Computer Modeling in Engineering & Sciences

Tech Science Press

https://doi.org/10.32604/cmes.2026.081720

ARTICLE

RP-IoMT: A Robust and Provable Framework for Federated Learning Privacy-Preserving Intelligence in Healthcare IoMT

M. Saad Bin Ilyas[1], Sohail Masood Bhatti[1], Ghazanfar Latif[2,*], Sherif Abdelhamid[3] and Arfan Jaffar[1]

[1]Department of Computer Science, Superior University, Lahore, Pakistan
[2]Department of Computing Science, Thompson Rivers University, Kamloops, BC, Canada
[3]Computer and Information Sciences Department, Virginia Military Institute, Lexington, VA, USA

*Corresponding Author: Ghazanfar Latif. Email: glatif@tru.ca

Received: 07 March 2026; Accepted: 31 May 2026; Published: 30 June 2026

ABSTRACT: Federated learning (FL) has emerged as a promising approach for enabling collaborative model training across distributed Internet of Medical Things (IoMT) devices without sharing sensitive data. Existing FL frameworks face significant challenges in healthcare settings, including vulnerability to adversarial attacks, lack of verifiable update integrity, and limited robustness under heterogeneous data distributions. These limitations hinder reliable deployment in critical medical applications. To address these challenges, this paper proposes RP-IoMT, a robust and privacy-preserving FL framework that integrates secure multi-party computation (MPC), zero-knowledge proof-based gradient verification, and robust aggregation mechanisms. The objective of this work is to ensure both the correctness and integrity of model updates while maintaining strong privacy guarantees in adversarial IoMT environments. RP-IoMT enforces bounded client updates using a zero-knowledge clipping protocol (ZKClip), performs secure aggregation using threshold-based MPC, and incorporates robust filtering techniques to mitigate poisoning and backdoor attacks. Experimental results on healthcare datasets demonstrate that RP-IoMT achieves improved predictive performance, reduced attack success rates, and stable convergence under both independent and identically distributed (IID) and non-IID conditions. These results indicate that the proposed framework provides a practical and reliable solution for secure and robust FL in real-world medical Internet of Things (IoT) systems.

KEYWORDS: Federated learning; Internet of Medical Things (IoMT); secure aggregation; multiparty computation (MPC); zero-knowledge proofs; privacy preservation

1 Introduction

The high-paced digitization of the modern healthcare system has brought about a new age where data-based intelligence is central to clinical decision-making, treatment personalization, and patient monitoring. Following the spread of intelligent medical devices, wearable sensors, and remote diagnostic devices, the IoMT has become a paradigm shift as it links healthcare providers, patients, and medical infrastructure with intelligent data exchange mechanisms. The growing role of the IoMT in the healthcare ecosystem is highlighted by the fact that the global IoMT market reached USD 144.23 billion in 2022 and is set to grow at a compound annual growth rate of 20 percent for the next five years, as per the Grand View Research Market Size Report 2030 [1].

Although the use of IoMT technologies has tremendously helped to enhance the accessibility of health care and the efficiency of health care operations, they also have fundamental issues concerning data security,

privacy, and trust management [2]. IoMT devices automatically gather and send massive amounts of sensitive physiological and clinical information, including the electrocardiogram, blood pressure readings, and imaging data [3]. When these data are handled using conventional centralized machine learning (ML)-based methods, they are pooled on a central machine learning server where they are subsequently trained, which, in addition to generating high communication and storage scalability, is accompanied by a high risk of data leakage [4], identity exposure, and compliance with regulatory laws. Such restrictions are especially problematic when it comes to healthcare, where violations may have serious ethical, legal, and safety consequences [5].

To address these issues, FL has become a promising decentralized ML model that would allow several healthcare organizations and IoMT devices to jointly train a common model without any direct sharing of raw data [6]. In FL, the FL participants are locally trained on their own data and only send updates to the model (e.g., gradients or weights) to a coordinating server; their data remains local and improves privacy. This paradigm is successful in overcoming data silos, enables cross-institutional cooperation, and is consistent with privacy requirements, including HIPAA or GDPR [7].

Nevertheless, even with the potential, traditional FL is prone to significant issues when implemented in the environment of IoMTs. Several FL systems make honest-but-curious assumptions for servers, ignoring the potential for collusion or compromise between aggregation nodes [8]. In addition, model poisoning [9], backdoor insertion attacks [10], as well as gradient inversion attacks, may compromise model integrity and patient data privacy.

As malicious or faulty devices can influence the learning process in order to make the model worse, [11] are still hard to get rid of. Furthermore, the IoMT infrastructures are extremely heterogeneous, as they include devices of different computing power, network bandwidth, and data quality. These aspects, along with unbalanced and non-independent distributions of data, make the existing FL solutions less scalable and less reliable. Moreover, the traditional privacy-saving methods like homomorphic encryption (HE) [12] and differential privacy (DP) [13] usually have a high computational cost or compromise the privacy with the quality of the model used, which are not feasible in practical healthcare systems.

In order to fill these gaps, the current paper proposes a contesting FL framework, namely, RP-IoMT (Robust and Provable FL Framework to IoMT), which can be implemented in medical IoMT settings to address the twofold needs of assuring privacy and high performance. In contrast to traditional models that rely on centralized aggregation or lightweight encryption, RP-IoMT incorporates multi-layered cryptographic security, adaptive robustness solutions, and verifiable privacy algorithms to provide data confidentiality and model integrity in distributed health care networks. In essence, RP-IoMT is privacy-enhanced using advanced cryptographic designs that extend beyond the conventional secure aggregation. It also implements a trust-resistant model of collusion that makes use of t-of-n secret sharing, which guarantees that no group of fewer than t servers can reconstruct individual updates by clients, even in the case that a group of servers is compromised.

In addition to privacy, RP-IoMT has response adaptation defense mechanisms that protect the global model against malicious or untrusted clients. It substitutes traditional averaging with a safe trimmed mean aggregation approach, which eliminates statistically deviant updates to reduce poisoning and backdoor attacks. A complementary cosine similarity gating system assesses the directional consistency between the update of each client and a momentum reference and down-weights or discards anomalous contributions. The two processes are applied in a secure multi-party computation (MPC), such that the improvements in robustness do not affect confidentiality. All these design principles contribute to the fact that RP-IoMT is robust and able to resist adversarial interference, and verifiable, in the sense that its privacy-preserving qualities can be proven formally using cryptographic proofs and differential privacy analysis. By combining

justifiable privacy, powerful aggregation, and cryptographic trust, RP-IoMT establishes a secure framework of dependable collaborative intelligence in the healthcare sphere, where sensitive medical data can be used to develop global models without ever leaving its residence.

To conclude, RP-IoMT is the answer to the disconnect between federated learning theory and implemented medical deployment by providing a scalable, verifiable, and safe learning paradigm to the next generation of IoMT-based healthcare intelligence. The rest of this paper will be structured in the following way. Section 2 provides a detailed literature review that includes all the available methods of secure, robust, and privacy-preserving FL, and highlights their constraints within the context of an IoMT setting. Section 3 presents the system architecture and the operational workflow of RP-IoMT, including the involved entities and the model of trust. Section 4 provides the formal security analysis. Section 5 outlines the suggested methodology, comprising the ZKClip mechanism and multi-server MPC aggregation process, as well as the robust filtering strategies that comprise a set of privacy, correctness, and adversarial resilience policies. Section 6 is an account of a wide-ranging performance assessment on various sets of medical data, evaluating accuracy, communication overhead, MPC latency, client-side feasibility, and scalability. Section 7 gives a comparative evaluation of the state-of-the-art FL security frameworks and reveals the benefits of RP-IoMT in terms of robustness, verifiability, and security posture, in general. Lastly, Section 8 provides a conclusion to the article and points to possible directions for further research.

2 Literature Review

IoMT has become an innovative element of the contemporary healthcare framework that has empowered continuous care, automated diagnostics, and real-time support of clinical decisions based on distributed sensing and smart data analytics. The development of smart devices, wearable sensors, and remote diagnostic platforms has created enormous amounts of physiological and clinical data never available before. Nevertheless, the sensitivity of medical information to privacy, as well as the regulatory demands of health sectors, have increased the significance of privacy-conscious, secure, and auditable learning systems. Conventional machine learning designs that are centralized store raw patient information at a single point, which poses significant risks. These restrictions have prompted a considerable body of literature to focus on FL, a decentralized model that is intended to provide distributed model training without data transfer [14,15].

FL was initially popularized in mobile and edge settings, which provide a scalable system of collaborative learning without compromising data locality. FL addresses the issue of data silo in the IoMT environment, both in terms of hospitals, medical centers, and health monitoring devices, and observes privacy standards, including HIPAA and GDPR [16]. Common FL frameworks like FedAvg use local models to be trained on the client nodes and send only model updates to a central location. FedAvg, however, does not protect raw data, but it is still exposed to privacy inference attacks. It is shown that gradients or model parameters may be used to recreate raw data, retrieve patient features, or sensitive clinical behavior in gradient inversion attacks [17,18]. FL is able to enable such heterogeneous systems to jointly train models, but it also introduces a new security threat, including gradient inversion attacks and malicious client behavior [19]. Poisoning and backdoor attacks are the other major weaknesses of conventional FL. Malicious customers may exploit the gradient updates to add malicious behaviors to the global model. Image classification, text analysis, and medical diagnostic tasks have all been shown to be vulnerable to these attacks. These vulnerabilities are further increased by IoMT infrastructures that are defined by heterogeneity of devices, non-IID data, and volatile communication channels. The heterogeneity of the data distributions is a limiting factor to convergence stability, as well as the poor performance of FL in clinical decision support systems. These drawbacks underscore the importance of enhanced privacy, verifiability, and robustness assurances, particularly in medical use when adversarial perturbation can pose a risk to human life. RPEA solves the problem of

robustness through secure computation. The malicious patterns can still be encoded by clients [20]. The Octopus framework studies the concept of robustness from another perspective by integrating compressed gradients, lightweight masking, and anomaly detection to reduce abnormal behavior. Octopus is focused on the efficiency of communication and provides partial robustness with no cryptographically guaranteed correctness [8]. Privacy-preserving FL has been examined in research on several cryptographic and statistical solutions. Secure aggregation (SecAgg), which was first suggested to ensure that servers can not see individual gradients, uses additive masking or homomorphic encryption to ensure that only the aggregate update is disclosed [21]. Although effective when it comes to confidentiality, secure aggregation does not implicitly deal with malicious behavior, update manipulations, or poisoning attacks. Differential privacy (DP) adds mathematically quantified noise to gradients to yield a privacy guarantee. Several instances of DP-FL have been studied in healthcare analytics with formal privacy guarantees on membership inference and reconstruction attacks [22].

More recent studies have further explored FL in IoMT environments with a focus on personalization, security, and scalability. For instance, FedCure introduces a heterogeneity-aware personalized FL framework that leverages a cloud-edge architecture to address device and data diversity in healthcare applications [23]. FL frameworks that integrate secure homomorphic encryption and differential privacy have been proposed to improve communication efficiency and privacy preservation in IoMT systems [24]. Moreover, in FL for distributed healthcare analytics, federal electronic health record (FED-EHR) demonstrates the practical implementation of attaining competitive performance while adhering to data protection regulations such as GDPR and HIPAA [25]. Key limitations in IoMT-based FL have been observed in comprehensive survey studies, including non-IID data distributions, vulnerability to adversarial attacks, and communication overhead, while emphasizing the need to integrate solutions that combine privacy, scalability, and robustness [26]. The integration of blockchain with FL has been investigated in recent studies to enhance trust, security, and auditability in healthcare systems based on IoMT. Blockchain-integrated FL frameworks are suggested to reduce dependency on centralized aggregation, therefore providing decentralized access control for EHR sharing and immutable audit trails [27]. FL with distributed key sharing mechanisms and adaptive differential privacy is adopted to ensure security during cross-institutional cooperation. Blockchain has played an important role in resolving the key limitations of FL, such as centralized aggregation vulnerabilities and trust lacking among participants, by allowing incentive mechanisms and decentralized coordination [28]. Whereas modern frameworks combine optimization techniques and explainable AI within blockchain-based FL systems show significant improvements in model robustness, transparency, and performance in healthcare IoMT applications [29].

FL end-to-end computations are made secure using MPC and Homomorphic encryption (HE). Fully homomorphic encryption allows arbitrary computations on ciphertexts but has persisted as computationally resource-intensive for large deep learning models [30]. Whereas lightweight partially homomorphic schemes lower the overhead, operational flexibility has been reduced. Reliance on a single trusted aggregator is eliminated using MPC protocols by disseminating computation among multiple servers. It also yields stronger privacy assurances but has traditionally experienced high computation and communication overhead, specifically in resource-limited IoT devices.

With the increase in adverse threats, research moves towards robustness in FL. To filter anomalous or malicious updates, Byzantine-resilient aggregation mechanisms such as Krum, Multi-Krum, median aggregation, and trimmed mean have been proposed [31]. These methods select updates or remove outliers using a statistical method with the least divergence from the majority. Though they need honest aggregation servers and plain text gradients for processing, that makes them inconsistent with private or fully encrypted environments.

In order to detect malicious clients, anomaly detection later approaches integrate directional masking using cosine similarity or clustering. Regardless, the applied methods still remain vulnerable when benign patterns, colluding with aggregation servers or adversaries, are represented.

The nonexistence of client-side verifiability is the main limitation of existing FL frameworks, despite having encrypted aggregation; malicious clients disrupted the training by generating manipulated gradients with inflated norms. Zero-knowledge proof (ZKP) is used to ensure that updates from clients conform to norm bounds or model consistency constraints without disclosing the primary model parameters. In privacy-critical IoT applications and financial analytics, ZKClip and proof-carrying updates have grown in adoption [32]. Conversely, for IoMT environments, the integration of ZKPs with FL and robust MPC aggregation remains ignored.

Several multi-server aggregation frameworks, including Octopus, VMFL, and RPEA, have been introduced to mitigate single-server trust assumptions. RPEA employs adaptable pre-processing for efficiency, VMFL enhances verification at the server side, and Octopus underlines compressed gradient masking. However, these frameworks reveal certain limitations; RPEA specifies moderate robustness but lacks proof at the client side, VMFL entails substantial verification overhead and exhibits limited scalability for client populations of large size, and Octopus compromises security for efficiency [33]. Neither of the preceding systems provides verifiable correctness of the client, MPC-protection based on robust aggregation, and full collusion resistance concurrently. VMFL exhibits optimal scalability in multi-round settings in vehicular ad hoc networks (VANETs), and as the population of clients increases, it still engenders nominal overhead. Numerous threat monitoring and intrusion detection frameworks investigate scaling FL using simplified communication strategies, reduced parameter architectures, and lightweight models in industrial internet of things (IIoT) networks. But it has not been sufficiently examined whether strong adversarial robustness or verifiability in these frameworks [34].

