Federated Learning for Smart Healthcare: A Survey

DINH C. NGUYEN*, School of Engineering, Deakin University, Australia
QUOC-VIET PHAM†, Korean Southeast Center for the 4th Industrial Revolution Leader Education, Pusan National University, Korea
PUBUDU N. PATHIRANA*, Network Sensing and Biomedical Engineering, School of Engineering, Deakin University, Australia
MING DING‡, Data61, CSIRO, Australia
ARUNA SENEVIRATNE§, School of Electrical Engineering and Telecommunications, University of New South Wales (UNSW), Australia
ZIHUAI LIN¶, Department of Engineering, The University of Sydney, Australia
OCTAVIA DOBRE‖, Faculty of Engineering and Applied Science, Memorial University, Canada
WON-JOO HWANG**, Department of Biomedical Convergence Engineering, Pusan National University, Korea

Recent advances in communication technologies and Internet-of-Medical-Things have transformed smart healthcare enabled by artificial intelligence (AI). Traditionally, AI techniques require centralized data collection and processing that may be infeasible in realistic healthcare scenarios due to the high scalability of modern healthcare networks and growing data privacy concerns. Federated Learning (FL), as an emerging distributed collaborative AI paradigm, is particularly attractive for smart healthcare, by coordinating multiple clients (e.g., hospitals) to perform AI training without sharing raw data. Accordingly, we provide a comprehensive survey on the use of FL in smart healthcare. First, we present the recent advances in FL, the motivations, and the requirements of using FL in smart healthcare. The recent FL designs for smart healthcare are then discussed, ranging from resource-aware FL, secure and privacy-aware FL to incentive FL and personalized FL. Subsequently, we provide a state-of-the-art review on the emerging applications of FL in key healthcare domains, including health data management, remote health monitoring, medical imaging, and COVID-19 detection. Several recent FL-based smart healthcare projects are analyzed, and the key lessons learned from the survey are also highlighted. Finally, we discuss interesting research challenges and possible directions for future FL research in smart healthcare.

CCS Concepts: • **General and reference → Surveys and overviews**; • **Security and privacy**; • **Computing methodologies → Distributed computing methodologies**; **Artificial intelligence**; • **Networks → Network services**;

Authors' addresses: Dinh C. Nguyen, cdnguyen@deakin.edu.au, School of Engineering, Deakin University, Waurn Ponds, Australia, 3216; Quoc-Viet Pham, vietpq@pusan.ac.kr, Korean Southeast Center for the 4th Industrial Revolution Leader Education, Pusan National University, Busan, Korea, 46241; Pubudu N. Pathirana, pubudu.pathirana@deakin.edu.au, Network Sensing and Biomedical Engineering, School of Engineering, Deakin University, Waurn Ponds, Australia, 3216; Ming Ding, ming.ding@data61.csiro.au, Data61, CSIRO, Sydney, Australia, 2015; Aruna Seneviratne, a.seneviratne@unsw.edu.au, School of Electrical Engineering and Telecommunications, University of New South Wales (UNSW), Sydney, Australia, 2052; Zihuai Lin, zihuai.lin@sydney.edu.au, Department of Engineering, The University of Sydney, Sydney, Australia, 2006; Octavia Dobre, odobre@mun.ca, Faculty of Engineering and Applied Science, Memorial University, Canada, A1B 3X5; Won-Joo Hwang, wjhwang@pusan.ac.kr, Department of Biomedical Convergence Engineering, Pusan National University, Gyeongsangnam, Korea, 50612.

0360-0300/2021/11-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Additional Key Words and Phrases: Federated learning, smart healthcare, privacy.

ACM Reference Format:
Dinh C. Nguyen, Quoc-Viet Pham, Pubudu N. Pathirana, Ming Ding, Aruna Seneviratne, Zihuai Lin, Octavia Dobre, and Won-Joo Hwang. 2021. Federated Learning for Smart Healthcare: A Survey. *ACM Comput. Surv.* 1, 1 (November 2021), 35 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The revolution in the Internet-of-Medical-Things (IoMT) has transformed the healthcare industry for improving the quality of human life [1]. In the smart healthcare environment, IoMT devices such as wearable sensors are widely used to collect medical data for intelligent data analytics enabled by artificial intelligence (AI) [2] to realize a plethora of exciting smart healthcare applications, such as remote health monitoring and disease prediction. For example, AI techniques such as deep learning (DL) have demonstrated their great potential in bio-medical image analytics for the early detection of chronic diseases by handling a large amount of health data to facilitate the provision of healthcare services [3].

Traditionally, smart healthcare systems often rely on centralized AI functions located at the cloud or the data center for health data learning and analytics. Given the increasing volumes of health data and the growth of IoMT devices in modern healthcare networks, this centralized solution is not efficient in terms of communication latency due to raw data transmission and cannot achieve high network scalability. Further, the reliance on such a central server or third party for data learning raises critical privacy issues, e.g., user information leakage and data breach [4]. This is particularly true in e-healthcare, where health-related information is highly sensitive and private subject to health regulations such as the United States Health Insurance Portability and Accountability Act (HIPPA) [5]. Moreover, in the future healthcare systems, such a centralized AI architecture may be no longer suitable because health data are not centrally located, but distributed over a large-scale IoMT network. Therefore, there is an urgent need to go toward distributed AI approaches for enabling scalable and privacy-preserving intelligent healthcare applications at the network edge.

In this context, federated learning (FL) has emerged as a promising solution for realizing cost-effective smart healthcare applications with improved privacy protection [6–9]. Conceptually, FL is a distributed AI approach which enables the training of high-quality AI models by averaging local updates aggregated from multiple health data clients, e.g., IoMT devices, without the need for direct access to the local data [10]. This potentially prevents disclosing sensitive user information and user preference, and thus mitigates privacy leakage risks. Moreover, since FL attracts large computation and dataset resources from a number of health data clients to train AI models, the health data training quality, e.g., accuracy, would be significantly improved which might not be achieved by using centralized AI approaches with less data and limited computational capabilities.

1.1 Comparison and Our Contributions

Driven by the recent advances of FL, many studies have been conducted to survey its related topics, including healthcare. For example, the works in [11] and [12] present the key FL concept and its enabling protocols and technical challenges in FL design and implementation. The survey in [13] discusses the security and privacy issues in FL systems, and described possible solutions for evaluations of malicious threats in FL networks. The integration of FL in mobile edge networks is investigated in [14], where challenges in FL implementation are explored, such as communication costs, resource allocation, and privacy and security. Meanwhile, the convergence of FL and Internet-of-Things (IoT) is explored in [15], by providing a survey on the technical issues in FL designs, such as sparsification, robustness, privacy, and scalability, along with a brief discussion of FL applications in IoT. Moreover, the authors in [16] present an overview of the FL applications in industrial IoT, where the focus is on the discussion of characteristics and fundamentals of FL, while the discussion of FL usage in healthcare is

Table 1. Existing surveys on FL-related topics and our new contributions.

Paper	Key topic	Recent Advances in FL				Taxonomy	Highlights
		Resource management	Security and privacy	Incentive mechanism	Personalized FL		
[11]	FL concept	✗	✗	✗	✗	None	A discussion of the architectures, algorithms, and data processing methods in FL systems.
[12]	FL concept	✓	✓	✗	✗	None	A survey of the FL concepts, technologies and associated learning approaches.
[13]	Security and privacy in FL	✗	✓	✗	✗	None	A review on the security and privacy issues in FL systems.
[14]	FL in edge networks	✓	✗	✓	✗	None	A survey on the integration of FL in mobile edge networks.
[15]	FL for IoT	✓	✓	✓	✗	None	A survey on the use of FL in IoT networks.
[16]	FL for IIoT	✓	✓	✗	✗	The discussion of FL in healthcare is very limited.	A survey on the combination of FL and IIoT, mostly focusing on technical issues in FL implementation.
[17]	FL for health informatics	✗	✗	✗	✗	The discussion of FL in healthcare is very limited.	A study on the FL architectures and models, with a very short introduction to healthcare.
[18]	FL for digital health	✗	✗	✗	✗	The discussion of FL in healthcare is very limited.	A discussion of technical issues and requirements of FL in digital health.
Our work	FL for smart healthcare	✓	✓	✓	✓	A holistic taxonomy is presented.	A comprehensive survey on the use of FL in smart healthcare, from motivations, requirements to FL designs and applications in a wide range of healthcare domains.

limited. The work in [17] mainly discusses the FL architectures and models with a very brief introduction to the roles of FL in healthcare informatics. Another study in [18] mostly considers technical issues and requirements of using FL in future digital health. However, the latest advances in FL such as resource-aware FL, secure and privacy-enhanced FL, incentive-aware FL, and personalized FL have not been fully explored. The comparison of the related works and our paper is summarized in Table 1.

Despite these research efforts, there is no existing work to provide a comprehensive survey of the applications of FL in smart healthcare, to the best of our knowledge. Moreover, a holistic taxonomy of the use of FL in emerging healthcare applications is still missing in the open literature. These limitations motivate us to conduct an extensive review of the integration of FL in smart healthcare. Particularly, we first identify the key motivations and highlight the requirements of using FL in smart healthcare. Then, we discover the latest advanced FL designs used for smart healthcare. Subsequently, we provide a state-of-the-art survey on emerging applications of FL in smart healthcare, such as electronic health record (EHR) management, remote health monitoring, medical imaging, and COVID-19 detection. The lessons learned from the survey are also summarized to provide readers with more insights into the use of FL in smart healthcare. Finally, the research challenges and future directions in FL-smart healthcare are highlighted. To this end, the key contributions of this article are summarized as follows:

(1) We present a state-of-the-art survey on the use of FL in smart healthcare, beginning with an introduction to the FL concept and discussions of the motivations as well as the technical requirements for the utilization of FL smart healthcare.
(2) We present the recently advanced FL designs that would be useful to federated smart healthcare applications, including resource-aware FL, secure and privacy-enhanced FL, incentive-aware FL, and personalized FL.
(3) We provide an updated review on the key applications of FL in smart healthcare via a wide range of key domains, namely federated EHRs management, federated remote health monitoring, federated medical imaging, and federated COVID-19 detection. The emerging real-world projects related to FL-healthcare use cases are provided, and the key lessons learned from the survey are also highlighted.
(4) Finally, we highlight interesting challenges and discuss future directions in FL-smart healthcare.

1.2 Structure of The Survey

The remainder of the article is organized as follows. Section 2 introduces the key principle of FL and describes the key FL types used in smart healthcare. The key motivations and technical requirements of the use of FL in smart healthcare are explained in Section 3. Subsequently, we present the advanced FL designs that are useful to federated smart healthcare in Section 4. In Section 5, we present a state-of-the-art review on the emerging applications of FL in smart healthcare, namely federated EHRs management, federated remote health monitoring, federated medical imaging, and federated COVID-19 detection. The real-world projects of FL implementation for smart healthcare via some practical use cases are highlighted in Section 6. We discuss the key challenges and future directions related to FL-Healthcare research in Section 7 and Section 8 concludes the article.

2 FL FOR HEALTHCARE: KEY PRINCIPLE AND CATEGORIES

In this section, we present the key principle of FL and describe the key FL types used in smart healthcare.

2.1 Key Principle

As shown in Fig. 1, the generic FL-smart healthcare process includes the following key steps.

(1) *System Initialization and Client Selection:* The aggregation server selects a healthcare analytic task, e.g., automatic medical imaging or human motion detection, along with model requirements such as task classification or task prediction, and learning parameters such as neural node numbers and learning rates. Moreover, the server selects a subset of clients that should be involved in the FL process.
(2) *Distributed Local Training and Updates:* Once the subset of the learning clients is determined, the server sends an initial model including an initial global gradient to the clients to trigger the distributed training. In every communication round, each client trains a local model using its own dataset and calculates its model update, e.g., the gradient in neural networks. Then, each client uploads its model update to the server for aggregation.
(3) *Model Aggregation and Download:* After receiving all updates from the selected clients, the server aggregates them by using an aggregation method. For example, we can use the model averaging approach in the Federated Averaging (FedAvg) algorithm proposed by Google [19], where the gradient parameters of local models are averaged element-wise with weights proportional to the sizes of the client datasets. Subsequently, the server calculates a new version of the global model and broadcasts it to all clients as the basis for further local model updates in the next learning round. The FL process is iterated until the global loss function converges or the desired accuracy is achieved.

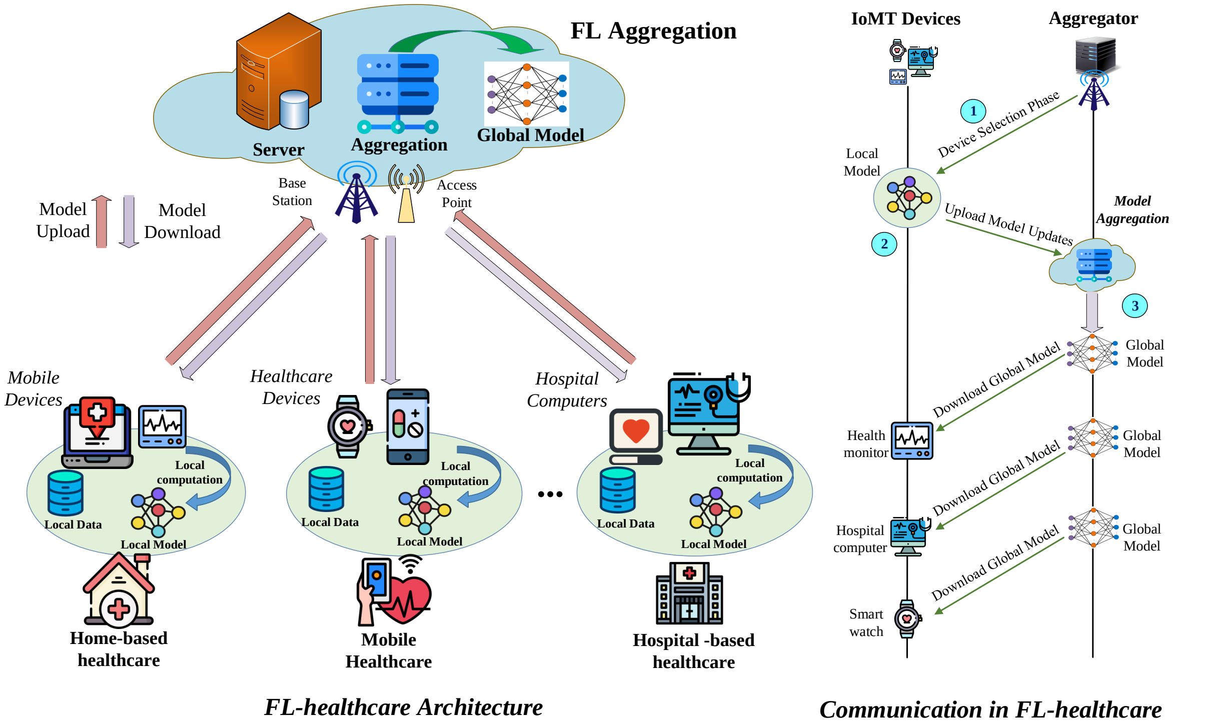
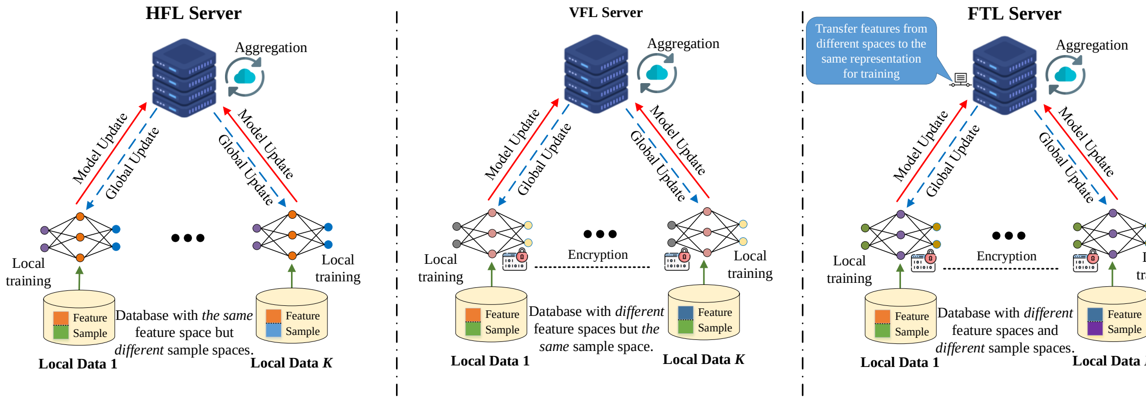


Fig. 1. The network architecture and communication process for FL-smart healthcare.

