electronics

Review

Federated Learning in Smart Healthcare: A Survey of Applications, Challenges, and Future Directions

Mohammad Nasajpour [1,*], Seyedamin Pouriyeh [1], Reza M. Parizi [2], Meng Han [3], Fatemeh Mosaiyebzadeh [4], Liyuan Liu [5], Yixin Xie [1] and Daniel Macêdo Batista [4]

[1] Department of Information and Technology, Kennesaw State University, Marietta, GA 30152, USA; spouriye@kennesaw.edu (S.P.); yxie11@kennesaw.edu (Y.X.)
[2] Decentralized Science Lab, Kennesaw State University, Marietta, GA 30060, USA; rparizi1@kennesaw.edu
[3] College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China; mhan@zju.edu.cn
[4] Department of Computer Science, University of São Paulo, São Paulo 05508-090, Brazil; fatemehm@ime.usp.br (F.M.); batista@ime.usp.br (D.M.B.)
[5] Decision & System Sciences Department, Saint Joseph's University, Philadelphia, PA 19131, USA; lliu@sju.edu
* Correspondence: mnasajp1@students.kennesaw.edu

Abstract: In recent years, novel technologies in smart healthcare systems have opened significant opportunities for diagnosis and treatment across various medical fields. Federated Learning (FL), a decentralized machine learning approach, trains shared models using local data from devices like wearables and hospital systems without transferring sensitive information, offering a promising solution to privacy challenges in areas such as cancer prediction, COVID-19 detection, drug discovery, and medical image processing. This literature survey reviews FL architectures (e.g., FedHealth, PerFit), applications, and recent advancements, demonstrating their impact on healthcare through enhanced predictive models for patient care. Key findings include improved accuracy in wearable-based diagnostics and secure multi-institutional collaboration, though limitations persist. We also highlight open challenges, such as security risks, communication costs, and data heterogeneity, which require further research attention.

Keywords: smart healthcare; federated learning; artificial intelligence

check for updates

Received: 13 March 2025
Revised: 17 April 2025
Accepted: 20 April 2025
Published: 25 April 2025

Citation: Nasajpour, M.; Pouriyeh, S.; Parizi, R.M.; Han, M.; Mosaiyebzadeh, F.; Liu, L.; Xie, Y.; Batista, D.M. Federated Learning in Smart Healthcare: A Survey of Applications, Challenges, and Future Directions. *Electronics* **2025**, *14*, 1750. https://doi.org/10.3390/electronics14091750

1. Introduction

Research on Artificial Intelligence (AI) has led the world into a journey of generating new technologies to allow machines to perform tasks similar to humans. This is technically based on the ability to learn from data [1,2]. One of the most important approaches to this concept is the application of machine learning (ML) algorithms. ML algorithms have shown superb results in different domains, including economics and finance [3,4], manufacturing [5,6], transportation [7,8], healthcare [9,10], cybersecurity [11,12], and many more [13]. Deep learning (DL), as a subset of ML, is also attracting more attention regarding its benefits for solving complex problems by learning from large datasets. It is a repetitive cycle of learning from data causing better performance [14].

However, the healthcare system presents unique challenges due to the sensitivity of patient data, which are generated by institutions, hospitals, and individuals. Regulations such as the Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR) impose strict policies to protect medical records, prioritizing patient privacy [15,16]. Consequently, health data are highly restricted, limiting access for research communities even with anonymization, as their usage remains tightly regulated [17]. While ML and DL rely on large, centralized datasets, these privacy constraints and the complexity of health data hinder traditional approaches, necessitating innovative solutions. Federated learning (FL) has emerged as a promising approach by enabling model training on decentralized data, keeping sensitive patient information local, such as on hospital servers or wearables, while sharing only model updates, often encrypted, with a central server [18]. This aligns with HIPAA's Privacy Rule by minimizing data exposure [15] and GDPR's data protection by design through secure aggregation [16], as seen in real-world applications like multi-hospital collaborations [19]. However, challenges such as ensuring the right to erasure of GDPR in distributed models persist [20]. Despite the growing adoption of Federated Learning (FL) to address these issues, a critical gap remains: existing literature lacks a comprehensive survey that systematically integrates FL architectures, applications, and challenges in smart healthcare, particularly in tackling privacy, regulatory compliance, and data heterogeneity. This survey aims to fill this gap by providing a comprehensive review of the role of FL in advancing secure, privacy-preserving healthcare solutions.

Additionally, the heterogeneity of health data can be a major hurdle in utilizing traditional ML in the healthcare domain [21]. For instance, variability in electronic health records (EHRs), such as differences in data formats, missing values, and unstructured clinical notes, often complicates the application of traditional ML models, which typically require standardized and homogeneous inputs [22]. Similarly, the diverse nature of medical imaging data—stemming from variations in acquisition protocols and patient demographics—can limit the generalizability of conventional ML approaches [23]. These challenges highlight the need for advanced techniques to address the complexity of healthcare data.

To date, a new privacy-preserving ML technique, called federated learning (FL) [24,25], has shown promising effects in data-sensitive domains, especially healthcare, by safeguarding more data rather than ML algorithms [18]. Unlike other distributed learning approaches such as data parallelism, where data are split across nodes but still require centralized aggregation of raw data, FL distinctly trains models locally on each participating device or institution and only shares model updates with a central server, never the raw data. The concept of FL was first coined by authors in 2017 [26] as a decentralized training method designed to improve data privacy. In essence, FL enables a global model to be trained by aggregating insights from locally trained models across various institutions, allowing sensitive data to remain at its source while still contributing to collaborative learning.

In the healthcare domain, the FL technique can be utilized to train a global ML model that incorporates health data from different sources (healthcare systems, hospitals) while addressing concerns about privacy of health data. It performs training on the client side for each individual data set without having any access to their data. Then, the local FL models captured from different sources are aggregated as a global FL model [18]. To be more specific, the FL can be implemented for various purposes regarding the distributions' situations including massively, not independent and identically, and unbalanced. First, an example of massive distribution can be smartphones that collect a huge amount of data through their sensor. A not identically and independently distributed distribution of data can also be the different types of patients. Furthermore, the amount of data can vary based on the source of data collection, such as the variety of patients in a hospital suffering from a specific disease [26]. In addition, FL can improve the healthcare infrastructure by integrating with medical device networks and improving healthcare delivery systems, particularly in urban and regional contexts where diverse data sources and connectivity challenges are prevalent [27]. In general, the domain that can benefit the most from FL techniques has been recognized as healthcare [18].

1.1. Motivation

As health data gain more attention from data science communities and authorities, various studies have been conducted on these data using different technologies. With the emergence of the FL technique, FL is becoming a major popular method of distributed learning from sensitive data. In the healthcare domain, FL can have a huge impact on healthcare systems by securely training models in different areas. Various research studies have already been conducted on this topic. This leads to a need for reviewing articles based on FL and healthcare. Moreover, we also review the existing challenges of FL in this domain. This study aims to integrate the recently adopted approaches on this subject. To the best of our knowledge, no comprehensive research has been conducted regarding FL and healthcare considering different architectures, applications, and challenges.

1.2. Methodology

1.2.1. Research Questions

We conduct a comprehensive overview of all the approaches proposed so far regarding FL in the healthcare domain. We begin the journey of this paper by discovering the main research questions (RQ) behind the scene. The first question fits into our introduction in order to offer a general idea to the reader.

RQ1: What is FL and how can it assist the health sector?

Next, the various proposed architectures based on FL lead us to the second question.

RQ2: What are the most common FL architectures that can be adopted for the healthcare domain?

After reviewing existing architectures, a high number of ML approaches in the different fields of healthcare have led us to the next question.

RQ3: Which sectors in the healthcare domain can benefit the most from FL technology and what are the state-of-the-art FL applications in those fields?

Although FL has shown promising results with respect to privacy compared to centralized ML techniques, it can be vulnerable to different challenges. This brings us to the last question in our review paper.

RQ4: What are the major challenges of using FL in the healthcare domain?

1.2.2. Search Strategy

We conducted a systematic search across various databases, including Google Scholar, SpringerLink, ScienceDirect, ACM Digital Library, IEEE Xplore, and arXiv, using specific search terms such as "federated learning", "federated machine learning", and "healthcare". The search process involved several stages:

- **Selection and Screening:** Initially, we screened papers based on their titles and abstracts, focusing on those related to federated learning and its application in healthcare. We applied the following inclusion criteria: (1) peer-reviewed articles, (2) studies that explicitly discuss federated learning or related techniques, and (3) research relevant to healthcare applications.
- **Exclusion Criteria:** We excluded papers that were not directly related to healthcare, those without empirical analysis or practical applications, and those published in languages other than English.
- **Evaluation:** After screening, we evaluated the full texts of the selected papers for quality and relevance, considering factors such as research methodology, sample size, and applicability to current trends in the field.

The search period covered November 2021 to December 2024, as this time frame captures the rapid evolution of federated learning applications in healthcare following the increased adoption of decentralized data approaches during the COVID-19 pandemic.

We selected this range to ensure the inclusion of foundational studies while maintaining relevance to recent advances in the field.

1.3. Difference with Other Review Papers

With the recent emergence of FL technology in healthcare, there are some review articles on this topic. Xu et al. [28] demonstrated the main challenges of FL with respect to recent advances in it. They also briefly described some of the FL applications within healthcare. Since overcoming FL challenges is vital for designing effective models, several authors have focused primarily on these challenges. The authors in [29] conducted a review that focuses on the advantages of FL for healthcare care and the associated challenges. In addition to security challenges, Kaissis et al. [30] provided an overview of appropriate technologies for better data protection with a focus on medical imaging. The purpose of this paper is to provide a comprehensive review of the latest federated learning approaches, achievements, and challenges in the healthcare domain.

1.4. Outline

The remainder of this paper is organized as follows: Section 2 covers different FL-based architectures applicable in healthcare. In the second Section 3, the applications and use cases of FL in the medical system are discussed. Then, Section 4 reviews the main challenges and considerations of the FL in the same domain. Section 5 provides some future research directions. Section 6 presents the concluding remarks.

2. Architectures

Due to recent research, various architectures and libraries have been adopted by federated learning (FL) technology. Horizontal FL (HFL) [18], Vertical FL (VFL) [18], Federated Transfer Learning (FTL) [31], and others represent general architectures designed to address specific problems across various domains. However, domain-specific approaches have also emerged. Here, we focus on architectures proposed for the healthcare system. To the best of our knowledge, nine key FL-based architectures, FedHealth, Federated-Autonomous Deep Learning (FADL), Ethereum Blockchain-based, PerFit, FEEL, DMFL-Net, FedCare, Sensor-based HAR, and FedHome are tailored to healthcare challenges and are discussed in detail below. Table 1 provides a comparative analysis of these architectures, summarizing their methodologies, strengths, limitations, and contexts of application to enhance clarity and facilitate quick comparisons.

Table 1. Quantitative Comparison of FL Architectures for Healthcare Systems.

Architecture	Methodology	Context of Application	Strengths	Limitations	Dataset	Perf. Metrics	Baseline Comparison
FedHealth [19]	FL + Transfer Learning (TL)	Wearable healthcare (e.g., activity monitoring)	Higher accuracy, personalization	Computationally intensive	UCI Smartphone	Acc = 98.8	Acc = 85 (CNN baseline)
PerFit [32]	FL + TL, Distillation	IoT healthcare (e.g., activity recognition)	Handles heterogeneity, high performance	Complex personalization process	MobiAct	Acc = 95.37 > FedAvg	Acc = 85 (cCNN baseline)
FedHome [33]	FL + Generative CNN Autoencoder	In-home elderly monitoring	Good performance, privacy	Imbalanced data handling	N/A	Acc = 95.41	Acc = 87.92 (CNN baseline)

Table 1. *Cont.*

Architecture	Methodology	Context of Application	Strengths	Limitations	Dataset	Perf. Metrics	Baseline Comparison
FADL [34]	FL + Neural Network	EHR-based mortality prediction	Higher accuracy, balanced models	Limited to structured EHR data	eICU	AUC = 0.79	AUC=0.75 (FL-Avg baseline)
Ethereum Blockchain [35]	FL + Blockchain, Encryption	Healthcare consortium data sharing	Strong privacy protection	High computational cost	N/A	Not specified	N/A
FEEL [36]	FL + Differential Privacy	Mobile healthcare (e.g., cancer detection)	High efficiency, privacy	Potential accuracy trade-off	Breast cancer	Acc = 86, F1 = 0.90	Acc = 88, F1 = 0.91 (Centralized Learning baseline)
DMFL-Net [37]	FL + Neural Network	COVID-19 and chest disease detection	High accuracy, fast classification	Specific to imaging data	CXR images	Acc = 92.25, F1 = 92.21	Acc = 90, F1 = 90 (default FL baselines)
FedCare [38]	FL + Split Learning	IoMT for rural/elderly monitoring	Reduced training time, scalability	Limited evaluation scope	N/A	Acc = 90.32	N/A
Sensor-based HAR [39]	FL + Homomorphic Encryption	Wearable devices (e.g., activity recognition)	Strong privacy, high accuracy	Encryption overhead	Sport, DaLiAC	Acc = 89.5	Acc = 94.6 (3D CNN baseline)

2.1. FedHealth

Within the enhancement of technologies, a new area of smart healthcare has been defined in order to perform various tasks such as diagnosing, monitoring, etc. The widespread adoption of wearable devices in smart healthcare has enabled machine learning algorithms to successfully train massive personal data from users [40]. Although smart healthcare itself is a great success regarding the issues in the healthcare industry, it faces two important challenges as well. The first one is that the data cannot be shared in various institutions due to the privacy concerns and regulations [41,42]. Another critical challenge is the lack of personalization, which happens when a specific trained model is distributed to various users with different characteristics. In order to solve these challenges in wearable devices, authors in [19] implemented a federated transfer learning framework named FedHealth. This framework is claimed to be capable of forming efficient machine learning models by collecting the data from separate institutions with respect to the privacy preserving of the users. By having the cloud built, transfer learning is adopted to personalize the model learning. Using FL, a cloud model of all the institutions' models is provided. The FedHealth architecture incorporates three different technologies, including wearable devices, transfer learning, and FL. To personalize the data collected from the institutions, the authors implement the transfer learning technique. They also apply a Convolutional Neural Network (CNN) in order to achieve a more generalized model. The UCI Smartphone dataset [43] was applied to evaluate the model. It includes regular activities such as walking, standing, etc., captured from 30 users. Moreover, it has been shown that the performance of FedHealth is higher than that of traditional algorithms.

The authors also demonstrate the positive effect of FedHealth on recognition accuracy. Moreover, they highlight the framework's potential for use in the healthcare system, particularly for real-time diagnosis of Parkinson's disease in the future.

Figure 1 illustrates the FedHealth architecture, showing wearable devices collecting local sensor data, training models on site, and sending encrypted updates to a cloud server for FL aggregation. The global model is then refined with TL for user-specific personalization, visually linking the process to privacy preservation and adaptability.

This workflow highlights how FedHealth addresses the dual challenges of restricted data sharing and personalization, though its reliance on deep neural networks may strain resource-constrained wearables, and its Parkinson's application lacks broader validation. These limitations suggest areas for further optimization in scalable deployments. The authors also showcase FedHealth's superior recognition accuracy and its potential for real-time healthcare applications, such as auxiliary diagnosis of Parkinson's disease, where it achieved average accuracies of 84.3% (arm droop) and 74.9% (postural tremor) across three hospitals, narrowing the gap with an ideal centralized model (92.6% and 83.1%, respectively). However, FedHealth is not without limitations. Its reliance on deep neural networks may pose computational challenges for resource-constrained wearable devices, potentially limiting scalability in low-power settings. Additionally, the framework assumes sufficient data similarity across users for effective transfer learning, which may not always hold true given the heterogeneity of real-world health data. These factors could impact its performance in broader deployments.

