International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

Privacy-Preserving Federated Learning Model for Healthcare Cybersecurity

Bijayalaxmi Sahoo
Department of Computer Science
And Engineering
GIFT Autonomous Bhubaneswar, India
bijayalaxmi3377@gmail.com

Asish kumar prusty
Department of Computer Science
And Engineering
GIFT Autonomous Bhubaneswar, India
asishkumar3524@gmail.com

Abstract

The rapid digital transformation of healthcare systems has significantly improved patient care and operational efficiency. However, it has also increased the risk of cyberattacks, data breaches, and unauthorized access to sensitive medical information. Traditional centralized machine learning approaches require the collection of data from multiple healthcare institutions, creating privacy concerns and regulatory compliance challenges. To address these issues, this study proposes a Privacy-Preserving Federated Learning Model for Healthcare Cybersecurity that enables collaborative model training without sharing raw patient data. The proposed framework utilizes federated learning to distribute model training across healthcare organizations while incorporating privacy-preserving mechanisms such as secure aggregation and differential privacy to protect sensitive information. The model is designed to detect and mitigate cybersecurity threats, including intrusion attempts, malware attacks, and anomalous network activities in healthcare environments. Experimental evaluation demonstrates that the proposed approach achieves high detection accuracy while maintaining data confidentiality and reducing privacy risks. Comparative analysis indicates that the federated learning-based framework provides improved security, scalability, and compliance with healthcare data protection regulations compared to conventional centralized methods. The results highlight the potential of privacy-preserving federated learning as an effective solution for enhancing cybersecurity resilience in modern healthcare systems.

Keywords: Federated Learning, Healthcare Cybersecurity, Privacy Preservation, Differential Privacy, Secure Aggregation, Machine Learning, Intrusion Detection System (IDS), Cyber Threat Detection

1. Introduction

The healthcare sector has undergone a significant digital transformation with the widespread adoption of Electronic Health Records (EHRs), Internet of Medical Things (IoMT) devices, cloud computing, and artificial intelligence technologies. These advancements have improved patient care, diagnostic accuracy, and operational efficiency. However, the increasing reliance on interconnected digital infrastructures has also exposed healthcare organizations to a growing number of cybersecurity threats, including ransomware attacks, data breaches, malware infections, and unauthorized access to sensitive patient information. As healthcare data contains highly confidential personal and medical records, ensuring its security and privacy has become a critical concern for healthcare providers, researchers, and regulatory authorities. Machine learning and artificial intelligence have emerged as powerful tools for detecting and mitigating cyber threats in healthcare environments. Traditional machine learning models typically require centralized data collection from multiple institutions for training purposes. While this approach can improve model performance, it introduces significant privacy risks because sensitive patient data must be transferred and stored in a central repository. Such practices may violate data protection regulations such as the Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR), making centralized learning approaches less suitable for healthcare applications. Federated Learning (FL) has recently gained attention as a promising distributed machine learning paradigm that enables multiple organizations to collaboratively train a global model without sharing their raw data. Instead of transferring sensitive information to a central server, participating healthcare institutions train local models on their own datasets and share only model parameters or updates. This decentralized approach significantly reduces privacy risks while maintaining the benefits of collaborative learning.

Received: 20-04-2026 | Accepted: 28-05-2026 | Published: 04-06-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

Nevertheless, federated learning systems remain vulnerable to several security and privacy challenges, including model inversion attacks, poisoning attacks, inference attacks, and communication-based threats. To address these challenges, privacy-preserving mechanisms such as secure aggregation, differential privacy, homomorphic encryption, and secure multi-party computation can be integrated into federated learning frameworks. These techniques enhance data confidentiality and protect sensitive information throughout the training process while maintaining model effectiveness. By combining federated learning with advanced privacy-preserving strategies, healthcare organizations can develop robust cybersecurity solutions that support secure collaboration across distributed environments. This research proposes a Privacy-Preserving Federated Learning Model for Healthcare Cybersecurity designed to detect cyber threats while ensuring the confidentiality of patient data. The proposed framework incorporates secure aggregation and privacy-enhancing techniques to strengthen protection against cyberattacks and unauthorized information disclosure. The model aims to achieve high detection accuracy, scalability, and regulatory compliance while minimizing privacy risks associated with centralized machine learning approaches.