2.2 Types of FL for Smart Healthcare

Reviewing the recent advances of FL algorithms used in smart healthcare, we categorize FL into three types as illustrated in Fig. 2.

- *Horizontal FL (HFL):* In HFL, the healthcare clients can participate in training a shared global model using their datasets which own the same feature space while having different sample spaces, as shown in Fig. 2(a). In this regard, local FL participants can adopt the same AI model (e.g., neural network-NN) for training their datasets. Subsequently, the server will combine the local updates transmitted from local participants to build a global update without the need for direct access to local data [20]. An HFL example in smart healthcare can be the detection of speech disorders where multiple users speak the same sentence (feature space) with different voices (sample space) on their smartphones and then the local speaking updates are averaged by a parameter server to create a global model for speech recognition.
- *Vertical FL (VFL):* VHL works on the federated training of health datasets which have the same sample space with different data feature spaces, as illustrated in Fig. 2(b). Particularly, to address the issue of data sample overlapping at distributed clients, solutions based on entity alignment can be employed by integrating with encryption techniques during the local training [21]. An example of VFL in IMoT applications can be the shared learning model among entities in a smart healthcare environment, e.g., hospitals and an insurance company. In this context, a hospital and an insurance company (different data feature) which serve patients (same sample space) can join a VFL process to cooperatively train an AI model using their datasets, e.g., historical medical records at hospitals and healthcare costs at the insurance company for intelligent healthcare decision making.


HFL Server
Aggregation
Model Update
Global Update
Local training
Feature
Sample
Database with the same feature space but different sample spaces.
Local Data 1
Local Data K
VFL Server
Encryption
Database with different feature spaces but the same sample space.
Transfer features from different spaces to the same representation for training
FTL Server
Database with different feature spaces and different sample spaces.

(a) Horizontal Federated Learning (HFL)

(b) Vertical Federated Learning (VFL)

(c) Federated Transfer Learning (FTL)

Fig. 2. Categories of FL used in smart healthcare.

- *Federated Transfer Learning (FTL):* Unlike VFL systems, FTL is provided to handle datasets with different sample spaces and different feature spaces, as shown in Fig. 2(c). By using a transfer learning method, feature values are calculated from different feature spaces to the same representation which is exploited to train local datasets [22]. Encryption techniques such as random masks are also useful to provide further privacy protection during the gradient exchange between clients and the server [23]. In smart healthcare, FTL can support disease diagnosis by collaborating countries with multiple hospitals that have different patients (sample space) with different therapeutic programs (feature space). In this way, FTL can enrich the shared AI model output for improving the accuracy of diagnosis.

3 MOTIVATIONS AND REQUIREMENTS OF USING FL IN SMART HEALTHCARE

In this section, we explain the key motivations and describe the technical requirements of the use of FL in smart healthcare in details.

3.1 Motivations

To highlight the motivations, we first identify the limitations of current healthcare systems, and then discuss the key benefits that FL can offer for healthcare.

3.1.1 Limitations of Current Smart Healthcare Systems.

- *Privacy Concerns:* As briefly discussed in the previous sections, the use of traditional AI-based approaches for realizing smart healthcare require open data sharing with cloud or data centres which makes health information vulnerable to privacy attacks. Indeed, adversaries can gain unauthorized access to the AI training centres for data retrieval, or third parties such as cloud providers can gain the control over the data and modify data patterns without the consent of users [24]. These bottlenecks would result in serious issues of data leakage and the user confidentiality can be compromised. Although the cloud servers with powerful computation capabilities can provide efficient data training and analytics, such a centralized AI-based smart healthcare solution comes at a heavy cost of privacy risks [25].
- *Lack of Datasets at Medical Sites:* In realistic healthcare systems, the dataset of a single medical site (e.g., a clinical lab) may be insufficient to run the AI model which can prevent a proper health data training [26]. This makes AI-based smart healthcare solutions inefficient and manual data analyses need to be taken,

but this incurs long data processing delays. A possible solution is to exchange the data between medical sites to support the data training, but given the institutional policies and growing user privacy concerns, it is not easy to obtain data from other sites to train the AI model [27]. Hence, how to solve the issues of dataset shortage is of paramount importance for smart healthcare system designs.

- *Limited Health Data Training Performance:* Due to the lack of datasets, the training at an single medical site cannot achieve the desired degree of accuracy, e.g., disease classification accuracy. The reasons behind this observation can come from the imbalance of data features and the insufficient data sizes. One can use data augmentation techniques such as generative adversarial networks (GANs) [28] to solve these issues, but it may still not have good diversity to build a comprehensive dataset for efficient data training. This is also one of the most critical challenge in applying AI in healthcare, where the training becomes more difficult due to the limited datasets.
- *High Costs in Health Data Training:* In the traditional AI-based smart healthcare systems, the offloading of health data to the cloud for execution incurs excessive network latency [29], especially when medical data often have large sizes (e.g., audio, images). Moreover, the health data transfer also consumes much network bandwidth, which is likely to cause network congestion when the number of devices increases. The offloading process also requires transmit power of medical devices which in turn poses new challenges on battery and hardware designs on devices.

3.1.2 Benefits of FL in Smart Healthcare. Based on the innovative operational concept, FL is able to bring many attractive benefits to advance smart healthcare, as explained below:

- *Data Privacy Improvement:* In the FL-based smart healthcare system, only the local updates such as model gradients are required by the central server for the AI training, while the local health data are kept at local medical sites and devices. This would reduce the risks of the leakage of sensitive user information to the external third-party, and thus providing a higher degree of user privacy [30]. Following the increasingly stringent health data privacy protection legislation, the capability of preserving health user information of FL is important for building sustainable and safe smart healthcare systems[1].
- *Reasonable Trade-off between Accuracy and Utility:* Compared with conventional centralized learning, FL is able to offer a reasonable trade-off between accuracy and utility along with privacy enhancement. Moreover, FL training retains the model generalizability at the cost of nominal accuracy loss. In return, FL can enhance the scalability of the smart healthcare system thanks to its distributed learning feature.
- *Low-cost Health Data Training:* By avoiding the offloading of huge data volumes to the server, FL can help reduce significantly communication costs, e.g., latency and transmit power, consumed by raw data transmission, as the model gradients generally have much smaller sizes compared to their actual datasets [32]. As a result, FL also save much network bandwidth and mitigate possibility of network congestion in massive healthcare networks.

3.2 Requirements

To realize the full potential of FL in smart healthcare, several requirements should be met as highlighted below:

3.2.1 Trusted Server. One of the most important entities in FL is the central server that is used to aggregate local model gradients to build the global model in each communication round. Although the FL concept can provide privacy protection by allowing users to keep their data at local sites during the training, it has been proven that the model updates might still contain health user-related information such as data features and image resolution that can be re-constructed by the curious global server [33]. As a result, user privacy can be put at risks during the

[1] However, it should be noted that FL cannot fully address the privacy problem in smart healthcare [31]. Dedicated privacy protection mechanisms thus need to be designed to enhance FL in healthcare networks.

training, which also makes FL vulnerable and discourages medical sites from joining the collaborative training. Based on this fact, a fundamental requirement to ensure reliable FL operations in smart healthcare is to build a trusted server for the data training coordination and model aggregation. The computation services provided by the global server need to ensure a transparent and reliable model aggregation under the agreement between the service provider and healthcare organizations such as local hospitals. This is particularly necessary for the smart healthcare domain, where health data are highly sensitive and the data computation outside the data sources must be trusted to provide reliable FL-based healthcare. To further build trust for the server, recent research efforts have been put to develop new solutions, such as building decentralized and trusted servers enabled by blockchain [34] or providing secure aggregation methods, which will be detailed in the following section.

3.2.2 Reliable Client-Server Communications. Another important requirement for FL-based smart healthcare is the reliable communication between local clients and the global server. The exchange of local model updates to the server may be risky due to external threats [35]. Indeed, an adversary can deploy data attacks to the communication channels established by the clients and the server to steal the updated information which can interrupt the model computation at the global server due to the insufficient local updates to build a global model. Further, an attack can make attempt to modify or change the local updates which leads to aggregation bias at the server. All of these security concerns require the provision of safe and reliable client-server communications before deploying the FL functions in the smart healthcare system. This also builds up trust for health users in joining the FL process to solve collaborative health tasks such as federated medical image analysis.

3.2.3 Computational Capability for Local Training at Health Clients. In FL-based mobile smart healthcare where mobile medical devices participate in the federated data training, their computational capability is a key concern. In fact, to realize federated smart healthcare, one needs to join in the multiple communication rounds to achieve a desired training performance. In this regard, certain medical devices such as lightweight smart watches may be unable to join the training in the long run due to their limited computation ability and less energy resources [36]. Without the involvement of multiple devices in the training, the FL concept becomes inefficient in smart healthcare, where the contributed computation from different devices are highly important to improve health data training. Hence, how to build computation-accelerated hardware for health devices is essential to build FL-based smart healthcare ecosystems.

3.2.4 Available Dataset at Health Clients. To obtain the desirable training performance in FL-based smart healthcare, the availability of datasets at clients is needed. One device or medical site is required to construct its own dataset based on its working environment, e.g., a smartphone can collect human motion data of patients under its working area. Each participating client can also prepare its own data features based on its collected dataset via data-driven techniques such as feature extraction, for its own local training. In this context, one of the key concerns in dataset preparation is the non-independent and identically distributed (non-IID) issue which potentially makes the FL training highly divergent in the data training. Several solutions to cope with non-IID issues thus need to be developed, e.g., creating an additional subset of datasets to allocate fairly among clients [37], aiming to ensure efficient data training in FL-based smart healthcare.

For designing an efficient FL system, one needs to take its related operational metrics into account, such as the size of the deep learning models, the size of the local data, convergence time, and required accuracy. These performance metrics depend on the device, network conditions, and calculation speeds of both clients and the server. Such quantitative evaluations can be found in a recent work [38].

4 ADVANCED FL DESIGNS FOR SMART HEALTHCARE

This section presents the advanced FL designs for smart healthcare, including resource-aware FL, secure FL, privacy-enhanced FL, incentive-aware FL, and personalized FL.

4.1 Resource-aware FL

Collaborative FL models are built based on local model updates of IoMT devices in the uplink, communications with the aggregation server, and the global model broadcasting in the downlink. In this regard, resource management plays an important role in improving the performance of FL-enabled healthcare applications.

A scheduling problem is investigated in [39] to minimize the total training time by deciding a set of IoMT devices. However, finding the optimal solution to the non-integer scheduling problem with massive devices is very challenging. To overcome this challenge, along with one caused by unknown channel state information between IoMT devices and the aggregation server, the multi-armed bandit (MAB) theory is adopted to find the solution. The experimental results show that the MAB approach results in significantly low training loss compared with several benchmarks as well as outperformance in terms of the training latency. A similar scheduling problem using the MAB theory is studied in [40]. In particular, two scenarios are considered: 1) when all IoMT devices are available and the data distribution is IID, and 2) when IoMT devices are not always available and the data distribution is non-IID. This work concludes an interesting observation on the relationship between the number of IoMT devices and convergence rate. Given the importance of resource management, joint optimization of device scheduling and resource allocation are proposed in the literature. For example, the work in [41] proposes a hierarchical federated edge learning framework, in which the model aggregation is partially done by immediate hospitals. Under this framework, a joint resource allocation and device association problem is formulated and then solved by an iterative algorithm. Another joint optimization of device scheduling and resource allocation framework is investigated in [42]. In particular, a lower bound on the training time is derived and a greedy algorithm is designed to jointly optimize bandwidth allocation and device scheduling. In [43], the concept of distributed coordinate descent from the optimization perspective is leveraged to analyze the effects of scheduling policies on the convergence rate of FL algorithm. Such analyses are important in scenarios where the aggregation server needs to connect with massive IoMT devices. Three scheduling schemes are examined in [43], including random scheduling, round-robin, and proportional fair, showing that the convergence rate of these scheduling policies is largely dependent on the signal quality threshold.

A promising direction towards resource-aware FL for smart healthcare is to optimize joint computing and radio resources. The work in [44] to considers a joint computation and communication resource allocation framework for multiple FL services, which is motivated by the success of AI-enabled multiple services and applications deployed simultaneously at mobile devices and the network edge (e.g., virtual reality, video streaming, and healthcare apps). The multiple FL service problem is then formulated to minimize the total training overhead in terms of energy consumption and completion time [45]. Two algorithms are designed, including a centralized algorithm based on the block coordinate descent method and a decentralized algorithm based on the Jacobi-Proximal alternating direction method of multipliers (ADMM) approach. Multiple FL services are also investigated in a recent work [46]. Different from [44] where diffident FL services running on the same devices, [46] considers that each device runs only one FL service and multiple FL services are available at multiple IoMT devices in the network. A bandwidth allocation scheme is designed to allocate bandwidth resources among clients running the same FL service and among FL services. A distributed algorithm is firstly proposed to guarantee fairness among FL services, following by an auction-based game-theoretic approach to balance fairness and performance.

4.2 Secure FL

It is important to study secure solutions for FL-enabled healthcare applications since there are potential security attacks in FL systems such as poisoning attacks, inference, backdoor attacks, malicious server, communication bottlenecks, and free-riding attacks [47–49]. These security attacks can be caused by numerous sources, including communication protocols, data manipulation, and aggregation algorithm [50]. In the context of smart healthcare, a number of solution approaches have been investigated over the last few years. To prevent unreliable updates

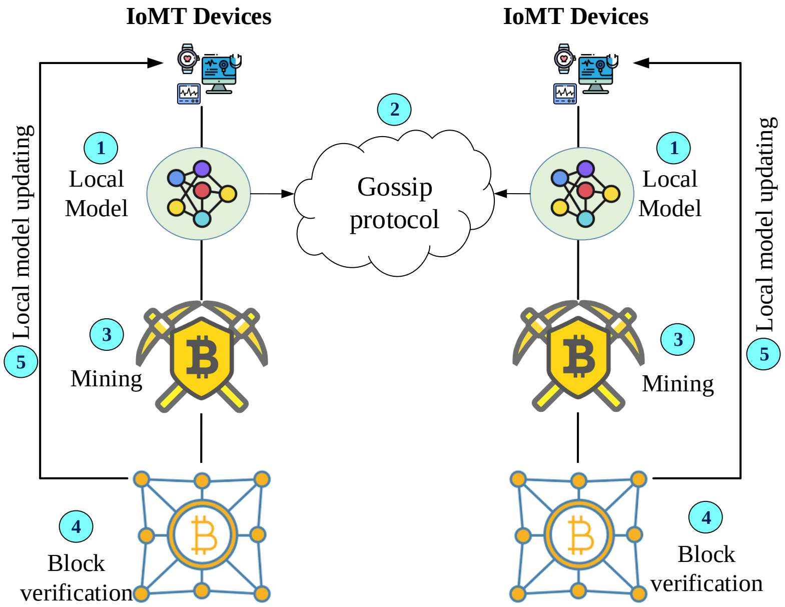


Fig. 3. Illustration of a single communication round of blockchain-enabled FL healthcare systems.

from untrusted devices, a new concept called reputation is introduced in the context of FL-enabled healthcare systems [51]. Such reliable device selection plays an important role in mitigating several security attacks. For example, a malicious IoMT device may inject poisonous data into its local data, and thus updating the fall local model may degrade the accuracy of the global model. Reliable device selection is also important in scenarios, where local FL models are trained using low-quality and noiseless data.

