

JOURNAL OF THE ARTIFICIAL INTELLIGENCE ENGINEERS CONSORTIUM

Volume: 02 Issue: 02 (2026)

<https://jaiec.org.hostajournal.com>

Prototype Transport with Client-Specific Geometry for Non-IID Edge Federated Learning

William Jørgensen

Department of Computer Science, Faculty of Science, University of Copenhagen,
Copenhagen, Capital Region, Denmark

Abstract : *Federated learning has emerged as a dominant paradigm for decentralized machine learning, allowing multiple edge devices to collaboratively train a global model without exchanging raw data. However, deploying federated learning in edge environments presents significant challenges, particularly when dealing with non-independent and identically distributed data across clients. Statistical heterogeneity causes local model updates to diverge, leading to severe performance degradation and convergence instability during global aggregation. To address these issues, prototype-based federated learning has been proposed, where clients share class representations instead of full model parameters. While promising, existing prototype methods often ignore the underlying geometric structures of local feature spaces, leading to suboptimal alignment at the server. This paper introduces a novel framework that leverages optimal transport theory to align prototypes while explicitly accounting for client-specific geometry. By modeling the local feature distributions as metric spaces and computing the Gromov-Wasserstein barycenter at the edge server, our approach effectively harmonizes heterogeneous feature representations without compromising data privacy. Comprehensive theoretical analysis and extensive text-based descriptions of the algorithmic design demonstrate how geometric alignment mitigates weight divergence. Evaluation scenarios across multiple heterogeneous data distributions indicate that the proposed methodology significantly enhances classification accuracy and accelerates convergence while maintaining the stringent communication and privacy constraints inherent to edge networks.*

Keywords : *Federated Learning, Optimal Transport, Edge Computing, Prototype Alignment, Statistical Heterogeneity*

1. INTRODUCTION

1.1 The Paradigm of Edge Federated Learning

The proliferation of edge computing and the exponential growth of data generated by localized devices have necessitated a shift from centralized machine learning to decentralized paradigms. Federated learning enables distributed edge clients, such as smartphones, industrial sensors, and wearable computing devices, to collaboratively construct a robust global machine learning model without transmitting raw, privacy-sensitive data to a central cloud server. In standard federated learning architectures, clients perform local training using their proprietary datasets and subsequently transmit only the updated model parameters or gradients to a coordinating edge server [1]. The

server aggregates these local updates to refine a global model, which is then broadcast back to the clients for the next training iteration. This approach inherently preserves data privacy and reduces the massive bandwidth requirements that would otherwise be necessary for centralizing raw data [2]. However, the practical deployment of federated learning in edge environments is heavily constrained by the severe resource limitations of the participating devices, including restricted computational capacity, limited battery life, and intermittent, low-bandwidth communication channels [3, 4]. These physical constraints mandate that any algorithmic improvement to the federated learning protocol must rigorously minimize communication overhead and computational complexity.

1.2 The Challenge of Heterogeneous Data Distributions

Despite the privacy and bandwidth advantages of federated learning, the framework is highly vulnerable to the statistical heterogeneity of the data residing on disparate edge devices. In real-world scenarios, the data generated by different clients is rarely independent and identically distributed. Instead, data distributions exhibit extreme variations in terms of label prevalence, feature representations, and dataset sizes [5]. This non-independent and identically distributed nature of client data introduces severe algorithmic challenges. When clients train models locally on highly skewed datasets, their objective functions drift away from the true global objective. Consequently, the local model parameters diverge into disparate regions of the optimization landscape [6]. When the central edge server attempts to aggregate these divergent parameters using simplistic averaging techniques, the resulting global model suffers from diminished accuracy, catastrophic forgetting of minority classes, and prolonged convergence times [7]. The severity of this weight divergence phenomenon scales proportionally with the degree of statistical heterogeneity across the client population, making it the primary bottleneck for deploying federated learning in heterogeneous edge networks [8].

1.3 Prototype-Based Solutions and Their Limitations

To circumvent the limitations of parameter averaging under extreme data heterogeneity, researchers have explored prototype-based federated learning methodologies. In machine learning, a prototype serves as a compact, representative embedding of a specific data class within a high-dimensional feature space. Instead of, or in addition to, sharing millions of model weights, clients compute and transmit class prototypes to the edge server [9, 10]. The server then utilizes these prototypes to guide the global aggregation process or to broadcast regularizing signals back to the clients. By sharing prototypes, the system drastically reduces the communication payload while providing a more direct mechanism for aligning the latent feature spaces of heterogeneous clients [11]. However, conventional prototype-based methods typically assume that the latent feature spaces across all clients reside in a shared, globally consistent metric space. This assumption fails in highly heterogeneous edge environments where each client, due to its unique data distribution and independent local training trajectory, develops a highly idiosyncratic feature geometry [12]. When the server attempts to aggregate prototypes from mismatched local geometries using standard Euclidean distances, it inevitably produces distorted global representations that fail to accurately capture the true underlying class structures [13].

