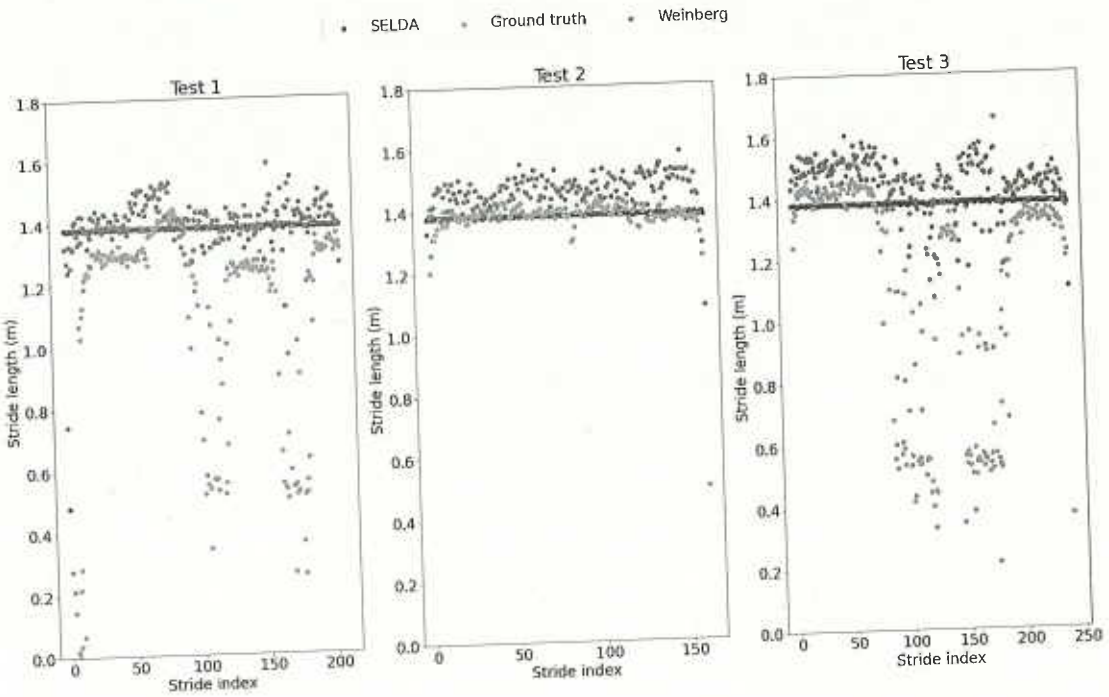


Fig. 8: Volunteer 1: Stride length predicted by SELDA (blue), by Weinberg (green) and the ground truth (orange)