Decentralized FL is a promising solution for secure federated smart healthcare to address the problem of untrusted parameter servers in centralized FL. The implementations of decentralized FL are typically based on gossip, consensus, and diffusion approaches. For example, a decentralized FL scheme is proposed in [52] using a segmented gossip aggregation method to enhance the performance of federated data training. Each data client can act as a worker that stochastically selects a few neighbour workers to transmit the model segment in every training iteration, aiming to optimize the utilization of bandwidth capacities of all clients. The simulation results show its superior performance in terms of reduced training time in practical network topology and bandwidth settings with only slight accuracy degradation, compared to centralized FL. A decentralized FL design for smart healthcare is proposed in [53] over a graph. In this case, the FL algorithm allows local clients to perform local updates for several iterations and then enables peer-to-peer communications among health clients such as distributed hospitals. In such a way, the delay of parameter exchange and communications among clients can be reduced since there is no need to transfer the models to the remote central server.

However, designing reliable user selection in smart healthcare systems faces several challenges such as no universal metric for reliability evaluation, no universal selection scheme, and no real-time user monitoring methods. To address these challenges, blockchain is adopted in [51] to manage FL users' reputation. Particularly, the integration of blockchain into federated settings decentralizes the FL that allows for the elimination of a single central server in the model aggregation [54]. The blockchain can coordinate the calculation of global model via the block consensus among participants in a peer-to-peer manner. The use of blockchain to prevent untrustworthy server and external attacks in decentralized FL-enabled healthcare systems is also studied in [55].

Here, blockchain helps avoid the need for a centralized server (i.e., located at hospitals or clinics), whereas in each global communication round, IoMT devices compete with each other to mine a block and then append a new block to their local ledgers.

The illustration of a single communication round of blockchain-enabled FL healthcare systems is shown in Fig. 3. In sum, the procedure in a global round of such a blockchain FL system consists of five main steps as the following.

(1) *Local training*: Each IoMT device trains the local FL model using its local data.
(2) *Model broadcasting and verification*: Each device adds its digital signature to the model and broadcasts the model to the other IoMT devices via some gossip protocols. The transaction of a device is then verified by all other IoMT devices in the network.
(3) *Mining*: Upon receiving the local models from the other devices, each IoMT device tries to mind the current block.
(4) *Block validation*: the current block, if verified, is added to local ledgers of IoMT devices.
(5) *Local model updating*: Upon receiving the verified, each device updates its local model and starts a new communication round.

The upper bound of the loss function is also analyzed in [55], which is shown to be a function of the communication rounds and other factors such as the training time of each local round, mining time, learning rate, data distribution. and computation time. Based on that derived upper bound of the loss function, the impact of the number of lazy devices is further evaluated. More lazy IoMT devices decrease the learning performance, e.g., no lazy client can achieve an accuracy of 85.53% while the accuracy is only 78.80% when 30% clients are lazy. The application of blockchain and game theory for the design of a reliable FL algorithm is studied in [56]. Particularly, the contract theory is used to encourage the participation of highly reputed IoMT devices and a public blockchain approach is employed to manage the reputation updates of FL devices in a decentralized and secure manner.

An important problem in securing FL is secure aggregation; that is, the multiparty aggregation at the central server is not revealed by external attackers. According to [57], secure aggregation approaches can be classified into secure communication protocols, dining cryptographers (DC) based secure aggregation, homomorphic threshold encryption, and pairwise masking. It is also suggested that a secure aggregation method needs to satisfy the following conditions:

(1) It can work with high-dimensional updates from users,
(2) It is communication-efficient even with massive users,
(3) It is robust to the user participant and unavailability (also known as user dropping-out),
(4) It can provide security guarantees under unfavorable environments such as unauthenticated transmission channels and resource-constrained edge nodes.

A seminal work on secure FL aggregation is conducted in [57]. To address the challenge caused by user dropping-out, FL users are required to share a part of their Diffie-Hellman keys with all the other users, and thus the pairwise can be completely recovered if some FL users are no longer active. Such a problem typically happens in the context of healthcare applications; for example, some patients only visit the hospital for a very short period of time. A double-masking scheme is also leveraged to protect the user privacy, i.e., the server cannot know the user's masks. To provide more practical solutions, [57] also proposes two variants to balance the security guarantee level and communication efficiency. Recently, a secure aggregation method, namely Turbo-Aggregate, is proposed in [58]. More specifically, the Turbo-Aggregate method comprises three main components, including multi-group circular strategy, additive secret sharing, and Lagrange coding. The first component is to divide the entire set of FL users into multiple groups and each group needs to send the aggregated model from itself and other groups to the next group. The second component is to enhance the user privacy via additive secret sharing, and the third component is to increase the robustness against user unavailability. There are three main advantages of Turbo-Aggregate: 1) the aggregation overhead is significantly reduced from $O(N^2)$ to $O(N\log(N))$ with N being the number of IoMT devices joining int the FL process, 2) up to 50% dropping-out rate can be tolerated by Turbo-Aggregate, and 3) the total running time is multiple times faster than the benchmarks (i.e., up to 40 times reported by the experiments). However, such advantages are only evaluated via numerical results and thus experiments on the real health datasets can be examined by using the secure aggregation methods proposed so far. A recent promising solution for model aggregation is called over-the-air aggregation, which exploits the properties of wireless channels to wirelessly aggregate local updates [59]. The over the air aggregation can be extended to take into account the security and privacy aspects of healthcare applications.

4.3 Privacy-enhanced FL

Despite the great potential in improving user data privacy, FL also has its own privacy concerns such as membership attacks, unintentional data leakage, generative adversarial network based inference attacks [50]. For example, based on the local model updates from an IoMT device, the attacker can infer and recover the local data using the construction attacks [60] and/or predict the existence of a data sample in the local training dataset such as blood type, disease name, gender information, and any other private information, which can be done via membership inference attacks [61]. This motivates the development of advanced solutions of privacy-enhanced FL designs for smart healthcare applications.

Differential privacy is a well-known concept that can be adopted to enhance the user privacy in wireless and mobile network. Motivated by this advantage, a number of research works have been devoted to study differentially private FL systems, for general domains as well as smart healthcare services. The work in [62] first considers that artificial noise can be added to the local dataset of IoMT devices to protect the user privacy. Then, a multi-dimensional cost problem is studied to design an incentive mechanism. Three types of costs are taken into account to evaluate the contribution of each IoMT device, which are computation cost, communication cost, and privacy cost. Experimental results showed that the proposed three-dimensional contract-based incentive approach with differential privacy can obtain lower training loss and higher training accuracy compared with the seminal FL. There are several works focusing on applying the differential privacy concept to federated health services. For example, the work in [63] proposes a method to balance the differential privacy guarantee and deep learning accuracy. In particular, the differentially-private stochastic gradient descent method is employed at the distributed healthcare datasets and a secure aggregation step using momorphic encryption is leveraged. The proposed method is then tested in the DR dataset[2] to detect diabetic retinopathy, and SqueezeNet is used as the main deep architecture. Experimental results show that the performance of the proposed method is close the that of the centralized approach, while significantly outperforming several benchmarks such as FL with parallel differential privacy, where different hospitals apply differential privacy on their datasets individually and FL with centralized differential privacy, where a trusted entity is required to apply differential privacy for all the patients. Another differently private FL for healthcare applications is investigated in [64]. Two learning tasks are considered which include the prediction of adverse drug reaction and mortality rate. A real health dataset with over 1 million training samples is tested, which contains many personally private information such as diagnosis results, prescription fills, and admission records. A promising observation from the experiments is that the higher the privacy level is, the lower the training accuracy is. As a result, more research works should be investigated in the future to improve both learning performance and data privacy. However, such differently private FL works in [63] and [64] are promising as their FL frameworks can avoid the need for transmitting a large amount of data from distributed health organizations to a central entity and also project the patient data against potential privacy attacks.

[2] The dataset is available at https://www.kaggle.com/c/aptos2019-blindness-detection/data.

In order to further enhance the performance of FL-enabled healthcare services, joint optimization of privacy-preserving and resource management is of importance [65]. The authors in [66] discuss that even the raw data is not needed to be sent to the central server, exchanging model updates from massive IoMT devices is not bandwidth-efficient. Accordingly, a bandwidth-efficient FL scheme is proposed in [66] with privacy guarantee. More specifically, only the sign of updated values is sent to the aggregation server, which then scales the updates and does the aggregation to update the global model. The differential privacy concept is also used to protect every data sample of all the hospitals, but at the same time it ensures that the hospital cannot infer the global model broadcast by the aggregation server. An EHR dataset is used to examine the performance of the proposed scheme, namely FL-SIGN-DP, in terms of in-hospital mortality rate. The FL-SIGN-DP scheme is also compared with a number of existing alternative schemes such as centralized learning, FL-SIGN (i.e., no differential privacy), standard FL, and standard FL with differential privacy. The FL-SIGN-DP is shown to consume a very small amount of bandwidth, i.e., FL-SIGN-DP sends only 1.76 Mb while the standard FL sends up to 56.48 Mb, while FL-SIGN-DP can maintain a reasonable level of privacy protection.

4.4 Incentive-aware FL

Conventional FL approaches require all IoMT devices to share local model updates with the aggregation server; however, it is not always available in practice. It is since the IoMT devices are usually limited in terms of computing resources, radio bandwidth, and privacy concerns for personal data and server trustworthiness, thus having no willingness to share their models. To incentivize the participation of more FL users and improve the performance of FL-enabled healthcare scenarios, incentive-aware FL solutions are necessary. According to a recent survey [67], incentive mechanisms designed for FL can be categorized by different aspects, including the device's data contribution, device reputation, and resource allocation.

- **Data contribution** is evaluated via two important metrics: data quantity and data quality. While data quantity is typically assessed using the Shapely value, data quantity refers to the size of the local model updates and training samples.
- **Device Reputation** is an important metric in designing incentive algorithms for FL. In general, reputation reflects the extent that FL users can provide high-quality data for model training and reliable local updates.
- **Resource Allocation** is an important phase of any incentive scheme as computation and communication resources need to be appropriately allocated to FL users so as to improve the FL performance.

Game theory, originally derived from economics and business studies, is an excellent tool for designing incentive mechanisms in wireless and mobile network over the last decade. Motivated by this excellence, several game-theoretic incentive strategies have been investigated for FL-based healthcare systems. For example, the work in [32] leverages the Stackelberg game to incentivize user participation in the FL learning process. In such a game, each IoMT device first receives an offer from the aggregation server located at the hospital and then decides whether or not to participate in the FL learning process. After that, the aggregation server updates the incentive strategy so as to maximize its utility, which can be defined based on the number of global communication rounds and the target learning accuracy [32]. These steps of IoMT devices and the aggregation server are repeated until a desirable performance is obtained. Such a Stackelberg game is promising and can be much extended in the context of smart healthcare. For example, a dynamic Stackelberg game can be studied to consider time-varying computing resources and channel connections between the IoMT devices and the aggregation server located at the hospital [68]. A notable limitation of Stackelberg game approaches is the requirement of the full knowledge of FL devices' contributions. To overcome this limitation, the work in [69] develops a deep reinforcement learning (DRL) method to help the aggregation server to determine the award strategy and the devices to decide their local training schemes. The main advantage of the proposed DRL approach is that the historical payment strategies can be utilized to deal with future scenarios, e.g., more patients/relatives will be interested in joining the FL learning process to receive valuable clinical information from the hospitals and/or insurance companies.

In the context of cross-silo FL, where different hospital/insurance companies try to collaboratively build a global model built at the third-party aggregation server, motivating different hospitals to contribute their resources to the global model so as to maximize the social welfare is an open problem. The work in [70] studies this problem in order to answer two key questions: how does each health organization allocate its computing resources for local training and how to compensate the local model of each health organization by other organizations computing resources. To overcome the difficulty of solving nonconvex problems, a distributed algorithm is investigated by using the non-cooperative game. The proposed game is proved to be not only Nash-equilibrium but also has nice properties such as social welfare, individual rationality, and no need for a third-party payment entity. Recently, the work in [71] proposes an efficient method to evaluate the contribution of data contributors such as wearable computing devices, smart phones, and health handsets. The proposed method in [71] is to overcome a critical challenge of conventional incentive FL designs using the Shapley value that the computing cost increases exponentially with the number of IoMT devices and the dimensionality of data features. Exploited the success of DRL in long-term evaluation, the contribution of data contributors is efficiently assessed via a DRL-based approach, showing its superior performance to the conventional incentive mechanism.

4.5 Personalized FL

To provide personalized services to IoMT users, the conventional FL faces several challenges. A very first challenge is that the global model only captures the statistical characteristics of different IoMT devices but is hard to provide distinct personal styles. For example, when learning to predict the disease, the same body weight and height from different people may have different meanings due to personal living environments and various external factors. Another challenge is heterogeneous computing resources and network conditions of different FL devices. Motivated by the importance of personalized FL models, recent years have witnessed several advances in personalized FL designs for smart healthcare.

