A multimodal approach for quantifying walking pace using chest-worn wearable sensors

A Multimodal Approach for Quantifying Walking Pace using Chest-worn Wearable sensors

Muhammad Fahim, Richard Gault, Baharak Ahmaderaghi, Lu Bai, Anna Jurek-Loughrey
School of Electronics, Electrical Engineering, and Computer Science
Queen's University, Belfast, United Kingdom
$\{m.fahim, richard.gault, b.ahhmaderaghi, l.bai, a.jurek\}$ $@qub.ac.uk$

***Abstract*—Quantifying the walking pace of older people is considered an essential measurement when evaluating functional mobility, the ability to live independently, and a predictor of adverse events such as falls. We hypothesize that existing sensors in chest-worn wearables can be utilized to predict walking pace accurately without the need for additional wearables. However, predicting the walking pace of an older person using a single triaxial accelerometer sensor poses challenges with age impacting the generation of acceleration signals for slow, normal, and fast-paced walking. We believe that adding another modality, such as electrocardiogram (ECG) signals, in conjunction with acceleration signals, can aid in determining the walking pace of an older person. Our proposed approach consists of a feature discovery network that is based on an autoencoder. This network encodes the ECG waves and accelerometer signals into a latent representation in an unsupervised manner. It is followed by a walking discriminator network based on feed-forward neural network to predict walking pace. The experiments are performed on clinical-grade wearable sensors from a public dataset "Growing Old TOgether Validation" (GOTOV) to evaluate the performance. The proposed multi-modal approach achieved an accuracy of 82%, which is 9% higher than processing a single accelerometer sensor data alone.**

***Index Terms*—Health Aging, Artificial Intelligence, Representational Learning**

I. Introduction

Walking is considered an affordable physical activity with a low risk of injury in old age. It can improve mental and physical well-being which consequently helps people to live independently for longer years and promotes healthy aging [1]. A change in a person's walking pace is an important indicator of potential health issues and can also suggest the need for timely changes to prevent potential injuries. The pace of a person's walk has a strong association with risks of falls [2], frailty [3], and cognitive decline [4] among the older population. Suitable and reliable tools are required to quantify walking pace to monitor vital signs and assess the overall status of a person's physical wellbeing [5].

Wearable devices are ubiquitous with great potential to improve the quality of life through their numerous applications; ranging from fitness tracking [6] to behavioural analysis [7]. Multimodal sensors such as triaxial accelerometers and electrocardiograms (ECG) are essential in body and wrist-worn devices available through clinical-grade wearables to ordinary smartwatches [8]. These affordable personal sensing devices enable a continuous and real-time collection of physiological data to provide insights into the level of a person's physical activity. The existing activity recognition models have robust performance to recognize the macro-level ambulatory activities such as walking, jogging, or running [9]. These models are trained with data from healthy and young adults, which causes bias to the disadvantage of the older population. In the case of older people, sensor data associated with walking pace has less variation in activity patterns due to the aging making it challenging to quantify correctly. Sensor data reflecting normal and slow pace walking are very similar, placing a challenge for classification models to recognize it accurately. Machine learning-based classifiers struggle to differentiate these micro-level activities. Furthermore, it becomes more complex in the case of chest-worn devices, which suffer from almost identical motion patterns as compared to the wrist-worn wearables. The acceleration signal struggles to distinguish the walking pace in older adults; may that be slow, normal, or fast. We believe that an appropriate consideration of onboard sensor modalities such as ECG and robust artificial intelligence (AI) algorithms can solve this challenge.