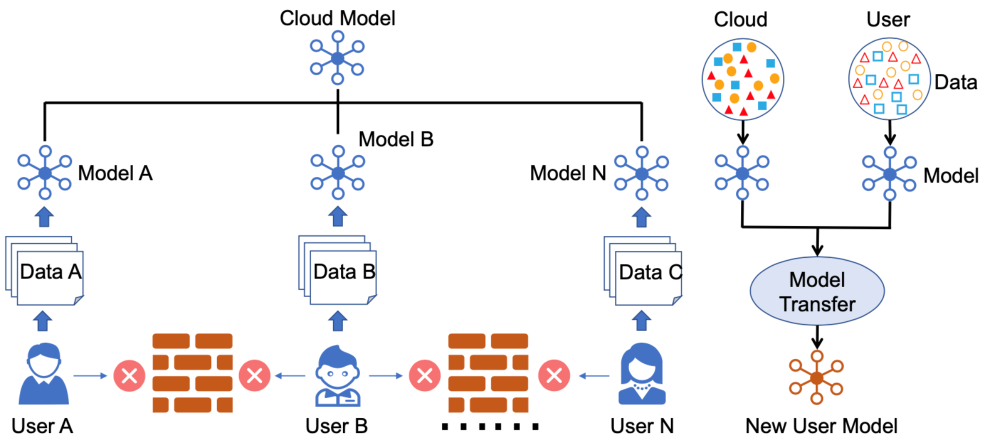

Figure 1. FedHealth framework architecture [19]. Wearable devices collect sensor data (e.g., accelerometer, gyroscope) and train local models. Encrypted updates are aggregated via federated learning on a cloud server into a global model, which is personalized using transfer learning for individual users, ensuring privacy and tailored healthcare applications like activity recognition and Parkinson's diagnosis.

2.2. PerFit

PerFit is another FL-based architecture proposed in [32]. In sort, PerFit is designed to be applied for the Internet of Things (IoT) devices, but it is also capable of assisting various devices that are applicable to the healthcare system. The heterogeneity challenges are discussed in the proposed model where they cover heterogeneity with respect to IoT devices, statistical variations, and model differences. Figure 2 depicts the PerFit architecture, showing IoT devices unloading local models to a cloud server, which aggregates them via FL and applies federated distillation for personalization. This three-stage process—unloading, learning, and personalization—illustrates how PerFit mitigates device heterogeneity and statistical skew in healthcare IoT, as validated on the MobiAct dataset [43] with 95.37% accuracy.

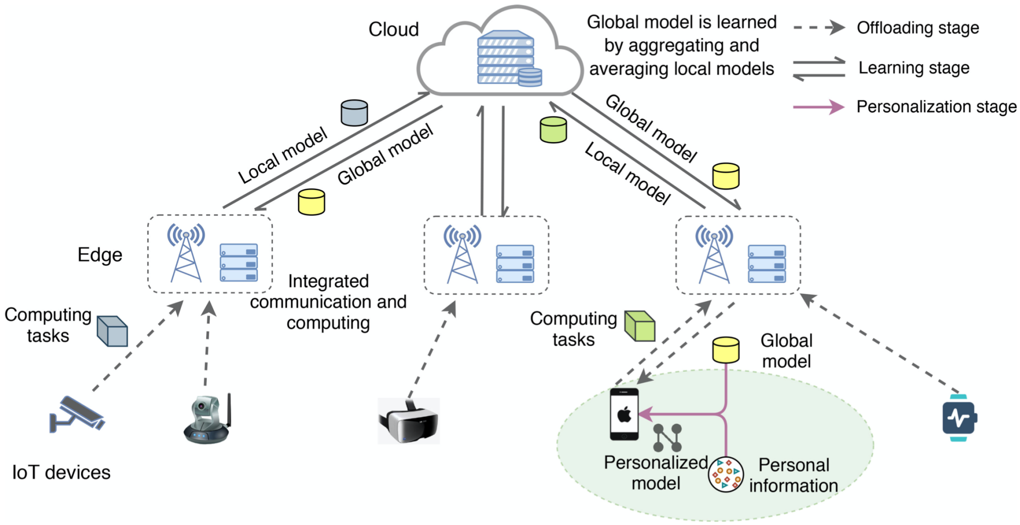

Figure 2. PerFit framework architecture [32]. IoT devices unload local models to a cloud server for federated learning aggregation, followed by federated distillation to personalize models for individual users, addressing device and data heterogeneity in healthcare applications like activity recognition.

2.2.1. Heterogeneity of Devices

There are a variety of IoT devices used in healthcare domain based on their characteristics such as hardware, network, etc. Due to the heterogeneity of IoT devices in healthcare, communication costs can critically lower the efficiency of the FL model.

In addition, loss of connection or energy can also negatively impact the FL system and must be considered. Later, in Section 4.1, we discuss the communication efficiency challenges and its solutions regarding FL and healthcare.

2.2.2. Statistical Heterogeneity

As mentioned previously, data distribution situations can highly impact the FL system. More specifically, the data captured from different devices can differ vastly in healthcare domain [28]. The paper demonstrates that the Federated Averaging (FedAvg) algorithm cannot solve the challenge of skewed data distribution, which leads to developing PerFit (we further explain the data skewing challenge in Section 4.3) [26].

2.2.3. Model Heterogeneity

In order to achieve an adaptable FL global model, all of the devices have to agree on following a specific framework. However, local models differ from each other when it comes to the IoT devices in healthcare. Consequently, all the model weights need to be shared, which is not acceptable due to privacy concerns regarding the medical data.

Overall, PerFit architecture, which is cloud-based FL, aims to address the aforementioned challenges by adopting personalized FL approaches (Figure 2) in three stages (below) to fully learn the model:

Unloading

The IoT devices send their learning models to the cloud for faster data processing.

Learning

Each of the devices and the cloud are calculated based on the data samples and build a local model. Then, the model information is transferred to the server to generate a global model. After a number of iterations, the appropriate global model is sent back to the cloud and devices for the next stage.

Personalization

Based on each device's local data, a personalized model is trained for further analysis.

The validation of the proposed framework is demonstrated by using a human activity recognition dataset called MobiAct [44], which is basically a collection of captured activities. The authors implemented FTL and Federated Distillation for personalizing the model, and claimed good performance of PerFit with respect to the aforementioned challenges [32].

2.3. FedHome

With the population's aging growth in the world, providing health assistance to the elderly is becoming crucial. Smart healthcare era has responded well enough to this essential growing need by adopting various types of IoT devices, specifically wearables. However, as we discussed previously, this era could face some major challenges, especially privacy of the collected data. Authors in [33] proposed FedHome, a personalized approach by adopting FL technology to better protect the user's data within the cloud server for in-home monitoring. They applied FL in order to avoid data leakage, which can be caused by uploading the sensitive data. This means the global model is trained using collected updated models, not the user's data. The FedHome architecture is basically designed with three main parts, including cloud, FL, and personalization technique. Since personalized data can impact the efficiency of the model, the authors implemented a Generative Convolutional Autoencoder (GCAE) approach within the FL framework. This autoencoder enables the framework to positively deal with the imbalanced data by learning the features. Moreover, CNN was chosen as the major algorithm for both encoder and decoder of the GCAE architecture.

FedHome was validated on a human activity recognition dataset for in-home monitoring, achieving 95.41% accuracy (Table 1), outperforming traditional CNNs (87.92%) and other FL methods [33]. Figure 3 illustrates the FedHome architecture, depicting wearable devices collecting activity data, training local models, and sending updates to a cloud server for FL aggregation. The GCAE then personalizes the global model, visually emphasizing how FedHome balances privacy and performance for elderly care. This workflow underscores its strengths in privacy preservation and communication efficiency, though its reliance on GCAE may increase computational demands, a limitation not fully explored in resource-constrained settings.

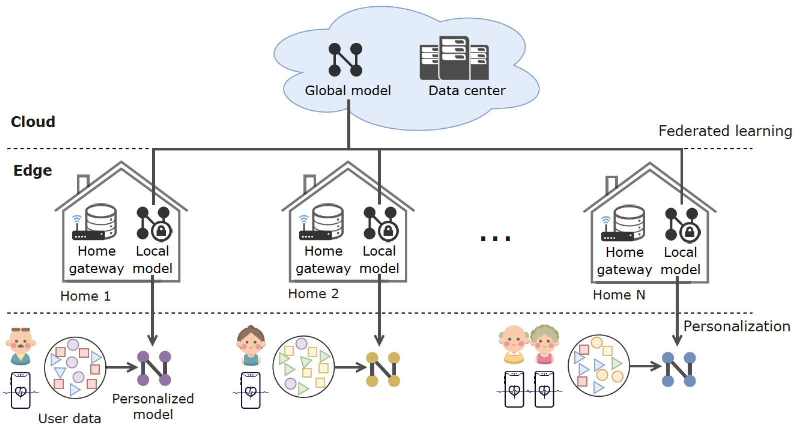

Figure 3. FedHome architecture overview [33]. Wearable devices collect activity data (e.g., motion, vital signs) and train local models. Updates are aggregated via federated learning on a cloud server into a global model, which is personalized using a Generative Convolutional Autoencoder (GCAE) to address imbalanced data, ensuring privacy and effective in-home monitoring for the elderly.

2.4. FADL

Electronic Health Records (EHRs) are defined as health data captured by various institutions or devices, constituting the patient's electronic medical history [45]. Using EHR data to develop different predictive ML models constitutes a major component of data-driven learning in healthcare. However, privacy and security regulations pose significant concerns in the utilization of EHR data. Authors in [34] demonstrated the weakness of a traditional FL approach, which unevenly distributed data for training, and proposed a more efficient FL-based model called Federated-Autonomous Deep Learning (FADL). They used an available EHR dataset of ICU records, captured from 58 different hospitals, to predict patient mortality rates [46]. The FADL has been tested over the ICU datasets. This framework was built on an artificial neural network with three fully connected layers. It was concluded that the FADL model has outperformed the traditional FL. Tested against traditional FL, FADL achieved superior performance, with an AUC of 0.79 compared to 0.75 for FL-Avg (Table 1), highlighting its ability to handle distributed EHR data effectively. Figure 4 illustrates the FADL workflow, showing local EHR data from hospitals being processed by autonomous deep learning models, with updates aggregated via FL into a global model on a central server. This visualization emphasizes how FADL mitigates uneven data distribution by enabling localized training while preserving privacy, a key improvement over traditional FL. Though effective for structured EHRs, its limitation lies in its specificity to such data, potentially restricting applicability to unstructured or multimodal datasets.

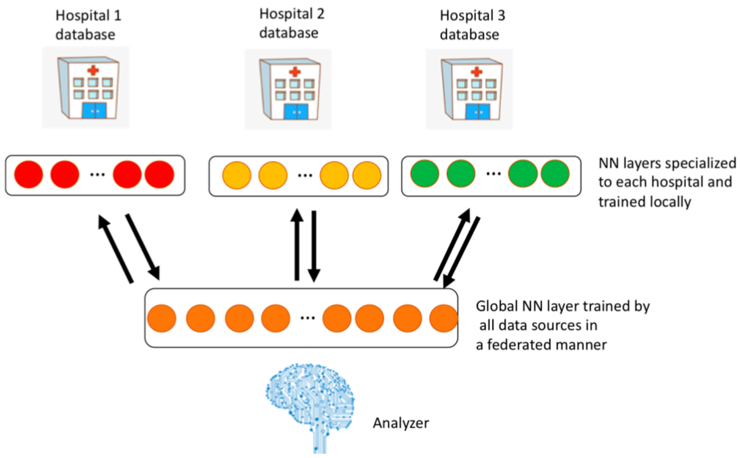

Figure 4. FADL framework architecture [34]. Hospitals process local EHR data (e.g., ICU records) using autonomous deep learning models with three fully connected layers. Model updates are aggregated via federated learning on a central server into a global model, enhancing mortality prediction while preserving patient privacy across 58 institutions.

2.5. Blockchain-Based Ethereum

The urgent demand for securing the privacy of health data has led researchers in [35] to propose an architecture of FL based on the Ethereum blockchain for the healthcare system. The architecture was developed in order to address different aspects including ethical, legal, economic, and technical. These aspects are integral to addressing the data privacy challenge. The distribution of data can have a significant impact on models. It can be divided into two major areas including mobile devices and a consortium. Here, the main focus of the paper is on a consortium in healthcare, which is a set of clinical institutions. It does not have the same large number of parties or institutions as mobile devices, but each institution can maintain a large amount of data. Moreover, having proper computational and storage resources and the ability to train at any time during a day with a reliable network can positively impact on the efficiency of the model. Figure 5 demonstrates the overview of different parties of the proposed system. Each party of the architecture is identified by its Ethereum account in order to collaborate with other parties. The architecture's setting is formed by three different characteristics as below sections.

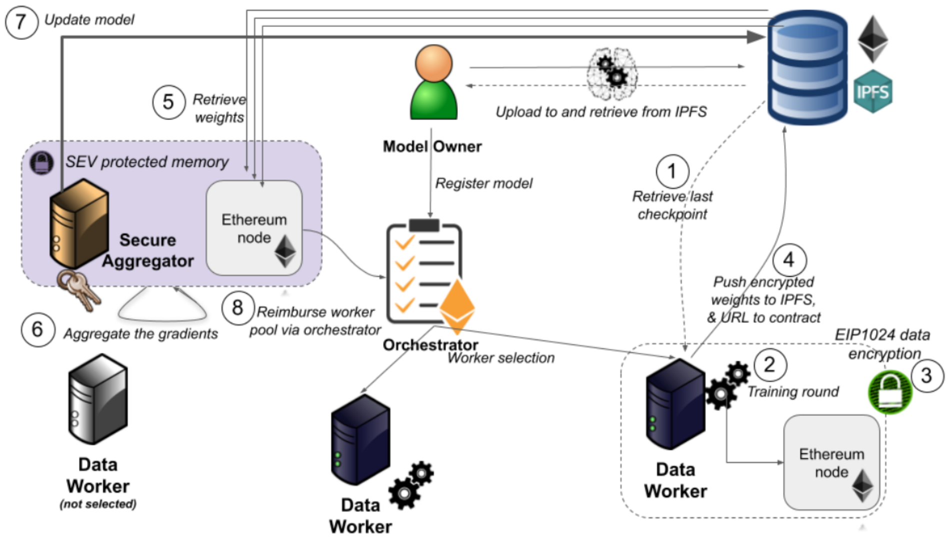

Figure 5. Ethereum blockchain architecture [35]. The Model Owner registers the model and retrieves checkpoints via the Orchestrator. Data Workers train local models, encrypt updates using EIP1024, and push them to IPFS. The Secure Aggregator, operating in SEV-protected memory, aggregates gradients, ensuring privacy and secure collaboration across healthcare institutions.

2.5.1. Data Policy

By implementing the Ethereum blockchain in the architecture, the data privacy challenge can be addressed within the healthcare consortium. Different tools have been applied within the Ethereum platform including Hyperledger Besu [47] and an Orchestrator smart contract [48] to perform different tasks such as viewing, transacting, communicating, and monitoring the network. Overall, these allow data owners to define the policies of accessibility to their data.

