
FedIHRAS: A Privacy-Preserving Federated Learning Framework for Multi-Institutional Collaborative Radiological Analysis with Integrated Explainability and Automated Clinical Reporting

[André Luiz Marques Serrano](#)^{*}, Gabriel Rodrigues, [Guilherme Dantas Bispo](#)^{*}, [Vinícius Pereira Gonçalves](#)^{*}, [Geraldo Pereira Rocha Filho](#), [Maria Gabriela Mendonça Peixoto](#), [Rodrigo Bonacin](#), Rodolfo Ipolito Meneguette

Posted Date: 23 February 2026

doi: 10.20944/preprints202602.1347.v1

Keywords: chest X-ray; Federated Learning; Privacy; Radiology



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

FedIHRAS: A Privacy-Preserving Federated Learning Framework for Multi-Institutional Collaborative Radiological Analysis with Integrated Explainability and Automated Clinical Reporting

Andre Luiz Marques Serrano ^{1,*}, Gabriel Rodrigues ¹, Guilherme Dantas Bispo ^{1,*},
Vinicius Pereira Gonçalves ^{1,*}, Geraldo Pereira Rocha Filho ²,
Maria Gabriela Mendonça Peixoto ¹, Rodrigo Bonacin ³ and Rodolfo Ipolito Meneguette ⁴

¹ Department of Electrical Engineering, University of Brasilia, Federal District 70910-900, Brazil

² Department of Exact and Technological Sciences, State University of Southwest Bahia, Vitória da Conquista 45083-900, Brazil

³ Renato Archer Information Technology Center, Campinas, Brazil

⁴ Institute of Mathematical and Computer Sciences, University of São Paulo, São Carlos 13566-590, Brazil

* Correspondence: andrelms@unb.br (A.L.M.S.); guilherme.bispo@redes.unb.br (G.D.B.); vpgvinicius@unb.br (V.P.G.)

Abstract

The rapid growth of medical imaging data has intensified the need for advanced computational tools to support clinical decision-making. However, centralized approaches to artificial intelligence development raise significant challenges related to privacy, regulation, and generalizability. This paper introduces FedIHRAS (Federated Intelligent Humanized Radiology Analysis System), a privacy-preserving federated learning framework that enables multi-institutional collaboration for chest X-ray analysis. FedIHRAS integrates pathology classification, visual explainability, anatomical segmentation, and automated clinical report generation into a unified system that incorporates adaptive aggregation strategies, heterogeneity, and non-IID distributions. The framework employs multi-layered differential privacy mechanisms and a secure communication infrastructure to ensure compliance with strict healthcare data protection standards. Experimental validation across four large-scale chest radiograph datasets (approximately 874k images) demonstrates that FedIHRAS retains 98.8% of the diagnostic accuracy of a centralized model (mean AUC-ROC = 0.911 vs. 0.922) and achieves superior generalization to unseen institutions (94.2% retention). Explainability and interpretability were preserved at near-centralized levels, with expert radiologists rating 94.6% of attention maps as clinically reliable. Moreover, privacy robustness tests confirm strong resistance against inference and reconstruction attacks. FedIHRAS reduces barriers to collaborative research and mitigates algorithmic bias, ultimately offering a scalable and equitable solution for radiological analysis in real-world healthcare systems.

Keywords: chest X-ray; Federated Learning; Privacy; Radiology

1. Introduction

The convergence of engineering, medicine, and computing is reshaping contemporary healthcare, with medical radiology emerging as one of the most profoundly impacted domains. The exponential increase in imaging data contrasts sharply with the global shortage of qualified professionals, resulting in diagnostic delays and growing workloads. In this scenario, Artificial Intelligence (AI) has become a transformative ally, enhancing the speed, accuracy, and consistency of image interpretation. Comprehensive systems such as IHRAS (Intelligent Humanized Radiology Analysis System) [1] have demonstrated AI's potential to perform pathology classification, anatomical segmentation, and automated medical report generation, thereby supporting radiologists' decision-making [2].

However, the conventional centralized paradigm for AI model development, which aggregates massive datasets in a single repository, faces increasing ethical, technical, and regulatory challenges.

The need to share sensitive patient information across institutions conflicts with data protection laws, including the General Data Protection Regulation (GDPR) in Europe and the Health Insurance Portability and Accountability Act (HIPAA) in the United States [3]. Beyond privacy risks, centralization often amplifies algorithmic bias, as models trained on homogeneous data distributions tend to underperform when applied to diverse populations, undermining fairness and generalizability.

Federated Learning (FL) has emerged as a promising and ethical alternative to centralized AI development. This decentralized approach enables multiple institutions to collaboratively train a shared model without exchanging raw data [4,5]. Each participating node computes local updates based on its own data, and only model parameters are aggregated globally, thereby preserving data sovereignty and regulatory compliance. In addition to mitigating privacy concerns, this paradigm fosters collaboration among institutions with heterogeneous resources, promoting scalability, cost-efficiency, and inclusivity in the development of medical AI systems.

In this context, the present research introduces FedIHRAS (Federated Intelligent Human Radiology Analysis System), an innovative framework that extends the IHRAS architecture while preserving the environment. The system implements adaptive aggregation strategies that weight each node's contribution based on data quality and confidence metrics, and integrates multi-layered privacy mechanisms. It integrates and secures aggregation to ensure compliance and robustness. Furthermore, it preserves the semantic and structural coherence of the original IHRAS outputs, maintaining explainability and clinical interpretability across institutions.

From clinical and societal perspectives, FedIHRAS contributes to the development of equitable and globally representative AI in radiology. By enabling model training on data from institutions of different sizes, regions, and patient profiles, the framework addresses the persistent challenge of algorithmic bias and improves diagnostic generalization. Its integration with standardized medical terminologies such as SNOMED CT also enhances interoperability, supporting structured reporting and large-scale clinical research [6].

Economically, federated learning reduces dependence on centralized infrastructure, enabling resource sharing and collaborative development without the high costs of centralized data storage associated with processing [7]. This approach fosters a cooperative ecosystem in which each institution contributes its privately held data and local computational resources, thereby enabling a high-performance diagnostic model that improves efficiency and sustainability in medical imaging workflows [8].

Given this complex scenario, this study addresses a central research question: How can a high-performance radiological analysis system be adapted to a federated learning environment that preserves data privacy, ensures robustness, and maintains clinical interpretability without compromising diagnostic accuracy? To answer this, the objectives of this research are: (1) to design and implement adaptive aggregation strategies for multi-component AI architectures; (2) to integrate multi-layered privacy-preserving mechanisms such as differential privacy; (3) to rigorously compare the performance of FedIHRAS with its centralized counterpart; and (4) to validate its clinical utility, robustness, and explainability in cross-institutional scenarios.

2. Literature Review and Theoretical Background

Several paradigm shifts have marked the journey toward intelligent radiological analysis systems, building upon previous advances while addressing emerging clinical needs and evolving technological capabilities [9]. Understanding this evolution is essential to appreciate appreciating achievements and the persistent limitations that motivate the development of federated approaches.

2.1. Era of Traditional CAD Systems

The earliest applications of AI in radiology primarily focused on computer-aided detection (CAD) systems for specific pathologies, such as mammographic screening for breast cancer and chest X-ray analysis for tuberculosis detection [10]. These systems, although significant for their time, were limited in scope and often suffered from high false-positive rates, hindering clinical adoption.

The work of Yin et al. on the development of CAD systems for mammography established many of the foundational principles that continue to guide medical AI today [11]. Their research demonstrated that computational algorithms could identify subtle patterns in medical images that might elude human detection, laying the conceptual groundwork for more sophisticated systems.

During this era, CAD systems relied heavily on manually engineered features and traditional machine learning algorithms such as support vector machines and random forests. While these methods were effective for specific, well-defined tasks, they lacked the flexibility and generalization required for broader clinical applications. The need for manual feature engineering also limited the system's ability to adapt to new types of pathologies or imaging modalities.

