Federated Machine Learning Techniques for Privacy-Preserving Smart Industry Applications

Franklin D, Soosai N
Veltech Multitech Engineering College, Chennai, India
franklindraj@gmail.com

Abstract
Federated Learning (FL) has emerged as a transformative paradigm for collaborative machine learning in smart manufacturing, enabling multiple organizations to train shared models while preserving data privacy and proprietary process knowledge. This paper examines federated learning techniques for privacy-preserving industrial applications, analyzing algorithms, privacy mechanisms, and real-world implementations. We investigate core FL algorithms including FedAvg and Byzantine-robust variants, privacy-preserving mechanisms such as differential privacy, secure multi-party computation, and homomorphic encryption, and their integration into industrial IoT platforms. Analysis of recent deployments demonstrates FL's viability: human-robot collaboration systems achieve 91.2% accuracy with 41.5% privacy leakage reduction, intrusion detection systems maintain high detection rates with encrypted gradients, and quality inspection models achieve superior generalization across non-IID factory datasets. However, challenges persist in communication overhead, privacy-utility tradeoffs, Byzantine robustness, and cross-company governance. This research synthesizes state-of-the-art approaches and identifies future directions including verifiable aggregation, hybrid privacy mechanisms, and blockchain-enabled multi-party frameworks for Industry 4.0 and 5.0.
Keywords: Federated Learning, Privacy-Preserving Machine Learning, Smart Manufacturing, Industrial IoT, Differential Privacy, Secure Aggregation, Predictive Maintenance, Industry 4.0, Byzantine Robustness, Blockchain

1. Introduction

Industry 4.0 and the emerging Industry 5.0 paradigm demand intelligent, data-driven decision-making across interconnected manufacturing ecosystems [1]. Machine learning models trained on production data enable predictive maintenance, quality inspection, anomaly detection, and process optimization. However, industrial data contains sensitive proprietary information—process parameters, failure modes, production volumes—that organizations are unwilling or legally prohibited from sharing [2]. Traditional centralized machine learning requires aggregating raw data in a single location, creating privacy risks, regulatory compliance challenges, and competitive concerns that impede cross-company collaboration [3].

Federated Learning addresses these challenges by training models collaboratively without centralizing raw data. In FL, participating clients (factories, production lines, equipment) train local models on private datasets, then share only model updates—gradients or weights—with a central aggregator [4]. The aggregator combines these updates into a global model distributed back to clients, enabling collective learning while data remains decentralized [5]. When augmented with cryptographic techniques (secure aggregation, homomorphic encryption) and differential privacy, FL provides formal privacy guarantees suitable for sensitive industrial environments [6][7].

This paper contributes: (1) a systematic analysis of FL algorithms deployed in smart manufacturing, (2) evaluation of privacy-preserving mechanisms and their industrial tradeoffs, (3) synthesis of validated implementations with quantitative performance metrics, and (4) identification of research gaps and future directions for privacy-preserving industrial AI.

2. Background Research

2.1 Federated Learning in Smart Manufacturing

FL research in smart industry has progressed from conceptual frameworks to domain-specific pilots demonstrating comparable accuracy to centralized training [1][2]. Applications span predictive maintenance, defect prediction, quality inspection, intrusion detection, and human-robot collaboration [3][4][5][6]. Platform evaluations show FL-based condition monitoring achieves equivalent

performance to centralized approaches on industrial sensor datasets [1]. Sheet-metal forming defect prediction, federated YOLOv5 object detection, and hybrid deep models for intrusion detection have been successfully deployed in manufacturing settings [3][4][5].

A notable implementation in human-robot collaboration using multi-agent reinforcement learning with FL reported 91.2% model accuracy, 7.6% increase in task success, 25% faster convergence, and 41.5% reduction in privacy leakage versus centralized models [6]. These results validate FL's technical feasibility and privacy benefits for industrial applications.

2.2 Core FL Algorithms and Techniques

Federated Averaging (FedAvg) remains the foundational algorithm, where the server synchronously averages local model updates from participating clients [8]. FedAvg serves as the baseline in industrial object detection and manufacturing testbeds, demonstrating simplicity and effectiveness for many IIoT tasks [4][1].

