Effective Sensor Selection for Human Activity Recognition via Shapley Value

Elisa Borella, Umut Berk Çakmakçı, Enrico Gottardis, Alessandro Buratto, Thomas Marchioro, and Leonardo Badia
Department of Information Engineering, University of Padova, Padova, Italy
Email: {elisa.borella.4, umutberk.cakmakci, enrico.gottardis}@studenti.unipd.it
{alessandro.buratto.1@phd., thomas.marchioro@, leonardo.badia@}unipd.it

2024 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv) | 979-8-3503-8501-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/METROLIVENV60384.2024.10615860

***Abstract*—Real-time human activity recognition is facing an ever-growing need for efficient sensor setup. Identifying a minimal sensor configuration can lead to cost savings and less intrusive equipment, ultimately improving the quality of the collected data. In this study, we introduce and assess a sensor selection approach that ranks sensors based on their relevance in human activity recognition (HAR) tasks. Our methodology utilizes the Shapley value – a widely adopted metric inspired by game theory – of sensor measurements to determine the importance of each sensor. To validate our approach, we assess the impact of sensor removal on the accuracy of XGBoost tree models, which are trained on a publicly available HAR dataset. Our experiments indicate that Shapley-based sensor ranking achieves a favorable cost-accuracy tradeoff allowing for a reduction by more than 50% in the number of exploited sensors without significantly affecting accuracy.**

***Index Terms*—Human activity recognition; Wearable sensors; Feature selection; Game theory.**

I. Introduction

Wearable sensing technology has emerged as a crucial tool for the collection and analysis of human data across various fields such as sport analytics, e-health, and more [1], [2]. One of the most significant applications of this technology is human activity recognition (HAR), which encompasses tasks related to identifying different activities using sensor data [3]–[5].

HAR is typically carried out by machine learning models using data gathered from wearable sensing devices [6]. These devices are often inertial measurement units (IMUs) that are attached to different parts of the body [7]. IMUs measure linear and angular acceleration in multiple directions, providing a comprehensive view of an individual's movements [8].

Recent years have seen a notable increase in real-time applications of HAR, where balancing the quality and quantity of collected data is paramount [9]. While a comprehensive sensor setup typically leads to more accurate predictions, it also necessitates enhanced processing capabilities, longer prediction times, more expensive equipment, and more obtrusive impact on the user [10], [11]. This tradeoff highlights the importance of carefully considering the requirements of individual applications when selecting the wearable sensor configuration.

This work was supported by the Italian PRIN project 2022PNRR "DIGIT4CIRCLE," project code P2022788KK.

In this paper, we introduce and evaluate a methodology to effectively choose a subset of sensors among those available for HAR applications. We rank different wearable sensing units, such as IMUs and temperature sensors, according to the importance of their measurements (i.e., the *features* captured by each sensor). Feature importance is evaluated using SHapley Additive exPlanations (SHAP) [12], a game theory-inspired metric that estimates the additive contribution of a feature to the final prediction of a machine learning model. Our approach can be used to limit the setup to a fixed number of relevant sensing units, thereby significantly decreasing computational complexity, and allowing for lightweight real-time management of data [13], [14].

The existing literature mostly focuses on offline extraction methods [15] or considers the optimization of sensor placement and control [16]. In contrast, our aim in this work is to develop a *zero-touch* form of control over the IMU units, which are independently placed prior to the data collection and over which we have no control at run time. We are interested in a limited selection of sensors based on their relevance and leveraging their redundant content in the decision process [17], [18]. The evaluation of the Shapley values is meant to highlight the most relevant sensors so as to decrease the handling of data without significantly affecting accuracy, as we will show [19].

We evaluate this proposed methodology on HuGaDB, a publicly available HAR dataset comprising data captured by 8 different sensing units for a total of 38 measurements per sample. We rank the sensing units in HuGaDB via our Shapley-based approach and gradually discard them, starting from the lowest ranked up to the highest.

Our evaluation suggests that a Shapley-based sensor selection provides a favorable tradeoff between number of adopted sensors and activity recognition accuracy, i.e., better accuracy compared to a random sensor selection. In particular, our effective choice mechanism allows for a decrease in the number of sensors used by 50% without significantly impacting the resulting accuracy, but at the same time heavily simplifying the coordination of data collection, and possibly allowing for a more systematic treatment [20]. This may especially be relevant whenever sensed data are coordinated among multiple sources, as is the case of crowdsensing [21], [22].