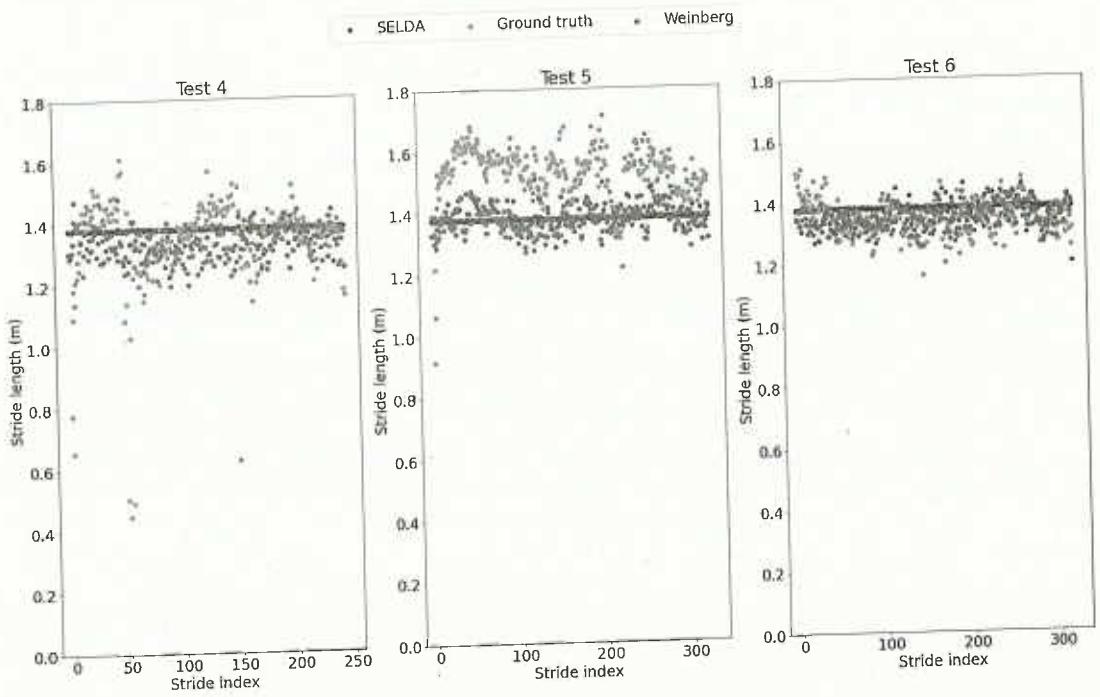


Fig. 9: Volunteer 2: Stride length predicted by SELDA (blue), by Weinberg (green) and the ground truth (orange)

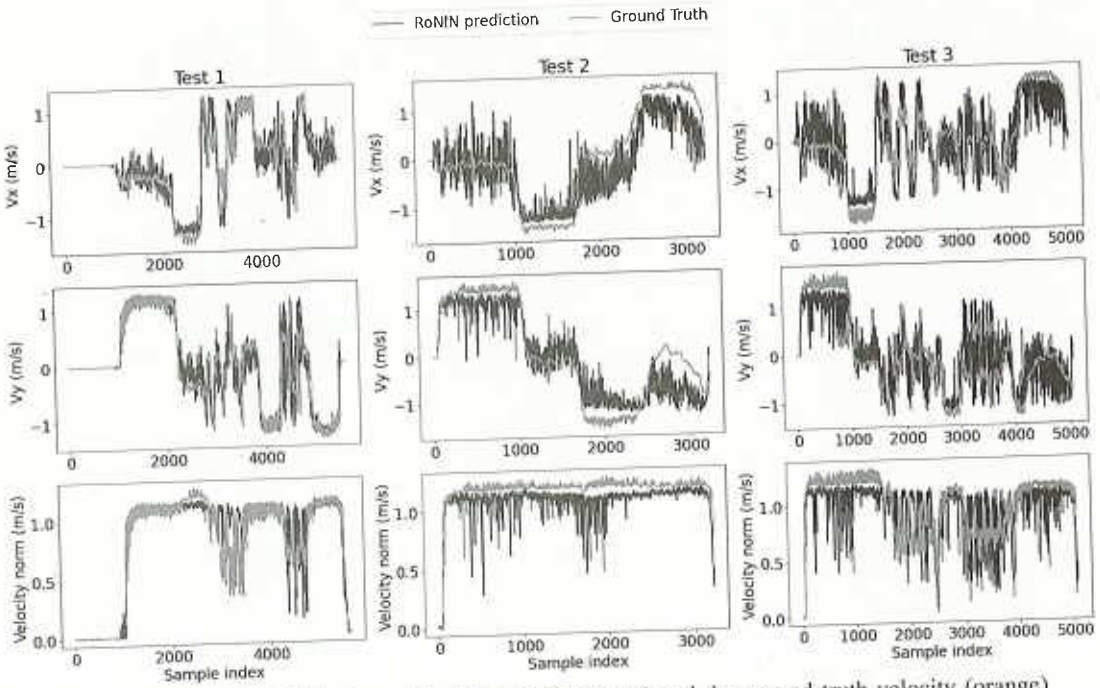


Fig. 10: Volunteer 1: Velocity predicted by RoNIN (blue) and the ground truth velocity (orange)