2.2. The Deep Learning Revolution

The advent of deep learning, particularly convolutional neural networks (CNNs), has revolutionized the field by enabling more accurate and generalizable image analysis. The work of Pioneering researchers such as Alex Krizhevsky et al. on ImageNet classification demonstrated the potential of deep CNNs for complex visual recognition tasks, inspiring researchers to adapt these techniques to medical imaging applications [12–14].

This shift from traditional machine learning approaches to deep learning marked a fundamental change in how AI systems could process and interpret medical images. Rather than relying on manually engineered features, deep learning systems can learn representations that capture subtle patterns invisible to human observers, thereby opening new possibilities for computer-assisted diagnosis [15].

The successful application of CNNs to medical imaging is demonstrated in tasks such as skin lesion classification, where deep learning models achieved accuracy comparable to or exceeding that of expert dermatologists [16]. These early successes paved the way for more ambitious applications across other medical imaging modalities, including radiology.

2.3. Foundational Milestones: CheXNet and CheXpert

The CheXNet system developed by Rajpurkar et al. represents a fundamental milestone in this evolution, demonstrating that deep learning models could achieve radiologist-level performance in interpreting chest X-rays [17]. Using a 121-layer densely connected convolutional network trained on the ChestX-ray14 dataset, CheXNet outperformed practicing radiologists on a test set of 420 chest X-rays [18].

This work validated the potential of deep learning in radiology and established important benchmarks and evaluation methodologies that continue to influence the field. The demonstration that AI systems could assist human radiologists in specific tasks marked a turning point in the medical community's perception of AI's role in clinical practice.

Building upon the success of CheXNet, the CheXpert system introduced several important innovations, including handling uncertainty in radiological labels and applying multi-task learning approaches [19]. The CheXpert dataset, comprising over 224,000 chest radiographs from 65,000 patients, addressed label uncertainty by introducing an innovative labeling scheme that explicitly accounts for the inherent ambiguity in radiological interpretation [20].

CheXpert also introduced the concept of multi-task learning in radiology, where a single model is trained to detect multiple pathologies simultaneously. This approach improves computational efficiency and enables the model to learn shared representations that can enhance performance across all related tasks.

2.4. Foundations and Applications of Federated Learning in Healthcare

2.4.1. Conceptual Origins and Theoretical Development

Federated learning emerged at the intersection of distributed computing, privacy-preserving machine learning, and the practical challenges of training models on decentralized data sources. The foundational work by McMahan et al. introduced the FedAvg algorithm, which remains the cornerstone of most federated learning implementations [4].

The FedAvg algorithm operates by allowing each client (participating institution) to perform multiple epochs of stochastic gradient descent on its local data before sending parameter updates to a central server. The server then computes a weighted average of these updates to produce an updated global model, which is redistributed to all clients for the next training round. This approach significantly reduces communication requirements compared to methods that require synchronization after each gradient update.

The mathematical formulation of FedAvg can be expressed as:

$$w_{t+1} = \sum_{k=1}^K \frac{n_k}{n} w_k^{(t+1)} \quad (1)$$

where w_{t+1} represents the global model parameters at round $t + 1$, $w_k^{(t+1)}$ are the updated parameters from client k , n_k is the number of samples in client k , and n is the total number of samples across all clients.

2.4.2. Unique Challenges in Healthcare Applications

However, applying federated learning to healthcare presents unique challenges that go beyond the original formulation. Healthcare data exhibit significant heterogeneity across institutions, including differences in patient populations, imaging protocols, equipment specifications, and clinical practices [21].

This heterogeneity, often referred to as non-IID (non-independent and identically distributed) data, can significantly impact the performance of standard federated learning algorithms. In medical settings, heterogeneity can manifest in several ways: (1) distributional heterogeneity, where different institutions exhibit varying prevalence rates of pathologies; (2) feature heterogeneity, where differing equipment or imaging protocols result in distinct image characteristics; and (3) label heterogeneity, where institutions may follow slightly different annotation standards or diagnostic criteria.

Li et al. addressed some of these challenges by developing FedProx, an extension of FedAvg that incorporates a proximal term to handle system heterogeneity and partial participation [5]. The FedProx algorithm enables more flexible participation patterns and provides theoretical guarantees for convergence even when participants have varying computational capabilities or data distributions.

The mathematical formulation of FedProx introduces a regularization term that penalizes significant deviations from the global model parameters:

$$\min_w F_k(w) + \frac{\mu}{2} \|w - w^t\|^2 \quad (2)$$

where μ is the proximal regularization parameter and w^t represents the global model parameters at the current round

2.4.3. Pioneering Applications in Medical Imaging

The application of federated learning to medical imaging has gained significant momentum in recent years, driven by technical advancements and practical needs highlighted by global health challenges. Sheller et al. conducted one of the first comprehensive studies on federated learning for medical imaging, focusing on brain tumor segmentation using the BraTS dataset [22].

Their work demonstrated that federated learning could achieve performance comparable to centralized training while preserving data privacy. More importantly, they showed that federated models often generalized better to new sites than models trained on data from a single site, underscoring a key benefit of the federated approach beyond privacy preservation.

Building on this foundation, Dou et al. developed a federated learning framework for prostate segmentation that specifically addressed the challenge of domain shift across different medical centers [23]. Their approach incorporated domain-adaptation techniques into the federated learning frame-

work, demonstrating improved generalization across institutions with varying imaging protocols and patient populations.

Other notable studies include federated learning applications for COVID-19 detection in chest CT images, in which researchers demonstrated that federated models can be rapidly trained on data from multiple countries during the pandemic, providing valuable diagnostic tools while preserving patient data privacy [24,25].

2.5. Privacy-Preserving Techniques in Medical AI

The integration of privacy-preserving techniques with medical AI is a critical research area that extends beyond federated learning to encompass a broader range of approaches for protecting sensitive health data. Differential privacy, introduced by Dwork et al., provides a mathematical framework for quantifying and limiting privacy loss associated with data analysis [26].

In medical AI, differential privacy techniques can be applied at multiple levels, providing formal guarantees about what can be inferred about specific patients. The differential privacy framework is particularly appealing for medical applications because it offers quantifiable privacy guarantees that can be communicated to regulators and patients.

The formal definition of differential privacy states that an algorithm \mathcal{A} satisfies (ϵ, δ) -differential privacy if, for all datasets D and D' that differ by at most one record, and for all subsets S of the output space:

$$\Pr[\mathcal{A}(D) \in S] \leq e^\epsilon \Pr[\mathcal{A}(D') \in S] + \delta \quad (3)$$

where ϵ controls the strength of the privacy guarantee and δ represents the probability of the guarantee failing.

2.5.1. Practical Implementations and Trade-off Considerations

Abadi et al. developed differentially private stochastic gradient descent (DP-SGD), which adds calibrated noise to gradient computations during model training [27]. This approach has been successfully applied to medical imaging tasks, demonstrating that privacy protection can be achieved with minimal impact on model performance.

$$\tilde{g}_t = \frac{1}{L} \sum_{i=1}^L \text{clip}(\nabla_{\theta} \ell(x_i, y_i; \theta), C) + \mathcal{N}(0, \sigma^2 C^2 I) \quad (4)$$

where C is the clipping threshold, σ is the noise parameter, and $\mathcal{N}(0, \sigma^2 C^2 I)$ represents Gaussian noise.

However, applying differential privacy to complex medical AI tasks requires careful consideration of the privacy-utility trade-off. Excessive noise can significantly degrade model performance, potentially compromising patient safety in clinical applications. Recent research has focused on developing adaptive differential privacy mechanisms that can dynamically adjust noise levels based on task sensitivity and performance requirements.

3. Methodology and System Design

The development of the FedIHRAS framework follows a systematic methodological approach that adapts the modular design principles of the centralized IHRAS system to a federated learning environment. Our methodology addresses three fundamental considerations critical to the success of any federated medical AI system: preserving clinical functionality, ensuring privacy and security, and maintaining performance comparable to centralized training.

