Federated Learning-Based Sports Motion Recognition Integrating Smart Sports Equipment Condition Monitoring And Personal Privacy Protection Mechanism

Xu Chu*

* *Corresponding author. E-mail: chuxu19860310@163.com*

Received July 21, 2026; accepted August 14, 2026

Wearable and edge-aware systems for sports motion recognition increasingly collect multi-source data (posture, acceleration, angular velocity) plus smart-equipment sensor signals (vibration, impact, temperature, wear). Centralized training exposes exercise habits and usage patterns while imposing high communication burden. This study proposes a scalable, privacy-preserving federated learning framework integrating motion recognition with equipment monitoring under non-IID distributions: edge devices preprocess and segment data via sliding window, train a lightweight CNN-BiGRU-Attention model locally, then prune, DP-perturb, and compress gradients before the server securely aggregates and redistributes the model to selected clients. Under non-IID conditions, FedAvg reaches 89.7% accuracy, FedProx 91.1%, and the proposed method 94.2%. Communication cost falls from 5.60 GB to 2.66 GB after 140 rounds (52.4% below FedAvg), and member-inference/gradient-inversion attack success rates fall to 15%/12%.

Keywords: sports motion recognition, federated learning, smart sports equipment monitoring, wearable sensors, differential privacy, secure aggregation, communication compression

http://dx.doi.org/10.6180/jase.202611_34.064

1. Introduction

Wearable sensors and mobile terminals have moved sports motion recognition from labs to real-world training, identifying actions (walking, running, jumping, squatting, stretching, throwing, kicking, cycling) and supporting quality assessment, exercise prescription, and abnormal-movement warning. Deep learning (CNN, LSTM, GRU, attention) is now a leading approach [1]. Sports equipment is also increasingly instrumented (rackets, insoles, bicycles, resistance machines) to monitor vibration, impact, temperature, and wear, motivating joint athlete-equipment sensing.

Sports motion recognition differs from ordinary image or text recognition in being both privacy-sensitive and naturally distributed: exercise data reveal not just actions but physical function, training habits, intensity, rehabilitation progress, and health status. In school sports, fitness, rehabilitation, and social platforms, centralized raw-sensor collection risks database leakage, unauthorized access, and behavioral profiling. Wearables also differ in sampling frequency, wearing position, battery condition, and sensing quality, increasing system heterogeneity [2], and equipment usage logs collected alongside motion data can further reveal training frequency and usage patterns, compounding privacy risk.

Federated learning mitigates these risks by decentralizing training to local clients, which train on local data and upload only model parameters or gradients for server aggregation, avoiding centralized raw-data storage and suiting mobile sensing environments. However, federated learning does not guarantee complete privacy: attackers can still infer client data through membership inference, gradient inversion, model-update analysis, and backdoor

injection even without raw data. Sports motion recognition federated systems must therefore further incorporate differential privacy, secure aggregation, robust aggregation, and communication compression.

Current research improves human activity recognition generalization via multimodal fusion, transfer learning, and cross-domain methods. Personalized recognition uses multi-source transfer learning to bridge inter-user action differences [3]; and semi-supervised domain generalization addresses insufficient labels and distribution bias [4]. These show the key difficulty is not just classifier accuracy but generalization across users, devices, and scenarios.

In federated activity recognition, attention meta-learning improves performance across diverse users [5]. Contrastive learning with adaptive control variables improves non-IID convergence [6]. However, existing work mostly focuses on a single model or mechanism, without addressing scalability, communication efficiency, attack resistance, and edge deployment together.

This paper extends the focus from a single recognition model to a scalable privacy-preserving federated information system for sports motion recognition. Unlike conventional federated HAR studies emphasizing classification accuracy or a single privacy mechanism, the proposed framework jointly considers Non-IID convergence, client reliability, privacy attack resistance, communication overhead, and edge deployment feasibility, coordinating temporal feature learning, adaptive client selection, differential privacy, secure aggregation, and update compression for large-scale wearable sensing.

1. A four-layer federated information system architecture integrating wearable sports motion sensing and smart equipment condition monitoring is constructed, including a wearable and equipment sensing layer, edge client layer, federated coordination layer, and sports application layer, supporting the local retention of personal motion and equipment-condition data and collaborative training of the global model.

2. A lightweight CNN-BiGRU-Attention temporal recognition model was designed to extract local temporal features, bidirectional dependencies, and key motion segments from multisource motion sensing data, balancing edge deployment efficiency and recognition accuracy.

3. A federated training mechanism integrating gradient pruning, differential privacy, secure aggregation, client adaptive selection, and model update compression is proposed to reduce the risk of personal motion feature leakage and communication burden on large-scale clients.

4. An in-depth system-level simulation experiment system is constructed to verify the effectiveness of the method from aspects such as IID/Non-IID convergence, algorithm performance comparison, client-scale expansion, communication cost, compression ratio, privacy budget, attack success rate, confusion matrix, and t-SNE feature distribution.