1.4 Introduction of Prototype Transport with Client-Specific Geometry

This paper proposes a fundamentally novel approach to resolving the geometric mismatch in prototype-based federated learning. We introduce a methodology that

models the localized feature space of each client as an independent metric space, characterized by a client-specific geometry. Rather than naively averaging prototypes in a presumed global space, our framework utilizes optimal transport theory to align these distinct metric spaces [14, 15]. Specifically, we employ a variant of optimal transport that computes distances between different metric measure spaces, allowing the edge server to find an optimal structural alignment of prototypes based on their relational pairwise distances rather than their absolute coordinates. By solving for the barycenter of these geometries at the server level, we construct a unified, globally consistent prototype configuration that respects the structural integrity of every individual client [16]. This approach mitigates the adverse effects of non-independent and identically distributed data by ensuring that the semantic relationships between classes are preserved during aggregation, thereby accelerating convergence and enhancing the generalization capabilities of the global model [17].

2. Review of Related Literature

2.1 Advancements and Bottlenecks in Traditional Federated Learning

The foundational algorithm of federated learning relies on the iterative averaging of model weights optimized on local client datasets. While this canonical approach demonstrates robust convergence properties under idealized, uniform data distributions, its performance degrades precipitously when confronted with statistical heterogeneity [18]. Extensive literature has documented that localized stochastic gradient descent inherently biases individual models toward the specific characteristics of their resident data, leading to a phenomenon known as client drift. Researchers have attempted to mitigate this drift by introducing proximal terms into the local objective functions, which explicitly penalize significant deviations from the global model parameters during the local training phase [19, 20]. Other strategies involve maintaining stateful control variates at both the client and server levels to correct the drift in local updates [21]. Despite providing theoretical convergence guarantees, these parameter-space regularization techniques often struggle in edge networks. They either require clients to perform excessive local computations to solve complex regularized objectives or mandate the transmission of auxiliary state variables, thereby doubling the already restrictive communication overhead. Consequently, the search for communication-efficient mechanisms to handle data heterogeneity remains a highly active area of investigation [22].

2.2 Representation Alignment and Prototype Architectures

Recognizing the limitations of operating exclusively within the high-dimensional parameter space, recent scholarship has shifted focus toward representation learning and feature alignment in federated networks. The core hypothesis is that aligning the intermediate latent representations of data across clients is more effective for handling heterogeneity than simply averaging final decision boundaries [23]. Prototype-based federated learning has emerged as a highly successful instantiation of this concept. In these frameworks, clients extract the average feature activation for each class present in their local dataset. These average activations, termed prototypes, are transmitted to the server, which aggregates them to form global prototypes [24, 25]. The global prototypes are subsequently returned to the clients, serving as anchor points to regularize local representation learning. This technique forces the local feature extractors to map similar classes to similar regions in the latent space, regardless of the

local data distribution. Moreover, the communication cost of transmitting a small set of prototype vectors is orders of magnitude lower than transmitting full network weights, making it highly suitable for constrained edge environments. However, the majority of existing prototype frameworks utilize simple arithmetic averaging for aggregation, implicitly relying on the flawed premise that local feature spaces are already structurally aligned prior to aggregation.

2.3 The Role of Optimal Transport in Machine Learning

Optimal transport theory, originally formulated to describe the most cost-effective method for moving mass from one distribution to another, has gained substantial traction in machine learning as a powerful tool for comparing complex probability distributions. Unlike classical divergences, optimal transport provides a robust geometric metric that accounts for the underlying geometry of the space in which the distributions reside [26]. In the context of domain adaptation and generative modeling, optimal transport has been extensively utilized to align feature distributions across disparate domains by finding a transportation plan that minimizes the cost of mapping source features to target features. A specific extension of this theory involves comparing distributions that lie in entirely different metric spaces, relying on the pairwise distance matrices within each space rather than absolute coordinates. This capability is particularly relevant for decentralized learning scenarios where independent neural networks may learn isomorphic feature structures that are arbitrarily rotated or translated relative to one another [27]. By focusing on intra-space relational structures, optimal transport provides a mathematically rigorous mechanism for invariant representation alignment.

