Power System Technology
Volume 50, No. 1, 2026
ISSN: 1000-3673

fedHAR: Testbed Development for User-Specific Human Activity Recognition Using Federated Learning

.

Kazi Md. Shahiduzzaman and Md Salah Uddin Yusuf
Department of Electrical and Electronics Engineering, Khulna University of Engineering & Technology, Khulna, Bangladesh
Tanvir Ahamed, Lutfar Rahman and Md. Sajjad Hossen
Department of Electrical and Electronics Engineering, Jatiya Kabi Kazi Nazrul Islam University, Trishal, Mymensingh, Bangladesh

Abstract

As wearable devices become essential for everything from fitness tracking to elderly care, traditional Human Activity Recognition (HAR) systems face significant privacy risks and a "one-size-fits-all" approach that fails to recognize our unique movement patterns. To solve this, we developed fedHAR, a privacy-preserving framework that uses Federated Learning to train machine learning models directly on your device, ensuring your sensitive data never leaves your hands. When tested on a realistic 12-client network using the visual and movement-based vioHAR dataset, fedHAR's baseline global model achieved an impressive 92% accuracy, rivaling traditional centralized systems. But we didn't stop at a general model; by empowering the system to automatically categorize hidden, unlabeled data through pseudo-labeling and seamlessly personalizing itself to each user's unique biomechanics, overall accuracy leaped from 88.01% to an outstanding 97.02%. Combined with a lightweight architecture that can evaluate movements in under half a second, fedHAR delivers a highly responsive, personalized, and inherently private solution ready for real-world deployment.

Key Words: Human Activity Recognition (HAR), Federated Learning, Personalization Pseudo-labeling, Visual-Inertial Odometry (VIO), Cloud-Edge Architecture

1. Introduction

Human Activity Recognition (HAR) has emerged as a cornerstone of pervasive computing, serving as an enabling technology for a wide range of applications, including personalized health monitoring, intelligent environments, elder care, and fitness tracking [1]. By analyzing data from sensors embedded in smartphones, smartwatches, or dedicated wearables, HAR systems can infer a user's activities, such as walking, sitting, or running [2]. Traditionally, these systems operate under a centralized paradigm, where raw sensor data from numerous users is aggregated and uploaded to a cloud server for model training [3]. While this approach benefits from a large volume of data, it introduces significant challenges that hinder its scalability and widespread adoption.

The centralized paradigm faces three primary obstacles: data privacy, communication overhead, and data heterogeneity [4, 5]. In our previous research work, we have found that in the context of decentralized HAR, a multi-tier approach such as EdgeFall [6] and smart helmet architecture [7], is designed as a three-tiered system (end-edge-cloud or cloud-edge-end). This architecture distributes the computational load between sensors (end devices), edge nodes, and cloud servers. The end devices collect raw data, the edge nodes perform localized processing and activity recognition, and the cloud handles model development and potentially aggregates learned patterns [6], [7]. Another key advantage of this decentralized setup is the ability to offload processing to the edge. This significantly reduces data transmission to the cloud, thus lowering latency, pre-serving privacy, and

optimizing power consumption, which is crucial for real-time applications such as fall detection [6], [7]. In general, these papers advocate for a decentralized HAR approach to overcome challenges in traditional centralized systems, leading to more efficient, private, and real-time fall detection and activity monitoring for elderly care. However, questions remain about how to handle a new user, offload the training process, and ensure model performance on a lightweight device. Currently, we are able to complete a simulation work on simulating a user-centric model for human activity recognition using trajectory data [8]. The results show a remarkable improvement in classification or detection tasks as well as computational efficiency. These results also motivate us to develop a testbed on the cloud-edge or edge-cloud side for developing a user-centric model and handling new users.

Recent advancements have demonstrated that fusing multi-modal data can overcome these hardware limitations. For example, our development of Visual-Inertial Odometry (VIO) based systems, such as the vioHAR [9, 21] framework, has proven that fusing visual data with inertial measurements using state estimators (e.g., VINS-Mono) generates highly accurate 3D spatial trajectories. By capturing complex human movements independently of device orientation, VIO-based tracking provides a vastly superior signal for activity differentiation compared to raw IMU data.

Despite the availability of high-fidelity trajectory data, centralized HAR models inherently struggle with inter-subject variability. Activity patterns vary widely across individuals due to differences in age, fitness levels, and physiological traits. A generic, "one-size-fits-all" global model trained on aggregated data frequently fails to capture these nuances. This limitation was starkly highlighted in the vioHAR [9, 21] study, where rigorous Leave-One-Subject-Out (LOSO) evaluations revealed massive performance drops when global models encountered completely unseen users. Conversely, the uActivity [8] framework demonstrated that shifting to a user-centric approach, building personalized models tailored to an individual's unique trajectory patterns, can propel activity classification and fall detection accuracy to over 99%.

However, achieving this level of personalization traditionally requires transmitting high-dimensional, sensitive trajectory data to centralized cloud servers. In continuous monitoring applications like elderly care, this centralized paradigm poses severe privacy risks under regulations like the GDPR, introduces unacceptable communication latency, and strains the bandwidth of resource-constrained wearable edge devices. Federated Learning (FL) offers a compelling solution to these challenges by enabling collaborative machine learning without requiring direct sharing of raw data [10–12]. In this decentralized approach, devices (clients) train a model locally on their private data. Instead of transmitting sensitive data, only model updates (e.g. gradients or weights) are sent to a central server for aggregation. This methodology not only safeguards user privacy but also reduces communication bandwidth by only transmitting a lightweight model update and inherently addresses data heterogeneity by allowing the model to adapt to local data distributions.

To bridge the gap between the accuracy of personalized trajectory modeling and the necessity of data privacy, this paper proposes fedHAR: a privacy-preserving, user-specific Human Activity Recognition system powered by Federated Learning (FL). By decentralizing the training process, fedHAR allows edge devices to collaboratively learn general activity patterns while retaining raw VIO data locally. Furthermore, to address the scarcity of labeled data on edge devices and the heterogeneity of individual behaviors, our framework integrates a threshold-based local personalization strategy and a high-confidence pseudo-labeling mechanism. Ultimately, this work provides a scalable, edge-efficient solution that delivers the precision of user-centric models without compromising user privacy.

The main contributions of this work are as follows:

1. Development of a Robust Cloud-Edge Testbed: We designed and implemented a complete FL testbed for HAR, simulating a realistic, decentralized environment with multiple clients.

2. Comparative Analysis of Aggregation Algorithms: We evaluated the performance of four prominent FL aggregation algorithms (FedAvg, FedOpt, FedProx, and FedAdam) to identify the most suitable method for HAR.

3. Integration of Pseudo-Labeling: We successfully integrated a pseudo-labeling technique to leverage the vast amount of unlabeled data available on edge devices, thereby enhancing the global model's performance.

4. Implementation of a Personalization Strategy: We developed and validated a personalization strategy that enables clients to fine-tune the global model, resulting in superior user-specific performance.

5. Performance and Efficiency Evaluation: We provide a thorough analysis of the system's performance metrics and computational efficiency, demonstrating its real-time feasibility for deployment on resource-constrained devices.

2. Literature

Early research in HAR primarily focused on traditional machine learning approaches. These methods often relied on handcrafted features extracted from sensor data, which were then fed into classifiers such as Support Vector Machines (SVMs), K-Nearest Neighbors (KNN), and Long Short-Term Memory (LSTM) [6]. For instance, feature engineering techniques involve calculating statistical metrics, such as the mean, standard deviation, and energy, of the sensor signals [6, 13, 14]. While effective for simple activities, these approaches struggled with complex and nuanced human behaviors, requiring significant domain expertise. The field advanced significantly with the introduction of deep learning techniques that could automatically learn hierarchical features from raw sensor data, eliminating the need for manual feature extraction. Convolutional Neural Networks (CNNs) have been particularly successful in this domain due to their ability to capture spatial features. In contrast, we have examined that Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are well-suited for modeling temporal dependencies in time-series data [6]. The majority of conventional HAR datasets and systems rely exclusively on IMU data, which limits their robustness in dynamic, real-world environments. Recent literature has increasingly focused on multi-modal sensor fusion to provide richer contextual data. Our previous vioHAR study[9, 21] illustrated the critical advantage of this approach by fusing camera metadata with IMU sensor data to track 3D human trajectories. Their findings proved that visual-inertial trajectory data provides a vastly superior baseline for classification; in rigorous LOSO cross-validation, the VIO method achieved a 73.96% accuracy, effectively doubling the 43.35% baseline accuracy of standard accelerometer and gyroscope data alone. Furthermore, the study highlighted the computational advantages of LSTM networks for edge computing, which trained nearly ten times faster than Random Forest algorithms while maintaining high predictive power.

