This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2024.3420891

A novel federated meta-learning approach for discriminating sedentary behavior from wearable data

Pedro H. Barros, Judy C. Guevara, Leandro Villas, Daniel Guidoni, Nelson L. S. da Fonseca and Heitor S. Ramos

***Abstract*—Characterizing and monitoring patient activities through time series data is critical for identifying lifestyle patterns that may impact health outcomes. Sedentary behavior is a significant concern due to its association with various health risks. This study introduces a lightweight supervised classifier for healthcare applications based on ordinal pattern transformation to detect sedentary behavior in federated learning scenarios. Our hypothesis is grounded on the idea that sedentary behavior exhibits distinct dynamics compared to other activities, and information descriptors derived from the transformation of ordinal patterns effectively capture these differences. Next, we proceed with the federated learning training. We train a neural network-based encoder locally and send the local models to a server. The federated learning process updates the encoder weights based on the encoded representations of the clients' data, enabling the model to learn from different participants. Finally, we personalize the model for the specific task of classifying sedentary behavior. Our approach utilizes a meta-learning framework, incorporating a Siamese neural network to learn a similarity space. We fine-tune the model in this step by further training the last neural network layer. This fine-tuning allows the model to adapt and specialize in accurately classifying sedentary behavior. We carry out a comprehensive analysis to support our hypothesis. We also extensively validated our proposal by comparing it with other methods over five different datasets. We obtain the best results using a smaller ML model compared with the best approaches in the literature. Specifically, our model has 78.73% times fewer parameters and consumes 48.67% times less energy than the best result in the literature.**

***Index Terms*—Wearable Data, Ordinal Patterns, Sedentary behavior, Federated Learning, Neural Network.**

I. INTRODUCTION

IN recent years, the importance of understanding the relationship between sedentary behavior and chronic health conditions has grown significantly [1]. Reducing sedentary behavior and promoting physical activity are essential for maintaining a healthy lifestyle. Machine Learning (ML) models trained on wearable device data offer a promising tool for identifying and addressing sedentary behavior, providing personalized recommendations and interventions to improve health. However, challenges such as data privacy and energy consumption hinder the widespread use of ML models in this context [2].

Federated learning (FL) offers a potential solution to these challenges by allowing ML models to be trained on decentralized data while preserving data privacy [3]. Data remains on local devices in FL, and only the local model parameters are shared and aggregated to form a global model [4].

Among distinct challenges in FL, client drift [5, 6] refers to the deviation of local models from the global model over time due to the heterogeneity of local datasets [7] and the varying training dynamics across clients. In the context of human activity data, different individuals exhibit unique patterns for the same activities, like walking [8]. The heterogeneity of local datasets and unique activity patterns are critical challenges in FL, especially for human activity data. Existing FL works predominantly focus on activity recognition or healthcare monitoring but often overlook sedentary behavior detection.

Detecting sedentary behavior in federated learning (FL) environments is a significant challenge due to the inherent data heterogeneity among clients, particularly the issue of label concept skew. Unlike label distribution skew, where different frequencies of activities are observed, but the feature-to-label relationships remain constant across clients, label concept skew involves different feature-to-label distribution for the same classes [9]. For example, the activities of jogging and playing soccer are the same sedentary class (high-intensity) but have different walking patterns. This variability reflects the unique ways individuals perform the same activities, making it challenging to develop a single global model that accurately captures sedentary behaviors across all users [9]. Thus, the literature still needs to develop novel FL techniques and methodologies that tackle the challenges associated with sedentary behavior detection.

This paper explores the effectiveness of using time series dynamics from wearable sensors in identifying sedentary behavior. Time series analysis is particularly relevant in this context, as the data collected by wearable devices inherently form time-ordered sequences. These sequences, or time series, provide valuable insights into various patterns of human activity. By analyzing these time series, mainly focusing on activity intensity, we can distinguish between different levels of physical engagement. The dynamics of the data change with varying activity intensities; low-intensity activities often produce data that resemble random time series, while higher-intensity activities tend to show more correlated dynamics.

To achieve this goal, we propose a new lightweight feder-

Pedro H. Barros and Heitor S. Ramos are with the Computer Science Departement, Federal University of Minas Gerais, Belo Horizonte, BH, Brazil. e-mail: {pedro.barros, heitor}@dcc.ufmg.br.

Judy C. Guevara, Leandro Villas ,and Nelson L. S. da Fonseca are with the Institute of Computing, UNICAMP, Campinas, Brazil.

Daniel Guidoni is with Computing Science Department, UFOP, Ouro Preto, Brazil.

Manuscript received XXXX XX, XXXX; revised XXXX XX, XXXX.

ated learning model for sedentary behavior classification by characterizing the wearable time series dynamics to identify the intensity of the activities and, thus, sedentary behavior. Our proposal uses the Ordinal Patterns (OP) transformation [10] to characterize the dynamics of the wearable time series in a discrete probability vector. Ordinal patterns are simple and flexible for describing the dynamics of time series data. By computing the probabilities of different ordinal patterns, we can obtain valuable insights about the behavior and patterns of the data [10]. To the best of our knowledge, this is the first federated approach to identify sedentary behavior.

The ***major novelty*** of our work is the proposal of a novel model that creates a discriminative latent space using features extracted from the ordinal patterns' empirical probability distribution in a federated scenario. OP analysis often focuses on local patterns or short subsequences within the time series. This localized approach requires less computational effort than global analyses, considering the entire time series [10]. Therefore, our proposal is a lightweight model compared to other models from the literature.

