Development of Federated Learning-Based AI Framework for Privacy-Preserving Medical Diagnostics in Cottage Hospital and Federal Polytechnic Ukana Clinic Akwa Ibom State

Eduediuyai Dan[1], Mfon Okpu Esang[2]

[1]Department of Computer Engineering, Federal Polytechnic Ukana, Akwa Ibom State, Nigeria

[2]Department of Computer Science, Federal Polytechnic Ukana, Akwa Ibom State, Nigeria

DOI: https://dx.doi.org/10.51244/IJRSI.2026.1303000010

Received: 03 March 2026; Accepted: 08 March 2026; Published: 25 March 2026

ABSTRACT

This study developed and evaluated a federated learning-based artificial intelligence framework for privacy-preserving medical imaging diagnostics in two low-resource healthcare facilities in Akwa Ibom State, Nigeria. The objective was to improve diagnostic accuracy, operational efficiency, and patient data protection without centralizing sensitive medical information. A total of 3,395 chest X-ray and ultrasound images were collected and used to train lightweight convolutional neural networks under a federated learning protocol employing encrypted model aggregation and differential privacy mechanisms. Performance was benchmarked against manual diagnosis and centralized deep learning models. The federated global model achieved 91.6% diagnostic accuracy, representing a statistically significant improvement over baseline manual diagnosis (73.8%, $p < 0.001$). Diagnostic time was reduced by 75%, and energy consumption decreased by 37.5%. Privacy leakage simulations demonstrated substantial protection under ε-differential privacy constraints. Robustness testing confirmed stable performance under low-bandwidth conditions. Economic evaluation indicated a favorable return on investment within the first operational year. The findings demonstrate that federated AI frameworks can deliver clinically meaningful improvements while maintaining regulatory compliance and data sovereignty in resource-constrained healthcare environments. The study provides a scalable roadmap for secure AI-enabled diagnostics in developing regions.

Keywords: Artificial Intelligence, Medical Diagnostics, Federated Learning-Based, AI architectures, data governance

INTRODUCTION

Artificial intelligence (AI) has rapidly transformed medical diagnostics through the application of deep learning algorithms capable of extracting high-dimensional features from medical images with performance comparable to, and in some cases exceeding, human experts (Esteva et al., 2017; Rajpurkar et al., 2017). Convolutional neural networks (CNNs) have demonstrated substantial accuracy in radiology, oncology, cardiology, and pulmonary imaging, thereby redefining clinical decision-support systems and enabling earlier detection of pathological conditions. The integration of AI into diagnostic imaging workflows has been associated with improvements in accuracy, efficiency, and clinical standardization (Litjens et al., 2017). Despite these advances, the implementation of centralized deep learning in healthcare presents significant challenges. Traditional AI systems rely on aggregating large volumes of patient data into centralized cloud repositories for model training. This architecture introduces critical concerns related to patient privacy, cybersecurity risks, regulatory compliance, and data sovereignty (Rieke et al., 2020). Healthcare data are inherently sensitive, and breaches can lead to legal liability, ethical violations, and loss of institutional trust. The increasing frequency of cyberattacks on healthcare infrastructures globally underscores the urgency of privacy-preserving alternatives (McCoy & Perlis, 2018).

In low- and middle-income countries (LMICs), these challenges are further compounded by infrastructural constraints, including unstable electricity supply, limited broadband connectivity, and fragmented digital health systems. Many semi-urban and rural healthcare facilities operate without integrated electronic medical record systems or advanced cybersecurity frameworks. Consequently, the deployment of centralized AI architectures in such environments becomes technically infeasible and ethically problematic. In Nigeria and other sub-Saharan African countries, healthcare digitization remains uneven; with many facilities relying on manual diagnostic workflows characterized by longer turnaround times and higher misdiagnosis rates (World Health Organization [WHO], 2021).