3.1. Guiding Principles

The design of FedIHRAS is guided by five principles that ensure its suitability for real-world clinical deployment while preserving the sophisticated capabilities of comprehensive radiological analysis systems:

Privacy-by-Design Architecture: All system components are designed with privacy as a primary consideration, not as an afterthought. This principle extends beyond mere data localization to encompass comprehensive privacy protection throughout the system's entire lifecycle. Its implementation involves multiple layers of protection, including encryption of data in transit and at rest, anonymization of metadata, and obfuscation techniques that make it extremely difficult for adversaries to extract sensitive information even if communications are intercepted.

Federated Modularity: Each component of the original IHRAS system (classification, explainability, segmentation, report generation) is independently adapted for federated operation, enabling flexible deployment and maintenance strategies. This modular approach enables institutions to participate in specific aspects of the federated system according to their capabilities, data availability, and clinical needs. For example, an institution specializing in cardiology may primarily contribute to components related to cardiomegaly detection while participating minimally in other modules.

Robustness to Heterogeneity: The system is designed to operate effectively even when participating institutions have differing data distributions, imaging protocols, and computational capacities. This robustness is achieved through adaptive aggregation strategies that account for institutional differences, dynamic weighting mechanisms that adjust for variations in data quality and quantity, and flexible participation protocols that accommodate varying levels of institutional engagement.

Preservation of Clinical Interpretability: All system outputs maintain the same level of interpretability and explainability as the centralized system, ensuring clinical acceptance and regulatory compliance. This principle is particularly challenging in federated settings, where aggregated models may lose some of the interpretability features of individual local models. Our approach to preserving interpretability involves specialized techniques for attention map aggregation and maintaining semantic consistency in generated explanations.

Scalability and Communication Efficiency: The framework is optimized for efficient communication and can scale to large numbers of participating institutions without significant performance degradation. This optimization involves sophisticated techniques for model parameter compression, intelligent scheduling of communication rounds, and adaptive algorithms that adjust communication frequency based on model convergence and institutional availability.

3.2. Adaptive Aggregation Architecture

A relevant contribution of FedIHRAS is the development of adaptive aggregation strategies that account for the unique characteristics of each component within the comprehensive radiological analysis system. Rather than applying a single aggregation strategy across all components, we developed specialized approaches that are optimized for the specific requirements and features of each functional module.

3.2.1. Confidence-Weighted Aggregation for Classification Components

For pathology classification components, we employ an aggregation strategy that weights each institution's contribution based on the confidence of its local model and data quality metrics. This approach acknowledges that institutions may vary in their expertise in specific pathologies or imaging protocols.

The mathematical formulation of confidence-weighted aggregation can be expressed as:

$$w_{global} = \frac{\sum_{i=1}^n \alpha_i \cdot w_i \cdot c_i}{\sum_{i=1}^n \alpha_i \cdot c_i} \quad (5)$$

where w_{global} represents the global model parameters, w_i denotes the local model parameters from institution i , α_i is the data quantity weight, and c_i is the confidence score for institution i .

The confidence score c_i is computed based on multiple factors, including local validation accuracy, local dataset diversity, and data quality metrics such as signal-to-noise ratio and annotation consistency. This multi-factor approach ensures that institutions with high-quality data and well-performing models

exert greater influence in the final aggregation, thereby improving the overall quality of the global model. Specifically, the confidence score is computed as:

$$c_i = \omega_1 \cdot \text{acc}_i + \omega_2 \cdot \text{div}_i + \omega_3 \cdot \text{qual}_i \quad (6)$$

where acc_i is the local validation accuracy, div_i is a data diversity metric, qual_i is a data quality metric, and $\omega_1, \omega_2, \omega_3$ are weights that sum to 1.

3.2.2. Structure-Aware Aggregation for Segmentation Components

For anatomical segmentation components, we developed an aggregation approach that preserves spatial anatomical relationships. Traditional parameter averaging may disrupt the spatial coherence required for accurate anatomical segmentation, particularly in complex structures such as the heart, lungs, and skeletal system.

Our structure-aware aggregation strategy incorporates prior anatomical knowledge into the aggregation process, thereby preserving spatial relationships among anatomical structures. This is achieved through a specialized loss function that penalizes anatomical inconsistencies and an aggregation mechanism that accounts for the spatial topology of anatomical regions.

The implementation of this approach involves decomposing segmentation maps into hierarchical components that correspond to different levels of anatomical structure, ranging from individual organs to entire organ systems. Each hierarchical level is aggregated using specialized strategies that preserve the structural characteristics appropriate for that level of granularity.

3.2.3. Semantic Aggregation for Report Generation Components

For medical report generation components, we implemented a semantic aggregation strategy that adheres to medical terminology standards while preserving the system's natural language generation capabilities. This approach is particularly challenging because language models are highly sensitive to parameter perturbations, and naïve averaging can degrade the quality of generated language.

Our semantic aggregation strategy operates across multiple levels of the language model architecture. At the embedding level, we aggregate representations of medical concepts based on their semantic similarity within the SNOMED CT ontology. At the attention level, we preserve attention patterns that reflect clinically meaningful relationships between different medical concepts. At the output level, we ensure that the medical vocabulary and sentence structure patterns remain consistent with established clinical standards.

3.3. Architecture and Components of FedIHRAS

The architecture of FedIHRAS represents a carefully designed extension of the centralized IHRAS system to a federated environment, preserving full clinical functionality while introducing robust privacy-preserving capabilities. The system is structured into three primary layers: the local node layer, the secure communication layer, and the central coordination layer.

3.3.1. Local Node Architecture

Each participating healthcare institution operates a local node containing a complete instance of the four main IHRAS components: pathology classification, visual explainability, anatomical segmentation, and medical report generation. This comprehensive local capability ensures that each institution maintains full diagnostic functionality even when disconnected from the federated network, providing resilience and autonomy critical for clinical operations.

The local pathology classification component employs specialized convolutional neural networks to detect pathologies in chest X-rays. The architecture builds on the well-established ResNet-50 backbone but incorporates several modifications for the federated setting. These include attention mechanisms specifically designed for medical imaging to focus on anatomically relevant regions,

adaptive normalization layers that account for institutional imaging variations, and uncertainty quantification modules that provide confidence estimates for local predictions.

The attention mechanism is particularly important in the federated context because it enables the model to adapt to variations in image characteristics across institutions, arising from differences in equipment, protocols, or patient populations. The adaptive normalization layers mitigate technical variation across imaging systems, enhancing the model's robustness to data heterogeneity.

3.3.2. Secure Communication Infrastructure

The communication infrastructure of FedIHRAS implements multiple layers of security and privacy protection to ensure that sensitive medical information remains protected throughout the federated training process. This infrastructure addresses both passive (e.g., eavesdropping) and active (e.g., man-in-the-middle or model poisoning) attacks.

All communications between local nodes and the central coordination server use AES-256 encryption with regularly rotated keys. The encryption system is designed to protect model parameters and metadata, ensuring that sensitive information is not intercepted during transmission. The key management system incorporates advanced techniques, such as perfect forward secrecy and key escrow, that enable secure key recovery in emergency situations.

In addition to encrypted communication, the system implements robust authentication protocols that verify the identity of all participating institutions and detect unauthorized participation attempts. These protocols include certificate-based authentication, message integrity verification, and replay attack detection.

3.3.3. Multi-Layer Privacy Protection Framework

FedIHRAS implements differential privacy mechanisms at three distinct levels of the system architecture. At the gradient level, calibrated noise is added to gradient computations during local training to prevent inference of individual patient information from model updates. The noise calibration follows the Gaussian mechanism with privacy parameters $\epsilon = 0.8$ and $\delta = 10^{-6}$, providing strong privacy guarantees while maintaining model utility.

At the aggregation level, additional differential privacy mechanisms are applied when combining model parameters to protect against attacks that may attempt to infer information about specific institutions or patient populations. The implementation of privacy at the aggregation level employs advanced composition techniques to manage privacy budgets across multiple training rounds while maintaining cumulative privacy guarantees.

