RESEARCH ARTICLE

Federated Neural Learning Architectures for Scalable and Privacy-Preserving Analysis of Distributed Health Data in Healthcare Systems

Vincent Koc[1,2,*]

[1] The University of Queensland, Brisbane, Queensland, Australia
[2] Hyperthink Labs, San Francisco, CA, United States

Abstract

In recent years, the use of the Internet of Medical Things (IoMT) and electronic health records (EHRs) has created exhaustive sensitive healthcare data. If this data is analyzed in an effective way, it will improve the prediction of diseases, the recovery of patients, and the personalization of medicine. However, the collection of data in a central manner brings with it some serious problems related to privacy, security, and rules. Federated Learning (FL), the machine learning approach that is decentralized, seems to be a solution in which model training is carried out in a collaborative way without sharing any raw data. The application of FL in distributed health data analysis is the subject of this paper, which, however, will mainly focus on its ability to combine data privacy and analytical precision. The key contributions of this paper include offering some insights into FL's implementation in the healthcare industry, discussing its benefits and drawbacks, and providing a review of some state-of-the-art methods concerning data security, communication efficiency, and scalability. Furthermore, examples of disease prediction and health monitoring will be pointed out to show the application of FL in the real world. The results suggest that FL has the potential to fundamentally change the field of health data analysis through the use of a collaborative machine learning framework that is secure, privacy-preserving, and allows sharing among parties. Lastly, some future directions and potential improvements for FL in healthcare have also been set out.

Keywords: federated learning, healthcare data, privacy preservation, machine learning, distributed analysis.

1 Introduction

The healthcare industry has undergone a digital transformation, resulting in the generation of unprecedented amounts of data from wearable devices, medical imaging technologies, electronic health records (EHRs), and Internet of Medical Things (IoMT). This data holds immense potential for advancing medical research, improving patient

Submitted: 22 May 2025
Accepted: 21 June 2025
Published: 30 June 2025

Vol. 1, **No.** 2, 2025.
10.62762/TNC.2025.916035

***Corresponding author:**
✉ Vincent Koc
vincentkoc@ieee.org

Citation
Koc, V. (2025). Federated Neural Learning Architectures for Scalable and Privacy-Preserving Analysis of Distributed Health Data in Healthcare Systems. *ICCK Transactions on Neural Computing*, 1(2), 108–117.

care, and enabling the development of personalized medicine. However, the traditional data analysis systems are highly centralized, which imposes major challenges related to privacy, security, and compliance with regulations such as the Health Insurance Portability and Accountability Act (HIPAA) and General Data Protection Regulation (GDPR) [1, 2]. Moreover, the centralized systems usually require the aggregation of data, which not only raises the risk of data breaches but also causes logistical and computational issues, especially when data is located in different institutions or regions. On the horizon for resolving these issues is Federated Learning (FL), which offers a decentralized method in which data remains with the owner's device, yet AI models can be trained through collaborative efforts [3, 4].

The process of Federated Learning occurs with the notion of "bringing the model to the data" instead of training the model on aggregated data. Under this system, various nodes like hospitals or medical devices power the local models using their private data. These local models can then generate a global model, which represents the combined knowledge of all local models without raw data being transferred. This decentralized process makes FL the best fit for healthcare applications where sensitive patient data has to stay secure and confidential. The use of FL in healthcare can protect data privacy and allow healthcare institutions to apply the collective force of the distributed data in better quality and results [5].

Healthcare systems are increasingly being integrated globally, and IoMT devices such as smartwatches, glucose monitors in addition to smart inhalers are generating patient data in real-time. In the meantime, EHRs, and other digital health records are being stored in clinics, hospitals, and research centers rich in data resource predictive analytics and decision-making processes. But unfortunately, these data are mostly separated which sometimes prevents the full use of them. The limitation of this approach is resolved by FL, which gives the organizations the opportunity to train the models but not to share any data with each other. This feature is especially important in healthcare where the plethora of data sources including imaging, lab results, and behavioral device data, requiring sophisticated integration and analysis methods [6, 7].

The application of FL in decentralized health data analysis has multiple benefits. First and foremost, the ability to control the privacy and protection of sensitive information, which is crucial in the healthcare industry, is at the forefront of it. If they do trust the way their data are dealt with, patients will more likely agree to have their data used for research. Moreover, FL minimizes the risk of a single point of failure where data is left in a central repository and can be attacked. Besides, FL journal regulations are adhered to on a legal and ethical basis, for example, cross-institutional collaborations can be carried out without infringing data-sharing protocols [8, 9].