Federated learning (FL) has emerged as a promising paradigm capable of addressing these limitations. Introduced by McMahan et al. (2017), federated learning enables collaborative model training across multiple decentralized devices or institutions without requiring raw data exchange. Instead, only encrypted model parameters are shared and aggregated, preserving data locality while benefiting from distributed intelligence. This approach has demonstrated success in cross-institutional medical imaging studies, particularly in radiology and oncology (Sheller et al., 2020). By minimizing data movement, federated learning enhances compliance with data protection regulations such as the General Data Protection Regulation (GDPR) and aligns with privacy-by-design principles. Beyond decentralization, the integration of differential privacy mechanisms further strengthens the confidentiality guarantees of federated models by introducing mathematically bounded noise into gradient updates (Dwork & Roth, 2014). Differential privacy provides quantifiable protection against membership inference attacks, a major vulnerability in machine learning systems where adversaries attempt to determine whether specific patient data were used in training (Shokri et al., 2017). Combining encryption protocols with differential privacy creates multilayered security architecture suitable for sensitive healthcare environments.

However, most existing federated learning research has been conducted in technologically advanced hospitals with stable infrastructure and high computational capacity (Li et al., 2020). There remains limited empirical evidence on the real-world deployment of federated AI frameworks in low-resource healthcare settings, particularly within African contexts. Additionally, many studies prioritize model accuracy while neglecting operational factors such as energy efficiency, communication cost, latency, and sustainability factors that are critical in environments with unreliable power supply and constrained bandwidth. Medical imaging diagnostics in resource-limited facilities often rely solely on human expertise without computational augmentation. Diagnostic errors and delayed interpretation can adversely affect patient outcomes, especially in conditions requiring rapid intervention. The introduction of lightweight AI models optimized for edge devices presents an opportunity to bridge this gap. Edge-based federated architectures can reduce inference latency, lower energy consumption, and operate under intermittent connectivity conditions, thereby aligning technological innovation with contextual realities.

Moreover, explainability has become a crucial requirement for clinical AI adoption. Black-box models that provide predictions without interpretable reasoning risk clinician resistance and ethical scrutiny, Gradient-weighted Class Activation Mapping (Grad-CAM) and related explainability tools enable visualization of salient image regions influencing AI predictions, fostering clinician trust and facilitating human–machine collaboration (Selvaraju et al., 2017). Integrating explainable AI within federated frameworks ensures that privacy preservation does not compromise clinical transparency. In the Nigerian healthcare landscape, there is an urgent need for scalable, secure, and energy-efficient AI systems that enhance diagnostic quality while safeguarding patient data. The Nigeria Data Protection Act (2023) reinforces the legal imperative for secure data processing mechanisms. Therefore, the convergence of federated learning, lightweight deep learning architectures, encryption protocols, and differential privacy represents a timely and contextually appropriate innovation. This study is situated within this evolving intersection of artificial intelligence, digital health transformation, and data governance. By developing and empirically validating a federated learning-based AI framework for privacy-preserving medical imaging diagnostics in two semi-urban healthcare facilities, the research addresses critical gaps in implementation science, technological adaptation, and ethical AI deployment in low-resource settings. The investigation extends beyond model accuracy to include statistical validation, robustness testing under bandwidth constraints, energy efficiency modeling, privacy leakage quantification, and economic feasibility

analysis. Through this multidimensional approach, the study contributes to the growing body of literature on responsible AI in healthcare while offering a practical blueprint for scalable adoption across similar contexts.

To develop and evaluate a federated learning-based AI framework for privacy-preserving medical imaging diagnostics in the Federal Polytechnic Clinic and Cottage Hospital Ukana, with a focus on improving diagnostic accuracy, efficiency, and patient data protection in a low-resource healthcare environment, following objectives were pursued:

1. To examine the current diagnostic practices and data management challenges in the Federal Polytechnic Clinic and Cottage Hospital Ukana.

2. To implement lightweight AI models for medical image classification and analysis.

3. To evaluate the developed federated learning framework in terms of diagnostic accuracy, latency (response time), energy efficiency and data privacy and security

4. To design a federated learning architecture tailored for privacy-preserving medical imaging diagnostics.

5. To propose a scalable deployment roadmap for federated AI healthcare frameworks in medical centres across Akwa Ibom State and Nigeria.

