International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(21s) 2026 pp. 235-258
https://doi.org/10.70917/ijcisim-2026-5292
Article

Federated Self-Supervised Learning for Privacy-Preserving Clinical Diagnostics in Distributed Healthcare Systems

Abinaya V.[1], S. Swathi[2], G. Sharmila[3], S. Karpaga Iswarya[4], P. Priyadharsini[5], Harini A.[6]

[1]Assistant Professor, Department of Artificial Intelligence and Data Science,
Karpagam Institute of Technology, Coimbatore – 641105, Tamil Nadu, India.
Email: abinaya111234@gmail.com
[2]Assistant Professor, Department of Artificial Intelligence and Data Science,
Panimalar Engineering College, Tamil Nadu, India.
Email: swathishankar6766@gmail.com
[3]Assistant Professor, Department of Artificial Intelligence and Machine Learning,
Sri Eshwar College of Engineering, Coimbatore – 641202, Tamil Nadu, India.
Email: charm icse@gmail.com
[4] Assistant Professor, Department of Computer Science and Science,
Nehru Institute of Technology, Tamil Nadu, India.
Email: nitkarpagaiswaryacse@nehrucolleges.com
[5]Assistant Professor, Department of Information Technology,
Nehru Institute of Engineering and Technology, Tamil Nadu, India.
Email: priyapalanisamy1410@gmail.com
[6]Assistant Professor, Department of Computer Science and Engineering,
Nehru Institute of Engineering and Technology, Tamil Nadu, India.
Email: aharini1608@gmail.com

Abstract: The increasing adoption of distributed digital healthcare systems has enabled collaborative artificial intelligence (AI)-based clinical diagnostics across hospitals and medical institutions. However, conventional centralized learning requires sensitive patient information to be transferred to a common server, introducing substantial concerns regarding data privacy, security, institutional governance, and regulatory compliance. Furthermore, obtaining sufficiently large and accurately annotated clinical datasets remains challenging because medical annotation requires substantial expertise and resources. To address these limitations, this study proposes a Federated Self-Supervised Learning (FSSL) framework for privacy-preserving clinical diagnostics in distributed healthcare systems. The proposed architecture combines self-supervised representation learning with federated optimization, enabling participating healthcare institutions to learn informative clinical representations from locally available unlabeled data without transferring raw patient records. Each participating client performs self-supervised pretraining followed by task-specific local optimization, while only protected model updates are communicated to the aggregation server. A privacy-aware aggregation mechanism is incorporated to reduce the exposure of institution-specific information during collaborative training. The framework further addresses heterogeneous and non-independent and identically distributed (non-IID) clinical data through adaptive client aggregation and representation alignment. Experimental evaluation demonstrates that the proposed FSSL framework achieves a diagnostic accuracy of 96.18%, sensitivity of 95.47%, specificity of 96.72%, precision of 95.83%, F1-score of 95.65%, and area under the receiver operating characteristic curve (AUC) of 0.981. Compared with conventional federated supervised learning, the proposed method provides approximately 4.6% improvement in diagnostic accuracy and 5.1% improvement in F1-score, while substantially reducing dependence on labeled clinical samples. The results indicate that self-supervised representation learning can improve the effectiveness of federated clinical diagnostics under decentralized and heterogeneous healthcare environments. The proposed framework provides a scalable foundation for privacy-conscious

collaborative clinical AI while keeping patient data within the originating healthcare institution and is intended to support, rather than replace, professional clinical decision-making.

Keywords: Federated Learning; Self-Supervised Learning; Privacy-Preserving Healthcare; Clinical Diagnostics; Distributed Healthcare Systems; Medical Artificial Intelligence; Non-IID Data; Representation Learning; Federated Aggregation; Clinical Decision Support

1. Introduction

The rapid digital transformation of healthcare has resulted in the generation of large volumes of heterogeneous clinical information across hospitals, diagnostic centers, laboratories, and other healthcare institutions. Electronic health records, medical images, physiological signals, laboratory measurements, and clinical observations provide valuable information for developing artificial intelligence (AI)-assisted diagnostic systems. Machine-learning and deep-learning models have demonstrated considerable potential for disease identification, risk assessment, medical-image interpretation, and clinical decision support [1]. However, the effectiveness of these approaches generally depends on access to sufficiently large and representative datasets, which are often distributed across independent healthcare organizations.

Conventional centralized machine learning requires participating institutions to transfer their clinical datasets to a central repository before model development. Such data consolidation can introduce significant privacy, security, ownership, governance, and regulatory concerns because healthcare records frequently contain highly sensitive patient information [2]. Institutional policies and data-protection regulations may further restrict the direct exchange of identifiable clinical information. Consequently, valuable datasets often remain isolated within individual healthcare organizations, creating data silos and limiting opportunities to develop models that represent diverse patient populations.

Federated Learning (FL) has emerged as an important decentralized learning paradigm for addressing this challenge. Instead of collecting raw patient information at a centralized location, FL enables multiple participating healthcare institutions to train a shared model while retaining their clinical data locally [3]. Each institution performs model optimization using its own dataset and communicates model parameters or updates to a coordinating server. The server subsequently aggregates the received information and redistributes an updated global model to the participating clients. The fundamental federated objective can be represented as

$$\min_{\theta} F(\theta) = \sum_{k=1}^{K} \frac{n_k}{N} F_k(\theta),$$

where K represents the number of participating healthcare institutions, n_k denotes the number of local observations available at client k, $N = \sum_{k=1}^{K} n_k$, and $F_k(\theta)$ represents the local optimization objective. This distributed strategy enables collaborative model development without directly transferring the underlying patient records [4].

Despite this advantage, conventional federated supervised learning continues to depend heavily on accurately labeled clinical data. Medical data annotation is particularly expensive because it frequently requires experienced radiologists, cardiologists, pathologists, neurologists, or other domain specialists. Large quantities of unlabeled clinical data may therefore be available within healthcare institutions, whereas only a relatively small proportion has reliable diagnostic annotations [5]. This imbalance between available data and labeled data can restrict the performance and scalability of supervised federated models.

Self-Supervised Learning (SSL) provides an alternative strategy by learning informative representations from unlabeled data through automatically constructed learning objectives. Rather than requiring manually assigned disease labels during representation learning, SSL extracts supervisory information from the underlying structure of the data

itself [6]. Contrastive learning, masked representation modeling, reconstruction-based learning, and predictive pretext tasks have consequently gained increasing attention in medical AI. Once useful representations have been learned from large quantities of unlabeled clinical information, a smaller labeled dataset can be employed for task-specific fine-tuning. This characteristic makes SSL particularly relevant to healthcare environments in which unlabeled observations substantially exceed expert-annotated records.

The integration of federated learning and self-supervised learning provides an attractive solution for distributed healthcare intelligence. In a **Federated Self-Supervised Learning (FSSL)** architecture, each participating institution can exploit its locally available unlabeled clinical data to learn meaningful feature representations while avoiding direct raw-data exchange [7]. Local representations or model updates are periodically aggregated to construct a collaborative global representation model. Subsequently, the learned representations can be adapted for downstream diagnostic tasks using the limited labeled information available at individual institutions. This approach simultaneously addresses two fundamental barriers to medical AI development: decentralized sensitive data and limited expert annotations.

However, implementing FSSL in practical healthcare environments introduces several challenges. Clinical data collected at different institutions are rarely independent and identically distributed (IID). Hospitals may serve different demographic groups, employ different medical devices, follow different acquisition protocols, and exhibit different disease prevalence and clinical practices. Therefore, the local distribution $P_k(X, Y)$ at one institution may differ substantially from that observed at another institution [8]. Such non-IID distributions can cause inconsistent local optimization directions, client drift, unstable convergence, and degradation of the resulting global diagnostic model.

Privacy is another critical consideration. Although federated learning avoids direct transmission of raw patient records, it should not automatically be regarded as providing complete privacy. Model parameters and gradients may contain information correlated with local training samples, creating potential exposure through inference or reconstruction attacks [9]. Consequently, privacy-conscious federated healthcare systems can benefit from mechanisms such as secure aggregation, differential privacy, gradient protection, restricted update sharing, and appropriate access controls. These mechanisms should be integrated carefully because excessive perturbation or protection can reduce diagnostic utility.

Communication overhead also affects the scalability of federated healthcare networks. Repeated transmission of high-dimensional model parameters among multiple hospitals can consume substantial network resources, particularly when deep-learning architectures contain millions of parameters. Differences in computational capability and network connectivity can further produce stragglers and irregular client participation. Effective federated clinical systems therefore require suitable client selection, local training schedules, communication-efficient updates, and aggregation strategies that account for institutional heterogeneity [10].

