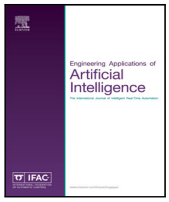
Engineering Applications of Artificial Intelligence 123 (2023) 106066

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai



HINNet: Inertial navigation with head-mounted sensors using a neural network

Xinyu Hou, Jeroen H.M. Bergmann *

Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, United Kingdom

ARTICLE INFO

Dataset link: HINNet_Dataset (Original data)

Keywords:
Machine learning
Inertial navigation
Pedestrian Dead Reckoning
Deep neural network
Inertial measurement unit
Wearable sensors

ABSTRACT

Human inertial navigation systems have been developing rapidly in recent years, and it has shown great potential for applications within healthcare, smart homes, sports, and emergency services. Placing inertial measurement units on the head for localisation is relatively new. However, it provides a very interesting option, as there are several everyday head-worn items that could easily be equipped with sensors. Yet, there remains a lack of research in this area and currently no localisation solutions have been offered that allow for free head-rotations during long periods of walking. To solve this problem, we present HINNet, the first deep neural network (DNN) pedestrian inertial navigation system allowing free head movements with head-mounted inertial measurement units (IMUs), which deploys a 2-layer bi-directional LSTM. A new 'peak ratio' feature is introduced and utilised as part of the input to the neural network. This information can be leveraged to solve the issue of differentiating between changes in movements related to the head and those that are associated with the walking pattern. A dataset with 8 subjects totalling 528 min has been collected on three different tracks for training and verification. The HINNet could effectively distinguish head rotations and changes in walking direction with a distance percentage error of 0.46%, a relative trajectory error of 3.88 m, and a absolute trajectory error of 5.98 m, which outperforms the current best head-mounted Pedestrian Dead Reckoning (PDR) method.

1. Introduction

Recent advances in sensor technologies and algorithms have enabled the development of localisation services that can be used for human monitoring. These algorithms aim to determine the position of an individual within a "global" reference frame. The most obvious applications within this space are focused on tracking for navigation purposes or for monitoring. These applications can be used in different environments, such as at home or in the hospital (Romme et al., 2014), as well as for keeping track of emergency services personnel (Ferreira et al., 2017). However, the unpredictability of human behaviour in dynamic environments makes this a complex task, especially for indoor/outdoor environments where there is little or no technical infrastructure.

Inertial navigation system (INS) is a human position tracking technology, which is agnostic to the exact infrastructure that is available in a specific environment. Unlike localisation methods with GNSS (Hsu et al., 2016), WiFi (Zhuang and El-Sheimy, 2015), cellular (Abdallah et al., 2022), Bluetooth (Yadav et al., 2019), Radio Frequency Identification (RFID) (Huang et al., 2016), Ultra-wideband (UWB) (Corrales et al., 2008), infrared (Namie and Nakagawa, 2013) or radar (Zhou et al., 2010), which all rely on external aiding signals, information, or infrastructure, the INS is a self-contained navigation technique and only requires Inertial Measurement Units (IMU). This makes it possible to navigate in places where external signals are severely affected or where there is no infrastructure available. Unfortunately, some of the use cases, such as those that can be found in healthcare, sports or emergency services, also come with a potential additional level of behavioural complexity. The INS is more sensitive to variations in behavioural patterns, as it builds-up the positional estimated from the information that is extracted from locomotion. These estimates can become more less accurate, as gait patterns become more diverse.

Traditional pedestrian inertial navigation methods utilise integration of measurements from accelerometers and gyroscopes, or apply model-based pedestrian dead reckoning (PDR). For example, Hou and Bergmann (2020a) adopted peak detection to detect steps, combined with a complementary filter (Mahony et al., 2008) to determine orientations and a Weinberg model to subsequently detect step lengths with head-mounted sensors. Kang and Han (2015) and Park et al. (2021) used IMUs in smartphones to track users also with a step and heading system, while (Hardegger et al., 2015) used a single foot-mounted IMU and a hip-worn smartphone for tracking with a zero velocity update, action recognition algorithm and Particle filter. Goyal

* Corresponding author.
E-mail address: jeroen.bergmann@eng.ox.ac.uk (J.H.M. Bergmann).

https://doi.org/10.1016/j.engappai.2023.106066
Received 20 September 2022; Received in revised form 31 January 2023; Accepted 25 February 2023
Available online 27 March 2023

et al. (2011) used extended Kalman filter (EKF) and Weinberg model with a waist-worn IMU. Foot-mounted IMUs can also be used, as shown by Fan et al. (2019), indicating that a range of possible sensor placements are available for inertial navigation. Diaz et al. (2019) provided a good overview of the methods in inertial tracking, with a particular focus on indoor tracking. It highlighted the ongoing problem caused by the accumulated drift error on the estimated position that is obtained by the sensors. A variety of calibration methods have been invented to increase the accuracy of pedestrian inertial navigation, such as the (Foxlin, 2005), Zero Angular Rate Update(ZARU) (Rajagopal, 2008), Heuristic Drift Reduction (HDR) (Borenstein et al., 2009). Although, those methods may have accurate results under experimental environments, their performance can reduce quickly in real-world scenarios with longer run times. Performance is also negatively affected if there are more varied users, if there is an increased unpredictability of the environment or if the user activities themselves becomes less predictable. Because model-based pedestrian inertial navigation methods represent the functional relationships between inertial measurement inputs and estimated localisation information, outputs with simple fitting equations with different parameters are needed for each person. This means the output suffers from sensor drifts and accumulated errors. For example, Weinberg step length model estimates the step length only with a fourth root of the difference between the maximum and minimum vertical accelerations in one step, which multiplies with a parameter k. This will be different for each participant (Weinberg, 2002).