The rest of this paper is organized as follows. In Section II, we review the fundamentals of HAR techniques and Shapley

values. The literature of related papers is discussed in Section III. Section IV outlines our proposed methodology, whereas Section V describes its quantitative evaluations on a dataset. We conclude the paper in Section VI.

II. Background

A. Human activity recognition and sensor selection

HAR is a subfield of machine learning aimed at automatically identifying human actions such as walking, running, climbing, and more, using measurements collected by sensors [4]. Given a tuple of measurements $x_i = (x_{i,1}, \dots, x_{i,F}) \in \mathbb{R}^F$, the objective is to correctly identify the corresponding activity y_i among a number of possible activities $\mathcal{Y} = \{1, \dots, m\}$. In many applications, measurements are analyzed as time series using sliding window methods or within specific time frames [23]. Nevertheless, in real-time situations, conducting time series analysis can be impractical, and it may be more advantageous to opt for sample-wise inference approaches. This paper concentrates on the latter approach, predicting the activity based on individual samples. This limits the computational complexity required by our approach while maintaining high accuracy levels [9].

Each sample includes different measurements collected by multiple sensing units $S_1, \dots, S_K$, which can be either IMUs, temperature sensors, or other sensory devices. Each sensing unit is characterized by the features collected by it, i.e., $\sigma_k = \{j_{k,1}, j_{k,2}, \dots\}$. Formally, the sensing units form a partition of the feature set $\mathcal{F} = \{1, \dots, F\}$, meaning that

$$\sigma_k \cap \sigma_{k'} = \emptyset, \forall k, k' \text{ and } \bigcup_{k=1}^{K} \sigma_k = \mathcal{F}. \quad (1)$$

B. Shapley value and SHAP

Shapley value is a metric inspired by cooperative game theory that measures the individual contribution of players in a coalition [19]. In machine learning, Shapley value is frequently used to evaluate the impact of individual features to the final prediction [20]. This can be evaluated through data processing packages available in the literature, such as SHapley Additive exPlanation (SHAP) [12]. SHAP scores for each feature $j \in \mathcal{F}$ are computed on a data point x_i as follows

$$\phi_j(x_i) = \sum_{S \subseteq \mathcal{F} \setminus \{j\}} \frac{|S|!(F-|S|-1)!}{F!} \left(v(x_{i,S\cup\{j\}}) - v(x_{i,S})\right) \quad (2)$$

where $v(x_{i,S})$ is the model's output for x_i when restricting the features to a subset $S \subseteq \mathcal{F}$.

In a binary classification problem, v is the positive class score computed by the model. In a multi-class classification problem, instead, a tuple of model outputs $v_1, \dots, v_m$ is used, one for each class, and a separated SHAP score is computed for each output: $\phi_{j,1}(x_i), \dots, \phi_{j,m}(x_i)$. A SHAP score can be either positive or negative, since the model's output can either increase or decrease. However, the impact is measured in terms of its absolute value. The overall SHAP value of a feature is determined by averaging the SHAP score magnitudes over a validation dataset $x_1, \dots, x_n$ and summing them across all classes, i.e.,

$$\phi_j = \frac{1}{n} \sum_{i=1}^{n} \sum_{\ell=1}^{m} |\phi_{j,\ell}(x_i)|. \quad (3)$$

Albeit widely used in academia, one main obstacle in the practical use of SHAP values is computational complexity [24]. In model-agnostic scenarios, evaluating SHAP has complexity $O(2^F)$, i.e., exponentially increasing with the number of features. Nevertheless, when the model structure is known, it is possible to take advantage of it and obtain a more efficient evaluation. This is the case for tree-based models, for which the SHAP value can be computed in $O(D^2)$, where D is the maximal depth of the tree, using the TreeSHAP algorithm [25].

III. Related work

Although many studies have highlighted the importance of effective sensor selection for HAR, few practical methods have been proposed in the scientific literature thus far.

For example, [6] discusses effective sensor *placement*, presupposing that data collection has a degree of freedom available beforehand, but in the end all of the sensors are used, which may be impractical for real-time control of multiple units [21]. Conversely, [3] suggest that jointly handling interconnected signals may benefit in compressing the reading, but this analysis is very application-specific to the actual features, which related to EMG and EEG data, and not multiple sensors.