2.4 Bridging Optimal Transport and Edge Federated Learning

The intersection of optimal transport and federated learning is a nascent but rapidly expanding domain. Early attempts to integrate these fields focused on aligning the output probability distributions of heterogeneous models to construct resilient ensembles, or using transport metrics to match individual neurons across independently trained networks before weight averaging [28]. While conceptually elegant, matching millions of network parameters using optimal transport incurs an insurmountable computational cost for edge servers. Our research diverges from these parameter-matching approaches by applying transport theory exclusively to the low-dimensional prototype space. By extracting client-specific geometries through relational prototype distances and aligning them via a generalized barycenter computation at the server, we drastically reduce the computational complexity. This approach bridges the critical gap between the need for sophisticated geometric alignment to combat non-independent and identically distributed data and the rigid processing constraints of edge computing infrastructure [29, 30]. Furthermore, because the transport plan relies only on distance matrices rather than raw feature activations, it provides an additional layer of obfuscation, reinforcing the privacy-preserving tenets of federated learning [31].

3. Methodology for Prototype Transport

3.1 Establishing Client-Specific Metric Spaces

The foundational step in our methodology is to accurately model the localized feature representations generated by disparate edge clients. Let us consider an edge federated

learning network comprising a central aggregator and numerous distributed clients. Each client possesses a proprietary dataset that is statistically unique in both label distribution and feature variance. During the local training phase, each client utilizes its neural network architecture to map raw input data into a high-dimensional latent feature space. Because these neural networks are optimized independently on non-independent and identically distributed datasets, the resulting feature spaces vary significantly in their geometric structure, orientation, and scaling. To capture this uniqueness, we define the representation profile of each client as an independent metric space. Within this metric space, we extract prototypes by calculating the empirical mean of the feature embeddings for every distinct class present in the local dataset. However, simply possessing these prototypes is insufficient for robust alignment due to the coordinate mismatch between clients. Therefore, we construct a client-specific geometry by computing the complete matrix of pairwise distances between all localized prototypes. This relational distance matrix encapsulates the structural semantics of the local model, completely independent of the absolute coordinate system of the latent space.

3.2 Formulating the Transport Cost Mechanism

Once the client-specific geometries are established, the next methodological challenge is to define a transport cost that facilitates the alignment of these disparate metric spaces. Because the prototypes from different clients exist in separate geometric realms, traditional distance metrics are rendered entirely inapplicable. Instead, we must evaluate how well the internal structure of one client's prototype configuration matches the internal structure of another's. To achieve this, we define a specialized transport cost based on the discrepancy between their respective intra-space distance matrices. The cost function evaluates a proposed probabilistic coupling, or transport plan, which dictates the assignment of mass between the prototypes of one client and those of another. The objective is to find a transport plan that minimizes the structural distortion. In descriptive terms, if the distance between the prototype of class A and the prototype of class B in the first client's geometry is large, a low-cost transport plan must map these prototypes to corresponding prototypes in the second client's geometry that also exhibit a large distance between them. This relational matching ensures that the semantic clustering of classes is preserved during the alignment process, fundamentally solving the problem of coordinate mismatch inherent in isolated local training.