2.5.2. Secure Aggregation

The number of parties for the FL setting within the healthcare system is pretty low, and it can be further reduced by the filters and rules for model selection. Since the institutions may know each other, posing significant security risks, the AMD's Secure Encrypted Virtualization (SEV) has been proposed within the architecture. This, as a trusted third party, enables better memory encryption to prevent attackers from accessing the private data.

2.5.3. Peer-to-Peer Transition

One of the most important aspects during training is to secure weights that are transferred between workers and Secure Aggregator (SA). The reason is that they contain sensitive private data. Since the institutions in the system may be interested in the raw updates from others, it is crucial to encrypt both traffic and communication across the network. In order to achieve this, the authors implemented Ethereum Improvement Proposal 1024 [49] for encrypting the new weights from each worker. The encryption is processed before the weights are transferred to the aggregator.

Although the aggregator provides proper privacy by hiding the data during the transmission of updates to the model, there is a possibility that an attacker can access the information by assuming the identity of institutions. To prevent this concern, the architecture implements two approaches as below.

Selecting Randomly

After gradient updates are computed by all the workers, SA randomly chooses weights in order to train the model. After the aggregation, neither does the SA remember the unpicked weights, nor do the workers know if they contributed. This approach is implemented to prevent training rounds from being reverse-engineered [50]. The architecture can also be enhanced by implementing Differential Privacy [51].

Audit Trail

Within the proposed architecture, the events of learning stages are captured using the audit trail, which is an important part of data privacy in the consortium. During the aggregation, each round is assigned a random nonce, and the encryption of data is applied by using the Diffie–Helmann [52] method, which exchanges the keys between the worker and SA. Consequently, if the mentioned keys remain private, there is no attack to be concerned about.

2.6. FEEL

With recent advances of AI in healthcare, hospitals have adopted mobile wearable devices to better link patients and health authorities. Training a neural network model can efficiently achieve better accuracy, which poses a great privacy risk due to the decentralized data collection. Moreover, training data on local datasets may not be very comprehensive when it comes to the matter of healthcare. Authors in [36] proposed a federated learning approach to only collect model updates instead of actual data by training the data locally. Although this model might achieve an appropriate accuracy, it can lead to inefficiency and security risks. They utilized edge computing to reduce communication overhead between the model and mobile devices. The model, called the Federated Edge Learning system (FEEL), also implemented differential privacy to better protect the data [36].

The authors demonstrated their proposed system by three major aspects as below:

- Mobile Healthcare Devices,
- Hospital Private Server,
- Cloud Data Center.

Figure 6 depicts the workflow of the FEEL system with respect to its design using three different modules. First, edge-based model training and health monitoring are performed by processing and analyzing the collected data. Then, they are sent to local hospitals to improve efficiency. Second, the local models are gathered in the cloud center in order to generate a global model. The global model is sent to local hospitals for several iterations to capture the data characteristics. Ultimately, the authors applied differential privacy to better secure the sensitive health data.

FEEL was evaluated using a breast cancer dataset to detect benignity or malignity of the cancer [53]. The data were distributed among 100 hospitals to assess the model's performance. Shortly, the model demonstrated appropriate results including efficient training, accurate diagnosis, and descent privacy protection [36].

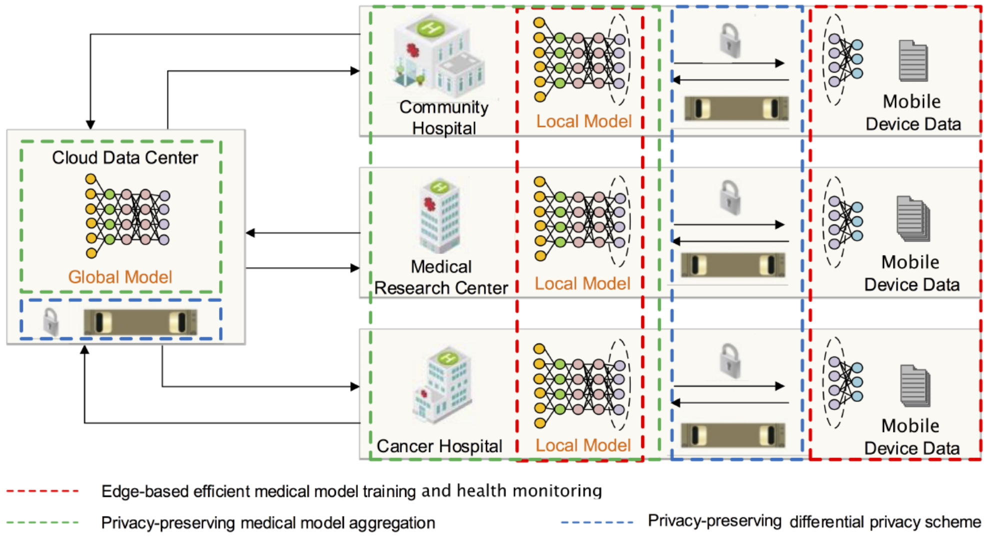

Figure 6. FEEL system implemented for mobile healthcare devices [36]. Mobile devices perform edge-based training, sending local model updates to hospital servers (e.g., Community Hospital, Cancer Hospital) via a differential privacy scheme. Local models are aggregated in the cloud data center into a global model, enabling efficient training, health monitoring, and privacy preservation across 100 hospitals.

2.7. DMFL-Net

The global impact of the COVID-19 pandemic as a significant healthcare crisis emphasizes the importance of timely detection to prevent widespread infection. Therefore, machine learning and deep learning methods are used on medical images like chest X-rays as a promising method for efficient COVID-19 detection and that of other chest diseases including lung cancer (LC), tuberculosis (TB), pneumothorax (PneuTh), and pneumonia (Pneu). However, privacy breaches in medical data raise concerns about using the centralized machine learning method in the healthcare system. To ensure medical data privacy, federated learning is a proposed solution that provides local training without sharing data for healthcare organizations. To address privacy issues, in [37], the authors proposed a decision-making-based federated learning network (DMFL-NET) framework for the classification of COVID-19-positive cases among various chest diseases such as lung cancer (LC), tuberculosis (TB), pneumothorax (PneuTh), and pneumonia (Pneu) using CXR images. The framework employs deep neural networks (DNNs) like DenseNet-169, VGG-16, and VGG-19 for feature extraction and classification, achieving 92.25% accuracy for COVID-19 detection (Table 1) across publicly available CXR datasets [37]. Similarly, recent advancements in FL for medical imaging have explored hybrid models to enhance multi-disease classification. For instance, Bilal et al. [54] proposed a hybrid model combining Extreme Learning Machines (ELMs) with quantum-inspired optimization for early multi-cancer detection, achieving high accuracies (96.98% for lung cancer) across diverse cancer types, demonstrating the potential of FL-based approaches in privacy-preserving medical diagnostics.

Figure 7 illustrates the DMFL-Net framework's process, detailing both server-side and client-side workflows. On the server side, the process begins with creating a task, collecting training times from clients, and updating waiting times to ensure all clients are synchronized before aggregating updates into a global model. On the client side, healthcare organizations download the task, train local models (e.g., on CXR data for COVID-19, LC, TB), evaluate model accuracy, and upload weights only if the local model outperforms the previous version, with a decision point to skip rounds if not. This visualization highlights DMFL-Net's decision-making mechanism for client selection and its use of waiting times to optimize communication efficiency, ensuring privacy through local training. However, the framework's reliance on DNNs may pose computational challenges for resource-limited clients, a potential limitation for broader adoption.

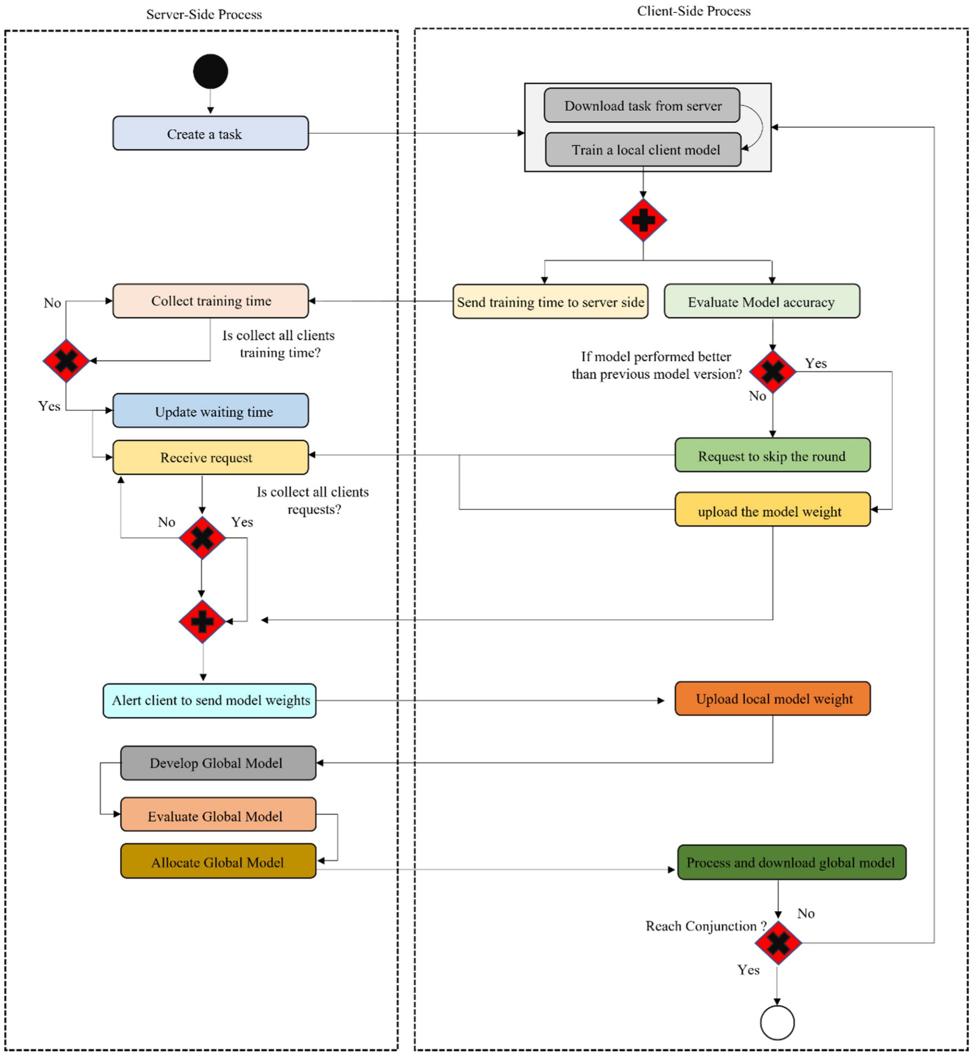

Figure 7. The overview of the DMFL-Net approach [37]. The server side creates a task, collects client training times, and aggregates updates into a global model after synchronization. The client side trains local models on CXR data (e.g., for COVID-19, lung cancer, tuberculosis), evaluates accuracy, and uploads weights only if performance improves, ensuring privacy, efficiency, and accurate chest disease classification.

The study presents a mechanism that considers the training time of each client. Clients are selected based on the performance of their local models, with updates sent to the server only when they improve the overall model. The central server calculates waiting times for each client based on the duration of their prior training cycles. The proposed method is evaluated in federated learning (FL) on aspects such as accuracy, recall, precision, F1-measure, specificity, and communication efficiency. The study primarily focuses on improving the identification of multiple-source chest X-ray images, fostering secure data exchange, and ensuring patient anonymity using the DMFL-Net framework. Additionally, deep neural network models, including DenseNet-169, VGG-16, and VGG-19, are used for efficient feature extraction and classification, specifically to differentiate COVID-19 from four other distinct chest disorders in publicly available CXR image datasets.

2.8. FedCare

The Internet of Medical Things (IoMT) devices are part of IoT systems that are capable of being deployed in the healthcare industry. These systems involve medical devices that autonomously establish communication and relationships while working collaboratively. IoMT devices aim to enhance real-time and remote health monitoring, which can be particularly beneficial in rural areas with a lack of healthcare centers and monitoring the elderly in the social system. These devices collect sensitive human body data, such as heart rate, blood pressure, and body temperature. In social systems, cameras are one of the most important IoMT devices for monitoring people. For instance, with the help of cameras, it is possible to obtain the vital signs of elderly people without requiring multiple sensors. However, these cameras collect users' facial videos and then train them at a central location or in the cloud, which could lead to privacy violations. To avoid privacy breaches and solve heterogeneity issues in federated learning, in [38], the authors proposed an FL-based IoMT framework in the healthcare environment to help rural people and monitor their health condition by camera-based IoMT devices. Figure 8 shows the architecture of the Fedcare framework for social IoMT devices.

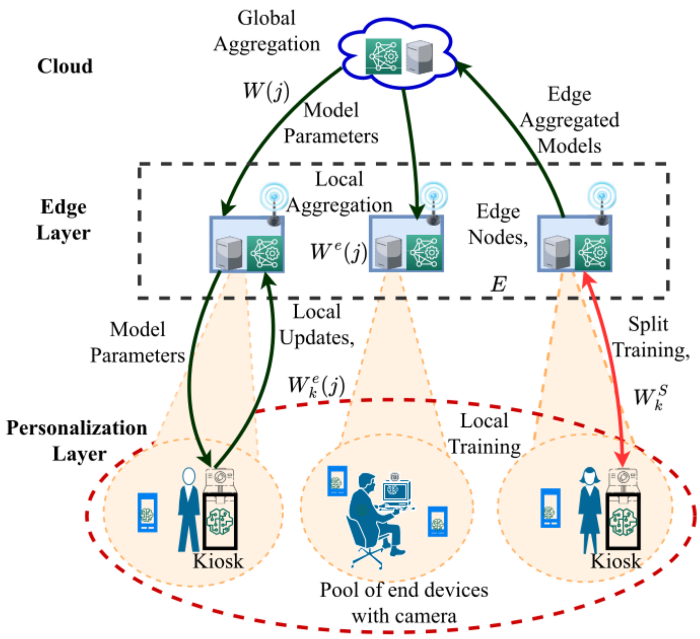

Figure 8. Fedcare framework for social IoMT devices [38]. IoMT devices (e.g., kiosks with cameras) in the personalization layer perform local training, sending updates $W_k^e(j)$ to edge nodes E. Edge nodes aggregate local models into $W^e(j)$, which are further aggregated in the cloud into a global model $W(j)$, using split training W_k^s to optimize resource use while ensuring privacy and efficiency for rural health monitoring.

In the FedCare system, each edge node selects a device for local training, which is responsible for data extraction. In the first phase, when the first round starts, each edge node selects the device for local training. During each round, the device remains connected to its node. Then, the central server, responsible for the global aggregation, identifies the node for this process, and another round begins. The performance of the proposed framework is evaluated, considering parameters such as accuracy, training time, CPU usage, memory consumption, and data rates. Remarkably, the FedCare framework achieves a global accuracy of 90.32%. Additionally, the proposed system significantly reduces training time, achieving efficiency with a concise duration of 3.6 hours.

2.9. Sensor-Based HAR

In [39], the authors propose a federated learning-based framework to address privacy concerns in the context of smart healthcare services utilizing the Internet of Medical Things (IoMT). The integration of wearable sensor-based devices connected to the Internet is vulnerable to security threats, particularly the risk of personally identifiable information being compromised. The paper introduces a federated learning approach that enables training models directly on-device data without centralizing it on servers to overcome the privacy issue. Also, the use of the bitwise XOR operator for data encryption is proposed to further safeguard biomedical data during transmission over the Internet. Figure 9 illustrates the high-level architecture of the sensor-based HAR framework, showing two servers (Server A and Server B) performing global averaging on encrypted updates from clients (Client A1 to An, Client B1 to Bm). Each client trains local models on sensor data (SGD), encrypts the data using the server's unique key (Key A or Key B), and uploads updates to the respective server. The servers cannot infer data from each other due to the distinct encryption keys, ensuring privacy during global model aggregation. This visualization highlights the framework's use of encryption to prevent unauthorized access and its ability to maintain privacy in IoMT applications, though the slight accuracy drop with FL suggests a need for further optimization in heterogeneous data settings.