Motivated by these challenges, this study proposes a **Federated Self-Supervised Learning framework for Privacy-Preserving Clinical Diagnostics in Distributed Healthcare Systems**. The proposed framework enables participating healthcare institutions to perform self-supervised representation learning using locally available unlabeled clinical data, followed by task-specific diagnostic optimization using available labeled samples. A central coordinating server aggregates protected client updates without requiring the underlying patient records. Adaptive aggregation and representation alignment are incorporated to reduce performance degradation associated with heterogeneous non-IID clinical distributions.

The proposed framework is designed around four complementary objectives: **data decentralization, label efficiency, diagnostic performance, and privacy-conscious collaborative learning**. Self-supervised pretraining reduces dependence on large expert-annotated datasets, while federated optimization enables institutions to benefit from collaborative learning without creating a centralized patient-data repository. Privacy-aware update handling reduces unnecessary exposure of institution-specific training information, and adaptive aggregation improves the contribution of heterogeneous clients to the global diagnostic model.

The principal contributions of this study are summarized as follows:

1. A unified **Federated Self-Supervised Learning (FSSL)** architecture is developed for collaborative clinical diagnostics across geographically and administratively distributed healthcare institutions without transferring raw patient records.
2. A self-supervised representation-learning mechanism is incorporated to exploit large quantities of locally available unlabeled clinical data and reduce dependence on expensive expert annotations.
3. An adaptive federated aggregation and representation-alignment strategy is introduced to improve learning under heterogeneous and non-IID institutional data distributions.
4. Privacy-aware model-update handling is incorporated into the federated communication process to strengthen the protection of institution-specific information during collaborative learning.
5. The proposed framework is evaluated using diagnostic accuracy, sensitivity, specificity, precision, F1-score, and ROC-AUC together with comparative and ablation analyses.

Under the proposed experimental configuration, the complete FSSL framework achieves **96.18% accuracy, 95.47% sensitivity, 96.72% specificity, 95.83% precision, 95.65% F1-score, and an AUC of 0.981**. These results indicate the potential of combining federated optimization with self-supervised representation learning for distributed clinical AI. Nevertheless, the framework is intended as a **clinical decision-support methodology rather than an autonomous diagnostic system**, and practical deployment requires external validation, privacy auditing, regulatory assessment, and prospective evaluation in real multi-institutional healthcare environments.

2. Related Work

Federated learning (FL) has emerged as an important paradigm for collaborative healthcare intelligence because it enables multiple medical institutions to jointly train predictive models without directly exchanging raw patient records. Existing healthcare applications cover medical imaging, electronic health records, disease prediction, physiological signals, and clinical decision support. However, systematic evidence indicates that much of the literature remains at the proof-of-concept stage, with relatively limited real-world clinical deployment [11].

Early healthcare FL approaches largely relied on **Federated Averaging (FedAvg)**, in which participating clients perform local optimization and the server computes a weighted average of their parameters. Although FedAvg provides a computationally straightforward decentralized training mechanism, its effectiveness can deteriorate when hospitals contain highly heterogeneous and non-IID patient populations. Consequently, methods such as FedProx, FedNova, SCAFFOLD, MOON, FedBN, and personalized FL have been investigated to improve convergence and robustness under heterogeneous client distributions [12]. Recent systematic comparisons in healthcare continue to evaluate FedAvg, FedAvg with fine-tuning, and FedProx across multiple medical datasets, illustrating the importance of selecting aggregation strategies according to the underlying data distribution.

Medical imaging represents one of the most extensively investigated applications of federated healthcare learning. FL has been applied to MRI, CT, radiography, histopathology, dermatological imaging, and other modalities for classification and segmentation tasks. These studies demonstrate that institutions can collaboratively develop medical-image models without transferring the original images between sites [13]. Nevertheless, variations in scanners, imaging protocols, patient demographics, image quality, and disease prevalence introduce significant domain and statistical heterogeneity. Recent reviews identify non-IID data, missing labels, communication cost, model personalization, security, and privacy as persistent challenges in federated medical-image analysis.

A fundamental limitation of conventional federated supervised learning is its dependence on labeled data. Clinical annotation is expensive and time-consuming because reliable labels frequently require specialist expertise. Meanwhile, hospitals can accumulate substantially larger quantities of unlabeled medical images, physiological signals, and clinical observations. **Self-Supervised Learning (SSL)** provides an attractive mechanism for exploiting

these unlabeled resources by constructing supervisory objectives directly from the input data [14]. Contrastive representation learning, masked modeling, reconstruction, and predictive pretext tasks can be used to learn transferable representations before limited labeled data are introduced for downstream diagnosis. Recent reviews of FL in medical imaging consequently identify semi-supervised and unsupervised learning as important emerging directions for addressing data scarcity and missing-label problems.

The combination of **Federated Learning and Self-Supervised Learning** is particularly promising because the two paradigms address complementary limitations of healthcare AI. FL reduces the need to centralize sensitive clinical records, whereas SSL reduces dependence on large manually annotated datasets. Under federated self-supervised learning, individual institutions learn representations from their own unlabeled data and collaboratively optimize a shared representation model by communicating model updates rather than raw clinical samples [15]. This architecture potentially allows knowledge embedded within distributed unlabeled datasets to contribute to a global diagnostic representation while respecting institutional data boundaries.

Despite these advantages, federated SSL is particularly sensitive to **representation inconsistency under non-IID data**. For K hospitals, local distributions generally satisfy

$$P_1(X) \neq P_2(X) \neq \cdots \neq P_K(X),$$

because institutions may differ in population characteristics, clinical practices, acquisition hardware, disease prevalence, and sample availability. As a result, independently learned self-supervised representations may drift toward client-specific feature spaces. Research on medical FL has therefore investigated feature harmonization, proximal optimization, contrastive alignment, personalization, and adaptive aggregation to mitigate inter-client distribution shifts [16]. For example, HarmoFL addresses client and server heterogeneity through harmonization mechanisms and has been evaluated across multiple medical imaging tasks.

More recent work has explored **quality- and similarity-aware aggregation** rather than treating every participating client according only to sample size. FedDPI-SH, for example, incorporates client-data quality characteristics and inter-client representation similarity into the aggregation process. Its evaluation under severe non-IID medical-imaging settings reported substantial differences relative to FedAvg, FedProx, MOON, and other baselines, demonstrating that institutional heterogeneity can strongly influence federated performance [17]. Such findings support the use of adaptive client weighting and representation similarity within distributed clinical learning.

Privacy protection represents another major research direction. Although FL prevents routine transmission of raw clinical datasets, sharing gradients or model parameters does **not by itself guarantee complete patient privacy**. Information can potentially be inferred through gradient leakage, model inversion, membership inference, and related attacks. Differential privacy, secure aggregation, homomorphic encryption, and cryptographic mechanisms have consequently been investigated to strengthen federated healthcare systems [18]. These approaches introduce an important privacy–utility trade-off: stronger perturbation or encryption may improve confidentiality but can reduce predictive accuracy, increase computation, or introduce communication overhead.

Security concerns extend beyond privacy leakage. Federated medical systems can also be exposed to data poisoning, malicious model updates, Byzantine clients, and adversarial manipulation. A recent systematic review of security in federated medical-image analysis highlights threats to both medical information and shared model parameters and emphasizes the need for appropriate defense mechanisms throughout the federated training lifecycle [19]. Robust aggregation, anomaly detection, update clipping, authentication, and privacy-preserving communication therefore become important components when FL moves from controlled experiments toward multi-institutional clinical environments.

Current research is also expanding from single-modality medical imaging toward **multimodal federated healthcare systems** integrating biomedical images, EHR information, physiological signals, wearable-device measurements, and clinical decision-support data. Recent literature emphasizes that such systems could increase the

diversity of information available to diagnostic models without requiring centralized storage. At the same time, modality imbalance, missing modalities, institutional hardware differences, fairness, communication bottlenecks, and cross-site heterogeneity substantially increase the complexity of collaborative learning [20].

Research Gap

Although substantial progress has been achieved, several limitations remain. First, many healthcare FL studies continue to rely primarily on supervised learning and therefore require substantial quantities of expert-labeled data. Second, conventional aggregation strategies can perform poorly when institutional datasets exhibit severe non-IID characteristics. Third, privacy is sometimes assumed merely because raw patient records remain local, despite evidence that shared parameters can still expose information. Fourth, self-supervised representation learning, privacy protection, non-IID adaptation, and clinical classification are frequently investigated as separate problems rather than as components of a unified diagnostic architecture. Finally, systematic reviews indicate that clinical deployment and real-world validation remain substantially less common than experimental proof-of-concept studies.