While robust data acquisition improves general model accuracy, centralized models fail to account for the unique physiological and lifestyle characteristics of individual users. The necessity of personalization is heavily supported by recent localized HAR studies. For instance, the uActivity[8] framework introduced a user-centric model deployed across a cloud-edge-end architecture specifically for elderly fall care. By utiliz- ing trajectory-based modeling to tailor the algorithm to specific users, uActivity[8] achieved an exceptional accuracy of 99.44% with a False Alarm Rate (FAR) of just 0.081%, drastically outperforming general ma- chine learning benchmarks. Similarly, while a global model in the vioHAR[9, 21] study achieved high accuracy on random data splits, its performance plummeted during testing on unseen users, rebounding to a near-perfect 99.16% only when transitioning to user-specific models.

The application of Federated Learning (FL) to HAR represents a recent and important evolution. Several studies have explored using FL to address the privacy concerns associated with centralized HAR systems. For example, some researchers have investigated FedAvg for training HAR models on mobile devices, demonstrating its ability to maintain privacy without a significant loss in accuracy [15–17]. Other works have focused on mitigating communication costs and model heterogeneity through client selection strategies or by modifying the model architecture. However, many existing FL-based HAR studies often overlook two critical challenges: the abundance of unlabeled data and the need for user-specific model personalization. A closely related piece of research work is FedHAR [18], a framework designed to address key challenges in Human Activity Recognition (HAR), including privacy preservation, limited labeled data, real-time processing, and heterogeneity among individuals. The proposed framework utilizes federated learning to maintain data on local devices, thereby protecting user privacy. To overcome the scarcity of labeled data, the framework introduces an algorithm that computes unsupervised gradients under a consistency training proposition and a strategy to aggregate these gradients. The study's extensive experiments on two real-world HAR datasets demonstrate that FedHAR outperforms other state-of-the-art methods with an average accuracy of 82.08%, and an F1-score is 80.90%. Additionally, the paper notes that when fine-tuning each unlabeled client, personalized FedHAR can achieve an average improvement of an additional 10% across all metrics. The average accuracy for person- alized FedHAR (PerFedHAR) across all clients is 92.68%, and the average F1-score is 92.32%. A client can significantly impact the performance of a global model with a small and unrepresentative dataset. However, many existing FL-based HAR studies often overlook two critical challenges: the abundance of unlabeled data and the need for user-specific model personalization. A client can significantly impact the performance of a global model with a small and unrepresentative dataset. Furthermore, a "one-size-fits-all" global model may not be the optimal solution for all users, given their varying physiological traits and behavioral patterns. The following Table 1 compares the proposed fedHAR with other related and recent reseach works related to HAR.

Table 1:

Hypothesized Performance of fedHAR Compared to Existing Federated Learning Approaches for HAR

Metrics	**Proposed fedHAR**	**Sozinov et al. (2018)[17])**	**Bettini et al. (2021)[19]**

Dataset	Visual Inertial Odometry Based HAR Dataset (vioHAR)	Heterogeneity HAR Dataset	MobiAct, WISDM
Number of Activities	8	6	5 from MobiAct, 6 from WISDM
Aggregation Technique	FedAvg, FedOpt, Fed- Prox, FedAdam	FedAvg	weighted average
Personalization Approach	Selective Saving (Threshold-Based)	Not implemented	Implemented using transfer learning strategy
Number of Clients	12	Not specified	Not specified
Model	FNN	DNN, softmax regres-sion model	Fully connected DNN, CNN
Rounds	100	1000	30
Unlabeled data handling technique	Pseudo Labeling	Not implemented	Semi-Supervised
Performance evaluation	Accuracy, Precision, Recall, F1-Score, FAR, System complexity for both user specific and global model	Achieved an accuracy of up to 89 % compared to 93 % in centralized train-ing	Evaluated personalization and generalization using F1 score

The key differences between the "FedHAR"[20] approach and the methodology outlined in our paper lie in the specific solutions implemented to address the challenges of real-world data developed in our lab (vioHAR). Here's a breakdown:

• The FedHAR[20] Approach: Most foundational FedHAR frameworks focus on the primary goal of federated learning: training a single, global HAR model collaboratively without centralizing raw data. They typically use a base aggregation algorithm, such as Federated Averaging (FedAvg). While this successfully handles the privacy and communication challenges, it often struggles with two significant issues:

1. Unlabeled Data: In real-world scenarios, a significant portion of the data on a user's device is unlabeled. Standard FedHAR doesn't have a mechanism to utilize this data.

2. Data Heterogeneity: A global model trained on data from many different users may not be accurate for an individual user, as people perform activities in slightly different ways. This is the "one-size- fits-all" problem.

• Our Paper's Approach: Our work goes beyond the basic FedHAR[20] framework by integrating two advanced techniques to solve these specific problems. This is where our unique contribution lies:

1. Pseudo-Labeling: We have introduced a semi-supervised learning technique where the global model is used to create "pseudo-labels" for the unlabeled data on each client. This effectively trans- forms a large amount of unused, unlabeled data into valuable training data, significantly improving the global model's performance.

2. Personalization: Our methodology includes a dedicated personalization strategy. Instead of relying solely on the global model, we enable each client to fine-tune a model using its own data, thereby creating a highly accurate, personalized model tailored to that specific user. This is a crucial step for real-world usability, as it ensures the HAR system is effective for each individual.

3. Real-time Processing with vioHAR[9, 21] Dataset: The implementation also prioritizes real-time processing capabilities, ensuring that the HAR system can deliver timely and actionable insights directly on the edge devices, which is critical for pervasive health monitoring and assistive technologies [6].

In essence, while FedHAR[20] provides the privacy-preserving foundation, our paper builds on that foundation with an advanced, multi-faceted solution that makes the HAR system more robust, efficient, and user-specific. We address the practical limitations of the general approach by incorporating pseudo-labeling and personalization, which is a major differentiator.

Even though recent advancements have made FL more suitable for HAR tasks, significant gaps remain. First, existing FL-HAR studies largely focus on idealized scenarios with homogeneous, strictly IMU- based data, failing to address the processing of complex, multi-modal VIO trajectory data at the edge. Second, the integration of semi-supervised pseudo-labeling combined with local personalization in a federated setup has not been thoroughly

investigated. This paper addresses these gaps by evaluating a realistic 12-client federated testbed utilizing the orientation-independent trajectory data from the vioHAR [9, 21] dataset. Our approach distinguishes itself from previous work by developing a comprehensive testbed that integrates a solution for both of these critical challenges. The introduction of a pseudo-labeling technique directly addresses the issue of unlabeled data, allowing the system to leverage a much larger portion of the available data to improve the global model. Simultaneously, our personalization strategy ensures that individual clients can fine-tune the global model to achieve a high degree of accuracy for their specific activity patterns, thereby bridging the gap between a robust global model and a highly accurate user-specific model. This holistic approach of developing a complete testbed with integrated solutions for these key challenges represents a significant contribution to the field of privacy-preserving HAR.

3. Methodology: The fedHAR Framework

This section delineates the architectural design, experimental setup, and evaluation metrics employed to validate the efficacy of our proposed federated learning framework for user-specific human activity recognition.

A. Proposed fedHAR Framework

The overall fedHAR system architecture, as illustrated in Fig. 1, operates in an iterative cycle between the central server and the clients.