Wearable motion recognition typically uses accelerometer, gyroscope, magnetometer, or multi-source IMU data segmented via sliding window and classified using machine learning or deep learning. Unlike video-based recognition, sensor-based recognition is less dependent on lighting, occlusion, or camera angle and doesn't collect images, suiting privacy-sensitive scenarios like training and rehabilitation. Traditional methods use handcrafted features (mean, variance, energy, peak, zero-crossing, frequency-domain) with SVM, random forest, or KNN; CNN, LSTM, GRU, Transformer, and attention models are now mainstream.

Sensor signals in sports motion recognition show clear temporal dynamics: walking/running have periodic rhythms, jumping has short high-amplitude impacts, squatting/stretching show posture changes, and throwing/kicking show local speed/direction changes — so a single static feature cannot stably represent the full motion. CNN extracts local motion patterns, BiGRU models temporal dependence, and attention weights key motion segments. Interpretability studies also show different sensor channels and hidden units contribute unevenly to activity recognition [2], motivating this study's use of attention and feature visualization.

Real-world deployment shows cross-user differences in amplitude, speed, rhythm, posture, and device placement, creating distribution shift that centralized training's same-distribution assumption ignores. Multi-source transfer learning [3] and domain generalization/semi-supervised methods [4] address this. In federated learning, it manifests as non-IID client data, where directly applying FedAvg causes client drift and unstable convergence [6]. This study constructs a Dirichlet Non-IID distribution and compares FedAvg, FedProx, FedNova, SCAFFOLD, and the proposed method.

Federated learning is a distributed machine learning paradigm enabling collaborative training across clients without sharing raw data. Federated transfer learning surveys note it promotes knowledge sharing while protecting data, though facing feature-space, distribution, system het-

erogeneity, and incremental-data challenges [7]. This study models sports motion recognition as distributed training in a large-scale mobile sensor information system, evaluating performance from 10 to 200 clients.

Therefore, this study regards sports motion recognition not only as a classification task but also as a scalable information system problem in which client number, communication cost, client availability, and edge inference latency must be evaluated together.

Federated learning still presents privacy risks despite avoiding raw-data upload: reviews note vulnerabilities to member inference, attribute inference, gradient leakage, poisoning, backdoor, and malicious-client attacks [8]. For sports motion recognition, recovering local motion patterns from gradients could reveal training participation, abnormal exercise habits, or partial motion trajectories.

Differential privacy limits a single sample's impact on model output by adding noise to gradients or updates. Adaptive DP federated frameworks reduce leakage risk via gradient pruning while maintaining accuracy [9], and other studies address DP mechanism design and privacy-budget control [10]. However, DP alone can degrade accuracy, especially with small or imbalanced wearable datasets, so this study combines DP with secure aggregation to reduce single-client update exposure.

This literature motivates the joint use of gradient clipping, Gaussian perturbation, and secure aggregation in the proposed framework, because differential privacy alone may reduce model utility when wearable sensor data are small, imbalanced, or highly heterogeneous.

Federated learning also faces client anomalies, malicious updates, and communication bottlenecks: model-update channels can become attack entry points allowing manipulation via malicious gradients [11]. Robust aggregation reduces anomalous updates' impact [12]; secure aggregation prevents the server seeing individual updates via masking, secret sharing, and encryption, with verifiable secure aggregation adding robustness for client disconnection and result verification [13].

Communication compression is important at scale: even lightweight recognition models require transmitting parameter updates over many communication rounds. This study uses Top-k sparsity and quantization compression to reduce uploaded parameters while maintaining accuracy through appropriate compression.

Accordingly, the proposed method combines Top-k sparsification with quantization compression so that the uploaded update size can be reduced while preserving the most informative gradient components for global aggregation.

2. Methods

2.1. Overall System Architecture

The proposed system has four layers (Figure 1): a wearable sensing layer (smartwatches, phones, IMUs, armbands, plus smart-equipment sensors reporting vibration, impact, temperature, wear); an edge client layer (local storage, preprocessing, training, gradient pruning, DP perturbation, compression); a federated coordination layer (client selection, secure aggregation, global updates); and a sports application layer (motion recognition, feedback, abnormal-motion warnings, equipment maintenance alerts).

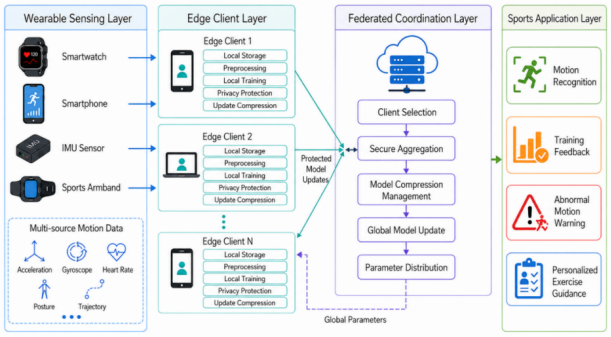

Fig. 1. Privacy-Preserving Federated Learning System Architecture for Sports Motion Recognition and Smart Equipment Condition Monitoring

Figure 1 shows the core data flow: wearable devices collect multi-source motion signals; each edge client performs local storage, preprocessing, and CNN-BiGRU-Attention training; updates are clipped, perturbed, sparsified, and quantized before upload; the server receives only protected aggregated updates via secure aggregation; and the updated global model is redistributed for real-time recognition and personalized feedback. The server therefore never accesses raw motion data or individual plaintext updates.