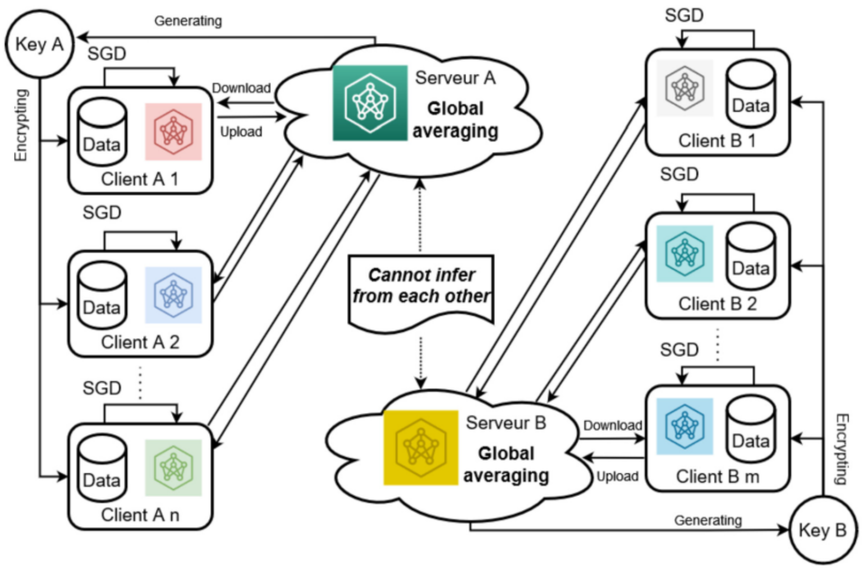

Figure 9. Architecture of sensor-based HAR federated learning approach [39]. Clients (A1 to An, B1 to Bm) train local models on sensor data using SGD, encrypt updates with server-specific keys (Key A, Key B), and upload them to Server A or Server B for global averaging. Distinct encryption keys prevent servers from inferring each other's data, ensuring privacy and security for physical activity recognition in IoMT applications.

The methods involve three-dimensional convolutional neural networks for physical activity recognition using various sensors. Also, the proposed encryption technique is then extended to both traditional federated learning and federated learning based on multi-key homomorphic encryption. Each server employs its unique encryption key to provide security for client data. This specific key prevents other servers from using the global model, which is trained on the original server, without the appropriate access key. Therefore, the hacker who obtains the trained model through networks cannot trace back the model and utilize it in a different database. Consequently, the proposed sensor-based HAR framework serves as an effective method to decrease the possibility of privacy breaches. Also, the experimental results show high accuracies of 94.6% and 94.9% (without federated learning) on the Sport and DaLiAC datasets, respectively. The proposed method slightly reduces accuracy to 89.5%, but it still shows promise compared to state-of-the-art methods using raw data alone.

3. Applications

Federated learning (FL) has emerged as a transformative technology, widely adopted across domains to enhance data privacy while enabling collaborative model training. In healthcare, FL addresses critical challenges by allowing institutions to train models locally without sharing sensitive patient data. Organizations such as Melloddy [55] and King's College London [56] have integrated FL into their workflows to improve privacy and predictive capabilities. This section surveys FL applications in healthcare, covering areas such as drug discovery, medical imaging, disease prediction, and more. Table 2 summarizes these applications, detailing their focus, approaches, techniques, datasets, and performance outcomes. To provide deeper insight, a critical assessment of their real-world applicability, maturity level, and scalability is presented at the end of this section.

Table 2. Applications of FL in Healthcare

Focus	Reference	Approach	Technique	Dataset
Drug	[57] [59] [61]	Cross-silo FL FL-QSAR Adverse drug reactions	DNN QSAR, HFL SVM, LM	AqSolDB [58] Kaggle datasets [60] LCED [62]
Mortality and stay time	[63] [64] [65]	CBFL Privacy of EHRs Data privacy	Encoder, K-means DP-SGD DP, LR, MLP	eICU [46] eICU [46] MIMIC-III [66]
Hospitalization	[67] [68]	COVID-19 Cardiac events	LASSO, MLP SVM, cPDS	MSHS [67] Boston Medical Center [69]
Preterm birth prediction	[70]	FUALA	RNN	Center Health Facts [70]
Brain segmentation	[71] [61] [74]	Whole brain segmentation Brain tumor segmentation Brain tumor segmentation	DNN FL, IIL, CIIL DP, DNN	MALC [72] BraTS [73] BraTS [75]
Functional MRI	[76]	Autism Spectrum Disorders	DP, MLP	ABIDE [77]
COVID-19 detection	[78] [68] [80]	CT scan Chest X-ray images Dynamic fusion	VGG, Resnet, etc. MobileNet, ResNet18, etc. GhostNet, ResNet50, ResNet101	CC-19 [78] COVIDx [79] CT , Radiography, Xray [81–83]
Medical records	[84] [86] [89]	Lung nodules detection Cardiovascular detection Thyroid image recognition	Vnet 3D, ResNet 3D-CNN DNN	LIDC [85] ACDC, M&M [87,88] Thyroid Nodule Clinical Data [89]
Patient similarity learning	[90] [91]	Privacy preserving Federated Patient Hashing	Hashing Hashing	MIMIC-III [66] MIMIC-III [66]
Phenotyping	[92] [94]	Privacy preserving Clinical data	Tensor Factorization, ADMM NLP, SVM	MIMIC-III, UCSD [66,93] MIMIC-III [66]
Communication overhead	[95]	Arrhythmia detection	DNN	PhysioNet 2017 [96]
Meta-analysis of brain data	[97]	PCA	ADMM	ADNI, PPMI MIRIAD, UK Biobank [98–101]

3.1. Drug Discovery

The healthcare industry is very dependent on the drug consortium, which attracts significant attention to it. In order to improve and perform better actions regarding the drug system, FL can be implemented. Two real time examples of implementing FL for drug purposes can be Melloddy [55] and King's College London [56], which aim to provide data privacy and predict recommended treatments, respectively.

3.1.1. Cross-Silo FL

Authors in [57] built a cross-silo FL framework with a database consisting of seven drug datasets. They intend to address the biased data issues, which can influence the accuracy of the training model due to the skewed data. This framework handles the model by having a coordinator server and collaborators who aim to support the FL client. These collaborative clients can include different data sources or silos including big pharmacies, biotech startups, and academic labs. The implementation of the model in this work consists of four steps as shown in Figure 10, which repeats during each round of training. First, the coordinator server transmits the last updated model to each one of the clients. Second, each client performs training locally and manages the updates, which is in general the local training. Next, in order to share the locally trained models with the server, each client encrypts its model and sends it to the server via a protocol. Lastly, after the models are received by the server, the global model is updated with regard to the local changes. The authors demonstrate good performance of the federated model regarding the significant biased values within their seven datasets. They evaluate their model by comparing it against a centralized method, which performs better in dealing with non-independent identically distributed (non-IID) data.

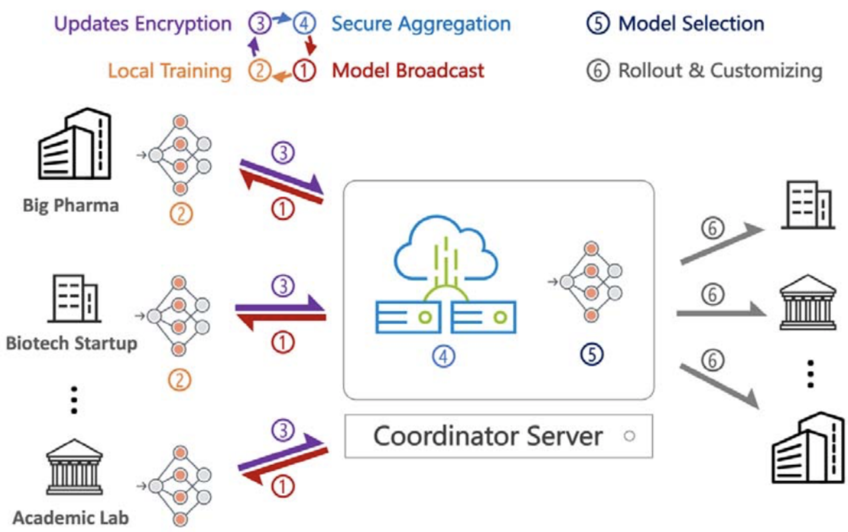

Figure 10. Federated learning for drug discovery [57]. The coordinator server broadcasts the model to clients (Big Pharma, Biotech Startup, Academic Lab) for local training. Clients encrypt updates, which are securely aggregated by the server, followed by model selection and rollout, ensuring privacy and addressing biased data challenges in cross-silo drug discovery.

3.1.2. FL-QSAR

Quantitative structure–activity relationship analysis (QSAR) has been applied for improving drug discovery. It is important that pharmaceutical institutions have a good collaboration together, which can enhance the QSAR performance. However, other factors including intellectual property can slow down this process. In short, QSAR has been applied to forecast different properties of compounds, which is considered an important initial step in drug discovery consortia. In the second paper, authors [59] proposed a new FL model, which is integrated with the QSAR called FL-QSAR. Within this platform, the horizontal FL architecture was applied on a database of fifteen datasets. Shaoqi Chen et al. provided results by making comparisons of different aspects, including comparisons between the horizontal framework and a traditional privacy-preserving framework, public and HFL framework and collaboration of the HFL and single client. The proposed study demonstrated the efficiency of FL-QSAR and HFL as a solution to the QSAR analysis challenges among different institutions.

3.1.3. Adverse Drug Reaction

Another research is proposed as a predictive approach for adverse drug reaction, which has been implemented using a FL framework. It uses the locally spread data over different institutions. This model can achieve a descent and efficient amount in each of the parameters including precision, recall, and accuracy [61].

3.2. Prediction

As the FL mission is needed to provide better integrity and privacy by training the model on the client side, it can be applied as a prediction methodology across various models in the healthcare domain.

3.2.1. Mortality and Stay Time Prediction

CBFL

When it comes to the healthcare domain, one of the most important parts is Electronic Medical Records (EMRs). They are used for different tasks such as disease prediction, a patient's response to treatment, etc. The EMRs have been analyzed and implemented by the traditional ML algorithms. Since EMRs are created by patients while they are in different institutions, storing EMRs in a centralized location is not applicable. Because of that, traditional ML algorithms are not a good fit for handling these data. As concerns are increasing regarding the security, privacy, and expenses, authors have applied FL. It is called community-based federated learning (CBFL), proposed for solving this issue. Authors have implemented this algorithm in order to predict mortality and hospital stay time of patients using drug features. Similar to the FADL framework, they have used the dataset of ICU records. To summarize CBFL, it basically makes predictions on the proposed dataset using a decentralized clustering with the implementation of FL. CBFL is more efficient than traditional FL approaches because in its implementation, EMRs are clustered into different communities. Moreover, each community has its own trained model. The main objective of this study is based on three parameters from the database, which are mortality, ICU stay time, and drug features. As an independent variable, drug features are used to predict other two parameters. Additionally, the evaluation of the model is performed by three different metrics including Area Under the Receiver Operating Characteristic Curve (ROC AUC), Area Under the Precision–Recall Curve (PR AUC), and communication leads. The model has shown a descent accuracy in prediction tasks regarding the non-IID challenge within a decentralized method compared to the centralized one.

Privacy of EHRs

Since the privacy of data is extremely important, authors in [64] proposed a federated averaging [26] framework with differentially private stochastic gradient descent (DP-SGD) [102] in order to predict mortality and hospital stay time. This study is performed on thirty one hospitals from an eICU database [46]. The authors demonstrated the benefits of their framework (FL with DP) using various experiments including different trainings (local, centralized, centralized with DP), and FL alone.

Data Privacy

Similarly, authors in [65] demonstrated superior performance of their FL framework compared to traditional approaches. The framework was implemented to predict patient mortality in hospitals. ICU data were used to train the global model. CoMind FL toolkit [103] was used to distribute training and testing data among the clients. Overall, maintaining data privacy can lead to high performance in FL applications.

3.2.2. Hospitalization Prediction

COVID-19

According to Vaid A. et al. [67], a FL technique was implemented in order to make predictions on the mortality rate of hospitalized patients who are infected with COVID-19. The electronic health records of patients including their past health records were used for evaluating the model. Classifiers such as Logistic Regression (LASSO) and Multilayer Perceptron (MLP) were deployed to train the dataset using three different distribution methods: local, pooled, and FL. The federated MLP and federated LASSO showed better performance compared to their local training counterparts.

Cardiac Events

A new framework has been developed to address the sparse Support Vector Machine problem. This framework enhances scalability and prevents the exchange of raw data within healthcare systems. Specifically, the authors proposed a FL model that predicts the likelihood of hospitalization for patients with cardiac diseases. This model uses Electronic Health Records distributed across various sources [69].

3.2.3. Preterm Birth Prediction-Federated Uncertainty-Aware Learning Algorithm (FUALA)

A large number of EHRs are being generated at each clinical institution everyday, and ML approaches have been widely developed for better analyzing data. However, due to privacy concerns and regulations, sharing data with the server for model training is inappropriate. As a result, FL enables secure training of models for prediction using EHRs, keeping data protected [18]. The concept of preterm birth has been critically affecting society by either babies' death or long-term disabilities. Moreover, it is costly to care and healthily deliver a baby [104]. In [70], the authors proposed a FL model to forecast preterm birth based on an EHR dataset from 50 hospitals. The prediction should be performed 3 months before delivery of the baby. The authors trained the models using Recurrent Neural Networks within FL setup. This approach improves the previously adopted FedAvg framework [26] to include uncertainty modeling. FUALA aims to predict preterm birth in two stages: assessing generalization performance and measuring the ultimate model. It demonstrates a more descent performance compared to other models.

3.3. Medical Imaging

Medical imaging is one of the valuable techniques for clinical analysis, which is widely adopted in the healthcare system. Privacy concerns in patient data have led to the deployment of FL, which enables models' training from different clinical institutions without data leakage [105].

3.3.1. Brain Segmentation

Medical segmentation is one of the most important approaches in medical imaging. In order to capture brain characteristics and measure its structure, brain magnetic resonance imaging (MRI) has been applied. Moreover, many different brain segmentation methods have been implemented [106]. For a comprehensive understanding of human brain segmentation methods using MRI, readers are encouraged to refer to the work in [106]. Since labeling the image data can be costly and time consuming, machine learning algorithms have been widely applied in different medical settings, especially in brain segmentation [107]. Although ML models have demonstrated promising results in medical imaging, they might not be the best fit due to limitations in the quality and variety of training data. Privacy regulations can be one of the reasons for this, preventing sharing patient data in a centralized manner. This issue can lead to the use of FL, which is compatible because it involves sharing only the model, not the data, and is decentralized.

Whole Brain Segmentation

A decentralized framework, called BrainTorrent, was proposed to facilitate peer-to-peer communications between multiple institutions [71]. As a result, the framework does not require a server for node communication, as depicted in Figure 11. Also, it trains the global model without sharing any raw data among institutions and allows for rapid updates to local models. The authors utilized the QuickNAT [108] architecture to segment brain scans within the network. Authors of [71] demonstrated BrainTorrent's efficient performance by comparing it with a traditional FL approach that uses a server. This comparison is based on results from two specifically defined experiments. Not only does the proposed approach eliminate the need for a server during training, but it also achieves similar performance to training on a pooled dataset aggregated from various institutions.

