IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 17 July 2022, accepted 23 August 2022, date of publication 26 August 2022, date of current version 2 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3201876

SURVEY

Adoption of Federated Learning for Healthcare Informatics: Emerging Applications and Future Directions

VISHWA AMITKUMAR PATEL[1], PRONAYA BHATTACHARYA[2], (Member, IEEE), SUDEEP TANWAR[2], (Senior Member, IEEE), RAJESH GUPTA[2], (Student Member, IEEE), GULSHAN SHARMA[3], PITSHOU N. BOKORO[3], AND RAVI SHARMA[4]

[1]Sardar Vallabhbhai Patel Institute of Technology, Vasad, Gujarat 388306, India
[2]Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat 382481, India
[3]Department of Electrical Engineering Technology, University of Johannesburg, Auckland Park 2006, South Africa
[4]Centre for Inter-Disciplinary Research and Innovation, University of Petroleum and Energy Studies, Dehradun 248001, India

Corresponding authors: Sudeep Tanwar (sudeep.tanwar@nirmauni.ac.in) and Gulshan Sharma (gulshans@uj.ac.za)

ABSTRACT The smart healthcare system has improved the patients quality of life (QoL), where the records are being analyzed remotely by distributed stakeholders. It requires a voluminous exchange of data for disease prediction via the open communication channel, i.e., the Internet to train artificial intelligence (AI) models efficiently and effectively. The open nature of communication channels puts data privacy at high risk and affects the model training of collected data at centralized servers. To overcome this, an emerging concept, i.e., federated learning (FL) is a viable solution. It performs training at client nodes and aggregates their results to train the global model. The concept of local training preserves the privacy, confidentiality, and integrity of the patient's data which contributes effectively to the training process. The applicability of FL in the healthcare domain has various advantages, but it has not been explored to its extent. The existing surveys majorly focused on the role of FL in diverse applications, but there exists no detailed or comprehensive survey on FL in healthcare informatics (HI). We present a relative comparison of recent surveys with the proposed survey. To strengthen healthcare data privacy and increase the QoL of patients, we proposed an FL-based layered healthcare informatics architecture along with the case study on FL-based electronic health records (FL-EHR). We discuss the emerging FL models, and present the statistical and security challenges in FL adoption in medical setups. Thus, the review presents useful insights for both academia and healthcare practitioners to investigate FL application in HI ecosystems.

INDEX TERMS Blockchain, federated learning, healthcare informatics, gradient, model aggregation.

I. INTRODUCTION

The healthcare industry has converged towards healthcare 4.0, which presents an integration of assisted technologies like Internet-of-Things (IoT), big data, and artificial intelligence (AI) to support the medical operations and has automated the health operations at massive scales. With personalization and emergence of digital wellness, the healthcare industry in near future would transition towards Healthcare 5.0, which would assure customized disease control, virtual and emotive care, assisted living, virtual clinics, and remote monitoring [1]. Healthcare 5.0 would integrate the diverse technologies of Industry 4.0, with disruptive technologies like fifth-generation (5G) and beyond networks, unmanned aerial vehicle (UAV) monitoring, UAV logistics (of health centers), augmented and virtual reality, and nanotechnology. On the downside, the amount of generated data in healthcare has been massive, and thus proper techniques are required to maintain the collection of generated data among the medical stakeholders. Electronic medical records (EMRs) can conveniently solve the problem of collecting data. According to Groves *et al.* [2] in 2005, usage of EMR was only 30 per cent

The associate editor coordinating the review of this manuscript and approving it for publication was Massimo Cafaro.

by hospitals and office-based physicians, which increased to more than 50% for physicians and nearly 75% for hospitals in 2011.

Healthcare 5.0 has shifted the dynamics of healthcare informatics (HI), which deals with resources, technicality, and methods to optimize data collection and retrieval. However, the management is easier said than done. Recently, with the surge of data privacy and confidentiality of health records, the sharing of data among different stakeholders should be authorized by the patient. Thus, with stringent rules on data sharing, the data is anonymized, and sensitive attributes of the patient are not released. Coupled with effective record maintenance, selecting precise models that help predict diseases and maintaining privacy for the records in HI ecosystems is a tedious task. Hence, privacy and security for personal information have become crucial tasks with the data sharing paradigm. Personal details of the patients should be only accessible to authorized users, and unauthorized users should be restricted in their usage to ensure privacy and security.

Centralized machine learning (ML) approach are often challenges by various threats like privacy and security issues. In case of central disk, network, and link failures, the central data is not retrievable. Such issues can create a significant problem in the medical space as the data accessed by HI very critical. For example, a significant threat is data vulnerability to cyberattacks, and crucial data can be lost [3]. Also, in centralized ML, data training is done from a single server source data, due to which optimal prediction of any disease is not possible.

To address the issues of data privacy and sharing among nodes, federated learning (FL) is a viable choice in HI. With FL, the training of models is conducted on the device level. The device gets trained from its data and sends updates to the main servers. The main server then aggregates the updates and again transfers them to the devices [4]. FL has a higher prediction accuracy than FL models, and allows personalization in model learning. However, privacy is a big concern when data is shared. Privacy-preservation techniques of FL are better than centralized ML techniques as in centralized ML possibility of gathering such extensive data and providing security to it is not possible [4].

In FL, the multiple sites can train different models in parallel to create an aggregated model, which improves the accuracy and computation of ML models. Two basic approaches are presented in FL, one data-parallel approach, and the other is the model-parallel approach. In a data-parallel approach, we consider the entire healthcare data distributed among different servers, and all the servers use a unified learning model. In the model-parallel approach, we use different models to train specific data portions. The usability depends on the specifics of the underlying application. Although FL allows parallel learning paradigms, it is often challenging due to statistical differences in local data, systems, and the models built to support the different data distributions. Different systems have different computing capacities, such as CPU-cycle rate, I/O rates, the learning times would be different at different sites. Moreover, data is quite heterogeneous, with sensitive attributes and non-independent and identical distributions (non-IID) of patients. The distributed ML algorithms executing at different servers have to rely on the aggregation approaches and the encryption standards applied while sharing the data to the global model to assure privacy.

As healthcare setups are resource-constrained, FL requires a compelling mix of resource optimization and security while training the local models. To improve the accuracy of local models, more data is required, and model learning tasks would require sufficient resources in terms of storage and energy to train. In such cases, the aggregation of local models is mainly assisted through a cloud server that can provide the required resources for effective training [11]. However, cloud-supported FL aggregation is limited in sharing the model parameters with the cloud server, which an adversary might intercept. Cloud-based communication also introduces high end-user latency, and thus focus have shifted towards edge-based aggregation models. Edge-based aggregation improves on the latency constraints and improves the context of local learning, as the edge servers are close to the local models. In some cases, a hybrid cloud-edge approach is also preferred, where specific healthcare models are trained and aggregated via edge nodes, and nodes with general data distributions are sent for cloud-based FL aggregation. Edge-cloud based FL aggregation would allow more local participants to train the global model collectively, and thus the convergence of global models is fast, as the learning rate improves significantly [12]. TABLE 1 presents the list of abbreviations and their intended meanings in the article.

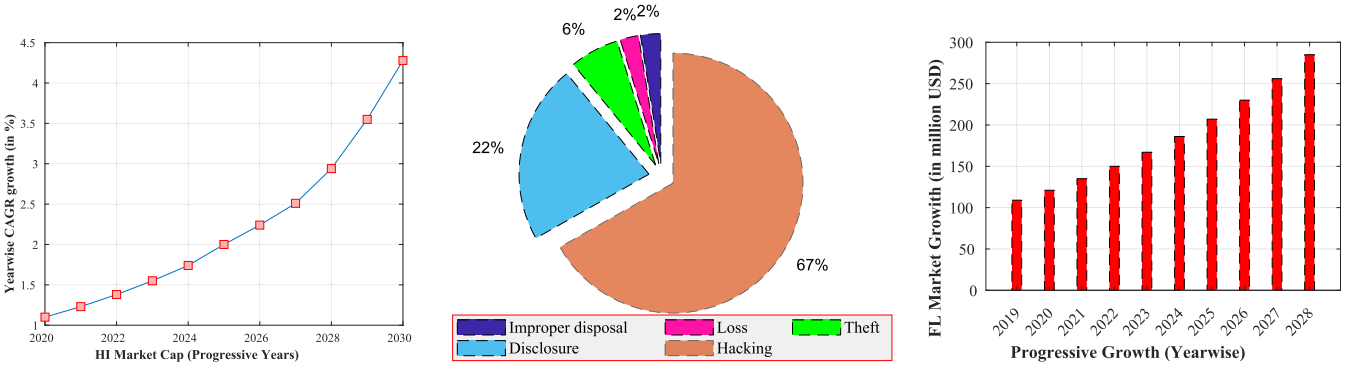
A. MARKET STATISTICS AND INDUSTRY TRENDS

In this subsection, we highlight the potential of FL adoption to HI. We investigate the predicted HI market cap by 2030 and analyze the privacy attacks on healthcare setups. We also investigate the FL adoption to the HI market in the future. FIGURE 1a shows the growth rate of FL from 2020 to 2030. As evident from the graph, the market cap of FL is continuously rising at an annual growth rate of 11.4% [5]. However, with decentralized healthcare analytics' popularity, the perimeter of data breaches has also increased. An adversary can locate multiple training regions to launch malicious training models that can jeopardize the sanctity of the global model. In a similar view, FIGURE 1b shows various types of data breaches in healthcare like hacking incidents, unauthorized access, theft, improper disposal, and loss [6], [13]. Thus, FL is viable to ensure HI privacy, confidentiality, and security. FIGURE 1c shows the emergence of FL in HI markets till 2030. The use of FL in healthcare is constantly growing to preserve the privacy of client's sensitive data [7].

Healthcare data is fragmented due to its complex nature, interconnected health processes, and entity interactions. It becomes difficult to assess the EMR sensitively. FL addresses the issues of sensitive data fragmentation, where local devices train the global model instance. Thus, FL is a

TABLE 1. Abbreviations and their descriptions.

Acronyms	Description	Acronyms	Description
4G-LTE	Fourth-Generation Long Term Evolution	HI	Healthcare Informatics
5G	Fifth-Generation	HIoT	Healthcare Internet-of-Things
5G-uRLLC	Fifth-Generation ultra-Reliable Low Latency Communications	HIPPA	Health Insurance Portability and Accountability Act
6G	Sixth-Generation	IoT	Internet-of-Things
6G-uRLLC	Sixth-Generation ultra-Reliable Low Latency Communications	LDPC	Low Density Parity Checks
AI	Artificial Intelligence	LDS	Local Data Store
ANN	Artificial Neural Networks	LSTM	Long-Short Term Memory
API	Application Programming Interface	ML	Machine Learning
CFL	Clustered FL	MLP	Multilayer Perceptron
CIFAR	Canadian Institute for Advanced Research dataset	MNIST	Modified National Institute of Standards and Technology dataset
CNN	Convolutional Neural Networks	MRI	Magnetic resonance imaging
COVID-19	Novel-Coronavirus-Disease-2019	non-IID	non-independent and identical distributions
DL	Deep Learning	P2P	Peer-to-Peer
DNN	Deep Neural Networks	PHR	Personal Health Records
DP	Differential Privacy	QSGD	Gradient Quantization and Encoding
EMNIST	Extended MNIST dataset	RPC	Remote Procedure Call
EMRs	Electronic Medical Records	SGD	Stochastic Gradient Descent
FedAvg	Federated Averaging	SMPC	Secure Multi-Party Computation
FedDist	Federated Distance	SVM	Support Vector Machines
FedMA	Federated Matched Average	TFF	TensorFlow Federated
FL	Federated Learning	TL-FL	Transfer-Learning FL
FLaaS	FL-as-a-Service	VFL	Vertical FL
HER	Electronic Health Records	WBAN	Wireless Body Area Networks
HFL	Horizontal FL		



(a) Predicted HI Market Cap Growth by 2030

(b) Different Sets of Privacy attacks on patient EHRs in HI

(c) FL adoption in HI and market growth in HI

FIGURE 1. Analysis of HI Markets and FL-driven Security to mitigate attack vectors in healthcare [5], [6], [7].

TABLE 2. Companies adopting FL as their solution.

Company name	Year	Domain	Objective	Benefit of integrating with FL
WeBank and Extreme Vision [8]	2019	Object detection	Safety monitoring solutions with the use of FL	High privacy and low cost of transmitting video data
Google [9]	2019	Edge devices	Virtual keyboard with FL for prediction of next word	High privacy and control over data
WeBank and Swiss Re [10]	2019	Finance	Analysis of data in finance and insurance	High data security
Owkin [10]	2020	Healthcare	Analysis of data in biomedical field	High data security is obtained with encryption
MELLODDY [10]	2021	Healthcare	Drug discovery with FL	High precision

good fit for HI, where both the data provider and data analyst can agree with the sharing rules defined. As public datasets do not contain explicit personal patient attributes, the data shared from global server models are free from risk. Due to this, many real-time project deployments of FL in healthcare and allied domains have started. TABLE 2 represents the companies that have adopted FL as a potential solution in their organizations. It also presents the project's domain and objectives and highlights the potential benefits of FL integration in such projects.

Some of the notable projects concerning FL are as follows. Webank and Extreme vision collaborated and formed a higher efficient object detection framework that detects objects through computer vision with also use of FL [8]. Google launched its gboard, which uses the FL model where locally trained models are aggregated and the global gradient is obtained [9]. For high precise financial analysis, Webank and Swiss Re used FL model [10]. Owkin used FL for security and privacy of user's data [10]. Furthermore, for drug discovery, MELLODDY also made use of the FL model [10]. Thus,

the deployment of FL in such real-life projects in diverse domains illustrates the requirement of secured data analytics at low computational power.

B. THE NECESSITY OF THE SURVEY

As indicated in above discussions, there is an exponential rise in EMR on a daily basis. Thus, all local setups have ample data, which can be analysed to form informed decisions. However, there are strict regulations on data sharing, with the general data protection regulation (GDPR) act in 2018. GDPR compels personnel to send data with comprehensive classification, and access grants. Moreover, EMRs are strictly guided by health insurance portability and accountability act (HIPPA), where personal indicators of patient cannot be shared without the consent of the patient. Thus, centralized models (cloud-based) for AI-driven health analytics are challenged with the quality of shared data (most shared data is anonymized, and explicit indicators are hidden), which allowed generic interpretability.

With the emergence of FL, the tradeoff between effective model learning vs sharing compliance is effectively handled. Data from heterogeneous local nodes are not shared with central sources. As all local nodes agree collaboratively on a training model, a customized global model is envisioned. The pretrained global model is then shared with local healthcare setups, and is trained with local data (no data is shared), which assures privacy regulations are maintained. This improves the efficiency of HI with FL as more data is gathered and precise models for detecting any disease can be done. Various inherent benefits of adopting FL, like disclosing sensitive data, i.e., achieved user data privacy techniques like differential privacy (DP), homomorphic encryption (HE), and secure multiparty computation (SMPC) can be achieved. In support of this, Rieke *et al.* [14] have reviewed many applications and advantages for medical imaging [14], whole-brain segmentation in magnetic resonance imaging (MRI), and other similar use-cases have been demonstrated. Thus, all types of stakeholders can be benefited when FL is integrated with HI.

Recently, the global FL research has been extensively conducted, but the existing researchers have studied FL in healthcare from a model viewpoint. However, the complexity of modern healthcare setups necessitates an umbrella survey that would present novel FL frameworks, tools, and effective case-studies of federated learning in healthcare informatics (FL-HI), which would assist stakeholders to design practical FL setups, which is still in its infancy. The proposed review highlights the adoption of FL in HI, and present a high-level FL-HI overview, and then delves deeper into the nuts and bolts of the architecture. As FL requires communication between local and global setups, security and privacy is equally important. We highlight the security and networking challenges, and practical frameworks, with an assisted decentralized FL-HI architecture underlying sixth-generation (6G) network services, which envisions a viable fit for Healthcare 5.0.

C. EXISTING SURVEYS

To date, researchers globally have proposed different surveys related to FL broad facets in healthcare applications. TABLE 3 presents the relative comparison of existing state-of-the-art surveys with the proposed survey on different proposed indicators like solution taxonomy, FL classification, architecture, privacy considerations, applications, metrics, and case-study. For example, Mothukuri *et al.* [15] discussed privacy and security concerns, achievements and impacts in FL. They presented a comparative study of threats in FL defence techniques, privacy preservation in FL, and approaches to enhance privacy preservation. The authors presented a detailed analysis of FL categorization, aggregation algorithms, security threats, and defensive techniques. However, they have not discussed any use case scenario as well as FL applications. Then, Xu *et al.* [16] reviewed various technologies, challenges, and privacy issues in FL-based healthcare systems. General solutions like consensus mechanisms and pluralistic solutions to statistical, system, and privacy challenges were discussed in this survey. A detailed analysis for communication efficiency was discussed into four groups-client selection, model compression, peer-to-peer (P2P) learning, and updates reduction. Later, Zhang *et al.* [17] surveyed the characteristics and applications of FL and divided it into five major parts- data partitioning, privacy mechanism, ML model, communication architecture, and system heterogeneity. The authors, however, have not mentioned the application of FL in HI ecosystems. Pfitzner *et al.* [18] proposed a survey that highlights the FL training process, and discussed the importance of uniform data collection in healthcare setups. The survey further discussed the neural network algorithms that are applied with the FL learning process. The survey indicates the FL model accuracy is heavily dependent on dataset characterstics. The survey delved into the hyperparameter tuning and optimization to reduce the number of epochs and improve the learning rate. Finally, different FL mechanisms are discussed, with open challenges. However, the survey did not focused on the security viewpoint in FL communication. To address this viewpoint, Prayitno *et al.* [19] proposed a survey that discusses the FL performance in medical imaging domain. Recent privacy and security challenges in data collection are discussed, and protection mechanisms are presented. Key aspects such as handle of data partitions, FL types, and non-IID characterstics are explored. Feature distribution skew characterstics, and hyperparameter tuning to support skewness is elaborated, with practical use-case examples. A privacy-preserving novel coronavirus disease-2019 (COVID-19) use case with the optimized classification and segmentation is proposed. However, the survey did not present any practical tools and frameworks to simulate the experimental parameters. Authors in [20] proposed a FL-based digital privacy health framework, that can collect data from collaborative hospital consortium setups. During the data collection, anonymization, and DP are induced into the data. The existing solutions and use-cases

TABLE 3. A comparative analysis of the proposed survey with existing state-of-the-art surveys in FL healthcare systems.

Author	Year	Objective	1	2	3	4	5	6	Pros	Cons
Weiss *et al.* [21]	2018	A review of mobile computing for the visually impaired	X	X	✓	X	X	X	better individual care of the patient	Privacy details are not mentioned
Yang *et al.* [22]	2019	A review federated ML with presented concept and applications	X	✓	✓	✓	✓	X	Reliability and detailed categorization of FL is mentioned	Case-study and proof-based solutions on HI is not proposed
Mothukuri *et al.* [15]	2019	A comprehensive literature review of security and privacy of federated learning	X	✓	✓	✓	X	X	Privacy, security, poisoning are explained in detail	Case-study and emerging applications are not discussed.
Xu *et al.* [16]	2020	Survey for FL in HI	X	X	✓	✓	✓	X	The survey proposes details of communication efficiency of FL models, presents the summary of recent works in FL in healthcare domain	The survey is mainly concentrated on biomedical space, and thus lacks in the discussion of different prediction models.
Rieke *et al.* [14]	2020	A review on digital health with FL	X	✓	✓	✓	X	X	Presents a clear compact scenario on the importance of healthcare stakeholders, and presents various FL model design choices	A detailed architecture, system model that ensures privacy of records is not discussed
Hakak *et al.* [23]	2020	FL in edge-assisted healthcare data analytics	X	X	✓	✓	✓	X	Efficient cloud based FL models are presented with focus on improved resource management through cloud services	As cloud models are prone towards high-end latency, single-point attacks, discuss on the same is not presented.
Aledhari *et al.* [4]	2020	A comprehensive study on FL, enabling technologies, protocols and applications	X	✓	✓	✓	✓	X	Techniques, market implementations, and business models are explained with adoption of FL in detail	Smooth network infrastructure and concerns about fairness-related issues are not mentioned
Mammen [24]	2021	Adoption of blockchain in asynchronous FL, one-shot FL, and FL-as-a-Service (FLaaS) is presented	X	✓	✓	✓	✓	X	Architecture of FL and recent deployments are presented properly	Application categorization, and detailed analysis is not presented
Zhang *et al.* [17]	2021	A review on Federated Learning	X	✓	✓	✓	✓	X	Categorization of FL, methods for solving heterogeneity, and privacy methods are discussed in detail.	Application for privacy, access restrictions in HI are not presented
Zheng *et al.* [25]	2021	A review on applications of FL in smart cities	✓	X	✓	✓	✓	X	Reliable, real-world applications of FL are presented in context of smart cities	Categorization of FL and associated metrics is not considered
Pfitzner *et al.* [18]	2021	The article surveys the background of FL learning, and its integration with ML and DL algorithms in healthcare	X	✓	✓	✓	X	✓	Extensive discussion on FL datasets are considered and presented, alongwith the challenges of each dataset (in terms of data split and heterogeneity) is discussed	Security and privacy issues with FL aggregation are not presented
Prayitno *et al.* [19]	2021	The surveys discusses the requirements of data-anonymity and security in FL-assisted healthcare	X	✓	✓	✓	X	X	The survey highlights the security attacks like model inversion and membership attacks	The survey does not address the research challenges of hybrid non-IID features
Long *et al.* [20]	2022	The surveys explores the issues and challenges in presenting an open-healthcare setups	X	✓	X	✓	✓	X	Specific focus is given to privacy rules and its protection and ownership in healthcare domain	A use-case of privacy-preservation scheme in FL-based healthcare is not discussed
Joshi *et al.* [26]	2022	Key FL algorithms for distributed model setups, alongwith their applications are discussed	✓	✓	✓	X	✓	X	FL tools, emerging frameworks, and data collection and preprocessing in FL is explicitly discussed	Security and privacy prospective of centralized and decentralized FL aggregation are not presented
Tedeschini *et al.* [27]	2022	The survey analysis different FL techniques in real-time medical setups	✓	✓	X	X	✓	✓	Network protocols for sensor-driven FL communication is presented, and novel frameworks of decentralized FL is discussed	The survey does not discuss the challenges of data heterogeneity, and focuses on pre-trained models and available datasets for segmentation purposes
Nguyen *et al.* [28]	2022	Recent advancements in FL for smart healthcare setups are discussed like resource-intensive FL, secure FL, and collaborative FL	X	✓	✓	✓	✓	X	Novel solutions of FL design, specifically resource-constrained FL networks, and blockchain-assisted FL is discussed	The survey does not discuss the security protocols to support the new FL networks
Proposed	2022	An exhaustive review on emerging FL applications, architecture, use-cases, and case-study in HI applications	✓	✓	✓	✓	✓	✓	Addresses gaps in earlier surveys through discussion of FL in HI from security, application, and architecture viewpoints. The proposed survey presents an umbrella view of emergence of FL in HI applications	-

Survey Parameters- 1. Solution Taxonomy, 2. FL classification, 3. FL schemes and frameworks, 4. Privacy consideration, 5. FL applications/deployments, 6. Accuracy Metrics.

are presented in the survey, with discussion on benchmark datasets.