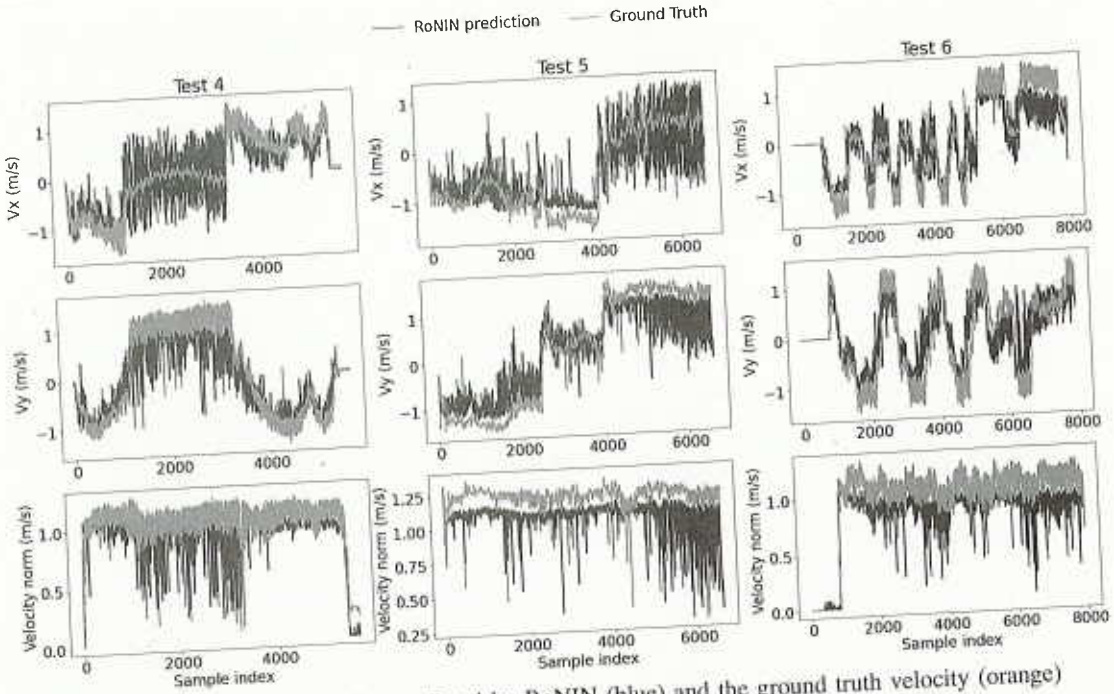


Fig. 11: Volunteer 2: Velocity predicted by RoNIN (blue) and the ground truth velocity (orange)

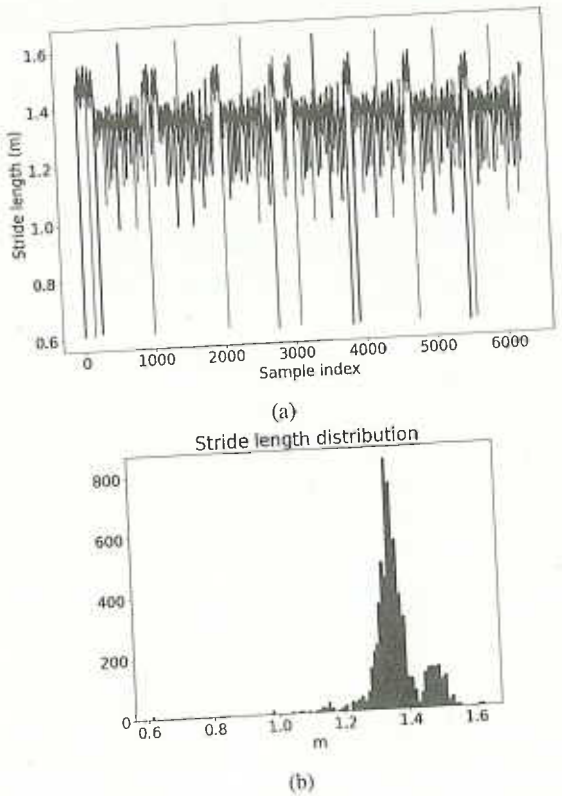


Fig. 12: SELDA training set: (a) stride length labels; (b) stride length distribution

A theoretical analysis of the fundamental assumptions that allow the two categories of AI-based methods to function properly is carried out. A state-of-the-art algorithm from each category (SELDA and RoNIN) and a classical signal processing-based algorithm (SmartWalk) are detailed and implemented for the 2,17 km experimental assessment in six scenarios and environments covering the diversity of pedestrians' mobility (open-sky, indoors, forest, urban, parking lot). The dataset is open-access.