2.2. Data Representation

Suppose there are N clients in the system. The local dataset of the i client is represented as:

In this notation, N denotes the total number of participating clients, i denotes the client index, D_i denotes the local dataset of client i, X_i^j denotes the j segmented motion sensor sample, y_i^j denotes its action label, and n_i denotes the number of local samples. These symbols are used consistently in the subsequent optimization formulas.

$$D_i = \{\left(X_i^j, y_i^j\right)\}_{j=1}^{n_i}, \quad i = 1, 2, \ldots, N. \tag{1}$$

Where X_i^j represents the j motion sensor sample from the i client, y_i^j represents the corresponding action category

label, and n_i represents the number of local samples from client i. A single motion sample is obtained by dividing the time window by a sliding time window:

$$X_i^j \in \mathbb{R}^{T\times C}, \tag{2}$$

This paper mainly considers triaxial acceleration and triaxial angular velocity (C=6 channels); if attitude angle or magnetometer signals are available, C can extend to higher dimension. When smart equipment condition sensors are available, additional channels — vibration amplitude, impact load, surface temperature, wear indicators — can be appended to the same client-side representation, jointly modeling athlete motion and equipment status without altering the federated optimization formulation.

$$\begin{aligned}\mathcal{Y} = \{&\text{walking}, \text{running}, \text{jumping}, \text{squatting},\\ &\text{stretching}, \text{throwing}, \text{kicking}, \text{cycling}\}\end{aligned} \tag{3}$$

2.3. Federated Optimization Objective

The objective of federated learning is to learn the global model parameter vector θ without centrally uploading the original motion sensor data. The global optimization objective is defined as:

$$\min_{\theta} F(\theta) = \sum_{i=1}^{N} \frac{n_i}{n} F_i(\theta), \quad n = \sum_{i=1}^{N} n_i. \tag{4}$$

The local objective function for client i is:

$$F_i(\theta) = \frac{1}{n_i} \sum_{j=1}^{n_i} \updownarrow \left(f_\theta\left(X_i^j\right), y_i^j\right). \tag{5}$$

In the formula, f_θ denotes the sports motion recognition model parameterized by θ, and $\updownarrow(\cdot)$ denotes the cross-entropy loss function. For the eight-class recognition task, the cross-entropy loss is written as:

$$\ell(\hat{y}, y) = -\sum_{c=1}^{8} \mathbb{I}(y = c) \log \hat{p}_c. \tag{6}$$

Here, $p(j, c)$ denotes the predicted probability that sample j belongs to action class c, $y(j, c)$ is the one-hot label indicator, and $1(\cdot)$ denotes the indicator function. This clarification ensures that the loss definition is reproducible and consistent with the eight-action classification setting.

2.4. Privacy Threat Model

This study considers honest-but-curious servers and external attackers: the server aggregates per protocol but may try to infer user information from updates, while external attackers may intercept updates or perform membership inference on the trained model. Privacy objectives: raw data never leaves the device; the server cannot observe any single client's real update; uploaded gradients are difficult to invert after pruning and noise perturbation; the privacy budget is adjustable to scenario needs; and the privacy mechanism should not excessively compromise training or scalability.

3. Method design

3.1. End-to-End Federated Training Process

Figure 2 shows the end-to-end training process: the edge device collects, normalizes, denoises, and segments sensor data via sliding window; the client trains a local CNN-BiGRU-Attention model; local gradients are clipped, perturbed by Gaussian DP noise, and compressed via Top-k sparsification and quantization; and the server performs secure aggregation, updates the global model, and redistributes parameters to selected clients — with raw motion data remaining on local devices throughout.

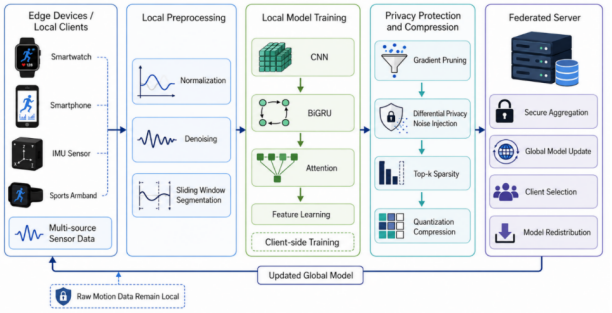

Fig. 2. Privacy-Preserving Federated Training End-to-End Flow

3.2. CNN-BiGRU-Attention Local Recognition Model

Sports motion signals contain both local impact features and temporal dependencies of the complete motion cycle. The study designed a lightweight CNN-BiGRU-Attention model as the client-side local recognition network. First, one-dimensional convolution is used to extract local motion patterns:

For reproducibility, the implemented local network contains one-dimensional convolutional blocks for local temporal pattern extraction, a two-layer BiGRU for bidirectional dependency modeling, an attention layer for key segment weighting, and a fully connected classifier for eight action categories. The hyperparameters used in the experiment are reported in Section 5.3.

$$H_c = \sigma\left(W_c * X + b_c\right), \tag{7}$$

Where W_c is the convolution kernel, b_c is the bias term, * denotes one-dimensional convolution, and $\sigma(\cdot)$ is the ReLU activation function. Subsequently, the convolutional features are input into a bidirectional GRU:

$$\overrightarrow{h_t} = GRU\left(x_t, \overrightarrow{h_{t-1}}\right), \quad \overleftarrow{h_t} = GRU\left(x_t, \overleftarrow{h_{t+1}}\right), \quad (8)$$

$$h_t = \left[\overrightarrow{h_t}; \overleftarrow{h_t}\right]. \quad (9)$$

To highlight the key action segments, attention weights are introduced.

$$e_t = v^T \tanh\left(W_a h_t + b_a\right), \quad \alpha_t = \frac{\exp\left(e_t\right)}{\sum_{k=1}^{T} \exp\left(e_k\right)}. \quad (10)$$

The final timing representation is as follows

$$z = \sum_{t=1}^{T} \alpha_t h_t. \quad (11)$$

The classification output was as follows:

$$\widehat{y} = Softmax\left(W_o z + b_o\right). \quad (12)$$

CNN extracts short-term impact and local oscillation features; BiGRU captures before/after stage dependencies; and attention enhances key action segments such as jumping, squatting, and throwing power phases — better suited to non-uniform temporal variation in sports movements than simple CNNs or unidirectional RNNs.

3.3. Federated Aggregation Mechanism

At the start of the t-th communication round, the server distributes the current global model θ_t to the selected client set S_t. Each selected client performs E local epochs and obtains a local update Delta θ_iˆt, which is then processed by privacy protection and compression modules before aggregation.

$$\theta_i^{t+1} = \theta^t - \eta \nabla F_i\left(\theta^t\right), \quad i \in S_t. \quad (13)$$

The server performs weighted aggregation based on the number of client samples.

$$\theta^{t+1} = \sum_{i \in S_t} \frac{n_i}{\sum_{k \in S_t} n_k} \theta_i^{t+1}. \quad (14)$$

Because sports motion data exhibit non-IID characteristics, this study does not directly use full client aggregation but instead performs adaptive client selection before each round of communication.

3.4. Adaptive Client Selection

Client selection considers data quality, online reliability, and communication cost simultaneously. The comprehensive score of the i client is defined as:

$$Score_i = \alpha Q_i + \beta R_i - \gamma C_i, \quad (15)$$

Where Q_i represents local data quality, R_i represents client online reliability, C_i represents communication cost, and α, β and γ are weighting parameters. The client set is determined by the following formula:

$$S_t = TopK\left(\{Score_i\}_{i=1}^{N}\right). \quad (16)$$

This strategy reduces the impact of low-quality, high-communication-cost clients while maintaining sample representativeness — important since large-scale platform clients may be online or offline at any time, making adaptive selection more practical than fixed full participation.

3.5. Differential Privacy and Gradient Pruning

To limit the impact of individual samples on model updates, the client first prunes the gradients before uploading.

$$\bar{g}_i = \frac{g_i}{\max\left(1, \frac{\|g_i\|_2}{C}\right)}. \quad (17)$$

Where C is the clipping threshold. After clipping, Gaussian noise with variance σ^2. $C^2 I$ is added to limit the contribution of any single sample to the uploaded update.

$$\widetilde{g}_i = \bar{g}_i + \mathcal{N}\left(0, \sigma^2 C^2 I\right). \quad (18)$$

The differential privacy mechanism can be represented as follows:

$$\Pr\left[\mathcal{M}(D) \in \mathcal{O}\right] \leq e^{\epsilon} \Pr\left[\mathcal{M}\left(D'\right) \in \mathcal{O}\right] + \delta, \quad (19)$$

In the differential privacy definition, D and D' are neighboring datasets that differ in one sample, ε is the privacy budget, and δ is the failure probability. A smaller ε provides stronger privacy protection but usually introduces larger noise and may reduce recognition accuracy.

3.6. Secure Aggregation Mechanism

Differential privacy reduces the risk of individual sample contribution exposure, but the server may still observe single client updates. Therefore, this paper further introduces secure aggregation. The mask update uploaded by client i is:

$$u_i' = u_i + r_i, \quad (20)$$

Where u_i is the actual update and r_i is the random mask. If multiple client masks satisfy the cancellation condition, the server can only obtain the aggregated result:

$$\sum_{i \in S_t} u_i^{'} = \sum_{i \in S_t} u_i + \sum_{i \in S_t} r_i = \sum_{i \in S_t} u_i. \tag{21}$$

Therefore, the server can complete global model updates but cannot recover any real updates from individual clients.