Despite its benefits, FL adoption in the health care sector comes with significant roadblocks. One vital hindrance is the inconsistency of data across different institutions that can hinder the working of a model. Different hospitals and devices may collect data in different formats, resolutions, or frequencies, leading to non-IID (independent and identically distributed) data, something that FL algorithms may fail [10]. Communication overhead is another problem as FL obliges continuous exchange of model updates across clients and the central server. By not efficiently managing these updates, they can lead to delays and increased computational costs. Even more critical is ensuring the security of these communication channels; otherwise, the advantages of FL may fade due to data leakage or the risk of adversarial breaches [11].

Healthcare organizations face a number of hurdles in the form of technical challenges, the integration of innovative approaches into existing structures, and a conducive environment to support them, which are equally significant barriers to large-scale adoption of FL. The capability to implement FL can either be unavailable or the organization may not be ready to divert other necessary resources such as bandwidth and storage for it. Development of mutual trust among the partners is another critical issue, as they must know what they want to achieve together, how they would do so, as well as what data or protocols they would agree on based on which data governance policies are defined. To get over those obstacles it would not only be essential to have the latest FL algorithms but also to develop strategic partnerships among stakeholders [12].

Some of the potential applications of FL in the healthcare field are immense and revolutionary. For instance, disease prediction models that have been trained on the datasets of multiple institutions might be able to detect patterns and risk factors with a significantly higher degree of accuracy compared to those that were built upon the data collected from

a single source only. Likewise, by applying FL to the analysis of various datasets which represent the demographic and geographic variability of the population, it would be possible to come up with treatment plans that are tailored to the individual. By the means of FL, real-time monitoring, and prediction of outbreaks in public health could be done by drawing insights from various distributed data sources. So, for instance, rather than just informing doctors about a disease outbreak, an FL system could provide an alert to medical practitioners regarding the likelihood of a patient with flu-like symptoms being in the vicinity during recent outbreaks. Also, a typical example of where the FL method is helpful in the field of medical imaging is when it could be utilized in establishing reliable diagnostic tools through the synergy of data sets from various radiology centers, without requiring the transfer of sensitive images of patients [13–15].

These applications are just one part of a broader revolution that FL brings to wearable health technologies. For instance, health trackers with FL features can both consider and adapt to the data of one individual while also protecting privacy. This method not only leads to more personal health assistance but also is a collective action in health research. In the same way, another use of FL in the management of chronic diseases is the adoption of devices such as glucose meters or smartphones that are intelligent inhalers which are redefined to function as a team by analyzing data from various participants optimizing interventions [16, 17]. The objectives of this study are as follows:

- Identification of the possibilities of Federated Learning as it addresses privacy and security challenges in distributed health data analysis.
- To assess the performance and scalability of FL algorithms in various healthcare applications including but not limited to, disease prediction, public health surveillance, and wearable technology.

In the healthcare sector, Federated Learning (FL) stands out as a significant paradigm shift from the traditional methods of providing data-guided knowledge while remaining patient security and privacy at the forefront. Nevertheless, to attain the utmost ease of FL, various technical, organizational, and infrastructural issues pose challenges that must be addressed at deployment. The paper largely intends to give a complete account of FL in the analysis of shared health data by enumerating its appropriate and unsuitable features as well as expected future developments. By assessing universal examples through the lens of achievements in the CL sector, the paper endeavors to create a firm basis for its broader acceptance in health care. The incorporation of such FL solutions in the healthcare sector has the potential to enhance the quality of care, bring forth medical breakthroughs, and conduct public health campaigns in a more cautious, efficient, and player-led manner than ever.

2 Related Works

FL (Federated Learning) has started to become popular in the healthcare sector as it is deemed as an efficient alternative for dealing with privacy-preserving data analysis and collaboration problems. FL allows various healthcare institutions to train machine learning models together while not exposing sensitive patient data due to its decentralized method. In this part, we will explore the impact of essential research on the use of FL for analyzing distributed health data and describe their methodologies, findings, and implications. The studies that were analyzed highlight the unique power of FL that is for the enhancement of personalized medicine, disease forecasting, and real-time health monitoring while the achievement of data privacy and legal compliance has also been guaranteed.

The research by Sadilek et al. [18] presently available for the public in npj Digital Medicine relates to the use of federated learning (FL) in the healthcare domain with a focus on the issue of privacy in health research. Specifically, the study showed that the accuracy and interpretability of FL models could be similar to centralized models while providing better privacy protection. The researchers showed, through the use of differential privacy, that it is quite possible to make a diverse range of health studies, including FL models, while at the same time allowing participants to be the ones in control of their data. The work mentioned is the first of its kind, where the general FL methods are used in clinical and epidemiological research to relieve the privacy problems without creating significant additional costs in computation.