SELDA uses the flat-foot instants of a foot-mounted tracker to segment inertial signals. It was found to be inefficient for labeling wearables' training datasets [42] and unrealistic in real-life situations. To better capture the walking changes with gait-driven AI methods, different walking dynamics (various stride lengths) should be added to the training dataset, and adopting a user-centric approach could be beneficial.

Compared to SELDA, the signal segmentation and labeling strategies of RoNIN show better learning. It estimates the walking distance up to a scale factor, which is found to be stable for the same individual and different from one individual to another. In addition, noisy predicted velocities and lack of precision in estimating turnings result in important positioning drifts.

SmartWalk shows that gait analysis can provide efficient corrections to trajectory estimation: better displacement estimates on stairs, smooth walking direction on straight lines, and correct turning angle estimation. It is worth noticing that significant changes in the device's yaw angle can indicate turnings. However, hypothesis GH5 still needs to be tested under diverse scenarios other than the "texting" case.

Globally, SELDA, RoNIN, and SmartWalk achieve 8 m, 22 m, and 17 m average positioning errors (RMSE) respectively,

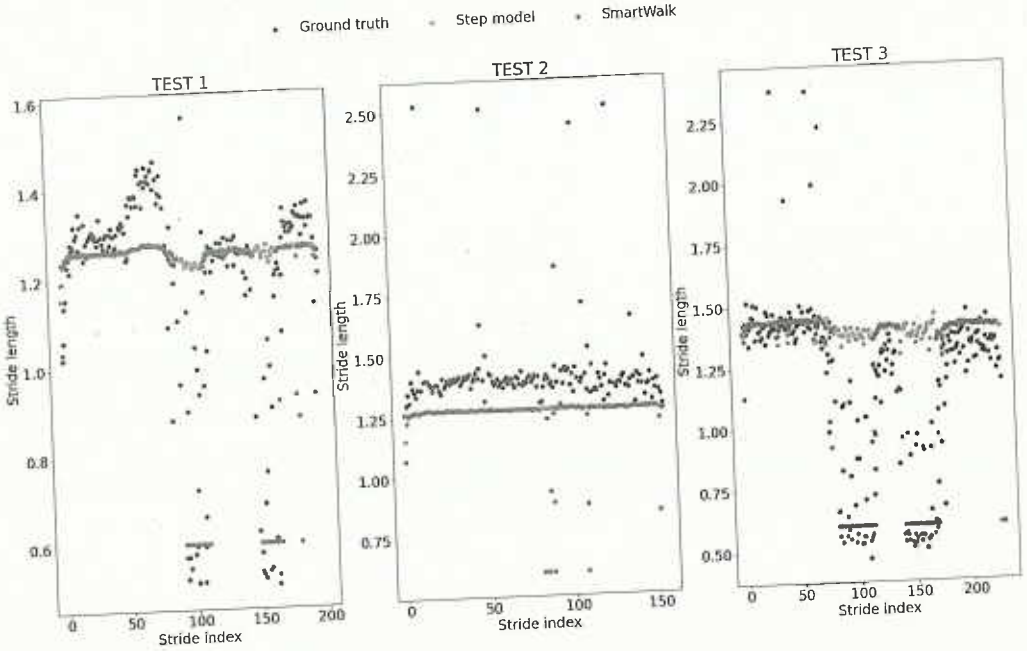


Fig. 13: Volunteer 1: Stride length predicted by SmartWalk's step model (orange), by step model corrected by an EKF (green) and the ground truth (blue)