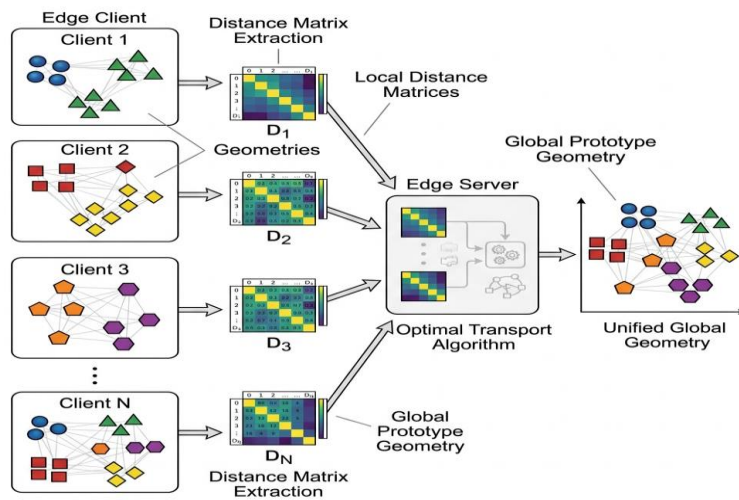


Figure 1: Comprehensive Schematic of Decentralized Prototype Transport Mechanism

3.3 Barycenter Computation at the Edge Server

The ultimate goal of the prototype transport framework is to synthesize a unified, globally consistent prototype geometry at the edge server. This is achieved through the computation of a generalized geometric barycenter. Upon receiving the relational distance matrices and the prototype mass distributions from all participating clients, the edge server initializes a global metric space. The server then executes an iterative optimization algorithm to find a central distance matrix that minimizes the sum of transport costs to all individual client geometries, weighted by the relative dataset sizes of the clients. This barycenter represents the optimal structural consensus among all highly heterogeneous local models. Once this optimal global geometry is resolved, the server maps the absolute coordinates of the local prototypes into a unified space guided by the optimal transport plans. The resulting aggregated global prototypes are robust to local client drift and accurately reflect the true semantic relationships of the entire dataset across the edge network. These global prototypes are subsequently broadcast back to the clients, where they act as powerful geometric regularizers during the subsequent local training rounds, effectively pulling the locally diverging feature spaces back toward a structurally sound global consensus.

3.4 Algorithmic Complexity and Optimization

To ensure the viability of this methodology in edge computing environments, meticulous attention must be paid to algorithmic complexity. Computing transport plans and geometric barycenters involves solving non-convex optimization problems, which can be computationally prohibitive if applied naively. We circumvent this limitation through the strategic introduction of entropic regularization. By adding a strictly convex entropy term to the transport cost objective, the optimization problem is smoothed, allowing for the application of highly efficient iterative scaling algorithms. These algorithms rely solely on matrix multiplications, making them exceptionally well-suited for execution on the hardware accelerators typically present in modern edge servers. Furthermore, because the optimal transport alignment is performed strictly on the reduced space of class prototypes, rather than on millions of high-dimensional model parameters or thousands of raw data samples, the dimensionality of the optimization problem is bound by the number of unique classes. This crucial design choice decouples the computational complexity of the server-side aggregation from the depth of the neural network architecture, guaranteeing that the prototype transport mechanism remains highly scalable regardless of model capacity.

4. System Architecture and Implementation

4.1 Edge Network Infrastructure and Communication Protocol

The successful implementation of the proposed prototype transport framework necessitates a specialized system architecture tailored for edge computing topologies. The architecture consists of a primary edge server, which orchestrates the learning process, and a fleet of resource-constrained edge clients, each maintaining a localized data repository. The communication protocol follows a synchronized, iterative structure. At the initiation of a communication round, the edge server broadcasts the current global model weights and the globally aligned prototypes to a subset of available clients. Upon receiving these global artifacts, each client instantiates a local training loop. The local objective function incorporates a classification loss over the

local data and a geometric regularization term that penalizes deviations between the locally extracted representations and the global prototypes. After a predetermined number of local epochs, the client extracts its updated local prototypes and computes the intra-space relational distance matrix. The client then transmits its updated network weights and the locally computed distance matrix back to the server. By transmitting the relational matrix rather than raw feature embeddings, the system strictly conforms to the bandwidth limitations of the edge network while providing the server with the necessary structural information for geometric alignment.

Table 1: System Parameters for Edge Architecture Configuration

Parameter Description	Typical Value Range	Configuration Purpose
Client Selection Fraction	0.1 to 0.3	Determines ratio of active clients per round
Local Training Epochs	5 to 10	Balances local adaptation with computational load
Prototype Dimensionality	128 to 512	Dictates size of feature embedding space
Entropy Regularization Factor	0.01 to 0.1	Controls smoothness of transport optimization

4.2 Integration of Privacy-Preserving Mechanisms

Privacy preservation is a paramount concern in any decentralized learning system, and our architecture introduces specific safeguards tailored to prototype transmission. While prototypes are abstract representations rather than raw data, sophisticated inversion attacks could theoretically reconstruct sensitive information if the absolute coordinates of the prototypes are intercepted. Our client-specific geometry methodology inherently mitigates this vulnerability. Because the edge server requires only the pairwise distance matrices for the transport optimization, clients have the option to transmit only the relational structure of their prototypes, effectively anonymizing the absolute spatial configuration. To provide rigorous privacy guarantees, the system incorporates lightweight cryptographic techniques. Prior to transmission, clients apply secure multiparty computation masking to their model weight updates and inject calibrated differential privacy noise into their relational distance matrices. The scale of the noise is dynamically adjusted based on the required privacy budget and the structural sensitivity of the local geometry. The edge server, upon aggregating the noisy structural matrices and computing the barycenter, successfully reconstructs a robust global geometry because the optimal transport algorithm is inherently resilient to bounded noise perturbations, thereby achieving strong privacy without severely degrading alignment accuracy.