To address the considerable aforementioned challenges, we proposed a multimodal approach that utilizes information from both ECG and triaxial accelerometer sensors to quantify the walking pace. The objective is to combine both modalities and evaluate the model performance for predicting the walking pace of older people in comparison with each modality being considered in isolation. The proposed approach consists of two networks a) feature discovery network and b) walk discriminator network. The feature discovery network is based on an autoencoder architecture to discover the representative features. The extracted features enhance the power of the subsequent walk discriminator network to differentiate the walking pace specifically in the case of older people. The performance of the model is evaluated on the publicly available dataset which confirms the multimodalities can assist the quantification of the walking pace. The key contributions of our multimodal approach are as follows:

- A feature discovery network is proposed by considering the heterogeneous sensor modalities to encode a single feature vector as a latent representation in an unsupervised fashion.
- A walk discriminator network is designed and fine-tuned by considering the latent space representation established by the feature discovery network. It walk discriminator

network performs the classification task and is based on a feed-forward neural network architecture.
- The experiments are performed on a dataset captured from a cohort of older people to evaluate the performance of the proposed multimodal approach. This research serves as a baseline for future research studies to develop robust models for older people and is timely with this cohort becoming larger in population across most countries.

This paper is organized as follows: Section II presents the analysis of existing models for quantifying walking pace and autoencoder-based models. In Section III, the details of the proposed model including the model learning, designed architecture, and data processing are presented. In Section IV, the results of the experimental study are presented and followed by a discussion of the findings. The conclusions are drawn in Section V.

II. Related Work

The walking pace in older age is linked to various health issues ranging from functional dependence to mortality [3]. Buckley *et al.* [10] determined the number of days required to quantify the walking activity with wearable devices in aged residential care (ARC) populations. To understand the mobility needs of sixty-five years or older residents in ARC, a wearable device was used to collect the data over a period of eight days. The results showed that participants required different levels of care based on their level of mobility which was recognized by the wearable device. Joddrell *et al.* [3] used LiDAR sensor as a low-cost technology in a home setting to measure the walking speed in the context of this metric being a significant indicator of frailty in older adults. Their study involves observation of twenty frail older adults for thirteen days. The statistical data analysis (i.e., Spearman's rank-order correlation test) was used to classify the walking speed and comparisons were made with manual timed walk test.

Zihajehzadeh *et al.* [11] combine a regression model and support vector machine to estimate the walking pace of young adults. The experiments were performed on ten healthy subjects where their activity on a treadmill was recorded using a smartwatch. The triaxial accelerometer, gyroscope and magnetometer from the smartwatch were utilized to recognize the walking speed of the participants. The experiment showed that combining multiple classifiers along with heterogeneous sensor data can improve the prediction of the ambulant activity. Silsupadol *et al.* [12] assess the gait during straight walking, turning, and walking speed in laboratory and free-living environments. They attached two smartphones to the back of participants with one positioned on a belt and the other placed in a bag. The participants in this study consisted of both younger and older adults. Their model is based on statistical techniques such as correlation and visual analysis of the signals. They confirmed the wearable sensors can contribute toward gait and walking speed recognition.

Ferrari *et al.* [13] propose an autoencoder architecture to learn the latent representation from high dimensional data for optimal electrode set identification in wearable electroencephalography (EEG) event monitoring. In this study brain rhythm is considered to evaluate the performance of the model. The model performance is measured in terms of F1-Score (i.e., 0.78). The autoenocder-based model enables the wearable devices to monitor real-life events. Wong *et al.* [14] developed a deep convolutional autoencoder for energy efficiency in wearables for smart healthcare services. The autoencoder based model reduce the error when compared with hand crafted features such as wavelets. Zhu *et al.* [15] designed variational autoencoder with Gaussian mixture to represent the latent space of normal emotion recognition and apply convolutional neural networks (CNNs) to detection the anomalies in emotions. Furthermore, they deploy the system in real-life setting to confirm the applicability using the low-cost wearable sensors.

Our work is inspired by the literature to extract the best possible representation of accelerometer and ECG signals as latent features and classifying them using a feed-forward neural network to quantify the walking pace of a given individual.