We focus on learning a latent feature space that captures the essence of the distribution of the ordinal pattern and enables effective similarity comparisons. We devise our methods using three steps. Firstly, we propose a new meta-learning approach to estimate a federated representation encoder function using a similarity space [11] derived from a Siamese neural network [12]. Secondly, utilizing federated learning, we aggregate the meta-learning local models to estimate a global encoder function. Finally, we use the global encoder function estimated by our federated meta-learning approach to fine-tune a sedentary behavior classification model and obtain a personalized model for each user.

We evaluated our proposal using five publicly available wearable datasets and show that our proposal creates *models* with reasonable size and efficient *energy consumption*. It requires 78.74% fewer parameters and 49.95% less energy consumption compared to other existing methods. The contributions of this paper are:

(i) hypothesizing that sensor data has distinct dynamics, where sedentary behavior exhibits dynamics similar to random noise, while vigorous activities show more correlated behavior. We validated this hypothesis using information theory descriptors derived from OP.
(ii) proposing a new federated meta-learning representation that separates regions where (dis)/similar OP distributions lie together, aiming to mitigate bias among local models;
(iii) conducting extensive evaluations using five publicly available datasets, showing superior performance in terms of F1-score across all datasets.

This paper is organized as follows. Section II overviews related works in FL environments. Section III introduces our proposal. Section IV describes the methodology used to analyze the data. Section V presents the main results. Finally, Section VI concludes the paper.

II. Related Work

Recent approaches have focused on automatic feature-based methods, primarily driven by the success of deep learning. These methods employ neural networks (NNs) to learn nonlinear representations of features automatically. Therefore, it is worth noting that these automatic feature-based approaches often outperform the traditional hand-crafted feature descriptors such as mean and spectral entropy.

Xiao et al. [8] designed a feature extraction network for each user, incorporating a convolutional neural network layer to discover local spatial features and a relation network that utilizes long-term memory and attention mechanisms to uncover global relationships in the data. The authors then attach a dense layer to classify the labeled data using the extracted features. Similarly, Ek et al. [13] proposed a federated self-supervised learning model for human activity recognition (HAR) in a pervasive computing application with heterogeneous datasets.

Traditional FL models face two fundamental challenges: poor convergence on heterogeneous data and a need for more personalized solutions. Motivated by these issues, researchers have proposed personalized federated learning approaches to address the statistical heterogeneity of client data distributions.

Loss regularization enhances the stability and generalization of machine learning models in federated learning, enabling improved convergence and personalized model generation by mitigating the effects of local updates. Li et al. [14] design FedProx to address heterogeneity in federated learning by introducing a proximal term (l_2-norm) to estimate the dissimilarity between the global FL model and local models to adjust the impact of local updates. Karimireddy et al. [5] proposed SCAFFOLD. This novel approach estimates the gradient update direction for server and client models using variance reduction to alleviate the client drift (divergence between the local and global models), which is used to correct the local update. Such variance reduction was extended in [15].

Recently, some novel meta-learning proposals aim at improving the learning algorithm by exposure to various tasks [4]. Meta-learning aims at creating a machine-learning model with minimal training examples to learn new concepts and skills. The authors [4] introduce a new federated meta-learning approach using Moreau envelopes. This paper includes an l_2-norm regularization loss that adjusts the trade-off between personalization and generalization performance.

Li et al. [7] proposed Meta-HAR, a federated representation learning framework for human activity recognition. The framework consists of a shared embedding network meta-trained by federated learning, capturing the heterogeneous data across devices and personalized classification networks tailored to each device for activity prediction using contrastive learning.

Zhang et al. [6] propose incorporating weight uncertainty into both the neural networks on the client devices and the central server using Bayesian neural networks. This uncertainty helps mitigate models' risk of noise fitting in the local datasets.

FedMask [16] empowers each device to acquire a personalized and structured sparse model by applying the learned binary mask to the fixed parameters of its local model. With FedMask, each mobile device can learn a personalized and structured sparse DNN, optimizing on-device efficiency.

This section highlights relevant work in federated activity classification and personalized federated learning, address-

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2024.3420891

ing challenges posed by the heterogeneity of user datasets. Feature-based approaches have shown potential in human activity recognition but need assistance with diverse activity types and signal distributions among users in traditional federated learning. Personalized federated learning combines federated learning with personalized adaptation to leverage diverse user data and improve generalization.

These approaches contribute to understanding personalized FL in human activity recognition. However, further research is needed to address the specific challenges of detecting sedentary behavior in federated learning, which still needs to be explored due to data heterogeneity. This heterogeneity stems from several factors, notably the mapping of human activities into sedentary behaviors, when considering these activities as super classes [9], i.e., for the same class, we would be significant variations in sensor data patterns. Personalized FL is crucial for tackling the unique challenges of sedentary behavior detection in this context.

In summary, our research addresses the detection of sedentary behavior via federated learning. To the best of our knowledge, this is the first federated approach to classify sedentary behavior. By introducing personalized federated learning methods and leveraging ordinal patterns as feature extractors, we contribute to developing practical and privacy-preserving approaches for mitigating sedentary behavior and promoting healthier lifestyles.

III. Our proposal

A. Proposed Algorithm

The proposed FL sedentary behavior classification method comprises three key steps, as illustrated in Figure 1. Our approach utilizes wearable time series data and ensures accurate sedentary behavior classification while prioritizing data privacy and battery efficiency through FL techniques.

In the first step, we employ a feature extraction method to transform raw time series data into a specialized feature space designed for capturing sedentary behavior patterns (see Section III-B). A meta-learning approach, using an encoder neural network in this new feature space, helps us achieve a more expressive and compact data representation (more details in Section III-C). It involves learning how to learn at a meta-level and using that knowledge to quickly adapt to specific tasks (representation learning) at a task level (classification) [7].

In the second step, we leverage federated learning to aggregate our encoder neural network. It preserves data privacy as only model parameters are shared between the central server and the devices, minimizing data transmission requirements.