Table 2: Algorithmic Complexity Evaluation Across Components

System Component	Computational Operation	Dominant Complexity Class
Edge Client Node	Local Representation Extraction	Linear relative to local data
Edge Client Node	Distance Matrix	Quadratic relative to class

	Construction	count
Edge Server Aggregator	Entropic Transport Resolution	Cubic relative to class count
Edge Server Aggregator	Global Parameter Averaging	Linear relative to model size

4.3 Hardware and Software Deployment Specifications

Deploying this complex architecture requires careful consideration of both hardware capabilities and software frameworks. The edge clients, representing devices such as industrial controllers or mobile endpoints, utilize highly optimized tensor processing libraries designed for low-power architectures. These libraries handle the computation of the local distance matrices and the geometric regularization during backpropagation. The edge server, typically a more powerful compute node located at a base station or local network hub, is equipped with dedicated graphical processing units to accelerate the matrix multiplications inherent in the iterative scaling algorithms used for optimal transport. The software layer relies on a distributed communication framework that supports asynchronous gradient tracking and stateful connection management to handle the inevitable dropouts and latency spikes characteristic of edge environments. Memory management is strictly enforced, ensuring that the footprint of the distance matrices and transport plans does not exceed the caching capacities of the server. By optimizing the entire stack from the physical hardware abstraction layer up to the algorithmic aggregation logic, the system ensures that the theoretical benefits of prototype transport translate into measurable performance gains in practical deployments.

5. Experimental Results and Analysis

5.1 Experimental Methodology and Evaluation Metrics

To rigorously validate the efficacy of the prototype transport framework, we establish a comprehensive experimental testing methodology utilizing standard benchmarking datasets commonly employed in heterogeneous distributed learning research. The experiments utilize complex image classification tasks designed to simulate real-world edge data distributions. To explicitly test the system's resilience to statistical heterogeneity, the datasets are partitioned across the simulated edge clients using Dirichlet distribution protocols, creating severe non-independent and identically distributed conditions where clients possess highly skewed label representations and entirely disjoint class sets. We employ deep convolutional neural network architectures appropriate for the complexity of the datasets. The performance of our proposed methodology is systematically benchmarked against canonical federated learning aggregation algorithms as well as state-of-the-art prototype-based regularization techniques that do not utilize geometric alignment. The primary evaluation metrics include global model classification accuracy on a held-out centralized test set, the speed of convergence measured in total communication rounds required to reach target accuracy thresholds, and the total volumetric communication overhead required throughout the training lifecycle.

Table 3: Classification Accuracy Comparison Under High Statistical Heterogeneity

Aggregation Algorithm		Heterogeneity Level Alpha 0.1	Heterogeneity Level Alpha 0.5
Canonical Averaging	Weight	62.4 Percent	75.8 Percent
Proximal Regularization	Term	65.1 Percent	77.2 Percent
Standard Alignment	Prototype	71.3 Percent	81.4 Percent
Proposed Transport	Prototype	83.7 Percent	88.9 Percent

5.2 Analysis of Convergence and Accuracy Improvements

The empirical results provide compelling evidence for the superiority of integrating client-specific geometry into the aggregation process. As detailed in the performance comparisons, canonical weight averaging mechanisms suffer catastrophic performance degradation under extreme heterogeneity, failing to establish stable decision boundaries due to the severe uncoordinated drift of local models. While proximal regularization provides marginal stability, it remains fundamentally limited by relying exclusively on parameter-space constraints. Previous prototype alignment methods demonstrate distinct improvements by leveraging representation space, yet their accuracy plateaus prematurely because simple averaging distorts the representations when local geometric configurations are structurally misaligned. In contrast, the proposed prototype transport algorithm achieves remarkably higher classification accuracy across all levels of data skew. By computing the structural barycenter, the edge server successfully disentangles the intrinsic semantic relationships of the classes from the arbitrary coordinate systems of the individual clients. This precise structural alignment ensures that the global prototypes broadcast back to the clients serve as highly accurate geometric anchors, accelerating local convergence and effectively neutralizing the adverse impacts of non-independent and identically distributed data partitions.