III. The Proposed Approach

The proposed approach considers two modalities from a single chest-worn wearable; specifically a triaxial accelerometer and ECG. Our overall approach consists of data preprocessing and two networks (a) feature discovery network and (b) walk discriminator network. Figure 1 provides an illustrative overview of the proposed approach.

A. Data Preprocessing

A clinical-grade wearable device from Equivital was used to capture the electrocardiogram (ECG) waveform and triaxial accelerometer signals [16]. Both sensors are synchronized over the defined window of interest. In Equivital, ECG signals are measured with the help of 2 leads/channels, denoted $E_{l1}(t)$ and $E_{l2}(t)$, and the accelerometer signal is measured along three axes referred to as $a_x(t)$, $a_y(t)$, and $a_z(t)$. Both sensor groups record at different frequencies, so the ECG signals are down-sampled to match the frequency of the accelerometer data during preprocessing. The downsampling process retains the original ECG values that coincide with the accelerometer data. The resultant input vector to the feature discovery network at time *t* is:

$$\mathrm{s(t)} = \begin{bmatrix} E_{l1}(t) \\ E_{l2}(t) \\ a_x(t) \\ a_y(t) \\ a_z(t) \end{bmatrix} \tag{1}$$

It covers non-overlapping window of 5 seconds of activity, which is extracted from the raw dataset to form input samples.

B. Feature Discovery Network

The feature discovery network is based on representational learning, where the goal is to extract the valuable representation of the multimodal sensor data. The network is based on a feed-forward autoencoder with encoder $E(s)$, latent features

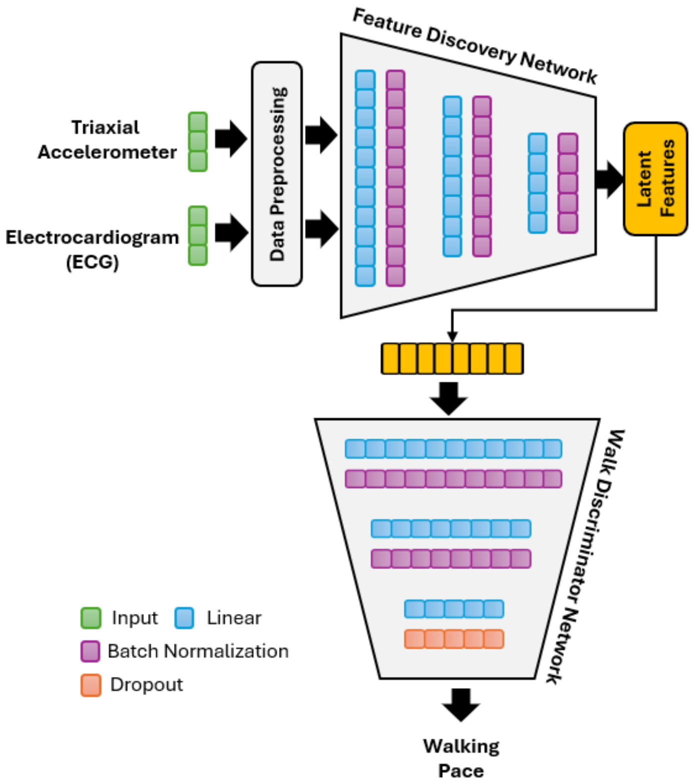

Fig. 1. The block diagram of the proposed model. It provides detail about the two networks (a) Feature Discovery Network (b) Walk Discriminator Network

(z), and decoder $D(s)$ components. The encoder maps the sensor data vectors $s \in \mathbb{R}^n$ to a latent representation $z \in \mathbb{R}^d$. It is defined by the following equation:

$$\mathrm{E(s)} : \mathbb{R}^{\mathrm{n}} \rightarrow \mathbb{R}^{\mathrm{d}} \tag{2}$$