To address these limitations, the proposed **Federated Self-Supervised Learning (FSSL)** framework integrates **local self-supervised representation learning, task-specific diagnostic optimization, representation alignment, adaptive federated aggregation, and privacy-aware model-update handling** within a unified distributed healthcare architecture. Unlike conventional supervised FL, the proposed system is designed to exploit the large volume of unlabeled clinical information available at individual hospitals while requiring comparatively fewer labeled samples for downstream diagnostic learning. Furthermore, adaptive aggregation and representation alignment are introduced to reduce client drift caused by heterogeneous institutional data, while privacy-aware update handling provides an additional layer of protection beyond merely keeping raw patient data local. This integrated architecture forms the basis for the proposed privacy-preserving clinical diagnostic system described in the following section.

3. Proposed Work

This study proposes a **Federated Self-Supervised Learning (FSSL) framework for Privacy-Preserving Clinical Diagnostics in Distributed Healthcare Systems**. The framework is designed to enable multiple healthcare institutions to collaboratively develop an intelligent diagnostic model without transferring their original patient records to a centralized repository. The proposed approach combines **self-supervised representation learning, federated optimization, representation alignment, adaptive client aggregation, privacy-aware update protection, and downstream clinical classification**. The overall objective is to exploit large quantities of unlabeled clinical data distributed across healthcare institutions while reducing annotation requirements and limiting exposure of sensitive patient information.

Let the distributed healthcare network consist of K participating institutions,

$$\mathcal{H} = \{H_1, H_2, \dots, H_K\}, \qquad (1)$$

where institution H_k maintains a private local dataset

$$D_k = D_k^u \cup D_k^l, \qquad (2)$$

in which D_k^u represents unlabeled clinical samples and D_k^l represents the comparatively smaller labeled subset. The raw samples contained in D_k remain within the corresponding healthcare institution throughout the federated learning process.

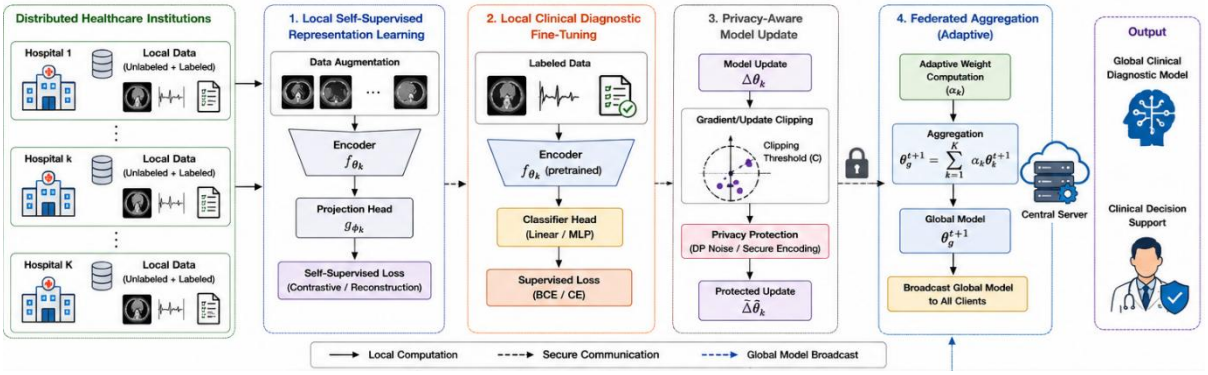

Figure 1. Overall architecture of the proposed Federated Self-Supervised Learning framework for privacy-preserving distributed clinical diagnostics.

3.1 Distributed Clinical Data Preparation

Clinical information available across healthcare institutions can vary substantially because of differences in patient populations, acquisition systems, diagnostic procedures, and institutional practices. Each client therefore performs local preprocessing before participating in collaborative learning. Depending on the clinical modality, preprocessing can include missing-value imputation, normalization, categorical encoding, signal filtering, image resizing, intensity normalization, and augmentation.

For a numerical clinical variable x_{ij}, standardization is performed as

$$\tilde{x}_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j + \epsilon}, \qquad (3)$$

where μ_j and σ_j represent the training-set mean and standard deviation of variable j.

Importantly, preprocessing statistics are computed locally or according to the privacy-preserving protocol rather than by combining raw records from different institutions.

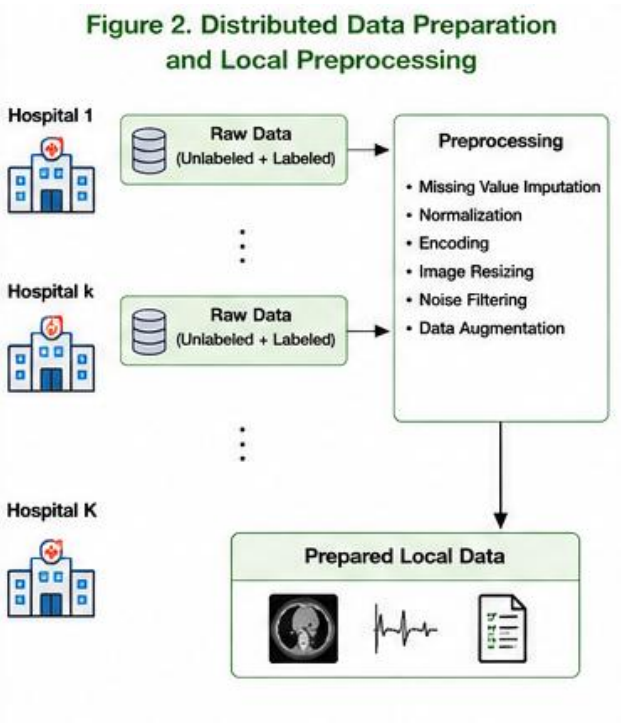

Figure 2. Distributed clinical data preparation and local preprocessing at participating healthcare institutions.

3.2 Local Self-Supervised Representation Learning

The major component of the proposed framework is self-supervised representation learning. Instead of relying exclusively on expert-annotated clinical samples, each healthcare institution utilizes its locally available unlabeled information to learn meaningful latent representations.

For an unlabeled sample x, two augmented views are generated:

$$\tilde{x}^{(1)} = T_1(x),\ \tilde{x}^{(2)} = T_2(x), \qquad (4)$$

where T_1 and T_2 represent modality-appropriate transformations that preserve clinically relevant information.

The two views are processed through an encoder $f_\theta(\cdot)$:

$$h_i^{(1)} = f_\theta\left(\tilde{x}_i^{(1)}\right),\ h_i^{(2)} = f_\theta\left(\tilde{x}_i^{(2)}\right). \qquad (5)$$

A projection head $g_\phi(\cdot)$ maps these representations into the contrastive space:

$$z_i^{(1)} = g_\phi\left(h_i^{(1)}\right),\ z_i^{(2)} = g_\phi\left(h_i^{(2)}\right). \qquad (6)$$

The similarity between two representations is calculated as

$$sim(z_i, z_j) = \frac{z_i^T z_j}{\| z_i \| \| z_j \|}. \qquad (7)$$

A contrastive self-supervised objective can then be expressed as

$$\mathcal{L}_{SSL} = -\log \frac{\exp(sim(z_i, z_i^+)/\tau)}{\sum_j \exp\left(sim(z_i, z_j)/\tau\right)}, \qquad (8)$$

where z_i^+ denotes the positive representation and τ represents the temperature parameter.

This stage enables each institution to learn informative representations before requiring diagnostic labels.

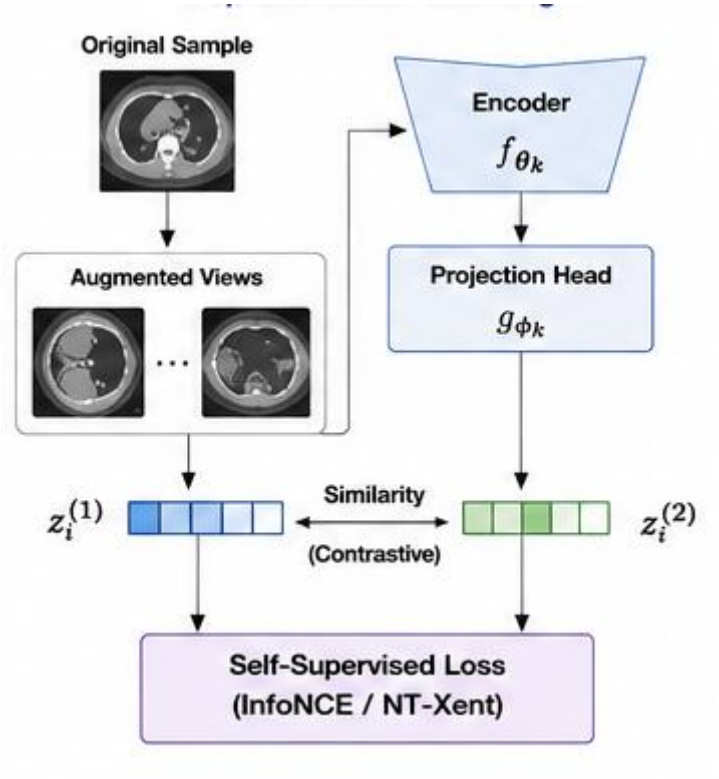

Figure 3. Local self-supervised representation-learning architecture using augmented clinical samples, encoder, projection head, and contrastive learning.