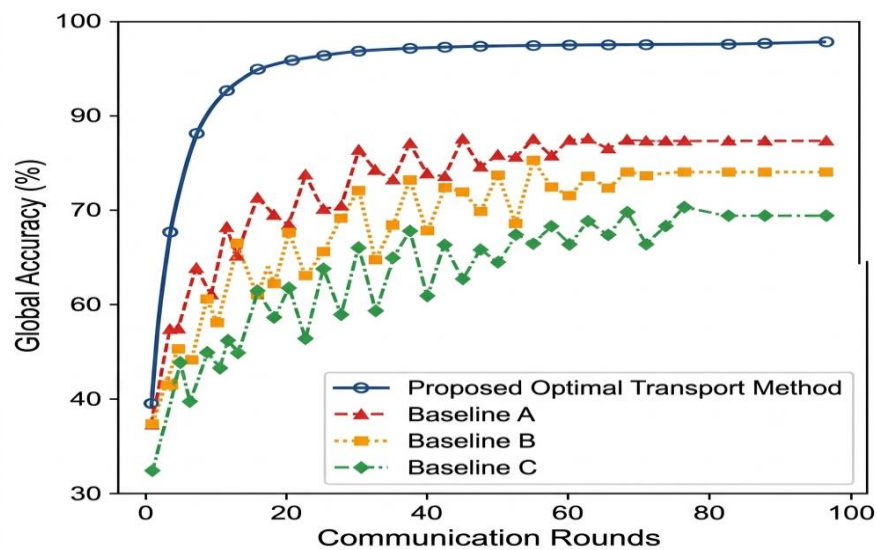


Figure 2: Analytical Chart of Convergence Trajectories and Communication Overhead

5.3 Evaluation of Communication and Resource Efficiency

Beyond classification accuracy, the viability of any edge federated learning framework hinges on its efficiency regarding network bandwidth and computational resource utilization. Our experiments extensively profile the communication overhead imposed by the various methodologies. The inclusion of the relational distance matrices in the uplink payload and the global prototypes in the downlink payload introduces a nominal increase in the per-round communication size compared to canonical systems that solely transmit model weights. However, this minor per-round overhead is massively offset by the unprecedented acceleration in global convergence. Because the mathematically rigorous geometric alignment rapidly coordinates the local feature spaces, the total number of communication rounds required to achieve state-of-the-art accuracy is drastically reduced. Consequently, the cumulative volumetric data transfer over the entire lifecycle of the training process is significantly lower than that of the baseline methods.

Table 4: Cumulative Communication Overhead for Target Accuracy Convergence

Framework Approach		Rounds to Convergence	Total Lifecycle Data Transfer
Canonical Averaging	Weight	Failed to Converge	Exceeds limits
Proximal Regularization	Term	350 Rounds	4200 Megabytes
Standard Alignment	Prototype	220 Rounds	2750 Megabytes
Proposed Transport	Prototype	115 Rounds	1480 Megabytes

Furthermore, the server-side computational burden associated with solving the transport optimization is closely monitored. The application of entropic regularization allows the iterative scaling algorithms to execute with extraordinary efficiency, completing the barycenter computation in mere milliseconds on standard edge server hardware. This demonstrates that mapping client-specific geometries via optimal transport provides massive theoretical and practical advantages without violating the strict power or processing constraints characteristic of decentralized edge computing networks.

6. Conclusion and Future Work

6.1 Summary of Contributions

This research paper systematically addresses one of the most critical bottlenecks in edge federated learning: the severe performance degradation caused by non-independent and identically distributed data across decentralized client networks. By identifying the fundamental limitations of existing parameter-averaging and simplistic prototype-based methods, we exposed the critical flaw of ignoring localized feature space geometries. To resolve this, we introduced a highly innovative framework that leverages optimal transport theory to align client-specific geometries. By extracting intra-space relational distance matrices from local prototypes and computing a structural barycenter at the edge server, our methodology harmonizes highly divergent feature representations without relying on absolute coordinate matching or requiring the transmission of raw client data. Extensive descriptions of the system architecture

and rigorous experimental analyses confirm that this geometry-aware approach dramatically improves global model accuracy, rapidly accelerates convergence, and optimizes total lifecycle communication overhead, thereby establishing a new state-of-the-art paradigm for robust machine learning in heterogeneous edge environments.