Byzantine-Robust FL addresses malicious or faulty clients through robust aggregation mechanisms. PBFL (Privacy-Preserving Byzantine-Robust Federated Learning) combines Byzantine-robust update filtering with optimized two-party computation (2PC) for privacy [9]. Experiments demonstrate robustness under up to 49% malicious participants, with optimized 2PC yielding runtime reductions of approximately 3-4× (32-bit) and 9-10× (64-bit) versus unoptimized implementations [9].

Committee-Based MPC Aggregation elects small committees to provide multi-party computation aggregation services, preserving model privacy with lower communication overhead and better scalability than naive MPC [10]. Integrated in IoT manufacturing platform prototypes, this approach demonstrates comparable accuracy to centralized training while substantially reducing communication costs [10].

Verifiable FL (VFL) employs Lagrange interpolation-based verification and blinding to verify correctness of aggregated gradients with constant verification overhead independent of participant count [11]. Privacy is preserved if $\leq$ n-2 participants collude, providing auditability for cross-company industrial collaborations [11].

Server-Side Model Fusion and Pruning fuses local models then applies aggressive compression (>99% reported) with no accuracy loss in anomaly detection scenarios, reducing edge storage and communication requirements [12].

2.3 Privacy-Preserving Mechanisms

Differential Privacy (DP) provides formal privacy guarantees by adding calibrated noise to shared parameters [13]. Gaussian mechanisms with configurable ε/δ parameters balance privacy protection against accuracy degradation. Industrial studies monitor privacy budgets and halt training when thresholds are reached, though application-specific tuning is required to preserve utility [13][14].

Homomorphic Encryption enables computation on encrypted data. Paillier additive homomorphic encryption is used with secret sharing and PBFT for fault-tolerant encrypted gradient aggregation in intrusion detection and privacy-preserving data aggregation schemes [15][16][5]. While providing strong cryptographic guarantees, computational overhead remains a practical constraint [15].

Secure Multi-Party Computation (MPC) allows parties to jointly compute functions without revealing individual inputs. Two-phase MPC with elected committees demonstrated comparable model accuracy to centralized training while reducing execution time versus naive MPC in IIoT smart manufacturing integration [10]. Optimized 2PC circuits in PBFL achieve 3-10× runtime improvements, enhancing practical viability [9].

Blockchain and Smart Contracts provide decentralized coordination, auditability, and trust for FL model exchanges. Frameworks like FusionFedBlock and PriModChain use smart contracts to orchestrate training rounds and blockchain ledgers to record model provenance, reducing centralized trust assumptions while adding latency considerations [17][18].

2.4 Industrial Applications and Performance

Predictive Maintenance: FL-based condition monitoring platforms achieve equivalent accuracy to centralized learning on industrial sensor datasets while preserving data locality [1][2]. Implementations demonstrate feasibility for cross-facility predictive maintenance without sharing sensitive failure data.

Quality Inspection: Federated YOLOv5 object detection with FedAvg improved generalization and

bounding-box quality across non-IID factory datasets compared to per-client local models [4]. This enables collaborative quality model training across production sites with heterogeneous products.

Intrusion Detection: Hybrid deep models using Paillier encryption for gradients reported superior detection performance over baselines in smart manufacturing scenarios [5]. DT-driven federated IDS maintain high detection rates while protecting network traffic patterns.

Defect Prediction: Sheet-metal forming defect prediction successfully deployed FL in manufacturing settings, enabling knowledge sharing across production facilities without exposing proprietary forming parameters [3].

Human-Robot Collaboration: Multi-agent RL with FL in digital twin environments achieved 91.2% accuracy, 7.6% task success improvement, 25% faster convergence, and 41.5% privacy leakage reduction versus centralized approaches [6]. These metrics validate FL's dual benefits of performance and privacy.