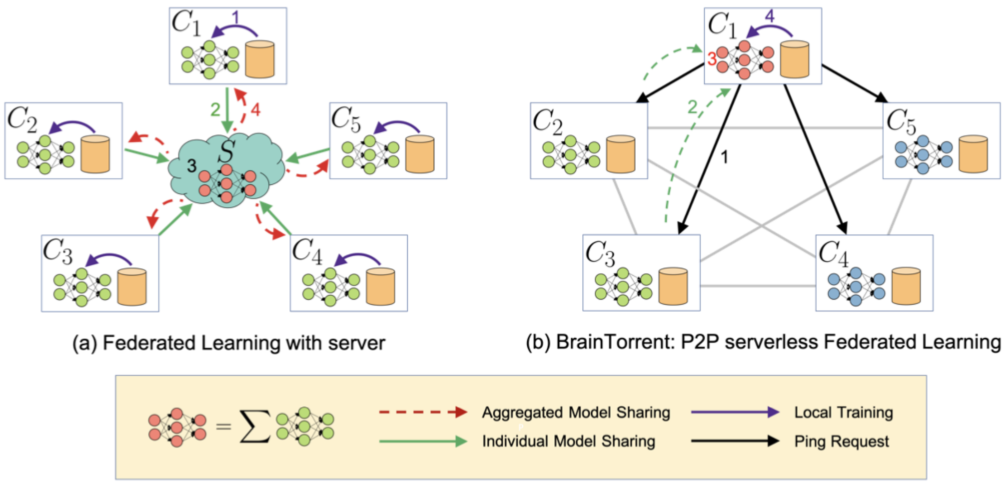

Figure 11. Peer-to-peer federated learning [71]. Part (**a**) shows traditional FL with a server S, where clients C_1 to C_5 send updates (red dashed arrows) for aggregation and receive the global model (green arrows). Part (**b**) depicts BrainTorrent's P2P serverless FL, where clients directly share models (green dashed arrows) after local training (purple arrows), using ping requests (black arrows) to coordinate, ensuring privacy and efficiency in brain scan segmentation.

Brain Tumor Segmentation

Research in [109] introduced the first use of FL for brain tumor segmentation, involving collaboration between various clinical institutions. The authors designed a deep learning (DL) model using FL in order to train various institutions without sharing data [109]. The model is trained on various brain images. The model's evaluation is based on comparisons between FL and two other collaborative learning methods: Institutional Incremental Learning (IIL) and Cyclic Institutional Incremental Learning (CIIL). In IIL, the client trains the model once according to its preference. CIIL essentially repeats the IIL process until the correct epoch number is achieved for each client. Overall, FL outperforms the proposed learning methods and achieves 99% model performance compared to when data are shared [109].

The authors in [74] proposed using differential privacy to secure patient data within a FL framework. They reduced the number of randomly selected parameters in each round. To evaluate the federated model, they compared it with centralized data training. It demonstrated superior FL performance due to shorter training times and the advantage of not sharing data.

3.3.2. fMRI—Autism Spectrum Disorders

As we discussed earlier, medical data are very sensitive when used for training models. As a result, data privacy is a major concern within the healthcare consortium. Various regulations and policies have been applied to the use of data based on characteristics including data content, identifiability, and more. These regulations can differ based on case-by-case scenarios [110,111]. Consequently, there is a lack of availability and quality in healthcare data. The authors in [76] proposed a privacy-preserving FL framework for functional MRI (fMRI) analysis, incorporating two domain adaptation methods to enhance neuroimage analysis. Gaussian and Laplace mechanisms have also been implemented to add noise and improve privacy [112]. The FL framework was compared across four different strategies. The main contribution of FL and domain adaptation was proposing a classification framework aimed at enhancing performance. It was claimed that the domain adaptation can enhance the performance of the FL framework. The authors aimed to detect Autism Spectrum Disorders or identify healthy controls. Overall, the proposed model can assist clinical institutions in securely training local models. This is particularly applicable when the diseases are rare and the number of patients is low.

3.3.3. COVID-19 Detection

The recent COVID-19 pandemic has resulted in a significant number of deaths, totaling 2,022,405 as of January 18 [113]. This has led researchers to apply various technologies to combat this infectious disease. As a result, surveys on the applicability of these technologies, including AI [114], ML [115], Big Data [116], and IoT [117], have been conducted. In this discussion, we explore various approaches to detecting COVID-19.

CT Scan

One of the most applicable approaches of medical imaging is the use of Computed Topography (CT) scans. As of the time of writing this paper, CT scans have been widely used for diagnosing the novel Coronavirus, which causes COVID-19 disease. Millions of people died due to this virus. Although machine learning analyses of COVID-19 patient CT scans have been widely studied, the importance of data privacy in these analyses must be considered. Authors in [78] have developed a framework for detecting COVID-19 using deep learning (DL) models within the CT scans, which can assist practitioners in faster detection of COVID-19. In deep learning model training, data leakage is a major concern. The study involves a large dataset from various institutions, comprising records of 89 patients (68 positive and 21 negative cases). Within this approach, they deployed various DL models for recognizing the patterns in patients' lung CT images. The training was performed using a Capsule Network, which resulted in better performance. In addition to preventing leakage of the data, FL was used. It facilitated data capture, model training, and sharing on a public network in a decentralized manner. This approach allows hospitals to share only the model weights and gradients via blockchain, not the patients' data. This study demonstrated better performance due to the better accuracy. Moreover, FL enhanced data privacy across various institutions.

Chest X-Ray Images

The medical fields, especially medical imaging, have been positively affected by deep learning techniques. However, these techniques can lead to significant data privacy leaks, particularly when collecting training data from various institutions. Given the critical need for COVID-19 detection, deploying deep learning within a federated learning framework can significantly enhance data privacy. According to [68], a FL framework was proposed for training data, with performance demonstrated by comparing four different image classification models: COVID-Net, ResNet18, ResNeXt, and MobileNet-v2. These models are implemented with and without FL. Within the implementation, ResNet18 achieved the highest accuracy across both training and testing datasets when using federated learning. Regarding the dataset's labels (which indicate the patient's condition, including normal, pneumonia, and COVID-19), ResNeXt showed the best performance in detecting COVID-19.

Figure 12 illustrates the FL framework for COVID-19 detection, showing a central server coordinating with multiple institutions (e.g., medical clinics, hospitals). Each institution trains local models on CXR datasets behind a firewall, ensuring data privacy, and sends parameter updates to the central server. The server aggregates these updates and broadcasts the updated global model back to the institutions for further training. This visualization highlights the framework's ability to maintain privacy by keeping data local while enabling collaborative model training for accurate COVID-19 detection, though its performance may vary with the heterogeneity of CXR data across institutions.

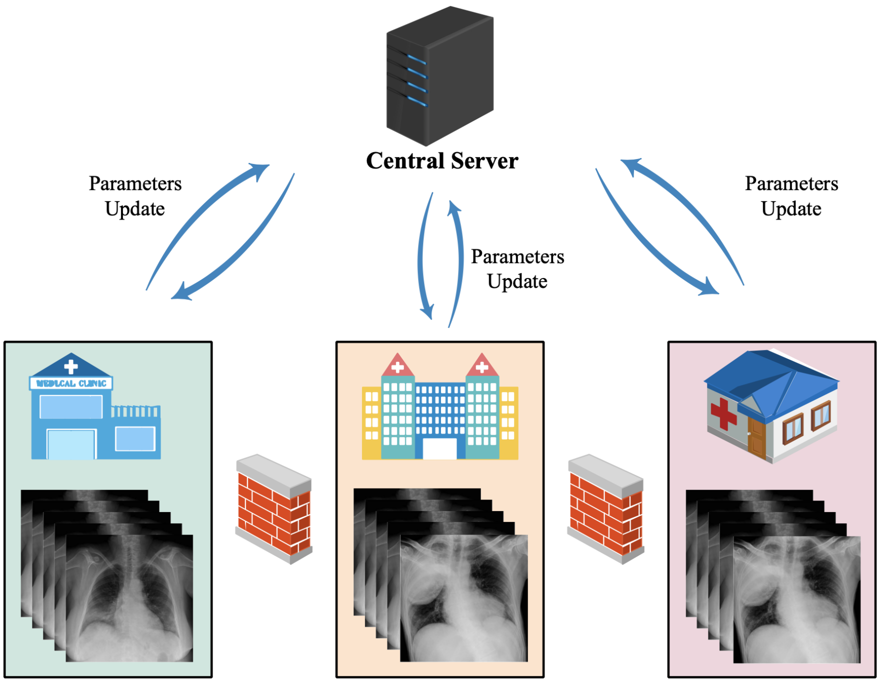

Figure 12. Federated learning for COVID-19 detection [68]. A central server coordinates with institutions (e.g., medical clinics, hospitals) which train local models on CXR datasets behind firewalls and send parameter updates. The server aggregates updates and broadcasts the global model, ensuring privacy and enabling collaborative training for accurate COVID-19 detection.

Dynamic Fusion

The two aforementioned examples were implemented in the default setting of FL, which can result in high communication costs. Since this can be a critical issue affecting model performance, a fusion-based FL approach has been proposed to detect COVID-19 cases more efficiently and with better communication [80]. The authors initially proposed a fusion-based FL architecture for detecting COVID-19 cases. Subsequently, they applied a dynamic fusion technique to manage participating clients. The trained local model does not upload any updates until it achieves improved performance. To evaluate performance, experiments were conducted using three different Convolutional Neural Network models: GhostNet, ResNet50, and ResNet101. The results demonstrated improved accuracy, reduced training times, and better communication performance compared to the default setting of FL.

3.3.4. Lung Nodules

With respect to the capabilities of CT scan images, authors in [84] proposed a FL approach for detecting lung nodules. The approach is based on two models: lung nodule detection and confirmation. By evaluating the model using a dataset of 1010 CT images, the authors demonstrated more time-efficient method for maintaining privacy of data in a decentralized matter of model distribution. In comparison to the existing models, the proposed approach indicated better prediction accuracy.

3.3.5. Cardiovascular Disease Detection

Cardiovascular disease diagnosis in the medical field is challenging because of data scarcity. The lack of sufficient data raises concerns about the generalizability of automated diagnosis studies across institutions. To solve these issues, federated learning can facilitate multi-center studies by enabling distributed training while protecting patient privacy. Authors in [86] presented a federated learning framework focused on cardiovascular magnetic resonance (CMR) imaging. They used data from four centers derived from M&M and ACDC datasets to detect hypertrophic cardiomyopathy. The study used a 3D-CNN network to train data along with various data augmentation techniques. Although the study used a small dataset (180 subjects from four centers), the experimental results demonstrated that the proposed federated learning framework achieved results similar to those of traditional centralized learning.

3.3.6. Thyroid Image Recognition

The author in [89] investigates the utilization of federated learning to analyze ultrasound images to determine whether thyroid nodules are benign or malignant. The aim of this study is to show whether the performance of a federated learning framework is comparable with centralized learning techniques while simultaneously preserving the privacy of medical data. They use a dataset which contains 8457 ultrasound images from six institutions. The dataset is used for both federated learning and centralized deep learning models. ResNet 50, VGG19, ResNext50, SE-ResNext50, and SE-ResNet50 models are employed as base models to compare with the federated learning model. Internal validation is conducted on a subset of images, while external validation is performed on images from another institution. The experimental results show that for internal validation, the area under the receiver operating characteristic (AUROC) curve ranges from 78.88% to 87.56% for federated learning and from 82.61% to 91.57% for centralized deep learning. For external validation, AUROC ranges from 75.20% to 86.72% for federated learning and from 73.04% to 91.04% for centralized deep learning. Based on the results of this study, federated learning appears to achieve comparable performance to centralized deep learning while employing decentralized data, suggesting that it could be useful for medical image analysis while safeguarding the privacy of patients.

3.4. Patient Similarity Learning

Adopting global frameworks is one of the principal approaches in order to analyze different institutions comprehensively without privacy concerns. However, among the proposed FL frameworks in healthcare, only a few approaches exist that address patient similarity. In general, it is based on developing algorithms in order to discover specific similarities in medical datasets or patients [118].

Privacy Preserving

Although patient similarity search can be tough to discover due to the issues of captured data, the authors in [90] proposed a privacy-preserving FL framework to efficiently find similar patients across different institutions. The implementation of their model uses the hashing technology. It enables discovering hash codes of patients' information across various hospitals. Authors performed the experiments based on two hashing technologies including multi-hash and uni-hash. The results demonstrated better usability for the multi-hash approach. Additionally, homomorphic encryption was applied to prevent security concerns while the FL system looks for similarities. This approach was evaluated based on the Multiparameter Intelligent Monitoring in the Intensive Care-III (MIMIC-III) database to indicate the performance of the algorithm within five different categories of diseases [66]. To recap, the proposed algorithm is a solution for efficiently finding the patients' similarities among different hospitals with respect to the privacy preservation within the FL system [90].

FPH

According to the study in [91], another framework for patient similarity search was proposed in a federated setting based on hashing technology. This framework, which is called Federated Patient Hashing (FPH), demonstrates patient similarity search among different institutions without sharing private data. In order to optimize federated setting, two learning methods, namely centralized and decentralized, have been adopted. Additionally, authors demonstrate the convergence of the model by depicting same behavior of different training samples. Similar to the previous research, authors in this study use the MIMIC-III database to evaluate the framework, which results in good performance of the model [66,91].

3.5. Phenotype Discovery

As previously mentioned in the above sections, the EHRs deal with different challenges such as skewed data, heterogeneity, etc. Data scientists usually prefer to work on concise concepts in the healthcare domain rather than complicated data. This is where phenotypes become involved in capturing specifically related details of patients. Phenotype is basically a disease and its subtypes. Moreover, the challenge of merging EHRs into the phenotypes still exists due to its limitations including slow and manual adoption [119]. Tensor factorization is an appropriate method to overcome different challenges that are applied to the phenotypes by capturing them from complex datasets [92,120]. However, in order to capture widely existing phenotypes, it is important to compute them on a dataset of multiple clinical institutions. In that case, data privacy is endangered due to the data sharing. Moreover, some approaches suggest adding noises to the models, which is not applicable for phenotypes. These noises, or technically differential privacy [121], can mislead clinical scientists with dangerous results. Herein, we discuss two different approaches with respect to privacy concerns.

3.5.1. Federated Tensor Factorization

Authors in [92] proposed a federated framework that aims to preserve privacy by using tensor factorization for horizontally partitioned data. This approach, TRIP, is developed for computing phenotypeswithout sharing patient data. After formulating the function of this framework, the authors attempted to solve the optimization problem using the ADMM (alternating direction method of multipliers) [122] algorithm. Within the study, ADMM divides the main problem into sub-problems for each institution. This process is performed by repeatedly updating the local factor matrices from individual components. As soon as all the local factors are updated within hospitals, the global factor is updated and sent to the hospitals. After a specific number of iterations, the framework's performance is evaluated within two different settings of central and local models. Accuracy and time are the major measurements of performance, which can be impacted by two issues such as nonzero values and skewed data. Results demonstrated an accurate performance even with small and uneven distribution of data. Consequently, TRIP can be deployed in large datasets for discovering phenotypes regarding the existing privacy concerns [92].

