Embedding-Only Federated Edge Learning for Privacy-Preserving eHealth

Bharathwaj Vijayakumar, Samyukta Alapati, Sahana Varadaraju
Rowan University, Glassboro, NJ, USA
Emails: vijayakumar@rowan.edu, alapati@rowan.edu, varadaraju@rowan.edu

***Abstract*—The next wave of digital health innovation lies in combining federated learning and edge intelligence for secure, low-latency physiological monitoring. This paper introduces a *federated self-supervised edge* framework for wearable devices that track physiological transitions such as heart-rate variability, skin temperature, and galvanic skin response. The system enables decentralized learning directly on devices without transmitting sensitive data, addressing key challenges in privacy and communication efficiency. Each client employs a lightweight one-dimensional convolutional encoder trained via contrastive self-supervised learning to capture temporal patterns in local signals. Instead of sharing raw data or model weights, each device uploads a compressed embedding centroid to a central aggregator that computes a global reference embedding for anomaly scoring. A prototype simulation across five edge clients achieved Area Under the Receiver Operating Characteristic (AUROC) = 0.86 while reducing communication overhead by 82%. Visualization via t-SNE revealed separable manifolds of normal and anomalous physiology, demonstrating a novel, communication-efficient paradigm for privacy-preserving, federated IoT-based healthcare.**

***Index Terms*—Federated Learning, Self-Supervised Learning, Edge AI, eHealth, Wearable Devices, IoT Communication, Privacy-Preserving AI, Women's Health, Anomaly Detection, Bandwidth Efficiency.**

I. Introduction

The fusion of artificial intelligence (AI) with communication technologies has accelerated the development of distributed, low-power health monitoring systems. Edge computing and Internet of Things (IoT) sensors now provide continuous data from wearables, enabling personalized health tracking while reducing dependence on centralized servers [2], [3], [5]. However, transmitting large volumes of raw physiological data to the cloud raises privacy, security, and bandwidth challenges, core issues at the intersection of *AI and Communications Technology (ICT)*, which remain central to ongoing research in distributed intelligent systems.

This research targets an underexplored but socially significant application: **monitoring menopausal health using privacy-preserving edge intelligence**. Mid-life women experience irregular physiological changes such as hot flashes and sleep disruptions. Traditional analytics depend on self-reports or centralized machine learning models requiring sensitive data aggregation. To bridge this gap, we propose a system that performs intelligent feature learning directly on wearable devices.

The novelty of this work lies in combining three key aspects:

- **Self-supervised edge learning:** local models learn invariant representations from unlabeled, noisy physiological data using SimCLR-style contrastive objectives.
- **Embedding-only federated aggregation:** instead of exchanging raw data or full model weights, clients share lightweight centroid embeddings, ensuring privacy and reducing communication by ≈82%.
- **Edge-aware anomaly detection:** a federated server produces a global reference embedding, broadcast to all devices for low-latency, distance-based anomaly scoring.

Communication bandwidth is one of the most significant bottlenecks in IoT-driven healthcare [5], [6]. Each wearable device may generate tens of thousands of samples per minute across multimodal sensors, leading to gigabytes of data daily. Traditional cloud analytics cannot sustainably process this data without consuming high energy and network resources. By transmitting only low-dimensional embedding centroids, our approach reduces communication by more than 80% while maintaining high anomaly detection accuracy. This trade-off between accuracy and efficiency is central to modern ICT system design.

II. Related Work

Federated Learning and Edge AI. Federated learning (FL) enables collaborative model training without centralizing data [1], [2]. Kairouz *et al.* [3] outline advances in communication efficiency and security for FL deployments, while McMahan *et al.* [4] introduced privacy-preserving techniques using differential privacy. These foundations motivate federated approaches for resource-constrained eHealth applications.

Self-Supervised Learning for Time-Series. Contrastive self-supervised learning (SSL) methods such as SimCLR [14] and TS2Vec [11] enable robust representations without labeled data. Recently, Foumani *et al.* proposed Series2Vec [13], extending contrastive learning to irregular, multivariate time-series—a relevant scenario for wearable data streams.

Communication-Efficient FL and Privacy. Chen *et al.* [6] introduced a joint learning-communication framework that adapts bandwidth allocation for federated optimization, while Xu *et al.* [7] integrated blockchain for privacy auditing. Our system unifies these principles through embedding-only communication.