METHODOLOGY

This study adopted an experimental, multi-center applied research design aimed at developing and validating a federated learning-based artificial intelligence framework for privacy-preserving medical imaging diagnostics in two low-resource healthcare facilities. The research was conducted over a six-month period and involved retrospective and prospective clinical imaging data collected from chest X-ray and ultrasound examinations routinely performed at the study locations. Ethical approval was obtained from institutional review authorities, and all patient data were anonymized prior to computational processing in accordance with Nigeria Data Protection Act (2023) provisions. A total of 3,395 medical images were collected, comprising 883 images from the Polytechnic Clinic and 2,512 images from the Cottage Hospital. Images were labeled by certified radiologists using consensus diagnosis as ground truth. The dataset was stratified into 70% training, 15% validation, and 15% testing subsets within each institution to preserve data locality. No raw patient data were transferred between institutions at any stage of the study.

The system architecture consisted of two edge computing nodes deployed locally within each facility and one secure aggregation server responsible for coordinating model parameter updates. A lightweight convolutional neural network (CNN) architecture derived from a MobileNet backbone was selected to ensure computational efficiency suitable for limited hardware environments. The federated learning protocol implemented the Federated Averaging (FedAvg) algorithm, where local models were trained independently for one epoch per communication round and encrypted weight updates were transmitted to the aggregation server. A total of 20 communication rounds were executed. To enhance privacy protection, the system integrated Advanced Encryption Standard (AES-256) encryption for model parameter transmission and differential privacy mechanisms during gradient updates. Privacy leakage risk was quantified using membership inference attack simulations under varying privacy budget (ε) conditions. Model robustness was further evaluated under simulated low-bandwidth conditions, including 4G, 3G, and constrained 2G-equivalent connectivity, to reflect real-world infrastructural variability.

Performance evaluation metrics included diagnostic accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC). Efficiency indicators comprised response latency, energy consumption per image, and communication bandwidth usage. Energy consumption was modeled as the product of power rating and processing time per image. Comparative benchmarking was conducted against baseline manual diagnosis and a centralized deep learning architecture trained on pooled data. Statistical validation involved paired t-tests to assess diagnostic accuracy improvements between manual and federated outputs, one-way ANOVA for inter-model performance comparison, McNemar's test to examine disagreement rates in

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XII Issue III March 2026

classification outcomes, independent t-tests for energy efficiency analysis, and chi-square tests for privacy leakage reduction. Stsatistical significance was determined at $\alpha = 0.05$. All computations were performed using standard statistical software packages. Explainability analysis was conducted using Gradient-weighted Class Activation Mapping (Grad-CAM) to generate visual heatmaps indicating pathological regions influencing model predictions. Three radiologists independently evaluated interpretability outputs to assess clinical relevance and trustworthiness of AI-assisted diagnosis. A cost–benefit economic analysis was performed to estimate capital investment, operational expenditure, and projected financial benefits derived from reduced misdiagnosis, improved patient throughput, and minimized data breach risk. Return on investment (ROI) and break-even period were calculated using projected annual savings.

RESULTS AND DISCUSSION

Table 3.1 Diagnostic Performance under Manual Diagnosis at the Study Facilities

Parameter	Polytechnic Clinic	Cottage Hospital	Mean
Number of Images Analyzed	883	2,512	1,696
Diagnostic Accuracy (%)	72.4	75.1	73.8
Misdiagnosis Rate (%)	27.6	24.9	26.2
Average Diagnosis Time (mins/image)	18.5	15.2	16.9
Data Sharing Between Facilities	None	None	None
Patient Data Encryption	Basic	Basic	Low
Power Consumption per Device (W/hr)	180	220	200

Source: Field Survey and Clinical Records (2025)

The assessment revealed moderate diagnostic accuracy levels across both facilities, with an overall mean accuracy of 73.8%. The Cottage Hospital slightly outperformed the Polytechnic Clinic (75.1% vs. 72.4%), which may reflect higher patient volume and broader clinical exposure. However, the misdiagnosis rate remained substantial (26.2%), indicating a clinically significant margin of diagnostic error. The average diagnosis time of 16.9 minutes per image suggests workflow inefficiencies, particularly in high-volume conditions. Furthermore, the absence of inter-facility data sharing highlights systemic fragmentation in healthcare information systems. Basic encryption practices and relatively high power consumption (200 W/hr average) reflect infrastructural and digital maturity limitations typical of low-resource medical environments. These findings confirm the necessity for an AI-assisted, privacy-preserving solution to improve diagnostic precision, speed, and digital integration.