The final step involves Personalization, where we fine-tune the sedentary behavior classification model by training a new last-layer neural network using local client data. This fine-tuning process enhances accuracy and adapts the model to individual user characteristics while retaining the shared knowledge acquired during the federated learning phase (more details in Section III-D).

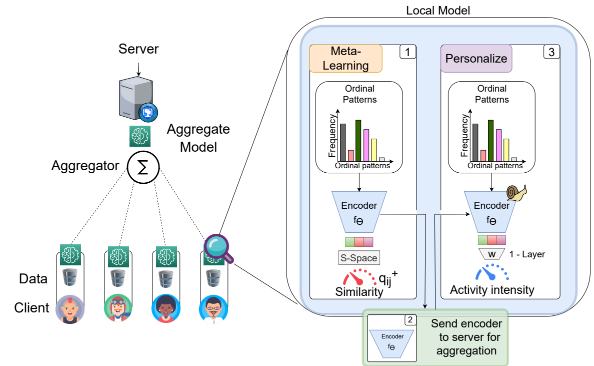

Fig. 1: A summary of the three-step process in our federated sedentary behavior classification approach, highlighting the key stages: Meta-learning with feature extraction and neural network encoding (Step 1), Federated learning training (Step 2), and Personalization with fine-tuning the encoder (represented by slug) with a low learning rate compared to the last neural network layer (Step 3).

This three-step process allows accurate and personalized sedentary behavior classification while respecting user privacy. The following subsections introduce notations for OP, data representation, federated learning, and mathematical definitions.

B. Feature Extraction

A time series is a discrete-time data sequence of data points indexed by time. More commonly, time series data are equally spaced in time. OP is a simple method of transforming time series that does not require any model assumption and can be applied to any arbitrary time series. The method is resistant to noise and invariant to nonlinear monotonic transformations. This approach involves mapping sliding windows of data points from a time series to vector symbols known as OP. The time series dynamics can be characterized by analyzing the frequency distribution of these OP [10].

Additionally, OP analysis involves transforming the original time series into symbolic sequences based on the order of values within a sliding window. This symbolic representation reduces the data dimensionality, making it computationally less demanding than raw time series data [10].

We hypothesize that sedentary activities have different dynamics. Low-intensity activities have a dynamic similar to random noise times series. In addition, when activities become more intense, the dynamics become more correlated. This characteristic can be captured using information theory descriptors obtained through OP. Therefore, OP is our feature extractor. Formally, given a time series $\mathcal{T}(t) = \{v_t\}_{t=1}^{n}$, an embedding dimension $D \in N$ and an embedding delay $\tau \in N$. At each instant t, we have a sliding window $w_t \subseteq \mathcal{T}(t)$ as $w_t = \{v_{t+i\tau}\}_{i=0}^{D-1}$. The sliding window w_t is mapped onto a vector symbol (ordinal pattern) $\boldsymbol{\pi}^D(w_t) = (R[v_{t+i\tau}])_{i=0}^{D-1}$ formed by the rank of its components, defined as $R[v_{t+i\tau}] = \sum_{k=0}^{D-1} \mathbb{1}(v_{t+i\tau} \geq v_{t+k\tau})$, where $1 \leq R[v_{t+i\tau}] \leq D$, and $\mathbb{1}$ is the indicator function: $\mathbb{1}[.] = 1$ if $[.]$ is true and 0 otherwise. In addition, $R(\mathtt{min}(w_t)) = 1$ and $R(\mathtt{max}(w_t)) = D$. Figure 2 shows all possible OP for $D = 3$.

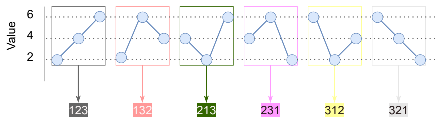

Fig. 2: Rank permutation mapping: The complete alphabet for D = 3 of the rank mapping technique is obtained by permuting all possible ranks.

For all $D!$ possible permutations $\boldsymbol{\pi}_i^D$, the probability of each ordinal pattern can then be estimated by simply computing the relative frequencies of the $D!$ possible permutations

$$p(\boldsymbol{\pi}^D) = \frac{|\boldsymbol{\pi}^D(w_i)|}{n-(D-1)\tau}, \tag{1}$$

where $|\boldsymbol{\pi}^D(w_i)|$ is the number of pattern observed of $\boldsymbol{\pi}^D(w_i)$.

Using ordinal patterns as feature extractors helps us capture the dynamics of sedentary behavior in wearable time series.

C. Latent Feature Representation

Supervised learning (SL) is a fundamental concept in machine learning, where a model learns a mapping between input data and corresponding output labels based on a labeled dataset. Let $\mathcal{D} = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^a$ be a training set consisting of a data examples, where $\boldsymbol{x}_i$ represents an OP distribution sample and y_i denotes the corresponding output label. In SL, the goal is to find a function $\ell : \mathcal{X} \to \mathcal{Y}$ that maps OP distribution from the input space $\mathcal{X}$ to the output space $\mathcal{Y}$. This function is parameterized by weights Ω, denoted as $\ell(\mathbf{x};\Omega)$. The function ℓ learns by minimizing the dissimilarity between the predicted output $\ell(\mathbf{x};\Omega)$ and the true output $\mathbf{y}$.

Consider the set $\mathcal{X}$ as the OP distribution space and the representation function $f_\Theta : \mathcal{X} \to \mathcal{Z}$, in which $f_\Theta(\boldsymbol{x}_i) = \boldsymbol{z}_i \implies \ell(\boldsymbol{x}_i) = \ell'(\boldsymbol{z}_i) = y_i$ and the function $\ell' : \mathcal{Z} \to \mathcal{Y}$, which maps the latent data into their respective labels. We defined the representation space $\mathcal{Z}$ called *latent feature space* from $\mathcal{X}$ as $\mathcal{Z} = \{\boldsymbol{z}_i\}_{i=1}^a$.