3.7. Communication Compression Mechanism

In large-scale federated learning, communication overhead is often much higher than local computation overhead. This paper employs Top-k sparsity and quantization compression to reduce uploaded parameters. Given a gradient vector g_i, only the k components with the largest absolute values are retained:

$$TopK(g_i) = \left\{ g_{i,m} \mid |g_{i,m}| \in top\text{-}k(|g_i|) \right\}. \tag{22}$$

A quantitative update is defined as

$$Q(g_i) = sign(g_i) \cdot \Delta \cdot round\left(\frac{|g_i|}{\Delta} \right). \tag{23}$$

The compressed cost of a single round of communication is

$$C_{comm}^{t} = \sum_{i \in S_t} \left(B_i^{index} + B_i^{value} \right), \tag{24}$$

Where B_i^{index} represents the sparse index transmission amount, and B_i^{value} represents the quantization parameter transmission amount. The total communication cost is:

$$C_{comm} = \sum_{t=1}^{T_c} C_{comm}^{t}. \tag{25}$$

3.8. Algorithm Description

Table 1 lists the main training steps of the proposed method.

4. Experimental design

4.1. Dataset and Simulation Setup

This experiment focuses on wearable sensor-based motion recognition in a federated simulation environment, with each client an independent user device. Samples include triaxial acceleration and angular velocity; action categories are walking, running, jumping, squatting, stretching, throwing, kicking, and cycling. Two partitions are used: IID for balanced comparison and Dirichlet-based Non-IID for cross-user heterogeneity.

$$p_i \sim Dir(\alpha), \tag{26}$$

Where represents the category proportion distribution of client, and controls the heterogeneity intensity. This paper mainly sets =0.3, representing a strong Non-IID distribution [14].

4.2. Comparison Methods

To verify effectiveness, experiments compare Local Training, Centralized CNN-BiGRU, FedAvg, FedProx, FedNova, SCAFFOLD, FedAvg-DP, and the full proposed method; Centralized CNN-BiGRU serves only as an upper-bound reference since it requires centralized data and doesn't meet privacy requirements.

4.3. Experimental Parameters

Client counts are set to 10, 20, 50, 100, 150, and 200, with 30% selected per round over 150 rounds, 5 local epochs, batch size 64, Adam optimizer, learning rate 0.001. The sliding window is 128 with six sensor channels; the CNN-BiGRU-Attention model has 1D convolutional feature extraction, two-layer BiGRU with attention weighting, and a fully-connected 8-class classifier. Privacy budget ϵ is set to 0.5, 1, 2, 4, 8, and No-DP; clipping threshold is 1.0; compression ratios are 0%, 25%, 50%, 75%, and 90%.

4.4. Evaluation Metrics

Evaluation metrics include accuracy, precision, recall, F1, communication overhead, attack success rate, and edge deployment latency: accuracy is correct-classification proportion; precision/recall/F1 assess class-imbalance quality; attack success rate measures successful membership-inference or gradient-inversion attacks; and edge latency is average per-sample preprocessing-to-classification time.

5. Results and discussion

5.1. Overall Recognition Performance Comparison

Table 2 compares methods on eight motion recognition tasks. Local Training generalizes poorly across clients; Centralized CNN-BiGRU is a non-private upper-bound reference; FedAvg learns useful shared representations without raw-data upload; FedProx, FedNova, and SCAFFOLD improve further by mitigating client drift; FedAvg-DP suffers accuracy loss from noise. The proposed method achieves the best balance of recognition performance and communication efficiency, reflecting synergy among attention-based temporal learning, adaptive client selection, privacy protection, and compression.

Table 1. Privacy-Preserving Federated Training Algorithm

Step	Operate	Illustrate
1	Initialize global model θ^0	Server initialization model parameters
2	Client status assessment	Calculate Q_i, $R_i$and C_i
3	Adaptive client selection S_t	Select the client according to Eqs. (15) and (16)
4	Local model training	Local updates using CNN-BiGRU-Attention
5	Gradient clipping	The gradient norm is constrained according to Eq. (17).
6	Differential privacy perturbation	Add Gaussian noise according to formula (18)
7	Update compression	Perform Top-k sparsity and quantization
8	Secure aggregation	The server only receives aggregated updates.
9	Global model update	Update global parameters according to formula (14)
10	Model distribution	Redistribute the global model to clients

Table 2. Comparison of Recognition Performance of Different Methods

Method	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F_1-score (%)**	**Relative communication overhead (%)**
Local Training	84.2	83.8	83.4	83.6	0.0
Centralized CNN-BiGRU	93.1	92.9	92.5	92.7	Centralized
FedAvg	90.4	90.0	89.6	89.8	100.0
FedProx	91.3	90.9	90.5	90.7	100.0
FedNova	91.7	91.3	90.9	91.1	100.0
SCAFFOLD	92.0	91.7	91.3	91.5	100.0
FedAvg-DP	88.7	88.3	87.7	88.0	100.0
Proposed	94.4	94.0	93.6	93.8	47.6

5.2. Convergence Analysis under IID and Non-IID Conditions

Figure 3 shows convergence under IID/Non-IID conditions: FedAvg converges smoothly under IID (92.2% after 150 rounds) but degrades under Non-IID (89.7%); FedProx's proximal constraints improve Non-IID accuracy to 91.1%; the proposed method reaches 94.2% under Non-IID with faster convergence in the first 60 rounds.