3.5.2. Clinical Data

Manual extraction of medical information by medical experts could be very time consuming. Moreover, natural language processing (NLP) has been widely adopted by scientists to better analyze medical data. As a task within the clinical NLP, automatic phenotyping has been used for capturing specific patients. Authors in [94] implemented a federated machine learning framework in order to address two important challenges in medical NLP. First, data sharing among different institutions for better model training is not applicable. In addition, mapping the medical information using patient representation learning or better classifier training also requires a massive dataset. This is shown in previously adopted models, including supervised [123] and unsupervised [124]. The framework adopts two stages of supervised patient representation learning and phenotyping extraction of clinical notes. The first one enables the applicability of training an Artificial Neural Network (ANN) on large datasets. Additionally, phenotyping captures the functions within the medical information, which enables disease prediction. Various experiments with different settings were conducted in the study, and it was found that the federated setting for both patient representation and phenotyping demonstrated appropriate performance [94].

3.6. Arrhythmia Detection

The IoT technology has been widely applied to the healthcare domain within different areas including health monitoring, chronic diseases, etc. [125]. Traditionally, machine learning algorithms have been used in order to train sufficient models on the server. Therefore, IoT devices should upload their data to the server, which brings privacy and security concerns. According to [126], the new technique of FL for training in a decentralized manner demonstrated different challenges including source of the energy for devices, low computational capacity, etc. Yuan et al. [95] proposed a FL framework for detecting arrhythmia disease that achieves better accuracy and reduces the communication bandwidth. The detection is achieved by monitoring cardiac arrhythmias captured by IoT devices.

3.7. Large-Scale Medical Data

Nowadays, a significant number of MRI images are being stored across various data centers, which have to be analyzed in order to better understand different diseases. However, sharing patient data from various data centers is not applicable due to the privacy concerns and regulations. Consequently, adopting different analysis methods including FL and meta-analysis [127] has enabled researchers to better work on data. Authors in [97] proposed a FL framework based onmedical imaging data in order to access and analyze the data without risking privacy. This study is mainly focused on comprehension of brain disorders with respect to the limitations of data exploitation. This approach implemented the ADMM [128] algorithm that decreases the number of iterations in order to prevent gradient-based optimization. Synthetic data were first adopted for evaluation, and then multiple medical datasets were deployed. Overall, the framework met the intended aims.

3.8. Critical Assessment

The FL applications surveyed in this section demonstrate significant potential in healthcare, yet their real-world applicability, maturity level, and scalability warrant further discussion to contextualize their practical impact.

Real-world Applicability: Deploying FL in healthcare settings faces practical challenges. For instance, drug discovery applications like Cross-silo FL [57] and FL-QSAR [59] rely on collaboration across institutions (e.g., pharmacies, labs), which requires standardized protocols, robust network infrastructure, and compliance with regulations like HIPAA or GDPR. Medical imaging applications (e.g., BrainTorrent [71], COVID-19 detection [78]) are promising for multi-center studies, but their adoption may be hindered by the need for compatible hardware (e.g., CT scanners, MRI machines) and trained personnel. Prediction tasks (e.g., CBFL for mortality [63]) and patient similarity learning [90] could integrate into hospital workflows, yet their effectiveness depends on the availability of high-quality, diverse EHR datasets, which vary widely across regions.

Maturity Level: The maturity of these applications varies. Some, like BrainTorrent [71] and FL-based COVID-19 detection [68], have been tested on real datasets (e.g., MALC, COVIDx) and show performance comparable to centralized methods, suggesting a transition from experimental to near-deployable stages. Others, such as FL-QSAR [59] and phenotyping via tensor factorization [92], remain largely theoretical or prototype-based, with limited evidence of real-world deployment. Arrhythmia detection [95] and thyroid imaging [89] show promise with IoT and ultrasound data, but their maturity is constrained by small-scale evaluations (e.g., PhysioNet, 8457 images), requiring broader validation.

Scalability: Scalability is a critical factor for FL's success in healthcare. Applications like Cross-silo FL [57] and dynamic-fusion COVID-19 detection [80] address communication overhead, making them scalable to larger networks of institutions, though computational costs rise with data heterogeneity. Brain segmentation [74] and cardiovascular detection [86] scale well with imaging data but may struggle with diverse hardware or small datasets (e.g., 180 subjects in [86]). Patient similarity [91] and phenotyping [94] frameworks handle large EHRs (e.g., MIMIC-III), yet their scalability is limited by non-IID data and processing complexity. Arrhythmia detection [95] offers bandwidth efficiency, but IoT device constraints (e.g., energy, computation) pose scalability challenges.

Overall, while these FL applications showcase innovative solutions to privacy and collaboration challenges, their real-world deployment requires overcoming infrastructural, regulatory, and validation hurdles. Future work should focus on standardizing frameworks, expanding testbeds, and addressing scalability bottlenecks to fully realize their potential in healthcare.

4. Challenges

While federated learning (FL) offers numerous advantages, such as decentralized training and enhanced privacy, it faces significant limitations and challenges, particularly within the healthcare industry. These challenges must be carefully addressed when designing and implementing federated training models. In this section, we explore key challenges and highlight potential strategies and ongoing research efforts aimed at overcoming them.

4.1. Communication Efficiency

As we previously discussed, a FL setting occurs when the training data are distributed over a significant amount of clients. The clients are not reliable, and they most likely do not have a fast connection. The distribution of training data divides communications into two parts, namely uplink and downlink [129]. Since different domains especially healthcare are in need of a cost-efficient setting for distributing training data and local models, there have been many efforts to enhance the communication efficiency. With that being said, simply the size of messages has to be reduced. Here, the messages are either the training datasets or local models [130]. In order to enhance the efficiency of communication, there are three segments that can be reduced including clients' numbers, updates' numbers, and size of updates. Based on the recent research, these three techniques have been addressed using four types of methods, as follows (see Figure 13).

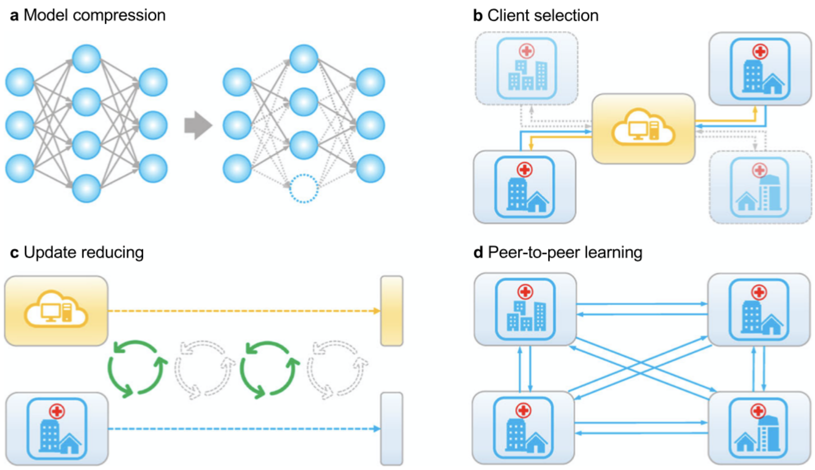

Figure 13. Methods for reducing bandwidth [28]. (**a**) Model compression simplifies neural networks to reduce variables. (**b**) Client selection chooses a subset of clients (e.g., hospitals) to participate. (**c**) Update reduction decreases the frequency of updates between clients and the server. (**d**) Peer-to-peer learning enables direct model exchange among clients, omitting the central server, to enhance communication efficiency in FL for healthcare applications.

4.1.1. Client Section

Despite the fact that selecting clients is one of the most effective factors in improving efficiency, it is also the most accessible method to decrease the bandwidth of communication. This means the communication costs can be reduced by lowering the number of clients or updating a specific parameter.

4.1.2. Model Compression

As the name suggests, the model mainly works towards compressing the data that are exchanged from the server to the client. This reduces the communication bandwidth and consequently the communication cost. The first compression method is based on structured updates. This means the model attempts to use fewer variables. There are different structured updates such as low rank and random mask. Further, a sketched update occurs when the update is learned by the model, encoded or simply compressed for sending to the server. It is decoded thereafter by the server [129].

4.1.3. Update Reduction

As mentioned before, reducing the number of updates can effectively decrease the communication costs. For example, a research has been proposed in order to reduce the communication costs of large distributed systems by averaging different models [131].

4.1.4. Peer-to-Peer Learning

One principle of FL is the need for a central server for distributing training data towards achieving a global model. Also, since the available networks are dynamic, accessing a specific central server is not practical. In addition, due to the high dependency on the central server, a trusted centralized model should be agreed upon by different clients. This can be a complex approach to implement. Therefore, implementing a fully decentralized model can be an appropriate solution in order to omit the central server [28].

Recent research has proposed several strategies to further enhance communication efficiency. For instance, gradient compression techniques, such as quantization and sparsification, have shown promise in reducing the size of updates without significantly compromising model accuracy [132]. Additionally, ongoing efforts explore adaptive communication protocols that dynamically adjust the frequency and size of updates based on network conditions and client reliability [133]. These approaches aim to balance efficiency with performance, a critical consideration for resource-constrained healthcare settings.

4.2. Privacy

It is always assumed that the number of clients in a FL setting is large. This indicates the potential presence of malicious clients. As a result of training data locally, there should not be any data leakage within this setting, but clients might gain access to the sensitive data from the shared model. Despite the fact that privacy has been vastly researched within FL settings, it is crucial to maintain a good outcome regarding the different aspects of model including accuracy, performance, etc. This can be one of the most important attempts to address privacy issues [130]. Consequently, here, we provide a brief description of the most important aspects when ensuring privacy in the FL model.

4.2.1. Performance

Within the application of FL, data privacy can be improved by only sharing the models and not the data. However, there are many other privacy concerns that have to be addressed in different domains particularly in healthcare. Various regulations are defined in order to prevent data leakage, which make ensuring data privacy even more challenging. It is because there is not a solution that can satisfy all the aspects. Some techniques have been deployed to ensure data privacy including privacy-preserving techniques and secure computations [134]. However, a trade-off here is that performance, such as accuracy, may decrease by applying these techniques including differential privacy (DP) [135] and secure multi-party computation [136]. It is crucial to ensure that this does not influence other aspects.

4.2.2. Level of Trust

Since medical data are highly important, data leakage has to be prevented. Without trust in the captured data, the healthcare system cannot count on the performance of the model [137]. To divide the trustworthiness of different parties within FL, there are two levels of trust, as follows.

Trusted

A trustworthy FL system can be considered when all parties agree to collaborate. As a result, the system is capable of preventing any malicious attempts for accessing sensitive information. This lowers the threats to the system, which means reducing the demand for countermeasures.

Non-Trusted

When it comes to a large number of parties, for instance, multiple institutions, it is unrealistic to establish an agreement on collaborating in a same manner. Parties may try to extract others' sensitive information, shut the system down, etc. With this scenario, it is important to implement security strategies to handle adversarial attacks effectively.

4.2.3. Information Leakage

Despite the use of FL for training medical data by maintaining the privacy, the model can contain private information that is used for local training. This can occur if the model updates, gradients, and adversarial attacks are reversed [138–140]. Additionally, leakage can be increased due to the captured changes in model updates by adversaries. In order to prevent this issue in the system, counter-measures (for instance adding noise inside the data) should be applied to ensure privacy preserving [74].

To mitigate privacy risks, ongoing research has focused on advanced cryptographic techniques and differential privacy enhancements. Homomorphic encryption, for example, allows computations on encrypted data, preventing sensitive information exposure during model updates [141]. Similarly, federated learning frameworks incorporating personalized differential privacy (DP) adjust noise levels based on individual client data sensitivity, improving the privacy–utility trade-off [142]. These strategies aim to safeguard healthcare data while maintaining model performance, though their computational overhead remains a challenge under active investigation.

4.3. Data Skewing

The data in real life are not always perfect, and it is a critical factor in training algorithms. Data skewing is primarily related to the distribution of the data for training. It can be caused by different factors, but mainly occurs due to missing classes, features, and values [143]. In addition, Imbalanced data are also a major factor contributing to data skewing in training distributions. With respect to these issues, the traditional FL methods can fail if they they are not taken into account. Authors in [143] have shown practical approaches in different machine learning algorithms in order to address the data skewing issues. In short, the server plays an important role in implementing the proposed techniques, such as limiting the training data exchange and bounds-aware fusion.

Addressing data skewing is an active area of research. Techniques such as data augmentation and synthetic data generation are being explored to balance skewed distributions across clients [144]. Moreover, federated optimization algorithms, like FedProx, introduce regularization terms to handle non-IID (non-independent and identically distributed) data, improving convergence and model robustness [145]. These strategies aim to ensure equitable model performance across diverse healthcare datasets.

4.4. Traceability

The importance of reproducibility is obvious in all algorithms used in healthcare. It means that a system should be able to keep track of different parts including events, history of accessing data, and changes made to the configuration. This can be applied during training processes [29]. One of the advantages of traceability is that overlapping between the training dataset and the testing dataset can be prevented by considering the history of previously trained data in the model [146]. As we previously discussed regarding the non-trusted systems, traceability requires integrity for better tracing different activities. It is also worth tracking the history of consumed resources and the contribution of each institution after achieving the optimal model.

4.5. System Architecture

Healthcare institutions are capable of computing large scale FL models due to their computational equipment. They usually have reliable throughput networks, which allows them to perform various approaches for achieving a highly consistent model. An example of approaches is the training of large models with large shared local models [147]. Within healthcare, the ability to perform larger training models can lead to various issues including data integrity, security of data transferred to resources, and design of node schedulers [29].

4.6. Data Heterogeneity

Heterogeneity itself can be described as a set of dissimilar contents. Similarly, data heterogeneity is the diversity of variable data distributions. With respect to the healthcare domain, this challenge occurs often due to the assortment of patients' captured data including diagnosis, treatments, etc. [148,149]. Authors in [150] proposed a FL framework to better train the model for handling the heterogeneous distribution of data among clients using optimization. Moreover, demographics should be considered. Overall, data heterogeneity can pose a challenge within the model. The distribution of data in FL models is often independent and identically distributed [151]. Additionally, it is critical to note that the global solution may differ from the final solution, which requires all the clients to agree on the proposed training before training the model [29].

4.7. Scalability

As FL systems expand to include more clients or larger datasets, scalability becomes a critical challenge, particularly in healthcare where institutions vary in computational resources and data volume. Managing an increasing number of clients can strain communication networks and central servers, leading to bottlenecks and delays. Furthermore, training larger models on heterogeneous hardware introduces complexity in resource allocation and synchronization. Recent efforts to address scalability include hierarchical FL architectures, where intermediate servers aggregate updates from clusters of clients before forwarding them to a global server [152]. Additionally, asynchronous FL approaches allow clients to update models independently, reducing synchronization overhead [153]. These strategies aim to support the growth of FL systems while maintaining efficiency and accuracy, though their application in healthcare requires further validation.

4.8. Interoperability

In healthcare, FL systems often involve collaboration across institutions with diverse data formats, protocols, and regulatory requirements. Interoperability—the ability of different FL systems to work seamlessly together—poses a significant challenge. Lack of standardized frameworks can hinder model sharing and aggregation, limiting the potential for large-scale federated networks. To overcome this, ongoing research explores the development of standardized FL protocols, such as the OpenFL framework, which facilitates cross-system compatibility [154]. Additionally, ontology-based approaches aim to harmonize heterogeneous healthcare data, enabling consistent interpretation across clients [155]. Enhancing interoperability could unlock broader collaboration, though it requires consensus on standards and robust testing across real-world settings.