From the latent representation (z), the original sensor data is approximately reconstructed using decoder $D(s)$. It is defined by the following equation:

$$\mathrm{D(z)} : \mathbb{R}^{\mathrm{d}} \rightarrow \mathbb{R}^{\mathrm{n}} \tag{3}$$

The discrepancy between the input s and reconstructed output, $\hat{s}$, is quantified using the mean squared error (MSE) loss:

$$\mathrm{MSE(s, \hat{s})} = \frac{1}{n} \sum_{i=1}^{n} (s_i - \hat{s}_i)^2 \tag{4}$$

During the training process, the parameters of the $E(s)$ and $D(z)$ are adjusted to minimize the loss function. As the loss decreases, it indicates that the reconstructed output is getting closer to the original input in a meaningful representation of the latent space. Later, during the test phase, only the encoder part is utilized to transform the input to latent features. The encoder part of the network is presented in Figure 1, as a feature discovery network. It consists of linear layers followed by batch normalization. Each layer provides a different level of abstraction as well as contributes towards the latent representation. The details of the architecture of the feature discovery network are presented in Table I.

TABLE I
THE ARCHITECTURE DETAILS OF FEATURE DISCOVERY NETWORK

Layer (Type)	Shape	Activation	Parameters
Input	1 × 650	-	-
Linear & Batch Norm.	650 × 1300	LeakyReLU	851500
Linear & Batch Norm	1300 × 650	LeakyReLU	848250
Linear	650 × 20	-	13020
Linear & Batch Norm	20 × 650	Linear	16250

C. *Walk Discriminator Network*

The walk discriminator network consists of a feed-forward multi-layer perceptron (MLP) to discriminate the walking pace of older people. The designed network consists of six layers (i.e., linear, batch normalization, and dropout) as shown in Figure 1 and detailed in Table II. The input to the network is the latent feature vector from the feature discovery network. Every layer of this MLP is fully connected to every subsequent layer.

TABLE II
THE ARCHITECTURE DETAILS OF WALK DISCRIMINATOR NETWORK

Layer (Type)	Shape	Activation	Parameters
Input	1 × 20	-	-
Linear & Batch Norm.	20 × 16	ReLU	460
Linear & Batch Norm	16 × 14	ReLU	606
Dropout	-	-	-
Linear	14 × 3	Softxax	45

A weighted sum is calculated over each neuron and a non-linear activation function is applied to pass the energy to the neurons in the next layer. It is calculated as follows:

$$y = \phi \left[f\left(\sum_{k=0}^{m} w^k f^k + b \right) \right] \tag{5}$$

where ϕ is the activation function, b is the bias and m is the number of neurons in each corresponding layer. The rectified linear unit (ReLU) is used as an activation function for all layers and softmax is used on the last layer to classify the output. The model is trained to minimize the cross-entropy loss:

$$\mathcal{L}_{CE}(\mathbf{\Theta}) = - \sum_{i \in C} y_i \cdot \log \hat{y}_i, \tag{6}$$

where c are the class labels, y_i is the original label, and $\hat{y}$ is the predicted label for walking pace; either slow, normal, or fast-paced. Minimizing the cross-entropy loss guides the model to make correct predictions and move closer to the underlying ground truth labels. Details about the model convergence and optimal numbers of hyperparameters will be discussed in Section IV.

IV. Experiment and Evaluation

A chest-worn wearable from publicly available dataset (GOTOV) [16] is utilized to measure the performance of our multimodal approach to categorise the walking pace of each person. It is collected in free-living conditions from 35 older people (14 female, 21 male) aged between 60 and 85 to quantify their lifestyle in terms of activity recognition and energy expenditure. The raw data are prepared according to the data processing information outlined in Section III-A.

The feature discovery network processes the multimodal synchronized input vectors from the ECG and accelerometer sensors to extract valuable information. The output of the feature discovery network is the latent representation with 20 features. These features are passed to our walk discriminator network to quantify the slow, normal, and fast-paced walk. The optimal hyperparameters of the feature discovery and walk discriminator network are presented in Table. III. These hyperparameters are the optimal hyperparameters according to model convergence.