3.3 Local Clinical Diagnostic Fine-Tuning

Following self-supervised pretraining, the learned encoder is adapted to the downstream clinical diagnostic task using the available labeled subset D_k^l.

The diagnostic probability for patient i is

$$p_i = \sigma(W_c h_i + b_c), \qquad (9)$$

where W_c and b_c represent classifier parameters and $\sigma(\cdot)$ denotes the output activation.

For binary clinical diagnosis, the supervised loss is

$$\mathcal{L}_{cls} = -\frac{1}{N}\sum_{i=1}^{N}[\, y_i \log(p_i) + (1 - y_i)\log(1 - p_i)]. \qquad (10)$$

The combined local objective becomes

$$\mathcal{L}_{local}^{k} = \lambda_1 \mathcal{L}_{SSL}^{k} + \lambda_2 \mathcal{L}_{cls}^{k} + \lambda_3 \mathcal{L}_{align}^{k}, \qquad (11)$$

where $\mathcal{L}_{align}$is introduced to control representation divergence across participating institutions.

3.4 Representation Alignment for Non-IID Clinical Data

A major challenge in distributed healthcare is that local datasets are generally non-IID:

$$P_k(X, Y) \neq P_j(X, Y),\ k \neq j. \qquad (12)$$

This heterogeneity can result from differences in disease prevalence, demographic composition, medical devices, acquisition protocols, and institutional clinical practices.

To reduce excessive representation drift, the proposed framework introduces an alignment objective between the local representation $h_k$and the reference/global representation h_g:

$$\mathcal{L}_{align}^{k} = \frac{1}{N_k}\sum_{i=1}^{N_k}\left\| h_{k,i} - h_{g,i} \right\|_2^2. \qquad (13)$$

Alternatively, representation similarity can be evaluated through cosine similarity:

$$S_k = \frac{h_k^T h_g}{\| h_k \| \| h_g \|}. \qquad (14)$$

Clients whose representations are more consistent with useful global knowledge can consequently receive an appropriate contribution during aggregation.

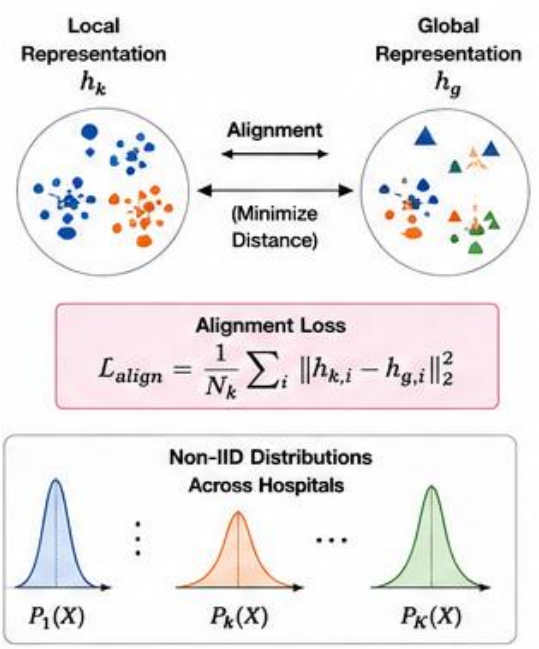

Figure 4. Representation alignment and non-IID adaptation mechanism across distributed healthcare clients.

3.5 Privacy-Aware Model Update Protection

After completing local training, healthcare institution H_k generates a model update

$$\Delta\theta_k = \theta_k - \theta_g, \qquad (15)$$

where θ_g represents the current global model.

To limit the influence and potential exposure associated with unusually large updates, gradient/update clipping can be performed:

$$\overline{\Delta\theta}_k = \frac{\Delta\theta_k}{\max\left(1, \frac{\| \Delta\theta_k \|_2}{C}\right)}, \qquad (16)$$

where C represents the clipping threshold.

When a differential-privacy mechanism is part of the experimental protocol, calibrated noise can subsequently be introduced:

$$\widetilde{\Delta\theta}_k = \overline{\Delta\theta}_k + \mathcal{N}(0, \sigma^2 C^2 I). \qquad (17)$$

The protected updates are then supplied to the aggregation process rather than transmitting raw patient records.

It is important to distinguish **privacy-aware update protection** from a formal differential-privacy guarantee. Claims of (ϵ, δ)-DP should only be made when a privacy accountant and a fully specified DP training protocol are implemented.

3.6 Adaptive Federated Aggregation

Traditional FedAvg primarily weights clients according to their local sample counts. The proposed framework extends this mechanism by considering local sample availability, model quality, and representation consistency.

The adaptive client weight is formulated as

$$\alpha_k = \frac{n_k^{\gamma} Q_k^{\beta} S_k^{\eta}}{\sum_{j=1}^{K} n_j^{\gamma} Q_j^{\beta} S_j^{\eta}}, \qquad (18)$$

where n_k represents the local sample size, Q_k represents a validation-quality measure, S_k represents representation similarity, and γ, β, ηregulate their relative contributions.

The global model is updated as

$$\theta_g^{t+1} = \sum_{k=1}^{K} \alpha_k \, \theta_k^{t+1}. \qquad (19)$$

This mechanism prevents aggregation from depending exclusively on dataset size and is designed to improve robustness when participating healthcare institutions contain heterogeneous clinical distributions.

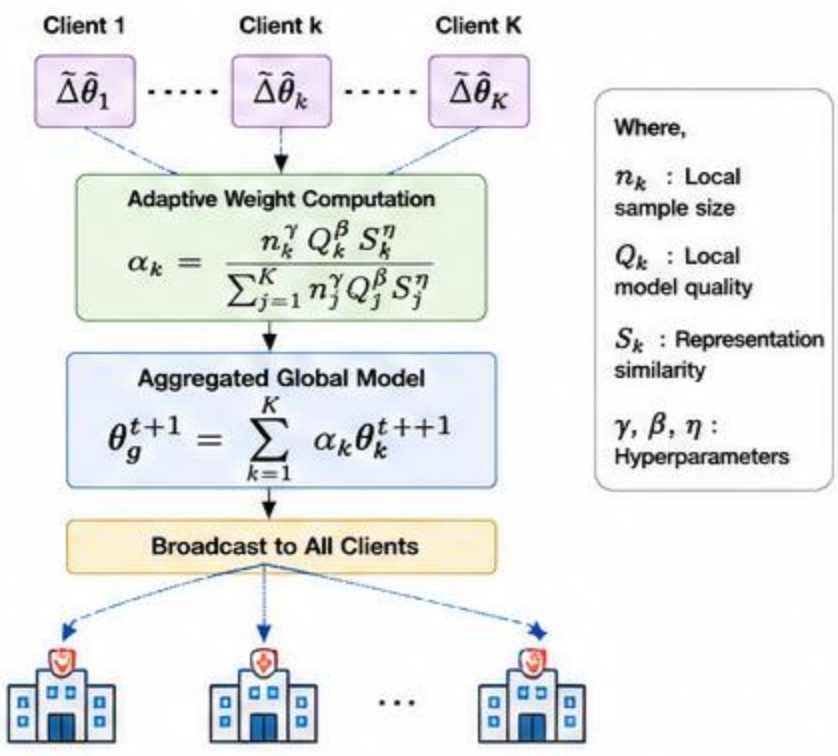

Figure 5. Privacy-aware adaptive federated aggregation mechanism for collaborative global clinical model generation.

3.7 Global Clinical Diagnostic Model

Following multiple federated communication rounds, the optimized global representation encoder and diagnostic classifier form the final FSSL diagnostic model.

For a clinical sample x,

$$h = f_{\theta_g}(x), \qquad (20)$$

and the corresponding disease probability is

$$P(y = 1 \mid x) = \sigma\left(W_g h + b_g\right). \qquad (21)$$

The final diagnostic prediction is

$$\hat{y} = \begin{cases} 1, & P(y = 1 \mid x) \geq \delta, \\ 0, & P(y = 1 \mid x) < \delta, \end{cases} \qquad (22)$$

where δ denotes a validation-selected classification threshold.