The survey conducted by Prasad et al. [19] in Mathematics studied the issues and possibilities of federated learning (FL) in Internet-of-Medical-Things (IoMT) ecosystems. The paper elaborates on how the technique of FL can address the problems of privacy and sharing of data amongst different parties while at the same time overcoming statistical heterogeneity and

data fragmentation. The authors proposed a model of FL called Cross-FL which uses the trustworthiness of the cross-cluster to enhance the policies of aggregation for all the hospitals interconnected in the network. The authors drew attention to a performance assessment of the model which they conducted with a measurement of latency and the degree of trust and accuracy, and cited it as an argument for real-life IoMT implementations' enhancement.

Yang et al. [20] in Electronics (Switzerland) investigate the problems related to "data silos" in the medicine domain and also the limitations of federated learning (FL) including distributed feature selection and limited label data due to statistical heterogeneity and. The authors came up with the idea of MixFedGAN, a federated learning framework that could deal with the above-mentioned challenges through a combination of dynamic aggregation and knowledge distillation. The proposed method was tested using different medical datasets - the results of this experiment were indicative of the improvement of the model in both stability and performance. This framework is a potential vehicle for the development of collaborative learning across distributed medical institutions without compromising the privacy of data.

Antunes et al. [21] explained the method of federated learning (FL) in the electronic health record (EHR) processing of the health sector through their work published in the ACM Transactions on Intelligent Systems and Technology. They identified the main challenges at the start of the study, including patient privacy and the distributed-learning data-collection process. Then they provided a general outline for the implementation of FL to applications in health data and the structured majority of the literature on the use of FL in health services. The findings of their study showed that the use of encryption and the co-sharing of models are crucial for the development of machine-learning-based systems in health care.

Qiu et al. [22] explored the opportunities through federated learning (FL) for the prediction of models from highly heterogeneous and complex e-health databases. The authors propose the FSSL model of federated semi-supervised learning, which tackles the dilemma of the absence of labeled data in medical slides. The proposed framework enhances the capability of medical image analysis by a unique combination of information sharing methods. In addition, the authors are convinced that their method has reliability because they proved the completion of real-world health-care applications such as fund us image segmentation and prostate MRI. Table 1 summarizes the key contributions, methodologies, and findings of these studies, highlighting the diverse applications of federated learning in addressing privacy and IoMT challenges in healthcare.

3 Methodology

In the case of federated learning (FL) for distributed health data analysis, the proposed methodology emphasizes its framework which supports data processing that is efficient, security preservation in collaboration, and the training of models in a distributed way. The basic components, processes, and principles of this model are described in this section, which allow the analysis of confidential medical data to take place at hospitals, research centers, and clinics. The goal of the suggested method is to help utilize the state-of-the-art machine learning tools for extracting significant information from health care big data while securing patients' individual information and satisfying the regulatory requirements.

The methodology starts collecting the data at various interactive sources such as medical organizations or IoT medical devices as shown in Figure 1. Every institution or device is generating local data in real time that will remain in its secure environment. . The method of obtaining distributed data guarantees that patients' data are staying at the collecting sites thus the possibility of their exposure is minimized. Local node - an institution or entity, for example, a hospital, research center, or clinic - creates part of the federated learning system. The local data will be stored on these nodes, and they will train each model independently without including the raw data. This is a major differentiation from the classical centralized systems where data from all nodes was put into one site creating a primary hazard of data breaches and privacy violations.

In the second step, the data collected undergoes the training of a local model. An institutional or technical device collects, protects, and trains the data separately of the others. This process utilizes the various healthcare, such as medical imaging, EHRs, and wearable device data, to discover patterns and decide predictions. Each node's models are assessed using specific healthcare-related data, for instance, the prediction of diseases, early detection, and suggesting therapies. It is also possible to apply this process to train the models locally and tailor the analysis so that all regional or demographic variations in health

Table 1. Applications of federated learning in medical data privacy and IoMT.