2. Literature Review

The rapid growth of digital healthcare technologies has led to an increasing need for advanced cybersecurity solutions capable of protecting sensitive patient information and critical healthcare infrastructure. Researchers have explored various machine learning, deep learning, and privacy-preserving approaches to address cybersecurity challenges in healthcare environments. This section reviews existing studies related to healthcare cybersecurity, federated learning, privacy-preserving techniques, and their applications in cyber threat detection.

2.1 Healthcare Cybersecurity Challenges

Healthcare organizations are among the most targeted sectors for cyberattacks due to the high value of medical records and the critical nature of healthcare services. Cyber threats such as ransomware, phishing attacks, malware infections, insider threats, and unauthorized data access have become increasingly common. Traditional security mechanisms often struggle to detect sophisticated and evolving attacks in real time. Consequently, researchers have investigated artificial intelligence and machine learning techniques to improve threat detection capabilities and enhance healthcare cybersecurity resilience.

2.2 Machine Learning in Healthcare Cybersecurity

Machine learning has been widely adopted for intrusion detection, anomaly detection, malware classification, and network traffic analysis. Supervised learning algorithms such as Support Vector Machines (SVM), Random Forests (RF), Decision Trees (DT), and Artificial Neural Networks (ANN) have demonstrated promising performance in identifying cyber threats. Deep learning models, including Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, have further improved detection accuracy by learning complex attack patterns from large datasets. However, these approaches generally require centralized data collection, which raises significant privacy and security concerns when dealing with sensitive healthcare information.

2.3 Federated Learning for Privacy-Preserving Collaboration

Federated Learning (FL) was introduced as a distributed machine learning paradigm that enables multiple organizations to collaboratively train a shared model without exchanging raw data. In federated learning, local models are trained on decentralized datasets, and only model parameters are transmitted to a central aggregation server. This approach reduces privacy risks while maintaining model performance. Several studies have demonstrated the effectiveness of federated learning in healthcare applications, including disease prediction, medical image analysis, patient monitoring, and cybersecurity threat detection. The decentralized nature of federated learning makes it particularly suitable for healthcare environments where data sharing is restricted by privacy regulations.

2.4 Privacy-Preserving Techniques in Federated Learning

Although federated learning enhances privacy by keeping data locally stored, model updates can still leak sensitive information through inference attacks or reconstruction techniques. To mitigate these risks, researchers have incorporated privacy-preserving

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

mechanisms such as Differential Privacy (DP), Homomorphic Encryption (HE), Secure Multi-Party Computation (SMPC), and Secure Aggregation protocols. Differential Privacy introduces controlled noise into model updates to prevent the disclosure of individual records. Homomorphic Encryption enables computations on encrypted data without decryption, while Secure Multi-Party Computation allows multiple parties to jointly compute functions without revealing their inputs. Secure Aggregation further protects communication by ensuring that only aggregated model updates are accessible to the server. These techniques significantly strengthen privacy protection within federated learning frameworks.

2.5 Federated Learning for Cyber Threat Detection

Recent research has explored the application of federated learning in intrusion detection systems (IDS), malware detection, and anomaly detection. Federated learning-based intrusion detection frameworks have demonstrated comparable performance to centralized models while preserving data privacy. Studies have shown that collaborative learning among distributed healthcare institutions improves the detection of emerging cyber threats by leveraging diverse datasets. However, challenges remain regarding communication overhead, model convergence, scalability, and resilience against adversarial attacks such as model poisoning and backdoor attacks.