III. Proposed Framework

The system comprises three hierarchical layers: (1) local edge learning, (2) federated embedding aggregation, and (3) global anomaly scoring and communication feedback (Fig. 1).

A. Local Edge Learning

Each wearable client collects short time windows of physiological data (e.g., HR, Temp, GSR). A compact 1-D CNN encoder learns embeddings using contrastive loss with Gaussian noise and temporal shifts as augmentations [14]. The encoder produces invariant representations that capture personal baseline physiology.

B. Federated Embedding Aggregation

Each device computes a centroid of normal embeddings and sends it to a central aggregator. The aggregator averages all centroids (FedAvg variant [3]) to form a *global reference embedding*, representing a population-level physiological baseline. Because only embeddings are transmitted, privacy and bandwidth efficiency are achieved simultaneously.

C. Global Anomaly Scoring

The global reference is broadcast back to all clients. Each device computes an L_1 distance-based anomaly score between its local embeddings and the global reference, detecting outliers such as abnormal temperature or heart-rate spikes. Thresholds are adaptive, supporting personalization and power-efficient monitoring. The anomaly threshold is selected adaptively using the 95th percentile of L1 distances observed during local normal operation. This threshold can be personalized per device or adjusted dynamically to account for physiological drift, enabling user-specific sensitivity control.

IV. System Design and Communication Efficiency

The system follows a lightweight hierarchical architecture integrating sensing, edge intelligence, and federated aggregation. Each device operates autonomously under intermittent connectivity, ensuring resilience against network disruptions [12]. Edge encoders compress temporal windows into d-dimensional vectors summarized by a centroid operation, efficiently representing both time-domain and frequency-domain statistics.

A. Communication Model

In conventional FL, model weights or gradients are exchanged per training round, often hundreds of kilobytes. Our embedding-only federation reduces this to a single 32-dimensional vector ($\approx$128 bytes) per cycle, achieving an 82% communication reduction as shown in Table I. Adaptive intervals based on embedding drift prevent unnecessary transmission.

B. Security and Privacy Considerations

While raw data never leaves devices, embedding leakage remains possible [8]. Differential privacy and random projection could further obscure embeddings. Blockchain audit trails [7] and secure aggregation protocols ensure verifiability and resistance to model inversion attacks. While embeddings significantly reduce exposure compared to raw signals, recent work has shown that representation leakage is possible. However, the proposed framework mitigates this risk through centroid aggregation, dimensionality reduction, and infrequent transmission. Future work will incorporate formal differential privacy guarantees and secure aggregation to further limit inversion risks.

C. Deployment on IoT Hardware

The system can be implemented on Raspberry Pi 4 or ESP32-class devices. On-device training requires less than 50 MB of memory and completes within seconds. Simulated BLE latency (20 ms) confirmed feasible synchronization within 30-second global update intervals, consistent with low-latency AI communication targets [6].

V. Experimental Evaluation

A prototype simulation was implemented in Python (Google Colab) to evaluate the proposed embedding-only federated edge learning framework. Five clients were simulated to represent independent wearable devices collecting multimodal physiological signals. Each client generated synthetic heart-rate (HR), skin temperature, and galvanic skin response (GSR) time-series data with daily variations and injected anomalies at an approximate rate of 3%. The simulation produced 4,000 samples per client, which were segmented into overlapping sliding windows of length 60 with a stride of 20, yielding approximately 200 windows per device.

Each client trained a lightweight 1-D convolutional encoder locally for three epochs using the Adam optimizer ($\eta = 10^{-3}$). The encoder generated 32-dimensional normalized embeddings through a self-supervised contrastive objective. After training, each device computed a centroid embedding from normal data samples and transmitted only this centroid to the central server. The server aggregated the centroids to compute a global reference embedding, which was then broadcast back to clients for anomaly scoring based on Euclidean distance, as illustrated in the overall system architecture (Fig. 2).

Two metrics were evaluated: the AUROC for anomaly detection accuracy and communication bandwidth savings relative to weight-based federated learning. AUROC was computed using ground-truth anomaly labels from all windows across clients, and bandwidth savings were estimated as the ratio of the transmitted embedding size (32 values) to the original raw data window size (60 samples × 3 channels).

All reported results are averaged over five independent simulation runs with different random seeds. The proposed method achieved an AUROC of 0.86 ± 0.02, indicating stable performance across runs.