Recently, machine learning has proved its impressive potential in solving tasks in a variety of fields like natural language processing (Devlin et al., 2018), image processing (Zhou et al., 2022), and healthcare (Porebski et al., 2018; Hou et al., 2019; Qi and Su, 2022). One of the key benefits consists of the ability to model complex non-linear relationships with large volumes of data. The precise identification of patterns or trends indicates the suitability for it to be utilised in pedestrian inertial navigation. Chen et al. (2018) adopted a 2-layer Bi-LSTM to learn human odometry with IMUs that were either hand-held, in a pocket, in a handbag or on a trolley. Herath et al. (2020) estimated horizontal positions and heading direction of a moving subject from a sequence of IMU sensor measurements using a phone with ResNet, long short-term memory (LSTM) and Temporal Convolutional Network (TCN). The feasibility and performance of human odometry estimation with machine learning have been demonstrated in these previous published studies. It also showcases the ability to collect this data with different sensor placements. The selection of placement should be routed in the applicability of that sensor for use in complex real-world activities.

The head is a location that provides potential universal applicability for sensor placement, as there are several everyday worn objects that could easily be merged with sensors. According to a survey about where people would like to wear sensor technologies, locations that would allow the sensor system to be small, discreet, and unobtrusive were preferred, with many people referring back to everyday objects. Bergmann et al. (2012) It also provides an acceptable location for use in healthcare, sports, or work. A clear example for sensor integration is the emerging use of smart glasses. Glasses are a requisite to people who need to correct for certain visual impairments, and there are already several smart glasses products on the market, such as Google Glasses, Vuzix Blade, Epson Moverio BT-300, etc. The field of smart glasses also links in well with the growing interest in providing tracking in virtual reality (VR) or augmented reality (AR). Glasses are also used in many industries to protect the eyes. Protective gear is often located on the head. Helmets, masks and mouthguards are some of the obvious devices that are used both in industry and sports. Smart mouthguards are already being equipped with sensors for a range of purposes (Davies et al., 2018; Hou and Bergmann, 2020a). For contact sports, watches or other wearable objects are not allowed, as it could cause injuries and therefore form a risk to players. However, the mouthguard is often required to be worn in order to protect the teeth, which makes it a very suitable piece of equipment for sensor integration. This integration would provide an unobtrusive, head-mounted sensor system that can be worn on-field. There is a range of other objects that can be used as well, such as earrings, earphones, hearing aids, caps, etc. This makes a head-mounted system very generalisable across use cases.

However, according to a recent systematic review, there are only a few papers (Hasan and Mishuk, 2018; Zhu et al., 2014) that are exploring inertial navigation with head-mounted sensors (Hou and Bergmann, 2020b). That can be attributed to the issue that rotational head motions (used to change the visual field) could be mistaken with body rotations (related to a change in walking direction). This is not or less of a problem for sensor attachments to the feet, waist or chest. The "heading" direction of the head may not always align with the body orientation, which can lead to an obvious position error in pedestrian localisation. The three papers with head-mounted sensors mentioned above are all based on traditional pedestrian inertial navigation methods with step-and-heading systems, which lack generalisability in more complex real world scenarios. None of them solved the problem of the misalignment of the head and body orientation, as they simply ignored head rotations during walking.

To solve this interesting problem, we propose the first deep neural network (DNN) framework for pedestrian inertial navigation with head-mounted sensors allowing for free head movements: HINNet. Previous studies on pedestrian navigation with head-mounted sensors all depended on traditional pedestrian inertial navigation methods with step-and-heading models, and only work when the head is facing forward. This paper is the first study to propose deep learning methods in pedestrian navigation with head-mounted sensors, and allow for free head rotations during walking. In this study, the relative trajectory error (RTE) and percentage for the HINNet and PDR are determined against a ground truth measurement during long-period walking trials. This paper introduces a DNN with a set of new features to address the issue of separating the head rotation signals from the walking data for head-mounted IMUs. This information is then used for positional tracking.

2. Methods

The HINNet system is summarised in Fig. 1. It provides an overview of how data from the head-mounted sensor is processed to estimate trajectories.