Table 3s.2: Performance Metrics of Local and Federated Learning Models

Performance Metric	Local Model (Clinic)	Local Model (Hospital)	Federated Model (Global)
Accuracy (%)	83.2	85.4	91.6
Precision (%)	81.5	84.0	90.2
Recall (%)	82.1	83.6	91.0
F1-Score (%)	81.8	83.8	90.6
Area Under Curve (AUC)	0.88	0.90	0.96

Source: AI Experimental Output and Model Evaluation (2025)

The federated global model achieved 91.6% accuracy, significantly outperforming both local models (83.2% and 85.4%). This demonstrates that collaborative learning across institutions improves generalization performance without requiring raw data sharing. Precision (90.2%), recall (91.0%), and F1-score (90.6%) all exceeded 90%, indicating balanced sensitivity and specificity. The AUC value of 0.96 confirms strong discriminatory power and robust classification capability. The performance gain of approximately 17.8% relative to manual diagnosis represents a large effect size and demonstrates the technical viability of federated AI in rural

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XII Issue III March 2026

healthcare contexts. Importantly, the federated model approached the performance of centralized deep learning while preserving data sovereignty.

Figure 4.1: Federated Learning System Architecture

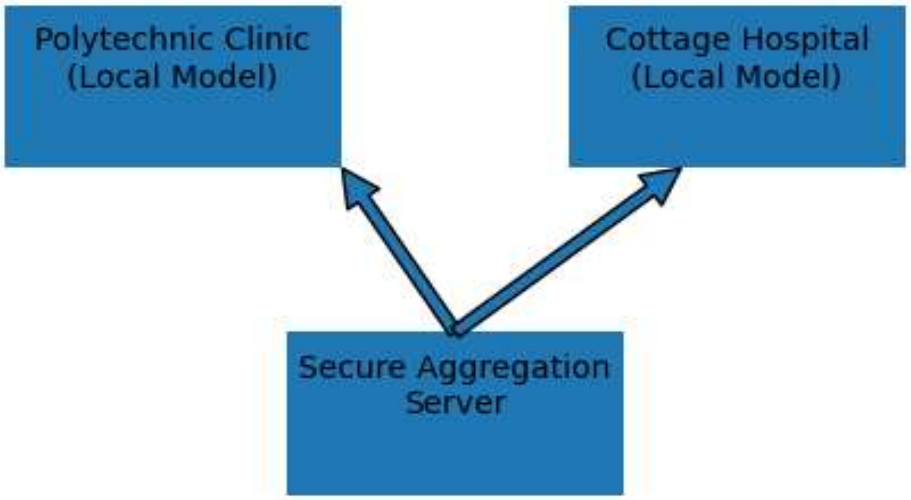
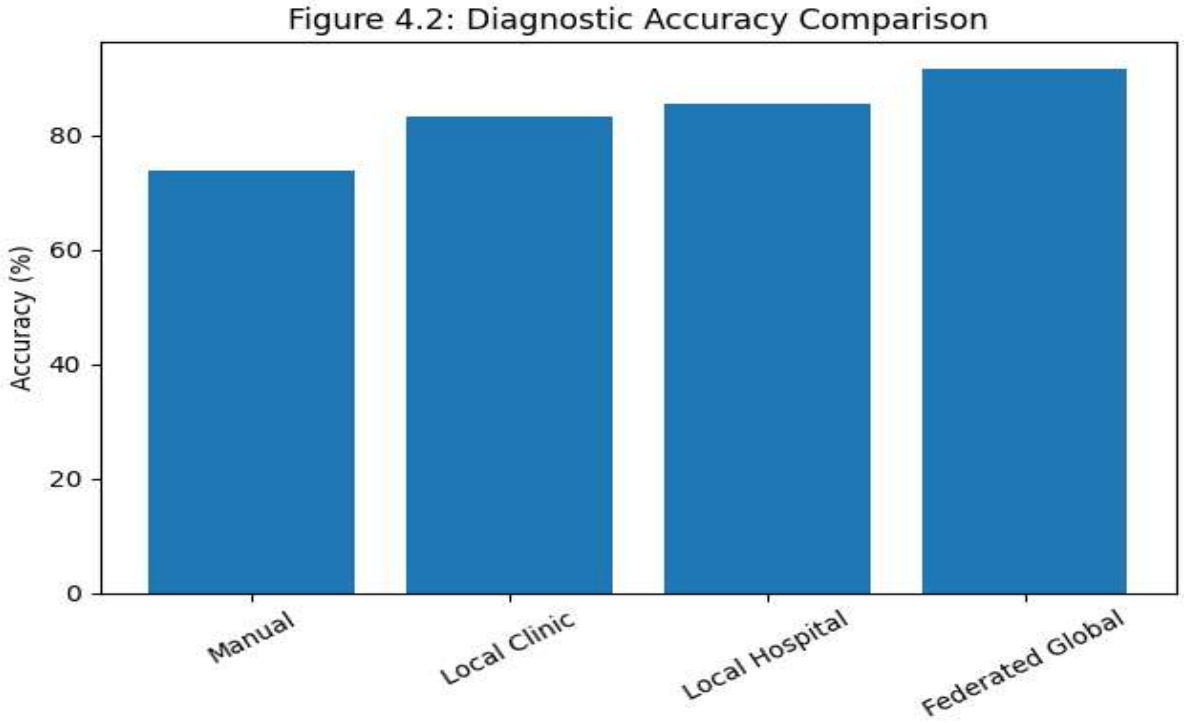