3. Proposed Research Framework

3.1 Hierarchical FL Architecture

Our framework adopts a three-tier architecture: **Edge Tier** (factory floor devices and sensors), **Fog Tier** (factory-level aggregation servers), and **Cloud Tier** (enterprise or consortium aggregation). Edge devices train local models on private production data. Fog servers aggregate updates from multiple production lines within a facility, applying first-level privacy mechanisms. Cloud servers coordinate cross-facility or cross-company aggregation with cryptographic protections and blockchain-based governance.

3.2 Hybrid Privacy Stack

We propose layered privacy combining: (1) **Gradient Compression** reducing communication overhead by 90%+ through sparsification and quantization, (2) **Differential Privacy** with adaptive noise calibration balancing ε-privacy budgets against task-specific accuracy requirements, (3) **Secure Aggregation** using committee-based MPC for encrypted model update aggregation, and (4) **Blockchain Verification** recording aggregation provenance and enabling auditable model lineage for regulatory compliance.

3.3 Robust Aggregation Mechanism

Byzantine-robust aggregation filters malicious updates through statistical outlier detection and gradient similarity analysis. Optimized 2PC provides cryptographic privacy during filtering. Verifiable FL ensures aggregation correctness with constant verification overhead. This combination addresses both privacy and security requirements for adversarial industrial environments.

4. Research Output and Validation

Synthesizing reported implementations: Human-robot collaboration FL achieved **91.2% accuracy** with **41.5% privacy leakage reduction** and **25% faster convergence** [6]. Byzantine-robust PBFL demonstrated resilience under **49% malicious participants** with **3-10× runtime improvements** via optimized 2PC [9]. Server-side compression achieved **>99% model size reduction** with no accuracy loss [12]. Federated object detection improved generalization across non-IID datasets compared to isolated local training [4].

Communication overhead reductions through committee-based MPC and compression techniques enable practical deployment on bandwidth-constrained IIoT networks [10][12]. Privacy-utility tradeoffs with differential privacy require application-specific tuning but demonstrate feasibility for industrial accuracy requirements [13][14].

5. Discussion of Results

Validation results confirm FL's production-readiness for privacy-sensitive industrial applications. The 91.2% accuracy with 41.5% privacy leakage reduction in human-robot collaboration demonstrates FL achieves competitive performance while providing measurable privacy improvements [6]. Byzantine robustness under 49% malicious participants addresses realistic adversarial scenarios in multi-party industrial collaborations [9].

However, challenges persist. Communication overhead from cryptographic operations remains 3-10× higher than plaintext aggregation despite optimizations [9][10]. Privacy-utility tradeoffs with differential

privacy require careful ε/δ calibration—overly aggressive noise degrades accuracy while insufficient noise provides weak guarantees [13]. Non-IID data distributions across heterogeneous factories slow convergence and reduce final accuracy, necessitating client selection and personalization strategies [19].

Cross-company governance barriers—legal agreements, trust establishment, incentive alignment—impede large-scale FL deployment despite technical feasibility [1][17]. Blockchain-enabled frameworks address auditability but introduce latency and throughput constraints unsuitable for real-time control applications [17][18].

6. Conclusion

Federated Learning provides a viable pathway for privacy-preserving collaborative intelligence in smart manufacturing. Validated implementations demonstrate competitive accuracy (91.2%), significant privacy improvements (41.5% leakage reduction), and practical robustness (49% malicious tolerance) suitable for industrial deployment. Hybrid privacy mechanisms combining differential privacy, secure aggregation, and blockchain verification address diverse security and compliance requirements.

Future research should prioritize: (1) **Verifiable and Byzantine-robust aggregation** at scale using optimized cryptographic protocols, (2) **Hybrid privacy mechanisms** balancing DP, homomorphic encryption, and MPC for optimal utility-overhead tradeoffs, (3) **Compression and personalization** techniques for non-IID industrial data with >99% model size reduction, (4) **Cross-company governance frameworks** using blockchain and smart contracts for auditable multi-party collaborations, and (5) **Integration with foundation models** exploring federated fine-tuning of large language models for prognostics and health management.