In edge-based analytics, Hakak *et al.* [23] presented edge-assisted data analytics framework that uses FL, using user-generated data to re-train local ML models. They have proposed an efficient, personalized edge-based federal learning approach with the help of three modules, namely, the cloud Module, edge Module, and application Module. They have not categorized the impact of cyberattacks by an adversary on the given modules. Aldehari *et al.* [4] presented an overview of FL protocols, frameworks, various applications, benefits, and challenges. The authors have compared HI's architectures, FL platforms, optimization techniques, and challenges. They discussed different FL use case scenarios, such as anomaly detection, drug discovery, and visual object detection. However, they have not presented fairness and smooth network orchestration related issues.

In smart cities and assisted infrastructures, Zheng *et al.* [25] discussed the key applications and latest results of FL. They presented a detailed analysis of FL's application and its challenges in various fields like aviation, financial, insurance, medical privacy, drug development, wireless communication, and sensor-based deployments. Then, Yang *et al.* [22] presented an FL-based architecture and for business domains. The survey reviewed training and evaluation metrics of different FL verticals. Also, they have mentioned detailed descriptions of applications of FL in various business models. But, the proof-of-concept based solutions were not discussed. Later, Weiss *et al.* [21] proposed a mobile application for the visually impaired with the help of ML algorithms. They have considered general object detection, specific object detection, depth and text extraction. They have proposed a detailed model that takes better care of individuals. However, they have not mentioned any architecture of FL, use-cases, privacy and security issues. Joshi *et al.* [26] discussed the potential of FL in healthcare, and explicitly presented FL algorithms like federated average (FedAvg), secured weighted association, and federated neural architecture search (FedNAS), with others. The FL topology is presented (placement of FL clients, and FL servers, and associated communication), and feature are engineered to support the requirements. Then, the survey shifted towards security viewpoint and discussed mechanisms like DP, secret key sharing, and secure multi-party communication (SMPC). Attack models and prevention are presented, and data and communication challenges are presented. Finally, a case-study of FL for chest X-rays for COVID-19 detection is proposed, and evaluation metrics are highlighted. Tedeschini *et al.* [27] proposed a survey of distributed FL in sensor-assisted medical setups. An FL architecture with IoT networking is proposed, that exploited the message queue telemetry transport protocol. As decentralized FL is presented, different consensus schemes like node FL, committee FL, and gossip protocols are presented. To validate the presented theory, the authors investigated the performance on a brain tumour segmentation problem, which is developed on the U-net model. Metrics like accuracy and latency in FL communication are evaluated, and discussed to realize the use-case performance in real local hospital setups. Authors in [24] presented various opportunities and challenges in the field of FL. The author has mentioned the adoption of blockchain in FL, asynchronous FL, one-shot FL, incentive mechanisms in FL, and FL-as-a-service(FLaaS). However, the detailed analysis and categorization of applications, security, architecture are not presented. Nguyen *et al.* [28] surveyed FL in Internet-of-Medical-Things (IoMT) setups, and studied the requirements of reliable client-server communication requirements. Advanced FL mechanisms in IoMT are discussed, which addresses the computational constraints of IoMT setups. Specifically, mechanisms like resource-aware FL, incentive FL, and privacy-enhanced FL are discussed. Training datasets to support the key mechanisms are explored, and the adoption of FL mechanisms in applicative use-cases like medical imaging, remote wellness monitoring, and COVID-19 diagnosis is presented. Finally, the security and non-IID issues in these use-cases are discussed, and future directions are presented. Xu *et al.* [29] proposed a low-latency automatic modulation recognition mechanisms at the backdrop of 5G communication networks to improve the accuracy of the DL models. A temporal CNN network is initiated, where principal component analysis is applied and uniform samples are achieved to reduce the computational complexity, which makes the scheme viable for lightweight setups. Rieke *et al.* [14] presented a solution for future of digital health and discussed the emerging challenges and considerations. They presented the concept of precise medical care to improve analytics through FL for precision medicine and for improving medical care. The security aspects, however, were not discussed.

D. IDENTIFICATION OF GAPS

Most FL surveys in healthcare domain have discussed FL frameworks and the general functioning of the models. As discussed in subsection I-B, in HI, both the dual issues of responsive and accurate analytics alongwith security and privacy of data-sharing is crucial. The proposed survey addresses the need of HI systems, and presents architectures, frameworks, tools, and statistical and security discussions to justify the completeness issues in earlier surveys. To validate the claim, a case-study of FL adoption to electronic health records (EHR) is presented, named as *FL-EHR*, and its potential is discussed. Thus, the proposed survey would help researchers globally propose novel solutions with FL adoption in HI.

E. NOVELTY OF THE SURVEY

The motivation and novelty of the survey is presented as follows.

- In the existing literature, different FL-based surveys are presented that focuses on secured analytics, and protective models and learning parameters [30]. Other sets of surveys are directed towards adversarial ML model to analyze threats towards encryption and secured access to healthcare models [31]. Schemes that combine lightweight security modules in HI to address a large corpus of big data with concerns of privacy and confidentiality have been proposed [32], [33]. As security and privacy play an essential role for data in the biomedical space, there is a need to present an end-to-end FL based survey in healthcare, with the building blocks that address decentralized HI analytics with the secured exchange.
- The current state-of-the-art FL in HI is still in development phases. The survey presents the existing methodologies used for the security and privacy of data. We discuss the use-cases, FL frameworks, and potential challenges not addressed in earlier surveys.

F. SURVEY CONTRIBUTIONS

The contributions of the proposed survey are as follows.

- We propose the basic fundamentals and technicalities of FL in HI, and presents the FL basics, the communication and security requirements of FL, and its adoption in HI through a high-level overview architecture of FL-assisted HI.
- We discuss the statistical and security challenges of adopting FL in HI systems, with the limitations of centralized HI model. Based on the limitations, we present the different FL aggregation techniques that are applied to HI ecosystems.
- We discuss the different application scenarios of FL with IoT-assisted healthcare networks, FL with DP, homomorphic-based FL solutions, cloud-based FL schemes in HI, and integration of FL with emerging communication networks like 5G and beyond (B5G) that orchestrates the networking requirements for HI.
- A solution-taxonomy is presented of adoption of FL in HI, based on review methodology section that selectively bifurcates out existing literature in terms of the proposed research questions of the study.
- A case-study named as *FL-EHR* is proposed that discusses a potential scheme of FL integration to collect EHR from different heterogeneous sources and form a resilient HI model that can be trained over local data sites.
- Open issues and challenges of the FL-based healthcare architecture are also discussed, with emerging directions.

G. FLOW OF THE ARTICLE

The article is divided into ten sections. Section II discusses the survey methodology. Section III discusses the fundamentals and technicalities of FL in HI. Section IV discusses the challenges in adopting FL-based HI. Section V discusses the state-of-the-art HI architectures. Section VI presents the conceptual overview of the FL-based HI aggregation. Section VII presents the solution taxonomy of FL application scenarios. Section VII presents the proposed FL-based HI architecture. Section IX discusses the proposed case study *FL-EHR*. Section X discusses the open issues and lessons learned, and finally, Section XI concludes the article with future works. FIGURE 2 shows the organization of the proposed survey.

II. SURVEY METHODOLOGY

The proposed survey on the adoption of FL in HI is designed based on the guidelines given by Kitchenham *et al.* [34], [35]. Based on this, the proposed study is divided into five possible stages/steps, such as Review plan, research questions identification, possible data sources, search criteria adopted, and the inclusion and exclusion criteria. The detailed description on the above-mentioned stages are described as follows.

A. REVIEW PLAN

The planning of the proposed survey was done systematically such as identification of possible research questions, identification of data sources, data collection, effective search strings/criteria, and inclusion and exclusion criteria for articles. To perform this study there is a need for some preplaning. The identification of relevant publications, white papers, technical blogs, books, and other works in FL has been considered for this survey that proved helpful for the proposed survey.

B. RESEARCH QUESTIONS

TABLE 4 shows the possible research questions the authors have identified in carrying out the proposed survey along with their objectives.

C. DATA SOURCES

We have considered a broad number of reputed and trusted literature for writing a detailed survey on FL-based HI. We referred only standard digital libraries and databases like IEEEXplore, Springer, ACM, Science Direct (Elsevier), ACM digital library, IET, Wiley online library, etc. Electronic data sources are also recommended for the literature survey, such as patents, related books, articles, technical reports, and technical blogs.

D. SEARCH CRITERIA

In the proposed survey, we explored related articles using the search criteria with keywords ''Federated Learning'', ''Federated Learning in Healthcare'', ''Federated Learning in Healthcare Informatics'', ''Need for Healthcare Informatics'', ''Security issues in Federated Learning-based Healthcare Informatics'' and other related keywords. FIGURE 3 shows the possible search strings used to explore the existing literature. We have majorly considered the papers that provide a survey on FL model as well as the centralized model in HI. Online references were also taken into consideration for better understanding.

E. INCLUSION AND EXCLUSION

After getting many papers, the papers that were relevant to the topic (by analyzing title, abstract, introduction, and conclusion) were only considered. The papers that included an introduction to FL and a centralized system and which have the same topic were taken into consideration. To make the survey interactive and effective, we preferred to include recent papers of the year 2021, 2022, and early access articles. For wider coverage and scope, we also considered tutorial approach articles, technical reports, patents, and other related materials. The filtering of papers and articles are based on the parameters mentioned above. Finally, we identified and focused on some papers having good citations.

III. FL-HI: FUNDAMENTALS AND TECHNICALITIES

In this section, we present the technicalities of FL and its convergence to HI. We begin the discussion with basic concepts of HI, and the requirements of HI, with the integration of FL-based HI. We present the requirement to adopt FL in HI

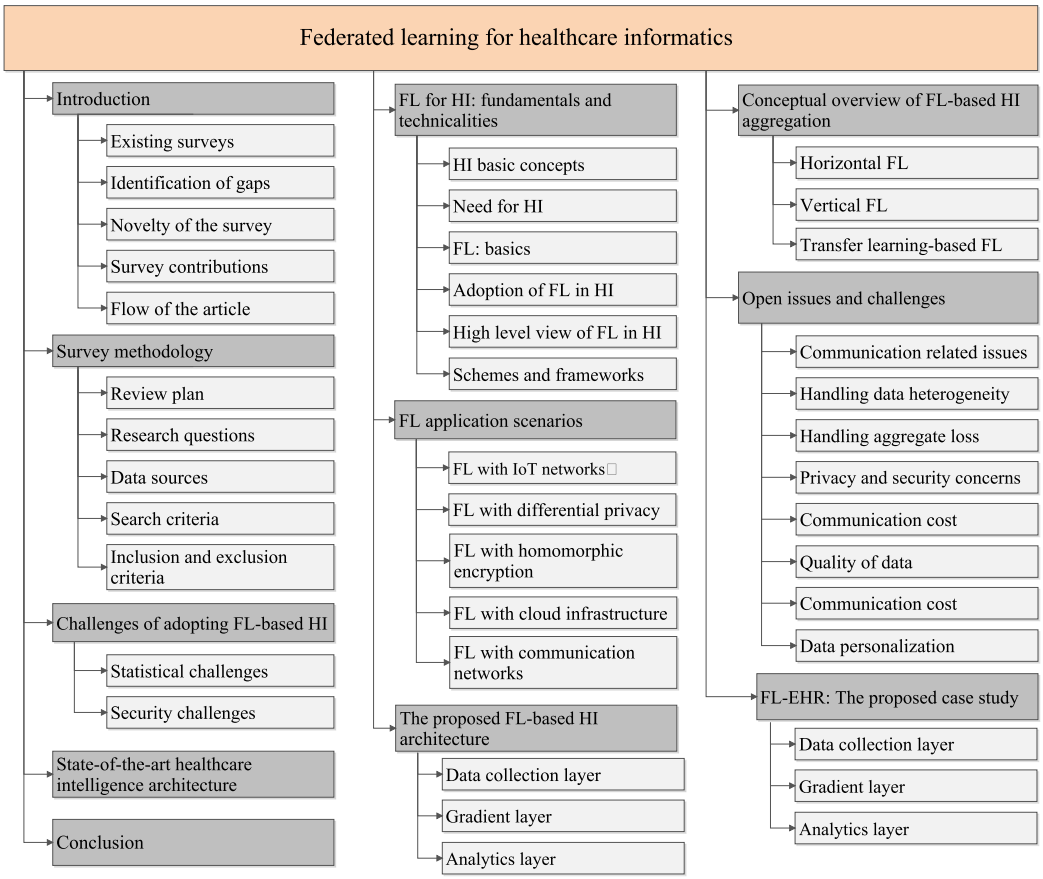


FIGURE 2. Survey organization.

TABLE 4. Identified research questions to perform the proposed survey.

Q. No.	Research Question	Objective
RQ 1	Why choose the FL model for healthcare systems over the centralized model?	To get an overview of FL technology and its advantages in the healthcare system.
RQ 2	What challenges and issues are affecting healthcare technology?	This question provided us with the knowledge of where bottlenecks are present and where improvement is required.
RQ 3	What are the advantages of using the FL model in HI	Here, we have listed the advantages with their usage in the biomedical space.
RQ 4	What are the loopholes in the centralized architecture of healthcare	The question aimed at providing us with a better solution for the problems in the current healthcare system with FL model.
RQ 5	What are other solutions that were helpful in the past?	Through this question, we were able to find more literature related to that area.
RQ 6	What do the past surveys have concluded?	It aimed in comparing our survey with other papers.

networks, and discuss the FL training process. Next, we discuss the communication and security requirements to set up FL in HI ecosystems. Next, we present a high level system design of different use-cases where FL fits in the healthcare domain. Finally, we discuss the training frameworks of FL. The details are presented as follows.

A. HI: BASIC CONCEPTS

Earlier, medical data was only limited to a particular institution. With more raw data, more accuracy, precision, and higher efficiency can be calculated. Earlier, in the centralized model, the sharing of the model was not there. Hence, lesser knowledge was there, resulting in lesser precision in the

Search String = "Federated Learning" + "X"
X:= "healthcare" OR "healthcare informatics" OR "need in healthcare informatics" OR "security issues" OR "privacy issues" OR "gradient weight compression" OR "applications" OR "open challenges"

FIGURE 3. Search strings.

model due to lesser input data. But, with HI, storage of data digitally is possible as a result of which shared users can use these large voluminous data and models can be made precise due to it as well as data can be obtained to remote users [36]. Much knowledge can be obtained from these data stored in EHR. Due to a large number of data, available doctors can obtain more knowledge with higher efficiency.

The integration of healthcare with technology gives rise to HI. According to Iyengar *et al.* [37], the relevance of healthcare and life sciences on information technology has been since 1950. Due to the digitization of healthcare faster, privacy-preserving and precise models can be made which are useful to both doctors and patients. For the protection of privacy and security, regulatory compliance requirements, such as the US health insurance portability and accountability act and EU GDPR, are working at various stages [37].

B. NEED FOR HI

Nowadays, a massive amount of medical data is generated from several medical institutions, patients, and healthcare professionals. But, it can be difficult to share this medical data of patients among different healthcare centres. For example, suppose there is an emergency regarding the patient's health and healthcare professionals need to share their health history with another hospital to get the advice of an expert. In that case, this leads to the need for HI. With the help of HI, data related to the patient's health history can be accessed across multiple healthcare centres using FL. An FL model is used to overcome the privacy and security issues while sharing the patient's sensitive data among multiple users [38].

As FL applied with HI ensures that patients' medical data is kept locally with the users, trained models from the local server can be transferred to the cloud to process the model further while preserving the sensitive data at the local server. Therefore, HI enables remote access of sensitive data across multiple hospitals digitally. Still, if it is applied with FL, then patients sensitive data can be secured by keeping it at the local server [39].

C. FL: BASICS

The convergence of healthcare 5.0 with AI has disrupted innovations in medical analysis, particularly in the way of data handling and distribution. Thus, healthcare analytics have become a separate domain of study, where different solutions of ML, and deep learning (DL) based analytics are presented [40]. Modern DL requires curated and big data sets to achieve accuracy and fairness. For example, a DL-based tumour detection algorithm requires a large database where the entire spectrum of possible tumour anatology and input types should be present. The challenge of data collections, and heterogeneous data formats, often form a roadblock to collecting massive data to meet the DL model requirements. Moreover, with rising privacy concerns, the data has highly sensitive attributes and requires proper data anonymization techniques before the public release of datasets. Thus, the collection of datasets requires time, effort, and medical resources, and have low business impact value, if the models are not fine-tuned to meet the dataset requirements.

To address these challenges, recently, FL is an emerging paradigm that provides solutions for data governance and maintains the privacy of training models as data exchange is not required. FL forms a consensus-based model training, where the local datasets are trained on mobile devices, or edge nodes, and thus data remains safe within organizational boundaries. It is generally based on the traditional machine learning model, but it works on the principle that the data generated from multiple devices must be secure using a decentralized architecture. Once data is trained, the parameters and local gradients are transferred, and the process is iterated to improve accuracy. Each device has its local copy of training data in FL, which they use to train the model locally. Then, the trained models from local nodes are sent to the cloud to process the models further and get a single model out of it.

Thus, a successful FL model can preserve the sensitive patient information identifiers, and optimally customize the model parameters according to data distribution, level of access, and patient requirements. However, FL models are faced with the challenge of optimal progression of the model accuracy, without sacrificing the privacy of the data. FL also handles the issue of expensive centralized training and communication bottlenecks on the central server that handles the analytics. FL is thus suitable for resource-constrained healthcare Internet-of-Things (HIoT) setups, as data collected through sensor nodes for different users can have high non-IID, and thus traditional ML and DL models would render ineffective on such data distributions.

An interesting use-case of FL to collaboratively train a model across different mobile devices was proposed by McMahan *et al.* [41], where the federated averaging (FedAvg) algorithm is proposed. In FedAvg, the clients, or local nodes, independently train the local data via deep neural networks and average the values. A particular use-case of the FedAvg algorithm is the GBoard word prediction in mobile devices. The challenge is to optimize the uploading and downloading of data between local nodes and global servers, and thus researchers have proposed solutions that optimize the communication requirements, which reduces the potential overheads. For example, Wang *et al.* [42] proposed a network control algorithm that dynamically changes the number of local updates before the model uploads, and forms

a resource tradeoff condition between computing power and network usage bandwidth. Sattler *et al.* [43] proposed a sparse binary compression algorithm after the model is learned, and sends the compressed gradients to global models to preserve bandwidth.

1) FL TRAINING PROCESS

FL follows the distributed ML training approach, where the direct data sharing with the global cloud healthcare server is not shared. Instead, periodic updates from local mobile devices (edge nodes, gateways, or local mobile with real-time monitoring applications) is shared to the healthcare cloud server. From the N different clients, we share the data periodically through a local aggregation process, and the global model parameters are updated. FIGURE 4 depicts the high-level overview of the FL training process. The entire FL training process can be divided into four phases, which are depicted as follows.

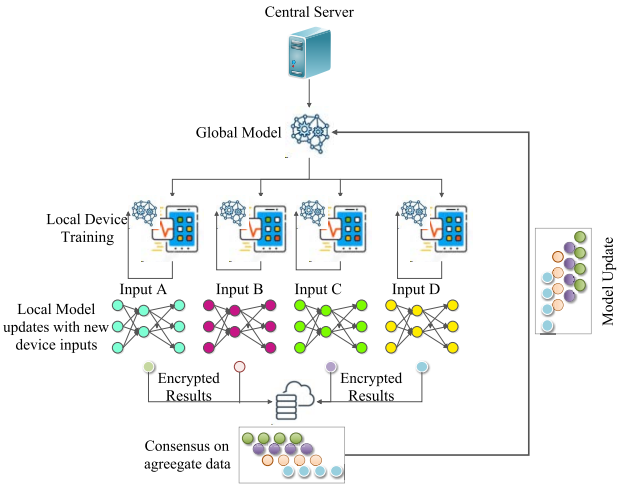


FIGURE 4. FL training process.