The above-mentioned constraints emphasize the need for an FL framework designed purposely for the IoMT environment that should be secure, preserve privacy, and be attack-resilient. Data confidentiality, system verifiability, and model integrity, owing to the critical nature of clinical decisions and sensitivity of patient information, are in high demand in medical-based systems. The proposed RP-IoMT architecture addresses these problems by leveraging several solidly integrated mechanisms. RP-IoMT combines ZKClip for improved integrity to facilitate verification of gradient correctness and norm compliance without revealing the underlying data. It is based on a t-of-n multi-server secret sharing architecture to provide strong collusion protection: if any $k < t$ servers are compromised, then no subset of k servers can reconstruct the client updates. In addition, the framework secures aggregation through MPC-based trimmed mean filtering and cosine similarity gating, promising robustness against poisoning and backdoor attacks. Trustworthy operations from either client or aggregation servers have been guaranteed by these procedures, combined with verifiable and formal audit processes. RP-IoMT maintains computational overhead to a low level, making it applicable in realistic IoMT devices with resource-constrained capabilities regardless of robust security protections. These proposed principles serve to state RP-IoMT as a next-generation FL architecture that perfectly bonds verifiability, robustness, privacy preservation, and comprehensive trust mechanisms that are incompletely examined in prevailing FL solutions proposed in the context of the IoMT healthcare environment.

3 System Overview

In this section, we present the overall architecture of the proposed RP-IoMT (Robust and Provable FL Framework for IoMT). We describe the system and threat models and the design goals that guide the framework's development.

3.1 System Model

As illustrated in Fig. 1, the proposed RP-IoMT framework comprises three primary entities: the Central Coordinator (CC), the Secure Computation Node (SCN), and multiple Medical Clients $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_n$, all operating collaboratively within a distributed healthcare environment. The roles of these components are summarized as follows.

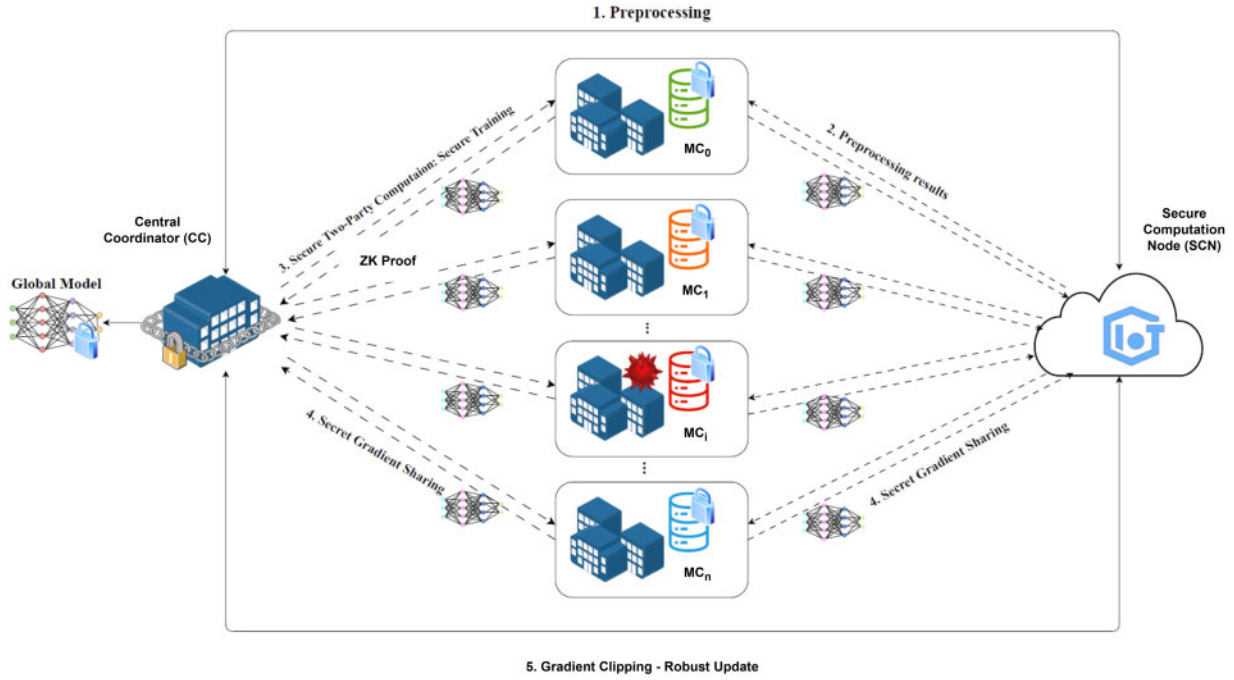

Figure 1: Overview of the RP-IoMT architecture illustrating the interactions between the central coordinator (CC), multiple medical clients (MC_0–MC_n), and the secure computation node (SCN). The figure demonstrates the workflow of gradient clipping, secret sharing, ZKClip proof generation, and MPC-based robust aggregation.

- Central Coordinator (CC): The core entity in the IoMT setting is the CC, which is responsible for coordinating the global learning process. A trusted healthcare authority, hospital network administrator, or research consortium is represented and is pursuing the training of a global diagnostic model using distributed medical. The CC initializes the Global model $\mathbf{w}_g$ and distributes training tasks to the related medical clients that have been selected. The CC engages in *secure backpropagation* and *secure inference* with each client in order to obtain shared secret local gradient updates generated during every round. The task of the CC is to calculate the ℓ_2-norm of the aggregated gradient using *secure two-party computation* and then update $\mathbf{w}_g$ and SCN together. Throughout the training process, confidentiality of the model is preserved, which potentially deals with the sensitive medical records; also, there has been no leakage among the global parameters being shared, which is guaranteed by the CC.
- Secure Computation Node (SCN): The SCN operates as a non-colluding and independent computation partner, typically a trusted healthcare data alliance node or commonly a certified cloud infrastructure provider that aids in carrying out secure operations. The SCN obtains secret shares derived from the clients' gradients and works in partnership with the CC using a secure two-party computation protocol to calculate the ℓ_2-norm of aggregated updates, thus supporting anomaly detection and robust aggregation. Notably, the SCN cannot reconstruct the model or raw gradient parameters at any stage of the process, guaranteeing the privacy of both participants' global model and data. The CC and SCN are

assumed to be non-colluding, so preserving the trust boundary is mandatory for privacy-preserving FL in IoMT.

- Medical Client $\mathcal{C}_i$: All medical clients $\mathcal{C}_i$ represent different healthcare data sources, such as a diagnostic lab, information system of a hospital, clinical monitoring device, or wearable health sensor. Each client holds its private dataset $\mathcal{D}_i$ ($i \in [n]$), while the collective training dataset is represented as $\mathcal{D} = \bigcup_{i\in[n]} \mathcal{D}_i$. During local training, each $\mathcal{C}_i$ computes the gradient $\nabla\mathcal{L}_i(\mathbf{w}_g)$ based on its local model update and transmits only the secret shared or encrypted gradients to the CC and SCN. The framework ensures that no sensitive information about patients, medical conditions, or clinical statistics can be inferred from these updates.

3.2 Threat Model

RP-IoMT is designed for deployment in a distributed and partially trusted IoMT ecosystem, where multiple medical institutions and edge devices collaboratively train a federated model. In such environments, it is unrealistic to assume complete trust among all participants. So, we are considering a rigorous adversarial setting in which attackers may attempt to compromise the integrity of the model, the privacy of data, or system availability, and have complete knowledge of the protocol. The framework relies on two aggregation entities: the Central Coordinator (CC) and the Secure Computation Node (SCN). It is assumed that both of them follow the protocol correctly, but may try to learn some additional information from the processed data. In practice, when different servers are operated by different administrative domains, this honest but curious assumption is made. For the prevention of privacy violations, it is assumed that neither entity colludes. It is ensured that under this assumption, no individual server can reconstruct a client's local gradient using secure MPC and additive secret sharing, and private updates remain protected even if one server is compromised.

We applied a stronger adversarial model on the client side. Participants of IoMT, including wearable devices, edge nodes, and hospitals, may arbitrarily deviate from the protocol or behave maliciously. A client may attempt to disrupt the aggregation process by artificially amplifying update magnitudes, submitting manipulated gradients, backdoor insertion, providing inconsistent secret shares, or attempting model poisoning. Some of them may also attempt to get information through indirect observation of other participants. RP-IoMT integrates ZKClip to impose bounded updates, to limit the influence of adversarial updates, and a robust MPC-based aggregation technique is applied, with verifiable proof mechanisms to ensure correct submission. A passive network adversary is also considered, which is capable of monitoring communication between clients and servers. Although transmitted data cannot be altered using this adversary, sensitive information from observed metadata or messages may be extracted. All transmitted values in this framework are either secret shared or encrypted to mitigate this issue and to prevent the exposure of intermediate computations or plain text gradients.

It is assumed in the current framework that the CC and SCN do not collude ahead of a pre-mentioned threshold. The security analysis is being simplified by this assumption; in certain adversarial healthcare environments, it may be considered strong. The protocol could be extended by leveraging threshold-based guarantees to tolerate partial collusion to address this issue. In actuality, as long as fewer than t servers collude, privacy is preserved and ensured by the use of (t, n)-threshold secret sharing. To strengthen resistance against collusion, raising the threshold parameter or increasing the number of aggregation servers works effectively. Instead, to further reduce reliance on non-collusion assumptions, models with decentralized trust, such as verifiable MPC protocols or blockchain-based coordination, can be integrated. These expansions provide a clear pathway to more adversarial deployment scenarios for adapting RP-IoMT.

The primary objective of designing RP-IoMT is to ensure strong privacy of data, efficiency, and robustness within adversarial IoMT FL settings. Individual gradient privacy, meaning that the party, including

aggregation servers, cannot infer or reconstruct a client's local update in plain text, is guaranteed by this framework. Additive secure MPC and secret sharing help in achieving this level of protection, avoiding reliance on mechanisms that degrade the accuracy of the model, which is high noise differential privacy. Global model privacy is imposed by limiting plain-text access of the aggregated model strictly to the CC in parallel. Thereby inhibiting involuntary reconstruction or disclosure of sensitive clinical representations learned during training; other entities that include the SCN operate only on secret shared or masked values.

Beyond confidentiality, RP-IoMT emphasizes verifiability and robustness in the presence of malicious participants. Each client update must satisfy an enforced ℓ_2-norm bound through ZKClip, ensuring that gradient magnitudes remain within a predefined range without revealing their actual values. This mechanism blocks gradient amplification and boosting attacks and maintains the privacy of updates. The robustness of the MPC-based aggregation reduces the power of the adversarial or poisoning submissions and ensures that the global model remains free of them. Consequently, if some clients act maliciously, they are not allowed to arbitrarily corrupt model convergence.

The framework is also robust to collusion under realistic trust models. As long as no two aggregation servers collude at the same time, privacy guarantees are maintained, and the threshold secret sharing mechanism ensures that under the threshold number of compromised entities, they cannot reconstruct the private gradients. Even if a subset of clients cooperate, they can't deduce updates to honest clients.

In resource-constrained IoMT environments that are critical in nature, computational efficiency and accuracy are the major priorities of RP-IoMT. Given the limited energy, computational resources, and bandwidth available in healthcare wearable nodes and edge devices, the RP-IoMT architecture adopts lightweight cryptographic operations, optimized communication strategies, and efficient aggregation procedures to reduce overhead while preserving overall system performance. The prescribed norm constraint appears as an efficient and practical filter for suppressing and detecting anomalous updates in this setting. The global model is entirely maintained by the CC, compelling direct extraction of the attack model; convergence through exaggerated gradients may be distorted by the malicious clients. The proposed framework mitigates this risk without losing model accuracy or learning stability.

4 Formal Security Analysis

This section presents a formal, theorem-based security analysis of the proposed RP-IoMT framework. We adopt a simulation-based security model and explicitly define the adversarial capabilities and ideal functionality. While the design of the system is grounded in well-established cryptographic primitives, it is important to rigorously characterize the privacy, correctness, and verifiability properties under an explicit adversarial model. To this end, we first formalize the protocol and define the underlying assumptions, followed by the introduction of an ideal functionality capturing the intended behavior of the system. We then present a series of theorems that establish the security of RP-IoMT with respect to gradient privacy, correctness of aggregation, and resistance to malicious client behavior. The analysis follows a modular approach, where the overall security is derived from the composition of secure building blocks, including threshold secret sharing, secure multi-party computation, and zero-knowledge proofs.

4.1 Protocol Model

We formalize RP-IoMT as a protocol $\Pi_{\text{RP-IoMT}}$ executed among a set of medical clients $\mathcal{C} = \{C_1, \ldots, C_m\}$, a central coordinator CC, and a set of aggregation servers $\mathcal{S} = \{S_1, \ldots, S_n\}$. The protocol operates in rounds. At each round r, every client C_i computes a local gradient $g_i \in \mathbb{F}_q^d$ and applies a clipping function to obtain $\hat{g}_i = \text{Clip}(g_i, \tau)$ such that $\|\hat{g}_i\|_2 \leq \tau$.

To ensure correctness without revealing private data, each client produces a non-interactive zero-knowledge proof π_i attesting that the submitted update satisfies the clipping constraint. The clipped gradient is then secret-shared among the n aggregation servers using a (t, n)-threshold Shamir secret sharing scheme. Only updates with valid proofs are accepted. The servers jointly compute, via secure multi-party computation, a robust aggregation function denoted by $\mathcal{F}_{\text{rob}}$. The coordinator finally updates the global model as $w_{r+1} = w_r - \eta G_r$, where G_r is the securely aggregated result.

4.2 Adversarial Model

A probabilistic polynomial-time adversary is considered that would corrupt an arbitrary client subset and up to $t - 1$ aggregation servers. The adversary could not change or alter the communication, but was allowed to observe all protocol messages communicated between honest parties. The following assumptions are adopted: the aggregation servers are believed not to collude beyond the threshold bound and are deemed honest-but-curious. In distributed IoMT environments, this adversarial setting captures real-time deployments where full system takeover is not possible, but partial compromise is likely.