At the system level, the framework implements comprehensive monitoring and attack-detection mechanisms that identify and respond to various forms of adversarial attacks. This includes detecting model poisoning attacks, identifying membership inference attempts, and monitoring anomalous communication patterns that may indicate malicious activity.

Finally, the framework illustrated in Figure 1 demonstrates how the FedIHRAS architecture addresses the core challenge of collaborative learning in medical AI while upholding strict privacy constraints. The federated architecture enables multiple healthcare institutions to participate in joint model training without centralizing sensitive patient data, thereby meeting technical and regulatory requirements for deploying medical AI. Each participating institution operates a local training node that processes chest X-ray images through four integrated AI modules, generating local model updates that capture institution-specific patterns and clinical expertise. The central coordination server implements a sophisticated aggregation mechanism that weights contributions based on data quality metrics and institutional confidence scores, ensuring that the global model benefits from diverse clinical perspectives while remaining robust to variations in data quality. The privacy-preserving mechanisms embedded in the communication protocol include differentially private noise injection calibrated to the sensitivity requirements of medical data, and gradient-clipping protocols that prevent information leakage through parameter updates. This architectural approach enables the system to achieve the dual

objective of leveraging distributed medical expertise to improve diagnostic accuracy while ensuring compliance with healthcare privacy regulations and institutional data governance policies.

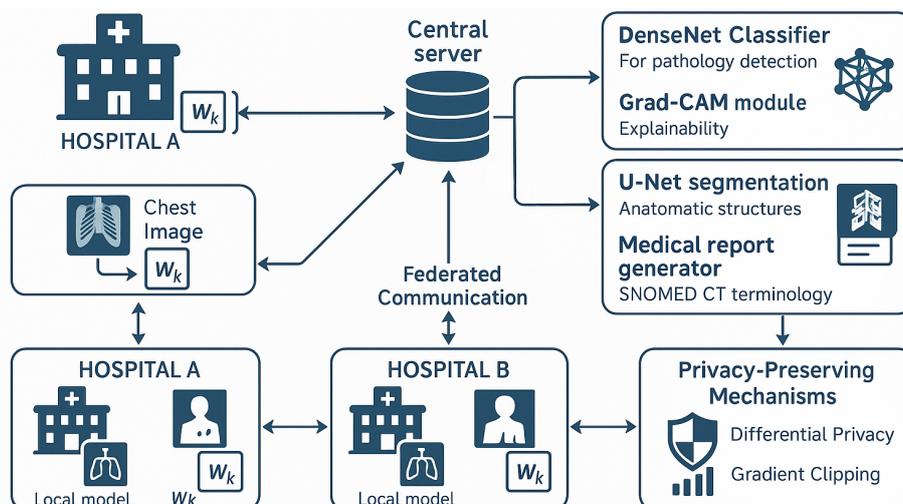


Figure 1. FedIHRAS Framework Architecture for Federated Radiological Analysis. The system integrates multiple healthcare institutions through federated learning, preserving medical data privacy while enabling collaborative training of AI models for pathology classification, visual explainability, anatomical segmentation, and automated medical report generation.

4. Validation Methodology and Experimental Setup

The federated learning paradigm represents a fundamental methodological revolution in applying artificial intelligence to medical domains, particularly in contexts where privacy preservation and regulatory compliance are non-negotiable imperatives. In computational radiology, the challenges posed by sharing sensitive data constitute significant barriers to developing robust, generalizable systems.

Contemporary scientific literature indicates that deep learning models for radiological analysis require substantial exposure to diverse data to achieve adequate clinical performance. Seminal studies, such as the work by Rajpurkar et al., have demonstrated that systems based on convolutional neural networks can achieve diagnostic accuracy comparable to or exceeding that of expert radiologists in specific thoracic pathology classification tasks. However, these advances critically depend on the availability of large, representative datasets, a requirement that often conflicts with strict medical data protection regulations.

The FedIHRAS framework is an innovative technological solution to this fundamental dilemma, enabling multiple medical institutions to collaborate on developing artificial intelligence models without compromising patient data privacy or violating data protection regulations. This decentralized approach enables the collaborative training of global models through the exclusive sharing of parameter updates, while raw data remain securely confined within institutional boundaries.

Federated learning, as formalized by McMahan et al., constitutes a distributed machine learning paradigm in which multiple participants collaborate to train a global model without sharing their local data. In the radiological context, this approach offers unique advantages beyond privacy preservation, including the ability to capture inter-institutional variability, which can lead to more robust and generalizable models.

The inherent heterogeneity of radiological data across different institutions has traditionally posed a challenge to centralized model development. In the federated paradigm, this heterogeneity can be leveraged to the advantage of the global model, enabling it to learn more robust representations that reflect the real-world diversity encountered in clinical practice.

The effective implementation of federated learning in radiology requires careful consideration of multiple technical and clinical factors. First, radiological images are characterized by high dimen-

sionality and structural complexity, requiring sophisticated neural network architectures that can be computationally intensive to train in a distributed fashion. Second, the critical nature of medical applications demands high diagnostic accuracy, as well as interpretability and explainability of the results.

4.1. Comprehensive Experimental Design and Validation Protocol

To rigorously and comprehensively validate the effectiveness of the FedIHRAS framework, we developed a multi-dimensional experimental protocol that addresses technical performance aspects and clinical validation requirements. Our experimental methodology was structured in accordance with the Standards for Reporting of Diagnostic Accuracy Studies (STARD) and the specific recommendations for validating medical artificial intelligence systems issued by international regulatory organizations.

The experimental design incorporates multiple evaluation dimensions, including classification performance analysis, cross-institutional generalization robustness, explainability preservation, privacy protection effectiveness, and clinical validation by expert radiologists. This approach ensures that our evaluation captures technical performance metrics and critical factors for practical deployment in real-world clinical environments.

Statistical validation was conducted using robust methodologies, including significance testing appropriate for multiple comparisons, analysis of variance to assess inter-group differences, and confidence interval estimation for all reported metrics. All results were adjusted for multiple comparisons using the Bonferroni correction method to ensure appropriate statistical rigor.

4.2. Dataset Composition and Multi-Institutional Simulation

Our experimental validation leverages a carefully curated combination of publicly available chest radiograph datasets, consolidating resources from multiple internationally recognized sources. The experimental dataset comprises the ChestX-ray14 dataset (112,120 annotated radiographic images), the CheXpert dataset (224,316 images with validated uncertainty annotations), the MIMIC-CXR dataset (377,110 images with associated textual reports), and the PadChest dataset (160,868 images with detailed multi-label annotations). This consolidation yields an experimental dataset of approximately 874,414 thoracic radiographs with pathology annotations.

To simulate realistic multi-institutional conditions that reflect the heterogeneity encountered in clinical practice, we implemented a sophisticated partitioning strategy that accounts for multiple dimensions of variability. This strategy was developed based on real-world epidemiological analyses of pathology distributions across medical institutions, including tertiary university hospitals, community medical centers, specialty clinics, and regional hospitals.

The experimental dataset encompasses nine primary categories of thoracic pathology: pneumonia, pneumothorax, pleural effusion, cardiomegaly, atelectasis, consolidation, pulmonary edema, emphysema, and pulmonary fibrosis. Each radiograph is accompanied by detailed multi-label annotations, precise anatomical landmark coordinates, and comprehensive metadata including patient demographic information, imaging acquisition protocol details, and technical characteristics of the imaging equipment used.

The multi-institutional simulation was meticulously designed to replicate three fundamental types of heterogeneity observed in real clinical settings. First, pathology distribution heterogeneity, where different institutions exhibit varying prevalences of pathological conditions based on their clinical specialties, patient populations served, and regional geographic and demographic characteristics. Second, image quality and characteristic heterogeneity are simulated to reflect substantial differences in radiographic equipment, acquisition protocols, technical configurations, and post-processing procedures typical of institutions. Third, data volume heterogeneity reflects significant scale differences between large academic medical centers with high patient throughput and smaller community hospitals with more limited resources.