1) *PHASE 1: Global model initiation*- In this phase, the initial setup parameters C_p, and the global model G_m is fixed. Now, the global healthcare server G_s decides on the number of local nodes to which it would communicate the information $\{C_p, G_m\}$. G_s, then broadcasts the information to all n local clients $\{C_1, C_2, \ldots, C_n\}$ to participate in the collaborative training process. All C_n now download $\{C_p, G_m\}$ from G_s, based on the networking channel statistics. Thus, in FL, the global updates might be propagated to local nodes at different time instants, depending on the communication channel specifics.
2) *PHASE 2: Local model updates*- Each participant C_n generates a local model L_m, with local parameters L_p, based on the local data distribution, specifics, and privacy requirements of the patients. Once C_n downloads $\{C_p, G_m\}$, it updates its own local parameters as L_p^0, which indicates the first iteration update. Consequently, the process is iterated k times, and the local update of L_p after k iterations, is defined as L_p^k. The parameter updates are fine-tuned to minimize the local loss function $min(F(L_p))$. The local loss function after k iteration can be defined as $min(F(L_p^k))$. The local models are then shared with the global server G_s through an secured aggregation policy.
3) *PHASE 3*: Local model aggregation- For aggregation of L_p, different algorithms like FedAvg [41], federated learning with personalization layer (FedPer) [44], federated learning with matched averaging (FedMA) [45], and federated distance (FedDist) [46] are proposed. In FedAvg algorithm, every local client C_n initializes a random neural network on local device in terms of local weights, layers, and neuron sets. When the on-device training is completed, the local model weights, denoted by $W(L_m)$ are communicated to G_s. The aggregation principle is based on weighted average, and thus is biased towards clients with more data, as they influence the global model. To address this limitation, FedPer is proposed, where every C_n sends the neural base layers to G_s, and retains the specifics of other layers. The reason of communicating only the neural base layers is trivial, as they represent the representational learning. Higher layer are more shifted towards decision specifics, and thus are not required by G_s. FedPer achieves more accuracy on distributed and non-IID data, but is limited that G_s might make a partial decision to favour any local C_k. Another approach is FedMA, which constructs a global shared model based on averaging values of hidden layers (convolutional layers, hidden states, and connected layer weights). FedMA is specifically useful on real-world datasets, and reduces the computational overheads. Finally, FedDist model computes the distance between neurons with similar features. The algorithm also identifies diverging neurons based on euclidean distance, and is suitable for sparse data, in which specific features are captured in small subsets, or data points.
4) *PHASE 4*: Global model aggregation- After G_s receives L_p from local participants based on a chosen aggregation strategy, it performs the computations to update the global model G_m. The updates are re-shared again with the local C_k, until the model reaches a convergence, and the global loss function is minimized. The global loss function, denoted by G_L^k is presented as follows.

$$min_k f(G_l) = \sum_{i=1}^{k} Q_k F_k(w) \tag{1}$$

where k is the total local participants, and $Q_k \geq 0$ is the relative impact of local participants on G_m, with the constraint $\sum_k Q_k = 1$. $F_k(w)$ denotes the expected loss in prediction values of the local sample inputs for any k^{th} local node. As any C_k is assumed to have q local samples s_q, thus $k = \sum_q s_q$. Overall, the relative impact of any C_k on G_m is denoted as $I_k = F_k(w)/q$.

2) COMMUNICATION AND SECURITY REQUIREMENTS

In this section, we discuss the impact of communication and security schemes in the FL process. The details are presented as follows.

1) Communication Requirements- In healthcare IoT setups, effective FL schemes are required to communicate data over a communication network, with the desired properties of low power and resource consumption. Thus, constrained FL schemes requires optimization in both local computations, aggregation, and the final global computations. The local device performance over low-powered channels (ZigBee, 6LowPAN, and Bluetooth) mainly depends on the size of local datasets, the available energy, and local computational resources (CPU and I/O requirements). In the same direction, the energy consumption by a local device that operates at frequency f, with dataset of size C, and performs k local iterations is given by Khan *et al.* [47] as follows.

$$E_l = k(\alpha \xi C f^2) \tag{2}$$

where E_l denotes the energy requirements of any l^{th} local node, α, and ξ are CPU and I/O dependent parameters, based on cycle rates required to process C. The local computation model time T_l is then presented as follows.

$$T_l = k(\frac{\xi C}{f}) \tag{3}$$

From the presented equations, there is an associated trade-off between the minimization of local model computation time and the energy requirements. As accuracy is a prime requirement in analytics, the dataset size becomes a dominant factor that drives T_l. Thus, fixed-sized datasets are generally preferred, so that the operating frequency f can be controlled to minimize E_l. However, there is a significant heterogeneity in the models, datasets, and resources, and thus a unified condition which optimizes all the parameters is difficult. Thus, authors in [48] presented the notion of relative local accuracy, denoted by ω, that shows that small values of ω have better local accuracy.

Once the learning parameters are aggregated, they are sent to the global model. The transfer time would depend on the channel access scheme, and the available radio link. Considering the uplink frequency as f_u, and the channel capacity as C, then the time taken to communicate the local model parameters of size d is presented as follows.

$$G_t = \sum_{I \in k} (df_u/C) \tag{4}$$

where G_t denotes the communication time to the global model, I is the iteration steps (converging towards the k^{th} step, and d is the distance parameter. The energy consumption, denoted by E_G would depend on the transmitting power P_t, and is given as follows.

$$E_G = \sum_{I \in k} (dP_t/f_u C) \tag{5}$$

2) Security Requirements- In FL schemes, security and privacy is of paramount importance. It is specifically more important in case of healthcare setups. However, FL assures privacy as data is trained locally, but still it suffers from different security concerns. In case the end-device and aggregation server exhibit a malicious behavior, then the local learning would be compromised. Thus, Nasr *et al.* [49] propose addition of noise to local learning models before sending the local parameters to the aggregation server. The scheme assures a local DP, and introduces the local data from p local participants, based on Gaussian and Laplace mechanisms. In Laplace mechanism, we add local noise based on Laplace distribution to the local datasets, however, it reduces the overall sanctity of the data. Another approach involves addition of DP based on exponential distribution, where the local learning model updates are mapped to an exponential distribution. Exponential distribution performs significantly better than Laplace forms. Similarly, the dataset may be mapped with a Gaussian distribution. As Gaussian distribution is addictive, it has significant benefits of simple analytics. The addition of noise generally improves the data privacy, but accuracy is generally compromised. Thus, there exists a tradeoff and convergence between privacy preservation and noise addition to the local models.

Another approach is through SMPC scheme, that encodes the local models. SMPC assures a cryptographic primitive that distributes a computation among multiple local nodes, where each local party is not able to see other local data, but each party jointly computes the security function based on send inputs. Thus, an adversary outside the SMPC ecosystem, would not be able to reconstruct the security function, as the initial and the partial computation conditions are not known.SMPC incurs high communication overheads, and might not be useful for high-privacy setups.

In such cases, a hybrid approach that combines adding noise and SMPC is generally preferred. Another approach is usage of lightweight security paradigms to secure the local devices an the assisted network setups. Lightweight authentication and access schemes should be designed that can authenticate and verify the updates at the aggregator node setups is important. Moreover, communication of encrypted learning updates to aggregation model should be supported with edge resource setups. A proper tradeoff is required between scheme performance and the security overheads to address the resource requirements of the healthcare ecosystems.

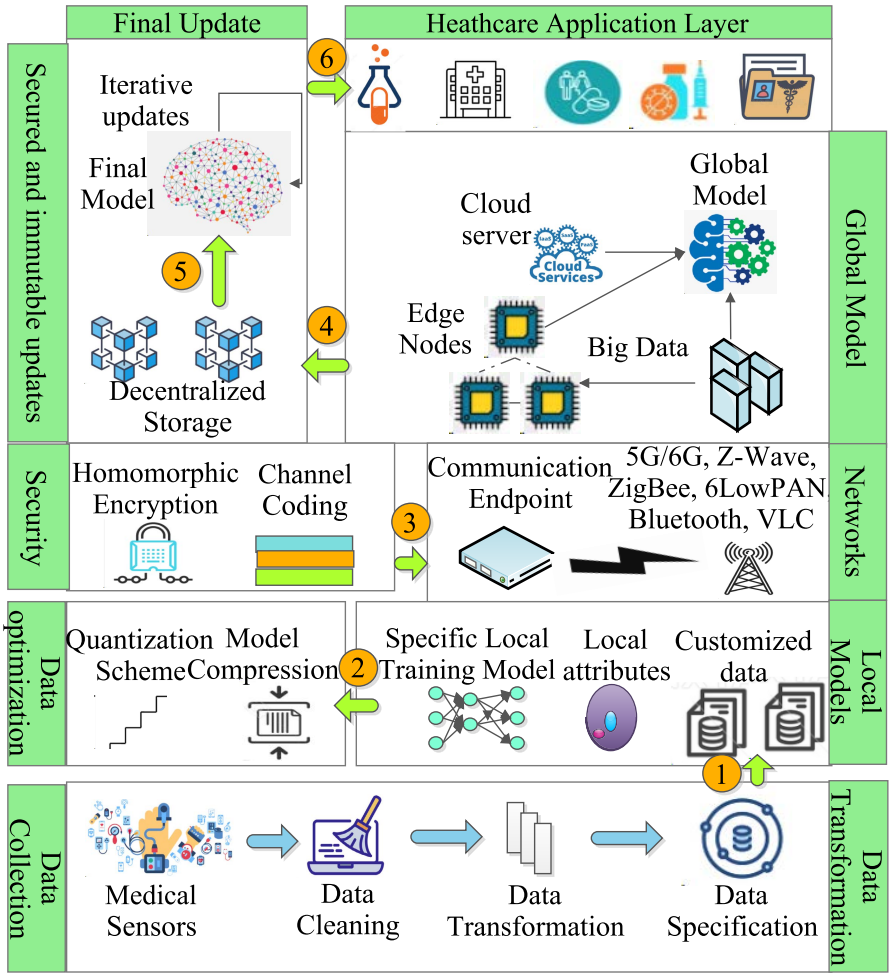


FIGURE 5. A high-level overview of FL-assisted Healthcare ecosystems.

D. ADOPTION OF FL IN HI

In 1996, the US congress formulated the HIPAA guideline for EHR storage [36]. HI is an emerging field in which all the records are maintained digitally. By 2007, EHRs which are initiated in Finland which were supposed to be fully functioning [36]. Australia is also taking steps towards EHR maintenance [36]. With digitization, the threat to security and privacy arises. If any violations are observed in security and privacy, then a maximum of ten years of jail sentence with a penalty of $ 250,000 is given [36].

With FL, these risks can be reduced to a great extend as FL provides a solution to the security and privacy questions of the users. With methods like HE, DP in FL models, users' data is much more secure. Furthermore, with emerging wireless technologies, such as B5G, the latency in communication models have significantly decreased. With FL, precision can be increased as more raw data is available, so a more precise model can be made, which can help doctors and medical staff predict diseases more accurately. The FL healthcare model is divided into three basic layers:

- *Data Layer*: The data Layer is the first layer where the data is collected from various sources like patients, Previous medical history, any recommendations, and the suggestions of doctors, nurses, etc. This data is trained into the individual model.
- *Integration Layer*: The individual models are concerted into local models, and various local gradients are computed in this layer. These local servers then aggregate the local data. Technologies like fourth generation (4G), 5G, and 6G can be used for better communication.
- *Analytical Layer*: The main server uses this aggregated gradient and converts it into a global model. The updates are also commuted when the clients send them.

According to the FL-based healthcare framework proposed by Vaid *et al.* [50], the central server initialized the main model by using random parameters. This model is sent to each client for updates and was trained for one epoch. Again, this model was sent to the main server and with an updated model for many epochs and from aggregator sent to various clients. In this way, the cycle is repeated many times.

E. HIGH LEVEL OVERVIEW OF FL-BASED HI

In this section, we present a high-level overview of FL-enabled healthcare architectures. FIGURE 5 presents the architecture design. At the lowest layer, we assume a medical data collection setup, where healthcare data is collected from primary setups, like wireless body area networks (WBAN), HIoT setups, and others. A secondary setup source would be recording the EHRs and personal health record (PHR) indicators. In the data collection process, we assume that the local nodes (or end-devices), collect data based on specific learning criteria. Once data is collected, we go through the data cleaning process, where missing data, outliers, and other values are added. Data collection follows a data transformation phase, that aligns the data according to the requirements of the learning model. After the data is cleaned, the end users, or healthcare users, may specify specific privacy constraints so that external linkage attacks are not possible. Sensitive attributes are hidden, and data specification phase builds the data according to the model required for learning. Thus, at local nodes, data is customized to meet the model requirements. The local datasets can be trained through different models (step 1) like feed-forward neural networks, convolutional neural networks (CNNs), support vector machines (SVM) classifiers. In case we require data analysis over a time-period, where data distribution follows a time-series, long-short term memory (LSTM) is a viable choice.

Once the data is trained through local learning model, the models are required to be compressed, and data quantization takes place (step 2). As model results are bulky in resource-constrained healthcare setups, a compressed model sent over the network would save space, as well as reduce the transmission time. For compression, normally, lossy FL models are preferred. For quantization, both the local and the global updates are effectively quantized via stochastic gradient descent (SGD), or gradient quantization and encoding (QSGD) schemes before sending to the aggregator node. Another quantization models proposed includes subtractive dithered lattice quantization, hexagonal lattice quantization, that performs better with lossy compressions. The amount of compression and quantization would depend on the network characteristics, and the resource banks at the disposal of the local nodes.

Once quantization and compression is complete, effective encryption and channel coding mechanisms are employed (step 3). As indicated, we can preserve privacy of healthcare data through SMPC scheme. A popular encryption technique employed in FL, HE, allows encryption between nodes where computations are performed without the requirement of decryption of the data. Once the data is encrypted, and finally when decrypted, it would produce the identical forms of the operations that are carried on the data itself. Thus, HE allows security and privacy of the data between multiple local nodes. It allows effective data sharing, with less computational overheads, and thus is useful for predictive analysis in healthcare setups. It also allows the data sent to the aggregation server to perform model updates. However, HE adds exponentiation overheads, which are complex operations. In case a large amount of data is encrypted, HE requires high network resources for transmission. A lightweight scheme, *BatchCrypt*, is proposed by Zhang *et al.* [51] that allows encryption on a batch of quantized gradients. In this case, a gradient-wise aggregation policy is designed at the aggregation server, and it combines both the quantization and encryption steps to improve on the communication requirements. *BatchCrypt* has the ability that preservers the quality of the aggregated model over cross-domain FL models. Once encryption is completed, channel coding is done that converts the data to be sent over communication medium. Normally, in wireless setups, channel coding should be designed with low latency principle. Different channel coding schemes are designed, like shortblock length codes that are designed specifically for shorter packets, and useful for fourth generation-long term evolution (4G-LTE), or 5G-ultra-reliable low-latency communications (5G-uRLLC) networks. Another specific schemes include low-density parity check schemes (LDPC), preferred in B5G networks, or 6G-uRLLC service. LDPC allows less errors while transmission [52]. Turbo codes are also specifically used along-with convolutional and polar coding schemes.

Once channel coding is completed, the updates are communicated to the aggregation server through a communication endpoint, that normally forms the basic networking stack for communication. For aggregation of data, we prefer the edge servers and cloud servers, so that they can address the resource requirements of the global server (step 4). To form trusted FL scheme, blockchain-based aggregation is a preferred choice, that allows local miners to add the data post verification to local mining pool [53]. Another schemes for aggregation includes FedAvg, or the FedMA approaches. Once data is aggregated, a blockchain-consensus scheme is executed to update the block proposal. Once block proposal is updated, the global model aggregation takes place, and the results are communicated to the final model (step 5). The final model is updated with the trusted local model learning updates, and the process is re-iterated till the convergence of final model is reached, and global model loss is minimized. Finally, the results of the model are communicated to different healthcare stakeholders like drug companies, research labs, hospitals, vaccine stakeholders, and patient EHRs/PHRs (step 6). The advantage of distributed aggregation is that it is secure against attacks of centralized aggregation, where the aggregation server might fail due to denial-of-service attacks on the server [54]. Moreover, centralized servers suffer from high-end latency, and computational bottlenecks, which is alleviated in the distributed approach.The only limitation is the computational complexity of the blockchain-mining process, as it involves a consensus algorithm for adding the block to the global chain. In such cases, effective low-powered consensus, and mining strategies are required to address the scalability of node additions.

F. SCHEMES AND FRAMEWORKS OF FL-HI

In this subsection, we present the different FL schemes applicable in healthcare ecosystems. Mainly, the FL schemes are designed based on two challenges- the statistical heterogeneity of local nodes, and the computational resource requirements of communicating the updates to the global server. Next, we present the different frameworks that are used for designing the FL schemes. The details are presented as follows.

1) FL SCHEMES

In this subsection, we present the different FL schemes applicable in healthcare ecosystems. As depicted, the goal of FL is to minimize the global loss function, where we consider the global model weights, and the local nodes as participants to the FL model update. A normal SGD update would require one communication round to update one local node, and thus it takes a long time for the global model to converge the results. Thus, a variant of SGD algorithm, federated SGD is normally employed, where all the end-devices compute a gradient $\eta_n = \nabla F_n(w)$, where F_n is the n^{th} round function, and is based on local datasets, and aggregate the update in a single round u_t. The next round, denoted by u_{t+1} is mapped to u_t as $u_{t+1} \rightarrow u_t \eta \sum_{i=1}^{N} \frac{k_n}{N} \eta_n$. Federated SGD improves the performance of local SGD by reducing the number of iterations for global model convergence. Another approach, FedAvg [41] involves a simple average on the local models at the aggregation server, and are particularly suitable for heterogeneous datasets. FedAvg tunes the local hyper-parameters so that the learning rate of the global model improves. However, with more local iterations, the model might get stuck at local minima, and convergence takes more time. Moreover, with higher learning rate would cause local models to not complete its learning in specified time, and thus, many local models are not able to participate in the collaborative learning process, that induces a biasness in the learning rate. A better approach was proposed in FedProx [55], that adds a local proximal term to the local loss function of the FedAvg model, which is defined as follows.

$$min_w q_n(\omega, \omega^t) = F_n(w) + \frac{\mu}{2||\omega - \omega^t||^2} \tag{6}$$

The proximal term ω handles the statistical heterogeneity of the local updates in a better manner, and thus less impact is on the model learning rate. At given time instant t, w denotes the model weights, and μ is the statistical variance.

Existing literature on FL schemes are presented as follows. Han *et al.* [56] investigated a robust zero-watermarking scheme based on federated learning for securing the healthcare data. Then, the edge-assisted healthcare framework is discussed for data analytics using FL in [23]. It majorly focuses on facilitating privacy in the system. Then, the authors in [57] explored the PFL-IU, a privacy-preserving FL scheme with irrelevant updates in E-healthcare applications to ensure the accuracy and robustness in the system.

Then, Li *et al.* [58] also discussed ADDetector, a privacy-preserving FL scheme to preserve the integrity of user's data. Later, Thwal *et al.* [59] presented an FL scheme combined with a DL approach in maintaining the healthcare system and its privacy efficiently. Then, the authors in [60] designed an FL-based dynamic contract mechanism to ensure the user's efficient participation and privacy in the healthcare system. Later, the authors in [61] presented an FL learning scheme combined with deep learning to mitigate the data integrity issues of the proposed scheme in [59] to preserve the data integrity and privacy in IoT-enabled healthcare systems using a secure access control mechanism.

2) FEDERATED LEARNING FRAMEWORKS

Now, we present the different frameworks for implementation of FL over healthcare setups.

1) PySyft [62]: PySyft is a framework designed in Python language and is based on PyTorch native interface. It assures that local learning is encrypted through SMPC and HE. Moreover, the users can include noise in the local datasets, and thus has variants to include DP. In PySyft, we include the local nodes as virtual machines, and a pointer tensor is used to locate the stored datasets.
2) FedML [63]: FedML is another framework that uses an application programming interface (API)- based library calls to implement FL. It is open-source, and uses FedML-core, that performs the operations of model training, and uses distributed remote procedure call (RPC) communication to send the updates to the aggregation server. Low-level API calls are made by the communication module between local clients and server, and the training module is implemented through PyTorch. It also provides a topology manager setup that allows the design of FL topologies and implement the communication model between the nodes.
3) TensorFlow Federated (TFF) [64]: TFF, presented by Google, is an open source framework, that can be used to train machine learning models on different datasets. For FL learning, TFF has introduced a FL-core API library, that provides an interface to make high-level calls to the local training models from the global models.
4) LEAF [65]: A specifically designed FL framework to support a variety of FL applications like meta-learning, multi-learning for different heterogeneous datasets. In LEAF, we can include open-source datasets like modified national institute of standards and technology (MNIST), anadian institute for advanced research (CIFAR), and many others, and also have a provision to make our own synthetic FL dataset. LEAF provides ready-made implementations of different FL aggregation algorithms like SGD, Federated SGD, FedAvg, and minibatch SGD. It allows ready-made call APIs to support the global learning, and improves the rate through defined set of hyperparameter tuning interface.

5) Paddle FL [66]: An open-source framework, PaddleFL, is used to offer FL implementations to support different verticals like IoT, computer vision libraries, natural language processing, with specific implementation of transfer and multi-task learning. FL allows job-scheduling and large scale distributed learning, powered by Kubernetes, and allows full stack development options.
6) OpenFed [67]: A comprehensive novel FL framework, OpenFed, is based on PyTorch, and provides an excellent FL toolkit, with rich libraries, and topology design. OpenFed allows automatic topology selection, with the power to decompose an entire FL scenario into individual atomic units. The framework can implement standard FL algorithms with rich configuration possibilities like partial-activation of local client nodes, sampling and dataset partition, and non-IID distribution. It allows the validation of different frameworks, and have third-party integration to other FL frameworks. It includes algorithms like SGD, FedAvg, with data augmentation, and local early stopping.

IV. CHALLENGES OF ADOPTING FL-BASED HI

In this section, we present the statistical and security challenges of adopting FL in HI. The details are presented as follows.