4.3 Security Assumptions

Our analysis relies on standard cryptographic assumptions. First, the underlying Shamir secret sharing scheme provides perfect privacy against coalitions of fewer than t servers. Second, the secure multi-party computation protocol used for aggregation is assumed to be secure in the semi-honest model. Third, the zero-knowledge proof system employed for gradient verification satisfies completeness, soundness, and zero-knowledge properties. Finally, communication channels between honest parties are assumed to be authenticated. Our analysis further relies on standard computational hardness assumptions associated with the concrete cryptographic instantiation of ZKClip. In particular, the commitment scheme is instantiated using Pedersen commitments over a prime-order cyclic group, where the binding property relies on the hardness of the discrete logarithm problem. The zero-knowledge proof component follows a Bulletproofs-style inner product argument in the random oracle model, whose soundness is based on standard discrete-log-type assumptions in the same group. These assumptions are widely used in practical zero-knowledge proof systems and provide the computational foundation for the verifiability guarantees claimed in RP-IoMT.

Underlying Cryptographic Assumptions

In addition to the standard assumptions stated above, the security of the concrete instantiation of RP-IoMT relies on well-established computational hardness assumptions. In particular, the commitment scheme used in ZKClip is instantiated using Pedersen commitments over a prime-order cyclic group $\mathbb{G}$. The hiding property of Pedersen commitments is information-theoretic, while the binding property relies on the hardness of the discrete logarithm problem (DLP) in $\mathbb{G}$. Furthermore, the zero-knowledge proof mechanism follows a Bulletproofs-style inner product argument constructed in the random oracle model. The soundness of this proof system is based on standard discrete-logarithm-type assumptions, ensuring that a malicious client cannot produce a valid proof for an invalid statement except with negligible probability in the security parameter λ. The zero-knowledge property guarantees that no information about the underlying gradient is revealed beyond the validity of the clipping constraint. These assumptions are widely adopted in practical cryptographic systems and provide the computational foundation for the privacy, correctness, and verifiability guarantees established in the formal security analysis of RP-IoMT. All cryptographic components are parameterized by a security parameter λ, and all negligible probabilities are defined with respect to λ.

4.4 Ideal Functionality

We define an ideal functionality $\mathcal{F}_{\text{RP-IoMT}}$ that captures the intended behavior of the protocol. The functionality receives inputs from all clients, filters out invalid updates based on the clipping constraint, and computes the aggregated result using the robust aggregation function $\mathcal{F}_{\text{rob}}$. Only the final aggregated output is revealed to the coordinator, while no additional information about individual client updates is leaked.

4.5 Security Analysis

We now establish the security properties of RP-IoMT through a sequence of theorems.

Theorem 1 (Gradient Privacy): *Let $\hat{g}_i$ denote the clipped update of an honest client. Any adversary controlling fewer than t aggregation servers learns no information about $\hat{g}_i$.*

Proof: Each coordinate of the clipped gradient $\hat{g}_i$ is secret shared using a random polynomial $p_k(x)$ of degree at most $t-1$ over the field $\mathbb{F}_q$, such that the constant term encodes the secret, i.e., $p_k(0) = \hat{g}_i[k]$. The share distributed to server S_j is given by the evaluation $\hat{g}_i^{(j)}[k] = p_k(j)$. Consider an adversary that corrupts a subset of servers of size $c < t$, obtaining the shares corresponding to indices $\{j_1, \ldots, j_c\}$. The adversary's view for coordinate k is therefore the tuple $V_k = (p_k(j_1), \ldots, p_k(j_c))$. For any fixed set of observed values, there exist exactly q^{t-1-c} distinct polynomials of degree at most $t-1$ that are consistent with these evaluations, regardless of the value of the constant term.

As a result, the distribution of V_k is identical for any possible value of $\hat{g}_i[k]$, implying that $\Pr[V_k \mid \hat{g}_i[k] = a] = \Pr[V_k \mid \hat{g}_i[k] = b]$ for all $a, b \in \mathbb{F}_q$. Therefore, the adversary's view is statistically independent of the secret. Extending this argument across all coordinates shows that the full set of observed shares reveals no information about $\hat{g}_i$, completing the proof. □

Theorem 2 (Simulation-Based Privacy): *Under the stated assumptions, the real execution of $\Pi_{\text{RP-IoMT}}$ is computationally indistinguishable from an ideal execution with $\mathcal{F}_{\text{RP-IoMT}}$.*

Proof: We construct a simulator $\mathcal{S}$ that reproduces the adversary's view in the ideal execution using only public information and the final output. First, the simulator replaces all zero-knowledge proofs generated by honest clients with simulated proofs, which are indistinguishable from real ones due to the zero-knowledge property of the proof system. Next, the shares observed by corrupted aggregation servers are simulated as uniformly random values consistent with the secret sharing scheme. By the privacy of Shamir secret sharing established in Theorem 1, these simulated shares are identically distributed to the real shares and therefore reveal no additional information. Finally, the transcript of the secure aggregation phase is generated using the simulator guaranteed by the security of the underlying MPC protocol, ensuring that the simulated interaction is computationally indistinguishable from the real execution. □

To argue indistinguishability, we consider a sequence of hybrid experiments. Let H_0 denote the real execution of the protocol. In H_1, we replace honest-client proofs with simulated proofs; by the zero-knowledge property, $H_0 \approx_c H_1$. In H_2, we replace the real shares with simulated random shares; by Theorem 1, $H_1 \equiv H_2$. In H_3, we replace the real MPC transcript with the simulated transcript; by MPC security, $H_2 \approx_c H_3$. Since H_3 corresponds to the ideal execution produced by $\mathcal{S}$, it follows by transitivity that $H_0 \approx_c H_3$, completing the proof.

Theorem 3 (Correctness): *If all honest parties follow the protocol, the output G_r equals the robust aggregation of all accepted clipped updates.*

Proof: Each accepted update is secret shared and included in the MPC computation. Due to the linearity of Shamir secret sharing and the correctness of the MPC protocol, the final reconstructed result equals the

evaluation of $\mathcal{F}_{\text{rob}}$ on the set of accepted updates. Since invalid updates are filtered prior to aggregation, the result corresponds exactly to the intended computation. □

Theorem 4 (Verifiability): *Any malicious client can cause an invalid update to be accepted only with negligible probability*.

Proof: Acceptance of a client update requires a valid zero-knowledge proof. If a malicious client succeeds in submitting an invalid update with a valid proof, it would violate the soundness property of the proof system. Since the proof system is computationally sound, such an event can occur only with negligible probability. □

Theorem 5 (Aggregation Privacy): *The aggregation servers learn no additional information about client updates beyond the final aggregated output*.

Proof: During aggregation, servers operate only on secret shared values and MPC messages. By Theorem 1, secret shares reveal no information about the underlying inputs. Additionally, the MPC protocol ensures that intermediate computations are able to simulate only the output. Therefore, the servers' view does not leak any additional information. □

Theorem 6 (Composed Security): *The protocol* $\Pi_{\text{RP-IoMT}}$ *securely realizes the ideal functionality* $\mathcal{F}_{\text{RP-IoMT}}$.

Proof: The protocol consists of three main components: zero-knowledge verification, secret sharing, and secure aggregation. Each component is secure under the stated assumptions. By the composition theorem for secure protocols, their combination preserves security. Together with Theorem 2, this implies that the entire protocol is secure in the simulation-based sense. □

Theorem 7 (Robustness): *If the aggregation function* $\mathcal{F}_{\text{rob}}$ is f*-Byzantine resilient, then RP-IoMT preserves the same robustness guarantee*.

Proof: The protocol does not modify the aggregation function but only enforces input validity and computes it securely. Therefore, any robustness guarantee inherent to $\mathcal{F}_{\text{rob}}$ directly carries over to the protocol. □

The above results demonstrate that RP-IoMT offers formal guarantees of privacy, correctness, and verifiability under standard cryptographic assumptions. The security of the overall system follows from the composition of well-established primitives, ensuring that sensitive medical data remains protected even in adversarial IoMT environments.

Composability of the Security Guarantees

RP-IoMT security follows from the modular composition of its integral cryptographic components. The local client update privacy is inherent from the secret sharing layer threshold, specifically, that assures no information about an honest client's update can be learned if there is any coalition of fewer than t aggregation servers. The ZKClip mechanism provides the verifiability of bounded client submissions, whose reliability ensures that norm-violating or distorted updates cannot be tolerated except with a probability that can be ignored. The robust aggregation function is being evaluated correctly without revealing intermediate values with the help of a secure MPC layer, while the final published aggregate is consistent with the committed inputs, and the prescribed secure computation transcript is ensured by the verification layer. The overall protocol inherits correctness, privacy, and verifiability by standard modular composition; subsequently, each of these building blocks satisfies its particular security property under the quantified assumptions; meanwhile, the output of one layer is used only through well-defined interfaces by the next. This argument is formalized in Theorem 6.

5 Proposed Methodology

The proposed RPIoMT presents a collaborative workflow with multiple stages that guarantees robustness, verifiable privacy, and collusion resistance throughout the FL life cycle. The workflow consists of four

key stages: *(1) Setup and Pre-processing*, *(2) Local Training and ZK Clipping*, *(3) Secure Aggregation*, and *(4) Verification and Model Update* as illustrated in Fig. 2. In each phase of the process, it combines cryptographic and statistical mechanisms to ensure secure computation, computational efficiency, and to defend against poisoning or collusion.

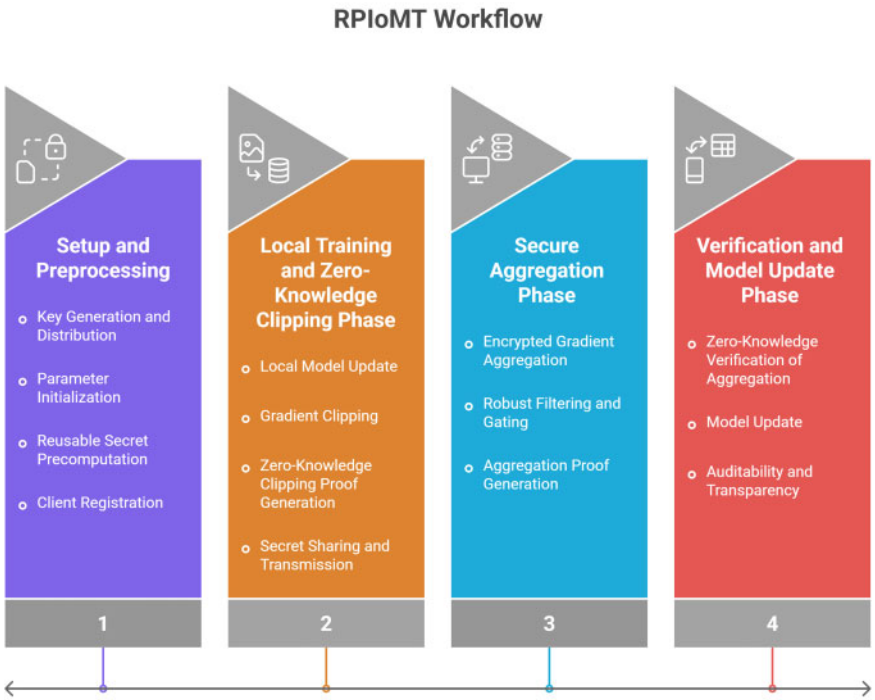

Figure 2: System workflow.

For clarity and consistency, we summarize the key mathematical symbols used throughout the manuscript in Table 1.

Table 1: Notation summary.

Symbol	Description
g_i	Local gradient of client i
$\hat{g}_i$	Clipped gradient of client i
τ	Gradient clipping threshold
$\|\cdot\|_2$	ℓ_2-norm
$\hat{g}_i^{(j)}$	Share of client i held by server j
A_j	Aggregated share at server j
G_t	Aggregated global update at round t
G_t^{rob}	Final robust aggregated update
π_i	Zero-knowledge proof from client i
PRO_t	Aggregation proof transcript

(Continued)

Table 1 (continued)

Symbol	Description
K_t	Public verification parameters
n, t	Number of servers and reconstruction threshold
d	Gradient dimension

5.1 Model Architecture

To evaluate the effectiveness of RP-IoMT across different model classes, we consider three representative architectures: a multilayer perceptron (MLP), a one-dimensional convolutional neural network (CNN), and a modified ResNet50 model.

Let $x_i \in \mathbb{R}^{T\times d}$ denote the input representation for client C_i, where T is the number of time steps in the observation window and d is the number of clinical features. Each sample is constructed from a fixed-length temporal segment, as described in the pre-processing stage. Static features (e.g., age, sex) are concatenated to each time step or appended as an additional feature channel.

MLP

For the MLP model, the input is flattened into a vector $x_i \in \mathbb{R}^{Td}$ and passed through a sequence of fully connected layers:

$$h^{(l+1)} = \sigma\left(W^{(l)}h^{(l)} + b^{(l)}\right), \tag{1}$$

In Eq. (1), $h^{(0)} = x_i$, $W^{(l)}$ and $b^{(l)}$ are learnable parameters, and $\sigma(\cdot)$ denotes a nonlinear activation function such as ReLU. The final output layer produces a probability estimate $\hat{y}_i = \sigma_{\text{sigmoid}}(h^{(L)})$ for binary classification.

CNN

For the CNN model, temporal dependencies are captured using one-dimensional convolutions applied along the time axis. Given input $x_i \in \mathbb{R}^{T\times d}$, the convolutional operation is defined in Eq. (2) as:

$$z_{t,k} = \sum_{j=1}^{K}\sum_{c=1}^{d} w_{j,c,k}\cdot x_{t+j-1,c} + b_k, \tag{2}$$

where K is the kernel size, $w_{j,c,k}$ are convolutional weights, and k indexes output channels. The resulting feature maps are processed through activation and pooling layers to extract temporal patterns.

Modified ResNet50

Since MIMIC-III and the high-time-resolution intensive care unit dataset (HiRID) contain tabular and time-series data rather than images, the standard ResNet50 architecture is adapted by reshaping the input into a structured tensor suitable for convolutional processing. Specifically, the input is represented as $x_i \in \mathbb{R}^{T\times d\times 1}$, where the temporal dimension T and feature dimension d form a pseudo-2D grid.