2.1. Roll and pitch compensation

The raw IMU sensor data was first transformed into a normalised coordinate system in which the z-axis is perpendicular to the horizontal plane, which was accomplished by Eq. (1) and Eq. (3), where a and ω are $3 * length_{data}$ vectors representing accelerometer data and gyroscope data.

$$a_{norm} = R_a^{-1} \cdot a_{raw} \tag{1}$$

$$R_a = R_x(\phi)R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ \sin\phi\sin\theta & \cos\phi & \sin\phi\cos\theta \\ \cos\phi\sin\theta & -\sin\phi & \cos\phi\cos\theta \end{bmatrix} \tag{2}$$

$$\omega_{norm} = R_\omega^{-1} \cdot \omega_{raw} \tag{3}$$

$$R_\omega = \begin{bmatrix} 1 & 0 & -\sin\theta \\ 0 & \cos\phi & \sin\phi\cos\theta \\ 0 & -\sin\phi & \cos\phi\cos\theta \end{bmatrix} \tag{4}$$

After compensation, the normalised z-axis of the IMU will be aligned with the gravity direction. The heading direction (yaw angle ψ) is considered in reference to the angular velocity ω_z and there is no expected gravity component on the normalised x-axis and y-axis, which will be used later for peak ratio calculation in Section 2.2.

In this part, the input $(a_{raw}, \omega_{raw})_6$ is the 6-dimensional raw IMU data, the output $(a_{norm}, \omega_{norm})_6$ is the 6-dimensional normalised IMU data.

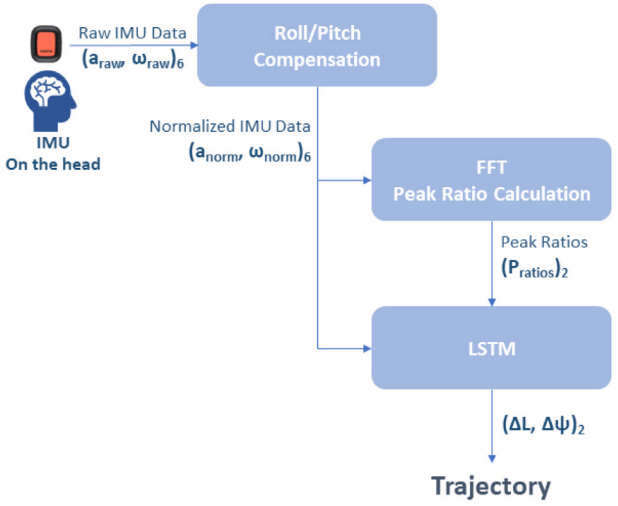


Fig. 1. Overview of the HINNet, a deep neural network for pedestrian inertial navigation with head-mounted sensors that allows for free head movements. The Inertial Measurement Unit (IMU) placed on the head is shown in the top left corner. After normalisation, the IMU Data is first transformed into frequency domain with a fast Fourier transform (FFT) to generate the peak ratio feature. The peak ratios and normalised IMU data form the input into the Long Short-Term Memory (LSTM) model.

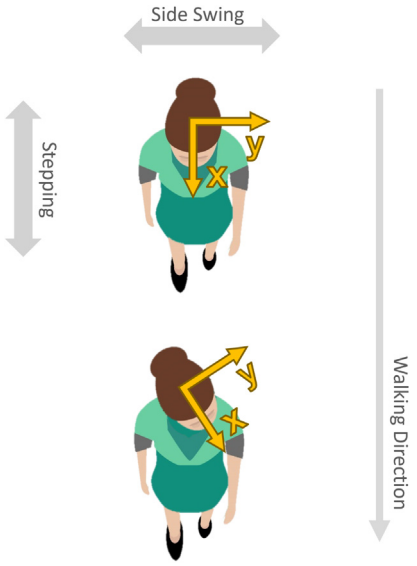


Fig. 2. The images show a person walking forward at two different arbitrary time points. The top image shows them with the head facing in the same direction as they are walking. The bottom image still depicts a forward walking direction, but with the head rotated to the left. Different magnitudes of side swing and stepping are projected on the x and y axes of the head-mounted IMU depending on the orientation of the head.

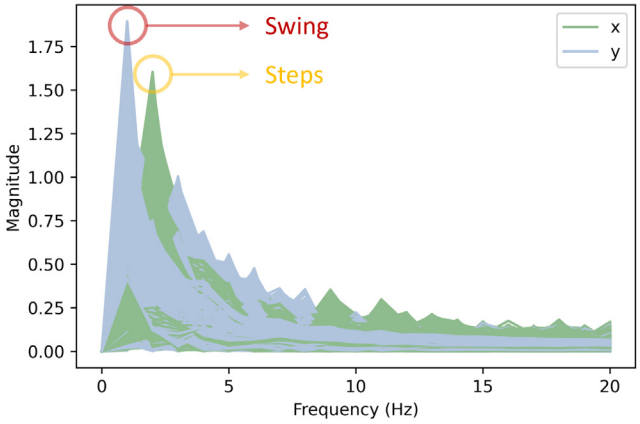


Fig. 3. Frequency spectrum of accelerations obtained from the x and y axes. This information was produced by a Fast Fourier Transform (FFT) taken from the normalised accelerations. The peak in the red circle is generated by swings and the peak in the yellow circle is produced by steps.

2.2. Feature of peak ratio

The key is to be able to distinguish the head rotation from the whole body rotation. This can be done by finding the difference in the walking patterns, which are recorded by the head-mounted IMU. There are two obvious kind of body movements during walking that can generate regular acceleration waves. They are generated by (i) the body swinging from left to right, (ii) the variation in linear acceleration in the front-to-back direction caused by stepping. The change in signals that are recorded on the x and y axis when the head is rotating is shown in Fig. 2. Only stepping will be recorded on x axis when the head is held "straight" during walking. Alternatively, both stepping and side swing will be measured on x axis when walking with the head facing sideways. This concept is used to help determine whether (i) a rotation occurred due to a head movement without any change in walking direction, or (ii) if the head relative to body remained aligned and a change in walking direction took place (Windau and Itti, 2016).