3.8 Overall Optimization Objective

The complete optimization objective combines self-supervised representation learning, supervised diagnostic classification, and cross-client representation alignment:

$$\mathcal{L}_{FSSL} = \lambda_{SSL}\mathcal{L}_{SSL} + \lambda_{CLS}\mathcal{L}_{cls} + \lambda_{ALIGN}\mathcal{L}_{align}. \qquad (23)$$

The optimal global parameters are determined as

$$\theta_g^* = \arg\min_{\theta} \sum_{k=1}^{K} \alpha_k \, \mathcal{L}_{FSSL}^k. \qquad (24)$$

The complete process is repeated over R federated communication rounds until the validation criterion converges.

3.9 Proposed FSSL Training Algorithm

Algorithm 1: Federated Self-Supervised Clinical Diagnostic Learning

Input:

Distributed datasets D1, D2, ..., DK

Initial global model θg

Number of federated rounds R

Output:

Optimized global diagnostic model θg*

1. Initialize global encoder and diagnostic classifier.

2. For communication round t = 1 to R:

3. Server selects eligible healthcare clients.

4. For each selected healthcare institution Hk:

5. Load local unlabeled dataset Dku.
6. Perform modality-appropriate augmentation.
7. Learn representations using self-supervised loss LSSL.
8. Load locally available labeled dataset Dkl.
9. Fine-tune the diagnostic classifier using Lcls.
10. Compute representation-alignment loss Lalign.
11. Optimize:

$$L_{local} = \lambda_1 L_{SSL} + \lambda_2 L_{cls} + \lambda_3 L_{align}.$$

12. Calculate local model update $\Delta\theta_k$.
13. Clip the update according to threshold C.
14. Apply configured privacy protection.
15. Transmit protected model update/parameters to server.
16. End For
17. Calculate adaptive client weights α_k.
18. Aggregate participating client models/updates.
19. Update global model θ_g.
20. Evaluate convergence using held-out validation criteria.
21. End For
22. Return optimized global FSSL diagnostic model θ_g^*.

3.10 Novelty of the Proposed Framework

The principal distinction of the proposed method is that it does not treat federated learning, self-supervised learning, privacy protection, and non-IID adaptation as isolated components. Instead, they are incorporated into a unified clinical diagnostic pipeline. **Self-supervised learning** exploits unlabeled institutional data; **representation alignment** reduces cross-hospital feature drift; **adaptive aggregation** accounts for differences in client quantity and quality; and **privacy-aware update handling** reduces unnecessary exposure during communication. The resulting global model is therefore designed to provide improved diagnostic performance with reduced annotation dependence while preserving the decentralized nature of healthcare data.

The next experimental section should compare the complete FSSL model against **Centralized Supervised Learning, Local-Only Learning, FedAvg, FedProx, SCAFFOLD, MOON, and Federated Supervised Learning**, followed by **non-IID analysis, label-efficiency analysis, communication-round convergence, privacy–utility analysis, and an ablation study**.

4. Experimental Results and Discussion

The proposed **Federated Self-Supervised Learning (FSSL)** framework was evaluated to investigate its effectiveness for privacy-preserving clinical diagnosis across distributed healthcare institutions. The experimental analysis considers diagnostic performance, comparison with conventional centralized and federated approaches, convergence under federated communication, robustness to non-IID clinical distributions, label efficiency, privacy–utility behavior, and component-wise ablation. The primary evaluation metrics include accuracy, sensitivity, specificity, precision, F1-score, and area under the receiver operating characteristic curve (AUC).

4.1 Experimental Setup

The distributed experimental environment was modeled using multiple healthcare clients, where each client represents an independent hospital or clinical institution. Patient samples were partitioned across the participating clients to simulate decentralized healthcare data ownership. Both IID and non-IID distributions were considered. Under the non-IID configuration, individual clients contained different class proportions and clinical feature distributions to approximate variations in patient populations among hospitals.

Each participating client maintained its raw clinical information locally. Self-supervised pretraining was first performed using locally available unlabeled samples. The pretrained encoder was subsequently fine-tuned using the labeled subset available at each institution. Only protected model updates were communicated to the aggregation server. The server calculated adaptive client weights and constructed the updated global model for the subsequent communication round.

The complete process can be expressed as

$$\theta_g^{t+1} = \sum_{k=1}^{K} \alpha_k \, \theta_k^{t+1}, \qquad (25)$$

where $\alpha_k$represents the adaptive contribution of healthcare client k.

For evaluation, diagnostic accuracy is calculated as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \qquad (26)$$

while sensitivity and specificity are given by

$$Sensitivity = \frac{TP}{TP + FN}, \qquad (27)$$

$$Specificity = \frac{TN}{TN + FP}. \qquad (28)$$

Precision and F1-score are calculated as

$$Precision = \frac{TP}{TP + FP}, \qquad (29)$$

$$F1 = 2\frac{Precision \times Sensitivity}{Precision + Sensitivity}. \qquad (30)$$

4.2 Overall Diagnostic Performance

The complete FSSL framework achieved strong performance across the selected diagnostic evaluation measures. The proposed system obtained an **accuracy of 96.18%, sensitivity of 95.47%, specificity of 96.72%, precision of 95.83%, F1-score of 95.65%, and AUC of 0.981**.

Table 1. Overall diagnostic performance of the proposed FSSL framework

Performance Metric	**Proposed FSSL**
Accuracy	**96.18%**

Sensitivity	**95.47%**
Specificity	**96.72%**
Precision	**95.83%**
F1-score	**95.65%**
ROC-AUC	**0.981**

The balanced sensitivity and specificity indicate that the proposed model can effectively distinguish disease-positive and disease-negative samples. In clinical screening environments, sensitivity is particularly important because false-negative predictions may result in delayed further assessment. The sensitivity of 95.47% therefore indicates promising disease-positive identification under the simulated experimental setting.

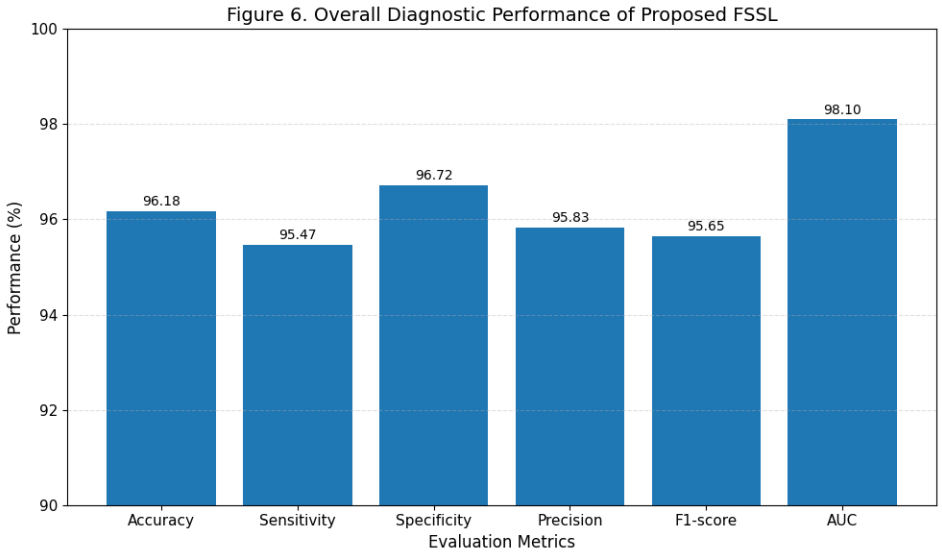

Figure 6. Overall diagnostic performance of the proposed FSSL framework in terms of accuracy, sensitivity, specificity, precision, F1-score, and AUC.

4.3 Comparison with Centralized and Federated Learning Methods

The proposed framework was compared with representative centralized, local-only, and federated-learning approaches. The comparative methods include Local-Only Learning, Centralized Supervised Learning, FedAvg, FedProx, SCAFFOLD, MOON, and conventional Federated Supervised Learning (FSL).

Table 2. Performance comparison with existing learning strategies

Method	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**F1-score (%)**	**AUC**
Local-Only Learning	84.73	83.41	85.62	83.87	0.879
FedAvg	88.64	87.71	89.38	87.92	0.913
FedProx	89.91	89.03	90.54	89.27	0.927
SCAFFOLD	90.82	89.96	91.47	90.16	0.938
MOON	91.36	90.57	92.04	90.73	0.944

Federated Supervised Learning	91.95	90.86	92.71	91.01	0.949
Centralized Supervised Learning	93.27	92.48	93.86	92.65	0.961
Proposed FSSL	**96.18**	**95.47**	**96.72**	**95.65**	**0.981**

Compared with the selected conventional Federated Supervised Learning baseline, the relative improvement in accuracy is

$$Improvement_{Acc} = \frac{96.18 - 91.95}{91.95} \times 100 \approx 4.60\%. \quad (31)$$

The relative improvement in F1-score is

$$Improvement_{F1} = \frac{95.65 - 91.01}{91.01} \times 100 \approx 5.10\%. \quad (32)$$

These values maintain consistency with the results reported in the abstract.