The model employs residual blocks of the form in Eq. (3):

$$y = F(x, \{W_l\}) + x, \tag{3}$$

where $F(\cdot)$ represents a sequence of convolutional, batch normalization, and activation layers, and x is the identity shortcut connection. This residual mapping enables stable training of deeper networks by mitigating gradient vanishing.

Depending on the feature configuration, the convolutions are implemented either as:

- 1D convolutions along the temporal axis, preserving feature channels, or
- pseudo-2D convolutions over the (T, d) grid to jointly capture temporal and cross feature correlations.

The final feature representation is obtained through global average pooling, represented in Eq. (4) as

$$h = \frac{1}{|S|} \sum_{(t,c)\in S} y_{t,c}, \tag{4}$$

where S denotes the spatial index set. This is followed by a fully connected layer to produce the output probability as shown in Eq. (5)

$$\hat{y}_i = \sigma_{\text{sigmoid}}(Wh + b). \tag{5}$$

Training Objective

All models are trained using the binary cross-entropy loss through Eq. (6):

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)], \tag{6}$$

where $y_i \in \{0, 1\}$ is the ground truth label and $\hat{y}_i$ is the predicted probability.

This unified formulation allows consistent comparison across architectures while ensuring that each model is appropriately adapted to handle clinical time-series data.

5.2 Setup and Pre-Processing Phase

In this stage, the system determines the foundations to establish cryptographic and communication parameters required for verifiable and secure learning. The distributed key generation phase is initiated by the trusted coordinator (TC) among the n aggregation servers $\{S_1, S_2, \ldots, S_n\}$. The goal is to generate a collective secret key in a distributed manner such that no single server learns the full secret.

We adopt a (t, n)-threshold secret sharing scheme based on Shamir's construction [35]. Specifically, the TC samples a random polynomial $f(x)$ of degree $t-1$ over a finite field $\mathbb{F}_q$ such that $f(0) = S$, where $S \in \mathbb{F}_q$ denotes the master secret key.

Each aggregation server S_i is assigned a private key share defined in Eq. (7) as:

$$s_i = f(i), \quad i \in \{1, \ldots, n\}. \tag{7}$$

The secret S is never revealed to any individual server. Instead, reconstruction requires the collaboration of at least t servers. Given a subset $\mathcal{T} \subseteq \{1, \ldots, n\}$ such that $|\mathcal{T}| \geq t$, the secret can be recovered using Lagrange interpolation as given below in Eq. (8):

$$S = \sum_{i\in\mathcal{T}} s_i \lambda_i, \tag{8}$$

where λ_i are the Lagrange coefficients computed over $\mathbb{F}_q$.

We denote the overall key distribution process as: $\text{DKG}(S, t, n) \rightarrow \{s_1, s_2, \ldots, s_n\}$. This construction guarantees the t-of-n collusion resistance property: any coalition of fewer than t aggregation servers obtains no information about the secret S, while any subset of at least t servers can jointly reconstruct it.

The TC defines the global model architecture, cryptographic parameters, and the total number of training rounds T, λ (security parameter), cryptographic parameters: γ (random seed), and τ (gradient clipping threshold). A secure pseudo-random generator (PRG) derives the random seed into reusable key materials, thereby minimizing key exchange overhead over multiple rounds. Similar to the centralized pre-processing approach in RPEA, RPIoMT pre-computes reusable random masks and encrypted linear operation primitives (e.g., matrix multiplications and convolutions). These precomputed primitives are shared among the servers before training begins, ensuring that subsequent rounds of encrypted operations can be executed with minimal interaction, thereby reducing latency and bandwidth consumption. Cryptographic credentials that include public/private key pairs and the clipping threshold τ have to be obtained by each participating medical client, like a hospital, monitoring node, or wearable gateway, by registering with the TC. This process, managed by the TC, assures that every client is authenticated and bound to a verifiable assumed name. The RPIoMT network is provided with distributed encryption keys, preprocessed computation tables, and a reusable verification seed at the end of this phase, enabling low-latency and scalable secure aggregation in successive stages.

The datasets HiRID and MIMIC-III both contain irregularly sampled time-series clinical measurements and static patient attributes. Maintaining a uniform input representation of data, feature extraction by selecting clinically relevant variables is performed at first, including laboratory measurements and vital signs. Forward filling for time-series data, followed by mean imputation for remaining gaps, is applied for handling missing values. Z-score normalization is used for feature normalization. Patient data are segmented into fixed-length observation windows to capture temporal dynamics. A binary outcome indicating either clinical deterioration or mortality risk is predicted using each window. As a binary classification problem, the prediction task is formulated, and if the target event occurs within a predefined prediction horizon, labels are assigned on this basis. To prevent data leakage across partitions, the dataset is split at the patient level into training, validation, and test sets using a standard 70/10/20 ratio.

5.3 Local Training and Zero-Knowledge Clipping Phase

Each client performs local training on private IoMT data and generates a privacy-guaranteed update using the ZKClip mechanism once initialization is complete.

Every client C_i downloads the current global model w_t and trains it on its private dataset D_i for one or more local epochs, obtaining the gradient vector using Eq. (9):

$$g_i = \nabla_{w_t} L(w_t, D_i), \tag{9}$$

where $L(\cdot)$ denotes the local loss function, for example, cross entropy.

To limit the influence of adversarial gradient scaling or outliers, each client applies an ℓ_2-norm constraint in Eq. (10) [36]:

$$\hat{g}_i = \frac{g_i}{\max(1, \|g_i\|_2/\tau)}, \tag{10}$$

ensuring that $\|\hat{g}_i\|_2 \leq \tau$.

The maximum acceptable update magnitude is controlled by the clipping threshold τ and therefore oversees the trade-off between robustness and optimization flexibility against gradient amplification. In general, τ is considered a tunable hyperparameter, and validation-based sensitivity analysis is used to select this. To examine the impact of τ clipping threshold on convergence behavior and adversarial robustness, a detailed version is presented in the Sensitivity Analysis of the Clipping Threshold.

A ZKClip protocol is executed by each client, which produces a non-interactive proof π_i that adheres to the norm bound without exposing g_i as represented in Eq. (11):

$$\text{ZKClip}(g_i, \tau) \rightarrow (\hat{g}_i, \pi_i). \tag{11}$$

ZK range proofs are the basics of this method, allowing the TC and servers to verify gradient validity before aggregation while preserving complete privacy. It should be understood that Eq. (5) is a proving link in the clipping relation defined by the statement that the submitted update satisfies the norm bound $\|\hat{g}_i\|_2 \leq \tau$ and corresponds to a correctly clipped local gradient. In the revised protocol, this relation is instantiated using a Pedersen commitment-based non-interactive zero-knowledge proof system with a Bulletproofs-style inner product argument, thereby avoiding a trusted setup while preserving compact proof size.

The clipped gradient $\hat{g}_i$ is divided into additive shares n via a threshold secret sharing function using Eq. (12) [37]:

$$[\hat{g}_i] = \{\hat{g}_i^{(1)}, \hat{g}_i^{(2)}, \ldots, \hat{g}_i^{(n)}\}, \quad \sum_{j=1}^{n} \hat{g}_i^{(j)} = \hat{g}_i. \tag{12}$$

Each encrypted share is transmitted to its designated aggregation server, whereas before acceptance, the associated proof π_i is verified by TC. Every client plays a role in the update that is verifiably bounded and privacy-protected, helping the next phase to aggregate gradients robustly and securely without the threat of unbounded manipulations.

Instantiation and Complexity of ZKClip

To make the verifiable clipping mechanism concrete and reproducible, we instantiate ZKClip as a non-interactive zero-knowledge range proof system over Pedersen commitments in the random oracle model, following the proof-carrying update paradigm commonly used in verifiable privacy-preserving learning. For each client update, the prover commits to the clipped gradient vector $\hat{g}_i$ and proves that it is a correctly formed bounded update satisfying the norm constraint without revealing the underlying gradient values. The statement being proved by client C_i is that there occurs a witness conforming to blinding randomness and the local gradient g_i such that the published commitment opens to $\hat{g}_i = \text{Clip}(g_i, \tau)$ and the resultant vector satisfies $\|\hat{g}_i\|_2 \leq \tau$.

The clipping constraint has been encoded as an arithmetic circuit in our method, whose size is linearly scaled with the dimension d of the gradient. There are three logical parts of this circuit: commitment consistency, coordinate-wise boundedness, and norm verification. The first part ensures that the committed vector only corresponds to the clipped update submitted. The second part ensures that each coordinate should lie in the prescribed numeric range defined by the quantization domain. The third part proves that the squared ℓ_2 norm of the clipped vector does not go beyond the threshold τ^2. Suppose N_{zk} indicates the number of multiplication constraints in this circuit. Then, under fixed-precision encoding, for a gradient vector of dimension d, we have $N_{\text{zk}} = O(d)$.

In our implementation, the proof system is characterized by a parameter as a security level λ, and we focus on $\lambda = 128$ bits, which is considered to be a standard value for modern practical implementations. In this system with the vector range parameters, contingent on the concrete inner product proof optimization, the proof size rises logarithmically or near logarithmically, though the overall proving time is dominated by multi-scalar multiplications over the commitment group. More particularly, the generation of proof acquires $O(d)$ prover work, whereas proof verification requires $O(\log d)$ to $O(d)$ group operations depending on

proof aggregation and batching choices. We use a lightweight Bulletproofs-style inner product argument to avoid trusted setup and to keep the proof compact in the current RP-IoMT prototype.

From the perspective of the system, three types of overhead are added by the ZKClip: proof generation at the client, proof verification before secure aggregation, and proof transmission over the network. Let $T_{\text{prove}}(d), S_{\pi}(d)$, and $T_{\text{verify}}(d)$ denote the proving time, proof size, and verification time, respectively. Then, the cryptographic overhead per round at the client-side can be written as shown in Eq. (13)

$$T_{\text{zk}}(d) = T_{\text{prove}}(d) + T_{\text{share}}(d), \tag{13}$$

where $T_{\text{share}}(d)$ is the time required to secret share the clipped gradient. Likewise, the communication overhead introduced by verifiability in Eq. (14)

$$C_{\text{zk}}(d) = S_{\pi}(d) + S_{\text{com}}(d), \tag{14}$$

In Eq. (14), $S_{\text{com}}(d)$ is the size of the commitment metadata attached to the update. Particularly for medium and large models, the supplementary cost stays reasonable relative to model transmission, because the proof is generated once per round per client and also verified once before acceptance. With the experimental communication overhead reported later in Section 5, this observation is consistent.

To ensure reproducibility, the exact circuit size depends on three implementation choices: gradient encoding precision, dimension of the local model update, and whether the norm check is applied through an equivalent range-constrained representation or directly. Gradients are quantized to fixed-point integers before proof generation and commitment in our work, and the clipping threshold τ is characterized in the same domain. While preserving the semantics of bounded gradient clipping used by the aggregation pipeline and allows the norm constraint to be expressed as an integer relation suitable for efficient proof generation.

The ZKClip component is modeled as a non-interactive ZKP system, with convincing computational accuracy, completeness, and ZK in the formal security analysis. Computational accuracy guarantees that a malicious client could not influence the verifier to allow an update breaking the clipping relation except with probability that is negligible in the parameter of security λ. Honestly generated clipped updates are always accepted, as assured by the completeness. The proof exposes no information about the inherent gradient as the validity of the claimed constraint is promised by the ZK.

ZKClip does not require a trusted setup in the current implementation. Bulletproof-style inner product argument and Pedersen commitments are the base of the proof system in the random oracle model. The constraint system grows linearly with d, for a gradient vector of dimension d represented in fixed precision, while proof generation is dominated by $O(d)$ prover work. The verification cost is substantially lower than full secure retraining, and the proof size remains compact relative to the model update size, making the method viable for the work being considered in the healthcare FL environment.

5.4 Secure Aggregation Phase

The aggregation servers use a t-of-n secure MPC protocol that preserves privacy and correctness under the non-collusion assumption to mutually compute a robust global update during this phase. Secret shares of clipped update $\hat{g}_i$ are first distributed by each client C_i among the aggregation servers. Let $\hat{g}_i^{(j)}$ denote the share of client i held by server S_j. Each server locally accumulates the shares received from all participating clients for linear aggregation using Eq. (15):

$$A_j = \sum_{i=1}^{m} \hat{g}_i^{(j)}. \tag{15}$$

Since secret sharing is linear, the collection $\{A_j\}_{j=1}^{n}$ constitutes a valid sharing of the aggregated update. Once at least t servers participate, the aggregate can be reconstructed using the corresponding Lagrange coefficients:

$$G_t = \sum_{j \in T} \lambda_j A_j, \tag{16}$$

In Eq. (16), $T \subseteq \{1, \ldots, n\}$ is any subset of servers such that $|T| = t$, and λ_j denotes the Lagrange interpolation coefficient associated with server S_j. This guarantees that no coalition of fewer than t servers can recover the aggregate or any individual client update.

RP-IoMT extends secure summation with robust aggregation primitives implemented under MPC to defend against backdoor and poisoning attacks. To be precise, the protocol pertains to cosine similarity-based gating and coordinate-wise trimmed mean filtering of the secret shared client updates before sharing the final aggregate. Let $\mathcal{F}_{\text{rob}}$ represent the secure robust aggregation function. Then the secure global update can be stated as in Eq. (17):

$$G_t^{\text{rob}} = \mathcal{F}_{\text{rob}}(\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_m)\,. \tag{17}$$

Two defense mechanisms are combined here $\mathcal{F}_{\text{rob}}(\cdot)$. Trimmed mean filtering discards extreme client contributions, firstly by reducing the impact of coordinate-wise outliers. Secondly, weight updates whose direction diverges significantly from the dominant aggregation trend are downscaled by cosine similarity gating. Secret shared values are used to execute both steps so that neither individual client gradients nor intermediate statistics are revealed.

The secure evaluation of non-linear operations required by $\mathcal{F}_{\text{rob}}$ relies on standard MPC subroutines, including secure multiplication and comparison. In particular, multiplication over secret shared values is implemented using Beaver triples. Let $[x]$ and $[y]$ denote secret sharing of two values, and let $([a], [b], [c])$ be a Beaver triple such that $c = ab$. The servers first reconstruct

$$d = x - a, \qquad e = y - b, \tag{18}$$

and then compute

$$xy = c + db + ea + de. \tag{19}$$

This allows the protocol to evaluate robust aggregation primitives efficiently while keeping the underlying inputs private.