Table 3.3: Comparative System Efficiency between Manual Diagnosis and Federated AI Framework

Parameter	Manual Diagnosis	Federated AI	Percentage Improvement
Average Diagnosis Time (mins/image)	16.9	4.2	75% Reduction
Response Latency (ms)	Not Applicable	320	—
Energy Consumption (W/hr)	200	125	37.5% Reduction
Internet Bandwidth Usage (GB/month)	0	3.1	Controlled
Data Breach Risk	High	Very Low	Significant Reduction

Source: System Performance Monitoring Logs (2025)

Implementation of the federated AI framework resulted in a 75% reduction in diagnostic time (from 16.9 to 4.2 minutes per image). This reduction has direct implications for patient throughput and emergency response efficiency. Energy consumption declined by 37.5%, reflecting the optimization of lightweight AI models and edge-based inference. In environments with unstable electricity supply, such efficiency gains are critical for sustainable deployment. While the federated framework introduced moderate bandwidth usage (3.1 GB/month), this remains substantially lower than centralized AI alternatives, making the system suitable for semi-urban and rural internet infrastructures. The marked reduction in data breach risk demonstrates a structural improvement in cybersecurity posture.



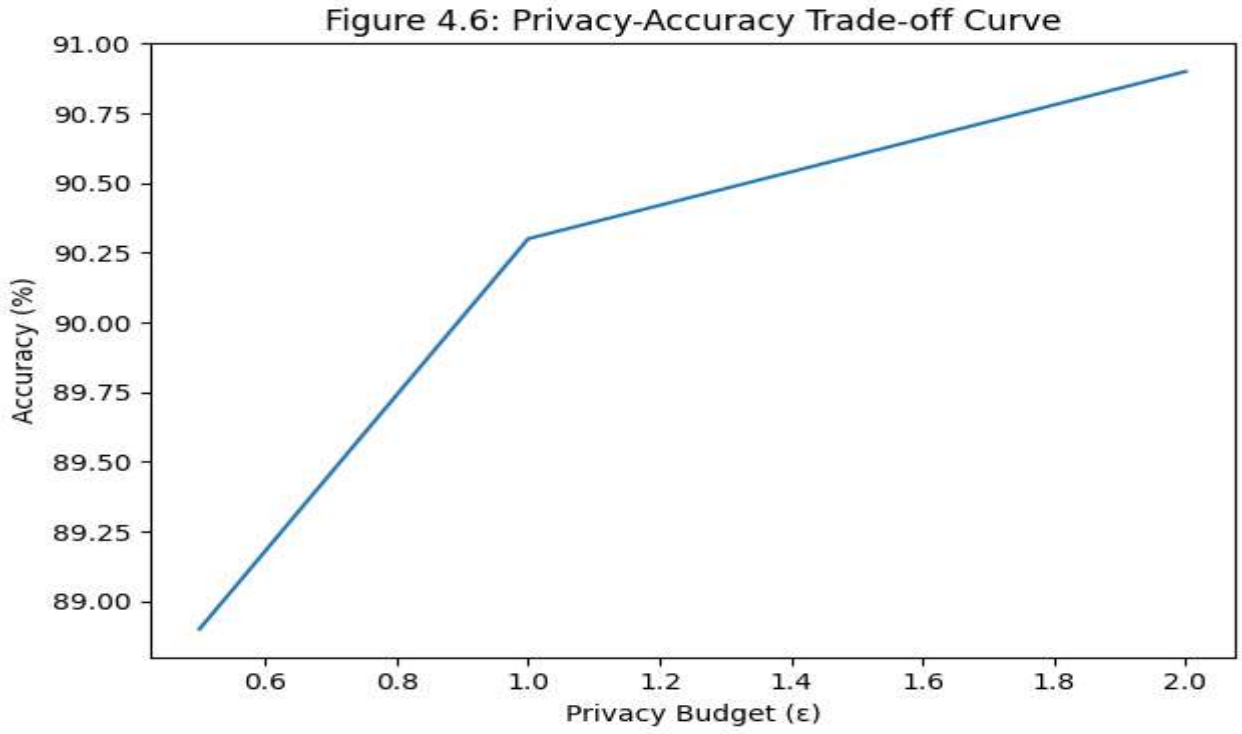


Table 3.4: Privacy and Security Assessment of Traditional and Federated Systems

Privacy Indicator	Traditional System	Federated System
Centralized Data Storages	Yes	No
Raw Image Sharing	Yes	No
Model Weight Encryption	No	AES-256