One of the very first works focusing on personalized FL for healthcare services is developed in [72]. This work firstly points out critical challenges related to personalized FL for healthcare services. The first challenge is that different people naturally have different physical characteristics, the kids' body temperature is typically higher than that of adults and adults have very few falling steps when walking and doing exercise, thus fewer falling samples in the training dataset. Another challenge is that deploying a single global model may not be suitable for a specific person. To overcome these challenges, [72] proposes an edge-cloud FL architecture, namely FedHome, for in-home healthcare services, where each local model is trained at home located at the network edge and the cloud is responsible for global model aggregation, as shown in Fig. 4. The training at the cloud server mainly relies on the distributed datasets, which may be different among different homes. The shared model only captures the common features of all health users, and thus it may perform poorly on a particular user. As a result, to achieve personalized in-home health monitoring, each user integrates the trained global model with his personal health data. To capture the personalized services, in-home IoMT devices download the global model from the cloud and trains its personalized model. The global model from the cloud is also used to generate a balanced dataset for all the devices in the network, thus decreasing the discrepancy between personalized models and the global model. Tested on the human activity recognition dataset, FedHome with the generative convolutional autoencoder at both the cloud and edges, can outperform several benchmarks in terms of training accuracy and model size. For example, FedHome can achieve the accuracy of 95.87% while the personalized FL with multi-layer perceptron (MLP) and convolutional neural network (CNN) can only achieve 92.31% and 91.77%, respectively. The potential of this FedHome model can be extended to other IoT applications such as personalized human activity in hospital settings and personalized driving in smart transportation [73]. However, this personalized FL model

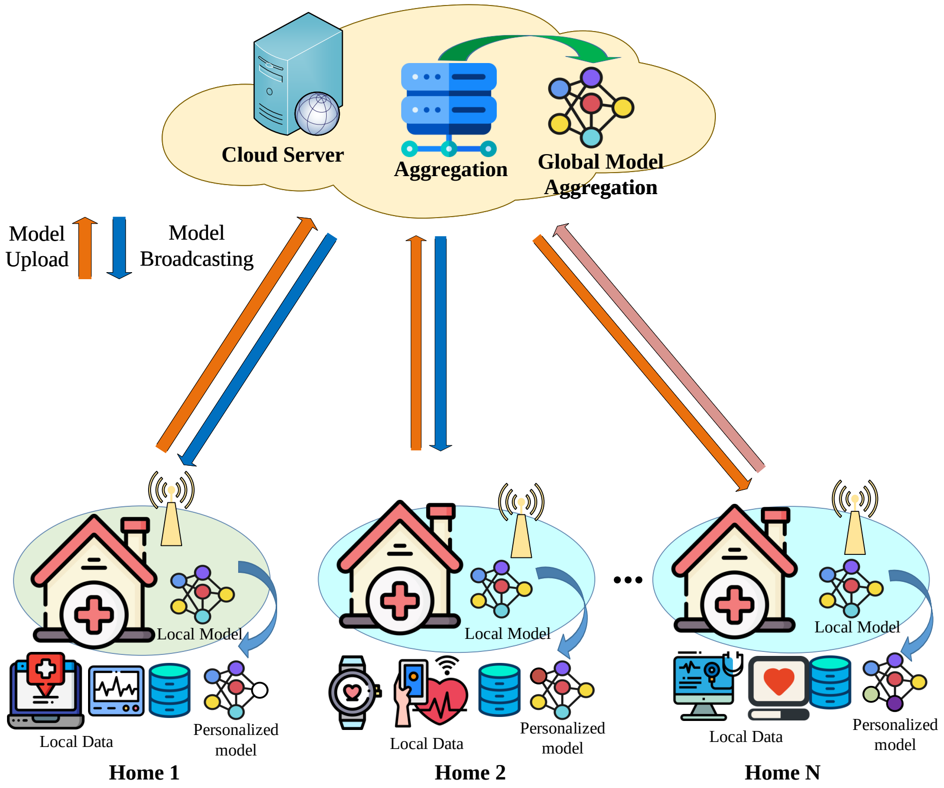


Fig. 4. An edge-cloud architecture for personalized FL enabled in-home health monitoring services [72].

also has limitations in terms of the imbalance in data distributions between the local data and the global data population. The reason for this is that users' personalized data is usually insufficient. The work in [74] identifies a new challenge in FL design, which is called label heterogeneity. Concretely, each FL device has its own definition of data labels and learning model, which is independent of the definitions in other devices and the central server. However, such a new challenge typically exists in health and activity tracking scenarios. For example, the blood type is labeled as bloodtype in the device A; however, which is labeled as bltype in the device B. This new challenge is solved in [74] by devising an α-weighted update. Particularly, the overlapping information of labels of different devices is aggregated at the central server. The proposed approach is examined on the Animals-10 dataset [75], showing that the classification accuracy is increased by about 16.7%. This approach is also tested on the human activity recognition dataset with an average increase of 9.153% 11.01% in [76] and [77], respectively. Recently, the work in [78] develops a personalized FL framework to predict the pain by using face images. Each device employs a lightweight CNN model and keeps the learning parameters and updates of the last layer locally. This is to 1) personalize the local pain estimation model and 2) increase the protection of data privacy against potential adversaries. The proposed model is evaluated by using the UNBC-McMaster Shoulder Pain dataset [79], showing that the proposed personalized FL with privacy preserving outperforms the conventional FL approach, which shares all the learning updates with the aggregation server, which in turn increases the probability of data breach.

Table 2. Summary of recent advances in FL designs for smart healthcare.

Theme	Ref.	FL type	FL clients	FL aggregator	Key contributions	Limitations / Potential directions
Resource Allocation	[43]	HFL	IoMT devices	Cloud server	Impact of scheduling policies on the convergence rate.	Only three base scheduling policies are considered.
	[44]	HFL	Smart devices	Edge server	A resource allocation for multiple FL services.	The proposed algorithms are not evaluated on the health dataset.
	[46]	HFL	Smart devices	Edge server	Each device runs only one FL service and bandwidth allocation schemes for intra- and inter-service is proposed.	The proposed algorithms are not evaluated on the health dataset.
Secure FL	[55]	HFL /VFL	Smart devices	Serverless	Bockchain for decentralized FL updates and the impact of lazy nodes are studied.	Blockchain may increase the computing requirements of smart devices.
	[57]	General FL	FL users	Aggregation server	A seminal approach of secure aggregation for FL.	The proposed method is examined on the non-health dataset.
	[58]	General FL	FL users	Aggregation server	An efficient secure aggregation method namely Turbo-Aggregate.	Only numerical results are reported.

4.6 Lessons Learned

In this section, we have reviewed the recent advances in FL designs for smart healthcare, stemming from five main themes: resource-aware FL, secure FL, privacy-enhanced FL, incentive-aware FL, and personalized FL. Some promising lessons learned are summarized in the following.

- **Training datasets**: Various designs have been investigated for general FL scenarios that may be applicable for smart healthcare services. However, it is observed that almost all the methods proposed in the literature are evaluated using non-healthcare datasets such as MNIST and Fashion MNIST. Evaluating FL designs using related healthcare datasets and from the healthcare perspective is necessary to better improve prospective FL-enabled healthcare services.
- **Performance tradeoff**: Numerous performance metrics have been used in the literature to design FL schemes for healthcare services. However, there is a tradeoff between metrics such as privacy and learning accuracy as experimented in [64]. This result suggests that the FL framework should be designed for specific healthcare applications. This also necessitates efficient FL schemes to take into account multiple learning objectives for healthcare applications.
- **Communication efficiency**: Exchanging extremely large health data among health organizations and between the distributed organization and the central site is very resource-consuming and may not be

Table 3. Summary of recent advances in FL designs for smart healthcare (continued).

Theme	Ref.	FL type	FL clients	FL aggregator	Key contributions	Limitations / Potential directions
Privacy-enhanced FL	[62]	HFL	IoMT devices	Cloud server	A privacy-preserving FL approach with multi-dimensional cost.	The propose approach is evaluated via numerical simulations.
	[63]	HFL	hospitals	Cloud server	A method is proposed to balance the differential privacy guarantee and deep learning accuracy.	Advanced encryption is needed for privacy enhancement.
	[64]	HFL	Smart phones	FL server	A study on the tradeoff between privacy protection and training accuracy.	The convergence of FL algorithms has not been verified.
	[66]	HFL	Hospital	Central server	A privacy-preserving resource allocation, namely FL-SIGN-DP, is developed.	AI techniques such as DRL can be adopted at the server to exploit the historical samples.
Incentive-aware FL	[68]	HFL	IoMT devices	FL server	A Stackelberg game is proposed to stimulate the participant of IoMT devices.	Results are reported using numerical results only.
	[70]	Cross-silo FL	Hospitals	Data centre	A non-cooperative game approach is used to devise an incentive mechanism for resource allocation.	The model is developed based on the assumption of IID data among hospitals.
	[71]	HFL	Smart device	FL server	A DRL-based approach to evaluate the contribution of data owners.	A reward model can be developed to encourage the participant of more FL users.
Personalized FL	[72]	HFL	In-home sensing devices	Cloud server	A edge-cloud FL architecture is proposed for in-home healthcare services.	The convergence of the proposed FL algorithm has not been reported.
	[74]	HFL	IoMT devices	FL server	An α-weighted update is proposed to mitigate the effect of label heterogeneity.	The convergence of FL algorithms has not been verified.
	[78]	HFL	Smart devices	FL server	A personalized FL approach is developed to predict the pain via face images.	The challenge caused by data imbalance should be further investigated.

economics-practical. Therefore, it is necessary to develop bandwidth-efficient and compressed FL methods for healthcare scenarios. For example, a sparse ternary compression framework is proposed in [80] to compress both uplink and downlink transmissions between the aggregation server and end FL participants. Such a compression method can be extended to healthcare scenarios, especially when the size of health data (e.g., X-ray, diagnosing results, and EMR photos) and model updates is extremely large.

- **Multi and personalized FL services**: Multi FL services are promising as more data will be generated at the network edge and each IoMT device can collect various types of data from various data sources. Personalized FL services will also play an important role in enhancing healthcare services as different patients have different health profiles and purposes. Joint healthcare services with general FL services are promising and should be further investigated in the future.

We summarize this section in the taxonomy Table 2 and Table 3 along with the key technical aspects of each reference work to provide more insights into the recently advanced FL designs for smart healthcare.

5 FL APPLICATIONS IN HEALTHCARE

In this section, we explore the emerging applications of FL in smart healthcare, namely federated EHRs management, federated remote health monitoring, federated medical imaging, and federated COVID-19 detection and diagnosis.

5.1 FL for Federated EHRs Management

In the past few years, AI/DL technologies have been widely used in the healthcare sector to gain insights into health issues and disease progression by learning digital medical information extracted from EHRs for facilitating diagnosis and assessments as well as promoting medical research activities. One of the challenges faced in such traditional AI techniques is privacy leakage during data analytics. Indeed, compared to other domains, EHRs data in healthcare systems are highly sensitive and private. The removal of metadata such as patient information is insufficient to preserve the privacy of patients, especially in the complex healthcare settings where multiple parties such as hospitals and insurance companies have access to the common healthcare database as part of their employment requirements, including data analysis and processing.

In fact, FL is able to provide much more reliable solutions for intelligent data analytics in EHRs management by exploiting AI functions for supporting healthcare services while adequately preserving user privacy based on the cooperation of multiple entities, e.g., patients and healthcare providers. For instance, a privacy-aware and resource-saving collaborative learning protocol based on FL is introduced in [65] for an EHRs analytic management system with the cooperation of multiple hospital institutions and a cloud server. Here, each hospital runs an NN using its own EHRs with the help of cloud computing. To ensure privacy for model parameters in the FL process, a lightweight data perturbation method is considered to perturb the training-related data, which thus can defend model memorization attack in the learning. Although attackers can obtain perturbed information of EHRs, it is hard to obtain or recover the original data. Furthermore, an FL-based approach is proposed in [81] to predict hospitalizations for patients diagnosed with heart diseases using their historic EHRs. Specifically, health data from an EHR system consisting of patients' smartphones and distributed hospitals are trained locally with respect to demographic information such as age, gender, and physical characteristics. Afterwards, the trained model parameters are aggregated by the cloud server for building a unified prediction model based on a global support vector machine (SVM) classifier. This aims to predict the future hospitalizations of patients due to heart diseases for hospital resource management (e.g., future estimation of treatment facilities) without disclosing the private dataset.

To improve the privacy of FL-based healthcare, the work in [64] employs FL to model the distributed EHR data learning among different hospital sites. A differential privacy-based solution is adopted in [82], where each local hospital adds an artificial noise to mask the local updates before offloading them to the server for sensitive information protection. Three ML models are employed, including perceptron, support vector machine (SVM), and logistic regression, to perform FL training, showing that a competitive accuracy performance is achieved by using differential privacy-based FL methods compared to traditional FL schemes. A federated HER management solution is also considered in [83], where 58 different hospitals are employed to collaborate the prediction of patient mortality without moving health data out of their silos. A new FL-based method called SplitNN is studied in [84] by implementing the vanilla FL configuration associated with split learning. Conceptually, in split learning, a deep neural network is divided into multiple parts, each of which is trained on a different client. The trained data may reside at one client or at multiple clients, but no client can retrieve other clients' data. By splitting the network into multiple sections that are then transmitted to distributed clients, the network training can be implemented by offloading the weights of the split layer, while raw data sharing is not required. For the case study in [84], each radiology client center trains a partial deep network and its neural parameters will be transmitted to a central server for model aggregation without sharing user information. To further improve the EHRs training performances in FL-smart healthcare, a statistic channel-based FL solution is considered in [85], where a small fraction of local gradients can be uploaded stochastically to the server from participating hospitals by choosing the channels with the most substantial variation. In other words, only the channel of neurons with a significant change in the training loop is used in the model aggregation, while ignoring neural channels with little change and less feature representation. A set of hospitals is adopted in simulations, comprising 30760 admissions with status information represented by alive or expired. A binary classification problem is formulated for mortality prediction using NNs in the FL simulation, showing an improvement in training convergence rates compared to traditional FL schemes. Moreover, to support the federated mortality prediction, patient clustering is adopted in [86], where patient subgroups capture similar diagnoses and geographical locations to train a neural network model to forecast mortality based on drug features. Another work in [87] focuses on complexity reduction in FL settings for EHR mortality prediction, by applying an adaptive boosting method named LoAdaBoost for increasing the efficiency of federated machine learning in both IID and non-IID data distribution scenarios. The study in [88] considers EHRs training in the FL-based healthcare with recurrent NNs (RNNs) for predicting preterm-birth 3 months using structured datasets with uncertainty training. 42 hospitals are employed in the test, and each hospital performs the update of model generalizations as the local weights in the FL process. Other works in [89], [90] also implement privacy-aware EHRs management solutions using FL for safe and reliable hospital networks in healthcare. A network of 20 hospitals is created in a case study in [90] for building a federated EHRs management solution to improve the efficiency of health data analytics by the federated process achieved by the collaboration of medical centres, healthcare providers, and patients. Particularly, to provide security for data training, FL can be combined with blockchain in the model update and data communications. As an FL example in Fig. 5, a number of hospitals cooperatively communicate each other via the blockchain to run DL algorithms by using local EHRs data. At each hospital, a DNN can be adopted for EHRs feature extraction and local modelling, where FL provides guidance for transmissions of model updates to perform the model aggregation at a common hospital. In this context, blockchain is used for secure model communication by providing an immutable data ledger managed by the hospitals and data centre [91]. The model update and transmission logs are recorded in the blockchain so that all entities can monitor the FL process in a transparent manner.