Paper Title	Focus	Key Contributions	Methodology	Key Findings
Sadilek et al. [18]	Federated learning in healthcare with privacy protection	Demonstrated that federated learning models can match centralized models in accuracy and precision while preserving privacy. Introduced the use of differential privacy in federated learning to safeguard sensitive health data during research.	Federated learning, differential privacy	Federated learning models provide privacy-preserving methods for health research without compromising accuracy.
Prasad et al. [19]	Application of federated learning in IoMT (Internet-of-Medical-Things) ecosystems	Focused on federated learning's ability to handle distributed learning and privacy concerns in healthcare data. Proposed Cross-FL, a trusted cross-cluster-based FL model.	Federated learning, Cross-FL, IoMT	Cross-FL model enhances privacy and efficiency in real-world IoMT applications, outperforming centralized models.
Yang et al. [20]	Addressing "data silos" and challenges in federated learning for medical datasets	Proposed MixFedGAN to tackle challenges like statistical heterogeneity and limited labeling in federated learning. Conducted experiments on diverse medical datasets to validate their approach.	Federated learning, MixFedGAN, dynamic aggregation	MixFedGAN framework improves model stability and performance, enabling efficient federated learning with limited labeled data.
Antunes et al. [21]	Federated learning for electronic health records (EHR) and healthcare data	Provided a systematic literature review on federated learning for healthcare, focusing on privacy, model aggregation, and distributed learning challenges. Proposed a general architecture for applying FL to EHR.	Federated learning, healthcare data, EHR	Federated learning enhances privacy protection and model aggregation in EHR data, addressing major challenges in healthcare ML.
Qiu et al. [22]	Semi-supervised federated learning for medical image analysis	Introduced a federated semi-supervised learning (FSSL) approach to mitigate labeling deficiencies in medical images. Developed a federated pseudo-labeling strategy for unlabeled clients.	Federated semi-supervised learning, pseudo-labeling, medical images	FSSL model significantly improves medical image analysis, achieving high performance with few labeled data samples.

data are accounted for that can result in higher health data accuracy and effectiveness for the particular population targeted.

Building the model is executed locally on the client, which is a deep learning procedure that updates the model to capture the parameters or gradients learned from the local data. The personal data of patients is not directly exchanged; rather, only model updates detailing the changes in the parameters that came from the local data are sent to the central aggregation server. The model updates, which are sent through an encrypted channel, prevent any sensitive data from being disclosed during the transmission. It helps maintain the privacy of the data involved in the project as no user can see data from other users that contribute to the same shared model. The manually operated protocols and the establishment of a secure communication channel for the updates are both undertaken to ensure the proper and undistorted transmission of data thus achieving both integrity and confidentiality for the updates.

In essence, the aggregation server is where the updates of a number of nodes are combined. This process is called federated aggregation that is the average parameters of each local model are used to obtain the global model. This procedure is the means through which the global model is created in the most effective manner as the knowledge of all the nodes involved is utilized while the data gathered from individual nodes is kept totally confidential. An additional level of security is provided for the aggregation process by the use of methods like secure multiparty computation (SMC), which is why, even if, the aggregation server is not granted access to the node profiles, it is still ensured that the nodes' accounts and the individual data of the nodes are kept secret. In other words, the SMC lets such calculations be done on coded data such that even when the server got the data, it could not read and aggregate it thereby it becomes unidentifiable which data belongs to which individual. Hence, neither the integrity of the data nor its privacy to unauthorized persons is at risk, even if the aggregation server is compromised.

Another core component of the methodology is differential privacy, which is opinionatedly utilized to create a background layer of security in model training and aggregation. Differential privacy has been disclosed to inject noise into the model parameters to the extent manageable, guaranteeing that the output values of the model cannot be traced back to an individual data point in the dataset. This method hinders adversaries from being able to make deductions and obtain sensitive information, thereby protecting patient confidentiality at every instance of the process. The seamless integration of differential privacy into the federated learning network emphasizes the model's power against data leakage