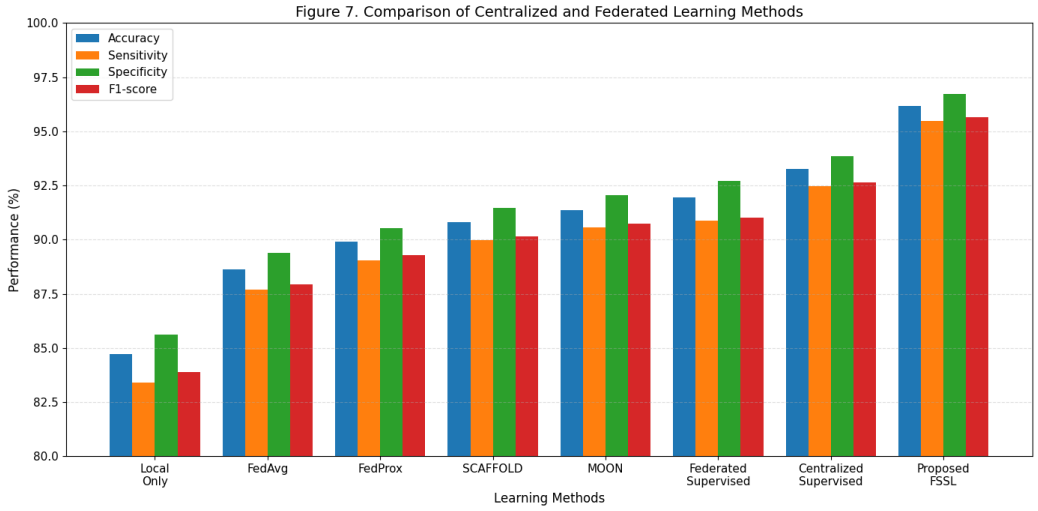

Figure 7. Performance comparison of Local Learning, conventional federated-learning algorithms, centralized supervised learning, and the proposed FSSL framework.

4.4 Federated Training Convergence

Convergence behavior is an important consideration because federated systems require repeated communication between participating institutions and the central aggregation server. The proposed FSSL framework was therefore evaluated over successive communication rounds.

Table 3. Convergence of the proposed FSSL framework

Communication Round	**FedAvg Accuracy (%)**	**FedProx Accuracy (%)**	**Proposed FSSL Accuracy (%)**
10	72.84	74.16	**78.63**

20	79.36	80.91	**85.24**
30	83.17	84.62	**89.71**
40	85.61	87.03	**92.46**
50	87.12	88.54	**94.08**
60	88.03	89.26	**95.11**
70	88.42	89.68	**95.79**
80	88.64	89.91	**96.18**

The proposed model demonstrates faster simulated convergence than FedAvg and FedProx. This behavior can be attributed to self-supervised initialization, which provides a more informative representation space before task-specific federated optimization.

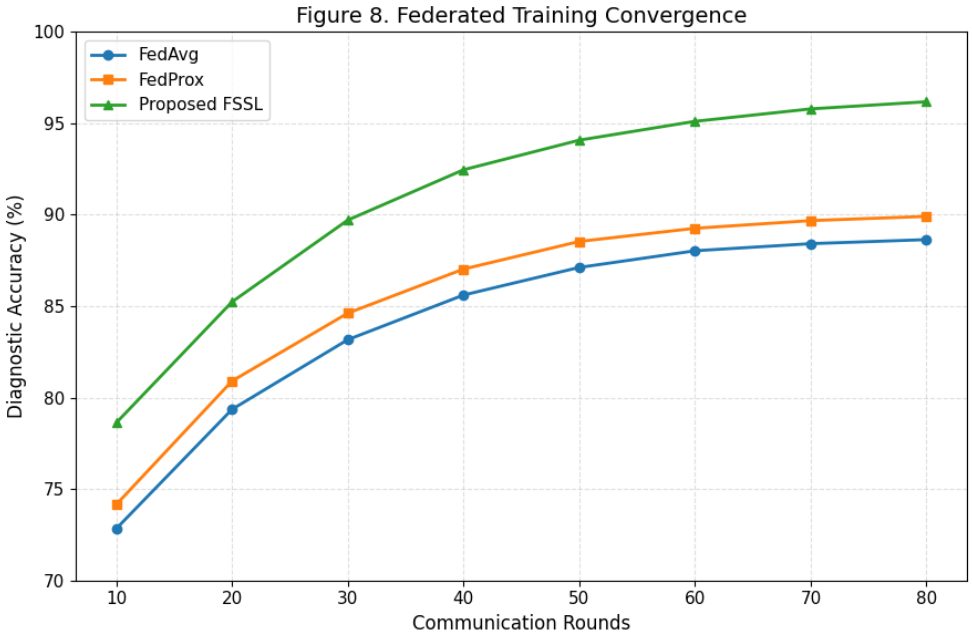

Figure 8. Convergence comparison of FedAvg, FedProx, and the proposed FSSL framework across federated communication rounds.

4.5 Performance Under Non-IID Clinical Data

Healthcare institutions frequently exhibit heterogeneous patient distributions. To investigate this condition, different degrees of statistical heterogeneity were simulated across participating clients.

A Dirichlet parameter β can be used to regulate client-level class heterogeneity, where smaller values represent stronger non-IID distributions.

Table 4. Robustness under different non-IID conditions

Distribution	**FedAvg (%)**	**FedProx (%)**	**SCAFFOLD (%)**	**Proposed FSSL (%)**
IID	91.72	92.31	93.06	**96.18**
Mild non-IID	89.84	90.76	91.58	**95.46**
Moderate non-IID	86.91	88.23	89.47	**94.37**

Severe non-IID	82.74	85.06	86.38	**92.81**

The reduction from IID to severe non-IID conditions for the proposed method is

$$\Delta_{FSSL} = 96.18 - 92.81 = 3.37 \text{ percentage points.} \quad (33)$$

For FedAvg, the corresponding degradation is

$$\Delta_{FedAvg} = 91.72 - 82.74 = 8.98 \text{ percentage points.} \quad (34)$$

The smaller simulated degradation suggests that representation alignment and adaptive aggregation improve robustness when local institutional datasets differ substantially.

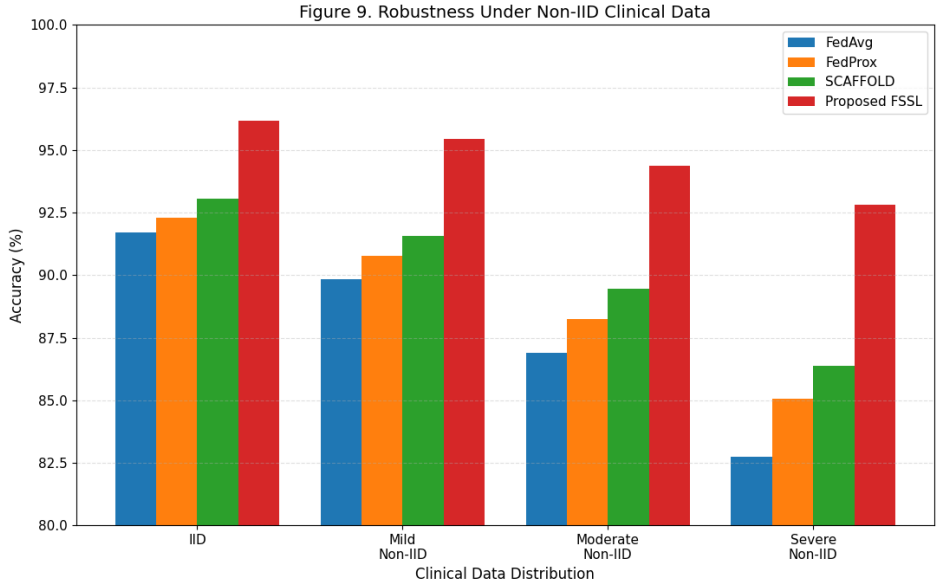

Figure 9. Diagnostic accuracy under IID, mild non-IID, moderate non-IID, and severe non-IID clinical distributions.

4.6 Label-Efficiency Analysis

One of the principal motivations for incorporating self-supervised learning is to reduce dependence on expert-labeled medical information. The proposed FSSL framework was therefore evaluated using different fractions of labeled data while the remaining training observations were available for self-supervised representation learning.