TABLE III
The feature discovery network hyperparameters

Hyperparameters	FDN: Value	WDN: Value
Number of epochs	150	500
Optimizer	Adam	Adam
Learning rate	0.01	0.01
Moment	-	0.1
Dropout	-	0.4
Batch Size	64	50

The model performance is measured in terms of precision, recall, accuracy, and F1-score. These metrics are calculated as Equations 7, 8, 9, and 10 respectively.

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{9}$$

$$F_1 - score = 2 \times \frac{P \times R}{P + R} \tag{10}$$

where TP represents true positives, TN represents true negatives, FP represents false positives, and FN represents false negatives. The obtained results to quantify walking pace using chest-worn wearable sensors are reported in Table IV.

Table IV, shows that our model achieve good performance to distinguish between the slow, normal and fast-pace walking. Furthermore, we also analyzed the walking pace recognition performance on a single accelerometer sensor and compared with our multimodal approach. The results are presented in Figures 2, 3, and 4.

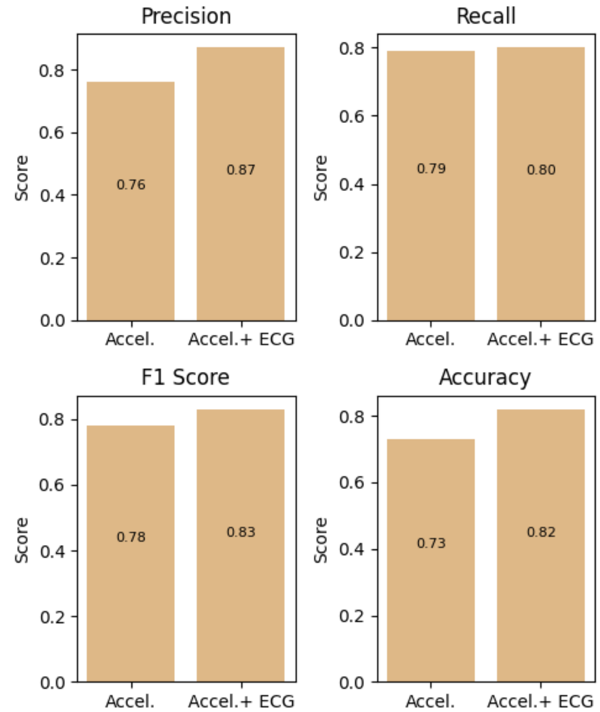

Fig. 2. The performance measures of slow walk with single modality (i.e, Accelerometer) and multimodality (i.e, Accelerometer and ECG).

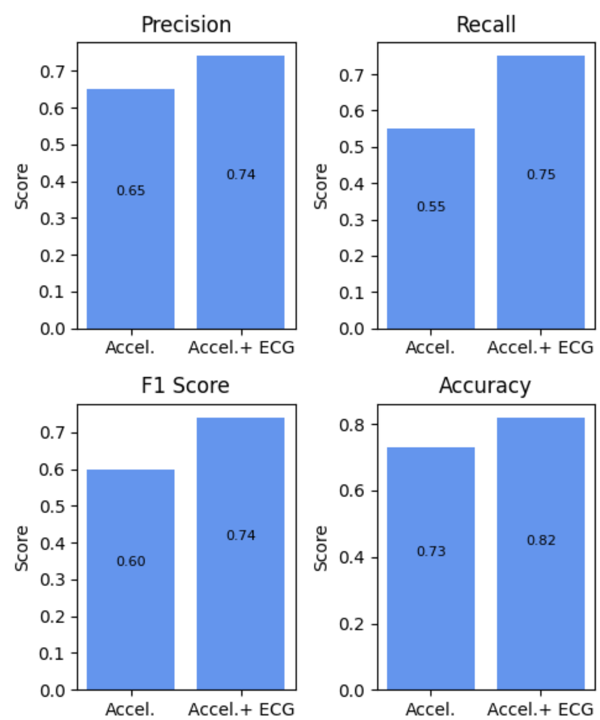

Fig. 3. The performance measures of normal walk with single modality (i.e, Accelerometer) and multimodality (i.e, Accelerometer and ECG).