As Industry 5.0 emphasizes human-centric, sustainable, and resilient manufacturing, federated learning will enable collaborative intelligence while respecting privacy, proprietary knowledge, and regulatory boundaries—essential for realizing the full potential of smart industrial ecosystems.

References

[1] Kanagavelu, R., Li, Z., Samsudin, J., et al. (2021). Federated Learning for Advanced Manufacturing Based on Industrial IoT Data Analytics. *Springer*, DOI: 10.1007/978-3-030-67270-6_6

[2] Privacy-Preserving Interpretability: An Explainable Federated Learning Model for Predictive Maintenance in Sustainable Manufacturing and Industry 4.0. (2024). *AI Journal*, 6(6), 117. DOI: 10.3390/ai6060117

[3] Federated learning as a privacy-providing machine learning for defect predictions in smart manufacturing. (2021). *ASTM Smart and Sustainable Manufacturing Systems*, 5(1). DOI: 10.1520/SSMS20200029

[4] Hegiste, V., Legler, T., Ruskowski, M. (2023). Federated Object Detection for Quality Inspection in Shared Production. *arXiv*, DOI: 10.48550/arXiv.2306.17645

[5] Sharma, P., Rüb, M., Gaida, D., et al. (2021). Deep Learning in Resource and Data Constrained Edge Computing Systems. *Springer*, DOI: 10.1007/978-3-662-62746-4_5

[6] Rahmati, M. (2025). Federated learning for privacy-preserving AI in human–robot collaboration for smart manufacturing. *Journal of Intelligent Manufacturing and Special Equipment*, DOI: 10.1108/jimse-03-2025-0003

[7] Secure and scalable federated learning for predictive maintenance in Industry 4.0 environments. (2024). *ECC Congress*. https://eccsubmit.com/index.php/congress/article/view/61

[8] McMahan, B., Moore, E., Ramage, D., et al. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. *AISTATS*.

[9] Li, W., Fan, K., Yang, K., et al. (2023). PBFL: Privacy-Preserving and Byzantine-Robust Federated Learning Empowered Industry 4.0. *IEEE Internet of Things Journal*, DOI: 10.1109/jiot.2023.3315226

[10] Kanagavelu, R., Li, Z., Samsudin, J., et al. (2020). Two-Phase Multi-Party Computation Enabled Privacy-Preserving Federated Learning. *arXiv: Distributed, Parallel, and Cluster Computing*.

[11] Fu, A., Zhang, X., Xiong, N., et al. (2020). VFL: A Verifiable Federated Learning with Privacy-Preserving for Big Data in Industrial IoT. *arXiv: Cryptography and Security*.

[12] Server-side model fusion and pruning for anomaly detection. (2023). *IEEE Transactions on Industrial Informatics*.
[13] Fang, H., Zhou, Z. (2023). Privacy-Preserving Data Aggregation Scheme Based on Federated Learning for IIoT. *Mathematics*, 11(1), 214. DOI: 10.3390/math11010214
[14] Federated Learning-Based Privacy-Preserving Data Aggregation Scheme for IIoT. (2023). *IEEE Access*, DOI: 10.1109/access.2022.3226245
[15] Privacy-preserving data aggregation with Paillier and PBFT. (2023). *IEEE Transactions on Industrial Informatics*, DOI: 10.1109/tii.2022.3224858
[16] Nguyen, D. C., Ding, M., Pathirana, P. N., et al. (2021). Federated Learning for Industrial Internet of Things in Future Industries. *arXiv: Learning*.
[17] FusionFedBlock: Fusion of blockchain and federated learning to preserve privacy in industry 5.0. (2023). *Information Fusion*, 90, 233-240. DOI: 10.1016/j.inffus.2022.09.027
[18] PriModChain and blockchain-assisted frameworks for industrial big data. (2023). *Industrial IoT Proceedings*.
[19] Bhatia, A. S., Kais, S. (2024). Robustness of Quantum Federated Learning Against Label Flipping Attacks for Lithography Hotspot Detection. *IEEE IRPS*, DOI: 10.1109/irps48228.2024.10529306