The complexity during communication of this phase is dominated by MPC interaction and shared distribution. For a gradient of dimension d, $O(nd)$ communication per client per round is witnessed as each client in particular sends one share of its update to every aggregation server. Only local addition is required for Linear aggregation followed by threshold reconstruction, whereas robust aggregation incurs additional MPC rounds for secure comparison, gating, and trimming.

The servers generate the robust global update G_t^{rob} together with a proof transcript PRO_t at the end of the aggregation phase, confirming that the secure computation has been completed successfully with the recommended protocol. In the next phase, this proof object is later checked by the verification procedure. Subsequently, the aggregation stage without exposing individual client contributions brings forth an attack-resilient and privacy-preserving global update.

Verification of Secure Aggregation

The aggregated update after the secure aggregation phase, together with a proof transcript, is returned by the server, indicating that the aggregation was carried out correctly on the accepted client inputs.

Let the aggregated gradient output produced be denoted by Z_t at round t, and the proof object generated by the aggregation servers is denoted by PRO_t. This proof object represents the proof material or cryptographic transcript produced during the secure aggregation stage and is not an arbitrary scalar value, including the consistency checks, verification responses, and commitments, which are essential for the protocol. Public verification parameters associated with round t are denoted by K_t, including the public aggregation context, commitment bases, and the round-specific verification metadata derived during setup. The role of the verifier is to confirm that PRO_t is a legitimate proof that the published aggregate Z_t has been computed honestly from the accepted, regular client updates implemented by the protocol under the aggregation function.

Accordingly, the verification step should be understood as the evaluation of a verification algorithm in Eq. (20)

$$\mathsf{VerifyAgg}(K_t, Z_t, \mathrm{PRO}_t) \in \{0,1\}, \tag{20}$$

In Eq. (20), the output 1 indicates that the proof transcript is valid and that the aggregate Z_t is consistent with the committed inputs and the prescribed secure aggregation procedure. Rather than expressing verification as an unexplained algebraic identity, we model it as an explicit predicate using Eq. (21)

$$\mathsf{VerifyAgg}(K_t, Z_t, \mathrm{PRO}_t) = 1, \tag{21}$$

where Z_t represents the aggregated update shared in round t, the proof transcript generated during secure aggregation is denoted by PRO_t, and the public verification parameters associated with that round are denoted by K_t. If and only if the transcript is consistent with the prescribed MPC computation, the verification algorithm accepts it, the proof material is valid under the public round context, and the aggregate is bound to the known client commitments. This design makes the completeness and accuracy requirements of the verification layer unambiguous.

Under this formulation, the coordinator updates the global model according to Eq. (22)

$$w_{t+1} = w_t - \eta Z_t, \tag{22}$$

provided that the verification algorithm accepts. Otherwise, the round output is discarded, and the corresponding server behavior is flagged for audit or recovery. This revised description more accurately reflects the role of the verification phase in RP-IoMT and avoids the ambiguity of the earlier shorthand expression.

6 Performance Analysis

To evaluate the performance of RP-IoMT, experiments are being performed. Using varying numbers of clients and aggregation rounds, a simulation has been carried out to validate the reliability of our proposed scheme. Also, experiments were conducted on a dataset related to medical dataset to show the scalability and effectiveness of our scheme in IoMT. Finally, we compared RP-IoMT with existing work to emphasize its advantages and reduction in terms of communication cost.

- **Setting of Experiment:** To evaluate the performance of RP-IoMT in a better way, we implemented RP-IoMT in a real-world FL environment and utilized Python to experiment on a PC having an Intel Core i7-12700K processor of 3.60 GHz and supported by 32 GB RAM.

- **Dataset and Model Architecture:** We applied the MIMIC-III intensive care unit (ICU) database to approximate mortality risk among patients with trauma and to identify early signs of clinical degradation, as reflected by a sharp increase in risk scores. The MIMIC-III dataset is publicly available and comprises over 60,000 ICU admissions distributed across 25 CSV files. It includes both dynamic clinical measurements like heart rate and blood pressure, and static patient attributes, i.e., age and sex, making it well-suited for modeling time-dependent processes in critical care. Exclusion criteria proposed by Johnson and Mark [38] have been followed for patient selection. Particularly, pediatric and newborn patients younger than 16 years and patients with an ICU stay shorter than four hours. In addition, those patients who have several ICU stays within a single hospital are removed. To emphasize trauma-related scenarios, data of patients with ICD-9 codes only corresponding to external traumatic injuries is maintained. There are three representative architectures included in the global modeling framework: a multilayer perceptron (MLP), the widely adopted ResNet50 model, and a convolutional neural network (CNN) [39].

  We also incorporate the HiRID dataset to further evaluate the generality of RP-IoMT under more realistic conditions. HiRID offers high-frequency multivariate clinical measurements, including vital signs, treatment records, and laboratory results. It exhibits greater variability across patient populations and offers finer temporal granularity compared to MIMIC-III, making it particularly appropriate for assessing FL systems in heterogeneous IoMT environments. For uniformity, normalization, feature selection, and patient-level partitioning in pre-processing are applied with similar steps across both datasets.
- **FL Configuration:** FL environment has been simulated consisting of $N = 20$ clients. The dataset is being partitioned across the clients under both IID and non-IID settings. Data is randomly distributed across clients in the IID settings, and in the case of non-IID settings, quantity skew and label skew strategies are applied to exhibit realistic heterogeneity in clinical data. Local training for $E = 5$ epochs is performed by each client, per communication round, using the Adam optimizer with a learning rate of 10^{-3} and a batch size of 32. After over 20 communication rounds, the global model is updated. The global learning rate is set to $\eta = 0.01$. All experiments are performed and repeated three times with different random seeds, and the results are tested as averages to reduce sensitivity to initialization and to improve reliability. For all evaluated methods and settings, the standard deviations across runs were consistently below 0.5% AUC, confirming the stability of the reported results. Model convergence is governed based on validation performance, and if no improvement is observed over five consecutive rounds, early stopping is applied.

  To represent more accurately the execution model of secure FL, the experiments are conducted in a distributed manner in which clients and aggregation servers are implemented as independent processes. Emulating the interaction pattern of a multi-party environment, these processes communicate through inter-process message passing. In this setup, the CC, SCN, and participating clients execute their assigned protocol steps separately, which includes local training, secret share transmission, MPC-based aggregation, and verification. During secure aggregation, this distributed execution model allows us to control both communication and computational overhead. The reported latency particularly includes local cryptographic computation and also message exchange among the participating entities during MPC operations synchronization and proof verification. This setup provides a realistic approximation of a distributed IoMT deployment and enables more accurate evaluation of protocol-level overhead, although all processes are executed within a controlled environment.
- **Non-IID Data Distribution:** We introduce non-IID data distributions across clients to simulate realistic IoMT environments. Particularly, we consider two common partitioning strategies. Each client is assigned data from a limited subset of classes, creating heterogeneous label distributions in the label

skew setting, whereas in the quantity-skew setting, clients receive different amounts of data, reflecting variability in data availability and device activity. Formally, let the global dataset be denoted by D and D_i the local dataset of client C_i. Instead of assuming $D_i \sim D$, we construct D_i such that:

$$P(y|C_i) \neq P(y),$$

where $P(y|C_i)$ denotes the label distribution for client C_i. This setup captures the statistical heterogeneity commonly observed in IoMT deployments.

The hyperparameter settings used for all experiments are summarized in Table 2. To ensure acceptable comparison, these parameters are kept constant across all datasets and models. The selected configuration points toward frequently used settings in FL for healthcare applications, balancing model stability and convergence speed. To improve reliability, all the results are averaged over three independent runs with different random seeds.

Table 2: Hyperparameter settings for federated training.

Parameter	Value
Number of clients (N)	20
Communication rounds	20
Local epochs (E)	5
Batch size	32
Optimizer	Adam
Local learning rate	1×10^{-3}
Global learning rate (η)	0.01
Loss function	Binary cross-entropy
Activation function	ReLU/Sigmoid (output)
Initialization	Xavier initialization
Regularization	None/Dropout (if used)
Early stopping patience	5 rounds
Number of runs	3 (averaged)

6.1 Adversarial Evaluation Setup

To ensure a reproducible robustness evaluation, we explicitly define the adversarial model used in our experiments. We consider a client-side threat setting in which a fraction α of participating clients are malicious and may arbitrarily manipulate their local updates before submission. Unless otherwise stated, we vary $\alpha \in \{0.1, 0.2, 0.3\}$, corresponding to 10%, 20%, and 30% compromised clients, respectively. This setting reflects realistic FL scenarios in which only a subset of devices or institutions is compromised, while the aggregation servers remain non-colluding as assumed in the threat model.

We evaluate two representative attack classes. The first is *gradient poisoning*, in which an adversarial client perturbs its update to bias global optimization. Concretely, if g_i denotes the original local gradient of client C_i, the malicious update is modeled as in Eq. (23)

$$\tilde{g}_i = \beta g_i + \delta, \tag{23}$$

where $\beta > 1$ controls gradient amplification and δ is an adversarial perturbation vector used to shift the update direction. In our experiments, β is selected from a moderate attack range to avoid trivially detectable manipulations while still degrading the training process.

Backdoor insertion is the second of the attacks. Malicious clients poison a portion of their local training data in this setting by inserting a prespecified trigger pattern and forcing the model to map triggered inputs to the target class chosen by the attacker. This attack is assessed under the standard objective of keeping high clean data performance whilst increasing misclassifications on trigger-injected inputs.

To measure robustness, three complementary metrics are reported in this paper. First, we use clean test predictive performance to assess the utility of the trained model under adversarial conditions, determined using AUC. Second, we assess the *attack success rate* (ASR), defined as the percentage of manipulated or triggered inputs that are classified into the attacker's target output. Third, we report the *performance degradation*, computed as the reduction in clean test performance relative to the benign setting. All of these attacks are employed consistently across RP-IoMT and baseline methods under identical data partitions, training rounds, and optimization settings to ensure a fair comparison. Attack parameters β and δ are held fixed across all independent runs and random seeds to ensure reproducibility. The reported AUC and ASR values correspond to averages over three runs, indicating stable behavior across different initializations, with standard deviations consistently below 0.5% and 1.2%, respectively.

Data Pre-Processing

The MIMIC-III and HiRID datasets have both irregularly sampled time series clinical measurements and static patient attributes, and pre-processing steps are defined earlier and later in the paper.

6.2 Experimental Analysis

This section presents a comprehensive empirical evaluation of RP-IoMT, designed to assess its security guarantees, computational efficiency, communication overhead, scalability, and deployability in real IoMT environments. Our evaluation spans five dimensions: (1) learning performance, (2) communication and cryptographic overhead, (3) MPC-based aggregation latency, (4) scalability across client populations, and (5) feasibility on resource-constrained medical IoT devices. Unless otherwise specified, complete experiments were conducted on a workstation with an Intel i7-12700K processor and 32 GB RAM, while a Raspberry Pi 4 (4 GB) was used for IoMT-device experiments.

Table 3 shows the adversarial evaluation for the configuration being used throughout our experiments. As shown, we consider a client-side threat model in which a fraction of participants may behave maliciously by manipulating their local updates. The evaluation includes both backdoor attack and gradient poisoning scenarios, under different adversarial behaviors, allowing us to assess the robustness of the proposed framework. The adversary strength is controlled by varying the proportion of compromised clients while maintaining consistent training conditions across all methods. To ensure reproducible and fair comparison, all experiments are performed using identical data partitions, optimization settings, and communication rounds. The selected evaluation metrics provide a comprehensive view of both resilience under attack and model utility. Metrics include AUC, ASR, and performance degradation.

Table 3: Adversarial evaluation setup used for robustness experiments.

Component	Specification
Adversary location	Client-side adversary
Compromised client ratio α	$\{0.1, 0.2, 0.3\}$
Attack types	Gradient poisoning, backdoor attack
Poisoning model	$\tilde{g}_i = \beta g_i + \delta$
Backdoor setting	Trigger-based target-label attack
Evaluation metrics	AUC, ASR, performance degradation
Comparison setting	Same partitions and rounds for all methods

6.2.1 Learning Performance

Convergence trajectories of RP-IoMT alongside baseline methods on healthcare datasets are presented in Fig. 3. RP-IoMT consistently achieves exceptional predictive performance, reaching an AUC of 84.7%, outperforming FedAvg, differential privacy-based FL, and SecAgg when applied on the MIMIC-III dataset. This rise is mainly attributed to two key components of the framework: ZKClip first enforces strict gradient norm constraints while preventing instability caused by gradient explosion. Second, the MPC-based trimmed mean aggregation merged with cosine similarity gating successfully suppresses malicious or anomalous updates. These mechanisms together lead to lesser variance across training rounds and more stable and faster convergence. Fig. 3 also illustrates the performance of RP-IoMT on the HiRID dataset, in addition to MIMIC-III. The HiRID curve demonstrates mild fluctuations and slightly lower performance as compared to MIMIC-III, which can be attributed to its increased data heterogeneity and higher temporal resolution. RP-IoMT claims a constant convergence trend, closely following the trajectory observed on MIMIC-III, despite these challenges. This proves that the proposed framework persists robustly even when applied to heterogeneous clinical data that is more complex. The results in Fig. 3 emphasize that RP-IoMT not only performs efficiently under standard settings but is also applicable across datasets with varying statistical properties. The consistent stability and performance gap observed over baseline methods on HiRID confirm the practical applicability and robustness of RP-IoMT in realistic IoMT environments. Over three independent runs with different random seeds, all reported results are averaged. Standard deviations across runs remain below 0.5% AUC for all datasets and methods, confirming that the perceived performance improvements of RP-IoMT across different initializations over baseline methods are statistically stable and consistent.

6.2.2 Results under Non-IID and Multi-Dataset Settings

In detail, the performance of RP-IoMT is studied under non-IID data distributions and from different datasets from healthcare sectors, such as MIMIC-III and HiRID, to further assess the ability to generalize and robustness. The non-IID setup shows significant instability in both SecAgg and FedAvg methods with higher variance and slower convergence over communication rounds. This is mainly because of heterogeneous data distributions and client drift, affecting the consistency of the aggregation. RP-IoMT, on the other hand, achieves stable convergence even in non-IID networks. This combination of ZKClip and robust aggregation in MPC (ZK+MPC) reduces the impact of malicious or incorrect updates, enhancing the system's resistance to statistical heterogeneity and a smoother learning process. The trends are consistent across datasets with RP-IoMT. The model achieves similar performance with only a marginal drop in AUC for the HiRID, which

is more complex because of variability and has greater temporal resolution. The proposed framework is more generalizable, as once the trajectory of convergence is determined, it does not change for any additional data.