TABLE IV
THE OBTAINED RESULTS OF PROPOSED MODEL USING MULTIMODAL DATA

Walking Pace	Precision	Recall	F1-Score	Accuracy
Slow	**0.87**	0.80	0.83	**0.82**
Normal	0.74	0.75	0.74	**0.82**
Fast	0.85	**0.91**	**0.88**	**0.82**

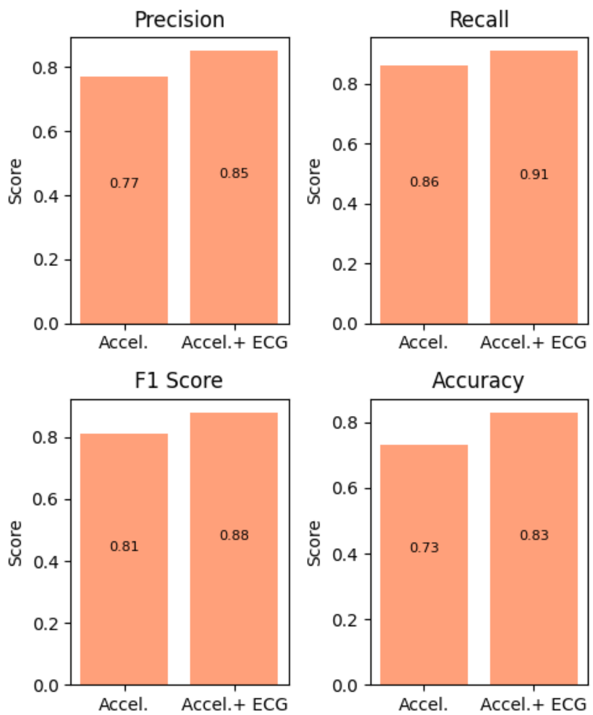

Fig. 4. The performance measures of fast walk with single modality (i.e, Accelerometer) and multimodality (i.e, Accelerometer and ECG).

The obtained results consistently demonstrate the superior performance of our multimodal approach (i.e., accelerometer + ECG) across all performance measures compared with only using accelerometer data. This approach not only leads to at least 9% more accurate results (Figures 2-4, but also brings stability to walking pace prediction.

V. Conclusion

In older adults, walking is a major physical activity that can improve well-being and promote healthy aging. Quantifying walking pace can contribute to observing behavioral changes in an older population. Our multimodal approach utilizes a chest-worn wearable device that captures both ECG and accelerometer through specially designed artificial neural networks. The obtained results show a 9% higher performance in accuracy when using both modalities as compared with using a single triaxial accelerometer. This demonstrates the feasibility of continuously classifying walking into slow, normal, and fast paces. In future work, we plan to expand this research further by building upon this approach to include the identification of daily routine activities and their association with behavioral changes.

References

[1] G. Valenti, A. G. Bonomi, K. R. Westerterp *et al.*, "Walking as a contributor to physical activity in healthy older adults: 2 week longitudinal study using accelerometry and the doubly labeled water method," *JMIR mHealth and uHealth*, vol. 4, no. 2, p. e5445, 2016.

[2] C. E. Adam, A. L. Fitzpatrick, C. S. Leary, A. Hajat, S. D. Ilango, C. Park, E. A. Phelan, and E. O. Semmens, "Change in gait speed and fall risk among community-dwelling older adults with and without mild cognitive impairment: a retrospective cohort analysis," *BMC geriatrics*, vol. 23, no. 1, pp. 1–11, 2023.