Three factors explain this: adaptive client selection reduces low-quality/unreliable/high-cost clients' negative effect; CNN-BiGRU-Attention captures shared local patterns, bidirectional dependencies, and key segments across users; and secure aggregation, DP, and moderate compression are jointly tuned to avoid training instability — consistent with prior findings that client skew slows Non-IID federated activity-recognition convergence.

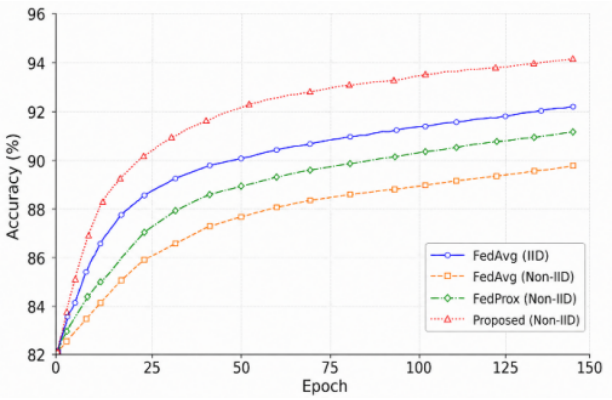

Fig. 3. Convergence curves of federated training under IID and Non-IID conditions

5.3. Scalability Analysis under Different Number of Clients

Figure 4 shows accuracy and per-round training time as clients grow from 10 to 200: FedAvg drops from 92.1% to 88.2%, FedProx from 92.5% to 89.4%, while the proposed method drops only from 94.8% to 93.8% — demonstrating stronger scalability, critical for systems with hundreds or thousands of users.

Table 3 shows that at 200 clients, FedAvg's single-round training time is 67s versus 39s for the proposed method, mainly from adaptive client selection and update compression reducing inefficient-client participation and communi-

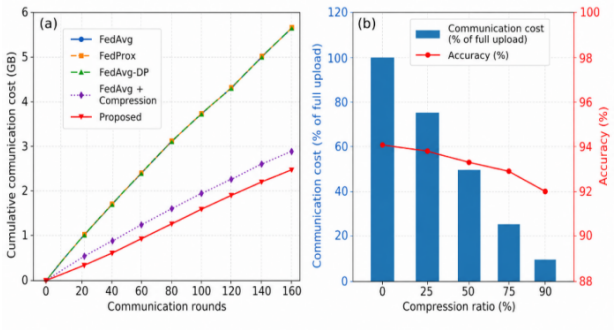

Fig. 4. Recognition Accuracy and Training Time under Different Client Scales

cation wait — directly connecting the method to scalable information-system requirements.

5.4. Communication Efficiency Analysis

Figure 5(a) shows cumulative communication cost: FedAvg, FedProx, and FedAvg-DP upload full updates each round, reaching ~5.60 GB after 140 rounds; FedAvg+Compression reduces this to 2.94 GB, and the proposed method further reduces it to 2.66 GB — a 52.4% reduction versus FedAvg, reflecting relative federated model-update transmission burden.

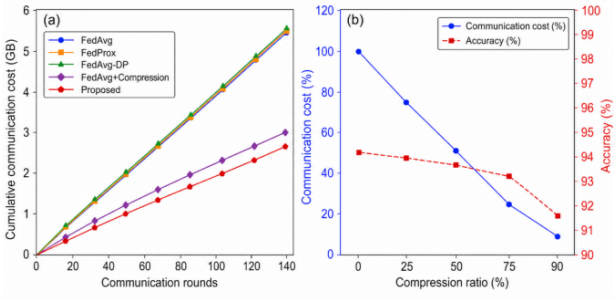

Fig. 5. Analysis of Communication Efficiency in Federated Training

The compression-ratio experiment (Figure 5b) shows cost falling from 100% to 25% as compression rises from 0% to 75%, with accuracy only dropping from 94.4% to 93.2% — a controllable loss. At 90% compression, accuracy drops to 91.8%, showing excessive compression destroys important gradient information, so moderate compression best balances accuracy and communication efficiency.

5.5. Analysis of Privacy Budget and Attack Success Rate

Figure 6(a) shows accuracy under different privacy budgets: at ϵ=0.5, protection is strongest but accuracy falls to 89.2%; at ϵ=4, accuracy reaches 93.4% (vs. 94.4% for No-DP), so ϵ=4 is chosen as a practical balance. Figure 6(b) shows differential privacy and secure aggregation provide complementary protection against membership inference and gradient inversion attacks.