The process begins with the server distributing a pre-trained global model. Each client then performs local training on its private dataset. This training process is enhanced by two key techniques:

• Pseudo-Labeling: To address the challenge of limited labeled data, we introduced a pseudo-labeling technique. After the first 20 communication rounds, when the global model has achieved a stable level of performance, clients use their local model to predict labels for their unlabeled data. Only predictions with a confidence score above 0.95 are used to create a pseudo-labeled dataset, which is then combined with the client's labeled data for subsequent local training.

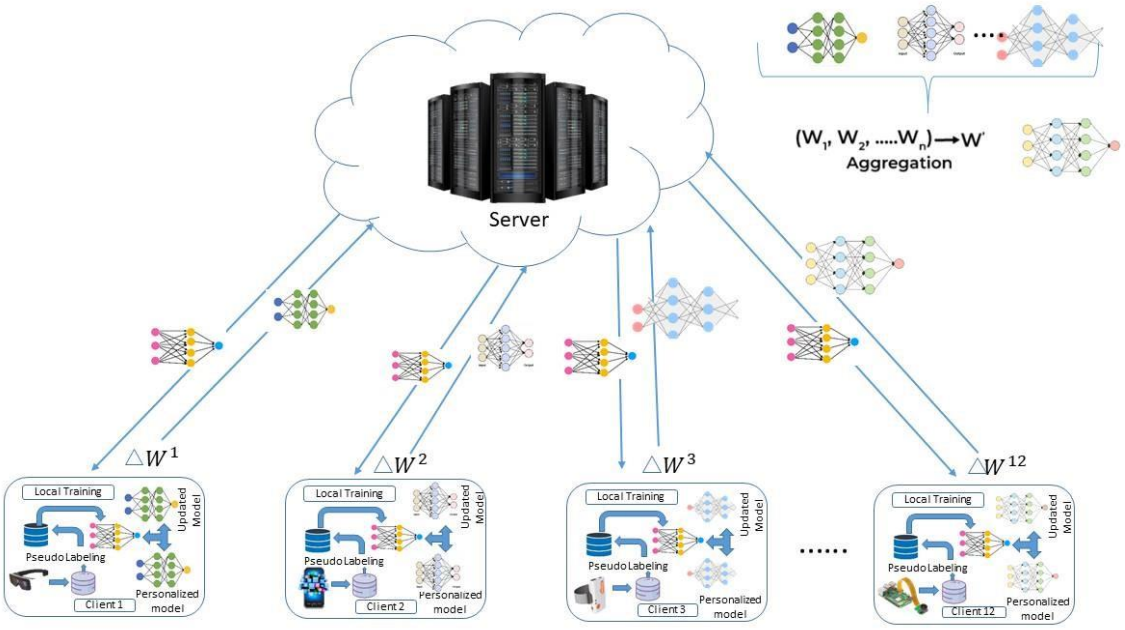

Fig. 1: Proposed fedHAR Framework

• Personalization: A personalization strategy is also integrated into the local training process. Each client evaluates its local model after training, and if the model's accuracy on a local validation set exceeds 95%, it is saved as a personalized model. This allows the system to adapt to individual user activity patterns.

After local training and personalization, clients send only the updated model parameters (weights) back to the server, preserving data privacy.

B. Dataset Description

We used the vioHAR dataset (developed in our lab), a novel HAR dataset that integrates visual and inertial data. The dataset was collected from 14 healthy adult subjects and features eight distinct activities: walking, running, jumping, sitting and standing, climbing stairs, descending stairs, falling forward, and falling backward. To simulate a non-IID (non-independent and identically distributed) scenario, the data was partitioned across 12 clients. Two clients were used to pre-train an initial global model, while the remaining 12 participated in the FL process. The total number of samples for each activity is not uniform. As shown in Fig. 2, the total number of samples for each of the eight activities varies significantly across the 14 clients, demonstrating a diverse and non-uniform data distribution.

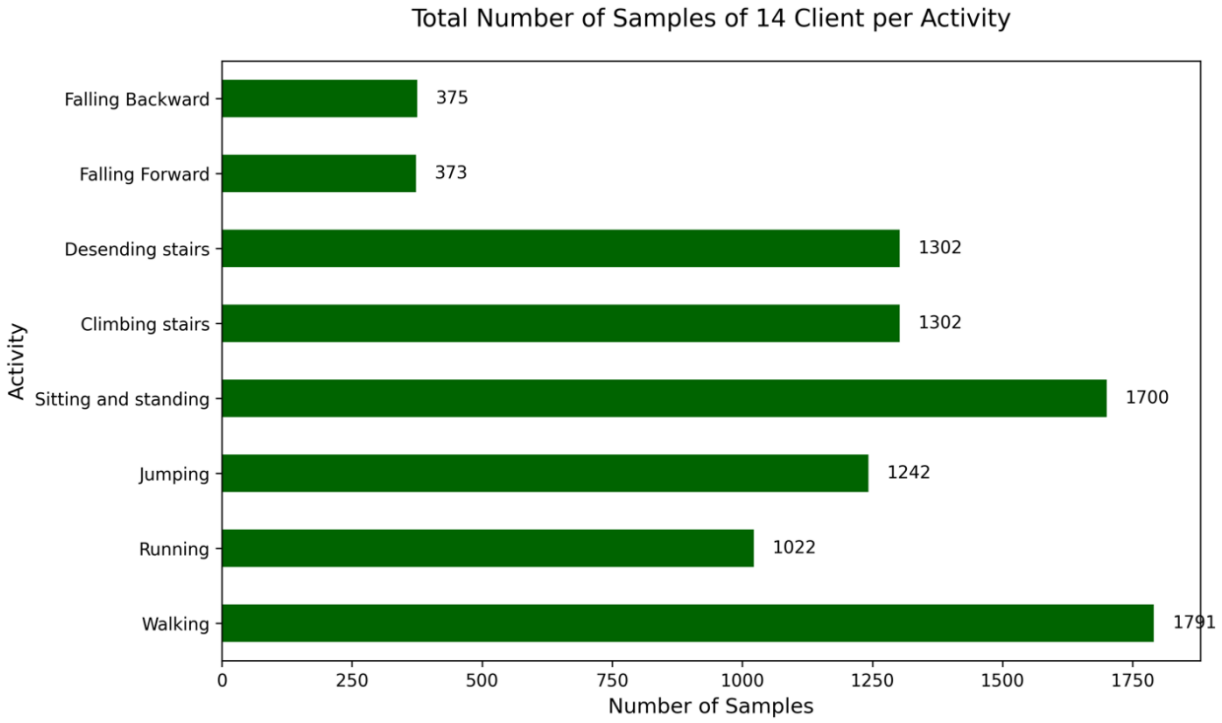

Fig. 2: Total Number of Samples of 14 Clients Per Activity

C. Global Model Aggregation

The central server aggregates the model updates from all participating clients to create a new global model for the next communication round. We evaluated four prominent aggregation algorithms to determine the most effective approach for our system:

• FedAvg (Federated Averaging): A simple but effective baseline that computes a weighted average of the client parameters.

• FedOpt (Federated Optimization): An optimization-based method that applies server-side optimizers to the aggregated updates to accelerate convergence.

• FedProx (Federated Proximal): A regularization-based approach that adds a proximal term to the local objective function to mitigate client drift in non-IID environments.

• FedAdam: A variant of FedOpt that uses the adaptive Adam optimizer on the server to improve convergence.