5.2 FL for Federated Remote Health Monitoring

Recently, there is an increasing demand for developing smart solutions for administering remote healthcare from hospital-centric to home-centric. FL can be used for facilitating in-home health monitoring, by training a global model from distributed homes under the control of a server, e.g., a cloud server, while preventing data leakage by keeping user data locally [72]. In this regard, the IoMT device (e.g., a user's mobile phone) at each home can

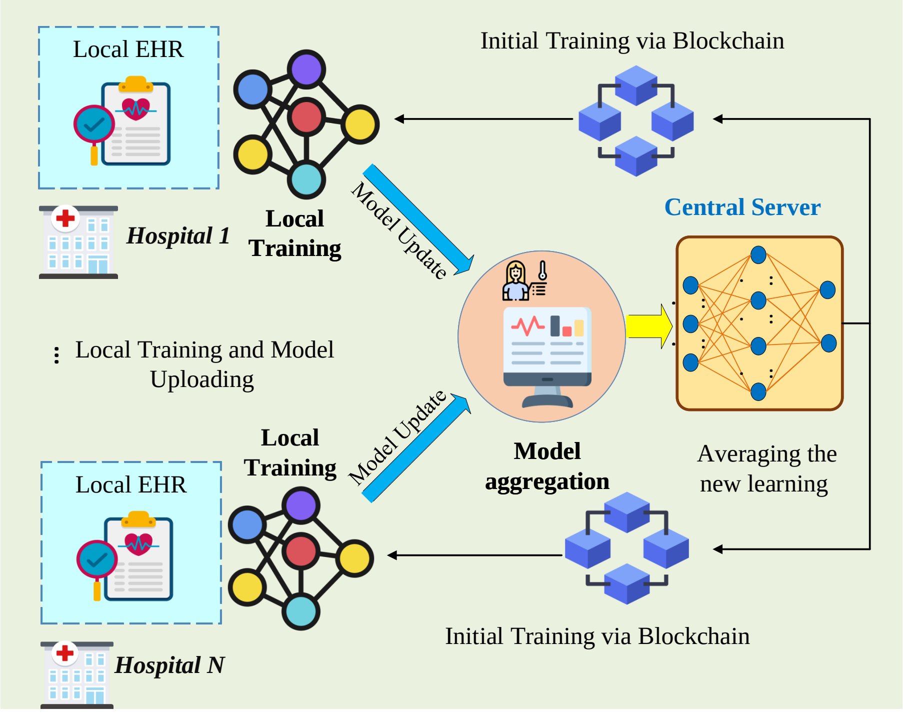

Local EHR
Hospital 1
Local Training
Initial Training via Blockchain
Model Update
Central Server
Local Training and Model Uploading
Model aggregation
Averaging the new learning
Local Training
Model Update
Local EHR
Initial Training via Blockchain
Hospital N

Fig. 5. Federated learning for collaborative EHRs analytics with blockchain.

learn a personalized model using convolutional neural networks (CNNs) by synthesizing a class-balanced dataset with its personal data and adjusting model gradients with the updated dataset in a fashion that the cloud and all homes update simultaneously. This not only addresses effectively imbalanced and non-IID data issues but also improves the personalized predictions. Extensive experiments are implemented using a realistic human activity dataset, showing that the FL-based approach can achieve a high accuracy of 95.41%, an increase by 7.49% over the standalone CNN scheme with low communication costs, in both the balanced and imbalanced data cases.

Additionally, an FLT scheme is proposed in [92] for wearable health monitoring where smart phones collaborate to train a shared CNN model with a cloud server for human activity recognition with privacy protection awareness. Since there is a large distribution divergence between models in the cloud and smartphones, transfer learning is utilized to make the training model more tailored which helps achieve personalization. Implementation results indicate the higher accuracy of the FL algorithm (5.3% enhancement) for wearable activity recognition compared to traditional methods. The potential of this FL scheme can be expanded significantly in other interesting healthcare and medical applications, e.g., health monitoring, fall prediction, and disease diagnosis. Another remote healthcare monitoring system based on FL is designed in [93] for obesity and comorbidity phenotyping control. To be clear, a two-stage federated natural language processing approach is proposed that allows the collaborative health data training using clinical notes from different hospitals or clinics without the need for sharing data. First, a patient representation model is built at each hospital for training a NN to predict the current procedural terminology code from the text of the notes. Then, a phenotyping ML model is built to perform collaborative training across multiple sites for the target phenotype for disease classification from three classes, namely presence, absence or questionable. Simulations for an FL setting with 10 hospital sites show a better precision and recall performance compared to non-FL approaches.

Moreover, a mobile activity monitoring approach based on FL is considered in [94] for supporting healthcare applications such as assisted living and fall detection. Here, a real-world heterogeneity human activity recognition dataset is employed which is distributed across mobile devices to perform federated NN training with four types of human activities, including sitting, walking, standing, and stairs up. A non-IID data environment is also established, where the activity data are split with disparities in both the predefined labels with these activity types and distributions in data. A disease prediction method is studied in [95] by taking the advantage of FL using a large nationwide health insurance dataset distributed over 99 medical sites (e.g., hospitals and clinical labs) from 34 states in the US. The data include EHRs of diabetes, psychological disorders, and ischemic heart diseases. By comparing with conventional approaches such as centralized learning, local training without federation, the FL method achieves a competitive performance in terms of high accuracy rates and privacy protection. A new FL scheme called FedMood is suggested in [96] for mood prediction and monitoring. The keyboard keystrokes such as the interval between two keystrokes are exploited for biometric identification that helps to predict depression through the analysis of keystroke habits of patients with depression. This is inherited from the fact that the depression patients often have different typing speeds compared to normal people. Particularly, mobile phones are employed to collect necessary information, e.g., key letters, special characters, and phone accelerometer values that can be trained via DNN and coordinated by a data server for aggregation. A holistic experiment is conducted for both IID and non-IID data cases, indicating a good mood estimation accuracy (above 85%) in comparison with local training and collaborative data sharing schemes. The work in [97] also builds an FL-based health monitoring solution for analyzing the treatment effect of patients from the distributed hospital network. Interestingly, an entity called personalized treatment effect estimator is created in each hospital. Each estimator can be classified in each subgroup, where personal treatment effect includes the outcomes on patient characteristics and site indicator is used to estimate the global treatment effect at the coordinating site.

5.3 FL for Federated Medical Imaging

Due to the privacy concerns, it is challenging to implement AI-based medical data imaging by fusing different medical institutions at a centralized entity. FL has emerged as a promising solution for supporting large-scale medical imaging tasks, by allowing to learn from multi-source datasets without the need for public data sharing. An FL approach is recently proposed in [98] with a focus on solving the variations among clients (e.g., hospitals) by transforming the images of all clients onto a common image space via the federated process with a cloud-based global classifier. This enables to build a multi-source diffusion-coefficient image dataset for supporting automated classification. A generative adversarial network (GAN) is employed at each institution that can generate synthetic image datasets and translate its raw images to the target image space, which addresses cross-client variation problems with privacy preservation. Through the simulations on prostate cancer-related images, the proposed FL scheme can yield an accuracy score of 0.9722, achieving a performance enhancement by 0.13% in comparison to non-FL schemes.

Moreover, an FL model for federated brain imaging is also suggested in [99], aiming to support brain tumour segmentation using deep neural networks (DNNs). Here, each federated client (e.g., MRI scan machine) has a fixed local dataset and reasonable computational resources to train the DNN structure and share the weight updates to the federated server for aggregation via a model averaging technique. Although FL can protect user privacy leakage, it is still vulnerable to misuse risks such as training sample reconstructions at the server. Then, a differential privacy technique is adopted to add noise to each node's training process, distort the updates and mitigate the exposure of information during the model exchange. Simulations are conducted using a brain tumour dataset containing multiparametric pre-operative MRI scans of 285 subjects, indicating a similar segmentation performance with the ideal centralized scheme, while a degree of privacy protection is achieved. Another federated brain imaging approach is suggested in [100] to take advantage of the magnetic resonance images (MRI) scans distributed across multiple clinical centres and institutions. In this case, an FL model is derived to simulate an end-to-end framework for data standardization, confounding factors correction, and measurement of variability of high-dimensional features, by the collaboration of medical sites and the central server. The MRI datasets along with covariates (e.g., age, sex information) are employed from 455 controls, 181 with non-progressive mild cognitive impairment (MCI), 208 progressive MCIc, 234 Alzheimer's disease, and 232 with Parkinson's disease. An approach using alternating direction methods of multipliers (ADMM) is also integrated for reducing the amount of iterations, aiming to mitigate training latency in the federated setting. In a similar direction, the authors in [101] also consider a federated multiâ€‘institutional collaborations for brain tissue-based MRIs analytics. A set of 10 medical institutions is involved where each entity runs an NN model to detect radiographically abnormal regions of each brain scanning image.

To facilitate X-Ray scan imaging in smart healthcare, an FL-based methodology is proposed in [102] for supporting diagnosis of acute neurological symptoms such as severe headache or loss of consciousness. Each hospital run a CNN-based DenseNet1212 model that can support feature propagation, encourage feature reuse, and minimize the number of neural parameters to train the X-ray image dataset provided by the Radiological Society of North America. To enhance privacy for FL-based medical imaging, differential privacy can be adopted by a Dopamine method [103]. It is assumed that patients only trust their local hospital, and hospitals are non-malicious and non-colluding, while the server might be honest but curious. To this end, the hospitals need to provide a privacy levels to patients in each local update round in a fashion that the server can build a desired global model with high accuracy rates while data leakage is minimized. Each hospital adds a Gaussian noise of variance to calculate the privacy loss in a way that the effect of each update on the local momentum is the same as its effect on the aggregated model to strike the balance between the privacy cost and accuracy loss. A CNN-based SqueezeNet model [104] is adopted to implement simulations, showing an increase of classification accuracy up to about 80% on the real-world medical image dataset. A method called FL-based Magnetic Resonance Imaging Reconstruction (FL-MR) is proposed in [105] for multi-institutional collaborations for MRI reconstruction. Here, the learned intermediate latent features from different hospitals are aligned with the distribution of the latent features at the target site. This is enabled by the collaborative training of local reconstruction networks at local sites and an adversarial domain identifier which aims to align the latent space distribution in the target domain. Moreover, a weakly supervised multiple instance learning approach is studied in [106] based on the FL concept for gigapixel whole slide images in computational pathology. Here, in each hospital silo, the tissue regions are automatically segmented to extract the image patches, which are then embedded into a low-dimension feature representation using a CNN. Hen, the hospital performs training of a weakly-supervised learning model using its own whole slide image dataset with slide-level and patient-level labels as data features, before sending the trained gradients to the server for averaging. To demonstrate the feasibility of the proposed FL scheme, a weakly-supervised classification task is taken in histopathology for diagnosis of two separate disease datasets, i.e., renal cell carcinoma and breast invasive carcinoma. Simulations show that the FL approach can achieve the balanced accuracy of disease detection up to 90% in various parameter settings, showing the potential of FL in decreasing the barriers to cross-institutional collaborations for facilitating pathology computation. Another work in [107] proposes an FL method for functional magnetic resonance imaging (fMRI) analysis, where a decentralized iterative optimization algorithm is designed with a randomization mechanism to coordinate the weight sharing process. A domain adaptation method is developed for cross-silo learning with labelled and unlabelled or semi-labelled images or text datasets. Simulation results show that the proposed model can boost neuroimage analysis performances and find reliable disease-related biomarkers by using multi-site data without data sharing. The real-world implementation of FL for medical imaging is also presented in [108] for breast density classification, showing that the FL scheme can perform 6.3% on average better than standalone training approaches.

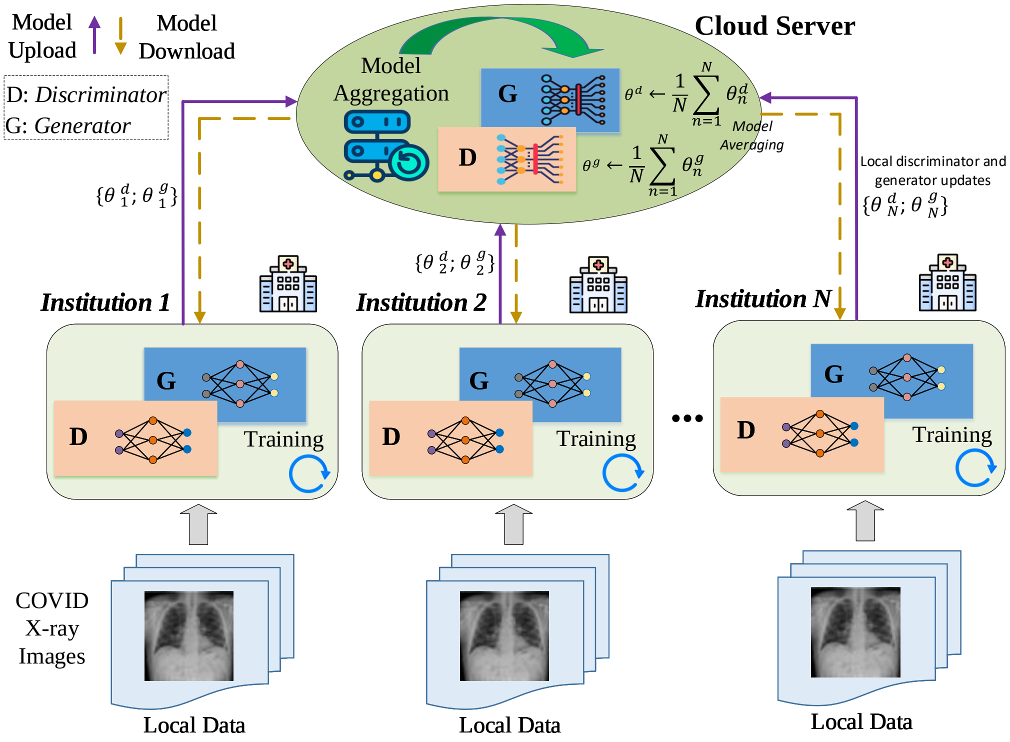

Model Upload
Model Download
D: Discriminator
G: Generator
Cloud Server
Model Aggregation
G
D
$\theta^d \leftarrow \frac{1}{N}\sum_{n=1}^{N}\theta_n^d$
Model Averaging
$\theta^g \leftarrow \frac{1}{N}\sum_{n=1}^{N}\theta_n^g$
Local discriminator and generator updates
$\{\theta_N^d; \theta_N^g\}$
$\{\theta_1^d; \theta_1^g\}$
$\{\theta_2^d; \theta_2^g\}$
Institution 1
Institution 2
Institution N
G
D
Training
•••
COVID X-ray Images
Local Data
Local Data
Local Data

Fig. 6. An FL framework with GANs for COVID-19 detection.

5.4 FL for Federated COVID-19 Detection and Diagnosis

Recently, COVID-19 has spread rapidly across the globe and become a major health concern of many countries [109, 110]. Many AI-based approaches have demonstrated significant promise for the early detection of COVID-19. DL-based techniques such as convolutional neural networks (CNNs) [110] have been widely used to identify COVID-19 cases by extracting essential features from chest X-rays. However, in the pandemic, collecting sufficient data to implement intelligent algorithms becomes more challenging and the user privacy concerns are growing due to the public data sharing with datacentres for COVID-19 data analytics [111]. FL can be an ideal candidate for assisting COVID-19 detection, by its federation and privacy protection features. In this way, each institution participates in training their DL model using their local COVID-19 images, e.g., X-ray and computed tomography (CT) images, and only model parameters such as gradients are exchanged while there is no need to share actual data and sensitive user information. For example, a number of hospitals can cooperatively communicate via the blockchain to run CNN updates locally for identifying CT scans of COVID- 19 patients [112]. At each hospital, a deep capsule network is developed to enhance image classification performance, while FL provides guidance for transmissions of model updates to perform the model aggregation at a data centre. Simulations from 34,006 CT scan slices (images) of 89 subjects verify a high COVID-19 X-ray image classification and low data loss in the FL algorithm running.

In a similar direction, FL is also used in [113] to provide privacy-promoting AI solutions for COVID-19 chest X-ray image analytics. Some preliminary experiments have been implemented, where multiple COVID-19 X-ray image owners run a CNN-based model for image classification, and then share the computed parameters with a data centre for mobile averaging while the data ownership of each user is guaranteed. Four state-of-the-art CNN models, e.g., MobileNet, ResNet18, MoblieNet and COVID-Net, are used in the federated setting for evaluation, where ResNet18 is proven with the best COVID-19 detection performance (98.06%) in federated X-ray image learning settings. For instance, an FL framework for COVID-19 detection is illustrated in Fig 6. Each institution runs a local GAN consisting of a discriminator and a generator based on CNNs to learn the COVID-19 data distribution using its own local image dataset. Then, the local GANs synchronize and exchange the learned model parameters for aggregation at a cloud server, which then returns a new version of a global model to all institutions for the next training round. This process is repeated until a desired accuracy is achieved, aiming to generate realistic COVID-19 images for the detection of COVID-19.