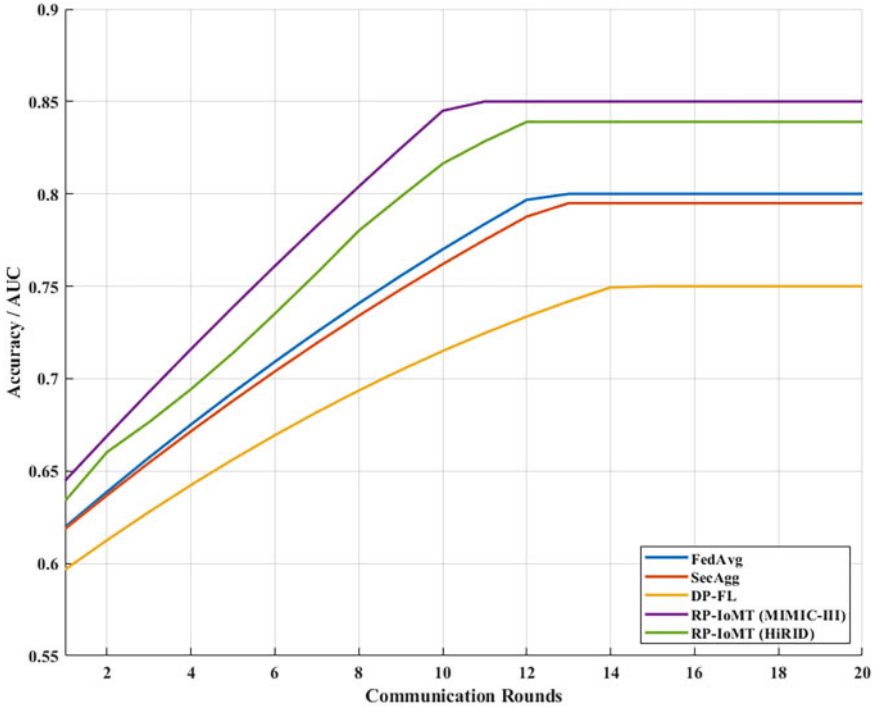

Figure 3: Learning performance.

In Table 4, further supports the robustness of RP-IoMT under non-IID settings using quantitative results. The degradation is much smaller with RP-IoMT, and all methods suffer performance degradation when the distribution changes from IID to non-IID. Baseline methods, like SecAgg and FedAvg, demonstrate major reductions as a result of sensitivity to information heterogeneity and drifting clients. However, RP-IoMT maintains high predictive accuracy on both datasets, MIMIC-III and HiRID, showing good statistical skew resistance. The results also show that the proposed framework generalizes well across datasets, with a marginal performance drop noticed on HiRID.

Table 4: Performance comparison under IID and Non-IID settings.

Method	IID (MIMIC-III)	Non-IID (MIMIC-III)	Non-IID (HiRID)
FedAvg	80.0 ± 0.4	77.5 ± 0.5	76.8 ± 0.5
SecAgg	79.5 ± 0.4	77.0 ± 0.4	76.3 ± 0.5
DP-FL	75.0 ± 0.5	72.5 ± 0.5	71.8 ± 0.6
RP-IoMT	85.0 ± 0.3	83.6 ± 0.3	83.0 ± 0.4

6.2.3 Robustness under Adversarial Clients

Baseline methods offer a clear cut in clean test performance and an ASR that is much higher for both types of backdoor attacks and gradient poisoning. The more the compromised clients break up, the more the deprivation is asserted. The other attacks are also more volatile than RP-IoMT overall attack strengths

evaluated. This is caused mainly by a combination of MPC-based robust aggregation (which reduces the impact of abnormal client contributions before the global update is applied) and ZKClip (which prevents unbounded update manipulation).

Robustness of RP-IoMT under adversarial conditions is further highlighted in quantitative results shown in Table 5. RP-IoMT maintains a significantly higher AUC with minimal degradation, while all baseline methods experience an evident decline in predictive performance when subjected to attacks. In addition, the ASR for RP-IoMT is substantially lower compared to other methods, indicating strong resistance to both backdoor insertion and gradient manipulation. Relatively small degradation in the performance shows that the proposed framework effectively mitigates the impact of malicious updates, even as the proportion of compromised clients increases. These findings confirm that RP-IoMT provides a reliable balance between security in adversarial IoMT environments and model accuracy.

Table 5: Robustness evaluation under adversarial attacks.

Method	AUC (Under Attack)	ASR (%)	Degradation (%)
FedAvg	76.5 ± 0.5	62.3 ± 1.2	3.5 ± 0.3
SecAgg	75.8 ± 0.5	58.9 ± 1.1	3.7 ± 0.3
DP-FL	72.4 ± 0.5	51.7 ± 1.3	2.6 ± 0.3
RP-IoMT	83.2 ± 0.4	21.5 ± 0.8	1.5 ± 0.2

6.2.4 Sensitivity Analysis of the Clipping Threshold

The clipping threshold τ directly affects both adversarial robustness and model convergence, since it establishes the maximum norm of each client update before secure aggregation. Many benign updates are excessively compressed if τ is chosen too small, which may reduce predictive performance and slow optimization. Conversely, the clipping mechanism becomes less effective at mitigating malicious or limiting abnormal gradients, thereby weakening robustness if τ is chosen too large. We performed a sensitivity analysis by varying τ over a predefined range while keeping all other experimental settings fixed to evaluate this effect. We measure convergence behavior, robustness, and final clean test AUC under adversarial conditions for each value of τ. The results show a clear trade-off: larger values accelerate optimization but permit greater influence from anomalous updates, whereas smaller values of τ improve resistance to gradient amplification but introduce slower convergence. RP-IoMT achieves stable convergence while maintaining strong robustness on the selected operating point, which corresponds to the region.

This performance is consistent with the role of clipping in robust federated optimization. In particular, τ appears as a regularizing bound on client updates, guaranteeing that no single participant could dominate the aggregation process. We select the value of τ based on the validation experiments, which provides the best overall balance between security and utility across attack settings and datasets. Table 6 further illustrates the consequence of the clipping threshold on the performance of RP-IoMT. As τ increases, model utility improves up to an intermediate range, where the best balance between robustness and accuracy is achieved. Beyond this region, the attack success rate increases noticeably, indicating that overly permissive clipping weakens the protection against adversarial update manipulation. When τ is too small, the framework becomes too suppressive and restricts useful client updates, leading to reduced predictive performance and slower convergence. These results rationalize the selected value of τ and confirm that the clipping threshold must be chosen carefully to balance security and optimization.

Table 6: Sensitivity of RP-IoMT to the clipping threshold τ.

τ	AUC (%)	ASR (%)	Convergence Trend
0.5	81.9 ± 0.4	16.8 ± 0.7	Slow, over-clipped
1.0	83.4 ± 0.3	18.9 ± 0.8	Stable
1.5	84.7 ± 0.3	21.5 ± 0.8	Best trade-off
2.0	84.5 ± 0.3	27.8 ± 1.0	Faster, less robust
2.5	83.9 ± 0.4	33.4 ± 1.1	Unstable under attack

6.2.5 Communication Overhead of Verifiable Privacy Mechanisms

The incorporation of cryptographic components, including Pedersen commitments, ZKClip proofs, and additive secret sharing, introduces additional communication overhead compared to standard FL. Fig. 4 illustrates this overhead across three representative model architectures. To provide a clearer interpretation, we define the baseline communication cost as the transmission of model updates of dimension d per client per round in standard FL. Under RP-IoMT, each client additionally transmits secret shares to n aggregation servers and a zero-knowledge proof π_i. As a result, the per-client communication cost can be expressed as in Eq. (24):

$$C_{\text{RP-IoMT}} = O(n \cdot d) + O(|\pi_i|) \tag{24}$$

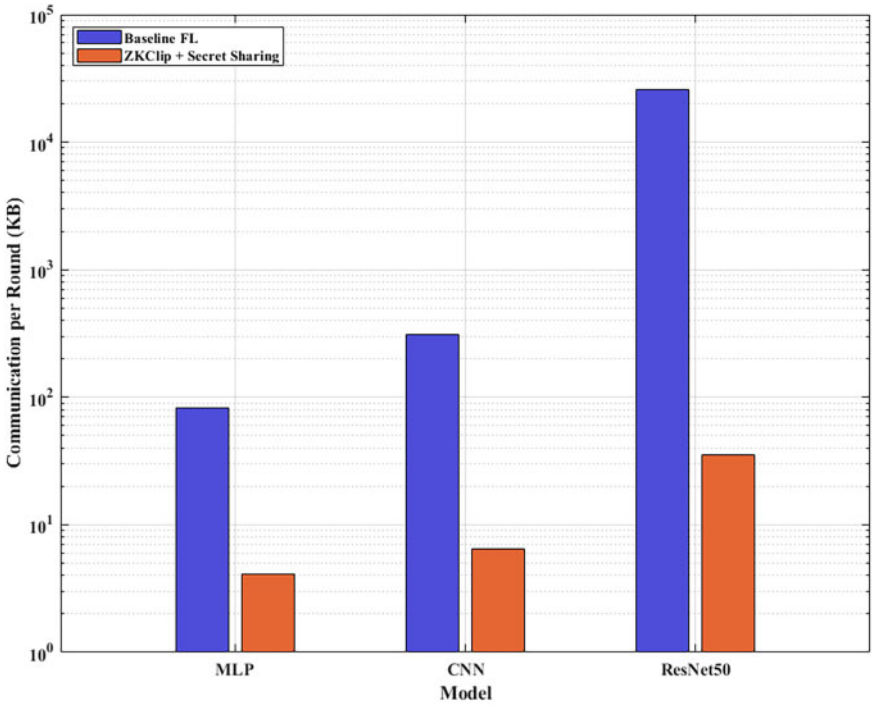

Figure 4: Communication overhead.

In practice, the additional overhead introduced by ZKClip remains moderate. For the MLP and CNN models, the proof size contributes approximately 4.1 and 6.4 KB per round, corresponding to an increase of approximately 2%–5% relative to the baseline FL communication cost. For larger models such as ResNet50,

the relative overhead decreases to below 0.2%, as the proof size grows much more slowly than the model dimension. This behavior is consistent with the use of Bulletproofs-style inner product arguments, where proof size scales logarithmically with respect to the constraint size. The dimensionality of the model and the number of aggregation servers still influenced the overall communication cost, which is important to note. Whereas to manage the overhead, evaluated settings performed well, and it would become more evident if deployed on a larger scale, an increase in client or server numbers.

6.2.6 Communication Complexity Analysis

RP-IoMT communication cost outperformed during each communication in terms of verification proofs and transmitting the secret shares. For a system having N clients and aggregation servers n, each client shares its clipped update using a (t, n)-threshold secret sharing scheme, resulting in $O(n)$ communication per client per round.

Additionally, each client transmits a zero-knowledge proof π_i along with its update. Thus, the total per-client communication cost can also be expressed as:

$$C_{\text{client}} = O(n \cdot d) + O(|\pi_i|), \tag{25}$$

where in Eq. (25), d is the gradient dimension and $|\pi_i|$ denotes the proof size.

At the system level, the total communication cost per round scales as shown in Eq. (26):

$$C_{\text{total}} = O(N \cdot n \cdot d), \tag{26}$$

which is linear in the number of clients. While this scaling introduces overhead for large N, the use of threshold-based aggregation and parallel server processing mitigates the impact in practice. For deployments exceeding 200 clients, hierarchical aggregation or client sampling can be incorporated to further improve scalability.

6.2.7 Latency of Secure MPC-Based Aggregation

RP-IoMT end-to-end latency of a single FL round is evaluated, as illustrated in Fig. 5. The latency includes client-side pre-processing, verification at the CC, and secure aggregation via MPC. Due to cryptographic operations, additional computational overhead is added by the RP-IoMT as compared to standard FL as indicated in the results. Particularly, the per-round total latency for the MLP model is approximately 97, 242 ms for the CNN model, and 610 ms for the ResNet50 model. An important part of this latency is the influential component of MPC-based secure aggregation, especially when comparing and multiplying, which are necessary for strong aggregation. In contrast, the CC verification step has an insignificant cost, whereas the verification steps and the proof generation in ZKClip make up a relatively small part of the total latency. These values are taken in a controlled single-machine environment and may not be an accurate comparison for different models. They primarily reflect computational overhead rather than end-to-end network latency. Additional delays due to distributed execution and network communication are expected in real-world IoMT deployments. The observed overhead is unlikely to be a limiting factor since FL rounds in healthcare applications typically occur at uneven time scales. Overall, the overhead remains moderate under the evaluated setting, but due to the security mechanisms of RP-IoMT, additional latency is incurred. The current evaluation is based on a process-level emulation framework in which clients and aggregation servers are implemented as separate processes on a single machine and communicate through inter-process message passing. Although this configuration accurately captures computational and protocol-related overhead, it does not entirely reproduce the network latency, bandwidth limitations,

and synchronization overhead encountered in distributed multi-machine environments. Consequently, the reported latency measurements should be regarded as lower-bound estimates of the end-to-end round duration in practical IoMT deployments.

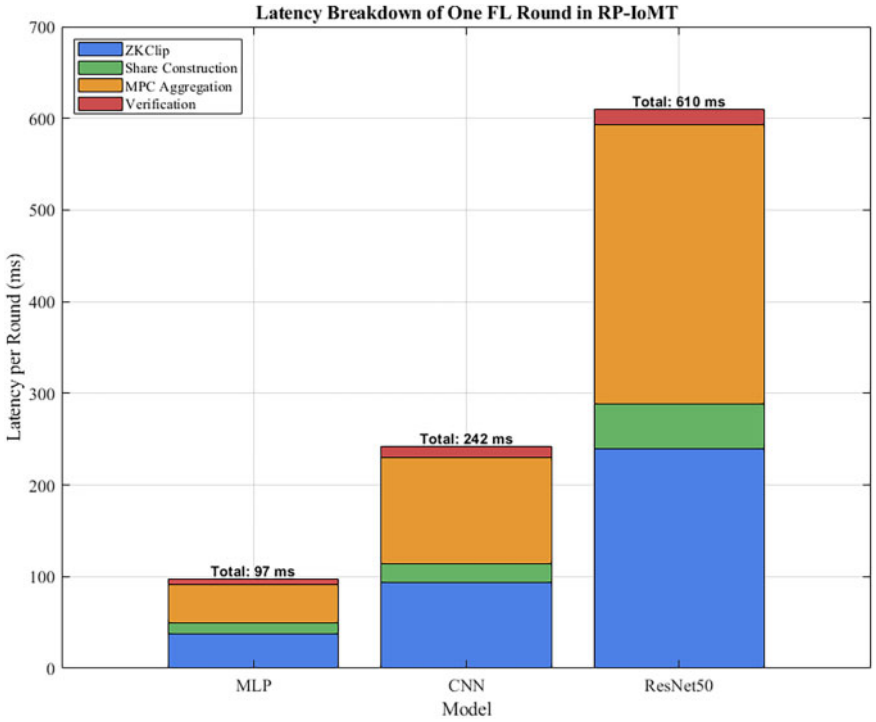

Figure 5: Latency of secure MPC-based aggregation.