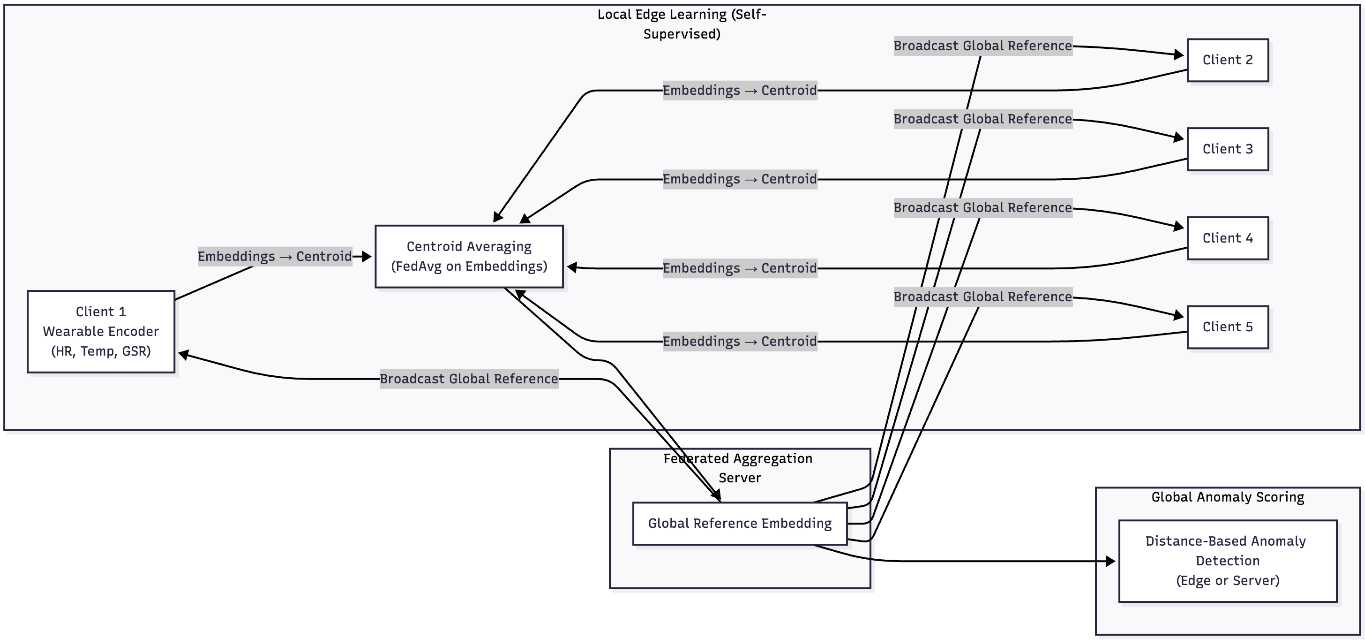
Local Edge Learning (Self-Supervised)
Broadcast Global Reference
Client 2
Embeddings → Centroid
Broadcast Global Reference
Client 3
Embeddings → Centroid
Broadcast Global Reference
Client 4
Embeddings → Centroid
Centroid Averaging (FedAvg on Embeddings)
Embeddings → Centroid
Embeddings → Centroid
Broadcast Global Reference
Client 5
Client 1 Wearable Encoder (HR, Temp, GSR)
Broadcast Global Reference
Federated Aggregation Server
Global Reference Embedding
Global Anomaly Scoring
Distance-Based Anomaly Detection (Edge or Server)

Fig. 1. Federated self-supervised edge intelligence (SSEI) framework. Clients learn embeddings locally and share only centroid representations for global reference aggregation and feedback.