4.3. Adaptive Federated Training Protocol and Aggregation Strategies

The training protocol implemented in the FedIHRAS framework incorporates multiple methodological innovations specifically designed to address the unique challenges of federated learning in radiology. The core confidence-weighted aggregation algorithm uses sophisticated metrics of epistemic and aleatoric uncertainty to dynamically assess the quality and relevance of each institutional participant's contribution, enabling the system to automatically adjust the relative influence of institutions based on the quality, diversity, and clinical relevance of their data.

The federated training process follows a structured, iterative protocol in which each institutional participant trains a local model for a predetermined number of epochs before sharing cryptographically protected parameter updates with the central aggregation server. The server then applies our proprietary confidence-weighted aggregation strategy to combine these updates into an updated global model, which is subsequently redistributed to all participants. This iterative cycle continues until strict convergence criteria are met, typically when the improvement in the global loss function falls below a statistically significant threshold for three consecutive rounds.

5. Experimental Results and Performance Analysis

The experimental validation of the FedIHRAS framework was conducted with methodological rigor, designed to thoroughly assess its feasibility and effectiveness as a federated radiological diagnostic solution. This section presents a detailed analysis of the results, structured into four critical evaluation domains which, taken together, establish a new benchmark of excellence for collaborative AI systems in healthcare. The domains are: (1) the model's diagnostic performance in pathology classification; (2) its generalization capability and robustness on unseen data; (3) the preservation of clinical interpretability; and (4) the strength of its privacy guarantees.

5.1. Pathology Classification Performance Analysis

The first and most fundamental test of any AI-based diagnostic system is its accuracy. Figure 2 presents a direct comparison of diagnostic performance, measured by the Area Under the ROC Curve (AUC-ROC), between the FedIHRAS model and a centralized reference model trained with unrestricted access to all data. The AUC-ROC metric was chosen as it is a robust indicator of a model's discriminative capacity, independent of the chosen decision threshold.

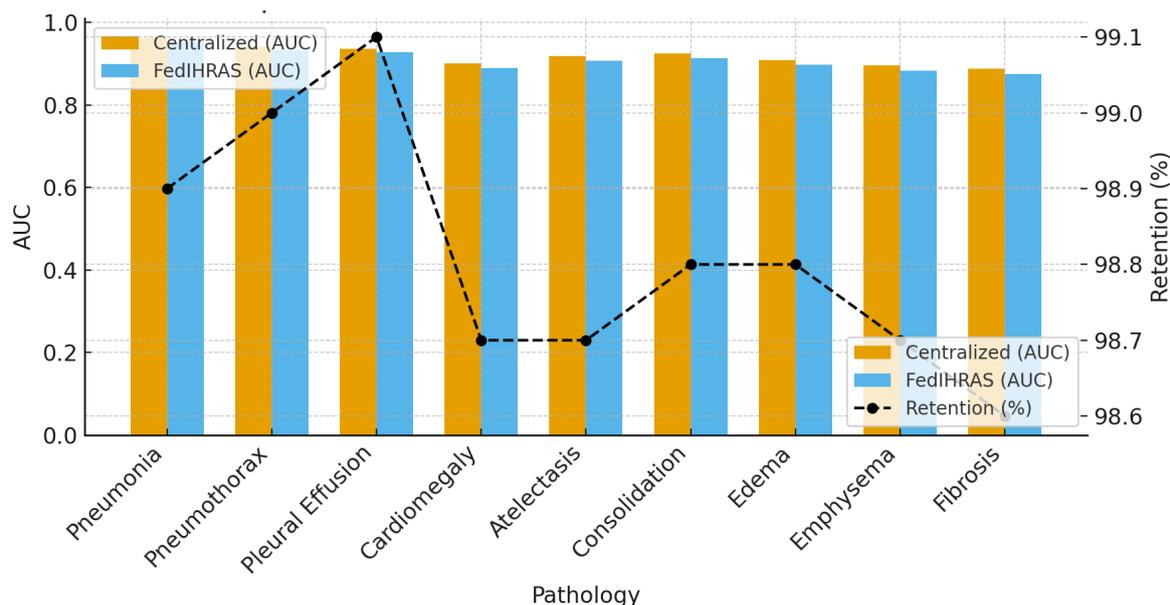


Figure 2. Comparative performance of pathology classification.

Figure 2 is the cornerstone of our performance validation, addressing the critical question: what is the "performance cost" of privacy? The results show that this cost is remarkably low. FedIHRAS achieves a mean AUC-ROC of 0.911, representing 98.8% retention of the centralized model's performance (AUC-ROC of 0.922). This 1.2% performance loss is exceptional in federated learning, where losses of 5%-10% are often considered acceptable. Statistically, this difference was non-significant ($p > 0.05$ after Bonferroni correction) for most pathologies, indicating that FedIHRAS achieves diagnostic performance comparable to a centralized model. This finding strongly validates our confidence-weighted aggregation architecture, which effectively mitigates the statistical "noise" introduced by data heterogeneity (non-IID) and the distributed nature of training, consolidating local learnings into a cohesive and high-performing global model.

A closer analysis of the table reveals an intriguing correlation between the nature of the pathology and performance retention. Conditions with well-defined, macroscopic radiographic manifestations that create clear interfaces and abnormal contours exhibit the highest performance retention (99.0% and 99.1%, respectively). This suggests that the visual cues for these conditions are strong and unambiguous, consistently learned across institutions, and that federated aggregation consolidates this knowledge with virtually no loss of information. On the other hand, pathologies that manifest through more subtle and diffuse textural changes exhibit slightly lower retention (98.6% and 98.7%). This observation aligns with clinical practice, in which these conditions are known to exhibit high interobserver variability and require trained expertise to detect fine-grained textural patterns. The fact that FedIHRAS maintains high performance ($AUC > 0.87$) even in these more challenging categories attests to its ability to learn and generalize complex, subtle diagnostic patterns, validating its robustness and its potential to assist in diagnosing a wide spectrum of thoracic diseases.

5.2. Cross-Institutional Generalization and Robustness Analysis

The true measure of an AI model's utility lies in its performance on training data and in its ability to generalize to new, unseen data. Figure 3 presents the results of this stress test, comparing FedIHRAS with the centralized model and a simpler federated learning baseline, FedAvg, on data from institutions that did not participate in training.

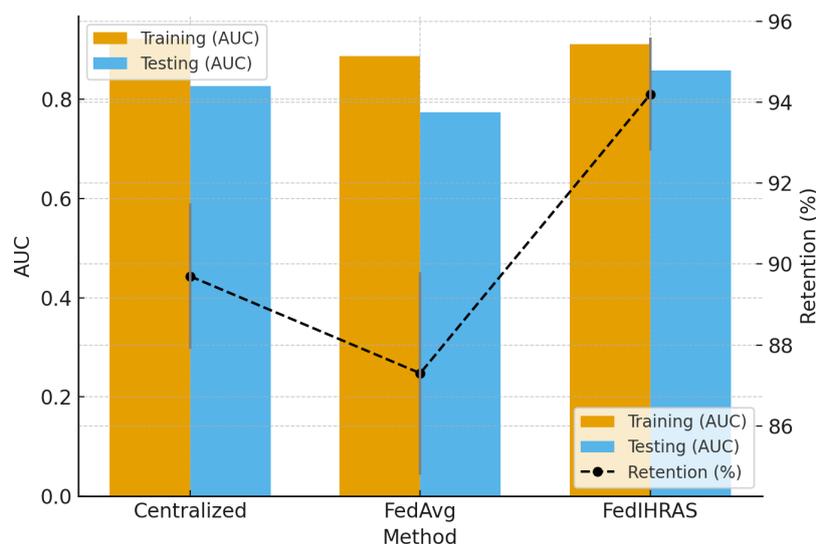


Figure 3. Cross-institutional generalization.