Siamese neural networks (SNN) are a specialized type of neural network architecture commonly used for tasks involving similarity or distance measurements between inputs [12]. A typical SNN approach is the distance d_{f_Θ} defined as $d_{f_\Theta}(x_i, x_j) = d(f_\Theta(x_i), f_\Theta(x_j))$, for a distance function d (e.g., euclidean distance or cosine similarity), and a neural network f_Θ [12]. Besides, a data pairwise $(\boldsymbol{x}_i, \boldsymbol{x}_j)$ are similar if $\ell(\boldsymbol{x}_i) = \ell(\boldsymbol{x}_j)$, i.e., the pair has the same label. Analogously, they are dissimilar if $\ell(\boldsymbol{x}_i) \neq \ell(\boldsymbol{x}_j)$ (different labels).

Following [3, 11], the representation space called *Similarity space*, or *S-space*, is a space built from the set $\mathcal{X} \times \mathcal{X}$ using SNNs. So, let f^S be the function $f^S : \mathcal{X} \times \mathcal{X} \to \mathcal{S}$, the *similarity space* is defined as $\mathcal{S} = \{\boldsymbol{s}_{ij}\}$. For *the similarity space* function f^S for a pair $(\boldsymbol{x}_i, \boldsymbol{x}_j)$, we define

$$\boldsymbol{s}_{ij} = f^S(\boldsymbol{x}_i, \boldsymbol{x}_j) = (\mid z_i^1 - z_j^1 \mid, \cdots, \mid z_i^n - z_j^n \mid). \tag{2}$$

Notably, vector $\boldsymbol{s}_{ij}$ is obtained by an element-wise process. It has the same dimension as $\boldsymbol{z}_i = f_\Theta(\boldsymbol{x}_i)$ and $\boldsymbol{z}_j = f_\Theta(\boldsymbol{x}_j)$, where z_i^n is the n-th feature of the i-th data example in a latent space representation $\mathcal{Z}$. Finally, we randomly sample the pair (x_i, x_j) to be used as input for our meta-learning model. In this context, we use a pair selection strategy where, given a training set $\mathcal{X}$ with a corresponding label set $\mathcal{Y}$, we select N pairs. Of these, $N/2$ pairs are similar (where x_i and x_j share the same label), and the remaining $N/2$ pairs are dissimilar (where x_i and x_j have different labels).

Our meta-learning proposal introduces the S-Space framework to represent pairwise input data in a latent and similarity space. We utilize two markers' sets: $\mathcal{M}^+$ for similarity and $\mathcal{M}^-$ for dissimilarity.

Therefore, we have k similarity markers and $n-k$ dissimilarity markers for the complete set of markers is denoted as $\mathcal{M} = \mathcal{M}^+ \cup \mathcal{M}^-$, with $\mathcal{M}^+ \cap \mathcal{M}^- = \emptyset$. We employ a discrete Cauchy distribution [11] as a kernel to measure the similarity between the S-vector $\boldsymbol{s}_{ij}$ and a marker $\boldsymbol{\mu}_m \in \mathcal{M}$ as

$$q_{ij}^m = \frac{\left(1 + ||\boldsymbol{s}_{ij} - \boldsymbol{\mu}_m||_2^2\right)^{-1}}{\sum_{\mu_{m'} \in \mathcal{M}} \left(1 + ||\boldsymbol{s}_{ij} - \boldsymbol{\mu}_{m'}||_2^2\right)^{-1}}. \tag{3}$$

The proximity of the S-vector $\boldsymbol{s}_{ij}$ to $\boldsymbol{\mu}^+$ in $\mathcal{M}^+$ signifies a higher degree of similarity between input data $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, while the distance to $\boldsymbol{\mu}^- \in \mathcal{M}^-$ represents dissimilarity. These markers are optimized during the process to estimate similarity or dissimilarity. We define $q_{ij}^+ = \sum_p q_{ij}^p$ for all $\boldsymbol{\mu}_p \in \mathcal{M}^+$ and $q_{ij}^- = \sum_n q_{ij}^n$ for all $\boldsymbol{\mu}_n \in \mathcal{M}^-$. By definition, q_{ij}^+ indicates the probability of a pair $(\boldsymbol{x}_i, \boldsymbol{x}_j)$ being similar (having the same labels). In contrast, q_{ij}^- represents the probability of a pair being dissimilar (having different labels). It should be noted that there are no overlapping elements between sets $\mathcal{M}^+$ and $\mathcal{M}^-$, ensuring that $q_{ij}^+ + q_{ij}^- = 1$.

The S-space optimization process simultaneously trains the markers' positions (set $\mathcal{M}$) and the Latent Feature Space with parameters Θ for the encoder functions. Therefore, we define the loss function $J(\mathcal{X} \times \mathcal{X}) = \sum_i \sum_j \left[u_{ij} \log q_{ij}^+ + (1-u_{ij}) \log q_{ij}^-\right]$, where $u_{ij} = 1$ if $\boldsymbol{x}_i$ has the same label as $\boldsymbol{x}_j$, and $u_{ij} = 0$ otherwise.

It is important to note that the loss function acts simultaneously in both the Latent Feature Space and the S-Space, grouping the elements $\boldsymbol{s}_{\boldsymbol{ij}}$ around their respective markers.

D. Personalized sedentarism classifier

We fine-tune the global model to adapt to each user's characteristics and preferences, as shown in Figure 1. The personalization process is performed after the meta-learning phase described in Section III-C. We designed a feedforward network $W_{\mathcal{X}_i}$ with one layer of dimensions d–c, where d is the latent feature dimension ($\mathcal{Z}$), and c is the neural network output (number of classes).

