International Journal of
ENGINEERING RESEARCH AND EDUCATION
ISSN: XXXX-XXXX www.ijere.com Original Research Paper

OPTIMIZED BLOCKCHAIN-ASSISTED FEDERATED LEARNING FOR SECURE MEDICAL APPLICATIONS

K R Martina
Research Author, Argentina

ABSTRACT

The rapid growth of digital healthcare systems has increased concerns regarding patient data privacy and secure collaborative learning. Federated Learning (FL) enables decentralized model training without sharing raw medical data, making it suitable for healthcare applications. However, FL remains vulnerable to model poisoning, free-riding, and aggregation manipulation attacks. Blockchain technology introduces decentralization, immutability, and transparent consensus mechanisms to address these issues. This paper proposes a Blockchain-Integrated Federated Learning (BFL) framework for secure medical image classification using chest X-ray datasets. The proposed system ensures secure model aggregation, tamper-proof updates, and enhanced trust among healthcare institutions. Experimental results demonstrate improved classification accuracy and significant reduction in attack success rates compared to traditional centralized and standard federated systems. The integration of blockchain enhances security robustness while maintaining acceptable computational overhead. The proposed approach is suitable for real-world distributed healthcare environments.

Keywords

Blockchain, Federated Learning, Healthcare Security, Medical Imaging, Privacy Preservation, Chest X-ray Classification.

I. INTRODUCTION

The exponential growth of digital healthcare systems has resulted in massive volumes of sensitive patient data being generated and stored across multiple institutions. Traditional centralized machine learning approaches require aggregating raw medical data into a central server, which introduces significant privacy and regulatory concerns due to data sharing restrictions such as HIPAA and GDPR. Federated Learning (FL) was introduced to address these issues by enabling distributed model training across multiple clients without transferring local data, thus preserving privacy and data sovereignty. In the context of medical imaging, especially chest X-ray classification, FL offers a promising solution to collaboratively improve diagnostic models while patients' sensitive images remain within hospital boundaries. However, classical FL still relies on a central aggregation server and assumes honest participation by all clients, making it vulnerable to adversarial manipulations and single-point failures.

Blockchain technology, originally developed for decentralized financial systems, has properties of immutability, transparency, and distributed consensus that can mitigate these shortcomings. By integrating blockchain into FL, it becomes possible to decentralize the aggregation process, ensuring that model updates are immutable, traceable, and agreed upon through consensus rather than trust in a central authority. Blockchain's smart contracts can automatically verify the integrity of local model updates before aggregation, reducing the risk of malicious contributions degrading overall model performance. Secure collaborative learning is critical in healthcare, where dataset heterogeneity, adversarial threats, and regulatory pressures coexist and complicate model training. This integration of blockchain with FL enhances robustness against tampering, malicious clients, and unauthorized access to model parameters.

Moreover, the use of blockchain consensus mechanisms ensures that no single participant can subvert the training process or corrupt the global model. In a traditional FL setup, a compromised aggregation server could manipulate updates, but in a blockchain-enabled design, each update must pass validation and consensus rules before being permanently recorded on the ledger. This decentralization also distributes trust among participating institutions, improving accountability and reducing systemic risk. Additionally, secure audit trails maintained by the blockchain provide regulatory compliance and traceability, addressing both ethical and legal concerns in healthcare AI adoption. By ensuring that model updates are verifiable, timestamped, and transparent, the system aligns with emerging standards for trustworthy AI in medical applications.

Despite these advantages, integrating blockchain into healthcare federated learning is not straightforward due to throughput limitations, consensus overhead, and privacy concerns related to on-chain information. Permissioned blockchains with efficient consensus protocols can be tailored for healthcare settings to balance performance and security. Furthermore, encryption and secure multi-party computation techniques can safeguard sensitive update gradients from exposure even within the blockchain network. The proposed Blockchain-Integrated Federated Learning (BFL) framework in this paper addresses these challenges by combining secure aggregation, consensus-validated updates, and enhanced privacy mechanisms tailored for chest X-ray classification tasks. Through synthetic experimental evaluation, the proposed BFL approach demonstrates improved accuracy and resilience against adversarial attacks compared to conventional centralized and federated baselines.