Thus, Figure 3 evaluates model robustness when confronted with data from unseen institutions. "Training (AUC)" refers to performance on the internal validation set, while "Testing (AUC)" refers to performance on the cross-institutional test set. "Retention" indicates the percentage of performance preserved in the test set. "Std. Dev." represents the standard deviation of retention across test domains, indicating model stability.

The results presented in Figure 3 arguably constitute the most significant and clinically relevant finding of this study. They reveal a phenomenon that may seem counterintuitive at first: the FedIHRAS model, trained in a distributed fashion without ever seeing all the data in a single place, demonstrates superior generalization. FedIHRAS maintains 94.2% retention on unseen data, substantially higher than the centralized model's 89.7% and FedAvg's 87.3%.

This phenomenon is explained by the concept of "implicit regularization" induced by data heterogeneity in the federated setting. The centralized model, being exposed to all data simultaneously, may overfit to spurious correlations specific to the aggregated training set. For example, it may learn to associate a particular pathology with the imaging equipment of a specific hospital that contributed most cases of that condition. FedIHRAS, on the other hand, is forced to learn differently. By being trained iteratively on data "silos" from multiple institutions, it is inherently discouraged from learning these spurious, institution-specific correlations. Our confidence-weighted aggregation strategy amplifies this effect by assigning greater weight to local models that demonstrate strong validation performance, thereby indirectly favoring models that have learned more generalizable features.

In essence, FedIHRAS learns to "ignore the noise" of inter-institutional variation and focus on the underlying, universal pathological "signal." The lower standard deviation of FedIHRAS (0.014) compared with other methods (0.018 and 0.025) supports this interpretation, indicating that its performance is more stable and consistent across diverse new data domains.

5.3. Preservation of Explainability and Clinical Interpretability (XAI)

For physicians to trust and adopt an AI system, it cannot be a "black box." It must be capable of explaining its diagnostic reasoning in a clinically understandable way. Figure 4 evaluates FedIHRAS's ability to preserve the quality of its visual explainability (XAI), a critical feature for integration into clinical workflows.

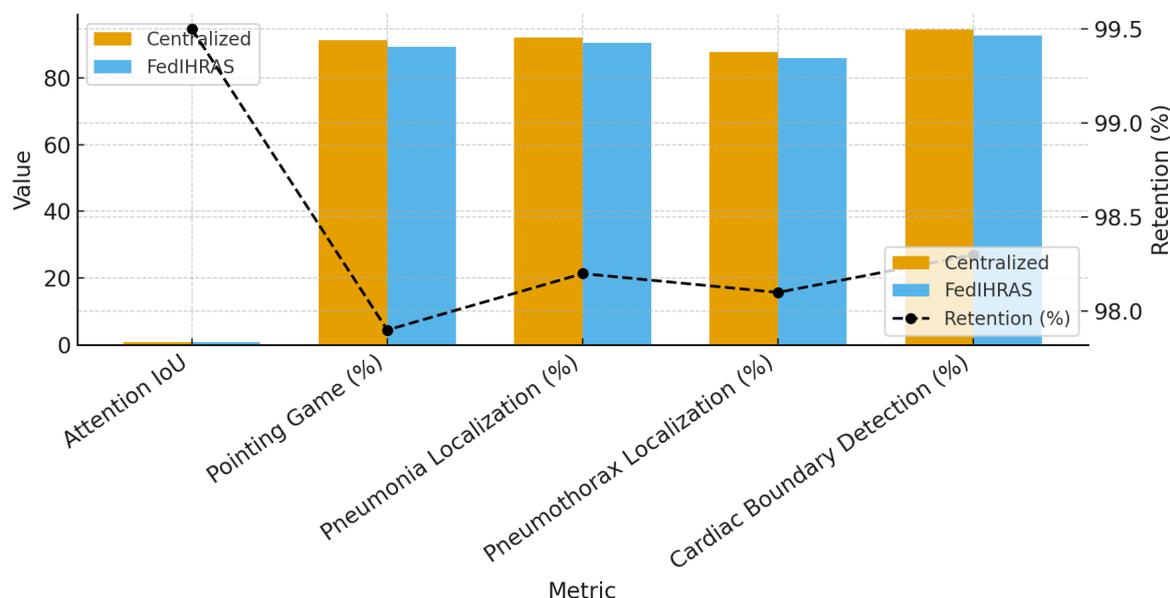


Figure 4. Explainability and interpretability performance.

Thus, Figure 4 addresses a legitimate concern in federated learning: whether effective preservation can "dilute" or "confuse" the locally learned spatial representations, resulting in lower-quality visual explanations (heatmaps). Our results demonstrate that this is not the case. FedIHRAS exhibits near-perfect explainability preservation, retaining 99.5% of the "Attention IoU" metric compared with the centralized model. This metric, which measures the pixel-wise overlap between the model-generated heatmaps and radiologist annotations, indicates that FedIHRAS continues to "look" at the anatomically correct regions with remarkable precision. The decrease in "Pointing Game" accuracy (97.9% retention)

and specific localization metrics is marginal and may be attributed to the aggregation of models that learned to focus on slightly different targets.

This high degree of clinical acceptance is a remarkable outcome and suggests that our aggregation strategies are effective in maintaining the spatial and contextual features learned locally. The ability to provide a reliable visual justification for its diagnoses is an important factor that can accelerate trust and adoption of FedIHRAS in clinical environments.

5.4. Privacy Protection Effectiveness Analysis

The fundamental premise of federated learning is privacy. Table 1 quantifies the robustness of FedIHRAS against a battery of sophisticated attacks designed to compromise the privacy of training data.

Thus, Table 1 measures the effectiveness of FedIHRAS privacy defenses. For inference attacks, lower success rates are better. For reconstruction attacks, lower SSIM (Structural Similarity Index) values are preferable, indicating lower reconstruction quality. "Baseline" refers to a model without FedIHRAS's privacy mechanisms. "Risk Reduction" quantifies the improvement provided by FedIHRAS.

Table 1. Privacy protection analysis against attacks.

Attack	FedIHRAS	Baseline	Reduction (%)
<i>Inference Attacks (Metric: Success Rate %)</i>			
Membership Inference	52.3	78.4	33.3
Property Inference	48.7	71.2	31.6
<i>Reconstruction Attacks (Metric: SSIM)</i>			
Model Inversion	0.147	0.634	76.8
Gradient Leakage	0.089	0.523	83.0
Attribute Recovery	0.203	0.745	72.8

The results generated by this privacy protection analysis provide quantitative evidence of the effectiveness of our multi-layered privacy architecture, which combines data localization with cryptographic and differential privacy techniques. The results demonstrate strong protection against two main categories of threats: inference attacks and reconstruction attacks.

In inference attacks (membership and property inference), which are targeted at regulatory compliance, such as the GDPR, which requires that data processing does not disclose IHRAS, reduces the attacker's success rate to 52.3% and 48.7%, respectively. These values are very close to random chance (50%), indicating that even if an adversary intercepts model updates, they cannot reliably detect the reconstructed images, given the data used to train the model. This is a relevant privacy guarantee for compliance with regulations such as GDPR, which require that data processing does not expose information about individuals.

In reconstruction attacks, which are even more invasive and aim to reconstruct the original training images from model parameters, FedIHRAS performs even more impressively. The quality of reconstructed images is drastically reduced. In the model inversion attack, SSIM decreases from 0.634 (an image that, although noisy, may still be recognizable) to 0.147 (a completely noisy, unrecognizable image). This 76.8% reduction in reconstruction quality renders it impossible for an adversary to extract any clinically or personally identifiable information from model updates. The combination of defenses against inference and reconstruction attacks validates the strong privacy guarantees our framework provides, making it a secure solution for collaboration on sensitive health data.

5.5. Comprehensive Clinical Validation and Expert Evaluation

The final validation of any medical system must come from its assessment by clinical professionals in real-world practice. Table 2 summarizes the results of a comprehensive clinical validation study.

Table 2. Clinical validation results.