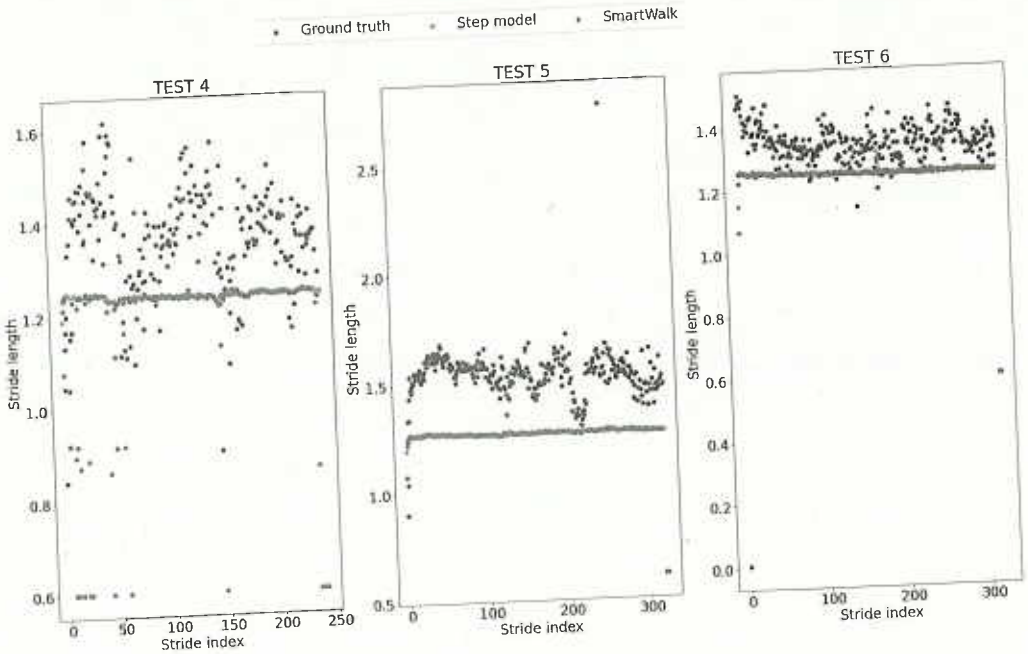


Fig. 14: Volunteer 2: Stride length predicted by SmartWalk's step model (orange), by step model corrected by an EKF (green) and the ground truth (blue)

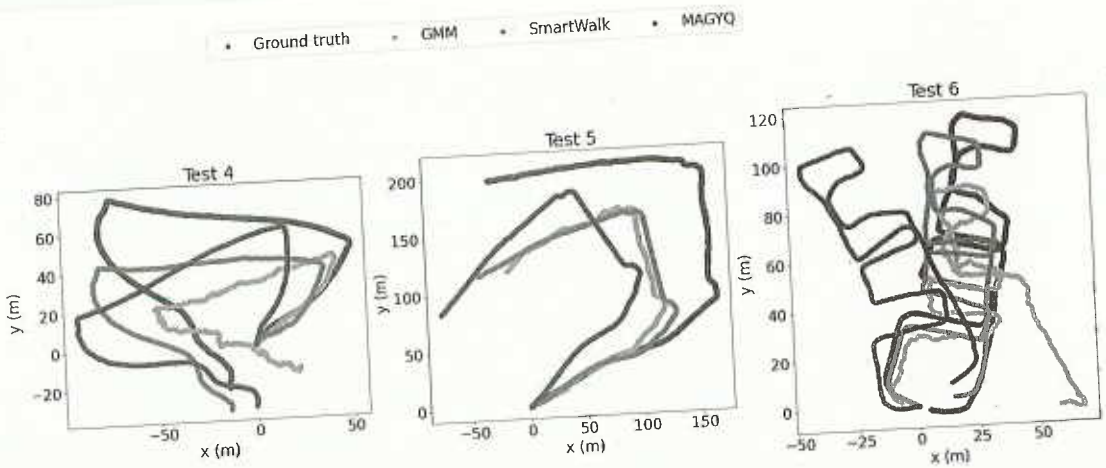


Fig. 15: Volunteer 2: Effeciency of the EKF correction: Ground truth (blue), GMM (orange), corrected GMM by MAGYQ (green) and MAGYQ (red)

on 6 testing tracks recorded with 2 volunteers in various environments, over a 2.17 km walking distance.