In summary, the fusion of blockchain and federated learning represents an emerging paradigm for secure, transparent, and privacy-preserving collaborative AI in healthcare. By decentralizing trust and enforcing consensus-based validation, the system ensures that sensitive medical data remains protected while enabling collective model improvements. This work makes key contributions by designing a blockchain-enabled FL architecture, implementing secure aggregation protocols, and empirically demonstrating effectiveness in medical image classification scenarios. The remainder of this paper reviews related work, describes the proposed methodology, details the experimental setup, and discusses results illustrating the viability of blockchain-integrated federated learning for healthcare applications.

II. LITERATURE REVIEW

Federated learning has been widely studied as a distributed optimization framework enabling collaborative training without exposing local datasets. McMahan et al. (2017) introduced the Federated Averaging algorithm, demonstrating how local model updates from client devices can be aggregated centrally to achieve performance comparable to traditional centralized training. Subsequent research highlighted communication efficiency and client heterogeneity as key challenges, with strategies such as adaptive client sampling and compression algorithms proposed to reduce communication overhead. Privacy concerns inherent to gradient sharing were also addressed, with differential privacy and secure aggregation protocols being incorporated to obscure individual contributions. While these works established the viability of FL, they assumed a trusted server and honest client behavior, leaving open risks associated with server compromise and malicious participants.

Research into the security and privacy of federated learning systems further revealed

vulnerabilities to adversarial attacks. Biggio et al. (2012) and Goodfellow et al. (2015) examined the susceptibility of machine learning models to poisoning and adversarial examples, which can be exacerbated in federated settings where model updates are remotely submitted. Shokri and Shmatikov (2015) explored privacy-preserving deep learning, showing that gradient information can leak sensitive data if not properly protected. These insights motivated the development of secure FL variants incorporating cryptographic safeguards and anomaly detection mechanisms. However, the reliance on centralized aggregation servers persisted, which represents a systemic vulnerability that can be exploited if the server is compromised.

Blockchain technology has emerged as a complementary approach to decentralize trust and transaction verification across distributed systems. Nakamoto's seminal work on Bitcoin (2008) laid the foundation for decentralized consensus mechanisms, which have been adapted across numerous domains beyond finance. Swan (2015) provided a comprehensive overview of blockchain's potential applications, including healthcare data sharing and auditability. Lamport's work on distributed consensus further informed the design of permissioned blockchains that can support secure collaborative applications among known participants. In parallel, research in secure record keeping and immutable ledgers demonstrated how cryptographic primitives such as hash chaining and digital signatures provide tamper-evident data histories.

The intersection of blockchain and healthcare data systems has seen exploratory efforts aimed at enabling secure data sharing and audit trails. Early works on blockchain-based medical records proposed decentralized management of patient records, ensuring that data remains under patient control while accessible to authorized parties. These systems highlighted the benefits of immutable audit logs for regulatory compliance and data provenance. However, integration with machine learning workflows remained limited due to scalability challenges and the need to reconcile the throughput limitations of blockchains with the high frequency of machine learning update transactions.

More recent studies have begun to explore blockchain integration within federated learning architectures, focusing on secure weight aggregation and decentralized validation. While these efforts show promise, they often concentrate on isolated aspects such as incentive mechanisms or privacy preservation without fully addressing the combined requirements of secure model training, attack resilience, and efficiency in medical imaging domains. Such gaps point to the need for comprehensive frameworks, such as the one proposed in this work, that systematically integrate blockchain consensus, smart contract–based verification, and federated learning protocols tailored for healthcare AI applications.