5. Future Directions

In this section, we discuss future directions to motivate researchers to address the challenges identified earlier, offering detailed and actionable recommendations for those entering the field of federated learning (FL) in healthcare. These directions aim to bridge current gaps and advance practical implementation.

5.1. Privacy and Security Issues

Security and privacy are critical concerns in the healthcare sector, where protecting patient data is paramount. While methods such as differential privacy and cryptographic techniques offer mechanisms to safeguard data, they introduce significant trade-offs, such as reduced accuracy and increased computational demands on healthcare devices involved in FL. The design of FL systems must therefore balance privacy assurance with system efficiency, a challenge that requires comprehensive solutions before practical medical deployment. One promising approach involves edge computing, which enhances data privacy and computational efficiency by processing data locally, reducing the need to transmit sensitive information across networks [156].

Future Research Questions: How can edge computing be optimized to minimize computational overhead while maximizing privacy in resource-constrained healthcare settings? What lightweight cryptographic protocols can be developed to secure FL without sacrificing model performance? Researchers could explore hybrid privacy-preserving frameworks that combine edge-based processing with adaptive differential privacy tailored to medical data sensitivity.

5.2. Communication Cost

Effective communication is crucial in FL due to its distributed nature and the need for multiple rounds of information exchange between clients and the central server to train a global model. In healthcare environments with numerous geographically dispersed clients, this can strain bandwidth and lead to bottlenecks, especially when network connections are unreliable. Techniques like gradient compression, sparsification, and quantization reduce communication cycles and speed up preprocessing [157], but their scalability across large networks remains underexplored.

Future Research Questions: How can communication-efficient techniques be scaled to support thousands of healthcare clients without compromising model convergence? Can adaptive compression algorithms dynamically adjust to fluctuating network conditions in real-time? Investigating scalable communication protocols, such as federated distillation or hierarchical aggregation, could provide practical solutions for large-scale healthcare FL deployments.

5.3. Heterogeneity Data

Medical data heterogeneity, arising from diverse sources like imaging, EHRs, genomics, and wearables, poses a significant challenge in FL's distributed setting. Increased heterogeneity often degrades performance, necessitating innovative solutions. Participant selection ensures maximum update aggregation by choosing clients strategically [158], while personalized FL, such as Federated Multi-Task Learning (FMTL), adapts models to client-specific data characteristics [159].

Future Research Questions: How can participant selection algorithms be optimized to balance computational load and data diversity in real-time? What meta-learning techniques can enhance personalized FL to handle extreme heterogeneity in healthcare datasets? Future work could focus on developing robust FMTL frameworks that dynamically adjust to client-specific nuances while maintaining global model integrity.

5.4. Scalability

As FL systems grow to include more clients and larger datasets, scalability emerges as a critical barrier, particularly in healthcare where institutions vary in computational resources. Hierarchical FL architectures [152] and asynchronous updates [153] address some issues, but their practical limits in massive networks are unclear.

Future Research Directions: Researchers should investigate how hierarchical FL can be adapted for dynamic healthcare networks with fluctuating client participation. What are the trade-offs of asynchronous FL in terms of accuracy versus scalability in real-world medical settings? Actionable steps include designing scalable FL simulators to test these approaches under realistic healthcare conditions and exploring resource-aware scheduling algorithms to optimize client contributions.

5.5. Standardization

Interoperability challenges in healthcare FL stem from diverse data formats, protocols, and regulations across institutions. Standardized frameworks like OpenFL [154] and ontology-based data harmonization [155] offer potential solutions, but consensus on standards remains elusive.

Future Research Questions: What standardized FL protocols can be universally adopted across healthcare institutions with varying regulatory requirements? How can ontology-based approaches be streamlined to reduce integration complexity? Researchers could develop open-source standardization toolkits and testbeds to facilitate cross-institutional collaboration, ensuring practical interoperability in FL deployments.

5.6. Model Explainability

While FL enhances privacy and scalability, the interpretability of its models crucial for healthcare decision-making remains underexplored. Black-box models may hinder trust and adoption by clinicians, especially when local updates obscure global model behavior [160].

Future Research Directions: How can explainable AI (XAI) techniques, such as SHAP or LIME, be integrated into FL to provide interpretable insights at both local and global levels? Can federated explainability frameworks balance privacy with transparency? Researchers should prioritize developing lightweight XAI methods tailored for FL, testing their efficacy in clinical settings to enhance trust and regulatory compliance.

5.7. Integrating FL with Emerging Technologies

Integrating federated learning (FL) with emerging technologies, such as blockchain, edge computing, and the Internet of Things (IoT), offers significant potential to enhance its capabilities. Blockchain, for example, with its decentralized nature, enables secure and transparent data exchanges in FL and eliminates the need for a transfer data to a central server while protecting against malicious attack. Along with the ability to improve security and scalability in FL, blockchain also ensures data integrity. It decreases latency through local data processing at local data devices or in edge devices [161]. Edge computing is another emerging technology that can integrate with FL for local processing. The ability of edge computing to reduce latency and preserve privacy is also a critical advantage, particularly in latency-sensitive applications like healthcare and smart cities [162]. Optimizing FMTL frameworks to balance global model performance is one of the critical areas for future research.

Future Research Questions: How can FL be seamlessly integrated into legacy health IT systems without disrupting workflows? What middleware solutions can bridge FL with existing EHR platforms while ensuring data security? Practical research could focus on designing plug-and-play FL modules and conducting pilot studies in hospitals to validate integration feasibility.

5.8. Hyperparameter Optimization

Hyperparameter selection plays a pivotal role in optimizing knowledge aggregation algorithms. However, in the context of FL, hyperparameter optimization presents novel challenges and remains a significant area of ongoing research. In FL, the optimal hyperparameter configuration may differ for each participant, depending on their data distribution. Indeed, it is crucial for each participant to optimize their hyperparameter settings based on their specific data properties [163]. One potential method for achieving hyperparameter optimization is through an auction mechanism. This approach can be implemented by offering incentives to clients, encouraging them to contribute to the hyperparameter optimization process through an auction mechanism [164].

Future Research Directions: How can decentralized hyperparameter optimization be scaled across heterogeneous healthcare clients? Can auction mechanisms ensure equitable participation without bias toward resource-rich clients? Future work could explore automated hyperparameter tuning frameworks using reinforcement learning or Bayesian optimization, validated in diverse medical FL scenarios.

6. Conclusions

In the smart healthcare domain, machine learning methods for disease detection have become increasingly prevalent, which raises profound concerns about the privacy and security of patient data. Due to federated learning's non-sharing of medically sensitive data, FL is emerging as a promising decentralized learning framework for medical applications. In this paper, we provide a systematic review of the growing use of federated learning (FL) in the area of smart healthcare. This paper comprehensively reviews the federated learning-based architecture associated with a smart healthcare environment. Also, the paper presents the application of federated learning in various fields such as disease prediction, drug discovery, treatment, and medical imaging. Moreover, we discuss the challenges of using federated learning and define some future research directions.

Author Contributions: Conceptualization, M.N., S.P. and F.M.; methodology, M.N., S.P., R.M.P., M.H., F.M., L.L., Y.X. and D.M.B.; investigation, M.N.; writing—original draft preparation, M.N.; writing—review and editing, M.N., S.P., R.M.P., M.H., F.M., L.L., Y.X. and D.M.B.; supervision, S.P. and D.M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research is part of the INCT of the Future Internet for Smart Cities funded by CNPq proc. 465446/2014-0, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, FAPESP proc. 14/50937-1 and proc. 15/24485-9. It is also part of the FAPESP proc. 21/06995-0. This research was additionally funded by the National Key Research and DevelopmentProgram of China under Grant No.2023YFB2704400.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016. Available online: http://www.deeplearningbook.org (accessed on 12 March 2025).
2. Mitchell, R.; Michalski, J.; Carbonell, T. *An Artificial Intelligence Approach*; Springer: Berlin/Heidelberg, Germany, 2013.
3. Ghoddusi, H.; Creamer, G.G.; Rafizadeh, N. Machine learning in energy economics and finance: A review. *Energy Econ.* **2019**, *81*, 709–727. [CrossRef]
4. Athey, S. The impact of machine learning on economics. In *The Economics of Artificial Intelligence: An Agenda*; University of Chicago Press: Chicago, IL, USA, 2018; pp. 507–547.
5. Wuest, T.; Weimer, D.; Irgens, C.; Thoben, K.D. Machine learning in manufacturing: advantages, challenges, and applications. *Prod. Manuf. Res.* **2016**, *4*, 23–45. [CrossRef]
6. Sharp, M.; Ak, R.; Hedberg, T., Jr. A survey of the advancing use and development of machine learning in smart manufacturing. *J. Manuf. Syst.* **2018**, *48*, 170–179. [CrossRef]
7. Bhavsar, P.; Safro, I.; Bouaynaya, N.; Polikar, R.; Dera, D. Machine learning in transportation data analytics. In *Data Analytics for Intelligent Transportation Systems*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 283–307.
8. Nguyen, H.; Kieu, L.M.; Wen, T.; Cai, C. Deep learning methods in transportation domain: a review. *IET Intell. Transp. Syst.* **2018**, *12*, 998–1004. [CrossRef]
9. Bhardwaj, R.; Nambiar, A.R.; Dutta, D. A study of machine learning in healthcare. In Proceedings of the 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC), Turin, Italy, 4–8 July 2017; IEEE: Piscataway, NJ, USA, 2017; Volume 2, pp. 236–241.

10. Shailaja, K.; Seetharamulu, B.; Jabbar, M. Machine Learning in Healthcare: A Review. In Proceedings of the 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 29–31 March 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 910–914.
11. Saharkhizan, M.; Azmoodeh, A.; Dehghantanha, A.; Choo, K.K.R.; Parizi, R.M. An Ensemble of Deep Recurrent Neural Networks for Detecting IoT Cyber Attacks Using Network Traffic. *IEEE Internet Things J.* **2020**, *7*, 8852–8859. [CrossRef]
12. Al-Abassi, A.; Karimipour, H.; Dehghantanha, A.; Parizi, R.M. An Ensemble Deep Learning-Based Cyber-Attack Detection in Industrial Control System. *IEEE Access* **2020**, *8*, 83965–83973. [CrossRef]
13. El Naqa, I.; Murphy, M.J. *What is Machine Learning*? Springer: Berlin/Heidelberg, Germany, 2015; pp. 3–11.
14. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
15. The HIPAA Privacy Rule. 2020. Available online: https://www.hhs.gov/hipaa/for-professionals/privacy/index.html (accessed on 8 December 2020).
16. General Data Protection Regulation (GDPR). 2018. Available online: https://gdpr-info.eu (accessed on 29 December 2020).
17. Rumbold, J.M.; Pierscionek, B.K. What are data? A categorization of the data sensitivity spectrum. *Big Data Res.* **2018**, *12*, 49–59. [CrossRef]
18. Yang, Q.; Liu, Y.; Chen, T.; Tong, Y. Federated machine learning: Concept and applications. *ACM Trans. Intell. Syst. Technol. (Tist)* **2019**, *10*, 1–19. [CrossRef]
19. Chen, Y.; Qin, X.; Wang, J.; Yu, C.; Gao, W. Fedhealth: A federated transfer learning framework for wearable healthcare. *IEEE Intell. Syst.* **2020**, *35*, 83–93. [CrossRef]
20. Voigt, P.; von dem Bussche, A. *The EU General Data Protection Regulation (GDPR): A Practical Guide*, 1st ed.; Springer: Cham, Switzerland, 2017. [CrossRef]
21. Ghassemi, M.; Naumann, T.; Schulam, P.; Beam, A.L.; Chen, I.Y.; Ranganath, R. A review of challenges and opportunities in machine learning for health. *AMIA Summits Transl. Sci. Proc.* **2020**, *2020*, 191. [PubMed]
22. Rajkomar, A.; Dean, J.; Kohane, I. Machine learning in medicine. *New Engl. J. Med.* **2019**, *380*, 1347–1358. [CrossRef]
23. Razzak, M.I.; Naz, S.; Zaib, A. Deep learning for medical image processing: Overview, challenges and the future. In *Classification in BioApps: Automation of Decision Making*; Springe: Berlin/Heidelberg, Germany, 2017; pp. 323–350.
24. Mothukuri, V.; Parizi, R.M.; Pouriyeh, S.; Huang, Y.; Dehghantanha, A.; Srivastava, G. A survey on security and privacy of federated learning. *Future Gener. Comput. Syst.* **2021**, *115*, 619–640. [CrossRef]
25. Aledhari, M.; Razzak, R.; Parizi, R.M.; Saeed, F. Federated Learning: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Access* **2020**, *8*, 140699–140725. [CrossRef]
26. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In *Artificial Intelligence and Statistics*; PMLR: Cambridge, MA, USA, 2017; pp. 1273–1282.
27. Hu, F.; Yang, H.; Qiu, L.; Wei, S.; Hu, H.; Zhou, H. Spatial structure and organization of the medical device industry urban network in China: evidence from specialized, refined, distinctive, and innovative firms. *Front. Public Health* **2025**, *13*, 1518327. [CrossRef]
28. Xu, J.; Glicksberg, B.S.; Su, C.; Walker, P.; Bian, J.; Wang, F. Federated learning for healthcare informatics. *J. Healthc. Inform. Res.* **2020**, *5*, 1–19. [CrossRef] [PubMed]
29. Rieke, N.; Hancox, J.; Li, W.; Milletari, F.; Roth, H.; Albarqouni, S.; Bakas, S.; Galtier, M.N.; Landman, B.; Maier-Hein, K.; et al. The future of digital health with federated learning. *arXiv* **2020**, arXiv:2003.08119. [CrossRef]
30. Kaissis, G.A.; Makowski, M.R.; Rückert, D.; Braren, R.F. Secure, privacy-preserving and federated machine learning in medical imaging. *Nat. Mach. Intell.* **2020**, *2*, 305–311. [CrossRef]
31. Liu, Y.; Kang, Y.; Xing, C.; Chen, T.; Yang, Q. Secure federated transfer learning. *arXiv* **2018**, arXiv:1812.03337.
32. Wu, Q.; He, K.; Chen, X. Personalized federated learning for intelligent iot applications: A cloud-edge based framework. *IEEE Comput. Graph. Appl.* **2020**, *1*, 35–44. [CrossRef]
33. Wu, Q.; Chen, X.; Zhou, Z.; Zhang, J. FedHome: Cloud-Edge based Personalized Federated Learning for In-Home Health Monitoring. *IEEE Trans. Mob. Comput.* **2020**, *21*, 2818–2832. [CrossRef]
34. Liu, D.; Miller, T.; Sayeed, R.; Mandl, K.D. Fadl: Federated-autonomous deep learning for distributed electronic health record. *arXiv* **2018**, arXiv:1811.11400.
35. Passerat-Palmbach, J.; Farnan, T.; Miller, R.; Gross, M.S.; Flannery, H.L.; Gleim, B. A blockchain-orchestrated Federated Learning architecture for healthcare consortia. *arXiv* **2019**, arXiv:1910.12603.
36. Guo, Y.; Liu, F.; Cai, Z.; Chen, L.; Xiao, N. FEEL: A Federated Edge Learning System for Efficient and Privacy-Preserving Mobile Healthcare. In Proceedings of the 49th International Conference on Parallel Processing-ICPP, Edmonton, AB, Canada, 17–20 August 2020; pp. 1–11.
37. Malik, H.; Naeem, A.; Naqvi, R.A.; Loh, W.K. DMFL_Net: A Federated Learning-Based Framework for the Classification of COVID-19 from Multiple Chest Diseases Using X-rays. *Sensors* **2023**, *23*, 743. [CrossRef] [PubMed]