Both categories are facing challenges. Gait-driven AI methods need to improve their robustness to deal with different device poses and user dynamics. Sampling frequency-driven AI methods need to reduce noises in their predictions. A direction for improvement is to fuse the two approaches to capture the user's dynamics with a fixed sampling frequency-based processing and correct the trajectory using estimated gait parameters. Finally, customizing the model parameters for each user is another promising strategy.

REFERENCES

[1] H. Fu, Y. Kone, V. Renaudin, and N. Zhu, "A survey on artificial intelligence for pedestrian navigation with wearable inertial sensors," in *2022 IEEE 12th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2022, pp. 1–8.

[2] N. A. Abiad, Y. Kone, V. Renaudin, and T. Robert, "Smartstep: a robust step detection method based on smartphone inertial signals driven by gait learning," *IEEE Sensors Journal*, vol. 22, no. 12, pp. 12 288 – 12 297, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9761967

[3] Q. Wang, L. Ye, H. Luo, A. Men, F. Zhao, and C. Ou, "Pedestrian walking distance estimation based on smartphone mode recognition," *Remote Sensing*, vol. 11, no. 9, 2019. [Online]. Available: https://www.mdpi.com/2072-4292/11/9/1140

[4] I. Klein and O. Asraf, "Stepnet—deep learning approaches for step length estimation," *IEEE Access*, vol. 8, pp. 85 706–85 713, 2020.

[5] M. Zhang, Y. Wen, J. Chen, X. Yang, R. Gao, and H. Zhao, "Pedestrian dead-reckoning indoor localization based on os-elm," *IEEE Access*, vol. 6, pp. 6116–6129, 2018.

[6] N.-Y. Liang, G. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *IEEE Transactions on Neural Networks*, vol. 17, pp. 1411–1423, 2006.

[7] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 785–794. [Online]. Available: http://doi.acm.org/10.1145/2939672.2939785

[8] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," *Advances in neural information processing systems*, vol. 30, pp. 3146–3154, 2017.

[9] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.

[10] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, pp. 81–106, 1986.

[11] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *IN PROCEEDINGS OF THE THIRTEENTH INTERNATIONAL CONFERENCE ON MACHINE LEARNING*. Morgan Kaufmann, 1996, pp. 148–156.

[12] C. Cortes and V. Vapnik, "Support vector networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.

[13] F. Gu, K. Khoshelham, C. Yu, and J. Shang, "Accurate step length estimation for pedestrian dead reckoning localization using stacked autoencoders," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 8, pp. 2705–2713, 2019.

[14] D. Yan, C. Shi, and T. Li, "An improved pdr system with accurate heading and step length estimation using handheld smartphone," *Journal of Navigation*, vol. 75, no. 1, p. 141–159, 2022.

[15] A. Krizhevsky, "Learning multiple layers of features from tiny images," Tech. Rep., 2009.

[16] Q. Wang, L. Ye, H. Luo, A. Men, F. Zhao, and Y. Huang, "Pedestrian stride-length estimation based on lstm and denoising autoencoders," *Sensors*, vol. 19, no. 4, 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/4/840

[17] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[18] Q. Wang, H. Luo, L. Ye, A. Men, F. Zhao, Y. Huang, and C. Ou, "Personalized stride-length estimation based on active online learning," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 4885–4897, 2020.

[19] Q. Wang, H. Luo, L. Ye, A. Men, F. Zhao, and Y. Huang, "Pedestrian heading estimation based on spatial transformer networks and hierarchical lstm," *IEEE Access*, vol. 7, pp. 162 309–162 322, 2019.