Formally, we consider a setting consisting of a global representation $f_\Theta : \mathcal{X} \to \mathcal{Z}$, which maps OP data points to a latent feature space, and client-specific heads $W_i : \mathcal{Z} \to \mathcal{Y}$. The model for the i-th user is defined as the function composition $W_i \circ f_\Theta(\boldsymbol{x})$, for a input $\boldsymbol{x}$. Based on [17], we split the model into base and personalized layers. The clients privately maintain the deep personalized layers for localized training to acquire task-specific personalized representations (W_i). In contrast, the base layers are shared with the FL server,

TABLE I: Characteristics of the datasets used in this work.

Dataset	Activities	Activities	Participants
BaSA [18]	7 daily activities	3 classes	15 volunteers
PAMAP2 [19]	18 daily activities	3 classes	9 volunteers
Open Dataset [20]	23 daily activities	2 classes	30 volunteers
Har UML 20 [21]	7 daily activities	3 classes	10 volunteers
mHealth [22]	12 daily activities	3 classes	7 volunteers

facilitating the learning of low-level features across clients (f_Θ).

Besides, we used mini-batch Stochastic Gradient Descent (SGD), where the learning rate is 0.01 and 0.001 for W_i and f_Θ, respectively. In addition, We initialize all weights of the neural network following [3]. Fine-tuning helps us adapt the model to individual user characteristics while retaining the shared knowledge acquired during the federated learning phase. Personalization enhances the accuracy and relevance of the sedentary behavior classification for individual users, providing a tailored experience.

IV. Experimental setup

A. Description of Dataset

This section provides the wearable datasets used in this study, each of which captures a wide range of physical activities individuals perform daily. Our analysis focuses exclusively on data generated by the accelerometer and gyro sensors, omitting additional sensor inputs. Table I offers a comprehensive overview of the datasets employed in this work. We provide a comprehensive overview of datasets used for sedentary behavior classification. These datasets encompass a variety of activities, participant groups, and sensor types, contributing to the study of sedentary behavior classification.

The Metabolic Equivalent of Task (MET) is the objective measure of the ratio of the rate at which a person expends energy relative to the mass of that person while performing some specific physical activity. In our study, MET values serve as critical thresholds to map human activities into sedentary classes within the federated learning model. The MET values of the Compendium of Physical Activities will be used as a reference for the activity's protocols: low-intensity (< 3 METs), moderate-intensity ($3-6$ METs), and vigorous activity (> 6 METs) [23]. Each client's dataset is preprocessed to associate activity readings with corresponding MET-based intensity classes in the federated learning context. Therefore, the federated model is trained to discern the presence of sedentary behavior and its intensity level.

B. Implementation Details

We employ the Flower framework [24] to implement our solutions in federated learning. This paper determined the number of clients in the federated learning environment by the dataset's participant count. We created a corresponding client in the federated learning environment for each participant in the selected dataset. Our model evaluation takes place on a central server, and we conduct training using an NVIDIA Quadro RTX 6000 GPU (24 GB) for 100 epochs. In each training round, the central server selects four clients to train the local model, where each model uses only the local data.

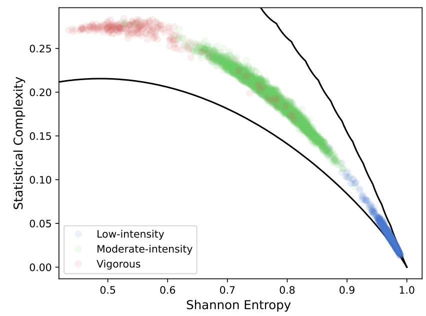

Fig. 3: Causality $H \times C$ plane for BaSA dataset.

In this paper, we randomly selected $N = 400$ pairs of data (200 similar and 200 dissimilar) for each local training epoch. The models are aggregated on the central server, which then distributes the aggregated model to the participating clients. The number of clients corresponds to the total number of participants in our federated learning environment.

This paper estimates the mean energy consumption for a single local training epoch of our personalized sedentarism classifier. We accomplish this through the integration of the pyRAPL toolkit[1], which utilizes Intel's RAPL technology to monitor energy usage across key CPU domains accurately. It is achieved by executing 40 independent training epochs for each model. The evaluation metric used in this work is the F1-score.

C. Parameters initialization and network architecture

We conducted a performance evaluation comparing our proposal with seven federated learning literature approaches (FedAvg [25], FedProx [14], pFedME [4], Meta-Har [7], FedDyn [15], pFedBayes [6], and FedMask [16]) , and we used two different neural network architectures: Our proposal and pFedBayes set network dimensions to d-512-n for all datasets, where d is the OP distribution dimension and n is the *latent space representation* dimension, and other approaches using architecture defined in [7], and we use $n = 64$ for this work.

Finally, the optimization model depends on some hyper-parameters (n, k), so we investigated which value of these variables could maximize the model's performance. We used the grid search technique for hyper-parameter optimization and found $k = 3$ similarity markers and $n - k = 3$ dissimilarity markers as the optimal settings. This finding indicates that our grid search spanned combinations of k and $n-k$ values from the set $\{1, 2, 3, 4, 5, 6\}$.

V. Results and Discussion

A. Causality Complexity-Entropy Plane

This section presents a comprehensive analysis of the results obtained using our proposed feature extractor. We employ the Causality Complexity-Entropy Plane ($H \times C$ plane) [10] to evaluate its effectiveness. This plane captures two key information theory descriptors: Shannon Entropy ($\mathcal{H}$) and

[1]https://github.com/powerapi-ng/pyRAPL

Statistical Complexity ($\mathcal{C}_{JS}$), which help quantify the disorder and the presence of correlational structures within the data.