Algorithm 1: Federated Learning with Pseudo-Labeling for HAR (FedAvg)

Require: Clients K, rounds R, local epochs E, batch size B, learning rate η, pseudo-label threshold τ, personalization threshold ρ, validation set D_k^{val} for each client k, pseudo-label start round R_{pseudo}.
Ensure: Global model w_G, personalized models $\{w_k^{\text{personalized}}\}_{k=1}^{K}$
Pre-train global model w_G^0 using labeled data from two clients
Initialize best accuracy $a_k^{\text{best}} \leftarrow 0$ for each client k
for $r = 0$ **to** $R - 1$ **do**
  Server broadcasts w_G^r to all clients
  for each client $k \in \{1, \ldots, K\}$ **in parallel do do**
    $w_k^r \leftarrow w_G^r$
    Initialize $D_k^{\text{u}} \leftarrow \emptyset$
    if $r \geq R_{\text{pseudo}}$ and $D_k^{\text{u}} \neq \emptyset$ **then then** {Delayed pseudo-labeling}
      Generate pseudo-labels for D_k^{u}:
      for each $x_i \in D_k^{\text{u}}$ **do do**
        Compute predictions: $\hat{y}_{i,c} = p(y = c | x_i; w_k^r)$ for $c \in \{1, \ldots, C\}$
        if $\max_c \hat{y}_{i,c} \geq \tau$ **then then**
          $\hat{y}_i \leftarrow \max_c \hat{y}_{i,c}$
          $\tilde{y}_i \leftarrow \arg\max_c \hat{y}_{i,c}$
          Add $(x_i, \tilde{y}_i)$ to D_k^{u}
        end if
      end for
    end if{End pseudo-labeling}
    Form $D_k^{\text{aug}} \leftarrow D_k^l \cup D_k^u$
    Train on D_k^{aug} for E epochs with batch size B:
    for each epoch $e = 1$ **to** E **do**
      for each batch $B \subset D_k^{\text{aug}}$ of size B **do**
        $L_k \leftarrow \mathcal{L}_{\text{aug}}(w_k^e; B)$
        $w_k^e \leftarrow w_k^e - \eta \nabla L_k$
      end for
    end for{End local training}
    Evaluate w_k^r on D_k^{val}: $a_k \leftarrow \text{accuracy}(w_k^r, D_k^{\text{val}})$
    if $a_k \geq \rho$ and $a_k > a_k^{\text{best}}$ **then then**
      $w_k^{\text{personalized}} \leftarrow w_k^r$
      $a_k^{\text{best}} \leftarrow a_k$
    end if
    Send w_k^r to server
  end for{End client loop}
  Server aggregates:
  $w_G^{r+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{N} w_k^r$ where $N = \sum_{k=1}^{K} n_k$
end for{End round loop}
return w_G^R, $\{w_k^{\text{personalized}}\}_{k=1}^{K}$ =0

D. Adapted Algorithm for fedHAR

The algorithm outlines a federated learning process that iteratively trains a global model while also creating personalized models for each client. The process starts with a pre-trained model, which clients use for local training. A key feature is "delayed pseudo-labeling," which allows clients to generate new training data from their unlabeled samples once the model's confidence exceeds a threshold. Concurrently, a personalization strategy saves a client's local model if its accuracy surpasses a predefined level. The server then aggregates the updated models to create a new global model for the next round. Algorithm 1 outlines the system's overall process. This algorithm represents a functional simulation of a horizontal Federated Learning (FL) system designed for daily activity and fall classification using vioHAR dataset [9, 21]. The core of the implementation relies on the client-server architecture, where the central server manages the federated training rounds, including client selection and model aggregation via different Federated optimization algorithm. The multiple clients independently handle their local, private data and update their local models, a basic Convolutional Neural Network (CNN), using the Stochastic Gradient Descent (SGD) with momentum optimizer. Communication between the server (using threads) and the clients (using processes) is established over Python sockets, with a utility layer ensuring reliable transfer of the large serialized model weight files, thereby simulating a complete, asynchronous, and communication-efficient distributed training pipeline.

4. Experimental Results and Comparative Analysis

This section provides a detailed analysis of the experimental outcomes, comparing centralized and federated learning paradigms, evaluating various aggregation strategies, and demonstrating the efficacy of personalization and pseudo-labeling for Human Activity Recognition (HAR).

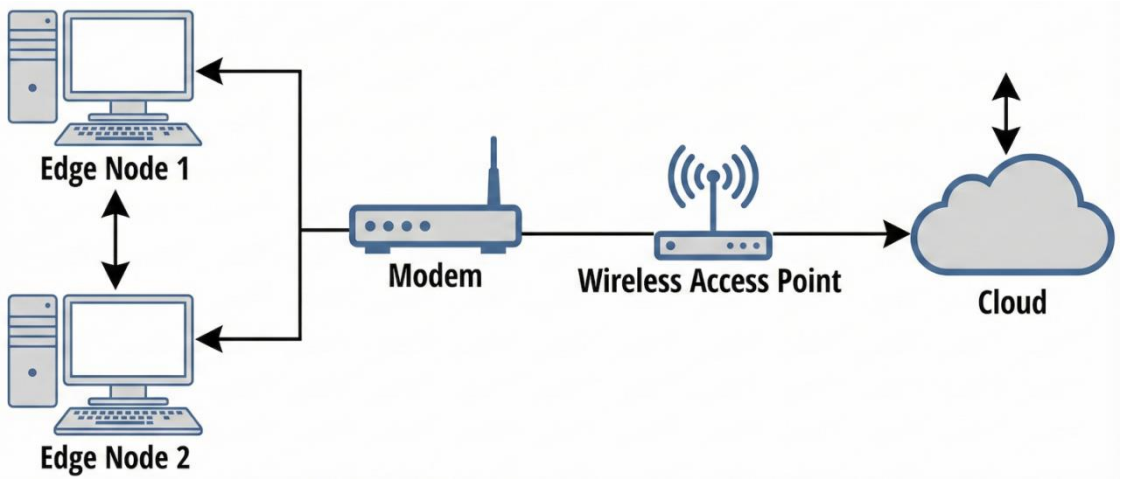

Fig. 3: Experimental setup to implement fedHAR on edge-cloud paradigm

A. Experimental Setup

To simulate a realistic, distributed federated learning environment, the experiments were conducted using three physical PCs connected over a local network. Fig. 3 illustrates the experimental setup. One PC served as the cloud, while the other two acted as edge nodes, each hosting six clients, for a total of 12 distributed clients. The configurations of these systems were as follows:

• Cloud: AMD Ryzen 5 56000G 3.90 GHz Processor, 32 GB DDR4 RAM, and 512 GB SSD. Its role was to serve as the cloud for aggregating model updates and managing communication.

• Edge Node 1: Intel(R) Core(TM) i5-7400 CPU @ 3.00GHz, 32.0 GB RAM, and 256 GB SSD. It hosted six federated learning clients.

• Edge Node 2: AMD Ryzen 5 3400G 3.70 GHz Processor, 16 GB RAM, and 228 GB SSD. It hosted six other federated learning clients.

The lab environment setup and computation process are shown in Fig. 4, ensuring that the experiments were conducted under conditions that closely mirrored a real-world decentralized network, accounting for hardware and network latency. The process, as depicted, is iterative and involves steps described in the Methodology.

• Cloud Workstation: The cloud workstation acts as the coordinator of the learning process. It begins by initializing a global model and then broadcasting copies of this model to a selected group of virtual clients.

• Local Training (Virtual Clients & Client PCs): Each "virtual client" and "client PC" represents a device or entity (like a mobile phone, computer, or a local server) that has its own local dataset. These clients train the model on their private, local data. A key aspect of this process is that the raw data remains on the client device and is never sent to the central server. Local training happens on individual devices, such as virtual clients and two client PCs, which are connected by a local network.

• Model Aggregation: After local training, each client sends its updated model parameters (such as weights or gradients) back to the central server. The central server receives these updates and aggregates them to create a new, improved version of the global model. The most common aggregation method is Federated Averaging (FedAvg), where the server computes a weighted average of the client updates.

(a) Experimental Setup in Lab Environment

(b) Runtime Instance of Computational Task in Cloud and Edge Nodes

Fig. 4: System Architecture showing the Experimental Setup in the lab environment and computational task distribution

• Global Model Broadcast: The central server then broadcasts this newly updated global model to a new set of clients for the next round of training. This cycle repeats until the global model reaches a desired level of accuracy or convergence.