[3] P. Joddrell, S. Potter, L. P. de Witte, and M. S. Hawley, "Continuous in-home walking speed monitoring in older people with a low-cost ambient sensor: Results of a feasibility study," *Technology and Disability*, vol. 33, no. 2, pp. 77–85, 2021.

[4] M. Quan, P. Xun, C. Chen, J. Wen, Y. Wang, R. Wang, P. Chen, and K. He, "Walking pace and the risk of cognitive decline and dementia in elderly populations: a meta-analysis of prospective cohort studies," *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, vol. 72, no. 2, pp. 266–270, 2017.

[5] T. Yates, C. Razieh, F. Zaccardi, A. V. Rowlands, S. Seidu, M. J. Davies, and K. Khunti, "Obesity, walking pace and risk of severe covid-19 and mortality: analysis of uk biobank," *International journal of obesity*, vol. 45, no. 5, pp. 1155–1159, 2021.

[6] D. Jin, H. Halvari, N. Maehle, and A. H. Olafsen, "Self-tracking behaviour in physical activity: a systematic review of drivers and outcomes of fitness tracking," *Behaviour & Information Technology*, vol. 41, no. 2, pp. 242–261, 2022.

[7] M. Fahim, V. Sharma, and T. Q. Duong, "A wearable-based preventive model to promote oral health through personalized notification," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2022, pp. 4282–4285.

[8] S. Paraschiakos, C. R. de Sá, J. Okai, P. E. Slagboom, M. Beekman, and A. Knobbe, "A recurrent neural network architecture to model physical activity energy expenditure in older people," *Data Mining and Knowledge Discovery*, vol. 36, no. 1, pp. 477–512, 2022.

[9] N. D. Nguyen, D. T. Bui, P. H. Truong, and G.-M. Jeong, "Classification of five ambulatory activities regarding stair and incline walking using smart shoes," *IEEE Sensors Journal*, vol. 18, no. 13, pp. 5422–5428, 2018.

[10] C. Buckley, A. Cavadino, S. Del Din, S. Lord, L. Taylor, L. Rochester, and N. Kerse, "Quantifying reliable walking activity with a wearable device in aged residential care: How many days are enough?" *Sensors*, vol. 20, no. 21, p. 6314, 2020.

[11] S. Zihajehzadeh, O. Aziz, C.-G. Tae, and E. J. Park, "Combined regression and classification models for accurate estimation of walking speed using a wrist-worn imu," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 3272–3275.

[12] P. Silsupadol, P. Prupetkaew, T. Kamnardsiri, and V. Lugade, "Smartphone-based assessment of gait during straight walking, turning, and walking speed modulation in laboratory and free-living environments," *IEEE journal of biomedical and health informatics*, vol. 24, no. 4, pp. 1188–1195, 2019.

[13] L. M. Ferrari, G. Abi Hanna, P. Volpe, E. Ismailova, F. Bremond, and M. A. Zuluaga, "One-class autoencoder approach for optimal electrode set identification in wearable eeg event monitoring," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021, pp. 7128–7131.

[14] J. Wong and Q. Zhang, "Deep convolutional autoencoder for energy-efficient smart health wearables in the era of big data," in *2022 IEEE International Performance, Computing, and Communications Conference (IPCCC)*. IEEE, 2022, pp. 202–206.

[15] J. Zhu, F. Deng, J. Zhao, D. Liu, and J. Chen, "Uaed: Unsupervised abnormal emotion detection network based on wearable mobile device," *IEEE Transactions on Network Science and Engineering*, 2023.

[16] S. Paraschiakos, R. Cachucho, M. Moed, D. van Heemst, S. Mooijaart, E. P. Slagboom, A. Knobbe, and M. Beekman, "Activity recognition using wearable sensors for tracking the elderly," *User Modeling and User-Adapted Interaction*, vol. 30, no. 3, pp. 567–605, 2020.