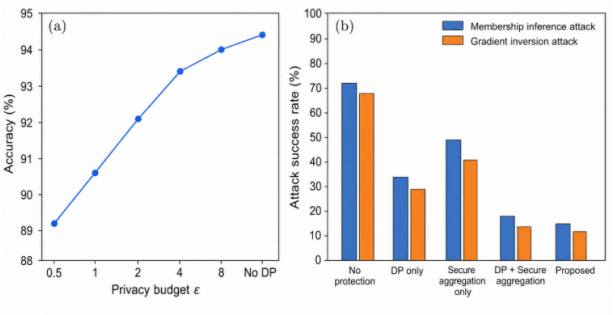

Fig. 6. Privacy Budget and Attack Success Rate Analysis

5.6. Confusion Matrix and t-SNE Feature Distribution Analysis

Figure 7(a) shows the confusion matrix: cycling has the highest recognition rate (96%), walking/throwing reach 95%, squatting/kicking 94%, running/stretching 93%, and jumping 92%, with main confusion between running/jumping and squatting/stretching due to similar acceleration peaks or slow posture changes. Figures 7(b-c) show the proposed method produces more compact feature clusters than FedAvg, though t-SNE serves as auxiliary rather than causal evidence.

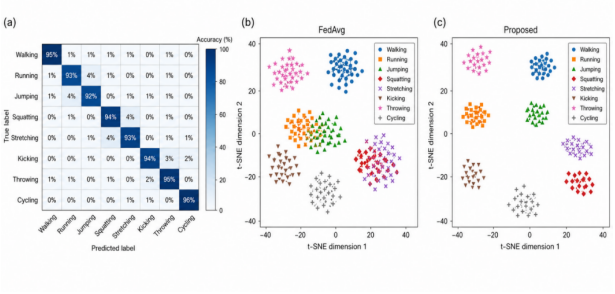

Fig. 7. Comparison of Confusion Matrix and t-SNE Feature Distribution of FedAvg/Proposed

5.7. Ablation Experiments

Ablation experiments (Table 4) show removing differential privacy slightly improves accuracy but sharply raises membership-inference risk; removing secure aggregation increases individual update exposure; removing compression returns overhead to 100%; and removing client selection reduces accuracy since low-quality or high-cost clients participate more often, increasing client drift and reducing global stability.

Table 3. Scalability Results under Different Client Scales

Number of clients	FedAvg accuracy (%)	FedProx accuracy (%)	Proposed accuracy (%)	FedAvg single round time (s)	Proposed single-round time (s)
10	92.1	92.5	94.8	6	4
20	91.6	92.0	94.6	9	6
50	90.9	91.4	94.4	18	11
100	89.8	90.7	94.2	34	20
150	89.0	90.0	94.0	49	29
200	88.2	89.4	93.8	67	39

Table 4. Ablation Experiment Results

Variants	Accuracy (%)	F_1-score (%)	Communication overhead (%)	Membership inference attack success rate (%)
Remove DP	95.0	94.5	48.2	36.8
Remove security aggregation	94.7	94.2	47.9	31.4
Remove compression	94.6	94.1	100.0	18.7
Remove client selection	92.8	92.3	63.5	19.4
Complete method	94.4	93.8	47.6	15.0

5.8. Edge Deployment Performance

Table 5 shows edge inference performance: model size is ~1.8 MB, with latency of 42 ms on a smartwatch, 18 ms on a smartphone, 25 ms on a Raspberry Pi, and 9 ms on an edge gateway — suitable for real-time inference, though real deployment must also consider battery, network instability, temperature, missing values, and sensor drift.

5.9. Discussion

The method's core value lies not only in recognition accuracy but in a privacy-preserving federated information system for large-scale wearable platforms. Versus centralized training, it avoids uploading raw motion data; versus standard FedAvg, it improves Non-IID convergence stability; versus DP alone, it combines secure aggregation and compression to balance privacy, utility, and scalability. The architecture is extensible to equipment condition monitoring, since equipment-side signals can be added at the sensing layer and client-side representation without altering the federated optimization or privacy design.

Applications include school physical education, smart fitness, and rehabilitation platforms: raw movement trajectories stay on local devices for end users; platform providers improve the global model without centralized collection; and institutions gain a privacy-aware technical route for sensitive motion data. Ethical deployment should include informed consent, local data control, clear privacy-budget configuration, and restrictions on unauthorized secondary use.

Limitations: experiments are simulation-based, while real systems face network interruption, battery limits, sensor drift, and compliance issues; the privacy budget is fixed rather than adaptive; only acceleration/angular-velocity data are used, leaving plantar pressure, EMG, heart rate, and visual features for future work; secure aggregation reduces exposure but malicious poisoning/backdoor attacks still need verifiable and robust aggregation [15]; and the equipment-monitoring extension is validated only conceptually.

6. Conclusions

This study addresses the challenges of privacy protection and large-scale distributed training in sports motion recognition by proposing a scalable, privacy-preserving federated information system that is architecturally extensible to smart sports equipment condition monitoring. The main conclusions are as follows:

1. The system employs a four-layer architecture—a wearable and equipment sensing layer, an edge client layer, a federated coordination layer, and a sports application layer—supporting multi-client collaborative

Table 5. Edge Deployment Performance

Equipment	Model size	Inference latency/ms	Memory usage / MB	Is it suitable for deployment in the real world?
Smartwatch	1.8 MB	42	38	Yes
Smartphones	1.8 MB	18	51	Yes
Raspberry Pi	1.8 MB	25	46	Yes
Edge gateway	1.8 MB	9	64	Yes

training without uploading raw motion or equipment-condition data.