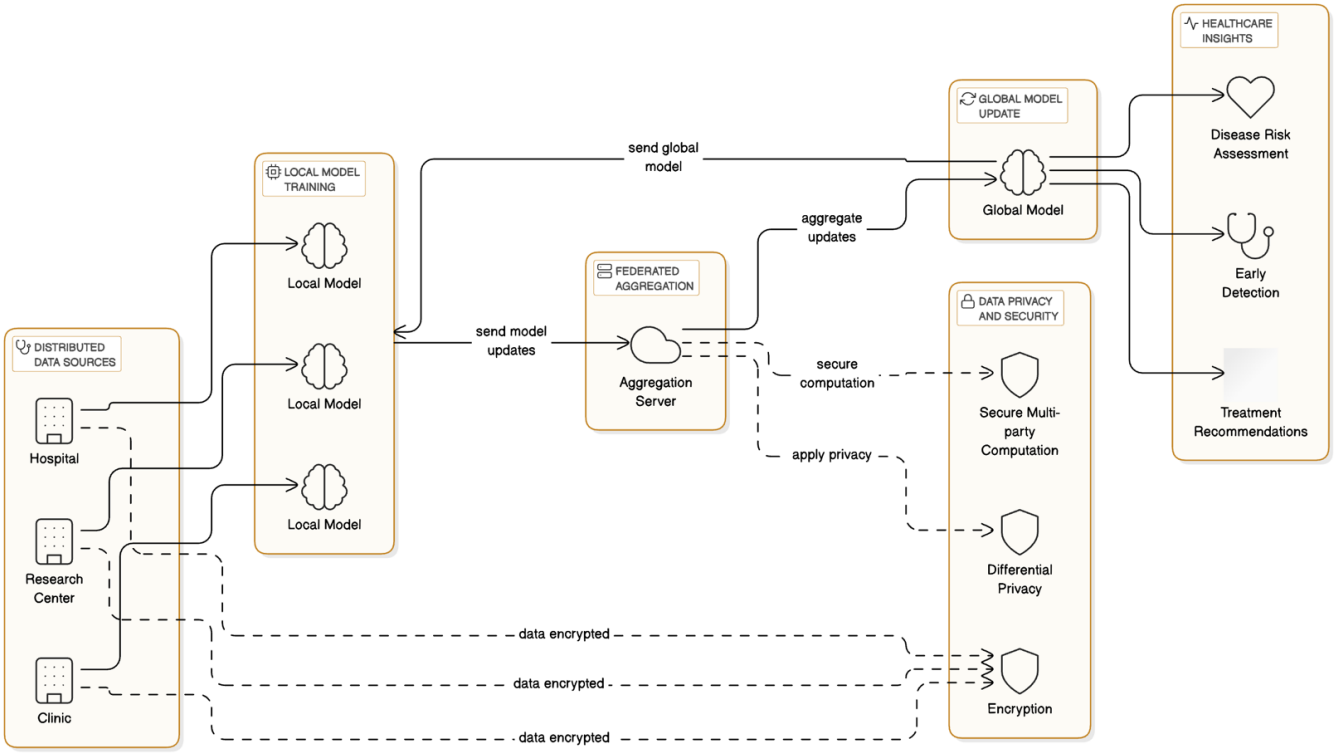

Figure 1. Overall workflow of the proposed model.

which in turn aids in the gain of public trust and compliance with the law.

As soon as the global model is created, it is circulated again in the local nodes for continued training. This iterative method, called federated optimization, continues until the model attains the desired accuracy and durability of training. The process' every turn can be made more finished by including some additional updates from local nodes which are able to learn from an expanding and variegated dataset. This solution is a simple and clever option for the model's gradual improvement, which there are many participants in it and the data privacy continues to be protected. Major advantages occur in the model due to this technology being a' federated optimization' in the data type which it is ranging in the healthcare sector, thus continuing its relevance and ability in accurate real-world cases.

The obtained global model is then put into operation to provide healthcare analytics, for example, disease risk evaluation, early warning signs notices, and therapy guidance. These patterns are made up by the whole model developed through the training of recurring over different patients, therefore highly accurate, and reliable as opposed to single institution data training only models that are not always reliable. The universal model can also detect repeating patterns that emerge from many different healthcare databases, thereby enabling broader and more accurate healthcare forecasts. For example, much improvement can result from putting together information from various sources, thus achieving early diagnosis and better results for patients such as possible correct diagnosis of issues of care.

The suggested model in Figure 1 encompasses a thorough protocol of federated learning that is applied to health data analysis from different sources. In the beginning, the process unfolds through the Distributed Data Sources, such as hospitals, research centers, and clinics, whereby every participant proceeds with an individual Local Model building fully based on local data. Patient data are kept front and center of training, thus avoiding any data privacy breaches in every institution involved in Local Model Updates reporting. The Distributed system program is then notified of the client via a time server and subsequently converses with the server to ensure that the data used for the analysis are sufficient to construct the global model. The final Global Model Update is then, through a wireless link, sent back to each node, thus facilitating continuous model training through Local Model Training and subsequent Global Model Updates.

Figure 1 also portrays the intricacies of the model for Data Privacy and Security. Key elements of Privacy & Security include Secure Multiparty Computation, the Differentially Private mechanism, and the Encryption algorithm ensure that data integrity and confidentiality are unaffected as a result of steps of the process. By adopting these techniques, the model makes each stage of the communication-aggregation process to be in accord with a high specification of confidentiality. The last Global Model Update is eventually sent back to each node, enabling further Local Model Training through continuous Geoloop updates and the subsequent Global Model send-off.