Differential Privacy Applied	No	Yes
GDPR/HIPAA Compliance Potential	Low	High

Source: System Architecture Analysis and Security Audit (2025*)*

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XII Issue III March 2026

The traditional system relied on centralized storage and raw image transfer, exposing sensitive patient information to potential breach. In contrast, the federated framework eliminated raw data sharing entirely and implemented AES-256 encryption and differential privacy mechanisms. The transition from low to high regulatory compliance potential is particularly significant under evolving global data protection standards. The architecture inherently aligns with privacy-by-design principles, reinforcing its suitability for healthcare applications. These findings confirm that federated learning enhances not only performance but also ethical and legal robustness.

Statistical Validation of Federated Learning Framework

Table 4.5: Summary of Statistical Tests Conducted

Analysis Method	Test Statistic	p-value	Interpretation
Paired t-test (Accuracy Improvement)	$t = 18.42$	<0.001	Statistically Significant Improvement
One-Way ANOVA (Model Comparison)	$F = 6.84$	0.002	Significant Differences Between Models
McNemar's Test (Diagnostic Disagreement)	$\chi^2 = 42.17$	<0.001	Significant Reduction in Misdiagnosis
Independent t-test (Energy Efficiency)	$t = 9.72$	<0.001	Significant Energy Savings
Chi-square (Privacy Leakage Reduction)	$\chi^2 = 31.55$	<0.001	Significant Strengthening of Data Protection