A. STATISTICAL CHALLENGES IN FL-ENABLED HI

As systems in FL learning are distributed and heterogeneous, each system's resources and computational capabilities vary, owing to the different usage of the CPU, memory, and I/O. Moreover, every system is connected through different networking links, and thus the communication latency varies. Due to this heterogeneity, the learning rate varies on other nodes, and some nodes might experience drop-outs at different learning iterations. This results in an inconsistent state of learning models, which levels up the data inconsistency. In healthcare setups, the sensor devices generate high quantities, and the data collection is not homogeneous. Thus, the collected data is non-identical, redundant, and inconsistent. Coupled with system capabilities, the problem further intensifies, and therefore the underlying statistical context while the learning models from the prediction tasks vary significantly.

The collected data points among the distributed nodes in terms of structure, labels, and distributions also differ. Thus, even the consistent global model, which the global model downloads, differs due to statistical differences on local nodes. The underlying statistical structure and associated distributions are not identical, which magnifies the model's complexity. Thus, empirical results and gradients might differ, and meta-learning is a preferred approach in such cases. A more native solution for better personalization results is a uniform computational device setup and device setup among local nodes.

To address the statistical heterogeneity issues, two techniques, meta-learning and multitask learning, are proposed by researchers. The FL framework is optimized in multitask learning to allow personalization through separate learning models, but the representation is a shared instance. The technique is proposed in Smith *et al.* [68], in which the authors designed a *MOCHA* framework that guarantees convergence for small FL networks. Once scaled to large networks, the problem is non-convex. Other approaches use a meta-learning framework that improves the learning rate through the multitask information through models like vanilla FedAvg, which averages the uneven distributions. Another approach is proposed by authors in [69], that uses agonistic FL, that exploits a min-max optimization on a centralized model for a given target distribution. The approach could be modelled with schemes like q-FFL [70], in which the nodes that suffer more losses are provided with some extra relative weight so that the overall variance is limited. However, the approach is not viable in case of high faults in systems, and in such cases, a local optimization solver is required. To improve the performance of FedAvg, a new model FedProx is proposed [71]. The idea is to improve the FedAvg through the underlying concept of dropping extreme values due to system and statistical constraints. In FedProx, a local partial update is performed across multiple devices that fine-tunes the local epochs. Moreover, the local updates are restricted in boundary conditions, and thus the model behaves under defined bound-box only. FedProx can be further combined with convex and uniform bounded gradients model, which heuristically models the statistical heterogeneity. However, the approaches increase the computational burden on the server, as resources like network bandwidth are greatly consumed. In such cases, auxiliary data can be prefetched through proxy servers to model the network latency. In medical scenarios, recent frameworks like federated AI technology enabler (FATE) is developed as an open source project by Webank's AI organization. In FATE, big data analytics is integrated with privacy regulations to instill high operational performance in medical systems. Another practical framework is Clara software defined kit (SDK), which is released by NVDIA edge AI platforms. In CLARA, collaborative node learning and training is carried out seamlessly to train the weights of the global model. To reduce the computational constraints, the partial model weights are shared with global model. This step also prevents the model inversion attacks on the medical ecosystems.

B. SECURITY CHALLENGES IN FL-ENABLED HI

Security in HI plays a significant role as privacy and integrity of data of users should be maintained. Existing methods for security include various technologies like blockchain, cryptographic primitives, biometrics any many more. Dwivedi *et al.* [36] proposed a security preserving framework which is a combination of biometrics, public key infrastructure, and smartcard technologies for a secure healthcare system. Public key infrastructure is a cryptographic method for encryption and decryption of data, whereas biometrics are the physical or behavioral traits of humans for their identification like fingerprints [36].

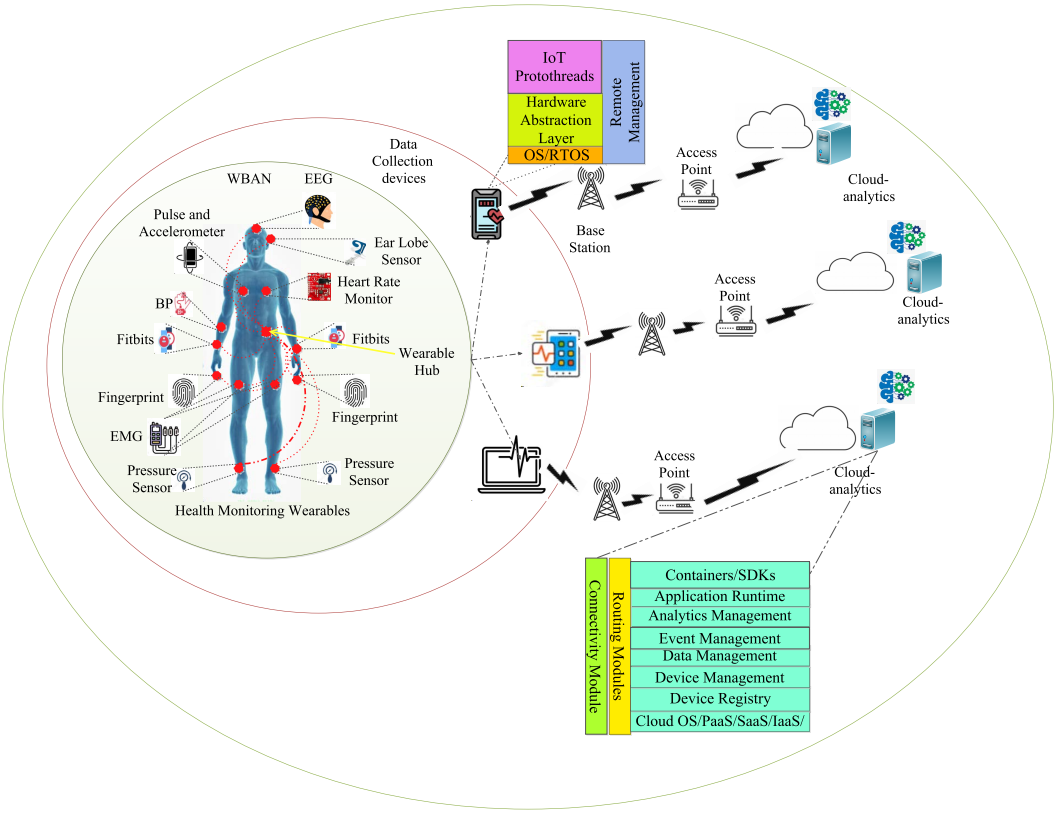


FIGURE 6. Centralized HI system architecture.

Moosavi *et al.* [72] has proposed a system for mobility enabled healthcare IoT systems for authorization and authentication of data using smart gateways. Hölbl *et al.* [73] reviewed blockchain methods for security in healthcare. The use of blockchain is rapidly increasing as sharing and management of data becomes very easy. Bodkhe *et al.* [74] surveyed the use-cases of blockchain in Healthcare 4.0, specifically in the management of volumetric health record data, and presented schemes to store the data into timestamped ledgers. A reference architecture of blockchain-based electronic record management system is presented, and open challenges of storage in terms of data inconsistency, record fragmentation, and security issues are discussed. Sun *et al.* [75] has proposed privacy and security preserving framework using cryptography which provides data integrity, privacy, access control, and accountability, and existing wireless network construction, which provides rapid retrieval of data.

FL brings privacy-preservation in HI, however, schemes are required to the aligned to be computationally applicable in resource-constrained ecosystems, energy-efficient, with low communication overhead. Privacy preservation requires two categories of privacy, one that addresses the privacy of the global updates, and one that addresses for the local updates. For the same, approaches like SMPC [76], and DP models [77] are designed. SMPC allows the federated nodes to calculate a joint function over their private data, and thus the adversary is not able to intercept the communication data. In DP, a noise function is generally added to the data, so the explicit and sensitive healthcare identifiers of patients are not exposed. In the same regard, k- anonymity models are designed, that links to specific data which can be mapped to k persons. Hence, it is not explicitly linked to a given patient. In FL, a secure aggregation technique is generally applied to guarantee low training loss and high privacy guarantees. The selection of effective hyperparameters simplifies the task and increases the accuracy of the local model. Authors in [78] proposed a local differential private algorithm with meta learning, that improves the personalization of the FL model. With DP, effective compression techniques are require that reduces the communication overheads, which is suitable to applications like IoT, vehicular networks [79], and healthcare setups.

V. STATE-OF-THE-ART HI ARCHITECTURE

This section explains the existing centralized healthcare architecture and its loopholes. It also depicts the comparative analysis of state-of-the-art FL-based centralized healthcare models. FIGURE 6 shows the traditional centralized HI model with three working layers such as physical, communication, and prediction layers. At the physical layer, sensors collect the person's healthcare information like body

temperature, sleep schedule, BP sensors, oxygen level, and so on. Researchers globally have proposed lightweight cryptographic mechanisms to secure the data access, and proposed effective data authorization mechanisms like signcryption mechanisms to reduce the computational constraints on the physical setups [85]. The data is then passed to the centralized intelligence server via an ultra-reliable low-latency wireless communication channel. The server then predicts the disease using various ML-based prediction models such as decision trees, random forest, and SVM, which sends back prediction results to the user for appropriate precautions [86]. The overall data of the physical layer reaches the mobile application, where the user can observe their data.

The communication layer transfers the data collected from the physical layer to the centralized server via a wireless communication channel faster. Here, various encryption standard protocols like data encryption standard (DES), advanced encryption standard (AES) and symmetric/asymmetric cryptographic algorithms are used to establish secure communication and prevent data leakage.

At the centralized server, the collected data is analyzed and predicted its accuracy using various ML algorithms such as artificial neural network (ANN), CNN, SVM, and many more. These models predict the type of disease the user is suffering from or any abnormal situations. The results are then passed to the user application so that they can view it.

TABLE 5 shows the comparative analysis of the centralized intelligence and FL models in healthcare.

A. ISSUES IN CENTRALIZED INTELLIGENCE MODELS

1) NETWORK CONNECTIVITY

In the traditional architecture, if a node loses its connection with the centralized intelligence server, the entire prediction model will fail as there is only one central server. Due to connectivity issues, if any node from the model loses its connection with the centralized intelligence server, it is not possible to retrieve the information. This results in an inaccurate finding due to a shortage of data. The model's requirement in healthcare needs to be highly precise and accurate, as doctors use them directly for diagnosis. For instance, in the smart healthcare system, the data from wearable devices do not reach the main server in connectivity issues. This leads to a fall in prediction accuracy, which is due to the loss of vital information.

2) CENTRALIZED NODE FAILURE

In a centralized model, if any failure occurs at the server node and no backup is available, which is the biggest limitation of the centralized model. Huge and voluminous EHRs are stored at the server and its failure can create severe issues, as EHRs can be lost due to failure at the server node. For instance, if the server node of any centralized model fails, the whole system becomes useless as, without the central server, no further processing can be done.

3) SERVER MAINTENANCE

The timely maintenance of a centralized server is essential as a plethora of data needs to be updated regularly for greater efficiency. But, in centralized servers, the maintenance becomes complex and tedious, which requires time, during which the server is occupied and cannot be used for other works. Thus, it creates a situation where the server cannot be used and the model becomes ideal for the time during which maintenance occurs.

4) COMMUNICATION

In a centralized intelligence model, there is no sharing of data between clients for efficient processing. Whereas, in FL, the model is trained with different data models aggregated to the cloud server, which has a wide variety of medical data and higher precision can be obtained. Doctors can also share their data with different healthcare centres to increase their model's accuracy as varied data is obtained. As discussed above, the data cannot be shared or distributed in a centralized intelligence model, resulting in no significant improvement. After a certain limit, the cost/benefit ratio becomes less than 1 in the server node despite upgrading hardware and software requirements of the server node [8].

5) SYSTEM SCALABILITY

With a centralized intelligence system, reliability cannot be achieved due to a limited number of connected resources, limiting the exchange of data. Thus, it also reduces the overall system's precision. With more resources, the cost of the system is increased, which restricts its scalability.

6) PRIVACY AND SECURITY

Privacy of medical data is of utmost importance, as it contains sensitive information such as the personal details of an individual, which requires high security. With time, the exponential increase in healthcare data breaches raises severe security concerns for patient data. In centralized systems, the data breaches can easily be launched at the central server, which reduces its privacy and security. This can be achieved with the inclusion of FL concept along with HE, cryptographic primitives, and data DP. Attacks like data and model poisoning can easily occur in a centralized intelligence system. Further, Fl uses data from the various localized models in encrypted form to obtain the optimal solution.

7) NETWORK LATENCY

Latency is also one of the major issues for centralized intelligence systems. The far nodes tend to have high latency, which decreases the overall efficiency of the system. In FL, the latency problem does not happen as the models are trained locally and then aggregated with the centralized model. In medical space, a time delay can result in fatal results, which restricts the adoption of a centralized intelligence system. From the aforementioned issues discussed, we infer that the centralized intelligence system is less effective than

TABLE 5. Comparative analysis of various centralized intelligence and FL models in healthcare.

Technique	Author	Year	Objective	1	2	3	4	5
Centralized intelligence	Pirnejad *et al.* [80]	2007	Reviewed challenges, strategies and complexities involved in building an inter-organizational communication network in healthcare with centralized and decentralized approaches	N	N	Y	N	N
Centralized intelligence	Kierkegaard *et al.* [81]	2011	Reviewed the effectiveness, reduce costs, accuracy, and completeness on EHR	N	N	Y	N	N
Centralized intelligence	Shanin *et al.* [82]	2018	Proposed a new EHR system device with low power and cost	N	N	Y	N	Y
FL	Xu *et al.* [16]	2020	Presented a detailed review on FL in HI	Y	Y	N	Y	Y
FL	Silva *et al.* [83]	2020	Proposed a Fed-BioMed framework with optimization methods	Y	Y	N	Y	Y
FL	Guo *et al.* [84]	2020	Proposed a framework for healthcare with higher privacy and efficiency	Y	Y	N	Y	Y

1. Privacy, 2. Security, 3. Maintenance, 4. Efficiency, 5. Cost efficiency
N: Not included and Y: Included

the FL model. The findings of various authors conflict in terms of cost efficiency, such as Shalnin *et al.* [82] suggested that the centralized system is cost-efficient, whereas Pirnejad *et al.* [80] and Kierkegaard [81] suggested the centralized models are cost-inefficient. So, FL achieves better security, privacy, connectivity, maintenance, scalability, and efficiency results. Disadvantages of the centralized healthcare model can be prevented with the help of the FL. Solutions to the issues of the centralized intelligence healthcare systems with FL is given as follows:

a: NETWORK CONNECTIVITY

As in centralized healthcare intelligence architecture, if any node losses its connection with the central server, then it can lead to a compromise result in terms of precision and accuracy parameters. These parameters are highly important in the medical field as doctors rely on these models for patient diagnosis. Instead, in FL based healthcare model, there are numerous clients. Also, a specific node does not affect any local gradient, which is computed with the help of individual gradients of any particular node or on any other criteria on which the number of individual clients is allocated to a local server. Thus, the global model is not affected by network connectivity issues of any node as an aggregation of overall data in the case of the FL-based healthcare model. Thus, with a large volume of clients and the aggregation policy of FL, network connectivity can never be an issue in FL based healthcare model.

b: FAILURE AT SERVER NODE

This becomes a significant issue in a centralized intelligence healthcare model as only one central server has all the gathered data. If it fails, the model remains of no use. Thus reliability on a central server is not recommended in the healthcare sector, which restricts its adoption. To overcome this issue, FL is used in which local servers are installed along with the centralized server, i.e., the global server. The data lost probability is highly reduced in FL-based healthcare models. The local servers have information about various clients, and this information is then collectively passed to the global server with encryption. Further, the global gradient is computed and pass it to the concerned local server. This reduces the risk associated with the patient's data. If anyone's server crashes, other servers can temporarily take their place till the maintenance of the server is completed.

c: SERVER MAINTENANCE

In centralized intelligence healthcare systems, the model updation and maintenance takes a long duration. Due to only one server, all the gathered data is accommodated needs to be updated and maintained, which becomes quite time-consuming. In the case of the FL-based healthcare model, various local servers are there which are timely updated. Also, the central server needs not to be updated at one time. This is because they directly give any updates by the clients to a local server, where local gradients are calculated accurately according to the updates. Thus, from local gradients, the global gradients will be updated automatically. If one local server is under maintenance, then the other server's work of finding local gradient is temporarily taken. Thus, the issue of maintenance can be easily solved by using an FL-based healthcare system.

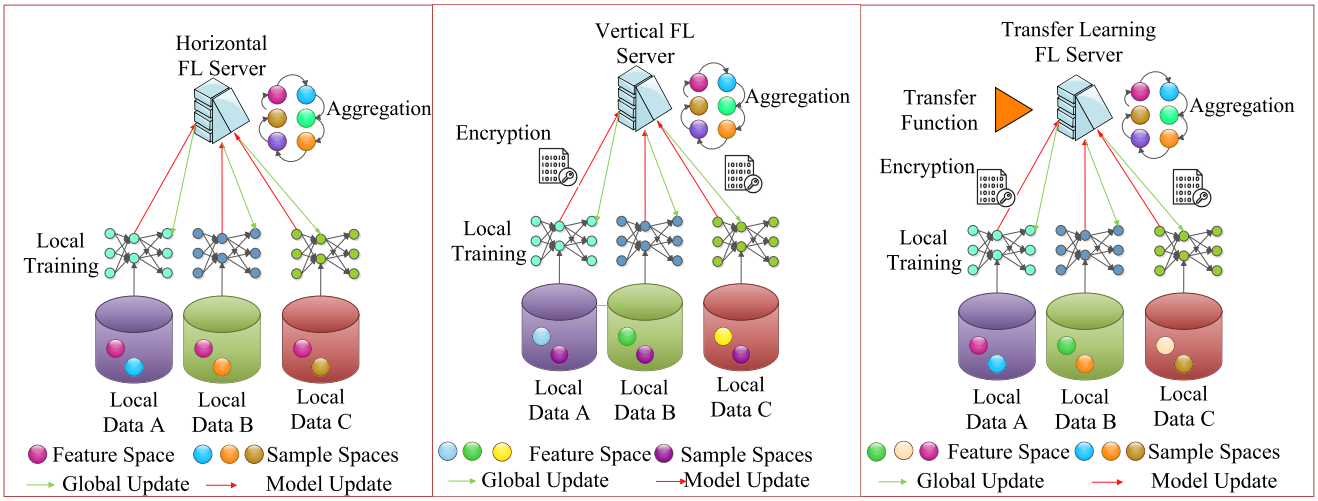


FIGURE 7. FL-HI aggregation conceptual overview: The left figure depicts the horizontal FL, centre figure depicts the vertical FL type, and the right figure indicates the transfer learning scheme in FL [87].

d: COMMUNICATION

In centralized systems, a limited number of nodes are there and data is not being shared with client nodes, which reduces the model's efficiency to a greater extent. Whereas, FL-based healthcare system has a large number of clients. In addition, the data is shared to find local gradients securely. Due to sharing data locally, we can get a local gradient that is not dependent on other models. Thus, an average gradient is obtained. Furthermore, many clients are sharing the data, and a huge volume of data is obtained, resulting in a highly precise and accurate model. Thus, the problem of communication is solved by the FL-based healthcare model.

e: SCALABILITY

With the limited number of resources, the efficiency of the model is not guaranteed to an extent. Centralized intelligence healthcare architecture is less scalable as the number of clients is less in numbers. Furthermore, the central main server cannot handle data more than its threshold capacity. Due to this, the scalability of the model decreases. In the FL-based healthcare model, we have a large number of clients, where the data is divided among local servers. Thus, better scalability can be achieved in the FL-based healthcare model, which was earlier a great concern in a centralized intelligence healthcare system.

f: PRIVACY AND SECURITY

In the medical biosphere, the security and privacy of data is a top priority. In a centralized intelligence healthcare system, frequent threats to data breaches affect the performance of the model. In an FL-based system, maintain the privacy and security of data while managing it at the local server and use security techniques like HE, cryptographic primitives, and deferential privacy. The individual data is processed at the local server and not being shared among the clients in lieu of data security and privacy.

g: NETWORK LATENCY

It is a prime concern in centralized intelligence healthcare systems. Using the FL-based healthcare model, this problem can be resolved as clients process their data at the local servers, which does not require fetching data from the remote centralized server.

VI. CONCEPTUAL OVERVIEW OF FL-HI AGGREGATION

In this subsection, we present the FL classification models, namely the horizontal FL (HFL), vertical FL (VFL), and transfer learning-based FL approach (TL-FL). FIGURE 7 presents the three scenarios. The classification models are based on the amount of data that is partitioned among different nodes and the assisted networking structure.