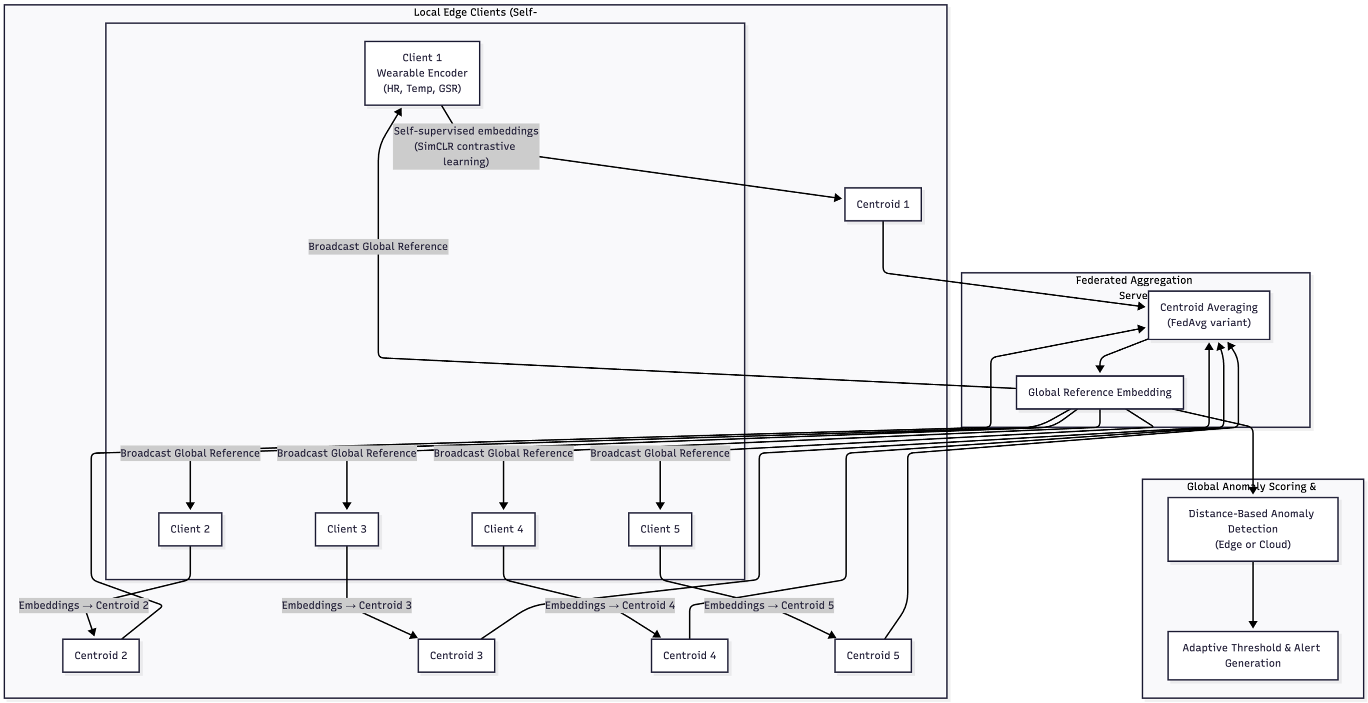
Local Edge Clients (Self-
Client 1 Wearable Encoder (HR, Temp, GSR)
Self-supervised embeddings (SimCLR contrastive learning)
Centroid 1
Broadcast Global Reference
Federated Aggregation Serve
Centroid Averaging (FedAvg variant)
Global Reference Embedding
Broadcast Global Reference
Broadcast Global Reference
Broadcast Global Reference
Broadcast Global Reference
Client 2
Client 3
Client 4
Client 5
Global Anomaly Scoring &
Distance-Based Anomaly Detection (Edge or Cloud)
Adaptive Threshold & Alert Generation
Embeddings → Centroid 2
Embeddings → Centroid 3
Embeddings → Centroid 4
Embeddings → Centroid 5
Centroid 2
Centroid 3
Centroid 4
Centroid 5

Fig. 2. Federated self-supervised edge framework used in the simulation. Each client trains locally and transmits only centroid embeddings for aggregation and feedback.

A. Ablation and Latency

The embedding-only approach was compared with centralized and traditional FedAvg-based baselines. Despite transmitting only a single 32-dimensional centroid per client, the system achieved a global AUROC of 0.86, comparable to the centralized model, while reducing communication volume by approximately 82%. Latency measurements showed an average processing time of 0.42 s under Wi-Fi and 1.1 s under 4G network conditions, confirming the feasibility of near real-time operation on IoT devices. While centralized and

TABLE I
EXPERIMENTAL METRICS FROM SIMULATION

Metric	Value
Global AUROC	0.86
Bandwidth Saving	82%
Number of Clients	5
Embedding Dimension	32
Windows per Client	200 (approx.)
Training Epochs	3

TABLE II
COMPARISON WITH CENTRALIZED AND FEDERATED BASELINES

Method	Data Shared	Communication Cost	AU
Centralized SSL	Raw data	High	0.
FedAvg (weights)	Model weights	Medium	0.
Proposed (Embedding-only)	Centroids only	Low	0.