The Fig. 4b illustrates the workflow of the lab environment distributed experimental setup, which consists of a central Cloud server (orchestrator) and distinct Edge Nodes connected via a wireless access point. In this Federated Learning framework, the process begins with the Cloud server initializing the global model and transmitting computational instructions or parameters downstream to Edge Node 1 and Edge Node 2. The Edge Nodes then execute local training using their private local data, as evidenced by the computational task logs. Once local processing is complete, the Edge Nodes transmit their computed model updates (gradients or weights) upstream back to the Cloud server. The Cloud aggregates these updates to refine the global model, completing a synchronized cycle of distributed intelligence without centralizing the raw data.

B. Performance Evaluation Metrics

We assessed the proposed fedHAR with the use of traditional machine learning methods and deep learning models, noting exceptional results. We implemented a 50-50 split between training and testing for this evaluation. This fair method guarantees that the model gets trained on a portion of the dataset. In contrast, the remaining portion evaluates its capacity to generalize to novel, unobserved data proficiently. We have adopt the following performance evaluating metrics (similar lin vioHAR):

• Accuracy is the proportion of correctly classified instances to the total and functions as a fundamental metric.

• A metric of precision is the extent to which affirmative forecasts are correct. Precision is the ratio of true positives to the sum of true and false positives.

• Recall quantifies the proportion of accurately predicted positive outcomes, calculated as true positives divided by the sum of true positives and false negatives.

• A crucial component of classifier evaluation is the F1-Score, a composite metric that combines accuracy and recall into a single metric.

• The False Alarm Rate (FAR), also known as the false positive rate, is a metric that measures the model's tendency to incorrectly predict the positive class. To better understand the model’s specificity, the false-positive rate is defined as the ratio of false positives to the sum of false positives and true negatives.

• Aggregation Processing Time is the time required by the server to generate an update for the global model to broadcast.

• The time required to complete Human activity recognition and classification (including human falls) is considered as Execution Time.

C. Centralized vs. Federated performance

The evaluation of the fedHAR system begins with a comparative analysis between the traditional centralized learning paradigm and the proposed federated framework. The centralized machine learning model with LSTM, which serves as the performance benchmark, was trained on a unified dataset where all sensor information from 14 subjects was aggregated on a single server. The following Fig. 5 shows that this model is capable of achieving a high accuracy of 94% and exhibits stable convergence as it benefits from access to the complete global data distribution. However, while centralized training offers high performance, it faces significant hurdles regarding data privacy, as it requires users to upload sensitive raw sensor data to a remote cloud. The centralized machine learning model with LSTM, which had access to the full dataset, serves as a performance benchmark.

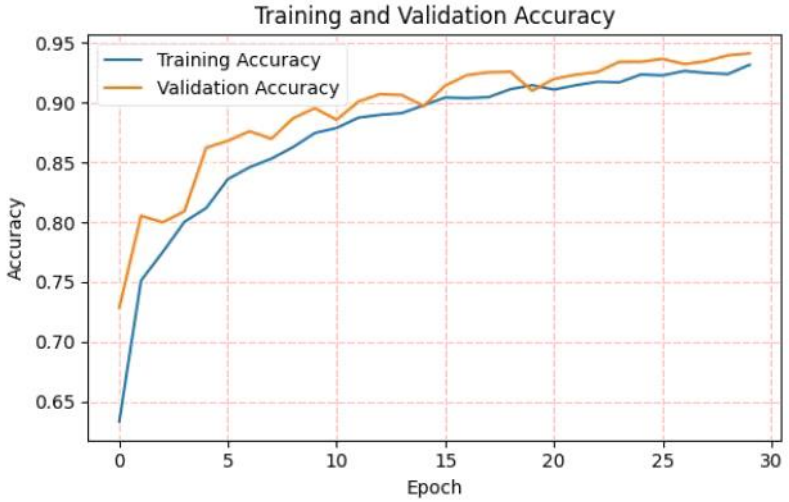

Fig. 5: Accuracy achieved by the LSTM-based centralized model

In contrast, the federated learning approach enables collaborative model training while keeping raw data strictly on the user's device. The experiments conducted on the 12-client testbed showed that federated models can achieve performance parity with centralized benchmarks while preserving data locality. The outcomes of the training process are shown by the following depictions (Fig. 6) with FedAvg, FedOpt, FedAdam, and FedProx approaches. Using the FedAvg algorithm, the global model reached a peak accuracy of 92% after 100 communication rounds. This indicates that decentralized training can effectively learn general activity patterns, such as walking, running, and jumping, without compromising user privacy. The slight discrepancy between centralized and federated accuracy (94% vs. 92%) is often attributed to the non-IID (non-independent and identically distributed) nature of the data partitioned across the clients. In real-world HAR applications, data heterogeneity arises from variations in how individuals perform the same activity, driven by their unique physiological traits and behavioral patterns. Despite these challenges, the federated approach proved robust, with algorithms like FedAdam achieving a significantly lower training loss of 0.23, suggesting a more stable and efficient convergence than simple averaging techniques. Table 2 summarizes the overall observations comparing the Centralized Paradigm and the Federated Learning (fedHAR) system.

D. Effectiveness of Aggregation Techniques

The choice of aggregation algorithm significantly influenced the performance and convergence of the fedHAR system. We evaluated four prominent techniques, FedAvg, FedAdam, FedOpt, and FedProx, over 100 communication rounds to determine the optimal strategy for decentralized human activity recognition.