We also remark that most HAR studies focus on sport activities [2], [7], where it is quite common to have multiple redundant sensors [5]. However, applying dimensionality reduction in these cases would be preferable than collecting huge amount of data that would make a real-time control impractical [8], [18]. Thus, identifying an effective yet limited subset of sensors for practical purposes is an extremely interesting issue. Moreover, as argued in [9], decreasing the number of sensors used helps in reducing battery consumption and prompts a better coordination with energy harvesting [26].

Finally, it is also to remark that data redundancy is often so deeply inherent in biometry sensors that even physical layer protections such as perturbation with artificial noise often fail to hide the anonymity of data [27]. While out of scope of the present analysis, a better understanding of the explainable characteristics of HAR can assist towards privacy goals.

The idea of using Shapley values to unveil the decision process based by ML is also not new, and many contributions used a similar approach to explanable AI [15], [19], [24]. However, a notable difference with the present contribution is that those approaches usually consider the importance of different features inside a dataset. The general aim is to derive an explanation of an intelligent model seen as a black box [20].

While applying a similar reasoning, our approach is different, as we are not interested in deriving an explanation of the adequacy of ML methodologies applied to IMU data for HAR, which is exhaustively proven by other contributions [10]. Instead, we propose to use this explanatory tools towards the identification of the most relevant sensors for an effective

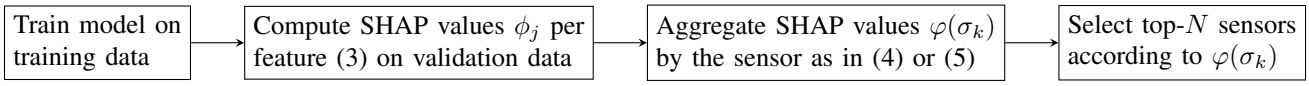

Fig. 1. The proposed SHAP-based sensor selection pipeline.

yet lightweight representation [13]. Computing the Shapley values of specific sensors, instead of just features, may also reflect better under dynamic measurements as HAR [23], allowing to identify the more critical nodes in detecting changing conditions, thereby improving resource allocation and/or energy efficiency [4].

IV. Methodology

Efforts to limit the amount of sensor reading without affecting the accuracy can be connected to several advantages, including the improvement of computational complexity as well as energy saving [26]. To some extent, this can also improve the experience for the users, if some sensors can be avoided or made less invasive [28]. For this reason, we propose a methodology for the selection of the top-N sensing units, i.e., a limited number N of sensors is used for the evaluation among those available, and they are chosen as the ones with highest Shapley values [15].

Our proposed methodology can be summarized as follows. First, we divide the available sensor data in training and validation sets. The training data is used to establish a machine learning model for activity classification. The validation data, instead, is employed to compute the SHAP values $\phi_1, \ldots, \phi_F$ for all features. After this step, SHAP values are aggregated by the sensor into values $\varphi(\sigma_k)$. The aggregation can be done by considering the sum of the feature SHAP values for each sensing unit σ_k, $k = 1, \ldots, K$, i.e.,

$$\varphi(\sigma_k) = \sum_{j \in \sigma_k} \phi_j \tag{4}$$

or by taking the maximum value

$$\varphi(\sigma_k) = \max_{j \in \sigma_k} \phi_j. \tag{5}$$

A mean aggregation of the feature SHAP values may also be considered, which is the same as the sum, but normalized by the number of features for each sensor.

The sensing unit can be ranked according to their φ value, and a top-N sensor setup can be obtained by choosing the N units with highest φ values. The complete sensor selection pipeline is summarized in Fig. 1. As demonstrated by the experimental results shown in the next section, this procedure is effective in finding the most relevant sensors.

V. Performance evaluation

A. Dataset

We evaluate our proposed pipeline on HuGaDB, a widely used HAR dataset [29]. The dataset comprises wearable sensor records collected from 18 individuals engaged in various activities such as walking, running, sitting, standing, and more. The dataset includes measurements obtained from six wearable IMUs, along with two EMG sensors. The IMUs were placed on both thighs, shins, and feet, while the EMG sensors were positioned on the quadriceps.