One full wave of a swing requires one stride (two steps) and a full stepping wave requires one step. The frequency difference between stepping and side swings gives us the possibility to visualise these two motions in the frequency domain. The spectrums of accelerometer sensor data in the horizontal plane (x-axis and y-axis) are shown in Fig. 3.

The relation of the two peaks in Fig. 3 can be represented by the peak ratio P_{ratio} in Eq. (5) (Windau and Itti, 2016).

$$P_{ratio} = \frac{P_{swing}}{P_{stepping}} \tag{5}$$

The applied method for this paper consists of the following steps: For each time point t, the accelerations on the x axis in a $2s$ window $[t-2s, t)$ prior to t are transformed into the frequency domain using a FFT. Then the two peaks P_{swing} and $P_{stepping}$ are extracted from the generated frequency spectrum. These are subsequently used to calculate a P_{ratio}^{before}, which represents the 2 s prior to t. The same operation is conducted on the $2s$ window $(t, t+2s]$ after time point t, for which we get a P_{ratio}^{after}. The gyroscope data is used to determine if a rotation took place at time point t. If the P_{ratio}^{before} and P_{ratio}^{after} are very similar then it is likely that the walking direction has changed. Essentially, only the "stepping" information should be projected on the x axis before and after rotation (see top image in Fig. 2). In the case of pure head rotation, the "side" swing projection on x axis will alter the P_{ratio}^{after}, which will create a different outcome when compared to the P_{ratio}^{before} (see bottom image in Fig. 2).

The $(P_{ratios})_2$ (see Fig. 1) consists of both the P_{ratio}^{after} as well as the P_{ratio}^{before}. This information is then used in the LSTM.

2.3. Deep neural network framework

After pre-processing and feature generation, a sequence of normalised IMU data and peak ratio features were fed into the neural network as input. The recurrent neural network (RNN) would be the best choice for this task, since it was designed to process sequential data or time series data, and has already been successfully applied in e.g. speech recognition (Sak et al., 2014) and machine translation (Wu et al., 2016). The LSTM was leveraged in this research instead of vanilla RNN to prevent the vanishing gradient problem. The LTSM is a special kind of RNN, capable of learning long-term dependencies (Hochreiter

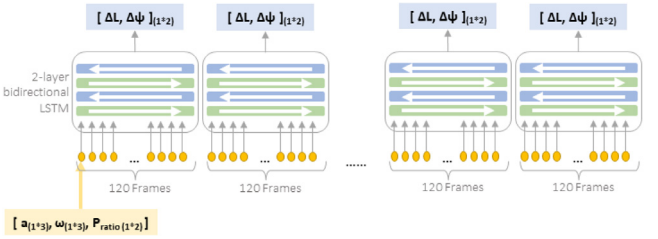


Fig. 4. Overview of HINNet neural network. In each window, 120 frames of normalised IMU data and peak ratios are sequentially inputted into a 2-layer bidirectional LSTM, which outputs the displacement and orientation variation.

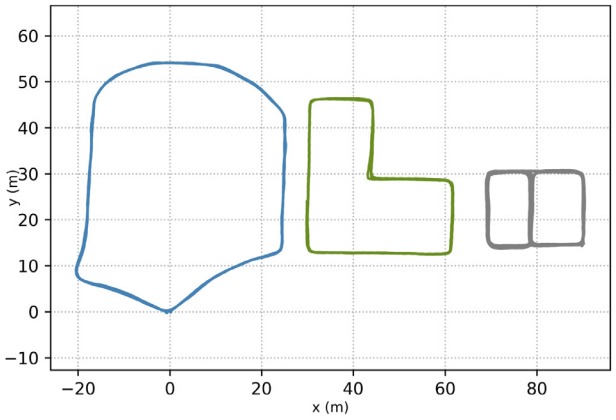


Fig. 5. The exact map of each of the walked paths is shown in this figure. The volunteers were asked to follow these paths during data collection.

and Schmidhuber, 1997). The LSTM works as a function f_θ mapping normalised accelerometer data a, gyroscope data ω and peak ratio features P_{ratios} to walking distance Δl and rotation $\Delta\psi$ over a window,

$$(a_{norm}, \omega_{norm}, P_{ratios})_{8*120} \xrightarrow{F_\theta} (\Delta l, \Delta\psi)_{2*1} \tag{6}$$

where a window length of 120 frames ($2s$) was used. To leverage each sample's past and future contexts, a bi-directional LSTM network was adopted in the study, with 2-layers stacked to add levels of abstraction of input observations over time. Each layer had 128 nodes. The framework overview is shown in Fig. 4.

After getting $(\Delta l, \Delta\psi)_2$ in every window, the whole trajectory could be generated by connecting them sequentially.