[20] H. Yan, Q. Shan, and Y. Furukawa, "Ridi: Robust imu double integration," 12 2017.

[21] S. Herath, H. Yan, and Y. Furukawa, "Ronin: Robust neural inertial navigation in the wild: Benchmark, evaluations, amp; new methods," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 3146–3152.

[22] J. Rose, J. G. Gamble, and J. M. Adams, "Human walking." Philadelphia: Lippincott Williams and Wilkins, 2006, p. 2.

[23] J. Perry and J. M. Burnfield, "Gait analysis." SLACK Incorporated, 2012, ch. 7.

[24] H. Weinberg, "Using the adxl 202 in pedometer and personal navigation applications," 2002.

[25] Q. Ladetto, "On foot navigation: Continuous step calibration using both complementary recursive prediction and adaptive kalman filtering," *Proceedings of ION GPS*, 01 2000.

[26] W. Pirker and R. Katzenschlager, "Gait disorders in adults and the elderly: A clinical guide," *Wiener klinische Wochenschrift*, vol. 129, 10 2016.

[27] C. Combettes and V. Renaudin, "Comparison of misalignment estimation techniques between handheld device and walking directions," in *2015*

International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2015, pp. 1–8.
[28] K. Kunze, P. Lukowicz, K. Partridge, and B. Begole, "Which way am i facing: Inferring horizontal device orientation from an accelerometer signal," in *2009 International SymposiuEnhancing the Performance of Pedometers Using a Single Accelerometerm on Wearable Computers*, 2009, pp. 149–150.
[29] Z.-A. Deng, G. Wang, Y. Hu, and D. Wu, "Heading estimation for indoor pedestrian navigation using a smartphone in the pocket," *Sensors*, vol. 15, no. 9, pp. 21 518–21 536, 2015. [Online]. Available: https://www.mdpi.com/1424-8220/15/9/21518
[30] M. Chowdhary, M. Sharma, Noida, A. Kumar, S. Dayal, and M. Jain, "Method and apparatus for determining walking direction for a pedestrian dead reckoning process," 2014.
[31] M. Kourogi and T. Kurata, "A method of pedestrian dead reckoning for smartphones using frequency domain analysis on patterns of acceleration and angular velocity," in *2014 IEEE/ION Position, Location and Navigation Symposium - PLANS 2014*, 2014, pp. 164–168.
[32] J. Scarlett, "Enhancing the performance of pedometers using a single accelerometer," Available: https://www.analog.com/en/analog-dialogue/articles/enhancing-pedometers-using-single-accelerometer.html. Accessed Apr. 25, 2022 [Online], Mar. 2007.
[33] Archeries, "Selda dataset," Available: https://github.com/Archeries/StrideLengthEstimation, Accessed Apr. 25, 2022 [Online], Jan. 11 2019.
[34] Y. Kone, N. Zhu, and V. Renaudin, "Zero velocity detection without motion pre-classification: Uniform ai model for all pedestrian motions (umam)," *IEEE Sensors Journal*, pp. 1–1, 2021.
[35] N. Zhu, V. Renaudin, M. Ortiz, Y. Kone, C. Ichard, S. Ricou, and F. Gueit, "Foot-mounted ins for resilient real-time positioning of soldiers in non-collaborative indoor surroundings," in *2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2021, pp. 1–8.
[36] S. Herath, H. Yan, and Y. Furukawa, "Ronin implementation and dataset," Available: https://github.com/Sachini/ronin, Accessed Apr. 25, 2022 [Online], Jan. 13 2022.
[37] "Android game rotation vector documentation," https://developer.android.com/reference/android/hardware/Sensor#TYPE_GAME_ROTATION_VECTOR. Accessed Apr. 25, 2022 [Online].
[38] V. Renaudin and C. Combettes, "Magnetic, acceleration fields and gyroscope quaternion (magyq)-based attitude estimation with smartphone sensors for indoor pedestrian navigation," *Sensors*, vol. 14, no. 12, pp. 22 864–22 890, 2014. [Online]. Available: https://www.mdpi.com/1424-8220/14/12/22864
[39] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015. [Online]. Available: https://arxiv.org/abs/1512.03385
[40] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *CoRR*, vol. abs/1803.01271, 2018. [Online]. Available: http://arxiv.org/abs/1803.01271
[41] J. Perul and V. Renaudin, "Head: smooth estimation of walking direction with a handheld device embedding inertial, gnss, and magnetometer sensors," *Annual of Navigation*, vol. 67, pp. 713–726, 2020.
[42] V. Renaudin, Y. Kone, H. Fu, and N. Zhu, ""physics" vs "brain": Challenge of labeling wearable inertial data for step detection for artificial intelligence," in *2022 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, 2022, pp. 1–4.
[43] V. Renaudin, M. Susi, and G. Lachapelle, "Step length estimation using handheld inertial sensors," *Sensors*, vol. 12, no. 7, pp. 8507–8525, 2012. [Online]. Available: https://www.mdpi.com/1424-8220/12/7/8507
[44] V. Renaudin, V. Demeule, and M. Ortiz, "Adaptative pedestrian displacement estimation with a smartphone," in *International Conference on Indoor Positioning and Indoor Navigation*, 2013, pp. 1–9.
[45] C. Combettes and V. Renaudin, "Walking direction estimation based on statistical modeling of human gait features with handheld mimu," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 6, pp. 2502–2511, 2017. [Online]. Available: http://ieeexplore.ieee.org/document/8078215/
[46] J. Le Scornec, M. Ortiz, and V. Renaudin, "Foot-mounted pedestrian navigation reference with tightly coupled gnss carrier phases, inertial and magnetic data," in *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2017, pp. 1–8.
[47] H. Fu, "Pedestrian inertial navigation dataset with wearable sensor," 2023. [Online]. Available: https://doi.org/10.57745/ZCBIIB