TABLE I
Description of HuGaDB dataset used in our evaluation

# samples	1 137 986
# activities	10
# sensing units	8
# features	38

A summary of the dataset is reported in Table I, which displays the total number of samples, activities, features, and sensing units. With "sensing units" we refer to the physical devices that were used to take measurements. Each sensing unit may comprise one or multiple sensors, which in turn may take multiple different measurements (i.e., features). For example, an IMU is a sensing unit comprising an accelerometer and a gyroscope, and yields a total of 6 features (the X, Y, and Z components of acceleration and angular acceleration). EMG sensors, on the other hand, are one-sensor units and produce one single measurement [8].

B. Experiments

To evaluate our approach, we split HuGaDB into training, validation, and test sets using an 80/10/10 ratio. We replaced null values in the features with the average value and performed a standard scaling of the features. All feature averages and standard deviations required for standard scaling and filling null values were computed on the training data. All the (x_i, y_i) pairs where label y_i was missing were removed.

We applied our methodology computing the SHAP values of all features and aggregating them by taking the maximum value for each sensor σ_k to compute $\phi(\sigma_k)$. We explain further this decision in the discussion section. The model used to compute the feature SHAP values was an XGBoost tree [30]. We opted for using the default hyperparameters rather than further splitting the training data in order to fine-tune them. The model was trained on the training data and SHAP values of the features were estimated using the validation data.

The overall SHAP weight of each feature is obtained by computing the Shapley values for all samples and classes, summing their magnitude across classes, and averaging across samples. The SHAP values of the top 20 features are shown in Fig. 2. The figure also displays the contribution of the different classes to the overall SHAP value of each feature. The name of the sensors follow the standard nomenclature of HuGaDB [29], i.e. `<sensor>_<unit>_<coordinate>`. For example, the feature with highest SHAP value is `acc_ls_z`, which

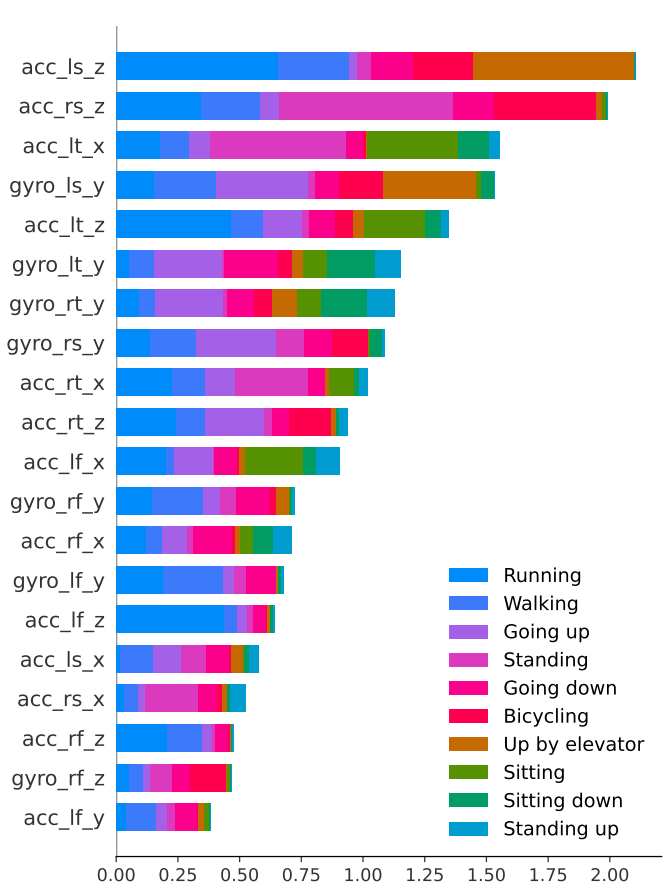

Fig. 2. Top 20 absolute SHAP values (summed across classes and averaged across samples).

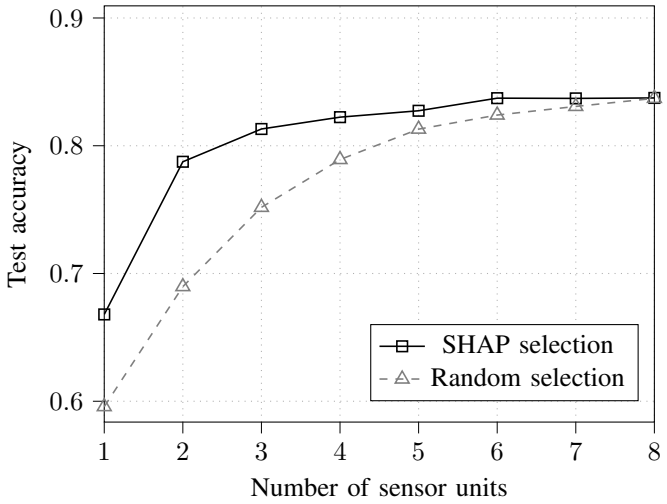

Fig. 3. Comparison of achieved test accuracy via Shapley selection versus random selection.