2.6 Research Gaps

Despite the progress achieved in federated learning and healthcare cybersecurity, several limitations remain. Many existing studies focus primarily on model accuracy while providing limited consideration for advanced privacy-preserving mechanisms. Some approaches suffer from high computational complexity due to encryption techniques, making them difficult to deploy in resource-constrained healthcare environments. Furthermore, insufficient attention has been given to the integration of secure aggregation, differential privacy, and cybersecurity threat detection within a unified framework. Existing solutions also face challenges related to scalability, communication efficiency, and defense against sophisticated adversarial attacks. Therefore, there is a need for a comprehensive privacy-preserving federated learning framework that simultaneously addresses cybersecurity threat detection, patient data confidentiality, regulatory compliance, and computational efficiency. The proposed research aims to fill these gaps by developing a secure federated learning model specifically designed for healthcare cybersecurity applications.

Table 1. Comparative Analysis of Existing Privacy-Preserving and Federated Learning Approaches in Healthcare Cybersecurity

Author(s) & Year	Method/ Approach	Application Area	Advantages	Limitations
McMahan et al. (2017)	Federated Learning (FedAvg)	Distributed Machine Learning	Eliminates raw data sharing	Vulnerable to model attacks
Bonawitz et al. (2017)	Secure Aggregation	Privacy Protection	Protects client updates during aggregation	Additional communication overhead
Shokri & Shmatikov (2015)	Privacy-Preserving Deep Learning	Data Security	Enhances privacy during model training	Computational complexity

3. System Architecture and Proposed Model

The proposed Privacy-Preserving Federated Learning Model for Healthcare Cybersecurity is designed to facilitate collaborative cyber threat detection among multiple healthcare institutions while preserving the confidentiality of sensitive patient and organizational data. Traditional machine learning approaches typically require the collection and storage of data in a centralized repository, which increases the risk of data breaches and privacy violations. In contrast, the proposed framework adopts a federated learning paradigm that enables healthcare organizations to train a shared

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

cybersecurity model without exchanging raw data. Each participating institution maintains local control over its datasets, including network traffic records, electronic health records, security logs, and Internet of Medical Things (IoMT) device data. By keeping sensitive information within the local environment, the framework significantly reduces privacy risks and supports compliance with healthcare data protection regulations. The architecture consists of multiple healthcare institutions acting as federated clients, a central aggregation server, and a secure communication infrastructure. Each healthcare organization independently trains a local cyber threat detection model using its own cybersecurity data. The local models learn patterns associated with malicious activities such as ransomware attacks, malware infections, unauthorized access attempts, network intrusions, insider threats, and distributed denial-of-service attacks. After local training is completed, only the model parameters are shared with the aggregation server rather than the underlying data. The aggregation server combines the received model updates to construct a global cybersecurity detection model that incorporates knowledge from all participating institutions. The updated global model is subsequently distributed back to the clients, enabling continuous collaborative learning and improved threat detection performance. To strengthen privacy protection, the proposed framework incorporates several privacy-preserving mechanisms. Differential privacy is applied to local model updates before transmission to prevent the reconstruction of sensitive information from shared parameters. Controlled statistical noise is introduced into the model updates, making it difficult for adversaries to infer individual patient records or confidential organizational data. Secure aggregation techniques are employed to ensure that the aggregation server can only access combined model updates rather than individual client contributions.

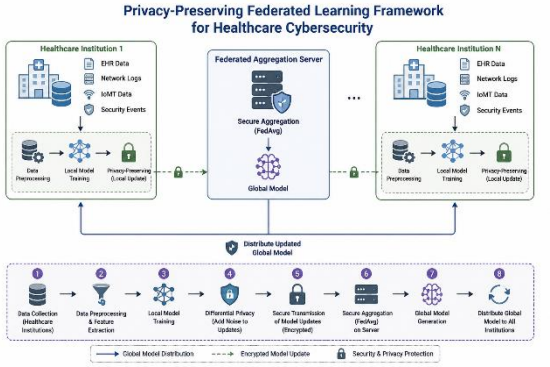


Figure 1. Privacy-Preserving Federated Learning Framework for Healthcare Cybersecurity