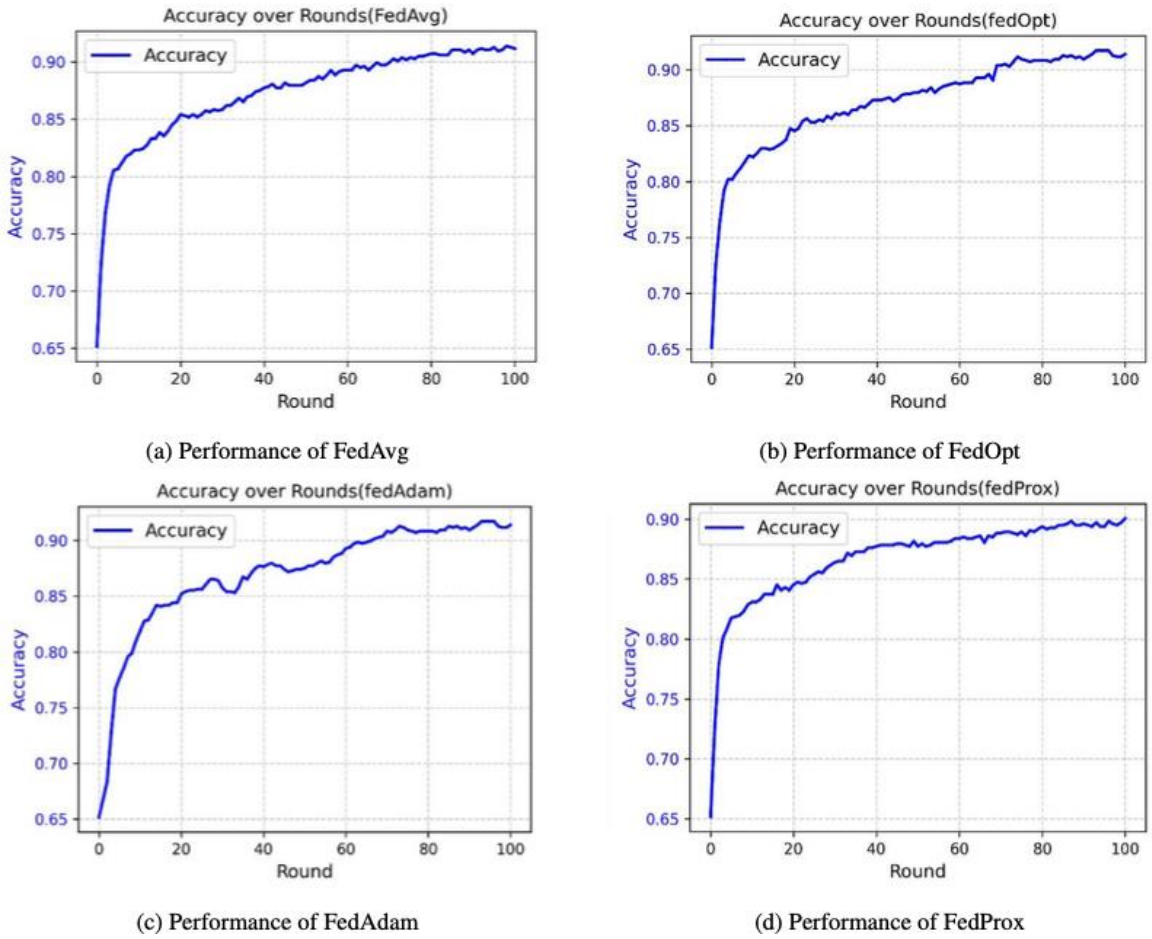

Fig. 6: Federated Accuracy and Loss per Round Using Different Federated Learning

Table 2: Comparison of Centralized and Federated Learning Paradigms for HAR

Feature	**Centralized**	**fedHAR (Federated Learning)**
Data Handling	Raw sensor data is aggregated and uploaded to a central cloud server.	Raw data remains strictly localized on the user s private device.
Peak Accuracy	Achieved a benchmark accuracy of 94 % using the full dataset.	Achieved parity with a peak global accuracy of 92 % using FedAvg.
Privacy Risk	High; exposure to data breaches due to central storage of sensitive info.	Low; only encrypted model updates (weights/gradients) are shared.
Communication	High; requires significant bandwidth to transmit large raw data volumes.	Efficient; minimizes overhead by only transmitting lightweight model updates.
Adaptability	Typically produces a general "one-size-fits-all" global model.	Highly adaptable through personalized local fine-tuning for each user.

From Table 3, we can observe that FedAvg stood out as the most effective method for achieving high global accuracy, serving as a robust baseline for the system. While it reached 92% accuracy, FedAdam offered superior stability. By applying server-side Adam optimization, it achieved the lowest global loss of 0.23, a key indicator of efficient convergence in heterogeneous environments.

The performance of these algorithms also varied in terms of computational overhead. For the lightweight model, FedAvg was the most efficient, requiring 254 seconds, whereas FedAdam required 291 seconds. This trade-off between convergence stability and execution time is a critical consideration for deploying real-time HAR on power-constrained edge devices. FedProx, though designed to mitigate client drift, appeared less well aligned with the specific data characteristics of the vioHAR dataset, resulting in the highest loss.

Table 3: Performance Comparison of Federated Learning Aggregation Techniques

Algorithm	Accuracy	Loss	Convergence and Key Observations
FedAvg	**92.0 %**	0.38	Most balanced method; combines architectural simplicity with the highest global accuracy.
FedAdam	91.3 %	**0.23**	Achieved the lowest training loss, indicating superior stability and faster convergence through adaptive server-side updates.
FedOpt	91.0 %	0.32	Demonstrated effective decentralized training, benefiting from server-side momentum optimization.
FedProx	90.0%	0.47	Exhibited the highest loss; the proximal regularization term was less effective for this specific data distribution.

E. User-Specific Personalization and System Reliability

A primary contribution of the fedHAR framework is the implementation of a personalization strategy to address the inherent challenges of user-specific activity patterns. Traditional centralized models often adopt a "one-size-fits-all" approach that fails to capture individual biomechanical nuances, a limitation clearly identified in Leave-One-Subject-Out (LOSO) evaluations in previous trajectory-based studies. To mitigate this, our framework enables each client to fine-tune a local model and save it as a "personalized model" once a 95% accuracy threshold is met on a local validation set.

To rigorously evaluate the reliability of these client-specific models, we tested them on new, unseen local data across a comprehensive suite of metrics: Accuracy, Precision, Recall, F1 Score, and False Alarm Rate (FAR). As detailed in Table 4, the personalized models achieved an exceptional average accuracy of 0.9702 (97.02%), substantially outperforming the global baseline.

Table 4: Performance of Client-Specific Models on New Data in Personalized Federated Learning

Clients	Accuracy	Precision	Recall	F1 Score	FAR
C01	0.9762	0.9762	0.9787	0.9755	0.0037
C02	0.9706	0.9706	0.9732	0.9706	0.0045
C03	0.9515	0.9470	0.9470	0.9467	0.0080
C04	0.9135	0.9375	0.9375	0.9210	0.0095
C05	0.9710	0.9699	0.9699	0.9697	0.0045
C06	0.9966	0.9969	0.9978	0.9972	0.0004
C07	0.9969	0.9966	0.9978	0.9972	0.0002
C08	0.9618	0.9650	0.9618	0.9618	0.0056
C09	0.9457	0.9560	0.9457	0.9452	0.0082
C10	1.0000	1.0000	1.0000	1.0000	0.0000
C11	0.9850	0.9864	0.9850	0.9849	0.0022
C12	0.9545	0.9600	0.9545	0.9520	0.0074
Average	**0.9707**	**0.9702**	**0.9703**	**0.9684**	**0.0045**

Beyond raw accuracy, the models demonstrated remarkable balance and precision. The average F1 Score of 0.9684 confirms that the personalized models maintain high harmonic mean values between precision and recall, ensuring that minority classes, such as critical fall events, are not overshadowed by frequent activities like walking or sitting. For example, Client C10 achieved perfect classification across all metrics (1.0000), and Clients C06 and C07 reached F1 scores of 0.9972.

Most importantly for applications in elderly care and continuous health monitoring, the system proved highly resistant to false positives. The average False Alarm Rate (FAR) was driven down to just 0.0045 (0.45%). This aligns with the findings from the user-centric uActivity[8] framework, confirming that local personalization not only maximizes recognition accuracy but also provides the operational reliability required for real-world, privacy-preserving edge deployment.

F. Personalization and Pseudo-Labeling

A primary contribution of the fedHAR system is the implementation of a personalization strategy combined with pseudo-labeling to address the inherent challenges of user-specific activity patterns and the scarcity of labeled data at the edge. Traditional centralized models often adopt a "one-size-fits-all" approach that fails to capture individual

biomechanical nuances, a limitation clearly identified in Leave-One-Subject-Out (LOSO) evaluations. In the vioHAR[21] study, LOSO testing revealed significant performance drops with precision falling to 81% for some subjects when the model encountered a completely unseen user. As per our experimental results on generalization capability with LOSO, we have observed that the centralized model is capable of reaching around 74% accuracy, with a balance accuracy of around 85% and a FAR of around 3.8% for both LSTM and RF. To mitigate this, our framework enables each client to fine-tune a local model, achieving an average accuracy of 97.02%, which substantially outperforms the global model's 84.08%. The findings are accumulated in Table 5.

Table 5: Detailed Impact of Personalization Across All Clients

Client ID	Global Model Accuracy (%)	Personalized Model Accuracy (%)	Absolute Improvement	Relative Improvement
C01	76.19	97.62	+21.43	+28.13%
C02	87.50	97.06	+9.56	+10.93%
C03	84.09	94.70	+10.61	+12.62%
C04	90.62	93.75	+3.13	+3.45%
C05	80.45	96.99	+16.54	+20.56%
C06	73.33	99.69	+26.36	+35.95%
C07	70.60	99.69	+29.09	+41.20%
C08	88.50	96.18	+7.68	+8.68%
C09	94.51	94.57	+0.06	+0.06%
C10	94.85	100.00	+5.15	+5.43%
C11	80.45	98.50	+18.05	+22.44%
C12	87.87	95.45	+7.58	+8.63%
Average	**84.08**	**97.02**	**+12.94**	**16.51%**

A detailed quantitative analysis of the personalization strategy reveals consistent performance gains across all participating clients, highlighting the limitations of a generic global model in accommodating unique biomechanical patterns. Every client in the 12-user cohort experienced an increase in classification accuracy after local fine-tuning, with 50% of clients (C01, C03, C05, C06, C07, and C11) demonstrating a relative improvement exceeding 10%. The most substantial benefits were observed among clients whose data distributions most closely diverged from the global average. Specifically, Client C07 exhibited the greatest relative improvement of 41.20%, corresponding to a massive absolute surge of 29.09 percentage points (from 70.60% to 99.69%). Similarly, Client C06 saw a 35.95% relative gain. Conversely, clients whose local data was already well-represented by the global baseline saw marginal but positive refinements. For instance, Client C09 experienced a minimal absolute improvement of just 0.06 percentage points (stabilizing at 94.57%), while Client C10 achieved a modest 5.15 percentage-point increase, reaching flawless 100.00% accuracy. Ultimately, these metrics confirm that the threshold-based personalization framework successfully bridges the gap between generalized federated learning and individualized activity patterns, ensuring robust recognition regardless of the user's initial baseline fit.