Moreover, a dynamic fusion-based FL method is proposed in [114] for CT scan image analysis to diagnose COVID-19 infections via two stages, including client participation and client selection. First, each client such as a medical institution make a decision to participate in the FL round based on the performance of the newly trained model. The central server also makes the decision to select which clients are permitted to update their local gradients by calculating the updating time. If a client cannot update its gradient within a predefined time interval, it is excluded from the FL aggregation. Another FL approach is designed in [115] for COVID-19 screening from Chest X-ray images by the collaboration of multiple medical institutions. The focus is on Chest X-ray images classification to identify COVID-19 from non-COVID-19 cases, where the feature extraction and the classification of X-ray images are performed based on a CNN to detect the COVID-19 disease. At each hospital, a CNN model is trained by allowing each X-ray image to be put into a convolutional layer and output the probability of COVID-19 infection, and then a central server is used to aggregate synchronously with the local institutions for building a strong classification model for COVID-19 detection without compromising significantly user privacy which is valuable in the pandemic [116]. Moreover, FL is also combined with DL to build a deep collaborative learning solution for detecting COVID-19 lung abnormalities in CT [117]. The internal datasets are collected from a total of 75 patients confirmed COVID-19 infection at three local hospitals in Hong Kong for FL simulations, and then the generalizability is validated on external cohorts from Mainland China and Germany.

5.5 Lessons Learned

The main lessons acquired from the review of the use of FL in smart healthcare applications are highlighted in the following.

- FL can be useful to build privacy-aware and resource-saving collaborative learning protocols for EHRs analytic management systems with the cooperation of multiple hospital institutions and a data centre [65]. Interestingly, the use of FL opens new opportunities for federated EHRs analytics among distributed medical institutions despite the strict privacy regulation due to its learning nature by only allowing model parameters to be exchanged while raw data are kept at local sites. From [64], [83], [88], we can find that FL is a very useful learning approach to accelerate the accuracy rates of AI model training thanks to the use of distributed data resources and computation capabilities of multiple silos.
- FL is also useful for facilitating in-home health monitoring, by training a global model from distributed homes under the control of a data server, while preventing data leakage by keeping user data locally [92], [93]. For example, FL can be employed to enable mobile activity monitoring [94] for supporting healthcare applications such as assisted living and fall detection. Compared to conventional approaches such as centralized learning, local training without federation, we find that the FL method achieves a competitive performance in terms of high accuracy rates and privacy protection in smart health monitoring applications [97].
- Moreover, we also realize that FL plays a significant role in supporting medical imaging applications, by fusing different medical institutions at a centralized entity via the federated data training process [98].

Particularly, FL is able to be deployed on medical devices such as MRI scan machine [99] which has enough local dataset and reasonable computational resources to compute AI updates to join the FL process with the cloud server. Recently, the roles of FL in medical imaging is also investigated in X-Ray scan imaging [102], which has the great potential for supporting diagnosis of acute neurological symptoms such as severe headache or loss of consciousness.

- In the COVID-19 pandemic with growing privacy concerns, FL is particularly useful to support COVID-19 diagnosis and detection, by coordinating massive hospitals to build a common AI model [112]. For example, we find that FL can be employed for COVID-19 screening from Chest X-ray images by the collaboration of multiple medical institutions. The feature extraction and the classification of X-ray images can be performed via collaborating among hospitals to detect the COVID-19 disease.

In summary, we list the emerging applications of FL in smart healthcare in the taxonomy Table 4 and Table 5 along with the key technical aspects of each reference work to provide more insights into the integrated FL-healthcare use cases.

6 REAL-WORLD PROJECTS OF FL IMPLEMENTATION IN SMART HEALTHCARE

In this section, we highlight several representative real-world projects of FL implementation in smart healthcare applications.

6.1 FL for Medical Imaging, US

The feasibility of FL in medical imaging has been investigated via a real-world experiment conducted at the University of Pennsylvania and 19 other institutions worldwide in the collaborative healthcare project [119]. Particularly, the Intel company has provided support to the FL-health project by leveraging the capabilities of Intel Xeon Scalable processors and Intel Software Guard Extensions (Intel SGX) for running FL functions at hospitals and the cloud server. This company also supports advanced DL algorithms along with strong hardware tools for accelerating the training, which helps building generalizable and state-of-the-art FL-healthcare models while improving the protection of sensitive patient data. Several preliminary real-world experiments has been conduced at the centre of Biomedical Image Computing and Analytics of the University of Pennsylvania, showing that the FL-based approach can achieve an training accuracy rate of up to 90% on image datasets, a competitive performance compared with the centralized scheme.

6.2 FL for Healthcare Collaboration, UK

As an effort to promote healthcare collaboration in the UK, the Nvidia corporation has recently joined the healthcare project with King's College London and Owkin to establish an FL platform for the National Health Service [120]. The Owkin Connect platform running on NVIDIA Clara enables AI algorithms to be trained at local hospitals in the UK under the management of a central server. Particularly, blockchain has been integrated to provide traceability and monitoring of all healthcare data used for model training and data communications. The project is initially connecting four of London's premier teaching hospitals, providing federated AI services to support important medical domains such as cancer, heart failure and neurodegenerative disease. The project is expected to be deployed at least 12 hospitals around the UK in 2021. Also, the NVIDIA company is developing on-device AI platforms for deploying FL functions on smart devices such as wearables to handle medical image and video processing at high data rates in future large-scale federated healthcare projects. Moreover, the UK Biobank, a large-scale biomedical study, has investigated the benefits of FL in real-world brain imaging under the support of the UK and French governments [100]. The focus of this project is the use of FL in the data analytics for two diseases, including Alzheimer and Parkinson. The MRI datasets along with covariates (e.g., age, sex information) are employed from 455 controls, 181 with non-progressive mild cognitive impairment (MCI), 208

Table 4. Taxonomy of FL applications for smart healthcare.

Applied domain	Ref.	FL type	FL clients	FL aggregator	Key contributions	Limitations
EHRs Management	[37]	HFL	Hospitals	Cloud server	A collaborative learning with FL for EHRs processing.	Convergence of the FL algorithm has not been verified.
	[65]	HFL	Smart phones	Data server	An FL scheme for predicting hospitalizations of patients.	The proposed model is simple and lacks detailed evaluations.
	[64]	HFL	Hospitals	Data server	An differential privacy-based FL for federated EHRs training.	The feasibility of differential privacy in real-world FL implementations should be considered.
	[84]	HFL	Radiology client center	Cloud server	An FL-based scheme with split learning for EHRs training.	The proposed model is simple and training complexity is not verified.
	[88]	HFL	Hospitals	Data server	Uncertainty FL for preterm-birth-related data analytics.	The complexity of local training at each hospital should be analyzed.
Health Monitoring	[72]	VFL	Smart phones	Cloud server	A personalized FL scheme for remote activity recognition monitoring.	Secure aggregation in FL communications has not considered.
	[92]	FTL	Wearable devices	Cloud server	A personalized FL scheme with transfer learning for human activity recognition.	Communication costs and training complexity have not been verified.
	[93]	HFL	Hospitals	Cloud server	A two-stage federated natural language processing approach for obesity analytics.	The privacy preservation for FL training should be considered.
	[94]	VFL	Mobile devices	Cloud server	An FL-based mobile activity monitoring approach for supporting healthcare applications.	The comparison with other non-IID-aware FL techniques has been ignored.
	[95]	HFL	Medical sites	Data centre	An FL-based approach for large nationwide health insurance data analytics in the US.	Issues related to federation agreement among medical sites need to be clarified in real-world settings.
	[96]	HFL	Mobile devices	Data centre	An FL-based method to predict mood using mobile devices.	The training resource usage and privacy should be considered in mobile FL.

progressive MCIc, 234 Alzheimerâ€™s disease, and 232 with Parkinsonâ€™s disease across multiple medical institutions in the UK. The preliminary results from the experiment shows that FL can improve the training

Table 5. Taxonomy of FL applications for smart healthcare (continued).

Applied domain	Ref.	FL type	FL clients	FL aggregator	Key contributions	Limitations
Medical Imaging	[98]	VFL	Hospitals	Cloud server	A variation-aware FL scheme for medical image construction.	Learning accuracy has not been investigated.
	[99]	HFL	MRI scan machines	Federated server	A privacy-enhanced FL scheme for brain imaging.	FL convergence has not been investigated.
	[100]	HFL	Medical sites	Data centre	An FL-based MRI analytic framework for brain imaging.	The practical aspects of the federated MRI training should be taken.
	[102]	HFL	Hospitals	Data centre	An FL scheme for supporting diagnosis of acute neurological symptoms.	The proposed model is too simple and more simulations need to be done.
	[103]	HFL	Hospitals	Data centre	An FL-based medical imaging approach with differential privacy for collaborative secure training.	Comparison of DL techniques in FL simulation has not been performed.
COVID-19 Detection	[112]	HFL	Hospitals	Data centre	An FL scheme for collaborative COVID-19 detection.	The proposed model is simple with a lack of detailed analysis.
	[113]	HFL	Hospitals	Data centre	A number of experiments for COVID-19 detection with FL.	The convergence of FL algorithms has not been verified.
	[114]	HFL	Data clients	Data centre	A dynamic fusion-based FL method for CT scan image analysis to diagnose COVID-19 infections.	Learning efficiency, e.g., latency, has not been verified.
	[115]	HFL	Medical institutions	Data centre	An FL-based solution for COVID-19 screening from Chest X-ray images.	Data loss caused by FL communications has not been considered.
	[117]	VFL	Hospitals	Aggregator	A federated deep learning method for detecting COVID-19 lung abnormalities in CT.	Training latency has not been analyzed.
	[118]	VFL	Medical institutions	Cloud server	An FL method for COVID region segmentation in chest CT using multi-national data from China, Italy, Japan.	Theorical analysis of FL communications and convergence should be provided.

performances in terms of better improved classification accuracy, which is highly applicable to medical image analytics in imaging-related disease diagnosis.

6.3 FL for International COVID-19 Project

Recently, a real-world FL project for COVID-19 region segmentation in chest CT with the participation of medical institutions from China, Italy, and Japan is presented in [118]. Specifically, a multi-national database consisting of 1704 scans from these three countries is collected for setting up the FL framework, including 736 scans of 700 patients from the First Affiliated Hospital of Hubei University of Medicine in Hubei, China, 496 scans of 244 patients from the Self-Defense Forces Central Hospital, Tokyo, Japan, and 472 scans of 147 patients from San Paolo Hospital, Milan, Italy. The FL training performance is evaluated via the experiments of semi-supervised segmentation of COVID regions in 3D CT. Compared with the traditional local training approach using 945 scans with the support of expert radiologists, the FL-based approach is able to capture better the ground truth shapes and has less false positives with the collaboration of training resources from multiple medical centres. Given the strict regulatory policy on data privacy, FL is proven as a promising solution for countries to collaborate in the COVID-19 segmentation and detection without worrying the user information leakage and lacks of dataset.

7 RESEARCH CHALLENGES AND FUTURE DIRECTIONS

This section presents the key challenges and future directions related to FL-Healthcare research.

7.1 Communication Issues in FL-based Smart Healthcare

As an important part between FL users and the aggregation server, communication plays an important role in FL enabled healthcare services. Indeed, proper communication resource allocation schemes can significantly improve the learning performance. This becomes more important when a large number of IoMT devices need to connect with the aggregation server for model updates in the uplink and model broadcasting in the downlink. In such a case, the aggregation server can employ efficient scheduling policies to select a suitable set of IoMT devices, as reported in many existing studies [39–41, 43]. Another critical challenge from the communication perspective lies in the dynamic and fast variations of wireless channels, thus affecting the reliability and quality of learning updates between IoMT devices and the aggregation server. A possible solution is to take into account the effect of user dropping out [57, 58] and considering more reliable design objectives such as outage probability and device availability.

7.2 Standard Specifications for Federated Healthcare Deployment

Although many promising results of FL-enabled healthcare services have been shown in the literature, there is no standard and universal to evaluate the performance of different approaches for the same problem. For example, various blockchain frameworks have been proposed for FL systems such as removing the need for a central server and managing the reliability of local updates from IoMT devices. However, it is difficult to compare such approaches since they are proposed for different (healthcare) scenarios and different network settings/datasets are used to evaluate their performance. In addition, there are critical challenges related to the universal existence and standardization of communication protocols, device hardware, deployment scenarios, and aggregation methods. Recently, a guideline on the architectural and design perspectives of FL is provided in IEEE Std 3652.1-2020 [121]. Moreover, this guideline also presents major concerns about FL such as privacy, security, performance efficiency, and economic viability, as well as evaluation schemes and performance metrics of FL systems.

7.3 Quality of Federated Healthcare Training Data

The training quality can be greatly degraded due to the heterogeneity of computational capabilities and data qualities of different hospital sites. A promising solution in this case is designing incentive mechanisms to motivate hospital/health organizations to use high-quality data for training as well as report reliable updates to the aggregation server. Game theory and blockchain are two important tools for designing incentive mechanisms

[32, 46, 55, 56, 68]. The training requirements (e.g., changes of data types, changes of learning rates, changes of training purpose-classification or regression) should be configured in a flexible manner so that the FL entity such as nurses, doctors, and patients can adjust their actions easily and appropriately. Such changes can affect the FL design and the underlying learning models (at the FL users and the aggregation server), which necessitates the development of adaptive FL approaches. AI is a promising tool as data from the past can be exploited to enhance the adaptability of FL model with future events. For example, DRL is leveraged in [71] to devise an incentive mechanism in case the conventional approaches perform worse when the feature dimensionality is sufficiently large.

7.4 Health Dataset Issues for Robust FL-based Health Data Analytics

In realistic healthcare scenarios, different clients may have different datasets such as text, images, audio, and time series, and different data content such as blood type, heart rate, face images, and body temperature. Conventionally, almost all FL approaches in the literature are examined on a single dataset with a limited number of features. For example, the work in [63] is tested on the diabetic retinopathy dataset while the work in [64] is evaluated on an EHR dataset, despite the fact that they are both proposed for privacy-preserving FL based healthcare services. Moreover, new heterogeneous FL approaches should be developed, where collaborating parties may have different models, but the central server can navigate this heterogeneity by private ensemble learning [122]. In this work, an inference strategy is proposed to enable participants to operate an ensemble of heterogeneous models, without having to explicitly join the data at a single location.

7.5 FL-based Healthcare in Next-Generation Networks

Although 5G networks are not fully available and commercially deployed over the world, there have been many research and development activities towards future 6G wireless systems [123]. 6G enables the availability of many applications such as Industry 5.0, intelligent healthcare, smart grid, holographic telepresence, and personalized body area networks. At the same time, many new technologies are introduced to meet much stricter 6G requirements such as blockchain, compressive sensing, THz and visible light communications, 3D networking, quantum communication, and large intelligence surface [123]. How to integrate FL functions on future 5G/6G medical devices, how to employ 6G devices, e.g., smart implants and wearables for large-scale FL-based healthcare, and what are new healthcare services enabled by 6G are open questions for future research. For example, future e-health services will be advanced by AI and FL capabilities, thus enhancing the quality of life and reducing the hospitalizations for patients [124].