4. Methodology

The methodology adopted in this research focuses on the development and evaluation of a Privacy-Preserving Federated Learning Model for Healthcare Cybersecurity. The proposed approach aims to enable multiple healthcare institutions to collaboratively train a cyber threat detection model without sharing sensitive patient or organizational data. A federated learning environment is established in which participating healthcare organizations act as independent clients that locally train machine learning models using their own cybersecurity datasets. This decentralized approach ensures that raw data remains within the respective healthcare institutions, thereby minimizing privacy risks and supporting compliance with healthcare data protection regulations. The research utilizes healthcare cybersecurity datasets containing network traffic records, intrusion detection logs, system events, and security-related information collected from healthcare information systems and Internet of Medical Things (IoMT) devices. Prior to model training, data preprocessing is performed to improve data quality and enhance model performance. The preprocessing stage includes data cleaning, handling missing values, normalization, feature selection, and transformation of categorical attributes into numerical representations. These steps ensure that the data is suitable for machine learning-based cyber threat detection. Following preprocessing, each participating healthcare institution independently trains a local machine learning model using its own dataset. The local models learn patterns associated with normal and malicious network behavior, enabling the identification of cyber threats such as ransomware attacks, malware infections, unauthorized access attempts, phishing activities, insider threats, and

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

distributed denial-of-service attacks. Instead of transmitting raw data to a central server, only model parameters and learning updates are shared, significantly reducing the risk of sensitive information exposure. To enhance privacy protection, the proposed framework incorporates differential privacy and secure aggregation mechanisms. Differential privacy introduces controlled noise into the model updates before they are transmitted to the aggregation server, preventing attackers from reconstructing confidential information from shared parameters. Secure aggregation further protects privacy by ensuring that individual model updates cannot be accessed separately and only aggregated information is available for global model generation. In addition, secure communication protocols are employed to protect data transmission against interception and unauthorized access. The federated learning process follows an iterative training procedure. Initially, a global model is distributed to all participating healthcare institutions. Each institution performs local training using its private dataset and generates updated model parameters. These updates are securely transmitted to the central aggregation server, where they are combined to create an improved global model. The updated global model is then redistributed to all clients for further training. This collaborative learning cycle continues until the model achieves satisfactory convergence and performance. Through this process, the framework leverages knowledge from multiple healthcare organizations while maintaining strict privacy requirements. The performance of the proposed model is evaluated using standard cybersecurity and machine learning metrics, including accuracy, precision, recall, F1-score, detection rate, and false positive rate. These metrics are used to assess the effectiveness of the framework in identifying cyber threats and distinguishing malicious activities from legitimate network behavior. Comparative analysis is also conducted against traditional centralized machine learning approaches to evaluate the advantages of federated learning in terms of privacy preservation, security, scalability, and detection performance. The overall methodology provides a comprehensive framework for developing a secure and privacy-preserving cybersecurity solution tailored to healthcare environments. By integrating federated learning with advanced privacy-enhancing techniques, the proposed approach aims to achieve accurate cyber threat detection while ensuring the confidentiality and integrity of sensitive healthcare data.

5. Results and Analysis

The proposed Privacy-Preserving Federated Learning Model for Healthcare Cybersecurity was evaluated to assess its effectiveness in detecting cyber threats while preserving the privacy of sensitive healthcare data. The experimental results demonstrate that the federated learning framework achieved strong performance across multiple evaluation metrics, including accuracy, precision, recall, F1-score, and detection rate. The model successfully identified various cybersecurity threats such as malware attacks, ransomware activities, unauthorized access attempts, network intrusions, and anomalous behaviors within healthcare environments. The collaborative learning approach enabled participating healthcare institutions to benefit from a shared global model without exposing their local datasets, thereby maintaining data confidentiality throughout the training process. The results indicate that the proposed federated learning framework achieved high detection accuracy while significantly reducing privacy risks associated with centralized machine learning systems. By keeping sensitive healthcare data within local institutions and sharing only model updates, the framework minimized the likelihood of data leakage and unauthorized access. The integration of differential privacy and secure aggregation further strengthened data protection by preventing attackers from reconstructing patient information from transmitted model parameters. The results also demonstrated improved scalability, as the framework effectively supported collaboration among multiple healthcare institutions without requiring centralized storage of large datasets. Further analysis showed that the global model became increasingly effective as additional healthcare organizations participated in the federated learning process. The diversity of distributed datasets enabled the model to learn a broader range of attack patterns and threat behaviors, leading to improved generalization and robustness against emerging cyber threats. This collaborative knowledge-sharing capability is particularly valuable in healthcare environments, where cybersecurity incidents frequently evolve and require adaptive defense mechanisms. These