The effectiveness of this approach is further supported by findings from the uActivity paper, which demonstrated that user-centric models significantly improve classification accuracy and computational efficiency in trajectory-based recognition tasks. To further enhance robustness in environments with limited manual annotations, pseudo-labeling was employed to utilize unlabeled sensor data. By using the global model to generate high-confidence labels ($\tau \geq 0.95$) for local unlabeled samples, the training set was augmented without requiring user intervention. This technique was especially beneficial for clients with limited labeled samples, increasing the baseline global accuracy from 83% to 89.9%. This semi-supervised mechanism ensures the system remains reliable for critical tasks such as fall detection, maintaining an average False Alarm Rate (FAR) of only 0.45%.

G. System Efficiency and Evaluation Time

The computational efficiency of the fedHAR system is a critical factor for its practical deployment on resource-constrained edge devices, such as smart helmets and wearable sensors. We evaluated the system's efficiency across two primary dimensions: the total time required for the federated training process across 100 communication rounds, and the local inference (evaluation) time required to classify test data for individual clients.

As shown in Table 6, implementing a lightweight model architecture significantly reduced computational burden without a substantial loss in accuracy. The lightweight FedAvg setup proved to be the most time-efficient, completing the entire federated learning process in just 254 seconds, compared to 271 seconds for its complex counterpart. Conversely, FedAdam required the most time (291 seconds) due to the added complexity of calculating adaptive server-side momentum updates.

Table 6: Total Time Required for Federated Learning Process (100 Rounds)

Model Architecture	**Aggregation Algorithm**	**Total Training Time (s)**
Complex	FedOpt	282
Complex	FedAvg	271
Lightweight	FedOpt	263
Lightweight	FedProx	264
Lightweight	**FedAvg**	**254**
Lightweight	FedAdam	291

Beyond training, real-time activity recognition relies heavily on low-latency inference. Table 7 details the local evaluation time for each of the 12 clients. The inference times remained remarkably consistent and well below the half-second threshold, ranging from 0.312 seconds (Client C04) to 0.421 seconds (Client C10). With an average evaluation time of approximately 0.35 seconds per client, the system demonstrates near real-time responsiveness, validating its suitability for continuous daily monitoring and time-critical applications such as fall detection.

Table 7: Inference Evaluation Time for Test Data per Client

Client ID	**Evaluation Time (s)**
C01	0.343
C02	0.356
C03	0.359
C04	**0.312 (Fastest)**
C05	0.350
C06	0.360
C07	0.368
C08	0.354
C09	0.350
C10	**0.421 (Slowest)**
C11	0.374
C12	0.368

5. Discussion

The experimental evaluation of the fedHAR framework provides compelling empirical evidence that high-fidelity, user-specific Human Activity Recognition (HAR) can be achieved without compromising data privacy or edge device efficiency.

A. The Privacy-Performance Parity

A central challenge in modern HAR systems is the trade-off between predictive accuracy and data privacy. Centralized models, while benefiting from comprehensive global data access, pose severe security risks by aggregating sensitive Visual-Inertial Odometry (VIO) trajectory data in cloud repositories. Our results demonstrate that this trade-off can be successfully mitigated. The centralized baseline achieved a 94% accuracy, whereas the decentralized fedHAR system closely trailed with a peak global accuracy of 92% (using FedAvg). This marginal 2% variance is a highly acceptable compromise given the profound privacy benefits. By transmitting only encrypted model weights rather than raw sensor data, fedHAR establishes a framework that complies with stringent data protection regulations (such as the GDPR) while maintaining the predictive power necessary for complex activity classification.

B. The Fallacy of the "One-Size-Fits-All" Model

The results starkly expose the limitations of relying exclusively on global generalized models for HAR. Human biomechanics are inherently heterogeneous; an activity performed by one individual may gen- erate a significantly different trajectory than the same activity performed by another. This was evident in our evaluation, where the global model struggled with outlier data distributions, yielding an accuracy of just 70.6% for Client C07. However, the implementation of threshold-based local personalization resolved this discrep- ancy entirely. By allowing the edge device to fine-tune the model to the specific user's movement patterns, the average network-wide accuracy surged from 88.01% to 97.01%. Aligning with findings from the prior uActivity[8] study, this confirms that future HAR systems deployed in highly diverse populations, particularly for sensitive applications like elderly care, must incorporate local personalization to ensure reliability.

C. Harnessing Unlabeled Data at the Edge

Wearable devices generate continuous, massive streams of sensor data, the vast majority of which remains unlabeled and traditionally unusable for supervised learning. The fedHAR framework effectively unlocked this "dark data" through semi-supervised pseudo-labeling. By utilizing the stabilizing global model to generate high-confidence labels ($\tau \geq 0.95$) for local unannotated samples, the system autonomously augmented its training pool. This mechanism drove a significant performance boost, elevating the baseline global accuracy from 83% to 89.9%. For scalable edge systems where user manual annotation is practically impossible, this semi-supervised approach is critical for continuous model improvement.

D. Algorithmic Trade-Offs in Edge Environments

Evaluating multiple federated aggregation strategies revealed a clear trade-off between computational efficiency and convergence stability. FedAvg emerged as the most resource-efficient algorithm, achieving the highest peak accuracy (92%) and completing the 100-round training process in just 254 seconds on a lightweight architecture. Conversely, FedAdam achieved the lowest global training loss (0.23), indicating superior stability and smoother convergence in noisy data environments, though it required a higher computa- tional overhead (291 seconds). This suggests that system designers must select aggregation protocols based on the specific hardware constraints of the target deployment, prioritizing FedAvg for severely battery-constrained devices, and FedAdam where network noise and data heterogeneity are the primary hurdles.

E. Real-Time Feasibility for Critical Healthcare

Ultimately, a HAR system designed for fall detection and elderly care is only viable if it can process data in real time. The adoption of a lightweight, 3-layer neural network within the fedHAR framework ensured that computational demands remained strictly bounded. With an average local inference time of just 0.35 seconds per client and an exceptionally low False Alarm Rate (FAR) of 0.45%, the system demonstrates near real-time responsiveness. This proves that complex, multi-modal VIO trajectory analysis can be practi- cally deployed on everyday, low-power wearables (such as smart glasses or head-mounted devices) to provide immediate, life-saving interventions.

6. Conclusions

In this study, we set out to address a pressing dilemma in human activity recognition (HAR): how to build highly accurate, real-time monitoring systems without compromising the privacy of users. By developing the fedHAR framework and testing it on the vioHAR dataset, we demonstrated that it is entirely possible to achieve the best of both worlds. Our findings show that decentralized federated learning can achieve performance levels rivaling those of traditional centralized models, while keeping sensitive raw sensor data strictly on the user's local device.

Through extensive testing of various aggregation strategies, we found that FedAvg acts as a highly reliable baseline, delivering a peak global accuracy of 92%. Meanwhile, FedAdam proved exceptional for stable convergence, yielding the lowest training loss of 0.23. To ensure our system could actually run on the types of low-power wearables people use every day, like smart helmets or glasses, we introduced a lightweight model architecture. This approach significantly reduced training time (to just 254 seconds) while maintaining high accuracy, demonstrating its readiness for edge computing.