3. Experiment conditions

3.1. Data collection site

The data collection was conducted in three different environments with different sizes and paths to ensure a broader ability to generalise. Fig. 5 shows the shapes and dimensions of trajectories of the three scenarios. The trajectories include straight routes, curves, and turnings with different angles, which increases the complexity and applicability of the tests.

3.2. Participants

There were 8 volunteers participating in the experiments, with the demographic information of the participants given in Table 1. According to the result from a recent systematic review (Hou and Bergmann, 2020b), 128 out of 145 papers about human inertial navigation used less than 8 subjects. The amount of subject in this study exceeds 88% papers in this domain. All participants signed a consent form before the data collection started, and they were given the opportunity to ask any questions before deciding to be involved in this study. Ethical approval was obtained from the University ethics committee and this experiment was part of a larger study (R70833/RE001).

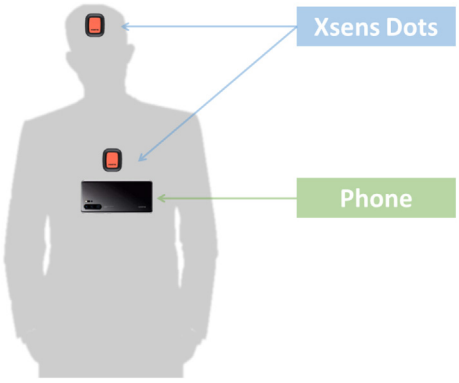


Fig. 6. Schematic of the sensor placement. A IMU (Xsens Dot sensor) was attached on the head using a strap, while another IMU is attached by applying a chest harness. This harness also holds a phone to collect positional reference data.

Table 1
Demographics of participants. (M = male, F = female)

Subject	Age (years)	Height (m)	Weight (kg)	Gender
1	42	1.79	73	M
2	27	1.65	60	F
3	25	1.79	77	M
4	24	1.77	62	F
5	48	1.75	85	M
6	49	1.60	58	F
7	30	1.70	66	M
8	21	1.65	65	F

3.3. Devices

Xsens Dot sensors (Xsens Technologies BV, Enschede, Netherlands) were used in the experiments. They logged the 3D angular velocity from a gyroscope, 3D acceleration from an accelerometer and 3D (local) magnetic field from a magnetometer. Data was collected at 60 Hz and transferred to a mobile phone by Bluetooth. An Xsens Dot was firmly attached to the forehead with a strap and another one was placed on the chest for reference. The placement is shown in Fig. 6. A phone was also firmly attached on the chest with straps.

An application developed in Unity 3D game engine (Unity Technologies, San Francisco, CA, USA) based on Google ARCore was installed in the phone for ground truth generation. The ground truth trajectories in Fig. 5 were generated using this information.

3.4. Experimental setup

Before the experiment started, participants were asked to put on the straps with sensors. They were requested to place one Xsens Dot on their head in such a way that it remained comfortably in contact with the head during walking. In each test, subjects were requested to walk the predetermined trajectories for 5 to 10 min. Each volunteer was tested between 4 to 16 times. In all the tests participants were instructed to walk at a normal and constant, speed whilst rotating their head in a random manner. A total of 79 datasets were collected across all experiments, with a total time of around 528 min. Each dataset contained both the data from the Xsens Dots, as well as from the phone.

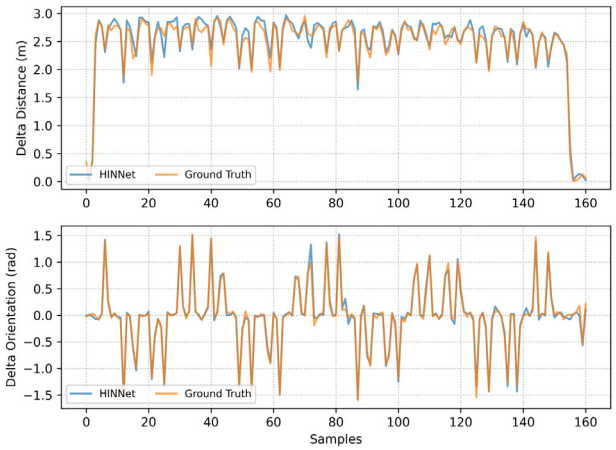


Fig. 7. The walking distance in each 2 *s* window (Delta Distance) and the variation of orientation in each 2 *s* window (Delta Orientation) from one of the test dataset. Orange lines represent the values from ground truth. Blue lines show the values estimated by HINNet.

4. Results

A recently published PDR system (Hou and Bergmann, 2020a) was selected as current baseline method for comparison, which is based on traditional step-and-heading systems, with peak detection to detect steps, complementary filter (Mahony et al., 2008) to determine orientations and Weinberg model to subsequently detect step lengths with head-mounted sensors. This particular method is the most accurate inertial tracking method available for head-mounted sensors, as the algorithm has been adapted to this specific sensor location, as shown in Hou and Bergmann (2020a). Current PDR algorithms are not agnostic to sensor placement, since the signals can drastically change depending on where the sensor is attached to the body. Percentage errors of total distances, relative trajectory errors (RTE) and absolute trajectory errors (ATE) were computed for the PDR and HINNet. RTE and ATE are standard position evaluation metrics in navigation (Zhang and Scaramuzza, 2018). RTE IS defined as the average root mean square error (RMSE) over a fixed time interval of 1 min. ATE is the RMSE between the whole ground truth trajectory and the estimated trajectory. These particular metrics have been widely utilised in other tracking studies, thus deemed to be a suitable output for this research.