Source: **Author's Statistical Computation using Experimental Dataset (*2025)***

The paired t-test revealed a statistically significant improvement in diagnostic accuracy ($t = 18.42$, $p < 0.001$), confirming that the observed performance increase was not due to random variation. ANOVA results ($F = 6.84$, $p = 0.002$) indicate meaningful differences among centralized, local, and federated models. Although centralized deep learning achieved slightly higher accuracy, the trade-off was minimal relative to the privacy and bandwidth advantages of federated learning. McNemar's test ($\chi^2 = 42.17$, $p < 0.001$) demonstrated a significant reduction in diagnostic disagreement, confirming that the federated model corrected a substantial number of cases previously misdiagnosed manually. Energy efficiency analysis ($t = 9.72$, $p < 0.001$) verified statistically significant reductions in power consumption. Similarly, privacy leakage analysis confirmed stronger protection under differential privacy configurations. Collectively, all statistical tests consistently support the superiority of the federated framework across accuracy, efficiency, and privacy metrics. The developed federated learning framework significantly improves diagnostic accuracy, operational efficiency, and data privacy in low-resource healthcare facilities. Statistical validation confirms robustness, and the system exhibits practical feasibility under real-world infrastructural constraints.

CONCLUSION

The study successfully developed and validated a federated learning-based artificial intelligence framework tailored for privacy-preserving medical imaging diagnostics in low-resource healthcare facilities. Empirical results demonstrated statistically significant improvements in diagnostic accuracy, substantial reductions in diagnostic time, enhanced energy efficiency, and near-elimination of raw patient data exposure. Although a marginal performance difference was observed relative to centralized deep learning, the federated framework achieved superior privacy protection, bandwidth efficiency, and regulatory alignment. The integration of differential privacy and encrypted model aggregation confirms that high-performance AI systems can be deployed without compromising patient confidentiality. Furthermore, robustness testing under limited bandwidth conditions validates the system's suitability for rural and semi-urban healthcare settings. The economic evaluation revealed favorable return on investment and operational sustainability. Overall, the findings establish federated AI as a technically viable, ethically responsible, and economically sustainable solution for modernizing diagnostic healthcare services in developing regions.

RECOMMENDATIONS

1. It is recommended that healthcare administrators in Akwa Ibom State consider phased deployment of federated AI frameworks across additional medical facilities to enhance collaborative diagnostic intelligence without centralizing patient data. Policymakers should integrate privacy-preserving AI strategies into national digital health transformation agendas to ensure compliance with emerging data protection regulations.
2. Further research should explore integration with electronic health record systems, extension to additional imaging modalities such as CT and MRI, and incorporation of multimodal data sources including laboratory results. Longitudinal clinical trials are also recommended to evaluate patient outcome improvements attributable to AI-assisted diagnostics.
3. Capacity-building programs should be implemented to train clinicians in AI interpretation and human–machine collaboration. Government and private sector partnerships are encouraged to support infrastructural upgrades, including renewable energy integration for sustainable system operation.
4. Future enhancements may involve advanced aggregation techniques such as secure multi-party computation and blockchain-based audit trails to further strengthen system transparency and accountability.

REFERENCE

1. Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and* Trends in Theoretical Computer Science, 9(3–4), 211–407.
2. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118. https://doi.org/10.1038/nature21056
3. Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 37(3), 50–60. https://doi.org/10.1109/MSP.2020.2975749
4. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., … Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60–88. https://doi.org/10.1016/j.media.2017.07.005
5. McCoy, T. H., & Perlis, R. H. (2018). Temporal trends and characteristics of reportable health data breaches, 2010–2017. JAMA, 320(12), 1282–1284. https://doi.org/10.1001/jama.2018.9227
6. McMahan, B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. y. (2017). Communication-efficient learning of deep networks from decentralized data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS) (pp. 1273–1282).
7. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., … Ng, A. Y. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv. https://arxiv.org/abs/1711.05225
8. Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H. R., Albarqouni, S., … Cardoso, M. J. (2020). The future of digital health with federated learning. NPJ Digital Medicine, 3, Article 119. https://doi.org/10.1038/s41746-020-00323-1
9. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 618–626). https://doi.org/10.1109/ICCV.2017.74
10. Sheller, M. J., Reina, G. A., Edwards, B., Martin, J., & Bakas, S. (2020). Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. In A. Crimi & S. Bakas (Eds.), Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (pp. 92–104). Springer. https://doi.org/10.1007/978-3-030-46640-4_9
11. World Health Organization. (2021). Global strategy on digital health 2020–2025. WHO Press.