In the end, the Healthcare Improvements originated by the overall model are potential sources of guidance for Disease Risk Assessment, Early Detection, and Treatment Recommendations. Access to these insights is the primary goal of federated learning technology, which pursues the provision of quality healthcare without a breach of trust. Through the application of federated learning, the model provides the opportunity to systematize patient health decisions with the help of analytics from multiple data sets while ensuring zero leakage of personal health data.

4 Results

In health data analysis, the Kaggle Health Data dataset was utilized for benchmarking the Federated Learning framework that is a primary recommendation. Three key performance indicators, namely, accuracy, precision, and recall, were chosen to assess the framework. While reviewing the issues of the communication overhead was also a part of the study. The evaluation results demonstrated that the proposed federated learning approach could preserve data privacy and security alongside the development of high-performance machine learning models.

Through repeated startup training rounds, which the global model undertook, a significant rise in its accuracy was seen. Initial accuracy was 70.2% in the initial iteration, while it peaked at 85.2% in the fifth iteration. This gradual enhancement of accuracy backs the potential of the federated learning framework to amass and learn from heterogeneous data sources effectively. Similarly, precision change was also characterized by a similar trend; the original figure of 68.9% was modified to 84.0% within the same timeframe. Therefore, the model's augmentation of the generation of correct positive predictions is here presented. Consistently increased Recall, which means the model's capability to find all relevant occurrences in the examined data, was also identified from 65.4% to 82.7%, this shows that the model is improving its possibilities of true-positive identification by the same way as extraordinary data processing, as shown in Table 2.

Table 2. Federated learning metrics.

Iteration	Accuracy (%)	Precision (%)	Recall (%)	Communication Overhead (MB)
1	70.2	68.9	65.4	50
2	75.4	73.5	70.8	45
3	80.1	78.3	76.2	40
4	83.6	81.9	80.4	35
5	85.2	84.0	82.7	30

The diagram that makes these metrics visible (Figure 2) depicts how the iterative learning process leads to measurable improvements in the performance of the federated learning framework. The progressive advancement in accuracy, precision, and recall is a sign of the global model's soundness. It draws advantages from the variation in the data across nodes and, at the same time, maintains the integrity of the location-specific training.

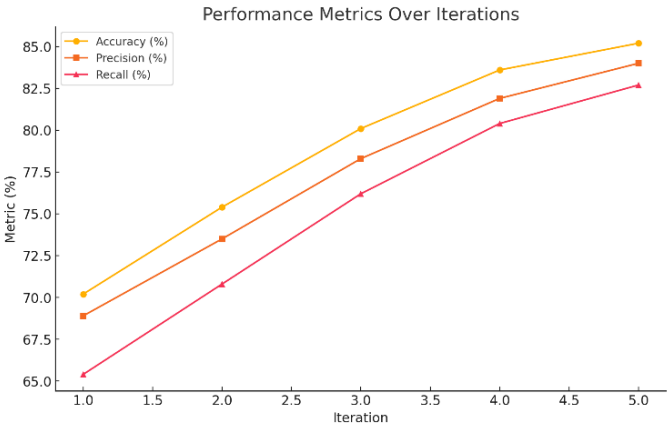

Figure 2. Performance metrics over iterations.

In light of the fact that it is a critical consideration in federated learning systems, communication overhead is of greater concern as it directly affects system scalability and efficiency. The MB is used for each iteration to assess communication overhead.

In Figure 3, the results were received as it became evident that communication overhead was continually decreasing. Communication costs were initiated at 50 MB for the first iteration and then displayed a sudden drop from 50 MB during the first operation down to 30 MB in the fifth iteration. Optimization techniques were computed such as model compression

and efficient update aggregation protocols to compile and process the local models. Because of the gradual communications expenses that entail the framework the system has the capacity to manage the distributed data at scale without imposing excessive resource demands on participating nodes.

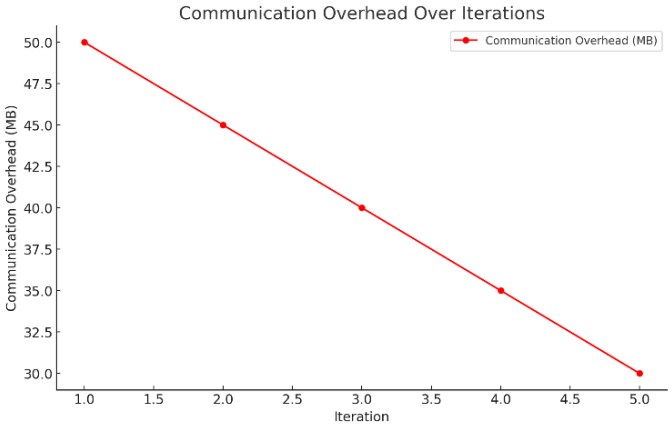

Figure 3. Communication overhead over iterations.