Fig. 7 shows the estimated and real walking distance (Delta Distance) and variation of orientation (Delta Orientation) in each 2s from one test dataset.

The estimated trajectories of different methods under the three scenarios are shown in Fig. 8.

RTE, ATE and percentage error of total distance of each method were summarised in Table 2.

An ablation study was conducted to investigate the contribution of peak ratio features in the neural network. A model without peak ratio feature input was trained with the same parameters as the one with the features. Results are shown in Table 3.

5. Discussion

We presented the first pedestrian inertial navigation system for a head-mounted IMU that allows for free head-movements by applying a DNN. HINNet performs much better than a dedicated PDR on walking distance estimation, with a distance error of 0.46%. The PDR requires tuning of specific parameters for each individual, while HINNet does not need human input to tune any parameters manually. The difference in performance can also be observed in the lower RTE and ATE of HINNet compared to that of the PDR. However, the biggest improvement accomplished by HINNet is possibility to eliminate head motion that might be due to "sightseeing" while walking. In Fig. 8, there are a lot of small curves in the PDR trajectories. These represent the head motions that are not related to the walking and thus they lead to false additional trajectories. While the trajectories estimated by HINNet are a lot more similar to the shape of the ground truth, which implies its ability to distinguish head rotations that are independent of walking. These promising results need to be considered carefully due to the relative small sample size.

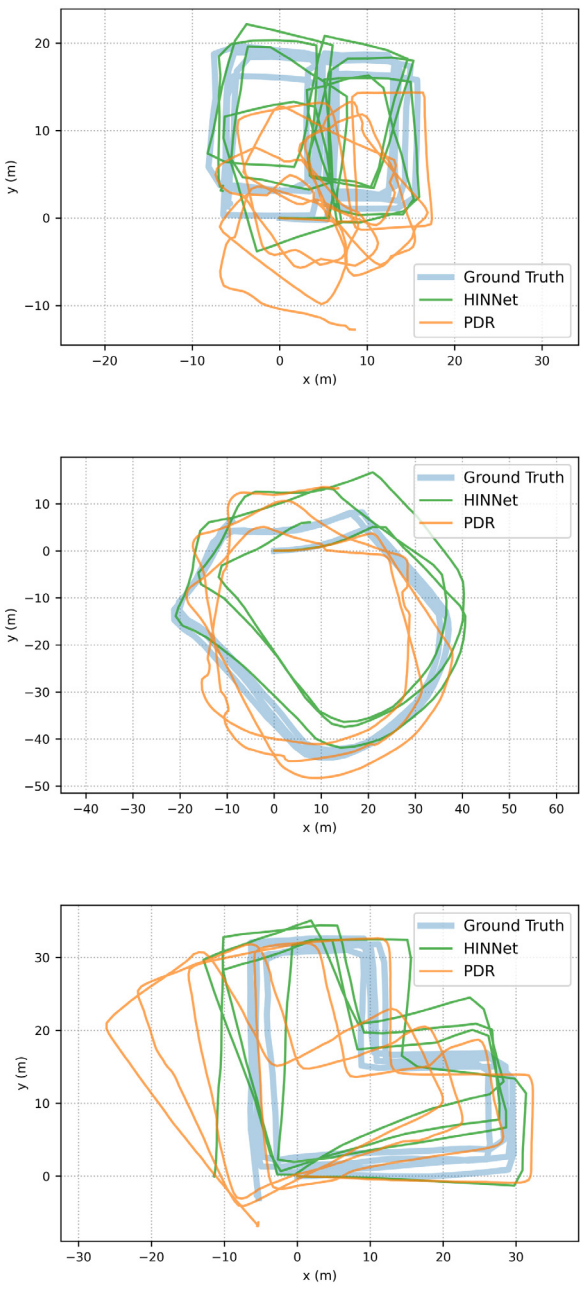


Fig. 8. Estimated trajectories of different methods. The light blue lines represent the ground truths. Green lines show trajectories estimated by the HINNet. Orange lines are the trajectories generated by the PDR.

The curves that are present in the PDR trajectories have different sizes and frequencies, which indicates the variation in head rotation behaviours between different volunteers. Interestingly, the data showed that when people were specifically told to rotate their head, head rotations were more frequent with larger angles comparing to previous tests where volunteers were just asked to walk as they normally do. It

Table 2
Relative trajectory error (RTE), absolute trajectory error (ATE), and percentage error of total distances of HINNet and PDR.

Methods	RTE (m)	ATE (m)	Distance error (%)
HINNet	3.88	5.98	0.46
PDR	5.76	9.89	6.05

Table 3
Ablation study on the peak ratio feature. Relative trajectory error (RTE), absolute trajectory error (ATE), and percentage error of total distances of HINNet with and without peak ratio feature.

Methods	RTE (m)	ATE (m)	Distance error (%)
With peak ratio	3.88	5.98	0.46
Without peak ratio	4.47	7.94	0.61

indicates that the instructions given to subjects should be composed carefully to avoid implications leading to biased data. Nonetheless, for this study the additional head rotation was requested to better assess the performance between the algorithms. It should also be noted that these increased head motions might very likely when people are e.g. sightseeing or navigating busy streets.