TABLE II
GLOBAL CLASSIFICATION METRICS FOR ALL, TOP-3, AND BOTTOM-3 SHAP-RANKED SENSING UNITS

	Accuracy	Balanced accuracy	Weighted F1-score
All sensors	0.8467	0.8058	0.8337
Top 3	0.8174	0.7724	0.8008
Bottom 3	0.6351	0.4808	0.5995

is the z-axis accelerometer value measured by the left-shin IMU, immediately followed by the z-axis accelerometer value measured by the right-shin IMU. Both these features seem to largely impact the prediction for walking and running, which is expected by sensors that are placed close to the knees.

As apparent from the figure, the top 20 features are all produced by the different IMUs, suggesting a limited contribution of EMG sensors to the model's prediction. The final ranking of the sensing unit obtained by the aggregation of the SHAP values was as follows: left shin, right shin, left thigh, right thigh, left foot, right foot, left-quadricep EMG, right-quadricep EMG. As expected, the low SHAP values of the EMG features caused them to be ranked last.

After determining the SHAP-based sensor ranking, we evaluated the performance of the top-N sensor setup for $N = 1, \ldots, 8$. We compared the accuracy (calculated on the test data) of the top-N SHAP-ranked features against N randomly selected features. The results of a random selection were obtained by averaging the results of a Monte Carlo simulation with 10 casual sensor rankings.

Fig. 3 shows the accuracy curves for SHAP-based and random selection. Clearly, both curves approach the same accuracy level for $N = 8$, since all sensing units are being used. The biggest difference can be observed for $N = 1$ and $N = 2$, where the ranking of the sensing units has the highest impact. Overall, these results suggest that SHAP values can indeed be used as an indicator for the most relevant sensors. Additionally, the use of just $N = 3$ sensors leaves the accuracy above 80% but reduces the management costs by more than 50%.

The validation of our SHAP-based sensor ranking was further confirmed through a comparison of the follow three setups: an overall setup utilizing all sensing units; a minimal setup consisting of the top-3 ranked sensing units; and a minimal setup consisting of the bottom-3 ranked sensing units. Evaluation was carried out using global metrics such as accuracy, balanced accuracy, and weighted F1-scores (Table II), as well as class-specific metrics like precision, true-positive rate (TPR), true-negative rate (TNR), and F1-score (Table III). The TPR and TNR are also known in literature as recall and specificity, respectively. It is worth mentioning that the balanced accuracy was computed as the average of the per-class TPR values (i.e., as the average recall) to adjust for any imbalance in the class distribution. Our findings indicate that the top-3 ranked sensing units achieve predictive performance similar to the complete setup on all classes, whereas the bottom 3 units exhibit worse results compared to randomly selected sensors.

TABLE III
CLASSIFICATION FOR EACH ACTIVITY USING ALL (ABOVE), TOP-3 (MIDDLE) AND BOTTOM-3 (BELOW) SHAP-RANKED SENSING UNITS