1) *Horizontal FL-* In HFL, we consider that the global server models are shared with all local clients and the clients then cooperatively train the server model with their local data instances. We consider that the local client nodes work with the same feature sets during the training but in different spaces. The benefit of using the same feature sets allows the clients to have a common ML and DL model, and thus the proposed updates converge faster as all nodes cooperatively propose the model values. In HFL, an adversary node, if authorized, might propose a common adversarial model which other client nodes would select and train the model. In such cases, the detection of anomalous nodes becomes crucial. Another approach is to secure the results with the addition of noise that pertubates the data.
2) *Vertical FL-* In VFL, we consider that all local clients are networked together, and opposite to HFL, now the clients have the same space but different feature sets. Thus, in VFL, the local clients can use their ML models, but the training has to be aligned at the same shared space set. As the space set is shared, the data samples

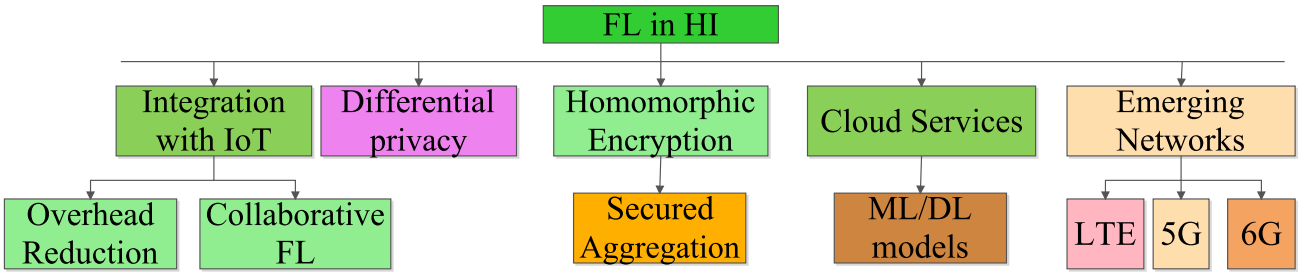


FIGURE 8. A solution taxonomy of FL-HI for different applicative verticals.

might overlap, and thus, different AI models can be integrated to train a single sample. A potential example is heterogeneous loan providers that can collaboratively process a single user loan, but can use different prediction models to assess the risk, client income stats, and other portfolios and train the model to propose interest rates. In such cases, using VFL, a customer gets selective interest rates from different loan providers and selects the most optimal choice.

3) *Transfer Learning-based FL-* In TL-FL, the sample space of VFL is extended to include details of more local clients with different feature sets and the same sample space. However, the feature is extended as TL extends features from different samples to local space, and then the model can be used to aggregately train the global model. In FL-TL, random and local masking-based encryption schemes are proposed, where the TL focuses on minimization of losses of the FL space. TL-FL based approach is suitable for EMRs as they are collected from different heterogeneous sources (hospitals, labs, patients, doctors, insurance agencies, and other stakeholders), and via TL, common features sets are extracted to be stored in the same shared space. This improves the learning rate significantly over the basic VFL approach.

VII. FL APPLICATION SCENARIOS: A SOLUTION TAXONOMY

This section discusses various scenarios where FL is applicable that can be helpful in healthcare applications. FIGURE 8 presents a solution taxonomy of FL applicative scenarios. The details are presented as follows.

A. FL WITH IoT NETWORKS

In this subsection, we discuss the integration of FL with the IoT for secure and efficient data processing. Various authors have given their frameworks for FL and IoT integration, such as Yaun *et al.* [88] proposed a framework in which there is a reduction in computational load and communication overhead from IoT devices. The proposed framework uses FL to train deep neural networks, which reduces network traffic by 99.8% compared to FedAVG and 90% in comparison with SplitNN. Then, the authors sparsified the gradients to reduce the network traffic. The limitation of the proposed framework is that it cannot be used for multiple IoT devices at a time, which can become an issue in healthcare systems, where more IoT devices need to be connected for greater accuracy. Further, the accuracy loss is also not mentioned based on theoretical analysis [89]. Then, Qayyum *et al.* [90] proposed a clustered FL (CFL)-based collaborative learning framework that processes the visual data obtained at the edge with a multi-modal ML model. The proposed model is capable of the diagnosis of COVID-19 with X-ray and Ultrasound imagery. The divergence in the training data improves the performance of the ML model. Improvement in F1-Score is achieved to 11% and 16% than the conventional FL. The proposed framework does not provide any personalized approaches for healthcare.

Later, Wang *et al.* [91] proposed an FL-based framework that integrates deep reinforcement learning with FL in mobile edge systems. Comparative analysis of centralized, distributed, and federated systems are given in detail. The artificial intelligence tasks are scheduled over the edge nodes and mobile devices collaboratively. The framework is not for heterogeneous scenarios and it does not give the best performance with FL. Then, Lim *et al.* [92] mentioned a comprehensive survey on FL, its opportunities, challenges, and solutions. They have also discussed the applications of FL for mobile edge network optimization. Detailed discussions on reduction in communication cost and resource allocation with FL has been discussed. Specific issues based on the individual model with FL are not mentioned. TABLE 7 shows the relative comparison of various state-of-the-art technologies for the integration of FL with IoT.

B. INTEGRATION OF FL WITH DP

This subsection focuses on achieving data privacy by integrating FL with DP techniques. The authors in [93] proposed an FL-based framework that learns from distributive data residing at different sites. It offers two levels of privacy protection with two real-world scenarios. However, differential FL provides privacy but can reduce the predictive capacity of the model. The model's drawback is that it cannot be applied to real-world applications. Then, Li *et al.* [94] proposed a model for brain tumour segmentation on the Brats dataset using DP techniques. Detailed comparative analysis on segmentation performance with and without FL has been highlighted.

TABLE 6. Relative comparison for the state-of-art technologies for integration of FL with DP.

Author	Year	Purpose	Key points	Limitations
Choudhury *et al.* [93]	2020	Adopted DP to prevent attacks and to enhance privacy sharing of data to centralized server during training process is not done	Applied their proposed framework on real world scenarios of 1 million patients.	Loss of model performance
Li *et al.* [94]	2019	Presented a framework enhancing privacy with DP in brain tumor segmentation	Higher accuracy is achieved	Usage of differentially private SGD is not mentioned
Xu *et al.* [16]	2020	Presented a survey on various FL system challenges,statistical challenges and privacy policies	Comparative analysis of recent work in FL	Improvement in precision of the model is not mentioned
Chamikaraa *et al.* [95]	2021	Proposed an algorithm to control the global perturbation parameter generation, whereas local data perturbation can be conducted by the distributed entities	Higher privacy and accuracy is achieved	Higher efficiency with vertical FL is not mentioned
Kanani *et al.* [96]	2021	Proposed an FL based solution for adverse event detection for mass scale vaccination programs.	Loss of privacy is prevented by adding personalization to additional security layers	All approaches of personalization are not mentioned

They also mentioned the comparison of various training and averaging algorithms. Further, to guarantee data privacy, the sparse vector technique has been used by the authors. Later, the authors of [16] provided a detailed survey on FL in healthcare on statistical challenges, security, and privacy. Detailed analysis on recent works of FL in HI and applications of FL in other fields were also mentioned. Many general issues like improvement in accuracy, increment in the number of devices having a high quality of data, and personalization of devices were not mentioned.

Then, Chamikara *et al.* [95] proposed an algorithm that can control the global perturbation parameter generation, whereas local data perturbation can be conducted by the distributed entities and enhance the privacy of the system. Impact on communication delay and number of distributed nodes is analyzed in detail. Comparative analysis of the proposed algorithm with the existing schemes based on the classification accuracy, time complexity, and attack resistance is also mentioned. Further, Kanani *et al.* [96] proposed a real-world application to detect adverse events in mass-scale vaccination programs. Detailed analysis of Federated fine-tuning algorithm with better performance and DP is mentioned. Further, the advantages of FL like robustness towards noise, personalized model, and FL for manufacturing were also given. The main focus of their proposed framework is to investigate the accuracy loss. TABLE 6 shows the relative comparison of various state-of-the-art technologies for the integration of FL with DP.

C. FEDERATED LEARNING WITH HE

This subsection highlights the role of HE in FL. Various authors have initiated their research in this particular field. For example, Chen *et al.* [97] proposed a framework with wearables that provides personalization and enhances privacy during aggregation. Their framework produces high accuracy compared to the traditional approach by 5.3%. They have used both deep learning and HE to avoid data leakage. The usage of the proposed framework in a real-world scenario is not mentioned in detail. Then, Malekzadeh *et al.* [98] proposed a system using deep neural networks on datasets using FL with differentially private stochastic gradient. A better trade-off between differential and other approaches is obtained by using secured aggregation. Accuracy is retained in this framework while retaining privacy. Further, accuracy can be increased by involving hospitals in one more iteration by 3%. The limitation of the framework is that it cannot be applied at a record level. Then, Ma *et al.* [99] proposed a framework using multi-key HE to improve the data privacy. However, this model is not applied in real-world scenarios. Comparative analysis of framework with paillier is mentioned in detail. The risk of privacy occurs as updates of individual devices are taken into consideration. The limitation of the framework is it cannot be applied to remotely participating devices. TABLE 8 shows the relative comparison of various state-of-the-art technologies for the integration of FL with HE.

D. FL WITH CLOUD INFRASTRUCTURE

In this section, we discuss the integration of FL with the cloud infrastructure. To support this, Rajendran *et al.* [101] aimed at determining whether the performance of ML models would be increased with FL while preserving the security. Performance of artificial neural networks increased by 3%. Cyclic and incremental FL models in real-world scenarios were also mentioned along with experimental results. Few features are only taken into consideration for testing and complex ML models were not taken into consideration. Also, the security and privacy analysis of data has not been analyzed in detail.

TABLE 7. Relative comparison for the state-of-art technologies for the integration of FL with IoT.

Author	Year	Purpose	Key points	Limitations
Yuan *et al.* [88]	2020	Reduced computational load load on IoT devices ,communication overhead ,and loss of accuracy by detecting arrhythmia detection	Decomposition of neural networks for computations and specifying gradients and activations	Not applicable to multi-sensor healthcare IoT.
Qayyum *et al.* [90]	2021	Clustered federated learning (CFL) framework for COVID-19 diagnosis	Better performance is observed in CFL than traditional FL	Not given how CFL performs in terms of number of samples per clients
Wang *et al.* [91]	2019	Proposed FL framework with mobile-edge computing for optimal performance	Experimental setup for computing offloading performance with FL	Distribution of computational load on edge devices is not mentioned
Lim *et al.* [92]	2020	Preserving privacy in mobile edge optimization	Detailed analysis on communication cost, data privacy and cost	Heterogeneity in FL is not mentioned

TABLE 8. Relative comparison for the state-of-art technologies for integration of FL with HE.

Author	Year	Purpose	Key points	Limitations
Chen *et al.* [97]	2021	Proposed a framework named FedHealth for wearable edges	Higher accuracy is obtained.	Application for a specific disease is not mentioned
Malekzadeh *et al.* [98]	2021	Proposed a framework Dopamine which can establishes higher trade-off between DP (DP) guarantee and DNN's accuracy	Higher accuracy and privacy preservation is there	Privacy analysis is not discussed
Ma *et al.* [99]	2021	Proposed a model which can detect elderly fall detection with FL with HE	Computational cost is reduced compared to Paillier-based federated learning	Not applicable in real-time scenarios
Li *et al.* [100]	2021	Presented a survey on various machine learning models with privacy preserving approaches	Comparative study on existing publishing studies is done in detail	Experimental framework is not mentioned
Xu *et al.* [76]	2019	Proposed an FL-based framework employing SMC protocol based on functional encryption for privacy preservation	Higher reduction in training time and data reduction volume is obtained	Use cases scenarios are not mentioned in detail

Then, Mansour *et al.* [102] proposed three methods such as user clustering, data interpolation, and model interpolation. The performance is demonstrated with the efficiency of the methods. Theoretical proofs based on personalized learning were also discussed. Further, these approaches are used on the EMNIST dataset. FL integration with cloud infrastructure was discussed briefly. Later, Mammen *et al.* [24] surveys the opportunities and challenges in FL. Further, data poisoning, backdoor, model poisoning attacks, and recent developments with FL are mentioned in detail. Also presented an analysis of the integration of FL with blockchain technology. Theor paper also lacks the security and privacy preservation scheme for data.

Then, Rieke *et al.* [14] provided a detailed survey on FL in the field of healthcare and discussed its implementation challenges. A detailed analysis of its impact on various stakeholders like healthcare providers, patients, manufactures, hospitals, and practitioners has been discussed. Further, the privacy and security bottlenecks were not addressed. Later, the authors of [103] have given the experimental findings on the improvement of the model's performance. They trained their system with CNN, LSTM, and MLP models and achieved the computing accuracy of 95% in 6 rounds of federated averaging. The framework proposes experimental results on IID and non-IID samples in a parametric study. Challenges and future opportunities of FL are mentioned. TABLE 9 shows the relative comparison of various state-of-the-art technologies for the integration of FL with cloud infrastructure.

E. FL WITH EMERGING COMMUNICATION NETWORKS

Researchers across the globe have focused on the integration of 6G with FL. Some are as, Liu *et al.* [104] presented a detailed survey on FL with 6G communication along with

TABLE 9. Relative comparison for the state-of-art technologies for integration of FL with cloud infrastructure.

Author	Year	Purpose	Key points	Limitations
Rajendran *et al.* [101]	2020	Framework that achieved higher statistical performance is proposed	Higher precision is achieved.	Simplistic model with few features for testing is not taken into consideration.
Mansour *et al.* [102]	2020	Proposed a FL based approach for personalization	Proof based algorithms are discussed	FL integration with cloud is not mentioned in detail
Mammen [24]	2021	Reviews challenges and opportunities in FL	Detailed explanation for FL with blockchain is mentioned	Experimental results are not mentioned
Rieke *et al.* [14]	2020	Given survey on future of digital health using FL is mentioned	Use-cases in FL in HI is mentioned in detail	Privacy-preserving techniques are not given in detail
Das [103]	2019	Presented experimental results on IID and non-IID samples in a parametric study	Increment in privacy and security	Use-cases are not explained with experimental results

TABLE 10. A relative comparison of state-of-art technologies for the integration of FL with 6G networks.

Author	Year	Purpose	Key points	Limitations
Liu *et al.* [104]	2021	Presented a survey on 6G communications with FL	FL learning methods for 6G and various applications of 6G with Flare mentioned in detail.	Experimental results are not mentioned
Nguyen *et al.* [87]	2021	A comprehensive survey on FL with IoT	Detailed analysis on 6G with FL is mentioned	Algorithms are not given in detail
Kaur *et al.* [105]	2020	Detailed review of FL with 6G and its integration with ML	A case study on smart biometric application at infrastructural and application level is discussed in detail	High accuracy is not obtained
Lu *et al.* [106]	2020	Proposed a model based on federated learning and blockchain to improve the efficiency .	Reliability ,security of the system, and data privacy is enhanced	Application of framework in healthcare sector is not given

its use cases. Comparative analyses of other studies on FL and use-case scenarios in the medical field were also not discussed. Then, Nguyen *et al.* [87] presented a comprehensive survey on FL with IoT. they presented a comparative study of existing surveys and the taxonomy of FL-IoT services were discussed. Also discussed the challenges of FL-IoT, applications of FL-IoT, and FL in smart healthcare with blockchain technology [107], [108]. Not discussed the integration security and privacy issues. Later, Kaur *et al.* [87] reviewed the communication system with ML techniques. Detailed analysis on the next-generation wireless communication systems and their applications. The performance of the proposed system with ML techniques such as artificial neural networks and reinforcement learning are discussed in detail. Further, Lu *et al.* [106] proposed a model based on FL and blockchain to improve network reliability, security, and privacy. The proposed framework uses blockchain and FL for digital twins networks. Experimental proofs are given in detail. Further, a framework for edge association is proposed in the paper. They formulated an optimization problem to improve their system's accuracy. Real-world datasets were used and improvements in latency and learning convergence were obtained. The study does not mention the application of FL in healthcare. Xu *et al.* [109] proposed a lightweight signcryption mechanism in the edge computing environment, that guarantees authentication and confidentiality of stored records. A certificateless signcryption using blockchain is proposed, such that the user identity is not compromised. The scheme eliminates illicit attacks from a malicious adversary with minimum computational overheads. TABLE 10 shows the relative comparison of various state-of-the-art technologies for the integration of FL with 6G networks [110].

F. ADVANCE FL MECHANISMS

In this subsection, we discuss some advanced FL mechanisms to support decentralized HI ecosystems. Modern healthcare setups are sensor-driven and collect patient readings from local WBAN environments. Such local setups are resource-constrained (in terms of storage and processing power), and thus emerging FL designs like resource-aware

FL, incentive-aware FL, and blockchain-based FL are proposed by researchers. The details are presented as follows.

1) RESOURCE-AWARE FL

Resource-aware FL is mainly designed to intelligently orchestrate the management of critical resource setups. In these setups, scheduling mechanisms play a critical part to transfer local model information collected from heterogeneous WBAN setups to an aggregation server (data collected from wearables to an edge/fog interface). Once data is sent, a joint optimization is proposed between local models to initiate the resource allocation for model training. Nguyen *et al.* [111] proposed an intelligent resource allocation framework for FL training for multiple local node setups, using a joint optimization of computation and communication costs of mobile devices and network parameters. The formulated problem minimizes the training overhead and lowers the energy consumption of local WBAN setups. Authors in [112] proposed a per-device FL service for multiple nodes and designed an AI algorithm to optimize the network bandwidth allocation to clients. A game-theoretic approach is presented to fine-tune the learning and fairness among the local clients.

2) INCENTIVE-AWARE FL

In limited resource setups, local nodes are memory and power-constrained, and hence are not interested to follow the conventional FL training mechanism, where local updates are collected securely, and send over to the global model. Thus, in incentive-aware FL, the main goal is to motivate the local setups to participate in the model training process. To exploit the same, incentive mechanisms are discussed for local nodes based on the amount of contributed data, the node reputation, and the number of allocated resources. In some cases, the global model gives more weightage to the quality of shared weights (which minimizes the iterations), over quantity (when a single node shares ample weights or gradient parameters). In some approaches, to optimize the incentive payoff, game-theoretic mechanisms are proposed. For example, Khan *et al.* [113] proposed a Stackelberg game model between local users, where each user is presented with an offer from the aggregator node. Based on the offer, the local nodes decide to participate in the FL training process. Once nodes are finalized, the aggregation server reveals the incentive strategy to the nodes, and nodes take corresponding actions to enhance their payoff function. On the other hand, the server also maximizes its utility function while proposing incentives. The process is repeated until the global model achieves a target accuracy (pre-defined), and the game assures that the number of training rounds is minimized, as both players (local nodes and aggregation server) want to maximize its payoff condition.

However, game formation schemes suffer from an inherent limitation of assuming that the aggregation server has full knowledge of local nodes, and their previous contributions to fix the incentives. To overcome this situation, authors in [114] proposed a deep reinforcement learning (DRL) mechanism that helps the aggregation server to formulate the awarding policy based on previous historic policies. This method is useful for cross-silo environments, where heterogeneous hospitals collaborate to build the global model and send their gradients to a third-party aggregation server. Based on past events, the DRL mechanism suggests the aggregation server fix an optimal incentive for the participants, which is quite promising and thus motivates them to participate in the learning process.

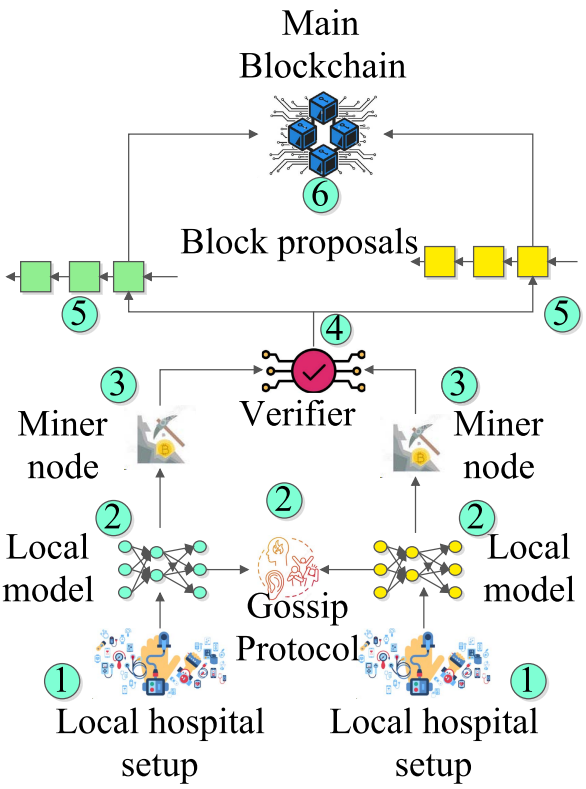


FIGURE 9. Gossip mechanism for blockchain-based FL.

3) BLOCKCHAIN-BASED FL

In FL learning process, security attacks like model poisoning, trapdoors, and free-grants are manipulated by malicious adversaries [115]. In FL model training, data is collected from unreliable sources, and thus a malicious adversary can inject unreliable data streams to local clients (through privilege escalation), which manipulates the global training process. Thus, device reputation is a critical aspect of the FL training process. To prevent these attacks, recent research has suggested the use of decentralized FL models in HI to secure the local parameters and assure trust in communicated parameters. Decentralized FL mechanisms typically employ mechanisms like gossip, diffusion, and consensus to design trusted clients. However, due to the inherent complexity of the collected data, reliable user selection is difficult. Moreover, in decentralized FL, all the local nodes form a peer-to-peer (P2P) learning network, and there is no need for the transfer of model parameters to the global server. Thus, in such

P2P FL networks, blockchain is a viable choice to induce transparency and immutability in sharing of trusted gradients. FIGURE 9 presents a blockchain-based FL setup that exploits a gossip mechanism for secured sharing of model weights. Gossip protocol, sometimes referred to as epidemic transmission protocol, ensures that updates are transmitted to all nodes in the network (same as flooding mechanism) [116].

The figure illustrates a single round of FL communication via blockchain. Initially, each local WBAN setup (device) trains its local model on the collected patient data (step 1) and forms the local model (step 2). Once all local nodes complete their training process, the training parameters are communicated to all other nodes using the gossip strategy (step 2). This step is often termed the model broadcasting step. The broadcasted data, along with the node's local data, is used to update the parametric weights, and improve the model learning. However, owing to the risk of false injection of false weights by an intruder, all the device model data is verified by elected miners in the blockchain network (step 3). The miner nodes are elected through a suitable low-powered consensus strategy that matches the end-application computational requirements [117]. The miners process the local broadcast as transactions and append them to local blocks for verification by verifier nodes (step 4). Once the unconfirmed transactions (in mempool) are verified, they are added to the local client blockchains. As transactions are verified and timestamped by the verifier and miner node, it places an inherent trust on another local node to use the model weights in the learning computations. Once all nodes synchronize their learning weights (sent through block proposals), the updated weights and communication parameters are recorded on the global blockchain (step 5). This greatly improves the upper bound of the model loss function, which depends on the data distribution time, and the number of participants. So *et al.* [118] investigated the aggregation overhead of the gossip protocol communication, and established that the overhead significantly reduces from $O(N^2)$ in centralized FL schemes, to $O(NlogN)$ in distributed schemes.