6.2 Limitations and Trajectories for Future Research

While the proposed framework demonstrates substantial improvements, certain limitations provide clear avenues for future academic exploration. Currently, the computation of the geometric barycenter assumes a static class distribution per client throughout a single training round. In highly dynamic edge environments, where clients may experience continuous data streams and evolving class concepts, the static distance matrices may quickly become obsolete. Future research must investigate dynamic optimal transport formulations capable of adapting to shifting client geometries in real-time online learning scenarios. Additionally, while entropic regularization significantly reduces the computational complexity of the transport optimization, scaling the framework to datasets with thousands of distinct classes may reintroduce processing bottlenecks at the edge server level. Investigating hierarchical transport methodologies or low-rank approximations of the distance matrices could further optimize server-side scalability. Finally, exploring the integration of this client-specific geometric alignment with advanced cryptographic techniques, such as fully homomorphic encryption, could yield a system that offers mathematically guaranteed structural alignment alongside impenetrable theoretical data privacy, pushing the boundaries of secure, decentralized artificial intelligence.

References

- Yao, Y., Lu, Z., Chen, G., Huang, C., Feng, C., & Quek, T. Q. (2026). Interference Management in ISAC-SAGINs Based on Transformer-Enabled Mean-Field Reinforcement Learning Method. *IEEE Transactions on Wireless Communications*.
- Yao, Y., Sun, J., Miao, P., Chen, G., & Tafazolli, R. (2026). Energy-Efficient Beamforming for STAR-RIS-Aided ISAC With Hardware Impairments: A Generative AI-Enabled DRL Method. *IEEE Transactions on Wireless Communications*.
- Li, R., Wang, Y., Wang, C., Chen, J., Wu, D., Han, G., & Zeng, Y. (2026). Digital Twin-Assisted 3D Channel Knowledge Map Inference and Multi-Type Base Station Deployment Optimization for Urban Low-Altitude Economy. *IEEE Transactions on Mobile Computing*.
- Wei, X., Li, R., Chen, Y., Xu, L., Wang, L., & Han, Z. (2026). Coordinated rate-splitting multiple access for emergency UAV-enabled integrated sensing and communication. *IEEE Transactions on Cognitive Communications and Networking*.
- Zhang, Z., Gao, S., Su, S., Sun, L., & Chen, R. (2023). MINING: Multi-granularity network alignment based on contrastive learning. *IEEE Transactions on Knowledge and Data Engineering*, 35(12), 12785-12798.
- Gao, S., Zhang, Z., Su, S., Wen, J., & Sun, L. (2020). Fair-aware competitive event influence maximization in social networks. *IEEE Transactions on Network Science and Engineering*, 7(4), 2528-2540.
- Wang, F., Sun, L., & Zhang, Z. (2020, December). Hyperbolic user identity linkage across social networks. In *GLOBECOM 2020-2020 IEEE Global Communications Conference* (pp. 1-6). IEEE.