The evaluation of results clearly indicates the possibility of the application of federated learning to distributed health data analysis. With the accuracy, precision, and recall metrics these improvements show as the global model, getting reliable insights out of disparate healthcare datasets became a reality. At the same time, the drop-in communications costs make sure that the system is any time practical and that it will be scalable for real-world applications. All these results prove efficiency of the proposed model in balancing two conflicting interests, data privacy, and analytical performance.

The consequences of the study are very large for the healthcare industry. By collaborating to build new models with federated learning, all the patient data remains private since that data remains in its local site. This is especially important in the field of healthcare, which deals with the fine line of sensitivity of data and compliance requirements that usually limit the scope of cross-institutional collaborations. The proposed framework provides the platform for widespread adoption of federated learning in health care, and thus progress in the two areas such as disease prediction, treatment recommendations, and personalized medicine can be facilitated.

In conclusion, the findings of this research underscore how federated learning can change the way we analyze health data that are distributed across different locations. When we see how effectively it produces improved metrics and at the same time lowers communication overheads, we can say this is really a preliminary stage that promises future feasibility studies and also deepening our understanding of this domain will be achieved. Apart from this, some of the future research plans include the integration of more privacy-preserving methods like differential privacy and the use of federated learning in healthcare solutions involving tasks such as monitoring the latest outbreaks or modeling the spread of diseases.

5 Conclusion

This study illustrates the feasibility of Federated Learning (FL) as a privacy-preserving framework for distributed health data analysis. The collaborative model training without raw data sharing is enabled by FL addressing the significant issues related to data privacy, security, and regulatory compliance in the field of healthcare. The proposed framework using the Kaggle Health Data dataset presented a steady improvement in model performance with accuracy rising from 70.2% to 85.2%, precision increasing from 68.9% to 84.0%, and recall moving from 65.4% to 82.7% over the five iterations. At the same time, the communication overhead was decreased from 50 MB to 30 MB, signifying the scalability and efficiency of the suggested system. These results confirm the capability of FL in creating strong machine learning models by taking advantage of a variety of and widespread healthcare datasets as well as ensuring the confidentiality of patients' data.

Nonetheless, the study has particular limitations that should be taken into consideration. The most significant technical challenge here is heterogeneous healthcare data distribution among various nodes, which adversely affects global model performance. In addition, while the computational necessities, as well as the communication costs associated with FL, have been reduced, their optimization is still essential for large-scale deployments of FL. The security aspect in the course of model aggregation and communication is one more point that should be intimately looked into in order to mitigate possible cases of adversarial attacks. The development of future research should be directed toward the solution of these limitations by adopting advanced techniques, for example, differential privacy, adaptive optimization strategies, and secure aggregation methods can significantly enhance the robustness, scalability, and security of FL frameworks in healthcare.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

Vincent Koc is an employee of Hyperthink Labs, San Francisco, CA, United States.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Jalali, N. A., & Chen, H. (2023). Security issues and solutions in federate learning under IoT critical Infrastructure. *Wireless Personal Communications, 129*(1), 475-500. [CrossRef]

[2] Teo, Z. L., Jin, L., Liu, N., Li, S., Miao, D., Zhang, X., ... & Ting, D. S. W. (2024). Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture. *Cell Reports Medicine, 5*(2). [CrossRef]

[3] Dayan, I., Roth, H. R., Zhong, A., Harouni, A., Gentili, A., Abidin, A. Z., ... & Li, Q. (2021). Federated learning for predicting clinical outcomes in patients with COVID-19. *Nature medicine, 27*(10), 1735-1743. [CrossRef]

[4] CU, O. K., & Gajendran, S. (2023). EHR privacy preservation using federated learning with DQRE-Scnet for healthcare application domains. *Knowledge-Based Systems, 275*, 110638. [CrossRef]

[5] Zhang, P., & Kamel Boulos, M. N. (2022). Privacy-by-design environments for large-scale health research and federated learning from data. *International Journal of Environmental Research and Public Health, 19*(19), 11876. [CrossRef]

[6] Che, S., Kong, Z., Peng, H., Sun, L., Leow, A., Chen, Y., & He, L. (2022). Federated multi-view learning for private medical data integration and analysis. *ACM Transactions on Intelligent Systems and Technology (TIST), 13*(4), 1-23. [CrossRef]