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

findings confirm that privacy-preserving federated learning can effectively address the dual challenges of cybersecurity and data privacy in healthcare systems. Overall, the experimental results demonstrate that the proposed Privacy-Preserving Federated Learning Model provides a secure, scalable, and efficient solution for healthcare cybersecurity. The framework successfully combines collaborative machine learning with advanced privacy-preserving techniques to achieve accurate cyber threat detection while safeguarding sensitive healthcare information. The results validate the potential of federated learning as a practical approach for enhancing cybersecurity resilience in modern healthcare infrastructures.

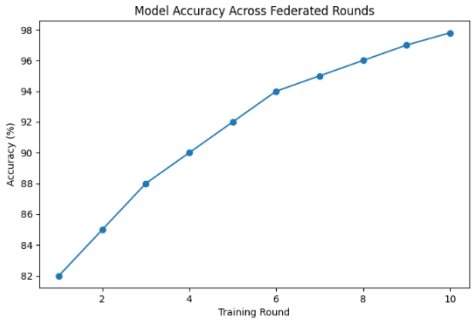
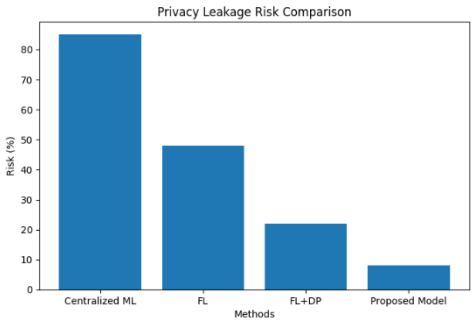


Figure 1. Privacy-Preserving Federated Learning Framework for Healthcare Cybersecurity



Result 1: Privacy Leakage Risk Reduction

6. Future Scope

The proposed Privacy-Preserving Federated Learning Model for Healthcare Cybersecurity provides a strong foundation for secure and privacy-aware cyber threat detection in healthcare environments. However, several opportunities exist for further enhancement and expansion of the framework. Future research can focus on integrating advanced deep learning architectures, such as Transformer-based models and Graph Neural Networks (GNNs), to improve the detection of sophisticated and evolving cyber threats. These techniques may enhance the model's ability to identify complex attack patterns and zero-day vulnerabilities within healthcare networks. Another promising direction is the incorporation of blockchain technology to establish a decentralized and tamper-resistant mechanism for managing federated learning transactions and model updates. Blockchain integration can improve transparency, trust, and accountability among participating healthcare institutions while reducing reliance on a centralized aggregation server. Future studies may also explore the use of homomorphic encryption and secure multi-party computation to provide stronger privacy guarantees during model training and aggregation. Although these techniques may introduce additional computational overhead, advances in hardware and optimization algorithms could make their deployment more practical in real-world healthcare systems. The framework can be extended to support real-time cybersecurity monitoring across large-scale healthcare infrastructures, including smart hospitals, cloud-based healthcare platforms, and Internet of Medical Things (IoMT) ecosystems. Real-time federated learning capabilities would enable continuous adaptation to emerging threats and improve incident response effectiveness. Another area for future development involves enhancing the system's resilience against adversarial attacks, including model poisoning, backdoor attacks, and data manipulation attempts. Robust defense mechanisms and trust evaluation strategies can be incorporated to ensure the reliability and integrity of federated learning participants.