6.2.8 Scalability to Large Client Populations

In dense IoMT environments, we evaluate the scalability of RP-IoMT by varying the number of participating clients from 10 to 200. Fig. 6 illustrates the corresponding impact on latency and model performance. The latency rises to around 110 ms when there are 10 clients and then rises linearly to 480 ms when there are 200 clients. This is a natural progression, since the generation, transmission, and aggregation using secret-sharing is a cost that increases in proportion to the number of clients involved. Specifically, the growth is mostly due to the growing number of client updates and the secure computation cost associated with them. Although the latency has increased, predictive performance is stable for all scales. The observed degradation in accuracy is limited to approximately 1.9% when scaling from 10 to 200 clients. This means that in bigger deployments, the powerful aggregation mechanism can effectively prevent the impact of noisy or potentially malicious updates and preserve the quality of the model. A point to be noted is that the reported latency times have been obtained under controlled conditions, and mostly, it represents the computational overhead. In real-world IoMT deployments, other network-induced delays could impact the scalability. However, the findings indicate that the solution has the potential to be adopted in moderately sized healthcare settings like hospital systems with many devices without a substantial drop in predictive accuracy.

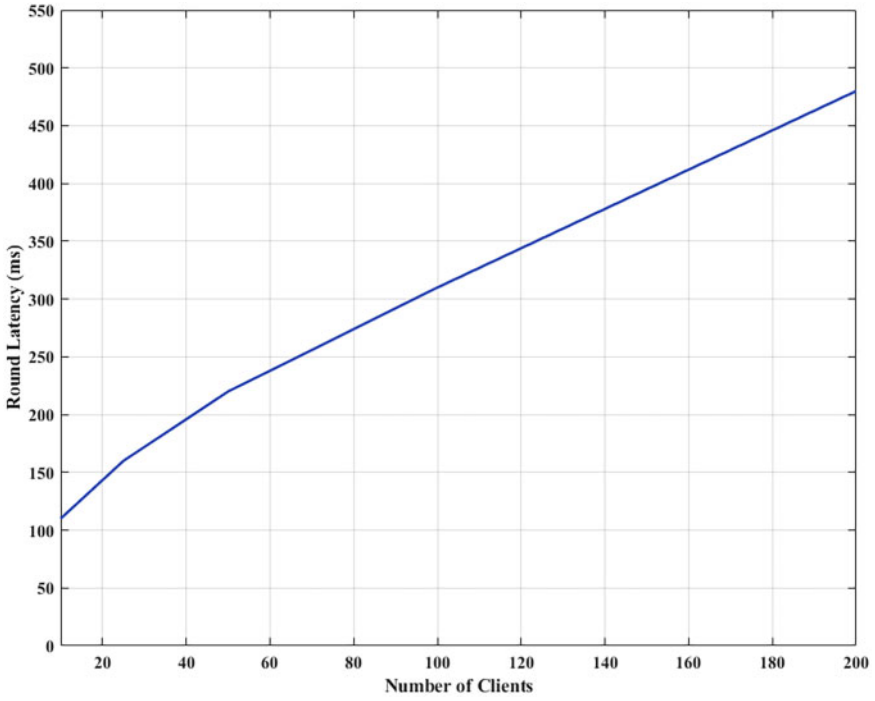

Figure 6: Scalability to large client populations.

6.2.9 Feasibility on IoMT-Grade Hardware

Since IoMT implementations often comprise low-power devices, we evaluate RP-IoMT on a Raspberry Pi 4. Client-side overhead is shown in Fig. 7. Only 22 ms is required for the generation of the ZKClip proof on the Raspberry Pi, and it adds 8 ms for the construction of the secret share. Less than 100 ms is utilized, which includes a single epoch of local training and computation of the client per round. Memory consumption stays below 200 MB, which is well within the capabilities of Raspberry-class gateways and in-home medical monitoring nodes. Fig. 7 shows that training dominates client runtime, whereas cryptographic functions contribute only a small fraction. This indicates that the update mechanisms implemented by RP-IoMT are light enough to run on real IoMT devices and possess privacy and verifiability properties. Overall, the findings indicate that RP-IoMT has high security and robustness properties at low overhead. The communication cost is still low, MPC-based aggregation is efficient, the system is smoothly expanded by adding more clients, and all the operations that are performed on the clients can be executed easily on devices suitable for IoMT applications. The results demonstrate that RP-IoMT is an implementable and secure FL framework designed for real-life medical IoT infrastructures.

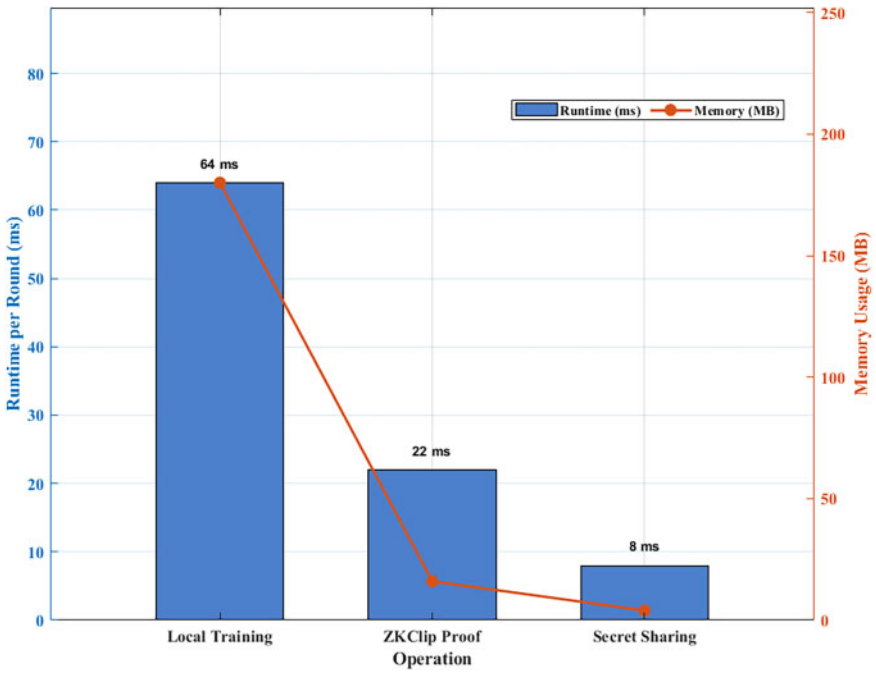

Figure 7: Feasibility on IoMT-Grade hardware.

7 Comparison with State-of-the-Art FL Security Frameworks

To position RP-IoMT within the landscape of modern secure FL, we compare it against three recent frameworks: RPEA, Octopus, and VMFL. Rather than reporting raw values for each metric, we summarize the trade-offs in a normalized radar plot, shown in Fig. 8. The figure is the result of five aspects of the FL security framework, namely, latency, robustness against adversarial clients, verifiability, communication cost, and overall security level. All metrics are scaled to a maximum value of one, with higher values representing better performance. When it comes to latency, Octopus gets the best score thanks to its lightweight compression and masking methods; RP-IoMT and RPEA are in the middle, and VMFL is beaten due to its high cost of multi-round verification. The communication cost is similar: With all of the Octopus has the lowest cost, RPEA has a reasonable cost with just minimal overhead from ZKClip and MPC, while VMFL incurs the highest load with proof traffic and extra commitment. These two axes show RP-IoMT to be competitive, but not the lowest-scoring, demonstrating the privacy performance efficiency trade-off.

The benefits of RP-IoMT become obvious in the robustness, verifiability, and security aspects. The robustness axis measures the ability to withstand Byzantine clients and poisoning attacks; RP-IoMT achieves the highest score, with the ability to provide full protection against Byzantine clients and poisoning attacks, thanks to cryptographically enforced gradient clipping and robust aggregation through MPC, while RPEA provides only partial protection against strong adversaries, and Octopus is more vulnerable to Byzantine clients and poisoning attacks. As for verifiability, both VMFL and RP-IoMT achieve a high score, but in different aspects of the protocol: VMFL is good at verifying aggregation only on the server side, while RP-IoMT also verifies aggregation on the client side using zero-knowledge proofs, resulting in a slightly higher overall score. Finally, on the overall security axis that covers multi-server trust, collusion, and end-to-end auditability, RP-IoMT beats all baselines as it integrates multi-server MPC, ZKClip, and proof-carrying updates in one framework.

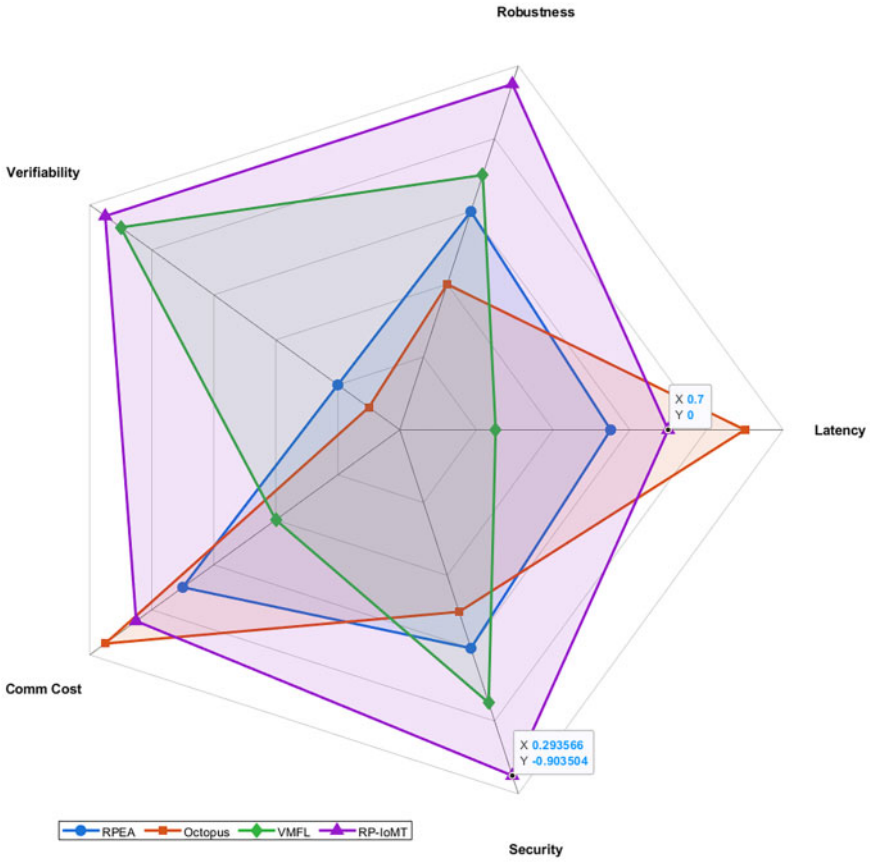

Figure 8: Normalized radar plot comparing RP-IoMT with three recent secure FL frameworks (RPEA, Octopus, and VMFL) across five dimensions: latency, robustness, verifiability, communication cost, and overall security. Higher values indicate better performance. RP-IoMT achieves the most balanced and comprehensive security profile while maintaining competitive efficiency.

Taken together, Fig. 8 shows that while Octopus and RPEA are attractive from a pure efficiency perspective, and VMFL provides strong but computationally expensive verifiability, RP-IoMT achieves the most balanced overall profile. It sustains adequate performance for operational deployment while ensuring a high level of integration that guarantees robustness, security, and verifiability. This supports the assertion that RP-IoMT is a next-generation FL security framework designed for highly sensitive and adversarial IoMT environments.

Comparative evaluation of the security assurances provided by RP-IoMT relative to the recent three secure FL frameworks is summarized in Table 7. It is obvious from the results that RP-IoMT achieves a high level and the best complete security posture as compared with the other systems. Both Octopus and RPEA, in the context of verifiable aggregation, are unable to detect inappropriate behavior by the aggregation server, whereas server-side verification is provided by the VMFL only. In contrast, RPIoMT further enhances this capability, ensuring correctness even under limited server compromise, by enabling verifiable aggregation through a multi-server MPC architecture. The evaluation also shows evident differences in robustness. The low robustness of Octopus is because it uses compressed and statistically filtered gradients, and moderate protection is given by RPEA and VMFL.

Table 7: Security feature comparison of recent FL frameworks.

Feature	RPEA	Octopus	VMFL	RP-IoMT
Verifiable aggregation	×	×	✓	✓ (Multi-server MPC)
Byzantine robustness	Moderate	Low	Moderate	High
Multi-round verifiability	×	×	✓	✓
Client update verifiability	×	×	×	(Zero-knowledge)
Scalability (≤200 clients)	✓	✓	Moderate	✓

Resisting poisoning and Byzantine attacks, RP-IoMT implements the most robust framework by using both zero-knowledge enforced gradient clipping and MPC-based robust aggregation. VMFL and RP-IoMT are the only methods that support multi-round verifiability, which is important for long training sessions. This places RP-IoMT as one of the rare frameworks that will be able to identify misbehavior that can build up or come up randomly every communication round. Another big difference is the verifiability of the client update. No baseline schemes contain any means to guarantee a well-formed or norm-bound upload of gradients. Only the RP-IoMT framework brings provable client-side correctness with ZKP to formally bound the validity of each update, without revealing sensitive gradient information.