Perhaps the most exciting takeaway from our research is the profound impact of local personalization. Human movement is inherently unique; a "one-size-fits-all" global model simply cannot capture the subtle differences in how individuals walk, run, or transition between postures. By allowing devices to fine-tune the global model to their specific user, we saw average accuracy leap from 88.01% to an impressive 97.01%. We also successfully tackled the challenge of limited labeled data by implementing pseudo-labeling, which allowed the system to automatically learn from untagged everyday movements and boosted our baseline global accuracy from 83% to 89.9%.

Ultimately, with inference times consistently clocking in under half a second per client, fedHAR proves to be more than just a theoretical concept. It is a highly responsive, privacy-preserving, and personalized system capable of providing life-saving interventions, such as real-time fall detection, in the real world.

While our current synchronous framework performs exceptionally well in a controlled testbed, real- world networks are inherently unpredictable. Wearable devices routinely face communication delays, battery constraints, and intermittent internet connections.

To bridge this gap, our future work will focus on transitioning to an asynchronous federated learning model. This will allow individual clients to upload their localized training updates independently, eliminating the need to wait for slower devices and drastically improving the system's overall scalability. Furthermore, we plan to explore adaptive pseudo-labeling thresholds and advanced data augmentation techniques. By doing so, we aim to make the

model even more resilient to rare or highly imbalanced activities, such as atypical fall events. Finally, testing the fedHAR system in completely unconstrained, dynamic environments with new, continuous user data will be our next major step in ensuring this technology is fully prepared for everyday deployment.

Data Availability Statement

The vioHAR dataset, along with the comprehensive code used for testing and performance evaluation, is publicly available on Figshare [9, 21]. This repository includes structured trajectory data, MATLAB tables, and implementation scripts for the various machine learning models discussed in this study [9, 21].

The dataset can be accessed directly at the following URL: https://figshare.com/s/911dc88878c5b5593e or via its Digital Object Identifier (DOI): https://doi.org/10.6084/m9.figshare.28517546.

References

[1] Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi, "Human activity recognition using inertial, physiological and environmental sensors: A comprehensive survey," IEEE Access, vol. 8, pp. 210 816–210 836, 01 2020. [Online]. Available: https://doi.org/10.1109/access.2020.3037715

[2] D. Bouchabou, C. Lohr, I. Kanellos, and S. M. Nguyen, "Human activity recognition (har) in smart homes," arXiv (Cornell University), 01 2021. [Online]. Available: https://arxiv.org/abs/2112.11232

[3] R. Liu, A. A. Ramli, H. Zhang, E. Henricson, and X. Liu, An Overview of Human Activity Recognition Using Wearable Sensors: Healthcare and Artificial Intelligence. Springer Science+Business Media, 01 2022, pp. 1–14. [Online]. Available: https://doi.org/10.1007/978-3-030-96068-1_1

[4] L. Xue, M. Madhyastha, R. Burns, M. Lee, and M. K. Marina, "Towards decentralized and sustainable foundation model training with the edge," ACM SIGEnergy Energy Informatics Review, vol. 5, pp. 17–25, 07 2025. [Online]. Available: https://doi.org/10.1145/3757892.3757895

[5] K. Raghavan, E. K. Narayan, and F. K. Prasad, "Scalable machine learning on cloud infrastructure: Opportunities and bottlenecks," 05 2025. [Online]. Available: https://philarchive.org/rec/EKTSML

[6] K. M. Shahiduzzaman and M. S. U. Yusuf, "Edgefall: a promising cloud-edge-end architecture for elderly fall care," International Journal of Power Electronics and Drive Systems/International Journal of Electrical and Computer Engineering, vol. 13, pp. 4721–4721, 04 2023. [Online]. Available: https://doi.org/10.11591/ijece.v13i4.pp4721-4733

[7] K. M. Shahiduzzaman, X. Hei, C. A. Guo, and W. Cheng, "Enhancing fall detection for elderly with smart helmet in a cloud-network-edge architecture," pp. 1–2, 05 2019. [Online]. Available: https://doi.org/10.1109/icce-tw46550.2019.8991972

[8] K. M. Shahiduzzaman, M. S. U. Yusuf, and M. S. Hossen, "uactivity: A user-specific human activity recognition and fall detection for elderly care," Engineering, Technology amp; Applied Science Research, vol. 15, no. 6, p. 30169–30174, Dec. 2025. [Online]. Available: https://etasr.com/index.php/ETASR/article/view/12832

[9] K. M. Shahiduzzaman, H. M. S. Yusuf, Md. Salah Uddin, and M. N. A. Munna, "vioHAR Output Dataset," 2 2025. [Online]. Available: https://figshare.com/articles/dataset/vioHAR_Output_Dataset/28515575

[10] C. Papadopoulos, K.-F. Kollias, and G. F. Fragulis, "Recent advancements in federated learning: State of the art, fundamentals, principles, iot applications and future trends," Future Internet, vol. 16, pp. 415–415, 11 2024. [Online]. Available: https://doi.org/10.3390/fi16110415

[11] X. He, X. Su, Y. Chen, and P. Hui, "Federated learning on wearable devices," pp. 613–614, 11 2020. [Online]. Available: https://doi.org/10.1145/3384419.3430446

[12] S. Ek, F. Portet, P. Lalanda, and G. Vega, "Evaluation and comparison of federated learning algorithms for human activity recognition on smartphones," Pervasive and Mobile Computing, vol. 87, pp. 101 714–101 714, 10 2022. [Online]. Available: https://doi.org/10.1016/j.pmcj.2022.101714

[13] R. A. Hamad, W. L. Woo, B. Wei, and L. Yang, Overview of Human Activity Recognition Using Sensor Data. Springer Nature, 01 2024, pp. 380–391. [Online]. Available: https: //doi.org/10.1007/978-3-031-55568-8_32

[14] M. S. Hidri, A. Hidri, S. A. Alsaif, M. Alahmari, and E. Alshehri, "Enhancing sensor-based human physical activity recognition using deep neural networks," Journal of Sensor and Actuator Networks, vol. 14, pp. 42–42, 04 2025. [Online]. Available: https://doi.org/10.3390/jsan14020042

[15] Y. Chen, L. Wang, J. Wang, and X. Qin, "Fedhealth 2: Weighted federated transfer learning via batch normalization for personalized healthcare," arXiv (Cornell University), 01 2021. [Online]. Available: https://arxiv.org/abs/2106.01009

[16] Y. Chen, J. Wang, C. Yu, W. Gao, and X. Qin, "Fedhealth: A federated transfer learning framework for wearable healthcare," arXiv (Cornell University), 01 2019. [Online]. Available: https://arxiv.org/abs/1907.09173

[17] K. Sozinov, V. Vlassov, and S. Girdzijauskas, "Human activity recognition using federated learning," 2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom), pp. 1103–1111, 12 2018. [Online]. Available: http://dx.doi.org/10.1109/bdcloud.2018.00164

[18] Yu, Z. Chen, X. Zhang, X. Chen, F. Zhuang, H. Xiong, and X. Cheng, "Fedhar: Semi-supervised online learning for personalized federated human activity recognition," IEEE Transactions on Mobile Computing, vol. 22, pp. 3318–3332, 12 2021. [Online]. Available: https://doi.org/10.1109/tmc.2021.3136853

[19] R. Presotto, G. Civitarese, and C. Bettini, “Semi-supervised and personalized federated activity recognition based on active learning and label propagation,” Personal and Ubiquitous Computing, vol. 26, no. 5, p. 1281–1298, Jun. 2022. [Online]. Available: http://dx.doi.org/10.1007/s00779-022-01688-8

[20] Yu, Z. Chen, X. Zhang, X. Chen, F. Zhuang, H. Xiong, and X. Cheng, “ FedHAR: Semi-Supervised Online Learning for Personalized Federated Human Activity Recognition ,” IEEE Transactions on Mobile Computing, vol. 22, no. 06, pp. 3318–3332, Jun. 2023. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/TMC .2021 .31368 53

[21] K.M. Shahiduzzaman, Md Salah Uddin Yusuf, Md Sajjad Hossen, and Md Nazmul Alam Munna, "vioHAR: A Visual-Inertial Odometry-based Human Activity Recognition System", Biophysics Reviews 2026. (Accepted)