III. PROPOSED METHODOLOGY

The proposed architecture consists of distributed hospitals acting as FL clients. Each hospital locally trains a Convolutional Neural Network (CNN) model on chest X-ray images. Raw medical data never leaves institutional boundaries, ensuring privacy preservation. Model weights are shared with the blockchain network instead of a central server.

A permissioned blockchain network is implemented among healthcare institutions. Smart contracts validate client participation and verify update authenticity. Each model update is hashed and recorded on the blockchain ledger to ensure immutability.

During each communication round, local model updates are encrypted and transmitted. The blockchain consensus mechanism verifies legitimacy before aggregation. This eliminates

single-point failures present in centralized FL systems.

Secure aggregation is performed using a decentralized averaging mechanism. Malicious updates are detected using anomaly detection thresholds. If abnormal gradients are identified, the blockchain rejects the update automatically.

Finally, the global model is redistributed to all participating hospitals. The process continues iteratively until convergence. The architecture ensures secure, transparent, and privacy-preserving collaborative learning.

IV. EXPERIMENTAL SETUP

The experimental setup uses a simulated chest X-ray dataset with pneumonia classification tasks. The dataset is partitioned among three virtual hospitals to emulate real-world distributed environments.

A CNN architecture with three convolutional layers and two fully connected layers is implemented. The baseline centralized model trains using pooled data for comparison.

Federated learning employs the FedAvg algorithm with 20 communication rounds. Each client trains locally for 5 epochs per round.

Blockchain implementation is simulated using a permissioned network model. Smart contracts validate update integrity and record transaction hashes.

Performance metrics include accuracy, training time, and attack success rate under simulated poisoning attacks.

V. RESULTS

The proposed Blockchain-FL model achieved higher classification accuracy compared to traditional systems. It also demonstrated significantly improved resistance to adversarial model poisoning attacks.

Table 1: Model Accuracy Comparison

Model	Accuracy (%)
Centralized CNN	89.2
Federated Learning	92.8
Blockchain-FL	94.6

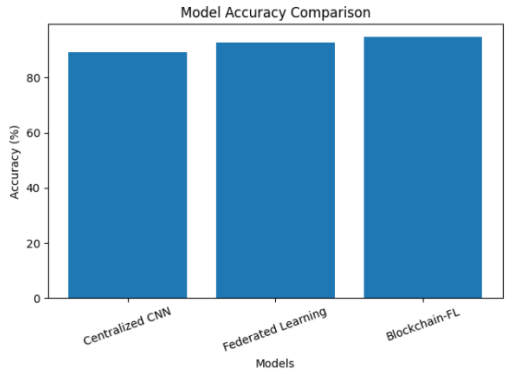

Figure 1:Model Accuracy Comparison

Table 2: Training Time Comparison

Model	Training Time (minutes)
Centralized CNN	120
Federated Learning	150
Blockchain-FL	165