Shannon Entropy measures the unpredictability of a system based on a probability distribution p (Equation 1), with lower values indicating regular or periodic processes and higher values suggesting uncorrelated stochastic processes [10]. Statistical Complexity complements entropy by quantifying the presence of correlational structures in OP distributions.

The H×C plane is an efficient characterization method, distinguishing between chaotic and stochastic dynamics in time series because probability distributions of OP have different shapes for different dynamics. In regular (periodic) processes, $\mathcal{H}$ and $\mathcal{C}_{JS}$ have values close to zero. On the other hand, uncorrelated stochastic processes have $\mathcal{H}$ close to one and $\mathcal{C}_{JS}$ close to 0. Correlated stochastic processes with f^{-k} power spectrum ($0 < k \leq 3$) have intermediate values [10]. Therefore, we focused on evaluating sedentary behavior through a machine-learning classification task.

We qualitatively evaluated our feature extraction proposal using the $H \times C$ plane with all activities in the BaSA dataset, as shown in Figure 3. For embedding dimension $D = 3$, the mean $\pm$ confidence interval (95% confidence level) Shannon entropy and Statistical Complexity $\mathcal{H} = 0.971 \pm 0.121$ and $\mathcal{C}_{JS} = 0.035 \pm 0.033$ for Low-intensity exercises. Similarly, $\mathcal{H} = 0.774 \pm 0.108$ and $\mathcal{C}_{JS} = 0.193 \pm 0.059$ for Moderate-intensity exercises and $\mathcal{H} = 0.559 \pm 0.121$ and $\mathcal{C}_{JS} = 0.264 \pm 0.047$ for vigorous exercises. Our $H \times C$ plane visualization demonstrates our feature extractor's capability to evaluate sedentary behavior effectively using wearable data. The patterns observed in Shannon entropy and Statistical Complexity across varying exercise intensities highlight our method's precision in classifying sedentary behavior, enhancing our understanding and informing future analysis.

B. Energy Consumption

This section estimates the energy consumption of various machine learning models using the pyRAPL framework. The experiments assess the energy efficiency of these models, considering the energy expended during one epoch, encompassing data preprocessing, feature extraction, and model training/evaluation. Table II compares different machine learning models based on their size (number of learnable parameters) and energy consumption, representing the energy consumption in joules (J) per epoch.

Firstly, the pFedBayes approach exhibited the highest energy consumption value. This model employs Gaussian approximations to estimate the weight, resulting in two parameters for each weight (representing the mean and standard deviation of the associated Gaussian). During training and prediction, this model utilizes neural network sampling via Markov Chain Monte Carlo (MCMC) for probabilistic estimation, which justifies its higher energy consumption.

Next, we analyzed our proposal and Meta-HAR, two meta-learning approaches presented in this article. Our proposal exhibited lower energy consumption compared to Meta-Har. Specifically, our model has 78.73% fewer parameters than Meta-Har, resulting in 48.67% less energy consumption. Compared to FedAvg, the literature model with the lowest energy consumption, our proposal consumes 22.55% less energy. This reduction in energy consumption can be attributed to the fact that our approach extracts features from time series using OP, which enables us to employ a much smaller neural network without the need for more complex feature extractors such as CNNs used in previous proposals. These results highlight the potential of our proposed method in achieving energy-efficient machine learning for sedentary classification applications. By reducing energy consumption, our approach can extend the battery life of wearable devices and enable energy-efficient systems in resource-constrained environments.

TABLE II: Comparison between different machine learning models based on their model size and energy consumption

Proposal	Model size (# of parameters)	Energy consumption (J)
FedAvg	504,244	12.828
FedProx	504,244	13.863
pFedME	504,244	13,687
Meta-Har	504,244	19,355
FedDyn	504,244	13,292
pFedBayes	214,404	27,185
FedMask	504,244	11,238
Our proposal	107,266	9.935

TABLE III: Ablation performance comparison (F1-Score)

Proposal	Dataset				
	BaSA	PAMAP2	Open dataset	Har UML20	mHealth
Meta-Har	0.9695	0.8358	0.8646	0.9099	0.9467
Meta-Har + OP	0.9789	0.8747	0.8798	**0.9197**	0.9373
Our proposal w/o OP	**0.9842**	0.9249	0.9347	0.9103	0.9671
Our proposal with OP	0.9823	**0.9417**	**0.9979**	0.8995	**0.9692**
local only (Mean)	0.9343	0.8822	0.99	0.8759	0.8437

C. Quantitative results

1) Ablation Analysis: In this section, we present an ablation analysis to evaluate the impact of OP on the performance of our proposed method compared to Meta-Har and the centralized learning model (local only). Table III summarizes the performance comparison across different datasets. We used the neural network described in Section IV-C for the proposal to use OP. Moreover, we use the CNN extract feature for proposals without OP.

Including OP in our proposal shows significant improvements across multiple datasets. We observe notable performance gains when comparing "Our proposal w/o OP" to "Our proposal with OP". Specifically, incorporating OP enhances the proposal's F1-score on PAMAP2, Open dataset, and mHealth by 1.78%, 6.33%, and 0.21%, respectively. Similarly, adding OP also improves performance when comparing "Meta-Har" to "Meta-Har + OP", including OP results in F1-score enhancements of 0.96%, 4.44%, 1.72%, and 1.06% for BaSA, PAMAP2, Open dataset, and Har UML 20, respectively. These results reinforce the significance of incorporating OP in the Meta-Har approach, indicating that they contribute to more accurate representations and improved performance.

In addition to the impact of OP, our proposed method outperforms both Meta-Har and Meta-Har + OP on four datasets. Specifically, "Our proposal with OP" achieves the highest F1-score across PAMAP2, Open dataset, and mHealth datasets, with improvements of 1.78%, 11.83%, and 3.29%.