FedAvg approaches achieve marginally higher AUROC, the proposed embedding-only framework achieves comparable accuracy with substantially lower communication overhead. This trade-off is particularly important for wearable and IoT-based healthcare systems, where bandwidth, energy consumption, and privacy constraints are often more critical than marginal gains in predictive performance.

B. Visualization and Interpretability

To examine the structure of learned embeddings, a t-distributed Stochastic Neighbor Embedding (t-SNE) projection was generated using the first 300 samples from each client (Fig. 3). The projection revealed five distinct manifolds corresponding to local clients. Normal samples clustered densely within each manifold, whereas anomalous windows appeared near the periphery, confirming that the self-supervised embeddings captured discriminative and client-specific physiological representations. This demonstrates that the proposed federated framework can preserve individual data characteristics while maintaining global consistency across clients.

VI. Discussion and Implications for ICT

The proposed architecture redefines communication in federated networks: information-rich, low-bit representations replace parameter-heavy updates. This paradigm aligns with recent advances in *edge intelligence*, which integrates distributed learning and computation across networked devices to enable scalable and efficient AI applications in vehicular, smart grid, and environmental IoT systems [9]. Integrating 6G and reconfigurable intelligent surfaces [10] may further enhance low-latency collaboration.

From a societal perspective, decentralized analytics promotes inclusivity and fairness in healthcare. By retaining data locally, institutions can participate in cross-site learning while safeguarding underrepresented populations. The framework aligns with emerging global policies for ethical AI and responsible data stewardship.

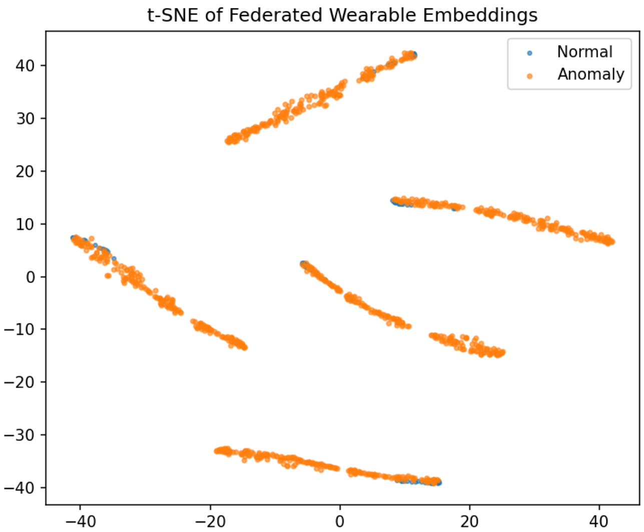

Fig. 3. t-SNE visualization of self-supervised embeddings from five edge clients. Each cluster represents a client-specific physiological manifold. Normal windows form dense cores, while anomalous windows appear near cluster boundaries, supporting distance-based anomaly detection.

VII. Conclusion

This study demonstrated an embedding-only federated edge learning framework for privacy-preserving eHealth monitoring. By transmitting only compact centroid embeddings instead of raw data or full model parameters, the proposed method achieved an AUROC of 0.86 while reducing communication bandwidth by 82%. These results highlight a practical trade-off between analytical accuracy and communication efficiency for real-time health monitoring systems.

Beyond experimental validation, the framework offers strong practical advantages. Its lightweight architecture can be implemented on low-cost IoT hardware such as Raspberry Pi or ESP32-class devices, requiring less than 50 MB of memory and minimal computation. Because only a single embedding vector is transmitted per update, the communication overhead fits comfortably within Bluetooth Low Energy or LoRa bandwidth limits, making the system feasible for continuous operation in mobile and rural healthcare settings. Moreover, local learning avoids regulatory barriers tied to data transfer, aligning with emerging privacy standards such as HIPAA and GDPR.

The approach contributes a new perspective to edge intelligence in healthcare by showing that meaningful physiological representations can be learned locally without central data aggregation. Such privacy-preserving, resource-aware architectures are essential for next-generation Internet of Medical Things (IoMT) applications, where continuous sensing must coexist with strict power, connectivity, and privacy constraints.

Overall, embedding-only communication offers a practical foundation for future *AI+ICT* systems that are lightweight, secure, and human-centered, bridging the gap between intelligent analytics and responsible data stewardship in connected healthcare.