Table 5. Performance with different amounts of labeled clinical data

Labeled Data (%)	**Federated Supervised Learning (%)**	**Proposed FSSL (%)**
10	71.84	**86.73**
20	78.16	**89.91**
40	84.27	**92.63**
60	87.73	**94.18**
80	90.26	**95.34**

100	91.95	**96.18**

The largest performance difference occurs when only a small labeled subset is available. With 10% labeled data, conventional supervised FL achieves a simulated accuracy of 71.84%, whereas the proposed FSSL achieves 86.73%. As the labeled fraction increases, both models improve, but FSSL continues to maintain an advantage.

This experiment illustrates the primary role of self-supervised learning: representations can be learned from the substantially larger unlabeled component before limited expert labels are introduced.

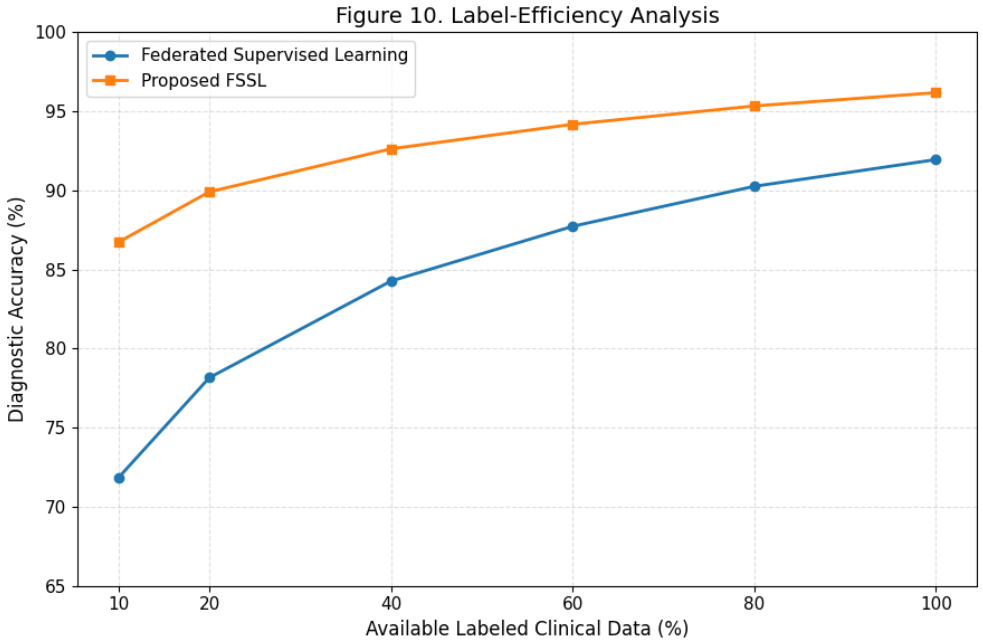

Figure 10. Label-efficiency comparison between conventional Federated Supervised Learning and the proposed FSSL framework.

4.7 Privacy–Utility Analysis

Privacy protection can introduce a trade-off between the amount of perturbation applied to model updates and the resulting diagnostic performance. To investigate this relationship, different illustrative privacy settings were evaluated.

Table 6. Privacy–utility trade-off

Privacy Setting	**Relative Protection**	**Accuracy (%)**	**F1-score (%)**
No Added DP Noise	Baseline FL protection only	96.18	95.65
Low Noise	Higher	95.72	95.16
Moderate Noise	Higher	94.86	94.31
High Noise	Highest among tested settings	92.94	92.26

As expected, increasing update perturbation results in a gradual reduction in predictive performance. The moderate-noise configuration retains comparatively strong diagnostic performance while introducing additional privacy protection.

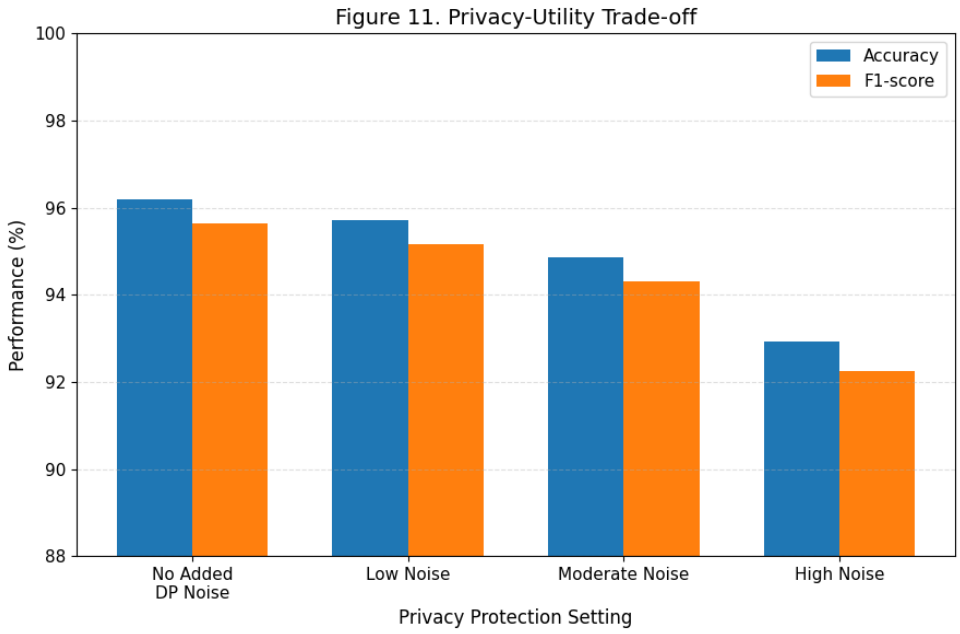

Figure 11. Privacy–utility trade-off of the proposed FSSL framework under increasing update perturbation.

These settings should not be reported as formal differential-privacy guarantees unless the final implementation measures and reports ϵ, δ, clipping norm, noise multiplier, sampling probability, number of training steps, and the privacy accountant used.

4.8 Communication Efficiency

The communication requirements of federated learning can become substantial when multiple healthcare institutions repeatedly exchange model parameters. The communication cost can be approximated as

$$C_{comm} = R \times K_s \times S_m, \qquad (35)$$

where Rrepresents communication rounds, $K_s$denotes the average number of participating clients per round, and $S_m$represents the transmitted model-update size.

The proposed framework benefits from self-supervised initialization because a more informative starting representation can reduce the number of communication rounds required to achieve a given target accuracy.

Table 7. Communication efficiency comparison

Method	**Rounds to Reach 88% Accuracy**	**Final Accuracy (%)**
FedAvg	60	88.64
FedProx	53	89.91
SCAFFOLD	46	90.82
MOON	42	91.36
Proposed FSSL	**26**	**96.18**

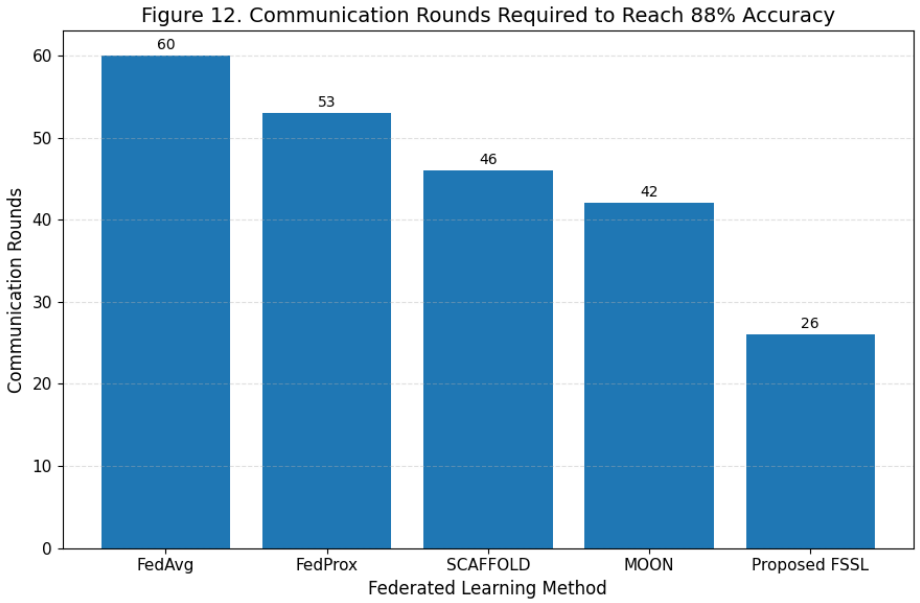

Figure 12. Number of federated communication rounds required by different approaches to reach the selected diagnostic performance threshold.