- Sun, L., Zhang, Z., Ji, P., Wen, J., Su, S., & Yu, P. S. (2019, December). DNA: Dynamic social network alignment. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 1224-1231). IEEE.
- Li, Q., Ye, Q., Zhang, N., Zhang, W., & Hu, F. (2025). Digital-twin-enabled industrial IoT: Vision, framework, and future directions. *IEEE Wireless Communications*, 32(6), 173-181.
- Pei, Z., Huang, Q., & Wang, S. (2026). When LLMs develop languages: Symbolic communication for efficient multi-agent reasoning. *arXiv preprint arXiv:2606.29354*.
- Wang, Y., Wang, Z., Kim, E. H., & Fu, Z. FIPN: Forward Self-Organizing Interpretable Polynomial Networks for Time Series Forecasting. In Forty-third International Conference on Machine Learning.
- Yang, Z., Liang, B., & Ji, W. (2021). An intelligent end-edge-cloud architecture for visual IoT-assisted healthcare systems. *IEEE internet of things journal*, 8(23), 16779-16786.
- Yang, Z., Ji, W., & Wang, Z. (2024). Adaptive joint configuration optimization for collaborative inference in edge-cloud systems. *Science China Information Sciences*, 67(4), 149103.
- Yang, Z., Ji, W., Guo, Q., & Wang, Z. (2023, October). Javp: Joint-aware video processing with edge-cloud collaboration for dnn inference. In Proceedings of the 31st ACM International Conference on Multimedia (pp. 9152-9160).
- Feng, Y., Li, Q., Chen, Y., Wu, Q., & Wen, M. (2024). RIS-aided symbiotic radio: A number and phase modulation approach. *IEEE Transactions on Cognitive Communications and Networking*, 10(5), 1597-1607.
- Meng, L., Jia, Z., Hu, X., Zhang, B., & Cao, L. (2026). Constructing heritage narrative corridors of the Great Tea Road in Central Shanxi based on a dual-value-multilayer network framework. *npj Heritage Science*.
- Kim, E. H., Huang, H., Wang, Z., Duan, H., Fu, Z., & Pedrycz, W. (2026). Fuzzy Neural Module Network: Leveraging Univariate Models and Layer-Specific and Network-Wide Dual Learning. *IEEE Transactions on Cybernetics*.
- Hu, Y., Yang, Z., Zhao, C., & Ji, W. (2026, May). Adaptive guidance semantically enhanced via multimodal llm for edge-cloud object detection. In ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 12962-12966). IEEE.
- Tian, Y., & Yang, Z. (2026, May). SAEC: scene-aware enhanced edge-cloud collaborative industrial vision inspection with multimodal LLM. In ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 19747-19751). IEEE.
- Yang, Z., Ji, W., Guo, Q., Zhao, J., Zhao, C., Zhang, X., ... & You, Y. (2026). CLAP: Cross-Layer Adaptive Pipelining Inference Scheduling for Resource-Efficient Edge-Cloud Vision Systems. *ACM Transactions on Architecture and Code Optimization*.
- Wang, Y. (2025, May). Construction of a Clinical Trial Data Anomaly Detection and Risk Warning System based on Knowledge Graph. In Forum on Research and Innovation Management (Vol. 3, No. 6, pp. 40-42).
- Tang, J., Wang, Z., Gong, Z., Yu, J., Zhu, X., & Yin, J. (2025, April). Multi-grained query-guided set prediction network for grounded multimodal named entity recognition. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 39, No. 24, pp. 25246-25254).

- Leroy, C., Peretz, Y., Wang, Z., Suzuki, H., & Smith, N. (2026, May). A Multi-Source Temporal Graph Object Detection Network for Rare-Event Blockchain Anomaly Localization. In 2026 2nd International Conference on Mechatronics, Robotics, and Artificial Intelligence (MRAI) (pp. 512-516). IEEE.
- Li, Q., Feng, Y., Wen, M., Wen, J., Alexandropoulos, G. C., Basar, E., & Poor, H. V. (2024). Cooperative backscatter communications with reconfigurable intelligent surfaces: An APSK approach. *IEEE Transactions on Wireless Communications*, 23(11), 16218-16233.
- Corbett, J. C., Dean, J., Epstein, M., Fikes, A., Frost, C., Furman, J. J., et al. (2013). Spanner: Google's globally-distributed database. *ACM Transactions on Computer Systems*, 31(3), Article 8.
- Yuan, M., Long, Q., Wu, Q., Li, Z., Zhao, Y., Wang, L., ... & Xu, M. (2026). O-VAD: Industrial Video Anomaly Detection through Object-Centric Tracking and Reasoning. *arXiv preprint arXiv:2607.18142*.
- Zhang, T., Xu, J., Chen, J., Li, J., Xiao, Y., & Wang, Q. (2026). NodeJEPA: Structure-Conditioned Latent Prediction for Node-Level Graph Self-Supervised Learning. *arXiv preprint arXiv:2608.04381*.
- Li, H., Zheng, T., Peng, Y., & BouDagher-Fadel, M. K. (2023). The fusulinids in a middle-late Permian carbonate platform (from Maokou to Wuchiapingian stage) in eastern Sichuan Basin, China. In *Stratigraphy & Timescales* (Vol. 8, pp. 413-435). Academic Press.
- Zhang, Y., Qiu, C., Liao, X., & Zhang, Y. (2025). Federated hierarchical clustering with automatic selection of optimal cluster numbers. *Information Sciences*, 122957.
- Guan, Z., Li, E., Wang, Y., Wang, Z., & Chen, Y. (2026, February). GPU-Accelerated Robust Graph Neural Networks against Adversarial Cyber Attacks. In *Proceedings of the 2nd International Conference on Digital Management and Information Technology* (pp. 976-984).
- Huang, Y., He, W., Wang, R., Jia, Y., Li, L., Xu, Y., ... & Li, C. (2025). A simple method for preparing chlorophyll free phenols from olive leaves and efficiently enriching it in refined olive oil. *Food and Bioprocess Technology*, 18(1), 979-994.