2. A CNN-BiGRU-Attention local recognition model, combined with adaptive client selection, gradient pruning, differential privacy, secure aggregation, and model update compression, achieves 94.4% accuracy and a 93.8% F_1 score on eight sports motion recognition tasks, maintaining 93.8% accuracy with 200 clients.

3. Communication overhead is reduced by approximately 52.4% compared with FedAvg, and the success rates of membership inference and gradient inversion attacks are reduced to 15% and 12%, respectively.

4. By extending the wearable sensing layer and the client-side data representation to include equipment-condition channels, the proposed framework provides a technical route toward jointly monitoring athlete motion and smart equipment status without centralizing raw data; validating this extension with real equipment-condition data is left to future work.

5. Future research will further extend the framework to real-world multi-device federated deployments, considering asynchronous federated learning, dynamic client disconnection, personalized model adaptation, adaptive privacy budgets, and robust defense against malicious clients, in order to build a more trustworthy, efficient, and scalable sports motion recognition and equipment monitoring information system.

References

[1] S. Zhang, Y. Li, S. Zhang, F. Shahabi, S. Xia, Y. Deng, and N. Alshurafa, (2022) *"Deep learning in human activity recognition with wearable sensors: A review on advances"* **Sensors 22**(4): 1476. DOI: 10.3390/s22041476.

[2] D. Navakauskas and M. Dumpis, (2025) *"Wearable sensor-based human activity recognition: Performance and interpretability of dynamic neural networks"* **Sensors 25**(14): 4420. DOI: 10.3390/s25144420.

[3] Q. Jia, J. Guo, P. Yang, and Y. Yang, (2024) *"A holistic multi-source transfer learning approach using wearable sensors for personalized daily activity recognition"* **Complex & Intelligent Systems *10***: 1459–1471. DOI: 10.1007/s40747-023-01218-w.

[4] J. Liu, W. Zhu, D. Li, X. Hu, and L. Song, (2025) *"Domain generalization with semi-supervised learning for people-centric activity recognition"* **Science China Information Sciences *68***: 112103. DOI: 10.1007/s11432-022-3860-y.

[5] Q. Shen, H. Feng, R. Song, D. Song, and H. Xu, (2023) *"Federated meta-learning with attention for diversity-aware human activity recognition"* **Sensors *23***(3): 1083. DOI: 10.3390/s23031083.

[6] I. Iwan, B. N. Yahya, and S.-L. Lee, (2025) *"Federated model with contrastive learning and adaptive control variates for human activity recognition"* **Frontiers of Information Technology & Electronic Engineering *26***(6): 896–911. DOI: 10.1631/FITEE.2400797.

[7] W. Guo, F. Zhuang, X. Zhang, Y. Tong, and J. Dong, (2024) *"A comprehensive survey of federated transfer learning: Challenges, methods and applications"* **Frontiers of Computer Science *18***: 186356. DOI: 10.1007/s11704-024-40065-x.

[8] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, (2021) *"A survey on security and privacy of federated learning"* **Future Generation Computer Systems *115***: 619–640. DOI: 10.1016/j.future.2020.10.007.

[9] Z. Chen, H. Zheng, and G. Liu, (2024) *"AWDP-FL: An adaptive differential privacy federated learning framework"* **Electronics *13***(19): 3959. DOI: 10.3390/electronics13193959.

[10] H. K. Tayyeh and A. S. A. Al-Jumaili, (2024) *"Balancing privacy and performance: A differential privacy approach in federated learning"* **Computers *13***(11): 277. DOI: 10.3390/computers13110277.

[11] N. Bouacida and P. Mohapatra, (2021) *"Vulnerabilities in federated learning"* **IEEE Access *9***: 63229–63247. DOI: 10.1109/ACCESS.2021.3075203.

[12] K. Pillutla, S. M. Kakade, and Z. Harchaoui, (2022) *"Robust aggregation for federated learning"* **IEEE Transactions on Signal Processing *70***: 1142–1154. DOI: 10.1109/TSP.2022.3153135.

[13] S. Zhou, L. Wang, L. Chen, Y. Wang, and K. Yuan, (2025) *"Group verifiable secure aggregate federated learning based on secret sharing"* **Scientific Reports *15***: 9712. DOI: 10.1038/s41598-025-94478-0.

[14] T.-M. H. Hsu, H. Qi, and M. Brown, (2019) *"Measuring the effects of non-identical data distribution for federated visual classification"* **arXiv**: DOI: 10.48550/arXiv.1909.06335.

[15] C. Yang and Y. Ma, (2026) *"Secure aggregation for heterogeneous enterprise data based on federated meta-learning"* **Scientific Reports *16***: 4484. DOI: 10.1038/s41598-025-34735-4.