TABLE IV: Performance comparison (F1-Score). The best results are in **bold**.

Proposal	Datasets				
	BaSA	PAMAP2	Open dataset	Har UML 20	mHealth
FedAvg.	0.7893	0.5398	0.4821	0.6981	0.6782
FedProx	0.9417	0.8961	0.8902	0.8984	0.9262
pFedME	0.8894	0.9378	0.8601	0.7927	0.9406
Meta-Har	0.9695	0.8358	0.8646	**0.9099**	0.9467
FedDyn	0.8950	0.7183	0.7885	0.8004	0.9385
pFedBayes	0.7738	0.8418	0.9628	0.8132	0.8409
FedMask	0.9242	0.8559	0.8793	0.7657	0.9420
Our proposal	**0.9823**	**0.9417**	**0.9979**	0.8995	**0.9692**

Finally, our FL model surpasses the performance of centralized models across all datasets, as shown in Table III.

In summary, OP significantly enhances the performance of our proposal, indicating the importance of incorporating them in time series analysis. Furthermore, our proposed method outperforms both Meta-Har and Meta-Har + OP, achieving a higher F1-score on various datasets and demonstrating the effectiveness of our combined approach.

2) General results: Table IV presents a comprehensive performance comparison of various federated learning methods based on the F1-Score metric across multiple datasets. In our comparative analysis, we observed that Meta-Har outperforms our proposed personalized federated learning approach on the Har UML 20 dataset, where Meta-Har achieves an F1-score of 0.9099. In contrast, our approach achieves a slightly lower F1-score of 0.8995. This finding suggests that Meta-Har, explicitly designed for human activity recognition, has certain advantages regarding sedentary behavior recognition in this dataset, although these advantages yield only marginally better results. However, it is essential to note that our proposed approach offers a distinct contribution by utilizing a simple feed-forward neural network instead of depthwise separable convolutions, which Meta-Har and several other existing approaches employ. This architectural difference makes our approach more accessible and easier to implement without sacrificing overall performance. Our proposal excels on other datasets, such as BaSA, where it achieves the highest F1-score of 0.9823, indicating its superiority over all other methods. Furthermore, our approach achieves an F1-score of 0.9979 on the Open dataset, demonstrating its effectiveness in handling sedentary behavior data from diverse sources.

Compared with other existing methods, our proposed personalized federated learning approach consistently ranks among the top performers across multiple datasets. It outperforms the baseline FedAvg, FedProx, pFedME, FedDyn, and pFedBayes on all datasets. Our approach achieves high F1-scores on the PAMAP2 dataset (0.9417) and the mHealth dataset (0.9692), evincing its robustness in capturing sedentary behavior information from different sources.

Therefore, the authors in [9] suggest that where activities under the same superclass vary significantly can be more challenging than usual in FL settings [6, 7, 14] because the high discrepancy between local data distributions, indicating higher heterogeneity. Refining our classification algorithms to manage these differences enhances model precision and reliability in federated learning settings, thus justifying our results. These results underscore the effectiveness and competitiveness of our proposed approach for personalized federated learning in sedentary behavior recognition. Figure 4 shows the latent space found by our proposal for the PAMAP2 dataset on the training dataset. We projected the *Latent Feature Space* into a 2-dimensional space using PCA for visualization. We have evidence that our proposal groups similar-intensity activities. These results corroborate our hypotheses that our proposal enhances class separability.

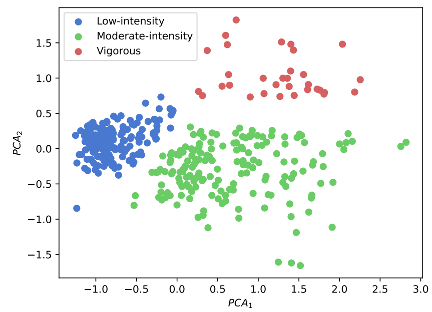

Fig. 4: Visualization of the latent space estimated by our approach for the PAMAP2 dataset.

VI. Conclusion

This paper introduces a personalized federated learning approach for sedentary behavior recognition. Our method leverages ordinal patterns (OP) and meta-learning in similarity space to capture the dynamics of sedentary activities. We hypothesized that low-intensity activities exhibit noise-like dynamics. In contrast, higher-intensity activities have more correlated dynamics, as evidenced by OP-based information theory descriptors. Our approach consistently outperformed baseline methods on multiple wearable datasets, achieving an F1-score of 0.9823 on the BaSA dataset and 0.9979 on the Open dataset, confirming our initial hypothesis.

Our work contributes to the field by offering a significant advantage in model simplicity, accessibility, and energy efficiency. Our model has 78.73% fewer parameters and consumes 48.67% less energy than the best-performing model in the literature. These findings emphasize the practicality of our approach for real-world applications aimed at promoting physical activity and improving health outcomes.

In conclusion, our personalized federated learning approach, combining OP and meta-learning, shows promise for accurate sedentary behavior recognition across diverse wearable datasets. Its superior performance and energy efficiency make it a viable solution for addressing sedentary behavior-related challenges and driving positive health outcomes. Future research may extend this approach to related domains and explore its applicability in broader contexts.

Acknowledgment

This work was partially funded by project "Hub de Inteligência Artificial em Saúde e Bem-Estar – VIVA BEM", funded by Samsung Eletrônica da Amazônia Ltda, under the terms of IT Law (Federal Law 8.248/91), São Paulo Research

Foundation (grant #2023/00721-1), Conselho Nacional de Desenvolvimento Científico e Tecnológico (grant #312682/2021-2), Fundação de apoio da UFMG – Fundep (grant # 29271) and Fundação de Amparo a Pesquisa do Estado de Minas Gerais (grant #APQ-00426-22).