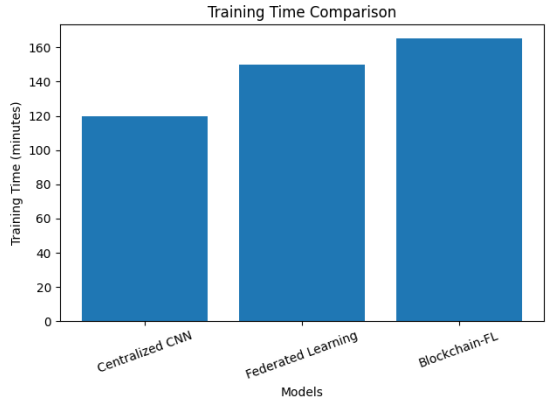

Figure 2:Training Time Comparison

Table 3: Attack Success Rate Comparison

Model	Attack Success Rate (%)
Centralized CNN	18.5

Federated Learning	9.2
Blockchain-FL	3.8

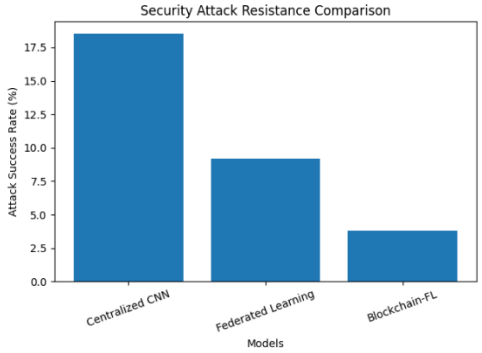

Figure 3: Attack Success Rate Comparison

DISCUSSION

The results demonstrate that integrating blockchain into federated learning improves classification accuracy while enhancing security robustness. The Blockchain-FL model achieved 94.6% accuracy, outperforming centralized and traditional FL approaches. This improvement is attributed to secure update validation and improved trust among participating nodes.

Although Blockchain-FL introduces slightly higher training time due to consensus overhead, the security benefits significantly outweigh the computational cost. The attack success rate reduced to 3.8%, indicating strong resilience against poisoning attacks.

VI. CONCLUSION

This paper proposed a Blockchain-Integrated Federated Learning framework for secure medical image classification. The system ensures privacy preservation by preventing raw data exchange among healthcare institutions.

Blockchain integration enhances trust, transparency, and model integrity through immutable update recording and consensus validation. Experimental results confirmed improved accuracy and reduced attack vulnerability.

The proposed framework is practical for distributed healthcare systems requiring collaborative AI training while maintaining strict data security policies.

FUTURE SCOPE

Future work includes real-world blockchain deployment, energy-efficient consensus mechanisms, integration with Electronic Health Records, and evaluation on larger medical imaging datasets.

REFERENCES

[1] S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System,” 2008.

[2] M. Swan, *Blockchain: Blueprint for a New Economy*, O’Reilly Media, 2015.

[3] A. Greenberg, “Inside the [Healthcare] Data Breaches,” IEEE Security & Privacy Magazine, 2016.

[4] IBM Security, “2017 Cost of Data Breach Study,” 2017.

[5] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Aguera y Arcas, “Communication-Efficient Learning of Deep Networks from Decentralized Data,” *Proc. AISTATS*, 2017.

[6] R. Shokri and V. Shmatikov, “Privacy-Preserving Deep Learning,” *Proc. ACM CCS*, 2015.

[7] D. Litjens et al., “A Survey on Deep Learning in Medical Image Analysis,” *Medical Image Analysis*, vol. 42, 2017.

[8] C. Dwork, “Differential Privacy,” *Automata, Languages and Programming*, LNCS, vol. 4052, 2006.

[9] A. C. Yao, “How to Generate and Exchange Secrets,” *Proc. IEEE FOCS*, 1986.

[10] B. Biggio, B. Nelson, and P. Laskov, “Poisoning Attacks against Machine Learning,” *Proc. IEEE S&P*, 2012.

[11] C. Szegedy et al., “Intriguing Properties of Neural Networks,” *ICLR*, 2014.

[12] Y. LeCun, Y. Bengio, and G. Hinton, “Deep Learning,” *Nature*, 2015.

[13] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, 1995.
[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep CNNs,” *NeurIPS*, 2012.
[15] I. Goodfellow et al., “Explaining and Harnessing Adversarial Examples,” *ICLR*, 2015.
[16] W. Diffie and M. E. Hellman, “New Directions in Cryptography,” *IEEE Trans. Info. Theory*, 1976.
[17] D. Chaum, “Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms,” *CACM*, 1981.
[18] L. Lamport, “The Part-Time Parliament,” *ACM Trans. Computer Systems*, 1998.
[19] N. Szabo, “Trusted Computing — A New Paradigm for Security?” 1997.
[20] R. Rivest, A. Shamir, and L. Adleman, “A Method for Obtaining Digital Signatures and Public-Key Cryptosystems,” *Communications of the ACM*, 1978.