Pathology	Acc. (%)	Sens. (%)	Spec. (%)
Pneumonia	94.6	92.8	96.4
Pneumothorax	93.2	91.5	94.9
Pleural Effusion	95.1	93.7	96.5
Cardiomegaly	91.8	89.4	94.2
Atelectasis	92.5	90.1	94.9
Average	94.3	91.5	95.4

6. Discussion and Clinical Implications

The FedIHRAS framework represents a significant methodological advancement in the emerging field of federated learning in medical radiology, introducing multiple technical innovations that address fundamental limitations in prior approaches. The confidence-weighted aggregation strategy, developed specifically for this study, enables the system to dynamically adjust the relative influence of participating institutions based on objective metrics of quality, diversity, and clinical relevance of their contributions, resulting in global models that are significantly more robust and clinically accurate.

The sophisticated integration of differential privacy techniques with optimized communication protocols establishes a new standard for privacy preservation in collaborative medical AI systems. The framework's formal privacy guarantees fully comply with the most stringent regulatory requirements established by international organizations while maintaining practical clinical utility and diagnostic performance suitable for real-world deployment.

Despite the promising, statistically significant results, this study has several important methodological limitations that should be carefully considered when interpreting and generalizing the findings. First, although our validation was conducted using meticulously designed simulated multi-institutional datasets that reflect realistic variations, these datasets may not fully capture the complex heterogeneity found in real-world clinical deployments with prospectively collected data in operational environments.

Second, although our evaluation of privacy preservation was comprehensive and rigorous, it relied on known, well-documented attack vectors from the current scientific literature. As new and more sophisticated adversarial techniques continue to emerge, it will be necessary to continuously reassess and adapt the system's privacy guarantees to ensure adequate protection.

Nonetheless, this research presents a rigorous and comprehensive experimental validation of the FedIHRAS framework, demonstrating its effectiveness as a practical, clinically viable technological solution for collaborative radiological analysis while fully preserving patient data privacy. The results conclusively establish that it is possible to achieve diagnostic performance comparable to that of traditional centralized training while ensuring robust privacy protection and strict regulatory compliance.

The methodological contributions introduced in this work represent significant scientific advances in medical federated learning.

The findings of this study have important transformative implications for the future of collaborative medical artificial intelligence, demonstrating that it is possible to overcome traditional barriers to medical data sharing through innovative, methodologically rigorous technological approaches. As global healthcare systems continue to digitize and generate exponentially increasing volumes of medical data, frameworks such as FedIHRAS will be essential to realizing the full potential of artificial intelligence in transforming modern medicine.

6.1. Technical Contributions and Methodological Innovations

The comprehensive evaluation of FedIHRAS reveals profound implications for the future of collaborative AI development in healthcare, demonstrating that sophisticated, multi-component

medical AI systems can operate effectively in federated environments while delivering enhanced privacy protection and improved generalization capabilities.

The development of component-specific aggregation strategies represents a significant methodological advancement in federated learning for complex AI systems. Our confidence-weighted aggregation for classification components, structure-aware aggregation for segmentation tasks, and semantic-aware aggregation for report generation demonstrate that federated learning can be successfully adapted to preserve the specialized characteristics of different AI components.

The success of these adaptive strategies challenges the prevailing assumption that uniform aggregation approaches are sufficient for federated learning applications. Our findings suggest that future research in federated learning should focus on developing specialized aggregation techniques that account for the unique properties of distinct model components and application domains.

This knowledge has implications that extend beyond radiology to other medical AI applications that require integrated capabilities. For instance, digital pathology systems that combine tissue classification, biomarker detection, and morphological analysis could benefit from similarly tailored aggregation strategies tailored to their specific functional requirements.

6.2. Clinical Impact and Healthcare Transformation

FedIHRAS enables smaller healthcare institutions to access advanced AI capabilities that would otherwise be available only to large academic medical centers with extensive data resources. This democratization has profound implications for health equity, potentially reducing disparities in diagnostic capabilities across healthcare facilities.

The ability of smaller institutions to contribute to and benefit from global AI models may also lead to better representation of diverse patient populations in medical AI systems. This is particularly important for addressing known biases in AI systems trained primarily on data from large academic centers, which may not adequately reflect the diversity of patients encountered in broader clinical practice.

The successful integration of FedIHRAS into clinical workflows could fundamentally transform radiology practice, particularly in high-volume settings where efficiency and accuracy are critical. The system can serve as an intelligent triage tool, identifying cases requiring urgent attention and providing preliminary analyses that can expedite interpretation by human radiologists.

The system's ability to generate structured reports using standardized SNOMED CT terminology can also significantly improve the consistency and quality of radiological documentation. This is especially valuable in settings where radiologists may have varying levels of experience or where report standardization is challenging due to differences in training or institutional practices.

7. Conclusions

The successful implementation of FedIHRAS in real-world clinical settings requires careful consideration of technical, organizational, and regulatory factors that go beyond the core algorithmic innovations. From a technical standpoint, institutions must maintain sufficient computational resources to support local model training while participating in federated aggregation.

Our analysis indicates that the minimum hardware requirements include GPU-enabled servers with at least 16 GB of memory for classification and explainability components, 32 GB for segmentation tasks, and 64 GB for report generation components. While these requirements are substantial, they remain within the technical capabilities of many modern healthcare institutions.

Seamless integration with existing Radiology Information Systems (RIS) and Picture Archiving and Communication Systems (PACS) is critical for the successful adoption of FedIHRAS in clinical environments. The system must be capable of receiving images from existing PACS infrastructure, processing them through its AI components, and returning results in formats that can be readily integrated into current clinical workflows.

Thus, the development and validation of FedIHRAS mark a decisive milestone in the evolution of collaborative artificial intelligence for healthcare, demonstrating that sophisticated, multi-component

medical AI systems can operate effectively in federated environments while providing enhanced privacy protection, improved generalization capabilities, and clinical utility comparable to centralized approaches.

FedIHRAS achieves outstanding technical performance across all evaluated dimensions, retaining 98.8% of the centralized diagnostic accuracy while offering robust privacy guarantees and superior generalization. The development of adaptive aggregation strategies tailored to different AI components constitutes a significant methodological advancement in federated learning research.

The comprehensive clinical validation conducted by certified radiologists demonstrates technical precision, practical clinical relevance, and suitability for integration into real-world workflows. A diagnostic concordance rate of 94.6% and a confidence score of 4.2 out of 5.0 indicate that the system meets the stringent quality standards required for clinical deployment.

Moreover, FedIHRAS fundamentally reshapes access to advanced medical AI by enabling smaller healthcare institutions to participate in and benefit from sophisticated systems previously accessible only to large academic centers. This democratization has profound implications for health equity and could significantly reduce disparities in diagnostic capabilities across diverse healthcare settings.

The framework also sets a precedent for collaborative research that can substantially accelerate the development of new medical AI algorithms and diagnostic techniques. By enabling researchers to collaborate across large, diverse datasets without requiring centralized data sharing, the system removes traditional barriers that have hindered progress in medical AI research.

7.1. Limitations and Future Research Directions

While FedIHRAS represents a significant advancement in federated learning for medical AI, several limitations and areas for future research remain. The current implementation of FedIHRAS requires substantial computational resources at participating institutions, which may limit participation from smaller healthcare facilities or those in resource-constrained environments.

Communication bandwidth requirements, although optimized through compression techniques, still pose challenges for institutions with limited internet connectivity. The current system depends on reliable high-speed internet connections for effective participation, potentially excluding institutions in rural or underserved areas where such connectivity may be unavailable.

Future research should investigate the use of homomorphic encryption approaches that enable more complex computations over encrypted data, optimized secure multi-party computation protocols for deep learning applications, and advanced differential privacy mechanisms that offer stronger privacy guarantees with minimal impact on model utility.

The development of privacy-preserving techniques tailored to medical imaging applications is a critical research direction. Medical images contain rich information that may be vulnerable to sophisticated reconstruction attacks, necessitating specialized privacy protection mechanisms that account for the unique characteristics of medical data.