	Precision	TPR	TNR	F1-score
Walking	0.9313	0.9432	0.9734	0.9372
Running	0.9843	0.9514	0.9991	0.9676
Going up	0.9152	0.9030	0.9887	0.9091
Going down	0.9024	0.8948	0.9886	0.8986
Sitting	0.9904	0.9928	0.9992	0.9916
Sitting down	0.9716	0.9396	0.9996	0.9553
Standing up	0.9239	0.8835	0.9987	0.9032
Standing	0.6815	0.9331	0.8939	0.7877
Up by elevator	0.6016	0.3322	0.9813	0.4281
Down by elevator	0.6806	0.2849	0.9914	0.4016
Walking	0.8853	0.9356	0.9536	0.9097
Running	0.9625	0.9347	0.9978	0.9484
Going up	0.8942	0.8697	0.9861	0.8818
Going down	0.8802	0.7788	0.9875	0.8264
Sitting	0.9856	0.9903	0.9989	0.9880
Sitting down	0.9326	0.9121	0.9989	0.9222
Standing up	0.9040	0.8689	0.9983	0.8861
Standing	0.6549	0.9255	0.8814	0.7670
Up by elevator	0.5927	0.2727	0.9841	0.3736
Down by elevator	0.5956	0.2355	0.9897	0.3375
Walking	0.6703	0.8670	0.8368	0.7561
Running	0.7851	0.5766	0.9903	0.6649
Going up	0.6045	0.4730	0.9583	0.5307
Going down	0.6199	0.4057	0.9707	0.4904
Sitting	0.8335	0.8043	0.9874	0.8187
Sitting down	0.4348	0.2747	0.9942	0.3367
Standing up	0.4571	0.2330	0.9949	0.3087
Standing	0.5623	0.8963	0.8302	0.6910
Up by elevator	0.5587	0.1975	0.9867	0.2919
Down by elevator	0.5670	0.0799	0.9961	0.1401

C. Discussion

a) Choosing the aggregation type: As per Section IV, SHAP values can be aggregated by sensing unit in three ways: taking the sum, the average, or the maximum. A 'max' aggregation can be advantageous when there is a great disparity in the number of features produced by different sensing units, as in the case of HuGaDB. In these cases, sensing unit with a single highly relevant features may score lower in the ranking compared to units with several irrelevant features. The 'sum' and 'mean' aggregation types might be more suitable for cases where the number of features is comparable across all sensing units.

b) Limitations of Shapley-based selection: While our experiments show that using Shapley-based selection can be effective for sensor selection, it may not guarantee an optimal configuration. For example, this method does not account for potential correlations between sensors [18]. Future research should explore combining Shapley evaluation with correlation-based selection techniques such as [5]. In general, a similar methodology can be used in a trans-directional way, encompassing multiple dimensionality reductions both in the number of sensors and the features used.

VI. CONCLUSIONS

We discussed a mechanism based on Shapley values to rank sensors collecting HAR-related features to discriminate the most relevant of them and achieve minimal management [5].

We proved the effectiveness of a selection mechanism based on said ranking, where using the highest Shapley-valued sensors significantly outperform a random dimensionality reduction, and obtaining an extremely favorable tradeoff, where less than half of the sensors can be used without significant losses in the resulting accuracy.

While this confirms the general validity of our approach, future work is still possible in the directions of expanding the dataset and dealing with heterogeneous sensors [13], [21]. Moreover, we plan to investigate the provision of concrete implementations of our proposal, for a possible use in real devices and applications.

REFERENCES

[1] M. Wu and J. Luo, "Wearable technology applications in healthcare: a literature review," *Online J. Nurs. Inform.*, vol. 23, no. 3, 2019.

[2] G. Aroganam, N. Manivannan, and D. Harrison, "Review on wearable technology sensors used in consumer sport applications," *Sensors*, vol. 19, no. 9, p. 1983, 2019.

[3] G. Cisotto, A. V. Guglielmi, L. Badia, and A. Zanella, "Joint compression of EEG and EMG signals for wireless biometrics," in *Proc. IEEE Globecom*, 2018.

[4] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1192–1209, 2012.

[5] G. Ioannou, A. Kazlouski, T. Marchioro, and M. Gijssel, "Minimal wearable setup via sensor correlation: a case study of field hockey players," in *Proc. CEUR Wkshps*, 2023.

[6] C. Xia and Y. Sugiura, "Optimizing sensor position with virtual sensors in human activity recognition system design," *Sensors*, vol. 21, no. 20, p. 6893, 2021.

[7] M. Kim and S. Park, "Golf swing segmentation from a single IMU using machine learning," *Sensors*, vol. 20, no. 16, p. 4466, 2020.

[8] K. Çoçoli and L. Badia, "A comparative analysis of sensor fusion algorithms for miniature IMU measurements," in *Proc. IEEE ISITIA*, pp. 239–244, IEEE, 2023.

[9] Z. Chen, L. Teng, L. Xu, J. Yu, and J. Liang, "MP-HAR: A novel motion-powered real-time human activity recognition system," *IEEE Internet Things J.*, vol. 11, pp. 7652–7663, Mar. 2024.