7.6 FL with Provable Privacy Guarantee

Despite the fact that FL has great potential in protecting the user data privacy, there are numerous privacy issues that should be addressed properly, especially in smart healthcare contexts due to the high sensitivity of the health-related data. According to [13], FL privacy issues can be categorized as membership inference attacks, unintentional information leakage, and generative adversarial network. For example, the attacker may misuse the global FL model to check the presence of a data sample in the FL health data set. Moreover, the patient's information can be inferred when the patient's device sends the local model updates to the central server located at the hospitals and health organizations. In order to design privacy-preserving solutions for FL-healthcare systems, making use of differential privacy, AI and advanced cryptography techniques is promising. The use of differential privacy to further enhance the privacy of FL systems is considered in many studies such as [62–64, 66]. Such studies should be further investigated for healthcare applications such as how artificial noise is added to the model updates from the patient devices and to the global model from the central server.

7.7 Security Issues in FL-based Smart Healthcare

In FL-based smart healthcare systems, several participants at the client side can act as attackers and try to send poisonous model updates or fake information to degrade the model aggregation. Further, an adversary can contaminate information about data features during the local data training or modify local updates during the model transmission between local clients and the central server. At the server side, an external adversary can deploy attacks to steal information about the aggregated global model, which leads to serious privacy concerns such as information leakage. How to address such security issues is a real challenge in FL-based smart healthcare systems. Several solutions should be considered, such as using differential privacy [125] to protect training datasets against data breach. In addition, developing secure aggregation techniques [126] are promising solutions to provide a double-masking structure for encrypting local updates and implementing the key sharing among clients and the central server, aiming to protect clients against data modifications and attacks.

7.8 Non-iidness and Data Quality in FL-based Smart Healthcare

To achieve a desirable training performance in FL-based healthcare systems, a critical problem to be addressed is the non-iidness of the medical datasets which potentially makes the FL training divergent in the training. For example, a hospital can have a higher distribution for a certain type of local diseases than other hospitals in different geographic areas. In this case, the label distributions differ across medical institutions, making them challenging to join the federated data training. Without addressing this non-iidness issue, the data training would significantly suffer from quality degradation or even diverge. Solutions to overcome the non-iid challenge thus need to be developed, e.g., creating an additional subset of datasets to allocate fairly among clients [37], aiming to ensure efficient data training in FL-based smart healthcare. Another promising approach is to implement the feature shift among heterogeneous clients [127], by using local batch normalization to adjust the feature distributions at the client side before averaging the local models. Quantitative metrics are needed to assess non-iid data in the FL-based smart healthcare sector, such as standard deviation, precision, and accuracy with respect to label/feature distribution skew and homogeneous partitions [128].

8 CONCLUSIONS

FL is an emerging collaborative AI approach that has sparked an extreme interest to realize privacy-enhancing and scalable smart healthcare networks and applications. Given the lack of a comprehensive survey on the FL-healthcare topic in the open literature, this paper has provided a detailed survey of the use of FL in smart healthcare. We have first introduced the FL concept and motivations as well as the technical requirements for the utilization of FL in the smart healthcare domain. The recently advanced FL designs that would be useful to federated smart healthcare have been then discussed. Then, the key applications of FL in smart healthcare have been explored and discussed in details, including federated EHRs management, federated remote health monitoring, federated medical imaging, and federated COVID-19 detection. The emerging real-world projects related to FL-healthcare use cases have been provided, and the key lessons learned from the survey have been also highlighted. Finally, we have highlighted interesting challenges and discussed future directions in FL-smart healthcare.

The application of FL in smart healthcare is still at its infancy and will quickly mature in the coming years for providing intelligent and privacy-enhanced health services. FL is expected to play a key role in realizing large-scale and collaborative healthcare systems and allow for a shift from centralized health data analytics to distributed healthcare operations with privacy awareness. We believe that this article will stimulate more attention in this emerging area and encourage more research efforts toward the full realization of FL in smart healthcare.

REFERENCES

[1] S. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K.-S. Kwak, "The Internet of Things for health care: a comprehensive survey," *IEEE Access*, vol. 3, pp. 678–708, 2015.

[2] D. C. Nguyen, P. Cheng, M. Ding, D. Lopez-Perez, P. N. Pathirana, J. Li, A. Seneviratne, Y. Li, and H. V. Poor, "Enabling AI in future wireless networks: A data life cycle perspective," *IEEE Communications Surveys & Tutorials*, pp. 553–595, 2021.

[3] B. Shickel, P. J. Tighe, A. Bihorac, and P. Rashidi, "Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis," *IEEE journal of biomedical and health informatics*, vol. 22, no. 5, pp. 1589–1604, 2017.

[4] S. Avancha, A. Baxi, and D. Kotz, "Privacy in mobile technology for personal healthcare," *ACM Computing Surveys (CSUR)*, vol. 45, no. 1, pp. 1–54, 2012.

[5] V. S. Cheng and P. C. Hung, "Health Insurance Portability and Accountability Act (HIPPA) compliant access control model for web services," *International Journal of Healthcare Information Systems and Informatics (IJHISI)*, vol. 1, no. 1, pp. 22–39, 2006.

[6] S. Warnat-Herresthal, *et al.*, "Swarm learning for decentralized and confidential clinical machine learning," *Nature*, vol. 594, no. 7862, pp. 265–270, 2021.

[7] M. J. Sheller, B. Edwards, G. A. Reina, J. Martin, S. Pati, A. Kotrotsou, M. Milchenko, W. Xu, D. Marcus, R. R. Colen *et al.*, "Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data," *Scientific reports*, vol. 10, no. 1, pp. 1–12, 2020.

[8] G. Kaissis, A. Ziller *et al.*, "End-to-end privacy preserving deep learning on multi-institutional medical imaging," *Nature Machine Intelligence*, vol. 3, no. 6, pp. 473–484, 2021.

[9] G. A. Kaissis, M. R. Makowski *et al.*, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nature Machine Intelligence*, vol. 2, no. 6, pp. 305–311, 2020.

[10] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, D. Niyato, and H. V. Poor, "Federated learning for industrial internet of things in future industries," *IEEE Wireless Communications*, 2021.

[11] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 2, pp. 1–19, 2019.

[12] 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," *IEEE Communications Surveys & Tutorials*, 2021.

[13] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Generation Computer Systems*, vol. 115, pp. 619–640, 2021.

[14] 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 Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 2031–2063, 2020.

[15] 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," *arXiv preprint arXiv:2009.13012*, 2020.

[16] M. Parimala, R. M. Swarna Priya, Q.-V. Pham, K. Dev, P. K. R. Maddikunta, T. R. Gadekallu, and T. Huynh-The, "Fusion of federated learning and industrial Internet of Things: A survey," *arXiv preprint arXiv:2101.00798*, 2021.

[17] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, "Federated learning for healthcare informatics," *Journal of Healthcare Informatics Research*, pp. 1–19, 2020.

[18] N. Rieke, J. Hancox, W. Li, F. Milletari, H. R. Roth, S. Albarqouni, S. Bakas, M. N. Galtier, B. A. Landman, K. Maier-Hein *et al.*, "The future of digital health with federated learning," *NPJ digital medicine*, vol. 3, no. 1, pp. 1–7, 2020.

[19] D. Ye, R. Yu, M. Pan, and Z. Han, "Federated learning in vehicular edge computing: A selective model aggregation approach," *IEEE Access*, vol. 8, pp. 23 920–23 935, 2020.

[20] Y. Cheng, Y. Liu, T. Chen, and Q. Yang, "Federated learning for privacy-preserving AI," *Communications of the ACM*, vol. 63, no. 12, pp. 33–36, 2020.

[21] S. Yang, B. Ren, X. Zhou, and L. Liu, "Parallel distributed logistic regression for vertical federated learning without third-party coordinator," *arXiv preprint arXiv:1911.09824*, 2019.

[22] S. Sharma, C. Xing, Y. Liu, and Y. Kang, "Secure and efficient federated transfer learning," in *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 2569–2576.

[23] J. So, B. Güler, and A. S. Avestimehr, "Turbo-aggregate: Breaking the quadratic aggregation barrier in secure federated learning," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 1, pp. 479–489, 2021.

[24] J.-J. Yang, J.-Q. Li, and Y. Niu, "A hybrid solution for privacy preserving medical data sharing in the cloud environment," *Future Generation Computer Systems*, vol. 43, pp. 74–86, 2015.

[25] C. Xu, N. Wang, L. Zhu, K. Sharif, and C. Zhang, "Achieving searchable and privacy-preserving data sharing for cloud-assisted e-healthcare system," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8345–8356, 2019.

[26] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Gan-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, 2018.

[27] M. Staffa, L. Sgaglione, G. Mazzeo, L. Coppolino, S. D'Antonio, L. Romano, E. Gelenbe, O. Stan, S. Carpov, E. Grivas *et al.*, "An OpenNCP-based solution for secure ehealth data exchange," *Journal of Network and Computer Applications*, vol. 116, pp. 65–85, 2018.

[28] Z. Che, Y. Cheng, S. Zhai, Z. Sun, and Y. Liu, "Boosting deep learning risk prediction with generative adversarial networks for electronic health records," in *2017 IEEE International Conference on Data Mining (ICDM)*, 2017, pp. 787–792.

[29] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, "BEdgeHealth: A decentralized architecture for Edge-based IoMT networks using blockchain," *IEEE Internet of Things Journal*, vol. 8, no. 14, pp. 11 743–11 757, 2021.

[30] Y. Zhao, J. Zhao, L. Jiang, R. Tan, D. Niyato, Z. Li, L. Lyu, and Y. Liu, "Privacy-preserving blockchain-based federated learning for IoT devices," *IEEE Internet of Things Journal*, vol. 8, no. 3, pp. 1817–1829, 2021.

[31] C. Ma, J. Li, M. Ding, H. H. Yang, F. Shu, T. Q. Quek, and H. V. Poor, "On safeguarding privacy and security in the framework of federated learning," *IEEE network*, vol. 34, no. 4, pp. 242–248, 2020.

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

[33] D. C. Nguyen, M. Ding, P. Q.-V, P. N. Pathirana, L. B. Le, A. Seneviratne, J. Li, D. Niyato, and H. V. Poor, "Federated learning meets blockchain in edge computing: Opportunities and challenges," *IEEE Internet of Things Journal*, vol. 8, no. 16, pp. 12 806–12 825, 2021.

[34] E. J. De Aguiar, B. S. FaiÃ§al, B. Krishnamachari, and J. Ueyama, "A Survey of Blockchain-Based Strategies for Healthcare," *ACM Computing Surveys*, vol. 53, no. 2, pp. 1–27, 2020.

[35] 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 *2019 IEEE symposium on security and privacy (SP)*, 2019, pp. 739–753.

[36] T. Nishio and R. Yonetani, "Client selection for federated learning with heterogeneous resources in mobile edge," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, 2019, pp. 1–7.

[37] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, "Federated learning with non-IID data," *arXiv preprint arXiv:1806.00582*, 2018.

[38] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. Kiddon, J. Konečnỳ, S. Mazzocchi, H. B. McMahan *et al.*, "Towards federated learning at scale: System design," *arXiv preprint arXiv:1902.01046*, 2019.

[39] B. Xu, W. Xia, J. Zhang, T. Q. Quek, and H. Zhu, "Online client scheduling for fast federated learning," *IEEE Wireless Communications Letters*, vol. 10, no. 7, pp. 1434–1438, 2021.

[40] W. Xia, T. Q. Quek, K. Guo, W. Wen, H. H. Yang, and H. Zhu, "Multi-armed bandit-based client scheduling for federated learning," *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, pp. 7108–7123, 2020.

[41] S. Luo, X. Chen, Q. Wu, Z. Zhou, and S. Yu, "HFEL: Joint edge association and resource allocation for cost-efficient hierarchical federated edge learning," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6535–6548, 2020.

[42] W. Shi, S. Zhou, Z. Niu, M. Jiang, and L. Geng, "Joint device scheduling and resource allocation for latency constrained wireless federated learning," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 453–467, 2021.

[43] H. H. Yang, Z. Liu, T. Q. Quek, and H. V. Poor, "Scheduling policies for federated learning in wireless networks," *IEEE transactions on communications*, vol. 68, no. 1, pp. 317–333, 2020.

[44] M. N. Nguyen, N. H. Tran, Y. K. Tun, Z. Han, and C. S. Hong, "Toward multiple federated learning services resource sharing in mobile edge networks," *IEEE Transactions on Mobile Computing*, 2021.

[45] Q.-V. Pham, M. Zeng, R. Ruby, T. Huynh-The, and W.-J. Hwang, "UAV communications for sustainable federated learning," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3944–3948, 2021.

[46] J. Xu, H. Wang, and L. Chen, "Bandwidth allocation for multiple federated learning services in wireless edge networks," *arXiv preprint arXiv:2101.03627*, 2021.

[47] V. Tolpegin, S. Truex, M. E. Gursoy, and L. Liu, "Data poisoning attacks against federated learning systems," in *European Symposium on Research in Computer Security*. Springer, 2020, pp. 480–501.

[48] M. Fang, X. Cao, J. Jia, and N. Gong, "Local model poisoning attacks to byzantine-robust federated learning," in *29th USENIX Security Symposium USENIX Security 20)*, 2020, pp. 1605–1622.

[49] 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, 2020, pp. 16 070–16 084.

[50] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Generation Computer Systems*, vol. 115, pp. 619–640, 2021.

[51] J. Kang, Z. Xiong, D. Niyato, Y. Zou, Y. Zhang, and M. Guizani, "Reliable federated learning for mobile networks," *IEEE Wireless Communications*, vol. 27, no. 2, pp. 72–80, 2020.

[52] C. Hu, J. Jiang, and Z. Wang, "Decentralized federated learning: A segmented gossip approach," *arXiv preprint arXiv:1908.07782*, 2019.

[53] S. Lu, Y. Zhang, and Y. Wang, "Decentralized federated learning for electronic health records," in *2020 54th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2020, pp. 1–5.

[54] J. Passerat-Palmbach, T. Farnan, M. McCoy, J. D. Harris, S. T. Manion, H. L. Flannery, and B. Gleim, "Blockchain-orchestrated machine learning for privacy preserving federated learning in electronic health data," in *2020 IEEE International Conference on Blockchain (Blockchain)*. IEEE, 2020, pp. 550–555.

[55] J. Li, Y. Shao, K. Wei, M. Ding, C. Ma, L. Shi, Z. Han, and H. V. Poor, "Blockchain assisted decentralized federated learning (BLADE-FL): Performance analysis and resource allocation," *arXiv preprint arXiv:2101.06905*, 2021.

[56] J. Kang, Z. Xiong, D. Niyato, S. Xie, and J. Zhang, "Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10 700–10 714, 2019.

[57] K. A. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, "Practical secure aggregation for federated learning on user-held data," in *NIPS Workshop on Private Multi-Party Machine Learning*, 2016.

[58] J. So, B. Güler, and A. S. Avestimehr, "Turbo-Aggregate: Breaking the quadratic aggregation barrier in secure federated learning," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 1, pp. 479–489, 2021.

[59] G. Zhu, Y. Du, D. Gündüz, and K. Huang, "One-bit over-the-air aggregation for communication-efficient federated edge learning: Design and convergence analysis," *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, pp. 2120–2135, 2021.

[60] L. Zhu and S. Han, "Deep leakage from gradients," in *Federated Learning*. Springer, 2020, pp. 17–31.

[61] M. Nasr, R. Shokri, and A. Houmansadr, "Machine learning with membership privacy using adversarial regularization," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2018, pp. 634–646.

[62] M. Wu, D. Ye, J. Ding, Y. Guo, R. Yu, and M. Pan, "Incentivizing differentially private federated learning: A multi-dimensional contract approach," *IEEE Internet of Things Journal*, 2021.