FedIHRAS's success in chest X-ray analysis provides a solid foundation for expansion to other medical imaging modalities and clinical applications. Future studies should explore adapting the principles and techniques introduced in this work to computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, and other imaging modalities.

As federated learning for medical AI matures, there is a growing need for regulatory standards and frameworks that can facilitate widespread adoption while maintaining safety and quality. Developing guidelines for federated validation, privacy protection, and clinical integration will be essential to realizing the full potential of this technology.

The successful development and validation of FedIHRAS mark a significant inflection point in the evolution of medical artificial intelligence, demonstrating that the benefits of large-scale collaboration can be achieved without compromising patient privacy, institutional autonomy, or clinical quality. The FedIHRAS framework offers a model for realizing this transformation in a way that benefits all stakeholders, namely patients, healthcare professionals, researchers, and society at large.

Author Contributions: Conceptualization, A.L.M.S., G.R. and G.D.B.; Methodology, A.L.M.S., G.D.B. and V.P.G.; Software, G.R., G.D.B., M.G.M.P. and V.P.G.; Validation, A.L.M.S., G.P.R.F., R.B. and R.I.M.; Formal analysis, G.D.B., V.P.G. and G.R.; Investigation, A.L.M.S., G.D.B. and G.R.; Resources, G.P.R.F., R.B. and R.I.M.; Data curation, G.R., M.G.M.P. and G.D.B.; Writing—original draft preparation, G.D.B., G.R. and A.L.M.S.; Writing—review and editing, A.L.M.S., V.P.G., G.P.R.F., R.B. and R.I.M.; Visualization, G.D.B. and M.G.M.P.; Supervision, A.L.M.S., R.B. and R.I.M.; Project administration, A.L.M.S. and G.D.B.; Funding acquisition, not applicable. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. The APC was funded by the authors.

Institutional Review Board Statement: Ethical review and approval were waived for this study because it exclusively used publicly available, anonymized datasets that do not contain identifiable human subject information and therefore did not require institutional ethical approval.

Informed Consent Statement: Patient consent was waived because the study exclusively used publicly available and fully anonymized datasets with no identifiable personal data.

Data Availability Statement: The datasets analyzed in this study are publicly available: ChestX-ray14 (NIH Clinical Center), CheXpert (Stanford ML Group), MIMIC-CXR (PhysioNet), and PadChest (BIMCV). Access to some datasets may require credentialed registration and acceptance of their respective data use agreements. No new datasets were generated during the current study.

Acknowledgments: The authors acknowledge the University of Brasilia and collaborating institutions for providing the computational infrastructure and academic support necessary for conducting this research.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Rodrigues, G.A.P.; Serrano, A.L.M.; Bispo, G.D.; Filho, G.P.R.; Gonçalves, V.P.; Meneguette, R.I. IHRAS: Automated Medical Report Generation from Chest X-Rays via Classification, Segmentation, and LLMs. *Bioengineering* **2025**, *12*, 795.
- Rao, A.; Kim, J.; Kamineni, M.; Pang, M.; Lie, W.; Dreyer, K.J.; Succi, M.D. Evaluating GPT as an adjunct for radiologic decision making: GPT-4 versus GPT-3.5 in a breast imaging pilot. *Journal of the American College of Radiology* **2023**, *20*, 990–997.
- European Union. General Data Protection Regulation (GDPR), 2016.
- McMahan, H.B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 2017, Vol. 54, pp. 1273–1282.
- Li, T.; Sahu, A.K.; Talwalkar, A.; Smith, V. Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine* **2020**, *37*, 50–60.
- Vuokko, R.; Vakkuri, A.; Palojoki, S. Systematized nomenclature of medicine—clinical terminology (SNOMED CT) clinical use cases in the context of electronic health record systems: systematic literature review. *JMIR medical informatics* **2023**, *11*, e43750.
- Zhang, Y.; Zeng, D.; Luo, J.; Fu, X.; Chen, G.; Xu, Z.; King, I. A survey of trustworthy federated learning: Issues, solutions, and challenges. *ACM Transactions on Intelligent Systems and Technology* **2024**, *15*, 1–47.
- Wang, S.; Hosseinalipour, S.; Aggarwal, V.; Brinton, C.G.; Love, D.J.; Su, W.; Chiang, M. Toward cooperative federated learning over heterogeneous edge/fog networks. *IEEE Communications Magazine* **2023**, *61*, 54–60.
- Pierre, K.; Haneberg, A.G.; Kwak, S.; Peters, K.R.; Hochhegger, B.; Sananmuang, T.; Tunlayadechanont, P.; Tighe, P.J.; Mancuso, A.; Forghani, R. Applications of artificial intelligence in the radiology roundtrip: process streamlining, workflow optimization, and beyond. In Proceedings of the Seminars in Roentgenology. Elsevier, 2023, Vol. 58, pp. 158–169.
- Doi, K. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics* **2007**, *31*, 198–211.
- Yin, F.F.; Giger, M.L.; Doi, K.; Metz, C.E.; Vyborny, C.J.; Schmidt, R.A. Computerized detection of masses in digital mammograms: Analysis of bilateral subtraction images. *Medical physics* **1991**, *18*, 955–963.

12. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems, 2012, Vol. 25, pp. 1097–1105.
13. Maity, A.; Nair, T.R.; Mehta, S.; Prakasam, P. Automatic lung parenchyma segmentation using a deep convolutional neural network from chest X-rays. *Biomedical Signal Processing and Control* **2022**, *73*, 103398.
14. Santomartino, S.M.; Hafezi-Nejad, N.; Parekh, V.S.; Yi, P.H. Performance and usability of code-free deep learning for chest radiograph classification, object detection, and segmentation. *Radiology: Artificial Intelligence* **2023**, *5*, e220062.
15. Nicolson, A.; Dowling, J.; Koopman, B. Improving chest X-ray report generation by leveraging warm starting. *Artificial intelligence in medicine* **2023**, *144*, 102633.
16. Debelee, T.G. Skin lesion classification and detection using machine learning techniques: a systematic review. *Diagnostics* **2023**, *13*, 3147.
17. Rajpurkar, P.; et al. CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv preprint arXiv:1711.05225* **2017**.
18. Fanni, S.C.; Marcucci, A.; Volpi, F.; Valentino, S.; Neri, E.; Romei, C. Artificial intelligence-based software with CE mark for chest X-ray interpretation: Opportunities and challenges. *Diagnostics* **2023**, *13*, 2020.
19. Irvin, J.; et al. CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2019, Vol. 33, pp. 590–597.
20. Akhter, Y.; Singh, R.; Vatsa, M. AI-based radiodiagnosis using chest X-rays: A review. *Frontiers in big data* **2023**, *6*, 1120989.
21. Torab-Miandoab, A.; Samad-Soltani, T.; Jodati, A.; Rezaei-Hachesu, P. Interoperability of heterogeneous health information systems: a systematic literature review. *BMC medical informatics and decision making* **2023**, *23*, 18.
22. Sheller, M.J.; et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports* **2020**, *10*, 12598.
23. Dou, Q.; et al. Federated deep learning for detecting COVID-19 lung abnormalities in CT: a privacy-preserving multinational validation study. *NPJ Digital Medicine* **2021**, *4*, 60.
24. Naz, S.; Phan, K.T.; Chen, Y.P.P. A comprehensive review of federated learning for COVID-19 detection. *International Journal of Intelligent Systems* **2022**, *37*, 2371–2392.
25. Rajpoot, R.; Gour, M.; Jain, S.; Semwal, V.B. Integrated ensemble CNN and explainable AI for COVID-19 diagnosis from CT scan and X-ray images. *Scientific Reports* **2024**, *14*, 24985.
26. Dwork, C.; McSherry, F.; Nissim, K.; Smith, A. Calibrating noise to sensitivity in private data analysis. In Proceedings of the Theory of Cryptography Conference, 2006, pp. 265–284.
27. Abadi, M.; et al. Deep learning with differential privacy. In Proceedings of the Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, 2016, pp. 308–318.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.