References

[1] M. T. Hamilton, D. G. Hamilton, and T. W. Zderic, "Sedentary behavior as a mediator of type 2 diabetes," *Diabetes and Physical Activity*, vol. 60, pp. 11–26, 2014.

[2] E. García-Martín, C. F. Rodrigues, G. Riley, and H. Grahn, "Estimation of energy consumption in machine learning," *Journal of Parallel and Distributed Computing*, vol. 134, pp. 75–88, 2019.

[3] P. H. Barros and H. S. Ramos, "A novel aggregation method to promote safety security for poisoning attacks in federated learning," in *2022 IEEE Global Communications Conference (GLOBECOM)*, 2022, pp. 3869–3874.

[4] C. T. Dinh, N. H. Tran, and T. D. Nguyen, "Personalized federated learning with moreau envelopes," in *Conference on Neural Information Processing Systems*, ser. NIPS'20, Red Hook, NY, USA, 2020.

[5] S. P. Karimireddy, S. Kale, M. Mohri, S. Reddi, S. Stich, and A. T. Suresh, "SCAFFOLD: Stochastic controlled averaging for federated learning," in *Proceedings of the 37th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, H. D. III and A. Singh, Eds., vol. 119, 2020, pp. 5132–5143.

[6] X. Zhang, Y. Li, W. Li, K. Guo, and Y. Shao, "Personalized federated learning via variational bayesian inference," in *International Conference on Machine Learning*. PMLR, 2022, pp. 26 293–26 310.

[7] C. Li, D. Niu, B. Jiang, X. Zuo, and J. Yang, "Meta-har: Federated representation learning for human activity recognition," in *Proceedings of the Web Conference 2021*, ser. WWW '21, 2021, p. 912–922.

[8] Z. Xiao, X. Xu, H. Xing, F. Song, X. Wang, and B. Zhao, "A federated learning system with enhanced feature extraction for human activity recognition," *Knowledge-Based Systems*, vol. 229, p. 107338, 2021.

[9] J. Zhu, X. Ma, and M. B. Blaschko, "Confidence-aware personalized federated learning via variational expectation maximization," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2023, pp. 24 542–24 551.

[10] O. A. Rosso, H. A. Larrondo, M. T. Martin, A. Plastino, and M. A. Fuentes, "Distinguishing noise from chaos," *Phys. Rev. Lett.*, vol. 99, p. 154102, Oct 2007.

[11] P. Barros, F. Queiroz, F. Figueiredo, J. A. D. Santos, and H. Ramos, "A new similarity space tailored for supervised deep metric learning," *ACM Transactions on Intelligent Systems and Technology*, vol. 14, 2022.

[12] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," in *ICML Deep Learning Workshop*, vol. 2, 2015.

[13] S. Ek, R. Rombourg, F. Portet, and P. Lalanda, "Federated self-supervised learning in heterogeneous settings: Limits of a baseline approach on har," in *2022 IEEE PerCom Workshops*, 2022, pp. 557–562.

[14] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," *Proceedings of Machine learning and systems*, vol. 2, pp. 429–450, 2020.

[15] A. E. Durmus, Z. Yue, M. Ramon, M. Matthew, W. Paul, and S. Venkatesh, "Federated learning based on dynamic regularization," in *International Conference on Learning Representations*, 2021.

[16] A. Li, J. Sun, X. Zeng, M. Zhang, H. Li, and Y. Chen, "Fedmask: Joint computation and communication-efficient personalized federated learning via heterogeneous masking," in *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, ser. SenSys '21, New York, NY, USA, 2021, p. 42–55.

[17] L. Collins, H. Hassani, A. Mokhtari, and S. Shakkottai, "Exploiting shared representations for personalized federated learning," in *International conference on machine learning*. PMLR, 2021, pp. 2089–2099.

[18] H. Leutheuser, S. Doelfel, D. Schuldhaus, S. Reinfelder, and B. M. Eskofier, "Performance comparison of two step segmentation algorithms using different step activities," in *2014 11th International Conference on Wearable and Implantable Body Sensor Networks*, 2014, pp. 143–148.

[19] A. Reiss and D. Stricker, "Creating and benchmarking a new dataset for physical activity monitoring," in *Proceedings of the 5th International Conference on PErvasive Technologies Related to Assistive Environments*. Association for Computing Machinery, 2012.

[20] W. Possos, R. Cruz, J. D. Cerón, D. M. López, and C. H. Sierra-Torres, "Open dataset for the automatic recognition of sedentary behaviors," *Studies in health technology and informatics*, vol. 237, p. 107—114, 2017.

[21] O. Barut, L. Zhou, and Y. Luo, "Multitask lstm model for human activity recognition and intensity estimation using wearable sensor data," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8760–8768, 2020.

[22] O. Banos, R. Garcia, J. A. Holgado-Terriza, M. Damas, H. Pomares, I. Rojas, A. Saez, and C. Villalonga, "mhealthdroid: a novel framework for agile development of mobile health applications," in *6th International Work-Conference, IWAAL 2014,*. Springer, 2014, pp. 91–98.

[23] B. E. Ainsworth, W. L. Haskell, S. D. Herrmann, N. Meckes, D. R. Bassett Jr, C. Tudor-Locke, J. L. Greer, J. Vezina, M. C. Whitt-Glover, and A. S. Leon, "2011 compendium of physical activities: a second update of codes and met values," *Medicine & science in sports & exercise*, vol. 43, no. 8, pp. 1575–1581, 2011.

[24] D. J. Beutel, T. Topal, A. Mathur, X. Qiu, J. Fernandez-Marques, Y. Gao, L. Sani, K. H. Li, T. Parcollet, P. P. B. de Gusmão, and N. D. Lane, "Flower: A friendly federated learning research framework," 2020.

[25] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017, pp. 1273–1282.