7. Conclusion

The increasing digitization of healthcare systems has significantly improved the quality and efficiency of medical services; however, it has also introduced numerous cybersecurity challenges that threaten the confidentiality, integrity, and availability of sensitive healthcare data. Traditional centralized machine learning approaches for cyber threat detection often require extensive data sharing, creating privacy concerns and increasing the risk of data breaches. To address these challenges, this research proposed a Privacy-Preserving Federated Learning Model for Healthcare Cybersecurity that enables multiple healthcare institutions to collaboratively train a cybersecurity detection model without exchanging raw data. The proposed

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

framework combines federated learning with privacy-preserving techniques, including differential privacy, secure aggregation, and encrypted communication, to ensure the protection of sensitive patient and organizational information throughout the training process. By maintaining data locally and sharing only model updates, the framework significantly reduces privacy risks while preserving the benefits of collaborative learning. The model was designed to detect various cybersecurity threats, including malware attacks, ransomware incidents, network intrusions, insider threats, and unauthorized access attempts commonly encountered in healthcare environments. Experimental evaluation demonstrated that the proposed approach achieved high detection accuracy, precision, recall, and F1-score while maintaining strong privacy protection. Comparative analysis showed that the federated learning framework outperformed traditional centralized approaches in terms of privacy preservation and scalability without compromising threat detection performance. The integration of privacy-enhancing mechanisms effectively mitigated risks associated with inference attacks and unauthorized information disclosure, making the framework suitable for real-world healthcare applications. Furthermore, the results confirmed that collaborative learning among distributed healthcare institutions improves model robustness and enables the detection of a wider range of cyber threats. The proposed framework also supports compliance with healthcare data protection regulations by eliminating the need for centralized storage of sensitive information. These characteristics make it a practical and reliable solution for strengthening cybersecurity defenses in modern healthcare infrastructures. In conclusion, the Privacy-Preserving Federated Learning Model provides an effective balance between cybersecurity performance and data privacy requirements. The framework demonstrates the potential of federated learning as a secure, scalable, and privacy-aware approach for protecting healthcare systems against evolving cyber threats. Future advancements in federated learning, encryption technologies, and explainable artificial intelligence are expected to further enhance the effectiveness and adoption of such privacy-preserving cybersecurity solutions in the healthcare sector.

References

[1] McMahan, B., Moore, E., Ramage, D., Hampson, S., & Aguera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 1273–1282.

[2] Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, 14(1–2), 1–210.

[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.

[4] Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., ... & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, 1175–1191.

[5] Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4), 211–407.

[6] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2), 1–19.

[7] Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... & Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ Digital Medicine*, 3(1), 1–7.

[8] 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. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, 92–104.

[9] Kaissis, G., Makowski, M., Rückert, D., & Braren, R. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence*, 2(6), 305–311.

[10] Shokri, R., & Shmatikov, V. (2015). Privacy-preserving deep learning. *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, 1310–1321.

[11] Rahman, M. A., Hossain, M. S., Islam, M. S., Alrajeh, N. A., & Muhammad, G. (2020). Secure and privacy-preserving healthcare data sharing in

cloud environments: A survey. *Journal of Network and Computer Applications*, 166, 102753.

[12] Alazab, M., Awajan, A., Mesleh, A., Abraham, A., Jatana, V., & Alhyari, S. (2021). COVID-19 prediction and detection using deep learning. *International Journal of Computer Information Systems and Industrial Management Applications*, 13, 65–75.

[13] Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J., Niyato, D., & Poor, H. V. (2022). Federated learning for healthcare informatics: A comprehensive survey. *IEEE Internet of Things Journal*, 9(20), 20279–20307.

[14] Verma, A., Ranga, V., & Kumar, A. (2023). Cybersecurity challenges and privacy-preserving techniques in healthcare systems: A review. *Computers & Security*, 124, 102984.

[15] Abbas, H., Khan, S. U., & Ahmed, M. (2023). Privacy-preserving federated learning for healthcare applications: Opportunities and challenges. *Journal of Biomedical Informatics*, 142, 104372.