[63] M. Malekzadeh, B. Hasircioglu, N. Mital, K. Katarya, M. E. Ozfatura, and D. Gündüz, "Dopamine: Differentially private federated learning on medical data," *arXiv preprint arXiv:2102.13314*, 2021.

[64] 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," *arXiv preprint arXiv:1910.02578*, 2020.

[65] M. Hao, H. Li, G. Xu, Z. Liu, and Z. Chen, "Privacy-aware and resource-saving collaborative learning for healthcare in cloud computing," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[66] R. Kerkouche, G. Acs, C. Castelluccia, and P. Genevès, "Privacy-preserving and bandwidth-efficient federated learning: an application to in-hospital mortality prediction," in *ACM Conference on Health, Inference, and Learning*, 2021.

[67] Y. Zhan, J. Zhang, Z. Hong, L. Wu, P. Li, and S. Guo, "A survey of incentive mechanism design for federated learning," *IEEE Transactions on Emerging Topics in Computing*, 2021.

[68] Y. Sarikaya and O. Ercetin, "Motivating workers in federated learning: A stackelberg game perspective," *IEEE Networking Letters*, vol. 2, no. 1, pp. 23–27, 2019.

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

[70] M. Tang and V. W. Wong, "An incentive mechanism for cross-silo federated learning: A public goods perspective," in *INFOCOM 2021 IEEE Conference on Computer Communications*, 2021.

[71] J. Zhao, X. Zhu, J. Wang, and J. Xiao, "Efficient client contribution evaluation for horizontal federated learning," in *The Second AAAI Workshop on Privacy-Preserving Artificial Intelligence (PPAI-21)*, 2021.

[72] Q. Wu, X. Chen, Z. Zhou, and J. Zhang, "FedHome: Cloud-edge based personalized federated learning for in-home health monitoring," *IEEE Transactions on Mobile Computing*, 2020.

[73] D. Yi, J. Su, C. Liu, and W.-H. Chen, "Personalized driver workload inference by learning from vehicle related measurements," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 1, pp. 159–168, 2017.

[74] G. K. Gudur, B. S. Balaji, and S. K. Perepu, "Resource-constrained federated learning with heterogeneous labels and models," *arXiv preprint arXiv:2011.03206*, 2020.

[75] "Animals-10 dataset," 2020. [Online]. Available: https://www.kaggle.com/alessiocorrado99/animals10

[76] G. K. Gudur and S. K. Perepu, "Resource-constrained federated learning with heterogeneous labels and models for human activity recognition," in *Deep Learning for Human Activity Recognition: Second International Workshop, DL-HAR 2020, Held in Conjunction with IJCAI-PRICAI 2020, Kyoto, Japan, January 8, 2021, Proceedings*. Springer Nature, 2020, p. 57.

[77] G. Krishna Gudur and S. K. Perepu, "Federated learning with heterogeneous labels and models for mobile activity monitoring," *arXiv e-prints*, pp. arXiv–2012, 2020.

[78] O. Rudovic, N. Tobis, S. Kaltwang, B. Schuller, D. Rueckert, J. F. Cohn, and R. W. Picard, "Personalized federated deep learning for pain estimation from face images," *arXiv preprint arXiv:2101.04800*, 2021.

[79] P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon, and I. Matthews, "Painful data: The UNBC-McMaster shoulder pain expression archive database," in *2011 IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, 2011, pp. 57–64.

[80] F. Sattler, S. Wiedemann, K.-R. Müller, and W. Samek, "Robust and communication-efficient federated learning from non-IID data," *IEEE transactions on neural networks and learning systems*, vol. 31, no. 9, pp. 3400–3413, 2020.

[81] T. S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I. C. Paschalidis, and W. Shi, "Federated learning of predictive models from federated electronic health records," *International journal of medical informatics*, vol. 112, pp. 59–67, 2018.

[82] K. Wei, J. Li, M. Ding, C. Ma, H. H. Yang, F. Farokhi, S. Jin, T. Q. S. Quek, and H. V. Poor, "Federated Learning With Differential Privacy: Algorithms and Performance Analysis," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3454–3469, 2020.

[83] D. Liu, T. Miller, R. Sayeed, and K. D. Mandl, “FADL: Federated-autonomous deep learning for distributed electronic health record,” *arXiv preprint arXiv:1811.11400*, 2018.

[84] P. Vepakomma, O. Gupta, T. Swedish, and R. Raskar, “Split learning for health: Distributed deep learning without sharing raw patient data,” *arXiv preprint arXiv:1812.00564*, 2018.

[85] R. Shao, H. He, H. Liu, and D. Liu, “Stochastic channel-based federated learning for medical data privacy preserving,” *arXiv preprint arXiv:1910.11160*, 2019.

[86] L. Huang, A. L. Shea, H. Qian, A. Masurkar, H. Deng, and D. Liu, “Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records,” *Journal of biomedical informatics*, vol. 99, p. 103291, 2019.

[87] L. Huang, Y. Yin, Z. Fu, S. Zhang, H. Deng, and D. Liu, “LoAdaBoost: Loss-based adaboost federated machine learning with reduced computational complexity on iid and non-iid intensive care data,” *Plos one*, vol. 15, no. 4, p. e0230706, 2020.

[88] S. Boughorbel, F. Jarray, N. Venugopal, S. Moosa, H. Elhadi, and M. Makhlouf, “Federated uncertainty-aware learning for distributed hospital EHR data,” *arXiv preprint arXiv:1910.12191*, 2019.

[89] P. Papadopoulos, W. Abramson, A. J. Hall, N. Pitropakis, and W. J. Buchanan, “Privacy and trust redefined in federated machine learning,” *Machine Learning and Knowledge Extraction*, vol. 3, no. 2, pp. 333–356, 2021.

[90] S. R. Pfohl, A. M. Dai, and K. Heller, “Federated and differentially private learning for electronic health records,” *arXiv preprint arXiv:1911.05861*, 2019.

[91] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, “Blockchain and edge computing for decentralized emrs sharing in federated healthcare,” in *GLOBECOM 2020-2020 IEEE Global Communications Conference*. IEEE, 2020, pp. 1–6.

[92] Y. Chen, X. Qin, J. Wang, C. Yu, and W. Gao, “Fedhealth: A federated transfer learning framework for wearable healthcare,” *IEEE Intelligent Systems*, vol. 35, no. 4, pp. 83–93, 2020.

[93] D. Liu, D. Dligach, and T. Miller, “Two-stage federated phenotyping and patient representation learning,” *arXiv preprint arXiv:1908.05596*, 2019.

[94] G. Krishna Gudur and S. K. Perepu, “Federated learning with heterogeneous labels and models for mobile activity monitoring,” *arXiv e-prints*, pp. arXiv–2012, 2020.

[95] D. Liu, T. A. Miller, and K. D. Mandl, “Confederated machine learning on horizontally and vertically separated medical data for large-scale health system intelligence,” *arXiv preprint arXiv:1910.02109*, 2019.

[96] X. Xu, H. Peng, L. Sun, Y. Niu, H. Ma, L. Liu, and L. He, “Federated depression detection from multi-source mobile health data,” *arXiv preprint arXiv:2102.09342*, 2021.

[97] X. Tan, C.-C. H. Chang, and L. Tang, “A tree-based federated learning approach for personalized treatment effect estimation from heterogeneous data sources,” *arXiv preprint arXiv:2103.06261*, 2021.

[98] Z. Yan, J. Wicaksana, Z. Wang, X. Yang, and K.-T. Cheng, “Variation-aware federated learning with multi-source decentralized medical image data,” *IEEE Journal of Biomedical and Health Informatics*, 2020.

[99] W. Li, F. Milletarì, D. Xu, N. Rieke, J. Hancox, W. Zhu, M. Baust, Y. Cheng, S. Ourselin, M. J. Cardoso *et al.*, “Privacy-preserving federated brain tumour segmentation,” in *International Workshop on Machine Learning in Medical Imaging*, 2019, pp. 133–141.

[100] S. Silva, B. A. Gutman, E. Romero, P. M. Thompson, A. Altmann, and M. Lorenzi, “Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data,” in *2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019)*, 2019, pp. 270–274.

[101] M. J. Sheller, B. Edwards, G. A. Reina, J. Martin, S. Pati, A. Kotrotsou, M. Milchenko, W. Xu, D. Marcus, R. R. Colen *et al.*, “Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data,” *Scientific reports*, vol. 10, no. 1, pp. 1–12, 2020.

[102] U. C. Srivastava, D. Upadhyay, and V. Sharma, “Intracranial hemorrhage detection using neural network based methods with federated learning,” *arXiv preprint arXiv:2005.08644*, 2020.

[103] M. Malekzadeh, B. Hasircioglu, N. Mital, K. Katarya, M. E. Ozfatura, and D. Gündüz, “Dopamine: Differentially private federated learning on medical data,” *arXiv e-prints*, pp. arXiv–2101, 2021.

[104] F. Ucar and D. Korkmaz, “COVIDiagnosis-Net: deep bayes-squeezenet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images,” *Medical Hypotheses*, vol. 140, p. 109761, 2020.

[105] P. Guo, P. Wang, J. Zhou, S. Jiang, and V. M. Patel, “Multi-institutional collaborations for improving deep learning-based magnetic resonance image reconstruction using federated learning,” *arXiv preprint arXiv:2103.02148*, 2021.

[106] M. Y. Lu, D. Kong, J. Lipkova, R. J. Chen, R. Singh, D. F. Williamsona, T. Y. Chena, and F. Mahmood, “Federated learning for computational pathology on gigapixel whole slide images,” *arXiv preprint arXiv:2009.10190*, 2020.

[107] X. Li, Y. Gu, N. Dvornek, L. H. Staib, P. Ventola, and J. S. Duncan, “Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: ABIDE results,” *Medical Image Analysis*, vol. 65, p. 101765, 2020.

[108] H. R. Roth, K. Chang *et al.*, “Federated learning for breast density classification: A real-world implementation,” in *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*. Springer, 2020, pp. 181–191.

[109] D. Nguyen, M. Ding, P. N. Pathirana, and A. Seneviratne, “Blockchain and AI-based solutions to combat coronavirus (COVID-19)-like epidemics: A survey,” *IEEE Access*, vol. 9, pp. 95 730–95 753, 2021.

[110] Q.-V. Pham, D. C. Nguyen, T. Huynh-The, W.-J. Hwang, and P. N. Pathirana, “Artificial intelligence (AI) and big data for coronavirus (COVID-19) pandemic: A survey on the state-of-the-arts,” *IEEE Access*, vol. 8, pp. 130 820–130 839, 2020.

[111] M. Loey, F. Smarandache, and N. E. M. Khalifa, “Within the lack of chest COVID-19 X-ray dataset: A novel detection model based on GAN and Deep Transfer Learning,” *Symmetry*, vol. 12, no. 4, p. 651, 2020.

[112] R. Kumar, A. A. Khan, S. Zhang, J. Kumar, T. Yang, N. A. Golalirz, Zakria, I. Ali, S. Shafiq, and W. Wang, “Blockchain-federated-learning and deep learning models for COVID-19 detection using CT imaging,” *arXiv:2007.06537 [cs, eess]*, 2020, arXiv: 2007.06537.

[113] B. Liu, B. Yan, Y. Zhou, Y. Yang, and Y. Zhang, “Experiments of Federated Learning for COVID-19 Chest X-ray images,” *arXiv:2007.05592 [cs, eess]*, 2020, arXiv: 2007.05592.

[114] W. Zhang, T. Zhou, Q. Lu, X. Wang, C. Zhu, H. Sun, Z. Wang, S. K. Lo, and F.-Y. Wang, “Dynamic fusion-based Federated Learning for COVID-19 detection,” *IEEE Internet of Things Journal*, pp. 1–1, 2021.

[115] I. Feki, S. Ammar, Y. Kessentini, and K. Muhammad, “Federated learning for COVID-19 screening from Chest X-ray images,” *Applied Soft Computing*, vol. 106, p. 107330, 2021.

[116] F. Qian and A. Zhang, “The value of federated learning during and post-COVID-19,” *International Journal for Quality in Health Care*, vol. 33, no. mzab010, 2021.

[117] Q. Dou, T. Y. So, M. Jiang, Q. Liu, V. Vardhanabhuti, G. Kaissis, Z. Li, W. Si, H. H. C. Lee, K. Yu, Z. Feng, L. Dong, E. Burian, F. Jungmann, R. Braren, M. Makowski, B. Kainz, D. Rueckert, B. Glocker, S. C. H. Yu, and P. A. Heng, “Federated deep learning for detecting COVID-19 lung abnormalities in CT: a privacy-preserving multinational validation study,” *npj Digital Medicine*, vol. 4, no. 1, p. 60, 2021.

[118] D. Yang, Z. Xu, W. Li, A. Myronenko, H. R. Roth, S. Harmon, S. Xu, B. Turkbey, E. Turkbey, X. Wang, W. Zhu, G. Carrafiello, F. Patella, M. Cariati, H. Obinata, H. Mori, K. Tamura, P. An, B. J. Wood, and D. Xu, “Federated semi-supervised learning for COVID region segmentation in chest CT using multi-national data from China, Italy, Japan,” *Medical Image Analysis*, vol. 70, p. 101992, 2021.

[119] “Federated learning for medical imaging,” 2021. [Online]. Available: https://www.intel.com/content/www/us/en/artificial-intelligence/posts/federated-learning-for-medical-imaging.html

[120] “Federated learning brings AI with privacy to hospitals,” 2021. [Online]. Available: https://healthcare-in-europe.com/en/news/federated-learning-brings-ai-with-privacy-to-hospitals.html

[121] Q. Yang, L. Fan, R. Tong, and A. Lv, “White paper-IEEE federated machine learning,” *IEEE White Paper*, 2021.

[122] C. A. Choquette-Choo, N. Dullerud, A. Dziedzic, Y. Zhang, S. Jha, N. Papernot, and X. Wang, “CaPC learning: Confidential and private collaborative learning,” *arXiv preprint arXiv:2102.05188*, 2021.

[123] C. de Alwis, A. Kalla, Q.-V. Pham, P. Kumar, K. Dev, W.-J. Hwang, and M. Liyanage, “Survey on 6G frontiers: Trends, applications, requirements, technologies and future research,” *IEEE Open Journal of the Communications Society*, vol. 2, pp. 836–886, 2021.

[124] L. Mucchi, S. Jayousi, S. Caputo, E. Paoletti, P. Zoppi, S. Geli, and P. Dioniso, “How 6G technology can change the future wireless healthcare,” in *2020 2nd 6G wireless summit (6G SUMMIT)*. IEEE, 2020, pp. 1–6.

[125] R. Hu, Y. Guo, H. Li, Q. Pei, and Y. Gong, “Personalized federated learning with differential privacy,” *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9530–9539, 2020.

[126] H. Fereidooni, S. Marchal, M. Miettinen, A. Mirhoseini, H. Möllering, T. D. Nguyen, P. Rieger, A.-R. Sadeghi, T. Schneider, H. Yalame *et al.*, “SAFELearn: secure aggregation for private federated learning,” in *2021 IEEE Security and Privacy Workshops (SPW)*. IEEE, 2021, pp. 56–62.

[127] X. Li, M. Jiang, X. Zhang, M. Kamp, and Q. Dou, “Fedbn: Federated learning on non-IID features via local batch normalization,” *arXiv preprint arXiv:2102.07623*, 2021.

[128] Q. Li, Y. Diao, Q. Chen, and B. He, “Federated learning on non-IID data silos: An experimental study,” *arXiv preprint arXiv:2102.02079*, 2021.