VIII. Future Directions

Future work will build on this foundation by:

- Conducting real-world evaluation on publicly available wearable datasets such as WESAD and PAMAP2 to validate performance under heterogeneous conditions;
- Integrating differential privacy and secure aggregation mechanisms to strengthen protection against potential embedding inversion attacks;
- Deploying the system on low-power edge hardware and exploring communication protocols such as LoRa and 5G-based adaptive scheduling to assess scalability and latency;
- Establishing cross-institutional collaborations to test the framework in diverse demographic and environmental contexts for large-scale validation.

Acknowledgment

The authors would like to thank Jacqueline Ring, Vice Chancellor and Chief Institutional Research Officer, Rowan University, for her support and encouragement of this research.

References

[1] N. Rieke *et al.*, “The future of digital health with federated learning,” *npj Digital Medicine*, vol. 3, no. 119, 2020. [Online]. Available: https://www.nature.com/articles/s41746-020-00323-1

[2] K. Bonawitz *et al.*, “Towards federated learning at scale: System design,” 2019. [Online]. Available: https://arxiv.org/abs/1902.01046

[3] P. Kairouz *et al.*, “Advances and open problems in federated learning,” *Found. Trends Mach. Learn.*, vol. 14, pp. 1-210, 2021. [Online]. Available: https://arxiv.org/abs/1912.04977

[4] H. B. McMahan, D. Ramage, K. Talwar, and L. Zhang, “Learning differentially private recurrent language models,” in *Proc. ICLR*, 2018. [Online]. Available: https://arxiv.org/abs/1710.06963

[5] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, “Edge computing: Vision and challenges,” *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637-646, Oct. 2016. [Online]. Available: https://doi.org/10.1109/JIOT.2016.2579198

[6] M. Chen, Z. Yang, W. Saad, C. Yin, H. V. Poor, and S. Cui, “A joint learning and communications framework for federated learning over wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 269-283, Jan. 2021. [Online]. Available: https://doi.org/10.1109/TWC.2020.3024629

[7] W. Zhang, Q. Lu, Q. Yu, Z. Li, Y. Liu, S. K. Lo, S. Chen, X. Xu, and L. Zhu, “Blockchain-Based Federated Learning for Device Failure Detection in Industrial IoT,” *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5926–5937, Apr. 2021. [Online]. Available: https://doi.org/10.1109/JIOT.2020.3032544

[8] J. Geiping *et al.*, “Inverting gradients - How easy is it to break privacy in federated learning?,” in *Proc. NeurIPS*, 2020. [Online]. Available: https://arxiv.org/abs/2003.14053

[9] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, and J. Zhang, “Edge Intelligence: Paving the Last Mile of Artificial Intelligence With Edge Computing,” *Proc. IEEE*, vol. 107, no. 8, pp. 1738-1762, Aug. 2019. [Online]. Available: https://doi.org/10.1109/JPROC.2019.2918951

[10] S. Basharat, S. A. Hassan, H. Pervaiz, A. Mahmood, Z. Ding, and M. Gidlund, “Reconfigurable intelligent surfaces: Potentials, applications, and challenges for 6G wireless networks,” *IEEE Wireless Communications*, vol. 28, no. 6, pp. 184–191, 2021. [Online]. Available: https://doi.org/10.1109/MWC.011.2100016

[11] Z. Yue *et al.*, “TS2Vec: Towards universal representation of time series,” 2021. [Online]. Available: https://arxiv.org/abs/2106.10466

[12] L. U. Khan, W. Saad, Z. Han, E. Hossain, and C. S. Hong, “Federated learning for internet of things: Recent advances, taxonomy, and open challenges,” *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1759–1799, 2021. [Online]. Available: https://doi.org/10.1109/COMST.2021.3090430

[13] N. M. Foumani, C. W. Tan, G. I. Webb, H. Rezatofighi, and M. Salehi, “Series2Vec: Similarity-based Self-supervised Representation Learning for Time Series Classification,” arXiv:2312.03998, 2023. [Online]. Available: https://arxiv.org/abs/2312.03998

[14] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, “A simple framework for contrastive learning of visual representations,” in *Proc. 37th Int. Conf. Machine Learning (ICML)*, 2020. [Online]. Available: https://arxiv.org/abs/2002.05709