VIII. THE PROPOSED FL-BASED HI ARCHITECTURE

In this section, we present the key technicalities of a decentralized FL based HI ecosystem. The centralized intelligence models are challenged in terms of network and security considerations. To address the limitations, a FL based learning model is applicable. FIGURE 10 presents the key architectural details that comprises of three layers, namely, the data collection layer, the gradient layer, and the analytics layer. Similar to centralized model, we consider that data collection is one from different heterogeneous sources and healthcare stakeholders like hospitals, clinics, body-area-networks assisted wearables, and EMRs. The collected data from N stakeholders are arranged into static and real-time collection, and are then stored into local data store (LDS). In general, we consider that n stakeholders collect their LDS, denoted as $\{LDS_1, LDS_2, \ldots, LDS_n\}$. Now based on a global server model, denoted as GS_m, it is downloaded and

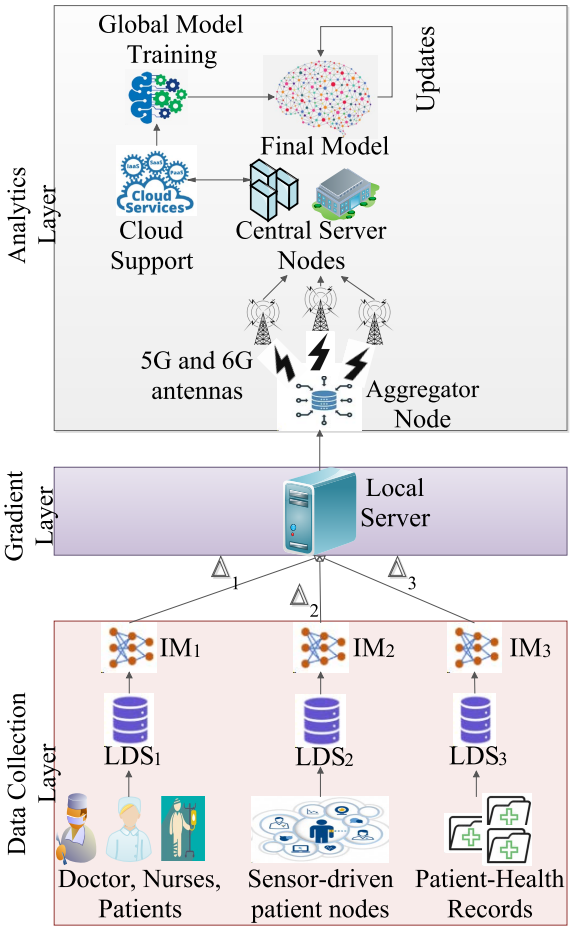


FIGURE 10. FL based healthcare architecture.

shared among all LDS. The layer-wise details are presented as follows.

A. DATA COLLECTION LAYER

At the data collection layer, we consider that there are local sources, who have local data stored in LDS. The stored data $\{D_1, D_2, \ldots, D_n\}$ is collected, and is applied on the shared GS_m through any FL classification model. Once the local training is complete, the data is collected, and training gradients are computed, denoted as $\{\Delta_1, \Delta_2, \ldots, \Delta_n\}$. The overall sample size is denoted as N, which is obtained as sum-total of overall sample size $\sum_{i=1}^{n} LDS_i$. For computation of gradient values, we consider a local instance module (IM), for all the n nodes, denoted as $\{IM_1, IM_2, \ldots, IM_n\}$. IM nodes computes the local learning model parameter ω, and computes the training cost as $\min_{\omega \in M^n} F(\omega) = \sum_{i=1}^{N} \frac{\gamma_n}{n} F_n(\omega)$, where $F_n(\omega)$ denotes the loss functions. Once loss functions are estimated at IM, we collect the gradient to be sent to gradient layer.

B. GRADIENT LAYER

At the gradient layer, the collected data is statistically averaged using the FedAvg model that minimizes the skewness of the learning models. At this point, we compute the optimal

condition G_{opt} to generate results to be communicated to the aggregator node via the 5G or 6G assembled antennas, with edge-computing support [40]. At gradient layer, we can add noise distribution samples for data perturbation. We consider n noise samples $\{NS_1, NS_2, \ldots, NS_n\}$ aare added which is generated from noise distribution $N(0, \sigma^2)$. Through the noise-addition, the user data privacy is maintained. In such cases, the most preferred model is the $\epsilon - DP$, where ϵ denotes the upper-bound of added noise. The noise is added from neighboring data-samples, defined under the bound e^{ϵ} [119]. Once noise is added, the gradients are sent to the antenna uplink channel with defined bandwidth B_U. We normally include a 5G or 6G-uRLLC service to minimize the end-point communication latency to the HI analytics layer [120].

C. ANALYTICS LAYER

At the analytics layer, the local gradients are presented to GS_m, that might offload task-sets to nearby cloud-services. The statistical difference $Diff = |U_{GM} - U_{LM}|$ is computed, where U_{GM} denotes the statistical weight average of global model, and U_{LM} denotes the weighted average for the local updates, and model edge weights are updated to minimize the bias. Via resilient uRLLC services, the communication latency is minimized, and updates to the global model are communicated faster, so the global model converges to optimal state with fewer iterations.

To explain the proposed FL-architecture in simple terms, at the primary level, healthcare consumers serve as clients for the model, which are hospitals, primary health centers, and outdoor patients. They are the main source of the data used to train the local and global models. The data is to be collected from mobile phones, smartwatches, smart bands, etc. The collected data are body temperature, respiratory rate, sleep tracking, heart rate, and blood pressure. The data processing is done at the edge device such as smartwatches and smart bands. In an FL-based healthcare model, individual edge devices train a local model using ML algorithms. Once the individual model is trained, its individual gradient is computed. These individual models are then passed to a local server, where it calculates the local gradient. If any update occurs, the slope of the local gradient is changed and accordingly, individual gradients would have also changed their slopes.

In FL, individual models at the same location are combined and the combined model is passed to the local server, where the local gradient is calculated. If the model has any kind of error in the data, then it is discarded. The large number of local gradients computed are then aggregated for the global server to obtain a global gradient. This aggregation of local gradients is possible with the help of algorithms like FedAVG, FedPROX, and FedCS. These algorithms have been used to achieve high security and privacy of the model with HE. The local gradients from the local servers are aggregated and passed to the centralized cloud server. The communication process evident here is the 6G technology. With 6G technology, ultra-reliable low latency communication can be achieved, which is quite necessary for a biomedical space.

The aggregated gradient is passed to the global server, where it is analyzed at the server. The global server forms a global gradient out of aggregated gradients. The global server then uses the aggregated gradient to form a global model. The global model is formed using ML algorithms to provide high accuracy and efficiency. This model can be adopted by various primary health centers and hospitals for disease prediction. The updates which are sent from various clients are also processed here. Thus, an updated model is obtained by each client from the global server. This proposed scheme can help in better privacy, security, accuracy, and efficiency in the model, with data not being shared among other clients. In the next section, we present the solution taxonomy of the FL in HI ecosystems

IX. *FL-EHR*: THE PROPOSED CASE STUDY

In this section, a case-study named as *FL-EHR* is proposed. The details are presented as follows. The proposed system (FIGURE 11) uses FL in HI for better privacy and security of the data of users. A centralized intelligence system deals with many limitations like data updation delay, lack of high precision and accuracy, lack of privacy and security, and handling limited resources. A centralized intelligence system can be implemented and set up easily. However, due to significant bottlenecks in the centralized system, we have proposed a new FL-based HI model that offers high privacy and security of the data to the users. In FL various resources are aggregated, which improves the precision and accuracy of the system. The proposed system achieves high precision and accuracy of data using methods like DP, HE, and cryptographic primitives. Privacy to a large extent is obtained with FL as the individual models are aggregated into one local model and the models are not shared among each other. With the usage of 6G technology, communications can be easily done in FL. Thus, with the usage of FL in HI, a secure and efficient model can be created in a biomedical space.

In the proposed system, we have n numbers of clients having their models encrypted with them. These models are not shared with other clients. The local gradients are calculated in the FL training scheme. Then all the local gradients are aggregated at the aggregation level with methods like FedAvg, FedProx, and FedSGD. The communication network used here is 6G, through which optimal and faster communication can be achieved. In 6G networks, effective optical wireless communication channels can be exploited for close and directed communication with the healthcare setups [121]. Interesting use-cases of 6G-assisted edge FL are proposed in applicative areas like vehicular networks, massive IoT aggregation networks, and other feilds [122], [123]. The aggregated gradient is then used in the cloud server for prediction purposes with ML models like artificial neural networks, CNN, etc. The global gradient is calculated out of it.

The prediction model is then passed through the aggregation layer with HE to FL training scheme to various clients.

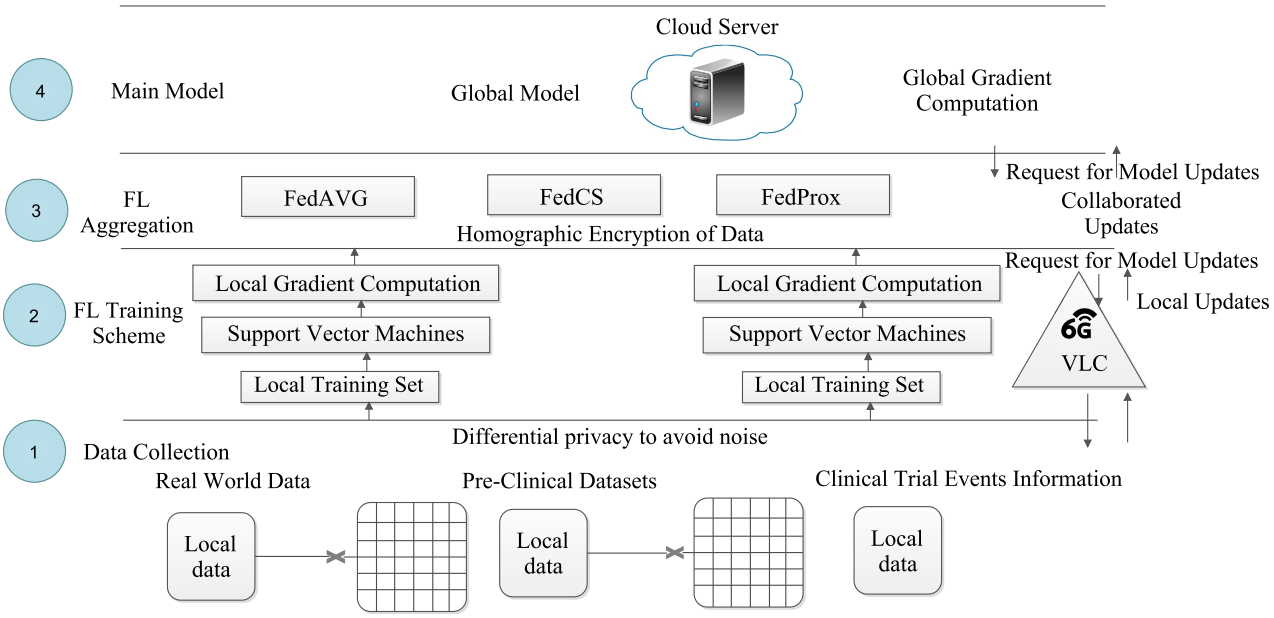


FIGURE 11. Proposed system architecture of FL based healthcare system.

The n clients send collaboratively request to the cloud for the model update. The request for model updates is satisfied by the centralized cloud server. The entire proposed model is bifurcated into diverse layers such as data collection, FL training, and data aggregation. The detailed description of each layer is discussed as follows.

A. DATA COLLECTION LAYER

The data collection layer consists of heterogeneous real-world data from smart wearables of the patients, pre-clinical tests records, clinical trial event information, and data from the hospitals up to n resources. Each model is initially trained at the individual level and their gradient is calculated. The individual data of n clients is not shared among each other to ensure the privacy and security of patients personal data. As medical data is quite sensitive, data privacy is considered an important aspect. DP is used for data privacy in this layer, which does not allow information sharing among other individuals, which ensures privacy of individual data. DP ensures privacy by adding noise to ensure the prevention of personal data. However, with the usage of DP, the enhancement in accuracy can be noted.

B. FL TRAINING LAYER

In this layer, the individual trained models having individual gradients are analyzed and trained locally again. Initially, the individual gradient computed of each model is aggregated with other individual gradients of the same localization. The localized models obtained are then passed to various ML algorithms like SVM and gradients are computed. The gradients of local models are known as localized gradients. Many localized models are obtained in this layer. Moreover, secure and high-speed communication is obtained by combining visible light communication techniques of optical wireless communication.

C. DATA AGGREGATION LAYER

The localized gradients are computed and aggregated in this layer using methods FedAVG, FedSGD and FedProx. The description of each method is as follows.

1) FedAVG

FedAVG is a method in which the model is updated locally and then aggregated with the global model at the cloud server. According to Muthukari *et al.*, initially, the local models are aggregated, then the global model is computed and the pass it to local clients. FedAVG achieves high accuracy among all the aggregation algorithms [15].

2) FedCS

FedCS is a protocol through which the central server manages the various clients and decides the clients who participate in the training task [16]. The models participate in the local model through FedCS and then models are aggregated to pass to the central cloud server.

3) FedProx

FedProx is optimized method than Fedavg for efficient performance. The variation in computational power and factors in devices participating in the FL training round is considered and a solution to deal with non-uniformities in local data is given by FedProx [15]. In this algorithm, different iterations at different heterogeneous devices are performed [4]. The HE ensures the data privacy and security by preventing its

leakage. HE achieves high level of privacy and security in addition to the accuracy of the model.

Overall, the proposed model enhances the performance of the healthcare system along with data security and privacy. This improves the prediction accuracy, which in turn enhances the quality of life of the patient.

The proposed *FL-EHR* case study integrates differential privacy which assures privacy in learning data, which is quite beneficial in real-world setups. Further, it removes the bias in data during the training process, that improves the model accuracy. Thus, compared to current healthcare setups, the proposed study improves the model validation, that forms the basis of effective analytics setups in heterogeneous medical ecosystems.

X. OPEN ISSUES AND LESSONS LEARNED

A. OPEN ISSUES AND CHALLENGES

Despite many advantages of FL, there are some challenges related to FL while working in the medical field. TABLE 11 describes the summary of open issues and challenges.

1) COMMUNICATION RELATED ISSUES IN FL-HI

As discussed in the aforementioned sections, there is a continuous exchange of model parameters (gradient weights) between the global and local client models for training purposes. It consumes substantial communication resources when the number of smart healthcare devices increases, i.e., more devices leads to more client models that require an enormous exchange of model parameters. The solution to this issue can be data compression, i.e., Compressing a model minimizes the number of communication cycles. Various model compression frameworks like sparsification [124] and quantization impressively minimizes the message size at each communication cycle [125]. These compression techniques manage to achieve high accuracy with low communication cost. Apart from sparsification and quantization, the neural network can also be used to reduce communication overhead (interference) in FL [126].

The authors in [127] investigated the redundancy in gradient that exchanges between the global and client models. The reduction/compression of gradient can significantly improves the communication bandwidth. Then, the authors in [128] proposed a Count Sketch-based gradient compression scheme for local clients before passing it to the global model. The communication performance of FL can be enhanced by compressing the uplink and downlink communications between the global and local client models. The authors in [129] proposed a novel communication protocol to compress the uplink and downlink communications while maintaining the robustness and reliability of the system. They have used optimal Golomb encoding technique to compress the uplink data and speeding up the training process for the global model. Above mentioned techniques significantly improves the performance quality of service in HI-based frameworks by minimizing the overall communication overhead. The efficiency of the communication system in FL-HI can further be improved by the incorporation of 6G communication network that offers ultra-low latency and extremely high reliability.

2) HANDLING DATA HETEROGENEITY IN FL-HI

Heterogeneous data is found in HI due to various diseases, treatments, and advances in the healthcare sector. The heterogeneous data obtained from various devices are used to train the model in FL. Messy data can create issues in the precision of the model. The model obtained should be of the highest accuracy and precision. It plays a vital role in monitoring and analyzing any patient as doctors rely on models for remote monitoring. The data is aggregated from various sources in FL, which can lead to biasing in optimization [4]. There exist many techniques in machine learning that models heterogeneity using meta and multitask learning methods. These methods and techniques can be applicable or extended in modeling the FL-based HI architecture. The authors in [126] and [131] presented a data heterogeneity solution with multitask learning to achieve quality results and predictions. Despite these solutions, handling heterogeneous data and heterogeneous modeling in HI with the increase in client models, some prime concerns are of need to be take care such as robustness, fairness, and scalability of heterogeneous HI data [126].

3) HANDLING AGGREGATE LOSS IN FL-HI

In FL-HI, the global model receives gradient weights from multiple client models to train itself. Some devices trained their model based on the devices having ample amount of data or some with their local updated. This creates variance in loss values at the global model and affect the overall system's performance. A solution to this issue is to assign high weight to the client model having high loss value. Another solution to handle the aggregate loss is to perform optimization on loss values at the local client model. These approaches minimizes the variance in model's performance.

4) PRIVACY AND SECURITY CONCERNS IN FL-HI

Though FL provides privacy and security, some attacks like data poisoning and Model-poisoning can affect users' security. Security in HI plays an essential role in HI as the data contains sensitive information. There are chances of data leakage while transferring the data. Moreover, many attacks like data poisoning degrade the model. In data poisoning, some amount of malicious data by a person and the model is tampered with to generate false parameters [4]. Through data poisoning, data manipulation takes place, resulting in bad quality of precision. Another type of attack is model poisoning, where the updated model is tampered resulting in privacy and security issues [4]. In model poisoning, the gradient is manipulated, which can cause a difference inaccuracy. Thus due to these attacks, the results are manipulated and accurate precision is not obtained. According to Jiang *et al.* [131] to Generative Adversarial networks can exact data from the server-side, which manipulates the privacy of the user's data.

TABLE 11. Summary of research challenges.

Challenges	Description	Possible Solution
Communication related issues in FL-HI	In FL-HI, the continuous exchange of weights between the global and the local models to accurately manage the healthcare data. This causes high communication overheads while uplinking and downlinking gradient/weights from the global and client models.	Gradient compression, sparsification, and quantization are the possible solutions that minimizes the communication cycle by compressing the gradient size. This also speedup the data processing.
Handling data heterogeneity in FL-HI	FL-HI is a distributed model, where the global model collects heterogeneous data from local client models to train itself. Heterogeneity in data may lead to an accuracy loss.	Meta and multitask learning techniques can be used to manage the data heterogeneity.
Handling aggregate loss in FL-HI	Aggregate weights received at the global FL-HI model from the client local models for training purpose. Due to biasness at the local model creates high variance in aggregate loss that affects the performance of the global model	Assign high weight to the local model having high loss value or using optimization techniques to achieve low aggregate loss.
Privacy and security concerns in FL-HI	The security and privacy of the gradient weights exchanged between the local and global models need to be secure from malicious or non-trusted users. Such activities can degrade the model's performance and minimize the prediction precision.	Traditional cryptography (homographic encryption) techniques and blockchain is a possible solution to share gradient weights between local and global models. AI-based solutions also exists that manages the security and privacy of the gradient weights.
Communication cost in FL-HI	In FL-HI, the many local client models are associated in aggregating global weights that persists high communication overheads. More the number of local client models, more will be the global model's accuracy for predicting disease.	Data compression and multi-stage optimization are one of the possible solutions in reducing the communication overheads.
Quality of HI data	Data quality must be guaranteed in HI for the accurate diagnosis of a disease. Data collected from different client models may not assure the quality in terms of missing values, high variance, etc.	There is a need for highly efficient data preprocessing methodologies are required at the local as well as global models.
FL-HI system efficiency	The efficiency of FL-HI degrades with the increase in the number of local client modes. It is because of the massive and highly noisy data. Low efficient HI system leads to a wrong decision on diagnosis selection.	Efficient preprocessing and malicious client machine detection. Malicious machines can broadcast dummy packets towards the global model with an intention degrade the system's performance.

a: HE-BASED FL-HI

While exchanging the gradient/learning weights/parameters between the global and the local client models, an attacker can sniff or modify the weights/parameters. As in HI, the patient's information is highly sensitive and due to the security and privacy issues, FL loses the trust. So, ensuring security and privacy is a primary concern, and HE helps achieve it. In this, the weights are encrypted with the shared secret key and forwarded to the global model. Encryption/decryption offers security but increases the computation cost. So, there is a requirement of a procedure that enhances the security without increasing the overall computation cost.

b: TRUSTFUL AGGREGATION IN FL-HI

A non-trusted user can participate in sending learning model to the global server, but it intentionally sends wrong weights to mislead the global model. It also leads global model towards the wrong aggregation of global weights [132]. The authors in [133] proposed a trusted aggregation system for FL, but it possesses high communication and computation overhead. In other works, the authors have given blockchain-based solutions to meet the trust criteria and mitigate the effect of non-trusted user [134]. Although, some literature were presented in lieu of the trustful aggregation, but not focused on latency and cost related issues (these are the important concerns in HI). So, there is a need for a protocol

that can perform trustful aggregation of gradient weights or learning models.

5) COMMUNICATION COST IN FL-HI

One of the major challenge in designing the FL-based HI system described by Zhang *et al.* [17] is communication cost. They described that the FL has many devices associated with it, which raises its communication cost. Communication overheads can be solved using data compression or sending only relevant information to the user [24]. Later, the authors in [135] given a multi-stage optimization solution to minimize the communication cost in FL. The same can be applicable to FL-HI. But, the multi-stage optimization technique is quite computationally expensive, which need to be addressed in the future prospects.