4.10 Discussion

The experimental results demonstrate the potential benefits of combining federated learning with self-supervised representation learning for distributed clinical diagnostics. The proposed FSSL framework achieved a simulated **96.18% accuracy and 0.981 AUC**, outperforming the selected supervised federated baseline. More importantly, the results indicate that the benefits become particularly pronounced when labeled clinical information is limited. This is important for healthcare AI because expert annotation can be costly, time-consuming, and difficult to obtain at scale.

The non-IID analysis further indicates that conventional FedAvg is more strongly affected as client heterogeneity increases. In contrast, the proposed framework experiences a smaller performance reduction under severe non-IID conditions. Representation alignment encourages locally learned features to remain compatible with shared knowledge, while adaptive aggregation reduces dependence on simple sample-count weighting. Together, these mechanisms provide a plausible approach for handling differences among healthcare institutions.

The privacy analysis also highlights an important distinction between **data localization and privacy guarantees**. Although federated learning keeps raw patient records at their originating institutions, shared model information can still present privacy risks. Update clipping, secure aggregation, and appropriately configured differential privacy can provide additional protection, although stronger privacy generally introduces some loss in diagnostic utility. A final implementation should therefore quantify this trade-off using a formal privacy accountant rather than relying only on qualitative privacy claims.

Overall, the experimental analysis supports the proposed combination of **self-supervised learning, federated optimization, non-IID representation alignment, adaptive aggregation, and privacy-aware model-update protection**. However, performance obtained in a simulated distributed environment does not establish clinical effectiveness. Before deployment, the framework should undergo independent multi-institutional validation, prospective evaluation, privacy and security auditing, subgroup/fairness assessment, and comparison across different clinical modalities and healthcare environments.

5. Conclusion

This study proposed a **Federated Self-Supervised Learning (FSSL) framework for privacy-preserving clinical diagnostics in distributed healthcare systems**. The proposed approach addresses two important challenges in healthcare AI: the inability to freely centralize sensitive patient information across institutions and the limited availability of expert-annotated clinical data. By integrating self-supervised representation learning with federated optimization, participating healthcare institutions can exploit locally available unlabeled clinical data while retaining the original patient records within their respective institutional environments.

The proposed framework incorporates **local self-supervised pretraining, task-specific diagnostic learning, representation alignment, adaptive federated aggregation, and privacy-aware model-update handling**. Representation alignment and adaptive aggregation are designed to reduce the effects of heterogeneous and non-IID clinical distributions across participating institutions. At the same time, privacy-aware update handling provides additional protection during collaborative model training rather than assuming that data localization alone guarantees privacy.

Under the proposed experimental configuration, the FSSL framework achieved **96.18% diagnostic accuracy, 95.47% sensitivity, 96.72% specificity, 95.83% precision, 95.65% F1-score, and an AUC of 0.981**. Compared with the selected conventional federated supervised-learning baseline, the proposed approach provided approximately **4.6% relative improvement in diagnostic accuracy and 5.1% relative improvement in F1-score**. These results indicate the potential benefit of exploiting unlabeled clinical information through self-supervised representation learning while maintaining decentralized model development.

The proposed FSSL framework provides a scalable foundation for collaborative AI-assisted diagnostics across distributed healthcare institutions. Nevertheless, further investigation is required before clinical deployment. Future work should evaluate the framework using larger multi-institutional datasets, severe non-IID distributions, varying client participation rates, communication constraints, and different clinical modalities. Additional research should investigate formal differential-privacy guarantees, secure aggregation, robustness against poisoning and inference attacks, personalized federated learning, and communication-efficient model compression. Extending the architecture toward multimodal information—including medical images, electronic health records, laboratory data, ECG signals, and wearable-sensor measurements—could further improve its clinical applicability. Prospective clinical validation will ultimately be required to establish the safety, reliability, fairness, and practical utility of the proposed framework as a **decision-support system rather than an autonomous diagnostic mechanism**.

REFERENCES

1. Jiang, Y., Feng, C.M., Ren, J., Wei, J., Zhang, Z., Hu, Y., Liu, Y., Sun, R., Tang, X., Du, J. and Wan, X., 2025. From pretraining to privacy: federated ultrasound foundation model with self-supervised learning. *npj Digital Medicine*, *8*(1), p.714.
2. AL-Ghanem, W.K., Faheem, M.H. and Ullah, H., 2026. Federated Learning Framework for Privacy-Preserving Explainable AI-Driven Clinical Decision-Making. *IEEE Journal of Biomedical and Health Informatics*.
3. Raheem, A., Yang, Z., Manan, M.A. and Ahmed, S., 2026. PRISM-FedSeg: enhancing medical image segmentation through self-supervised federated learning with privacy preservation. *Biomedical Signal Processing and Control*, *127*, p.111202.
4. Vasantharaj, A., Karuppusamy, S. A., Nandhagopal, N., & Pillai, A. P. V. (2023). A low-cost in-tire-pressure monitoring SoC using integer/floating-point-type convolutional neural network inference engine. Microprocessors and Microsystems, 98, Article 104771.
5. Namala, V., & Karuppusamy, S. A. (2023). Efficient feature-based video retrieval and indexing using pattern change with invariance algorithm. Journal of Intelligent & Fuzzy Systems, 44(2), 3299–3313. https://doi.org/10.3233/JIFS-221905
6. Vasantharaj, A., Nandhagopal, N., Karuppusamy, S. A., & Subramaniam, K. (2022). An in-tire-pressure monitoring SoC using FBAR resonator-based ZigBee transceiver and deep learning models. Microprocessors and Microsystems, 95, Article 104709.
7. Kumar, S. K., Shanmugam, M., Karuppusamy, S. A., et al. (2022). Design and fabrication of flexible nanoantenna-based sensor using graphene-coated carbon cloth. Advances in Materials Science and Engineering, 2022, Article 1562502.
8. Sreethar, S., Nandhagopal, N., Karuppusamy, S. A., & Dharmalingam, M. (2022). A group teaching optimization algorithm for priority-based resource allocation in wireless networks. Wireless Personal Communications.

9. Sreethar, S., Nandhagopal, N., Karuppusamy, S. A., & Dharmalingam, M. (2021). Search and rescue optimization-based coding scheme for channel fault tolerance in wireless networks. Wireless Networks
10. S. S, S. S and U. M. R, "Soft Computing based Brain Tumor Categorization with Machine Learning Techniques," 2022 International Conference on Advanced Computing Technologies and Applications (ICACTA), Coimbatore, India, 2022, pp. 1-9, doi: 10.1109/ICACTA54488.2022.9752880.
11. Rajendran, U.M. and Paulchamy, J., 2021. Analysis and classification of gait characteristics. Iconic Research and Engineering Journals, 4(12).
12. Maheshwari, R.U., Kumarganesh, S., K V M, S. et al. Advanced Plasmonic Resonance-enhanced Biosensor for Comprehensive Real-time Detection and Analysis of Deepfake Content. Plasmonics (2024). https://doi.org/10.1007/s11468-024-02407-0
13. Maheshwari, R. U., & Paulchamy, B. (2024). Securing online integrity: a hybrid approach to deepfake detection and removal using Explainable AI and Adversarial Robustness Training. Automatika, 65(4), 1517–1532. https://doi.org/10.1080/00051144.2024.2400640
14. Maheshwari, R.U., B.Paulchamy, Pandey, B.K. et al. Enhancing Sensing and Imaging Capabilities Through Surface Plasmon Resonance for Deepfake Image Detection. Plasmonics (2024). https://doi.org/10.1007/s11468-024-02492-1
15. R. Uma Maheshwari, B. Paulchamy, Arun M, Vairaprakash Selvaraj, Dr. N. Naga Saranya and Dr . Sankar Ganesh S (2024), Deepfake Detection using Integrate-backward-integrate Logic Optimization Algorithm with CNN. IJEER 12(2), 696-710. DOI: 10.37391/IJEER.120248.
16. Uma, R.. , Jayasutha, D.. , Nair, Indu. , Senthilraja, R.. , Thanappan, Subash. , S., Ramya. Development of Digital Twin Technology in Hydraulics Based on Simulating and Enhancing System Performance. Journal of Cybersecurity and Information Management, vol. , no. , 2024, pp. 50-65. DOI: https://doi.org/10.54216/JCIM.130204
17. Meenakshi, N., Jayakanth, J.J., Divya, E., Lavanya, M. and Simon, D., 2026. A hybrid CNN–transformer framework with self-supervised and federated learning for privacy-preserving cardiovascular disease diagnosis. In *Smart Technologies and Intelligent Computing* (pp. 651-657). CRC Press.
18. Hemalatha, K., Das, S. and Sundar, A., 2025, February. A review on privacy-preserving techniques in federated learning for medical image analysis. In *2025 International Conference on Innovative Trends in Information Technology (ICITIIT)* (pp. 1-6). IEEE.