Hanyuan FU received her engineer's degree in fluid mechanics and rotating machinery from Arts et Métiers ParisTech (France) in 2017 and worked as a mechanical engineer in the automotive industry. In 2021, she received her master's degree in autonomous mobile systems from Université d'Evry (France). Currently, she's a Ph.D. student in laboratory GEOLOC, Université Gustave Eiffel (France). Her research interests are inertial pedestrian navigation and individual walking characteristics.

Valérie RENAUDIN (Member, IEEE) received the M.Sc. degree in geomatics engineering from ESGT in 1999 and the Ph.D. degree in computer, communication, and information sciences from EPFL in 2009. She was the Technical Director at Swissat Company, Switzerland, developing real-time geopositioning solutions based on a permanent global navigation satellite system (GNSS) network, and a Senior Research Associate with the PLAN Group, University of Calgary, Canada. She is the Research Director at Gustave Eiffel University. She is currently leading the GEOLOC Laboratory, and the team dedicated to positioning and navigation for travelers in multimodal transport. Her research interests include outdoor/indoor navigation using GNSS, and inertial and magnetic data, particularly for pedestrians. She has been a member of IEEE Society since 2013, of the steering committee of the "Indoor Positioning and Indoor Navigation" international conference and the founding editor in Chief of the IEEE Journal of Indoor and Seamless Positioning and Navigation.

Yacouba KONE (Member, IEEE) received the M.Sc. degree in data sciences modelization statistics from Bretagne Sud University in 2019. He worked as a consulting data analyst/manager at the world bank in Burkina Faso in 2017, on the project RBF (result-based financing). He is a Research Engineer with the GEOLOC Laboratory, Gustave Eiffel University.

Ni ZHU is a researcher at the laboratory GEOLOC of University Gustave Eiffel. She received her engineering degree in aeronautic telecommunications from the Ecole Nationale de l'Aviation Civile (ENAC) in 2015 and the Ph.D. degree in science of information and communication from the University of Lille in 2018. Her recent research is specialized in GNSS channel propagation modeling in urban environments, positioning integrity monitoring for terrestrial safety-critical applications, multi-sensory fusion techniques for indoor/outdoor pedestrian positioning assisted by artificial intelligence.