Although HINNet could effectively differentiate head motions, limitations still exist when head and body rotate separately compared to synchronised motion. Overlaps of head and body movements lead to a confusion in estimation and can benefit from further information produced by other sensors (e.g. Magnetometer, GPS) to maintain a correct heading direction.

It should also be noted that the position error is still accumulating over time although the head rotation error has been eliminated. To improve this, additional information or complementary sensors could be leveraged, such as Bluetooth RFID, GPS, etc. Another option would consist of using self calibration methods like the HeadSLAM (Hou and Bergmann, 2022).

Increasing the dataset is important if we aim to increase the accuracy of neural inertial navigation system in the future. The scale of parameters of the DNN could then be enlarged to contain more information in a higher dimension, thus allowing more complex movements from a variety of people to be correctly estimated. In the natural language processing area, one of the biggest language model OpenAI GPT-3 has 175 billion parameters and was trained on dataset with 499 billion tokens, making its language processing ability approximate a real human (Brown et al., 2020). If a similar size of dataset would be adopted in human motion estimations, a variety of unimaginable tasks may be fulfilled by a much larger model. This study has shown that this research direction can be very promising for human position tracking.

Besides the above research directions, future studies could also focus on developing robust tracking methods to include more complex motions like running, jumping, sitting, etc. Furthermore, extending the tracking system to a three-dimensional space would also provide real-world monitoring benefits, as it would allow for incorporation of information related to e.g. stair climbing or walking on slopes.

6. Conclusion

This study highlighted the performance increase that can be obtained by considering the sensors placement and selecting suitable algorithms that can overcome identified biases in localisation that result from these placements. It showed that the HINNet could cope with head rotations for head-worn sensors outperforming the current best head-mounted Pedestrian Dead Reckoning (PDR) method.

CRediT authorship contribution statement

Xinyu Hou: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Jeroen H.M. Bergmann:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

HINNet_Dataset (Original data) (Github)

References

Abdallah, A.A., Jao, C.-S., Kassas, Z.M., Shkel, A.M., 2022. A pedestrian indoor navigation system using deep-learning-aided cellular signals and ZUPT-aided foot-mounted IMUs. IEEE Sens. J. 22 (6), 5188–5198. http://dx.doi.org/10.1109/JSEN.2021.3118695.

Bergmann, J.H., Chandaria, V., McGregor, A., 2012. Wearable and implantable sensors: The patient's perspective. Sensors 12 (12), 16695–16709.

Borenstein, J., Ojeda, L., Kwanmuang, S., 2009. Heuristic reduction of gyro drift for personnel tracking systems. J. Navig. 62 (1), 41–58.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al., 2020. Language models are few-shot learners. Adv. Neural Inf. Process. Syst. 33, 1877–1901.

Chen, C., Lu, X., Markham, A., Trigoni, N., 2018. IONet: Learning to cure the curse of drift in inertial odometry. Proc. AAAI Conf. Artif. Intell. 32 (1), http://dx.doi.org/10.1609/aaai.v32i1.12102, URL https://ojs.aaai.org/index.php/AAAI/article/view/12102.

Corrales, J.A., Candelas, F.A., Torres, F., 2008. Hybrid tracking of human operators using IMU/UWB data fusion by a Kalman filter. In: 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction. HRI, IEEE, pp. 193–200.

Davies, W., Ye, H., Bergmann, J., 2018. Unobtrusive bioanalytics for impact–related sport activities. In: International Conference on Applied Human Factors and Ergonomics. Springer, pp. 285–293.

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Diaz, E.M., Ahmed, D.B., Kaiser, S., 2019. A review of indoor localization methods based on inertial sensors. Geogr. Fingerpr. Data Create Syst. Indoor Position. Indoor/Outdoor Navig. 311–333.

Fan, Q., Zhang, H., Pan, P., Zhuang, X., Jia, J., Zhang, P., Zhao, Z., Zhu, G., Tang, Y., 2019. Improved pedestrian dead reckoning based on a robust adaptive Kalman filter for indoor inertial location system. Sensors 19 (2), 294.

Ferreira, A.F.G.G., Fernandes, D.M.A., Catarino, A.P., Monteiro, J.L., 2017. Localization and positioning systems for emergency responders: A survey. IEEE Commun. Surv. Tutor. 19 (4), 2836–2870.

Foxlin, E., 2005. Pedestrian tracking with shoe-mounted inertial sensors. IEEE Comput. Graph. Appl. 25 (6), 38–46.

Goyal, P., Ribeiro, V.J., Saran, H., Kumar, A., 2011. Strap-down pedestrian dead-reckoning system. In: 2011 International Conference on Indoor Positioning and Indoor Navigation. IEEE, pp. 1–7.

Hardegger, M., Roggen, D., Tröster, G., 2015. 3D ActionSLAM: wearable person tracking in multi-floor environments. Pers. Ubiquitous Comput. 19 (1), 123–141.

Hasan, M.A., Mishuk, M.N., 2018. Mems IMU based pedestrian indoor navigation for smart glass. Wirel. Pers. Commun. 101 (1), 287–303.