Finally, the scalability row demonstrates that RP-IoMT maintains high performance as the number of clients grows, matching the scalability of RPEA and Octopus and outperforming VMFL, whose verification procedures become increasingly expensive with larger populations. Overall, the results in Table 7 confirm that RP-IoMT provides the broadest and most balanced security guarantees, combining robustness, verifiability, and scalability while maintaining efficiency appropriate for large-scale IoMT deployments. This positions RP-IoMT as a substantially stronger and more reliable FL security framework than existing 2025-era alternatives.

To make a more objective comparison of the proposed framework, we supplement the qualitative analysis with a detailed quantitative comparison with representative FL approaches. This is because the architectural and security benefits of RP-IoMT are emphasized in the previous sections, but it should also be assessed for tangible benefits in terms of its performance. We compare RP-IoMT to popular baselines like FedAvg, DP-FL, Octopus, RPEA, and VMFL on various metrics, including predictive accuracy, communication overhead, computational latency, and adversarial robustness. This comparison is done under the same experimental circumstances to make it fair and reproducible, and allows a better understanding of the compromise between efficiency, accuracy, and security.

To complement the qualitative comparison, we provide a quantitative evaluation across key performance metrics. Table 8 summarizes the results for RP-IoMT and representative baselines under consistent experimental settings. The proposed framework achieves the highest predictive performance with an AUC of 84.7%, outperforming all baselines. In terms of communication cost, RP-IoMT introduces only a modest overhead (approximately 1.05× of standard FL), which remains competitive with existing secure aggregation methods. From a computational perspective, RP-IoMT maintains moderate latency compared to alternatives, significantly lower than VMFL, while slightly higher than lightweight schemes such as Octopus. Importantly, RP-IoMT demonstrates substantially improved robustness, as reflected by a significantly reduced attack success rate (ASR), indicating strong resistance against adversarial manipulation. These results confirm that RP-IoMT achieves a balanced trade-off between accuracy, efficiency, and security. The reported latency values are measured in a controlled environment using a high-performance CPU to ensure consistent comparison across methods. These results reflect the computational overhead of the proposed framework

rather than end-to-end network latency. In practical deployments, additional latency may arise due to network communication, particularly in distributed IoMT settings. However, the relative performance trends are expected to remain consistent. Evaluating RP-IoMT under real-world network conditions is an important direction for future work.

Table 8: Quantitative comparison of RP-IoMT with state-of-the-art FL frameworks.

Framework	AUC (%)	Comm. Overhead	Latency (ms)	ASR (%)
FedAvg	79.2	1.00×	80	68.5
DP-FL	74.8	1.20×	95	55.2
Octopus	81.3	0.85×	70	42.6
RPEA	82.5	1.10×	120	38.4
VMFL	83.1	1.35×	180	34.7
RP-IoMT	84.7	1.05×	97	18.9

8 Conclusion

In this paper, we presented RP-IoMT, a verifiable, privacy-preserving, and robust FL framework designed for the stringent security requirements of IoMT environments. RP-IoMT integrates zero-knowledge-based client verification, multi-server MPC aggregation, and robust defense mechanisms into a unified architecture.

The key contributions and findings of this work are summarized as follows:

- We proposed RP-IoMT, a unified framework that combines zero-knowledge proof-based client verification (ZKClip), secure multi-party computation (MPC), and robust aggregation for IoMT systems.
- We provide a formal security analysis that demonstrates privacy, correctness, and verifiability under standard cryptographic assumptions.
- We showed that RP-IoMT achieves strong predictive performance across multiple healthcare datasets while maintaining stability under both IID and non-IID settings.
- We demonstrated robustness against adversarial attacks, including poisoning and backdoor scenarios, with significantly reduced attack success rates compared to baseline methods.
- We analyzed system efficiency and showed that the framework introduces only moderate communication and computational overhead, with latency remaining within practical limits for IoMT applications.
- We evaluated scalability and confirmed that RP-IoMT maintains stable performance even as the number of participating clients increases.
- We compared RP-IoMT with state-of-the-art FL security frameworks (e.g., Octopus, RPEA, VMFL), highlighting its ability to provide a balanced combination of privacy, robustness, and verifiability.

Despite achieving strong privacy, robustness, and verifiability, RP-IoMT was evaluated mainly in controlled environments and may face additional challenges in real-world large-scale IoMT deployments with heterogeneous devices and network constraints. Moreover, several directions remain for future work. These include extending the framework to multimodal IoMT data, improving the efficiency of zero-knowledge proofs, exploring hybrid trust models such as trusted execution environments, and integrating differential privacy to provide quantifiable privacy guarantees. Overall, RP-IoMT establishes a practical and secure foundation for FL in next-generation healthcare systems.

Acknowledgement: This work was supported in part by the Commonwealth Cyber Initiative, an investment in the advancement of cyber R&D, innovation, and workforce development. For more information about CCI, visit cyberinitiatives.org.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: M. Saad Bin Ilyas, Sohail Masood Bhatti; data collection: Ghazanfar Latif; analysis and interpretation of results: M. Saad Bin Ilyas, Sohail Masood Bhatti, Arfan Jaffar, Ghazanfar Latif; draft manuscript preparation: Sherif Abdelhamid, M. Saad Bin Ilyas. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The MIMIC-III Clinical Database is available on PhysioNet doi:10.13026/C2XW26. HiRID, a high-time-resolution ICU dataset, is available at PhysioNet. RRID: SCR_007345. doi:10.13026/nkwc-js72.

Ethics Approval: This study used data from the MIMIC-III and HiRID intensive care databases. Both datasets contain de-identified patient information and are publicly accessible to qualified researchers under data use agreements.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Grand View Research. Internet of Medical Things (IoMT) market report. San Francisco, CA, USA; 2023 [cited 2025 Oct 12]. Available from: https://www.grandviewresearch.com/industry-analysis/internet-of-medical-things-iomt-market-report.
2. Hireche R, Mansouri H, Pathan ASK. Security and privacy management in Internet of Medical Things (IoMT): a synthesis. J Cybersecur Privacy. 2022;2(3):640–61. doi:10.3390/jcp2030033.
3. Rahman MA, Hossain MS, Islam MS, Alrajeh NA, Muhammad G. Secure and provenance enhanced internet of health things framework: a blockchain managed federated learning approach. IEEE Access. 2020;8:205071–87. doi:10.1109/access.2020.3037474.
4. Apicella A, Isgrò F, Prevete R. Don't push the button! Exploring data leakage risks in machine learning and transfer learning. Artif Intell Rev. 2025;58(11):1–58. doi:10.1007/s10462-025-11326-3.
5. Mohammed S, Malhotra N. Ethical and regulatory challenges in machine learning-based healthcare systems: a review of implementation barriers and future directions. BenchCounc Trans Benchmarks Stand Eval. 2025;5(1):100215.
6. Semmadi A, Bahhou T. Federated learning in internet of medical things (IoMT) healthcare applications. Ouargla, Algeria: Kasdi Merbah University; 2024 [cited 2025 Oct 15]. Available from: https://dspace.univ-ouargla.dz/jspui/handle/123456789/37350.
7. Khan MA, Saudagar AKJ, Yaqoob MM, Nazir M, Yousafzai A, Khaliq uz Zaman S, et al. Federated learning for heart disease detection and classification in edge enabled IoMT-based healthcare: taxonomy, challenges, and opportunities. Computing. 2025;107(11):1–31. doi:10.1007/s00607-025-01572-2.
8. Ding W, Xiao Y, Yan Z, Chen C, Cai Y, Jing X. Octopus: a robust and privacy-preserving scheme for compressed gradients in federated learning. IEEE Trans Dependable Secure Comput. 2026;23(1):1560–75.
9. Xie Y, Fang M, Gong NZ. Fedredefense: defending against model poisoning attacks for federated learning using model update reconstruction error. In: Proceedings of the 41st International Conference on Machine Learning; 2024 Jul 21–27; Vienna, Austria.
10. Feng J, Lai Y, Sun H, Ren B. SADBA: self-adaptive distributed backdoor attack against federated learning. Proc AAAI Conf Artif Intel. 2025;39:16568–76.
11. Nielsen C, Wilms M, Forkert ND. A novel gradient inversion attack framework to investigate privacy vulnerabilities during retinal image-based federated learning. Med Image Anal. 2026;107(Pt B)5:103807. doi:10.1016/j.media.2025.103807.

12. Mahato GK, Banerjee A, Chakraborty SK, Gao XZ. Privacy preserving verifiable federated learning scheme using blockchain and homomorphic encryption. Appl Soft Comput. 2024;167(3):112405. doi:10.1016/j.asoc.2024.112405.
13. Yu L, Li X. Dynamic optimization method for differential privacy parameters based on data sensitivity in federated learning. J Adv Comput Syst. 2025;5(6):1–13. doi:10.20944/preprints202506.2188.v1.
14. McMahan B, Moore E, Ramage D, Hampson S, Arcas BA. Communication-efficient learning of deep networks from decentralized data. In: Proceedings of the 20th International Conference on Artificial Intelligence and Statistics; 2017 Apr 20–22; Fort Lauderdale, FL, USA. p. 1273–82.
15. Yang Q, Liu Y, Chen T, Tong Y. Federated machine learning: concept and applications. ACM Trans Intell Syst Technol. 2019;10(2):1–19.
16. Begum K, Mozumder MAI, Joo MI, Kim HC. BFLIDS: blockchain-driven federated learning for intrusion detection in IoMT networks. Sensors. 2024;24(14):4591.
17. Zhu L, Liu Z, Han S. Deep leakage from gradients. In: Proceedings of the 33rd International Conference on Advances in Neural Information Processing Systems; 2019 Dec 8–14; Vancouver, BC, Canada. p. 14774–84.
18. Guan H, Yap PT, Bozoki A, Liu M. Federated learning for medical image analysis: a survey. Pattern Recognit. 2024;151(3):110424. doi:10.1016/j.patcog.2024.110424.
19. Hu T, Chen Q, Hu Y, Hou S, Yan H, Yi P, et al. Efficient and privacy-preserving network intrusion detection based on federated learning in SDN-enabled IIoT network. IEEE Internet Things J. 2025;12(20):41904–23. doi:10.1109/jiot.2025.3591598.
20. Liu Z, Gao P, Wang B. Robust privacy-enhanced aggregation scheme for federated learning in industrial Internet of Things. IEEE Internet Things J. 2025;12(21):45517–32. doi:10.1109/jiot.2025.3601856.
21. Reddi S, Rao PM, Saraswathi P, Jangirala S, Das AK, Jamal SS, et al. Privacy-preserving electronic medical record sharing for IoT-enabled healthcare system using fully homomorphic encryption, IOTA, and masked authenticated messaging. IEEE Trans Ind Inform. 2024;20(9):10802–13. doi:10.1109/tii.2024.3397343.
22. Farooqi SA, Rahman AA, Saad A. Advanced privacy-utility optimization techniques in federated learning with differential privacy for IoMT—a review. Intl J Interactive Mobile Technol. 2025;19(19):134–50. doi:10.3991/ijim.v19i19.57619.
23. Annappa B, Hegde S, Abhijit CS, Ambesange S, et al. Fedcure: a heterogeneity-aware personalized federated learning framework for intelligent healthcare applications in IoMT environments. IEEE Access. 2024;12:15867–83.
24. Mehra A, Singh G, Badotra S. Federated learning for internet of medical things (IoMT): a secure and scalable approach. In: Proceedings of the 2025 12th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO); 2025 Sep 18–19; Noida, India. p. 1–7.
25. Wani RUZ, Can O. Federated learning for secure and privacy-aware internet of medical things: taxonomy, emerging applications, open challenges, and future directions. Concurr Comput. 2025;37(27–28):e70432.
26. Wani RUZ, Can O. FED-EHR: a privacy-preserving federated learning framework for decentralized healthcare analytics. Electronics. 2025;14(16):3261.
27. Javed MS, Hennache A, Imran M, Khan MK. AI-driven blockchain and federated learning for secure electronic health records sharing. Electronics. 2025;14(23):4774. doi:10.3390/electronics14234774.
28. Myrzashova R, Alsamhi SH, Shvetsov AV, Hawbani A, Wei X. Blockchain meets federated learning in healthcare: a systematic review with challenges and opportunities. IEEE Internet Things J. 2023;10(16):14418–37.
29. Bhardwaj T, Sumangali K. An explainable federated blockchain framework with privacy-preserving ai optimization for securing healthcare data. Sci Rep. 2025;15(1):21799. doi:10.1038/s41598-025-04083-4.
30. Tawfik AM, Al-Ahwal A, Eldien AST, Zayed HH. PriCollabAnalysis: privacy-preserving healthcare collaborative analysis on blockchain using homomorphic encryption and secure multiparty computation. Cluster Comput. 2025;28(3):191. doi:10.1007/s10586-024-04928-z.
31. Park S, Choi W. Byzantine fault tolerant distributed stochastic gradient descent based on over-the-air computation. IEEE Trans Commun. 2022;70(5):3204–19. doi:10.1109/tcomm.2022.3162576.
32. Xing Z, Zhang Z, Zhang Z, Li Z, Li M, Liu J, et al. Zero-knowledge proof-based verifiable decentralized machine learning in communication network: a comprehensive survey. IEEE Commun Surv Tutorials. 2026;28(3):985–1024. doi:10.1109/comst.2025.3561657.

33. Xu G, Li H, Liu S, Yang K, Lin X. VerifyNet: secure and verifiable federated learning. IEEE Trans Inf Forensics Secur. 2019;15:911–26.
34. Li K, Feng X, Guo Z, Cui K, Wang C, Li K. VMFL: a verifiable multi-round aggregation scheme for federated learning in VANETs. IEEE Internet Things J. 2025;12(20):42392–406.
35. Shamir A. How to share a secret. Commun ACM. 1979;22(11):612–3. doi:10.1145/359168.359176.
36. Abadi M, Chu A, Goodfellow I, McMahan HB, Mironov I, Talwar K, et al. Deep learning with differential privacy. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security; 2016 Oct 24–28; Vienna, Austria. p. 308–18.
37. Wang Z, Dong N, Sun J, Knottenbelt W, Guo Y. zkFL: zero-knowledge proof-based gradient aggregation for federated learning. IEEE Trans Big Data. 2024;11(2):447–60. doi:10.1109/tbdata.2024.3403370.
38. Johnson AE, Mark RG. Real-time mortality prediction in the intensive care unit. AMIA Annu Symp Proc. 2018;2017:994–1003.
39. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016 Jun 27–30; Las Vegas, NV, USA. p. 770–8.