6) QUALITY OF HI DATA

The authors in [16] mentioned a challenge regarding the quality of data. As the data in the FL model comes from various clients, the quality of the data is not guaranteed. In biomedical space, dirty samples can even harm the data. The data collected can be not uniform; thus, the precision of the model is hampered due to raw, unclean data. Strategies should be incorporated to collect uniform data, which provides better precision to the model. Thus, there is a requirement for an efficient preprocessing model that makes the incoming healthcare data clean and uniform. One of the possible solution is to disburse incentives to the local client HI model for providing the best learning model.

7) FL-HI SYSTEM EFFICIENCY

The performance efficiency of FL-HI is hindered to a large extent as massive data sources are there [25], which increases the algorithm complexity. Noisy datasets can often temper the convergence of the FL algorithm. Further, network communication also hinders its performance. Due to heterogeneous data, gradients can be affected, which leads to lesser precision in the model.

8) DATA PERSONALIZATION

The gradients in FL are calculated from various clients and then aggregated, which does not achieve personalization. The biomedical data is of various clients can lead to a reduction in personalized healthcare in wearables. Aggregating data from various clients results a specific problem solution is not always obtained. Further, more power is consumed when FL models are trained locally. So, power-efficient models should be designed.

B. LESSONS LEARNED

We have provided a detailed study on the applicability of FL in the healthcare sector. We come across some positive and negative lessons learned for the integration of FL in HI, which are as follows.

- In FL-HI, the accurate and efficient data processing capabilities of both global and local client models/servers is critical. Any kind of mishandling in the data preprocessing is intolerable as it leads to the wrong diagnosis, which can harm the patient's health. Here, the local models are smart healthcare devices that collect patients' sensitive healthcare information and aggregate it to the local server. Similarly, the global model collects gradient weights from various local servers for training purposes.
- It is necessary to secure the gradient weights or learning models that are exchanged between the global and local models/servers. The healthcare data of any patient is extremely important and any modification to that data may lead to adverse effects, i.e., health deterioration due to wrong data and misdiagnosis. So, the security of such data in the FL-HI model is of utmost importance. Cryptographic primitives and blockchain technology are plausible solutions to achieve high security and privacy.
- In the remote diagnosis of a patient, we can not tolerate even a millisecond of the delay that can cause health deterioration. In the case of non-critical diseases, we can bear delay, but when it comes to critical diseases, i.e., stroke, a delay can cause the patient's death. To prevent this, there is a need for a faster communication channel, i.e., 6G that possesses $< 1ms$ of delay, the data rate in 10's of Gbs, and reliability of 99.99999%.
- Another lesson learned from the survey is the importance of data reduction. If the data is too large, it's better to compress the data that reduces the number of exchange communication cycles and helps achieve better and faster results.

XI. CONCLUSION AND FUTURE WORKS

This paper presents a comprehensive survey on the adoption of FL in smart healthcare informatics. The rise in adoption of decentralization concept in smart healthcare ecosystem, raises alarming concerns about the patient's data privacy and security, which is remotely aggregated over the open communication channel. However, with strict privacy regulations, the shared data has tend to become more cryptic and anonymized for AI models. With less data, the models tend to become biased, and thus a fair accuracy judgement is not possible. Thus, centralized learning models for HI have become bulky, and these models do not support accurate decisions. Thus, the integration of FL allows decentralized local learning in HI without the requirement of data sharing that maintains the data privacy and confidentiality. This paper discusses a comprehensive survey on FL emergence as a potential solution to support HI. A tutorial approach is followed where the background and technicalities of HI and FL learning models is discussed. Next, a high-level overview of FL-assisted HI, and its communication and security requirements are discussed. We also present the different challenges in the adoption, and a comparative analysis with various past surveys is presented to give better solutions. Open issues and challenges

are mentioned so that researchers can find solutions to the questions of FL integration to HI. Finally, we discuss the lessons learned, and the concluding remarks are presented.

As part of the future scope, the authors would investigate the performance of FL systems on different healthcare datasets, with varying degrees of sensitivity and anonymity. The authors would propose a DP model for FL-HI that combines noise sources from heterogeneous nodes in the proposed framework and evaluates system performance under defined parameter sets. This would result in high secrecy of medical setups, with the advantage of accurate and verified diagnosis.

REFERENCES

[1] E. Mbunge, B. Muchemwa, S. Jiyane, and J. Batani, "Sensors and healthcare 5.0: Transformative shift in virtual care through emerging digital health technologies," *Global Health J.*, vol. 5, no. 4, pp. 169–177, Dec. 2021.

[2] *The'Big Data'Revolution in Healthcare: Accelerating Value and Innovation*. Accessed: Jul. 17, 2022. [Online]. Available: https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/the-big-data-revolution-in-us-health-care

[3] G. S. Aujla, A. Jindal, R. Chaudhary, N. Kumar, S. Vashist, N. Sharma, and M. S. Obaidat, "DLRS: Deep learning-based recommender system for smart healthcare ecosystem," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[4] M. Aledhari, R. Razzak, R. M. Parizi, and F. Saeed, "Federated learning: A survey on enabling technologies, protocols, and applications," *IEEE Access*, vol. 8, pp. 140699–140725, 2020.

[5] *Market Research Firm*, MarketsandMarkets, Pune, India, 2020.

[6] S. Alder, Ed., *2020 Healthcare Data Breach Report: 25% Increase in Breaches in 2020*, HIPPA J., Compliancy Group, South West England, U.K., Mar. 2021.

[7] *Federated Learning Solutions Market*, Markets Markets Res. Pvt Ltd., Pune, India, 2021.

[8] Y. Liu, A. Huang, Y. Luo, H. Huang, Y. Liu, Y. Chen, L. Feng, T. Chen, H. Yu, and Q. Yang, "FedVision: An online visual object detection platform powered by federated learning," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 13172–13179.

[9] A. Hard, K. Rao, R. Mathews, S. Ramaswamy, F. Beaufays, S. Augenstein, H. Eichner, C. Kiddon, and D. Ramage, "Federated learning for mobile keyboard prediction," 2018, *arXiv:1811.03604*.

[10] M. Tang and V. W. S. Wong, "An incentive mechanism for cross-silo federated learning: A public goods perspective," in *Proc. IEEE Conf. Comput. Commun.*, May 2021, pp. 1–10.

[11] D. Saraswat, A. Verma, P. Bhattacharya, S. Tanwar, G. Sharma, P. N. Bokoro, and R. Sharma, "Blockchain-based federated learning in UAVs beyond 5G networks: A solution taxonomy and future directions," *IEEE Access*, vol. 10, pp. 33154–33182, 2022.

[12] A. Verma, P. Bhattacharya, Y. Patel, K. Shah, S. Tanwar, and B. Khan, "Data localization and privacy-preserving healthcare for big data applications: Architecture and future directions," in *Emerging Technologies for Computing, Communication and Smart Cities*, P. K. Singh, M. H. Kolekar, S. Tanwar, S. T. Wierzchoń, and R. K. Bhatnagar, Eds. Singapore: Springer, 2022, pp. 233–244.

[13] A. Jindal, A. Dua, N. Kumar, A. V. Vasilakos, and J. J. P. C. Rodrigues, "An efficient fuzzy rule-based big data analytics scheme for providing healthcare-as-a-service," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[14] N. Rieke, J. Hancox, W. Li, F. Milletarì, H. R. Roth, S. Albarqouni, S. Bakas, M. N. Galtier, B. A. Landman, K. Maier-Hein, S. Ourselin, M. Sheller, R. M. Summers, A. Trask, D. Xu, M. Baust, and M. J. Cardoso, "The future of digital health with federated learning," *npj Digit. Med.*, vol. 3, no. 1, pp. 1–7, Dec. 2020.

[15] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Gener. Comput. Syst.*, vol. 115, pp. 619–640, Feb. 2021.

[16] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, "Federated learning for healthcare informatics," *J. Healthcare Informat. Res.*, vol. 5, no. 1, pp. 1–19, Mar. 2021.

[17] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, "A survey on federated learning," *Knowl.-Based Syst.*, vol. 216, Mar. 2021, Art. no. 106775.

[18] B. Pfitzner, N. Steckhan, and B. Arnrich, "Federated learning in a medical context: A systematic literature review," *ACM Trans. Internet Technol.*, vol. 21, no. 2, pp. 1–31, Jun. 2021.

[19] C.-R. Shyu, K. T. Putra, H.-C. Chen, Y.-Y. Tsai, K. S. M. T. Hossain, W. Jiang, and Z.-Y. Shae, "A systematic review of federated learning in the healthcare area: From the perspective of data properties and applications," *Appl. Sci.*, vol. 11, no. 23, p. 11191, Nov. 2021.

[20] G. Long, T. Shen, Y. Tan, L. Gerrard, A. Clarke, and J. Jiang, *Federated Learning for Privacy-Preserving Open Innovation Future on Digital Health*. Cham, Switzerland: Springer, 2022, pp. 113–133.

[21] M. Weiss, M. Luck, R. Girgis, C. Pal, and J. P. Cohen, "A survey of mobile computing for the visually impaired," 2018, *arXiv:1811.10120*.

[22] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–19, 2019.

[23] S. Hakak, S. Ray, W. Z. Khan, and E. Scheme, "A framework for edge-assisted healthcare data analytics using federated learning," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2020, pp. 3423–3427.

[24] P. M. Mammen, "Federated learning: Opportunities and challenges," 2021, *arXiv:2101.05428*.

[25] Z. Zheng, Y. Zhou, Y. Sun, Z. Wang, B. Liu, and K. Li, "Applications of federated learning in smart cities: Recent advances, taxonomy, and open challenges," 2021, *arXiv:2102.01375*.

[26] M. Joshi, A. Pal, and M. Sankarasubbu, "Federated learning for healthcare domain—Pipeline, applications and challenges," *ACM Trans. Comput. Healthcare*, vol. 2022, pp. 1–15, May 2022.

[27] B. C. Tedeschini, S. Savazzi, R. Stoklasa, L. Barbieri, I. Stathopoulos, M. Nicoli, and L. Serio, "Decentralized federated learning for healthcare networks: A case study on tumor segmentation," *IEEE Access*, vol. 10, pp. 8693–8708, 2022.

[28] D. C. Nguyen, Q.-V. Pham, P. N. Pathirana, M. Ding, A. Seneviratne, Z. Lin, O. Dobre, and W.-J. Hwang, "Federated learning for smart healthcare: A survey," *ACM Comput. Surveys*, vol. 55, no. 3, pp. 1–37, Apr. 2023.

[29] Y. Xu, G. Xu, C. Ma, and Z. An, "An advancing temporal convolutional network for 5G latency services via automatic modulation recognition," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 6, pp. 3002–3006, Jun. 2022.

[30] Y. Jin, X. Wei, Y. Liu, and Q. Yang, "Towards utilizing unlabeled data in federated learning: A survey and prospective," 2020, *arXiv:2002.11545*.

[31] L. Lyu, H. Yu, and Q. Yang, "Threats to federated learning: A survey," 2020, *arXiv:2003.02133*.

[32] H. Ghayvat, S. Pandya, P. Bhattacharya, M. Zuhair, M. Rashid, S. Hakak, and K. Dev, "CP-BDHCA: Blockchain-based confidentiality-privacy preserving big data scheme for healthcare clouds and applications," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 5, pp. 1937–1948, May 2022.

[33] P. Bhattacharya, P. Mehta, S. Tanwar, M. S. Obaidat, and K.-F. Hsiao, "HeaL: A blockchain-envisioned signcryption scheme for healthcare IoT ecosystems," in *Proc. Int. Conf. Commun., Comput., Cybersecur., Informat. (CCCI)*, Sharjah, United Arab Emirates, Nov. 2020, pp. 1–6.

[34] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering—A systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, 2009.

[35] S. Keele, "Guidelines for performing systematic literature reviews in software engineering," Keele Univ., Durham Univ., Durham, U.K., Joint Rep. EBSE 2007-001, 2007.

[36] A. Dwivedi, R. K. Bali, M. A. Belsis, R. N. G. Naguib, P. Every, and N. S. Nassar, "Towards a practical healthcare information security model for healthcare institutions," in *Proc. 4th Int. IEEE EMBS Special Topic Conf. Inf. Technol. Appl. Biomed.*, Apr. 2003, pp. 114–117.

[37] A. Iyengar, A. Kundu, and G. Pallis, "Healthcare informatics and privacy," *IEEE Internet Comput.*, vol. 22, no. 2, pp. 29–31, Mar./Apr. 2018.

[38] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, "Federated learning for healthcare informatics," *J. Healthcare Informat. Res.*, vol. 5, no. 1, pp. 1–19, Mar. 2021.

[39] P. Bhatttacharya, K. Patel, M. Zuhair, and C. Trivedi, "A lightweight authentication via unclonable functions for industrial Internet-of-Things," in *Proc. 2nd Int. Conf. Innov. Practices Technol. Manage. (ICIPTM)*, Gautam Bhuddh Nagar, India, Feb. 2022, pp. 657–662.

[40] P. Bhattacharya, S. Tanwar, R. Shah, and A. Ladha, "Mobile edge computing-enabled blockchain framework—A survey," in *Proceedings of ICRIC 2019*, P. K. Singh, A. K. Kar, Y. Singh, M. H. Kolekar, and S. Tanwar, Eds. Cham, Switzerland: Springer, 2020, pp. 797–809.

[41] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Arcas, "Communication-efficient learning of deep networks from decentralized data," 2016, *arXiv:1602.05629*.

[42] S. Wang, T. Tuor, T. Salonidis, K. K. Leung, C. Makaya, T. He, and K. Chan, "Adaptive federated learning in resource constrained edge computing systems," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 3, pp. 1205–1221, Jun. 2019.

[43] F. Sattler, S. Wiedemann, K.-R. Müller, and W. Samek, "Sparse binary compression: Towards distributed deep learning with minimal communication," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Budapest, Hungary, Jul. 2019, pp. 1–8.

[44] M. G. Arivazhagan, V. Aggarwal, A. K. Singh, and S. Choudhary, "Federated learning with personalization layers," 2019, *arXiv:1912.00818*.

[45] H. Wang, M. Yurochkin, Y. Sun, D. Papailiopoulos, and Y. Khazaeni, "Federated learning with matched averaging," 2020, *arXiv:2002.06440*.

[46] S. Ek, F. Portet, P. Lalanda, and G. Vega, "A federated learning aggregation algorithm for pervasive computing: Evaluation and comparison," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Kassel, Germany, Mar. 2021, pp. 1–10.

[47] L. U. Khan, M. Alsenwi, Z. Han, and C. S. Hong, "Self organizing federated learning over wireless networks: A socially aware clustering approach," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Barcelona, Spain, Jan. 2020, pp. 453–458.

[48] S. R. Pandey, N. H. Tran, M. Bennis, Y. K. Tun, A. Manzoor, and C. S. Hong, "A crowdsourcing framework for on-device federated learning," 2019, *arXiv:1911.01046*.

[49] M. Nasr, R. Shokri, and A. Houmansadr, "Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning," in *Proc. IEEE Symp. Secur. Privacy (SP)*, San Francisco, CA, USA, May 2019, pp. 739–753.

[50] A. Vaid *et al.*, "Federated learning of electronic health records improves mortality prediction in patients hospitalized with COVID-19," *MedRxiv*, pp. 1–21, Aug. 2020, doi: 10.1101/2020.08.11.20172809.

[51] C. Zhang, S. Li, J. Xia, W. Wang, F. Yan, and Y. Liu, *BatchCrypt: Efficient Homomorphic Encryption for Cross-Silo Federated Learning*. Berkeley, CA, USA: USENIX Assoc., 2020, ch. 1, pp. 1–14.

[52] P. Bhattacharya, U. Bodkhe, M. Zuhair, M. Rashid, X. Liu, A. Verma, and R. K. Dewangan, "Amalgamation of blockchain and sixth-generation-envisioned responsive edge orchestration in future cellular vehicle-to-anything ecosystems: Opportunities and challenges," *Trans. Emerg. Telecommun. Technol.*, vol. 2021, p. e4410, Dec. 2021.

[53] R. Gupta, A. Kumari, and S. Tanwar, "A taxonomy of blockchain envisioned edge-as-a-connected autonomous vehicles," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 6, pp. 1–32, Jun. 2021.

[54] N. S. Patel, P. Bhattacharya, S. B. Patel, S. Tanwar, N. Kumar, and H. Song, "Blockchain-envisioned trusted random oracles for IoT-enabled probabilistic smart contracts," *IEEE Internet Things J.*, vol. 8, no. 19, pp. 14797–14809, Oct. 2021.

[55] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," 2018, *arXiv:1812.06127*.

[56] B. Han, R. Jhaveri, H. Wang, D. Qiao, and J. Du, "Application of robust zero-watermarking scheme based on federated learning for securing the healthcare data," *IEEE J. Biomed. Health Informat.*, early access, Oct. 29, 2021, doi: 10.1109/JBHI.2021.3123936.

[57] H. Chen, H. Li, G. Xu, Y. Zhang, and X. Luo, "Achieving privacy-preserving federated learning with irrelevant updates over E-health applications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[58] J. Li, Y. Meng, L. Ma, S. Du, H. Zhu, Q. Pei, and X. Shen, "A federated learning based privacy-preserving smart healthcare system," *IEEE Trans. Ind. Informat.*, vol. 18, no. 3, pp. 2021–2031, Mar. 2022.

[59] C. M. Thwal, K. Thar, Y. L. Tun, and C. S. Hong, "Attention on personalized clinical decision support system: Federated learning approach," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Jan. 2021, pp. 141–147.

[60] W. Y. B. Lim, S. Garg, Z. Xiong, D. Niyato, C. Leung, C. Miao, and M. Guizani, "Dynamic contract design for federated learning in smart healthcare applications," *IEEE Internet Things J.*, vol. 8, no. 23, pp. 16853–16862, Dec. 2021.

[61] H. Lin, K. Kaur, X. Wang, G. Kaddoum, J. Hu, and M. M. Hassan, "Privacy-aware access control in IoT-enabled healthcare: A federated deep learning approach," *IEEE Internet Things J.*, early access, Sep. 15, 2021, doi: 10.1109/JIOT.2021.3112686.

[62] T. Ryffel, A. Trask, M. Dahl, B. Wagner, J. Mancuso, D. Rueckert, and J. Passerat-Palmbach, "A generic framework for privacy preserving deep learning," 2018, *arXiv:1811.04017*.

[63] C. He, S. Li, J. So, X. Zeng, M. Zhang, H. Wang, X. Wang, P. Vepakomma, A. Singh, H. Qiu, X. Zhu, J. Wang, L. Shen, P. Zhao, Y. Kang, Y. Liu, R. Raskar, Q. Yang, M. Annavaram, and S. Avestimehr, "FedML: A research library and benchmark for federated machine learning," 2020, *arXiv:2007.13518*.

[64] X. Zhu, J. Wang, Z. Hong, T. Xia, and J. Xiao, "Federated learning of unsegmented Chinese text recognition model," in *Proc. IEEE 31st Int. Conf. Tools with Artif. Intell. (ICTAI)*, Nov. 2019, pp. 1341–1345.

[65] S. Caldas, S. M. K. Duddu, P. Wu, T. Li, J. Konečný, H. B. McMahan, V. Smith, and A. Talwalkar, "LEAF: A benchmark for federated settings," 2018, *arXiv:1812.01097*.

[66] I. Kholod, E. Yanaki, D. Fomichev, E. Shalugin, E. Novikova, E. Filippov, and M. Nordlund, "Open-source federated learning frameworks for IoT: A comparative review and analysis," *Sensors*, vol. 21, no. 1, p. 167, Dec. 2020.

[67] D. Chen, V. Tan, Z. Lu, and J. Hu, "OpenFed: A comprehensive and versatile open-source federated learning framework," 2021, *arXiv:2109.07852*.

[68] O. Marfoq, G. Neglia, A. Bellet, L. Kameni, and R. Vidal, "Federated multi-task learning under a mixture of distributions," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 34, 2021, pp. 1–14.

[69] M. Mohri, G. Sivek, and A. T. Suresh, "Agnostic federated learning," in *Proc. 36th Int. Conf. Mach. Learn.*, vol. 97, K. Chaudhuri and R. Salakhutdinov, Eds., Jun. 2019, pp. 4615–4625.

[70] T. Li, M. Sanjabi, A. Beirami, and V. Smith, "Fair resource allocation in federated learning," 2019, *arXiv:1905.10497*.

[71] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," in *Proc. Mach. Learn. Syst.*, vol. 2, I. Dhillon, D. Papailiopoulos, and V. Sze, Eds., 2020, pp. 429–450.

[72] S. R. Moosavi, T. N. Gia, E. Nigussie, A. M. Rahmani, S. Virtanen, H. Tenhunen, and J. Isoaho, "End-to-end security scheme for mobility enabled healthcare Internet of Things," *Future Gener. Comput. Syst.*, vol. 64, pp. 108–124, Mar. 2016.

[73] M. Hölbl, M. Kompara, A. Kamišalić, and L. N. Zlatolas, "A systematic review of the use of blockchain in healthcare," *Symmetry*, vol. 10, no. 10, p. 470, Oct. 2018.

[74] U. Bodkhe, S. Tanwar, P. Bhattacharya, and A. Verma, "Blockchain adoption for trusted medical records in healthcare 4.0 applications: A survey," in *Proc. 2nd Int. Conf. Comput., Commun., Cyber-Secur.*, P. K. Singh, S. T. Wierzchoń, S. Tanwar, M. Ganzha, and J. J. P. C. Rodrigues, Eds. Singapore: Springer, 2021, pp. 759–774.