Herath, S., Yan, H., Furukawa, Y., 2020. Ronin: Robust neural inertial navigation in the wild: Benchmark, evaluations, & new methods. In: 2020 IEEE International Conference on Robotics and Automation. ICRA, IEEE, pp. 3146–3152.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.

Hou, X., Bergmann, J., 2020a. A pedestrian dead reckoning method for head-mounted sensors. Sensors 20 (21), 6349.

Hou, X., Bergmann, J., 2020b. Pedestrian dead reckoning with wearable sensors: A systematic review. IEEE Sens. J. 21 (1), 143–152.

Hou, X., Bergmann, J., 2022. HeadSLAM: pedestrian SLAM with head-mounted sensors. Sensors 22 (4), 1593.

Hou, X., You, J., Hu, P., 2019. Predicting drug-drug interactions using deep neural network. In: Proceedings of the 2019 11th International Conference on Machine Learning and Computing. pp. 168–172.

Hsu, L.-T., Gu, Y., Huang, Y., Kamijo, S., 2016. Urban pedestrian navigation using smartphone-based dead reckoning and 3-D map-aided GNSS. IEEE Sens. J. 16 (5), 1281–1293. http://dx.doi.org/10.1109/JSEN.2015.2496621.

Huang, J., Yu, X., Wang, Y., Xiao, X., 2016. An integrated wireless wearable sensor system for posture recognition and indoor localization. Sensors 16 (11), 1825.

Kang, W., Han, Y., 2015. SmartPDR: Smartphone-based pedestrian dead reckoning for indoor localization. IEEE Sens. J. 15 (5), 2906–2916. http://dx.doi.org/10.1109/JSEN.2014.2382568.

Mahony, R., Hamel, T., Pflimlin, J.-M., 2008. Nonlinear complementary filters on the special orthogonal group. IEEE Trans. Automat. Control 53 (5), 1203–1218.

Namie, H., Nakagawa, M., 2013. Development of indoor positioning system by using the infrared rays data communication tags for pedestrian navigation. IEEJ Trans. Electron. Inf. Syst. 133 (4), 713–721.

Park, S., Lee, J.H., Park, C.G., 2021. Robust pedestrian dead reckoning for multiple poses in smartphones. IEEE Access 9, 54498–54508.

Porebski, S., Porwik, P., Straszecka, E., Orczyk, T., 2018. Liver fibrosis diagnosis support using the Dempster–Shafer theory extended for fuzzy focal elements. Eng. Appl. Artif. Intell. 76, 67–79.

Qi, W., Su, H., 2022. A cybertwin based multimodal network for ECG patterns monitoring using deep learning. IEEE Trans. Ind. Inform. 18 (10), 6663–6670. http://dx.doi.org/10.1109/TII.2022.3159583.

Rajagopal, S., 2008. Personal Dead Reckoning System with Shoe Mounted Inertial Sensors (Master's Degree Project). Citeseer, Stockholm, Sweden.

Romme, J., van den Heuvel, J., Dolmans, G., Selimis, G., Philips, K., De Groot, H., 2014. Measurement and analysis of UWB radio channel for indoor localization in a hospital environment. In: 2014 IEEE International Conference on Ultra-WideBand. ICUWB, IEEE, pp. 274–279.

Sak, H., Senior, A.W., Beaufays, F., 2014. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In: INTERSPEECH. pp. 338–342.

Weinberg, H., 2002. Using the ADXL202 in pedometer and personal navigation applications. Analog Devices AN-602 Application Note, 2, (2), pp. 1–6.

Windau, J., Itti, L., 2016. Walking compass with head-mounted IMU sensor. In: 2016 IEEE International Conference on Robotics and Automation. ICRA, IEEE, pp. 5542–5547.

Wu, Y., Schuster, M., Chen, Z., Le, Q.V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., et al., 2016. Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv preprint arXiv:1609.08144.

Yadav, R.K., Bhattarai, B., Gang, H.-S., Pyun, J.-Y., 2019. Trusted k nearest Bayesian estimation for indoor positioning system. IEEE Access 7, 51484–51498.

Zhang, Z., Scaramuzza, D., 2018. A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS, IEEE, pp. 7244–7251.

Zhou, C., Downey, J., Stancil, D., Mukherjee, T., 2010. A low-power shoe-embedded radar for aiding pedestrian inertial navigation. IEEE Trans. Microw. Theory Tech. 58 (10), 2521–2528. http://dx.doi.org/10.1109/TMTT.2010.2063810.

Zhou, J., Yang, T., Chu, W., Zhang, W., 2022. Underwater image restoration via backscatter pixel prior and color compensation. Eng. Appl. Artif. Intell. 111, 104785.

Zhu, Y., Zhang, R., Xia, W., Jia, Z., Shen, L., 2014. A hybrid step model and new azimuth estimation method for pedestrian dead reckoning. In: 2014 Sixth International Conference on Wireless Communications and Signal Processing. WCSP, IEEE, pp. 1–5.

Zhuang, Y., El-Sheimy, N., 2015. Tightly-coupled integration of WiFi and MEMS sensors on handheld devices for indoor pedestrian navigation. IEEE Sens. J. 16 (1), 224–234.