[10] N. G. Nia, E. Kaplanoglu, A. Nasab, and H. Qin, "Human activity recognition using machine learning algorithms based on IMU data," in *Proc. IEEE BioSMART*, 2023.

[11] S. Yfantidou, C. Karagianni, S. Efstathiou, A. Vakali, J. Palotti, D. P. Giakatos, T. Marchioro, A. Kazlouski, E. Ferrari, and Š. Girdzijauskas, "LifeSnaps, a 4-month multi-modal dataset capturing unobtrusive snapshots of our lives in the wild," *Scientific Data*, vol. 9, no. 1, p. 663, 2022.

[12] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Adv. Neur. Inf. Proc. Syst.*, vol. 30, 2017.

[13] G. Cisotto, A. V. Guglielmi, L. Badia, and A. Zanella, "Classification of grasping tasks based on EEG-EMG coherence," in *Proc. IEEE Healthcom*, 2018.

[14] R. Hazra, M. Banerjee, and L. Badia, "Machine learning for breast cancer classification with ANN and decision tree," in *Proc. IEEE IEMCON*, pp. 0522–0527, 2020.

[15] L. Merrick and A. Taly, "The explanation game: Explaining machine learning models using Shapley values," in *Proc. CD-MAKE*, pp. 17–38, Springer, 2020.

[16] Y. Yiğitbaşı, F. Stroppa, and L. Badia, "Optimizing real-time decision-making in sensor networks," in *Proc. IEEE DeSE*, 2023.

[17] B. Dutta, A. Krichel, and M.-P. Odini, "The challenge of zero touch and explainable AI," *J. ICT Standardization*, vol. 9, no. 2, pp. 147–158, 2021.

[18] A. Zancanaro, G. Cisotto, and L. Badia, "Modeling value of information in remote sensing from correlated sources," *Comp. Commun.*, vol. 203, pp. 289–297, 2023.

[19] D. Scapin, G. Cisotto, E. Gindullina, and L. Badia, "Shapley value as an aid to biomedical machine learning: a heart disease dataset analysis," in *Proc. IEEE CCGrid*, pp. 933–939, 2022.

[20] S. Mirzaei, H. Mao, R. R. O. Al-Nima, and W. L. Woo, "Explainable AI evaluation: A top-down approach for selecting optimal explanations for black box models," *Information*, vol. 15, no. 1, p. 4, 2023.
[21] T. Marchioro, A. Kazlouski, and E. P. Markatos, "Practical crowdsourcing of wearable iot data with local differential privacy," in *Proc. ACM/IEEE Conf. Internet Things Design Implem. (IoTDI)*, pp. 275–287, 2023.
[22] L. Badia and N. Borra, "Crowdsensing strategies inspired by choir management analyzed via game theory," in *Proc. IEEE UEMCON*, pp. 0862–0868, 2020.
[23] J. He, Z. Guo, L. Liu, and Y. Su, "Human activity recognition technology based on sliding window and convolutional neural network," *J. Electron. Inf. Technol*, vol. 44, pp. 168–177, 2022.
[24] G. Van den Broeck, A. Lykov, M. Schleich, and D. Suciu, "On the tractability of SHAP explanations," *J. Artif. Intell. Res.*, vol. 74, pp. 851–886, 2022.
[25] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee, "From local explanations to global understanding with explainable AI for trees," *Nat. Machine Intell.*, vol. 2, no. 1, pp. 56–67, 2020.
[26] N. Michelusi, K. Stamatiou, L. Badia, and M. Zorzi, "Operation policies for energy harvesting devices with imperfect state-of-charge knowledge," in *Proc. IEEE ICC*, pp. 5782–5787, 2012.
[27] T. Marchioro, A. Kazlouski, and E. P. Markatos, "User identification from time series of fitness data.," in *Proc. SECRYPT*, pp. 806–811, 2021.
[28] S. Nehra and J. L. Raheja, "Unobtrusive and non-invasive human activity recognition using Kinect sensor," in *Proc. IEEE ICAN*, pp. 58–63, 2020.
[29] R. Chereshnev and A. Kertész-Farkas, "HuGaDB: Human gait database for activity recognition from wearable inertial sensor networks," in *Proc. AIST*, pp. 131–141, Springer, 2017.
[30] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. ACM SIGKDD*, pp. 785–794, 2016.