[75] J. Sun, X. Zhu, C. Zhang, and Y. Fang, "HCPP: Cryptography based secure EHR system for patient privacy and emergency healthcare," in *Proc. 31st Int. Conf. Distrib. Comput. Syst.*, Jun. 2011, pp. 373–382.

[76] R. Xu, N. Baracaldo, Y. Zhou, A. Anwar, and H. Ludwig, "HybridAlpha: An efficient approach for privacy-preserving federated learning," in *Proc. 12th ACM Workshop Artif. Intell. Secur. (AISec)*, New York, NY, USA, 2019, pp. 13–23.

[77] Y. Zhao, J. Zhao, M. Yang, T. Wang, N. Wang, L. Lyu, D. Niyato, and K.-Y. Lam, "Local differential privacy-based federated learning for Internet of Things," *IEEE Internet Things J.*, vol. 8, no. 11, pp. 8836–8853, Jun. 2021.

[78] J. Li, M. Khodak, S. Caldas, and A. Talwalkar, "Differentially private meta-learning," 2019, *arXiv:1909.05830*.

[79] D. Reebadiya, T. Rathod, R. Gupta, S. Tanwar, and N. Kumar, "Blockchain-based secure and intelligent sensing for autonomous vehicles activity tracking beyond 5G networks," *Peer-Peer Netw. Appl.*, vol. 14, pp. 1–18, Sep. 2021.

[80] H. Pirnejad, R. Bal, A. P. Stoop, and M. Berg, "Inter-organisational communication networks in healthcare: Centralised versus decentralised approaches," *Int. J. Integr. Care*, vol. 7, no. 2, pp. 1–12, May 2007.

[81] P. Kierkegaard, "Electronic health record: Wiring Europe's healthcare," *Comput. Law Secur. Rev.*, vol. 27, no. 5, pp. 503–515, Sep. 2011.

[82] F. Shanin, H. A. A. Das, G. A. Krishnan, L. S. Neha, N. Thaha, R. P. Aneesh, S. Embrandiri, and S. Jayakrishan, "Portable and centralised E-health record system for patient monitoring using Internet of Things (IoT)," in *Proc. Int. CET Conf. Control, Commun., Comput. (IC)*, Jul. 2018, pp. 165–170.

[83] S. Silva, A. Altmann, B. Gutman, and M. Lorenzi, "Fed-BioMed: A general open-source frontend framework for federated learning in healthcare," in *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*. Cham, Switzerland: Springer, 2020, pp. 201–210.

[84] Y. Guo, F. Liu, Z. Cai, L. Chen, and N. Xiao, "FEEL: A federated edge learning system for efficient and privacy-preserving mobile healthcare," in *Proc. 49th Int. Conf. Parallel Process. (ICPP)*, Aug. 2020, pp. 1–11.

[85] I. Ullah, N. U. Amin, A. Almogren, M. A. Khan, M. I. Uddin, and Q. Hua, "A lightweight and secured certificate-based proxy signcryption (CB-PS) scheme for E-prescription systems," *IEEE Access*, vol. 8, pp. 199197–199212, 2020.

[86] A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar, and S. Mishra, "Decision tree and SVM-based data analytics for theft detection in smart grid," *IEEE Trans. Ind. Informat.*, vol. 12, no. 3, pp. 1005–1016, Jun. 2016.

[87] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, "Federated learning for Internet of Things: A comprehensive survey," 2021, *arXiv:2104.07914*.

[88] B. Yuan, S. Ge, and W. Xing, "A federated learning framework for healthcare IoT devices," 2020, *arXiv:2005.05083*.

[89] S. Tanwar, M. Mittal, B. Agarwal, and L. M. Goyal, *Energy Conservation for IoT Devices: Concepts, Paradigms Solutions*. Singapore: Springer, 2019.

[90] A. Qayyum, K. Ahmad, M. A. Ahsan, A. Al-Fuqaha, and J. Qadir, "Collaborative federated learning for healthcare: Multi-modal COVID-19 diagnosis at the edge," 2021, *arXiv:2101.07511*.

[91] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen, and M. Chen, "In-Edge AI: Intelligentizing mobile edge computing, caching and communication by federated learning," *IEEE Netw.*, vol. 33, no. 5, pp. 156–165, Sep. 2019.

[92] W. Y. B. Lim, N. C. Luong, D. T. Hoang, Y. Jiao, Y.-C. Liang, Q. Yang, D. Niyato, and C. Miao, "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031–2063, 3rd Quart., 2020.

[93] O. Choudhury, A. Gkoulalas-Divanis, T. Salonidis, I. Sylla, Y. Park, G. Hsu, and A. Das, "Differential privacy-enabled federated learning for sensitive health data," 2019, *arXiv:1910.02578*.

[94] W. Li, F. Milletar, D. Xu, N. Rieke, J. Hancox, W. Zhu, M. Baust, Y. Cheng, S. Ourselin, M. J. Cardoso, and A. Feng, "Privacy-preserving federated brain tumour segmentation," in *Machine Learning in Medical Imaging*, H.-I. Suk, M. Liu, P. Yan, and C. Lian, Eds. Cham, Switzerland: Springer, 2019, pp. 133–141.

[95] M. A. P. Chamikara, P. Bertok, I. Khalil, D. Liu, and S. Camtepe, "Privacy preserving distributed machine learning with federated learning," *Comput. Commun.*, vol. 171, pp. 112–125, Apr. 2021.

[96] P. Kanani, V. J. Marathe, D. Peterson, R. Harpaz, and S. Bright, "Private cross-silo federated learning for extracting vaccine adverse event mentions," 2021, *arXiv:2103.07491*.

[97] Y. Chen, X. Qin, J. Wang, C. Yu, and W. Gao, "FedHealth: A federated transfer learning framework for wearable healthcare," *IEEE Intell. Syst.*, vol. 35, no. 4, pp. 83–93, Jul. 2020.

[98] M. Malekzadeh, B. Hasircioglu, N. Mital, K. Katarya, M. E. Ozfatura, and D. Gündüz, "Dopamine: Differentially private federated learning on medical data," 2021, *arXiv:2101.11693*.

[99] J. Ma, S.-A. Naas, S. Sigg, and X. Lyu, "Privacy-preserving federated learning based on multi-key homomorphic encryption," 2021, *arXiv:2104.06824*.

[100] Q. Li, Z. Wen, Z. Wu, S. Hu, N. Wang, Y. Li, X. Liu, and B. He, "A survey on federated learning systems: Vision, hype and reality for data privacy and protection," 2019, *arXiv:1907.09693*.

[101] S. Rajendran, J. S. Obeid, H. Binol, R. D. Agostino, K. Foley, W. Zhang, P. Austin, J. Brakefield, M. N. Gurcan, and U. Topaloglu, "Cloud-based federated learning implementation across medical centers," *JCO Clin. Cancer Informat.*, vol. 5, pp. 1–11, Dec. 2021.

[102] Y. Mansour, M. Mohri, J. Ro, and A. T. Suresh, "Three approaches for personalization with applications to federated learning," 2020, *arXiv:2002.10619*.

[103] A. Das and T. Brunschwiler, "Privacy is what we care about: Experimental investigation of federated learning on edge devices," in *Proc. 1st Int. Workshop Challenges Artif. Intell. Mach. Learn. Internet Things*, Nov. 2019, pp. 39–42.

[104] Y. Liu, X. Yuan, Z. Xiong, J. Kang, X. Wang, and D. Niyato, "Federated learning for 6G communications: Challenges, methods, and future directions," *China Commun.*, vol. 17, no. 9, pp. 105–118, Sep. 2020.

[105] J. Kaur, M. A. Khan, M. Iftikhar, M. Imran, and Q. E. Ul Haq, "Machine learning techniques for 5G and beyond," *IEEE Access*, vol. 9, pp. 23472–23488, 2021.

[106] Y. Lu, X. Huang, K. Zhang, S. Maharjan, and Y. Zhang, "Low-latency federated learning and blockchain for edge association in digital twin empowered 6G networks," *IEEE Trans. Ind. Informat.*, vol. 17, no. 7, pp. 5098–5107, Jul. 2021.

[107] R. Chaudhary, A. Jindal, G. S. Aujla, N. Kumar, A. K. Das, and N. Saxena, "LSCSH: Lattice-based secure cryptosystem for smart healthcare in smart cities environment," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 24–32, Apr. 2018.

[108] R. Kakkar, R. Gupta, S. Tanwar, and J. J. P. C. Rodrigues, "Coalition game and blockchain-based optimal data pricing scheme for ride sharing beyond 5G," *IEEE Syst. J.*, early access, Dec. 1, 2021, doi: 10.1109/JSYST.2021.3126620.

[109] G. Xu, J. Dong, C. Ma, J. Liu, and U. G. O. Cliff, "A certificateless signcryption mechanism based on blockchain for edge computing," *IEEE Internet Things J.*, early access, Feb. 15, 2022, doi: 10.1109/JIOT.2022.3151359.

[110] R. Gupta, D. Reebadiya, and S. Tanwar, "6G-enabled edge intelligence for ultra-reliable low latency applications: Vision and mission," *Comput. Standards Interfaces*, vol. 77, Aug. 2021, Art. no. 103521.

[111] M. N. H. Nguyen, N. H. Tran, Y. K. Tun, Z. Han, and C. S. Hong, "Toward multiple federated learning services resource sharing in mobile edge networks," 2020, *arXiv:2011.12469*.

[112] J. Xu, H. Wang, and L. Chen, "Bandwidth allocation for multiple federated learning services in wireless edge networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2534–2546, Apr. 2022.

[113] L. U. Khan, L. U. Khan, S. R. Pandey, N. H. Tran, W. Saad, Z. Han, M. N. H. Nguyen, and C. S. Hong, "Federated learning for edge networks: Resource optimization and incentive mechanism," *IEEE Commun. Mag.*, vol. 58, no. 10, pp. 88–93, Oct. 2020.

[114] Y. Zhan, P. Li, Z. Qu, D. Zeng, and S. Guo, "A learning-based incentive mechanism for federated learning," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6360–6368, Jul. 2020.

[115] H. Wang, K. Sreenivasan, S. Rajput, H. Vishwakarma, S. Agarwal, J.-Y. Sohn, K. Lee, and D. Papailiopoulos, "Attack of the tails: Yes, you really can backdoor federated learning," in *Advances in Neural Information Processing Systems*, vol. 33, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds. Red Hook, NY, USA: Curran Associates, 2020, pp. 16070–16084.

[116] G. Saldamli, C. Upadhyay, D. Jadhav, R. Shrishrimal, B. Patil, and L. Tawalbeh, "Improved gossip protocol for blockchain applications," *Cluster Comput.*, vol. 25, no. 3, pp. 1915–1926, Jun. 2022.

[117] U. Bodkhe, D. Mehta, S. Tanwar, P. Bhattacharya, P. K. Singh, and W.-C. Hong, "A survey on decentralized consensus mechanisms for cyber physical systems," *IEEE Access*, vol. 8, pp. 54371–54401, 2020.

[118] J. So, B. Guler, and A. S. Avestimehr, "Turbo-aggregate: Breaking the quadratic aggregation barrier in secure federated learning," *IEEE J. Sel. Areas Inf. Theory*, vol. 2, no. 1, pp. 479–489, Mar. 2021.

[119] C. Wang, C. Ma, M. Li, N. Gao, Y. Zhang, and Z. Shen, "Protecting data privacy in federated learning combining differential privacy and weak encryption," in *Science of Cyber Security*, W. Lu, K. Sun, M. Yung, and F. Liu, Eds. Cham, Switzerland: Springer, 2021, pp. 95–109.

[120] P. Bhattacharya, D. Saraswat, A. Dave, M. Acharya, S. Tanwar, G. Sharma, and I. E. Davidson, "Coalition of 6G and blockchain in AR/VR space: Challenges and future directions," *IEEE Access*, vol. 9, pp. 168455–168484, 2021.

[121] P. Bhattacharya, A. Singh, A. Kumar, A. K. Tiwari, and R. Srivastava, "Comparative study for proposed algorithm for all-optical network with negative acknowledgement (AO-NACK)," in *Proc. 7th Int. Conf. Comput. Commun. Technol. (ICCCT)*, New York, NY, USA, 2017, pp. 47–51.

[122] V. A. Patel, P. Bhattacharya, S. Tanwar, N. K. Jadav, and R. Gupta, "BFLEdge: Blockchain based federated edge learning scheme in V2X underlying 6G communications," in *Proc. 12th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence)*, Noida, India, Jan. 2022, pp. 146–152.
[123] R. Saha, S. Misra, and P. K. Deb, "FogFL: Fog-assisted federated learning for resource-constrained IoT devices," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8456–8463, May 2021.
[124] H. Wang, S. Sievert, S. Liu, Z. Charles, D. Papailiopoulos, and S. Wright, "ATOMO: Communication-efficient learning via atomic sparsification," in *Advances in Neural Information Processing Systems*, vol. 31, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2018.
[125] M. Khodak, M.-F. Balcan, and A. Talwalkar, *Adaptive Gradient-Based Meta-Learning Methods*. Red Hook, NY, USA: Curran Associates, 2019, ch. 1, pp. 1–12.
[126] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 50–60, May 2020.
[127] Y. Lin, S. Han, H. Mao, Y. Wang, and W. J. Dally, "Deep gradient compression: Reducing the communication bandwidth for distributed training," in *Proc. Int. Conf. Learn. Represent.*, 2018, pp. 1–14.
[128] D. Rothchild, A. Panda, E. Ullah, N. Ivkin, I. Stoica, V. Braverman, J. Gonzalez, and R. Arora, "FetchSGD: Communication-efficient federated learning with sketching," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 8253–8265.
[129] F. Sattler, S. Wiedemann, K.-R. Müller, and W. Samek, "Robust and communication-efficient federated learning from non-i.i.d. Data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 9, pp. 3400–3413, Sep. 2020.
[130] L. Corinzia, A. Beuret, and J. M. Buhmann, "Variational federated multi-task learning," 2019, *arXiv:1906.06268*.
[131] J. C. Jiang, B. Kantarci, S. Oktug, and T. Soyata, "Federated learning in smart city sensing: Challenges and opportunities," *Sensors*, vol. 20, no. 21, p. 6230, Oct. 2020.
[132] L. U. Khan, W. Saad, Z. Han, E. Hossain, and C. S. Hong, "Federated learning for Internet of Things: Recent advances, taxonomy, and open challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1759–1799, 2021.
[133] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, "Practical secure aggregation for privacy-preserving machine learning," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, New York, NY, USA, 2017, pp. 1175–1191.
[134] A. Kumari, R. Gupta, and S. Tanwar, "Amalgamation of blockchain and IoT for smart cities underlying 6G communication: A comprehensive review," *Comput. Commun.*, vol. 172, pp. 102–118, Apr. 2021.
[135] C. Hou, K. K. Thekumparampil, G. Fanti, and S. Oh, "Reducing the communication cost of federated learning through multistage optimization," in *Proc. Int. Conf. Learn. Represent.*, 2022, pp. 1–49.

VISHWA AMITKUMAR PATEL received the Bachelor of Technology degree in computer engineering from the Sardar Vallabhbhai Patel Institute of Technology. She is currently working on presenting solutions to integrate federated learning in emerging domains, such as healthcare, vehicular networks, and emerging communication networks. Her research interests include the IoT, federated learning, and blockchain.

PRONAYA BHATTACHARYA (Member, IEEE) is currently working as an Assistant Professor with the Computer Science and Engineering Department, Institute of Technology, Nirma University, Ahmedabad, India. He has over eight years of teaching experience. He has authored or coauthored more than 70 research papers in leading SCI journals and top core IEEE COMSOC A* conferences. Some of his top-notch findings are published in reputed SCI journals, such as IEEE Journal of Biomedical and Health Informatics, IEEE Transactions on Vehicular Technology, IEEE Internet of Things Journal, IEEE Transactions on Network Science and Engineering, IEEE Transactions on Computational Social Systems, IEEE Transactions of Network and Service Management, IEEE Access, IEEE Sensors, *IEEE Internet of Things Magazine*, *IEEE Communication Standards Magazine*, *ETT* (Wiley), *Expert Systems* (Wiley), *FGCS* (Elsevier), *OQEL* (Springer), *WPC* (Springer), ACM-MOBICOM, IEEE-INFOCOM, IEEE-ICC, IEEE-CITS, IEEE-ICIEM, IEEE-CCCI, and IEEE-ECAI. He has 1061 citations to his credit with an H-index of 17 and an i10-index of 25. His research interests include healthcare analytics, optical switching and networking, federated learning, blockchain, and the IoT. He has been appointed at the capacity of a Keynote Speaker, a Technical Committee Member, and the Session Chair across the globe. He was awarded eight best paper awards in Springer ICRIC-2019, IEEE-ICIEM-2021, IEEE-ECAI-2021, Springer COMS2-2021, and IEEE-ICIEM-2022. He is a Reviewer of 21 reputed SCI journals, such as IEEE Internet of Things Journal, IEEE Transactions on Industrial Informatics, IEEE Transactions of Vehicular Technology, IEEE Journal of Biomedical and Health Informatics, IEEE Access, IEEE Network, *ETT* (Wiley), *IJCS* (Wiley), *MTAP* (Springer), *OSN* (Elsevier), *WPC* (Springer), and others.

SUDEEP TANWAR (Senior Member, IEEE) is currently working as a Professor with the Computer Science and Engineering Department, Institute of Technology, Nirma University, India. He is also a Visiting Professor at Jan Wyzykowski University, Polkowice, Poland; and the University of Pitesti, Pitesti, Romania. He has authored two books, edited 13 books, and more than 270 technical papers, including top journals and top conferences, such as IEEE Transactions on Network Science and Engineering, IEEE Transactions on Vehicular Technology, IEEE Transactions on Industrial Informatics, IEEE Wireless Communications, IEEE Network, ICC, GLOBECOM, and INFOCOM. He initiated the research field of blockchain technology adoption in various verticals, in 2017. His H-index is 50. He actively serves his research communities in various roles. His research interests include blockchain technology, wireless sensor networks, fog computing, smart grid, and the IoT. He is a member of the Technical Committee on Tactile Internet of the IEEE Communication Society. He is a Senior Member of CSI, IAENG, ISTE, and CSTA. He has been awarded the Best Research Paper Awards from IEEE GLOBECOM 2018, IEEE ICC 2019, and Springer ICRIC-2019. He has served many international conferences as a member of the Organizing Committee, such as the Publication Chair for FTNCT-2020, ICCIC 2020, and WiMob2019; a member of the Advisory Board for ICACCT-2021 and ICACI 2020; the Workshop Co-Chair for CIS 2021; and the General Chair for IC4S 2019 and 2020 and ICCSDF 2020. He is serving on the editorial boards for *Frontiers of Blockchain*, *Cyber Security and Applications*, *Computer Communications*, the *International Journal of Communication Systems*, and *Security and Privacy*.

RAJESH GUPTA (Student Member, IEEE) received the Bachelor of Engineering degree from the University of Jammu, India, in 2008, and the master's degree in technology from Shri Mata Vaishno Devi University, Jammu, India, in 2013. He is currently a full-time Ph.D. Research Scholar with the Computer Science and Engineering Department, Nirma University, Ahmedabad, Gujarat, India. He has authored/coauthored some publications (including papers in SCI indexed journals and IEEE ComSoc sponsored international conferences). Some of his research findings are published in top-cited journals and conferences, such as IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS, *IEEE Network Magazine*, *IEEE Internet of Things Magazine*, *Computer Communications*, *Computer and Electrical Engineering*, *IJCS* (Wiley), *ETT* (Wiley), *Physical Communication*, IEEE ICC, IEEE INFOCOM, IEEE GLOBECOM, IEEE CITS, and many more. His research interests include device-to-device communication, network security, blockchain technology, 5G communication networks, and machine learning. His H-index is 23 and i10-index is 32. He was also a recipient of Doctoral Scholarship from the Ministry of Electronics and Information Technology, Government of India, under the Visvesvaraya Ph.D. Scheme. He was a recipient of Student Travel Grant from WICE-IEEE to attend IEEE ICC 2021 in Canada. He has been awarded Best Research Paper Awards from IEEE ECAI 2021, IEEE ICCCA, and IEEE IWCMC 2021. His name has been included in the list of Top 2% scientists worldwide published by the Stanford University, USA. He was felicitated by Nirma University for their research achievements, in 2021. He is also an Active Member of the ST Research Laboratory (www.sudeeptanwar.in).

GULSHAN SHARMA received the B.Tech., M.Tech., and Ph.D. degrees. He is currently working as a Senior Lecturer with the Department of Electrical Engineering Technology, University of Johannesburg. He is also a Y Rated Researcher with NRF South Africa. His research interests include power system operation and control and application of AI techniques to the power systems. He is working as an Academic Editor of *International Transactions on Electrical Energy System* (Wiley) and a Regional Editor of *Recent Advances in E & EE* (Bentham Science).

PITSHOU N. BOKORO received the M.Phil. degree in electrical engineering from the University of Johannesburg, Johannesburg, South Africa in 2011, and the Ph.D. degree in electrical engineering from the University of the Witwatersrand, in 2016. He is currently an Associate Professor with the University of Johannesburg. His research interests include modeling and reliability prediction of insulating materials and dielectrics, power quality, and renewable energies. He is a Senior Member of the South African Institute of Electrical Engineers.

RAVI SHARMA is currently working as a Professor with the Centre for Inter-Disciplinary Research and Innovation, University of Petroleum and Energy Studies, Dehradun, India. He is passionate in the field of business analytics and worked in various MNCs as a Leader of various software development groups. He has contributed various articles in the area of business analytics, prototype building for startup, and artificial intelligence. He is leading academic institutions as a Consultant to uplift research activities in inter-disciplinary domains.