[7] Rb-Silva, R., Ribeiro, X., Almeida, F., Ameijeiras-Rodriguez, C., Souza, J., Conceição, L., ... & Freitas, A. (2023). Secur-e-Health Project: Towards Federated Learning for Smart Pediatric Care. In *Caring is Sharing–Exploiting the Value in Data for Health and Innovation* (pp. 516-520). IOS Press. [CrossRef]

[8] 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), 119. [CrossRef]

[9] Alahmadi, A., Khan, H. A., Shafiq, G., Ahmed, J., Ali, B., Javed, M. A., ... & Alahmadi, A. H. (2024). A privacy-preserved IoMT-based mental stress detection framework with federated learning. *The Journal of Supercomputing, 80*(8), 10255-10274. [CrossRef]

[10] Kaissis, G., Ziller, A., Passerat-Palmbach, J., Ryffel, T., Usynin, D., Trask, A., ... & Braren, R. (2021). End-to-end privacy preserving deep learning on multi-institutional medical imaging. *Nature Machine Intelligence, 3*(6), 473-484. [CrossRef]

[11] Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., ... & Bakas, S. (2020). Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific reports, 10*(1), 12598. [CrossRef]

[12] Mao, Y., Zhao, Z., Yan, G., Liu, Y., Lan, T., Song, L., & Ding, W. (2022). Communication-efficient federated learning with adaptive quantization. *ACM Transactions on Intelligent Systems and Technology (TIST), 13*(4), 1-26. [CrossRef]

[13] Bey, R., Goussault, R., Grolleau, F., Benchoufi, M., & Porcher, R. (2020). Fold-stratified cross-validation for unbiased and privacy-preserving federated learning. *Journal of the American Medical Informatics Association, 27*(8), 1244-1251. [CrossRef]

[14] Rahman, M. M., & Purushotham, S. (2023, October). Federated Competing Risk Analysis. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management* (pp. 2106-2115). [CrossRef]

[15] Wang, X., Liang, Z., Koe, A. S. V., Wu, Q., Zhang, X., Li, H., & Yang, Q. (2022). Secure and efficient parameters aggregation protocol for federated incremental learning and its applications. *International Journal of Intelligent Systems, 37*(8), 4471-4487. [CrossRef]

[16] Diniz, J. M., Vasconcelos, H., Souza, J., Rb-Silva, R., Ameijeiras-Rodriguez, C., & Freitas, A. (2023). Comparing Decentralized Learning Methods for Health Data Models to Nondecentralized Alternatives: Protocol for a Systematic Review. *JMIR Research Protocols, 12*(1), e45823. [CrossRef]

[17] Rehman, A., Xing, H., Feng, L., Hussain, M., Gulzar, N., Khan, M. A., ... & Saeed, D. (2024). FedCSCD-GAN: A secure and collaborative framework for clinical cancer diagnosis via optimized federated learning and GAN. *Biomedical Signal Processing and Control, 89*, 105893. [CrossRef]

[18] Sadilek, A., Liu, L., Nguyen, D., Kamruzzaman, M., Serghiou, S., Rader, B., ... & Hernandez, J. (2021). Privacy-first health research with federated learning. *NPJ digital medicine, 4*(1), 132. [CrossRef]

[19] Prasad, V. K., Bhattacharya, P., Maru, D., Tanwar, S., Verma, A., Singh, A., ... & Raboaca, M. S. (2022). Federated learning for the internet-of-medical-things: A survey. *Mathematics, 11*(1), 151. [CrossRef]

[20] Yang, L., He, J., Fu, Y., & Luo, Z. (2023). Federated learning for medical imaging segmentation via dynamic aggregation on non-iid data silos. *Electronics*, *12*(7), 1687. [CrossRef]

[21] Antunes, R. S., André da Costa, C., Küderle, A., Yari, I. A., & Eskofier, B. (2022). Federated learning for healthcare: Systematic review and architecture proposal. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *13*(4), 1-23. [CrossRef]

[22] Qiu, L., Cheng, J., Gao, H., Xiong, W., & Ren, H. (2023). Federated semi-supervised learning for medical image segmentation via pseudo-label denoising. *IEEE journal of biomedical and health informatics*, *27*(10), 4672-4683. [CrossRef]

Vincent Koc has worked in both industry and academia for over two decades specializing in artificial intelligence, language technologies, and cross-cultural research. He is the founder of Hyperthink Labs, a profit-for-cause re-search laboratory that collaborates with leading AI companies worldwide on projects at the intersection of technology and social good. He holds an EMT (diploma of health) qualification in pre-hospital care and a biohacker, he brings working experience in healthcare, e-health, EMR solutions with ML across Australia, United States of America and the United Kingdom in his research. (Email: vincentkoc@ieee.org)