38. Gupta, A.; Misra, S.; Pathak, N.; Das, D. Fedcare: Federated learning for resource-constrained healthcare devices in iomt system. *IEEE Trans. Comput. Soc. Syst.* **2023**, *10*, 1587–1596. [CrossRef]
39. Pham, C.H.; Huynh-The, T.; Sedgh-Gooya, E.; El-Bouz, M.; Alfalou, A. Extension of physical activity recognition with 3D CNN using encrypted multiple sensory data to federated learning based on multi-key homomorphic encryption. *Comput. Methods Programs Biomed.* **2024**, *243*, 107854. [CrossRef]
40. Al-Turjman, F.; Baali, I. Machine learning for wearable IoT-based applications: A survey. *Trans. Emerg. Telecommun. Technol.* **2019**, *33*, e3635. [CrossRef]
41. Liang, P.P.; Liu, T.; Ziyin, L.; Salakhutdinov, R.; Morency, L.P. Think locally, act globally: Federated learning with local and global representations. *arXiv* **2020**, arXiv:2001.01523.
42. SÜZEN, A.A.; ŞİMŞEK, M.A. A Novel Approach to Machine Learning Application to Protection Privacy Data in Healthcare: Federated Learning. *Namik Kemal Tip Derg.* **2020**, *8*, 22–30.
43. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In Proceedings of the International Workshop on Ambient Assisted Living, Vitoria-Gasteiz, Spain, 3–5 December 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 216–223.
44. Vavoulas, G.; Chatzaki, C.; Malliotakis, T.; Pediaditis, M.; Tsiknakis, M. The MobiAct Dataset: Recognition of Activities of Daily Living using Smartphones. In Proceedings of the ICT4AgeingWell, Rome, Italy, 21–22 April 2016; pp. 143–151.
45. CMS. Electronic Health Records. 2012. Available online: https://web.archive.org/web/20120915155253/http://www.cms.gov/Medicare/E-Health/EHealthRecords/index.html (accessed on 23 April 2025).
46. Pollard, T.J.; Johnson, A.E.; Raffa, J.D.; Celi, L.A.; Mark, R.G.; Badawi, O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci. Data* **2018**, *5*, 180178. [CrossRef]
47. Jones, R. Hyperledger Besu. 2020. Available online: https://wiki.hyperledger.org/display/BESU/Hyperledger+Besu (accessed on 15 December 2021).
48. Shbair, W.; Steichen, M.; François, J. Blockchain orchestration and experimentation framework: A case study of KYC. In Proceedings of the IEEE/IFIP Man2Block 2018-IEEE/IFIP Network Operations and Management Symposium, Taipei, Taiwan, 23–27 April 2018.
49. Alabi, T. Ethereum Improvement Proposal-1024 - Add web3.eth.encrypt and web3.eth.decrypt Functions. Technical Report. 2018. Available online: https://github.com/ethereum/EIPs/pull/1098 (accessed on 15 December 2020).
50. Melis, L.; Song, C.; De Cristofaro, E.; Shmatikov, V. Exploiting unintended feature leakage in collaborative learning. In Proceedings of the 2019 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 19–23 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 691–706.
51. Abadi, M.; Chu, A.; Goodfellow, I.; McMahan, H.B.; Mironov, I.; Talwar, K.; Zhang, L. Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 25–27 October 2016; pp. 308–318.
52. Boneh, D. The decision diffie-hellman problem. In Proceedings of the International Algorithmic Number Theory Symposium, Portland, OR, USA, 13–18 June 2004; Springer: Berlin/Heidelberg, Germany, 1998; pp. 48–63.
53. Mangasarian, O.L.; Wolberg, W.H. Cancer diagnosis via linear programming. In *Technical Report*; University of Wisconsin-Madison Department of Computer Sciences: Madison, WI, USA, 1990.
54. Bilal, A.; Shafiq, M.; Obidallah, W.J.; Alduraywish, Y.A.; Long, H. Quantum computational infusion in extreme learning machines for early multi-cancer detection. *J. Big Data* **2025**, *12*, 1–48. [CrossRef]
55. Rhodes, C. Perfect Harmony: Pharma's MELLODDY Consortium Joins Forces with NVIDIA to Supercharge AI Drug Discovery. 2019. Avaiable online: https://blogs.nvidia.com/blog/2019/08/08/pharma-melloddy-ai-drug-discovery-consortium/ (accessed on 3 November 2020).
56. nvidia. King's College London is Working to Bring AI Benefits to the Point of Care. 2019. Avaiable online: https://www.nvidia.com/content/dam/en-us/Solutions/data-center/gated-resources/hc-casestudy-kings-college-london.pdf (accessed on 3 November 2020).
57. Xiong, Z.; Cheng, Z.; Liu, X.; Wang, D.; Luo, X.; Chen, K.; Jiang, H.; Zheng, M. Facing small and biased data dilemma in drug discovery with federated learning. *BioRxiv* **2020**.
58. Sorkun, M.C.; Khetan, A.; Er, S. AqSolDB, a curated reference set of aqueous solubility and 2D descriptors for a diverse set of compounds. *Sci. Data* **2019**, *6*, 143. [CrossRef]
59. Chen, S.; Xue, D.; Chuai, G.; Yang, Q.; Liu, Q. FL-QSAR: a federated learning based QSAR prototype for collaborative drug discovery. *bioRxiv* **2020**. [CrossRef] [PubMed]
60. Ma, J.; Sheridan, R.P.; Liaw, A.; Dahl, G.E.; Svetnik, V. Deep neural nets as a method for quantitative structure–activity relationships. *J. Chem. Inf. Model.* **2015**, *55*, 263–274. [CrossRef]

61. Choudhury, O.; Park, Y.; Salonidis, T.; Gkoulalas-Divanis, A.; Sylla, I.; k Das, A. Predicting Adverse Drug Reactions on Distributed Health Data using Federated Learning. In Proceedings of the AMIA Annual Symposium Proceedings. American Medical Informatics Association, Washington, DC, USA, 16–20 November 2019; Volume 2019, p. 313.
62. IBM. IBM Explorys Electronic Health Record (EHR) Database. Available online: https://pophealth.ucsf.edu/ibm-marketscan-research-databases-access (accessed on 23 April 2025).
63. Huang, L.; Shea, A.L.; Qian, H.; Masurkar, A.; Deng, H.; Liu, D. Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *J. Biomed. Inform.* **2019**, *99*, 103291. [CrossRef]
64. Pfohl, S.R.; Dai, A.M.; Heller, K. Federated and Differentially Private Learning for Electronic Health Records. *arXiv* **2019**, arXiv:1911.05861.
65. Sharma, P.; Shamout, F.E.; Clifton, D.A. Preserving patient privacy while training a predictive model of in-hospital mortality. *arXiv* **2019**, arXiv:1912.00354.
66. Johnson, A.E.; Pollard, T.J.; Shen, L.; Li-Wei, H.L.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.; Celi, L.A.; Mark, R.G. MIMIC-III, a freely accessible critical care database. *Sci. Data* **2016**, *3*, 1–9. [CrossRef]
67. Vaid, A.; Jaladanki, S.K.; Xu, J.; Teng, S.; Kumar, A.; Lee, S.; Somani, S.; Paranjpe, I.; De Freitas, J.K.; Wanyan, T.; et al. Federated learning of electronic health records to improve mortality prediction in hospitalized patients with COVID-19: machine learning approach. *JMIR Med. Inform.* **2021**, *9*, e24207. [CrossRef]
68. Liu, B.; Yan, B.; Zhou, Y.; Yang, Y.; Zhang, Y. Experiments of federated learning for covid-19 chest x-ray images. *arXiv* **2020**, arXiv:2007.05592.
69. Brisimi, T.S.; Chen, R.; Mela, T.; Olshevsky, A.; Paschalidis, I.C.; Shi, W. Federated learning of predictive models from federated electronic health records. *Int. J. Med. Inform.* **2018**, *112*, 59–67. [CrossRef]
70. Boughorbel, S.; Jarray, F.; Venugopal, N.; Moosa, S.; Elhadi, H.; Makhlouf, M. Federated Uncertainty-Aware Learning for Distributed Hospital EHR Data. *arXiv* **2019**, arXiv:1910.12191.
71. Roy, A.G.; Siddiqui, S.; Pölsterl, S.; Navab, N.; Wachinger, C. Braintorrent: A peer-to-peer environment for decentralized federated learning. *arXiv* **2019**, arXiv:1905.06731.
72. Landman, B.; Warfield, S. MICCAI 2012 workshop on multi-atlas labeling. In Proceedings of the MICCAI Grand Challenge and Workshop on Multi-Atlas Labeling, CreateSpace Independent Publishing Platform, Marrakesh, Morocco, 7–11 October 2012.
73. Bakas, S.; Akbari, H.; Sotiras, A.; Bilello, M.; Rozycki, M.; Kirby, J.S.; Freymann, J.B.; Farahani, K.; Davatzikos, C. Segmentation Labels for the Pre-operative Scans of the TCGA-GBM collection [Data Set]. 2017. Available online: Available online: https://www.cancerimagingarchive.net/analysis-result/brats-tcga-gbm/ (accessed on 12 March 2025).
74. Li, W.; Milletarì, F.; Xu, D.; Rieke, N.; Hancox, J.; Zhu, W.; Baust, M.; Cheng, Y.; Ourselin, S.; Cardoso, M.J.; et al. Privacy-preserving federated brain tumour segmentation. In Proceedings of the International Workshop on Machine Learning in Medical Imaging, Shenzhen, China, 17 October 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 133–141.
75. Bakas, S.; Reyes, M.; Jakab, A.; Bauer, S.; Rempfler, M.; Crimi, A.; Shinohara, R.T.; Berger, C.; Ha, S.M.; Rozycki, M.; et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. *arXiv* **2018**, arXiv:1811.02629.
76. Li, X.; Gu, Y.; Dvornek, N.; Staib, L.; Ventola, P.; Duncan, J.S. Multi-site fmri analysis using privacy-preserving federated learning and domain adaptation: Abide results. *arXiv* **2020**, arXiv:2001.05647. [CrossRef]
77. Di Martino, A.; Yan, C.G.; Li, Q.; Denio, E.; Castellanos, F.X.; Alaerts, K.; Anderson, J.S.; Assaf, M.; Bookheimer, S.Y.; Dapretto, M.; et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* **2014**, *19*, 659–667. [CrossRef]
78. Kumar, R.; Khan, A.A.; Zhang, S.; Wang, W.; Abuidris, Y.; Amin, W.; Kumar, J. Blockchain-federated-learning and deep learning models for COVID-19 detection using CT imaging. *arXiv* **2020**, arXiv:2007.06537. [CrossRef]
79. RSNA. RSNA Pneumonia Detection Challenge. 2018. Available online: https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/ (accessed on 24 December 2020).
80. Zhang, W.; Zhou, T.; Lu, Q.; Wang, X.; Zhu, C.; Wang, Z.; Wang, F. Dynamic Fusion based Federated Learning for COVID-19 Detection. *arXiv* **2020**, arXiv:2009.10401. [CrossRef]
81. Zhao, J.; Zhang, Y.; He, X.; Xie, P. COVID-CT-Dataset: a CT scan dataset about COVID-19. *arXiv* **2020**, arXiv:2003.13865.
82. Kaggle. COVID-19 Radiography Database. 2018. Available online: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database, (accessed on 24 December 2020).
83. Figure 1 COVID-19 Chest X-ray Dataset Initiative. 2020. Available online: https://github.com/agchung/Figure1-COVID-chestxray-dataset, (accessed on 24 December 2020).
84. Baheti, P.; Sikka, M.; Arya, K.; Rajesh, R. Federated Learning on Distributed Medical Records for Detection of Lung Nodules. In Proceedings of the VISIGRAPP (4: VISAPP), Valletta, Malta, 27–29 February 2020; pp. 445–451.

85. Armato, S.G., III; McLennan, G.; Bidaut, L.; McNitt-Gray, M.F.; Meyer, C.R.; Reeves, A.P.; Zhao, B.; Aberle, D.R.; Henschke, C.I.; Hoffman, E.A.; et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Med. Phys.* **2011**, *38*, 915–931. [PubMed]
86. Linardos, A.; Kushibar, K.; Walsh, S.; Gkontra, P.; Lekadir, K. Federated learning for multi-center imaging diagnostics: a simulation study in cardiovascular disease. *Sci. Rep.* **2022**, *12*, 3551. [CrossRef] [PubMed]
87. Bernard, O.; Lalande, A.; Zotti, C.; Cervenansky, F.; Yang, X.; Heng, P.A.; Cetin, I.; Lekadir, K.; Camara, O.; Ballester, M.A.G.; et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved? *IEEE Trans. Med. Imaging* **2018**, *37*, 2514–2525. [CrossRef]
88. Campello, V.M.; Gkontra, P.; Izquierdo, C.; Martin-Isla, C.; Sojoudi, A.; Full, P.M.; Maier-Hein, K.; Zhang, Y.; He, Z.; Ma, J.; et al. Multi-centre, multi-vendor and multi-disease cardiac segmentation: the M&Ms challenge. *IEEE Trans. Med. Imaging* **2021**, *40*, 3543–3554.
89. Lee, H.; Chai, Y.J.; Joo, H.; Lee, K.; Hwang, J.Y.; Kim, S.M.; Kim, K.; Nam, I.C.; Choi, J.Y.; Yu, H.W.; et al. Federated learning for thyroid ultrasound image analysis to protect personal information: Validation study in a real health care environment. *JMIR Med. Inform.* **2021**, *9*, e25869. [CrossRef] [PubMed]
90. Lee, J.; Sun, J.; Wang, F.; Wang, S.; Jun, C.H.; Jiang, X. Privacy-preserving patient similarity learning in a federated environment: development and analysis. *JMIR Med. Inform.* **2018**, *6*, e20. [CrossRef]
91. Xu, J.; Xu, Z.; Walker, P.; Wang, F. Federated patient hashing. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 6486–6493.
92. Kim, Y.; Sun, J.; Yu, H.; Jiang, X. Federated tensor factorization for computational phenotyping. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017; pp. 887–895.
93. Wah, C.; Branson, S.; Welinder, P.; Perona, P.; Belongie, S. The Caltech-UCSD Birds-200-2011 Dataset. California Institute of Technology, CNS-TR-2011-001. 2011. Available online: http://www.vision.caltech.edu/datasets/cub_200_2011/ (accessed on 12 March 2025).
94. Liu, D.; Dligach, D.; Miller, T. Two-stage federated phenotyping and patient representation learning. *arXiv* **2019**, arXiv:1908.05596.
95. Yuan, B.; Ge, S.; Xing, W. A Federated Learning Framework for Healthcare IoT devices. *arXiv* **2020**, arXiv:2005.05083.
96. Goldberger, A.L.; Amaral, L.A.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* **2000**, *101*, e215–e220. [CrossRef]
97. Silva, S.; Gutman, B.A.; Romero, E.; Thompson, P.M.; Altmann, A.; Lorenzi, M. Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 270–274.
98. Alzheimer's Disease Neuroimaging Initiative. 2020. Available online: http://adni.loni.usc.edu (accessed on 26 December 2020).
99. Malone, I.B.; Cash, D.; Ridgway, G.R.; MacManus, D.G.; Ourselin, S.; Fox, N.C.; Schott, J.M. MIRIAD—Public release of a multiple time point Alzheimer's MR imaging dataset. *NeuroImage* **2013**, *70*, 33–36. [CrossRef]
100. Parkinson's Progression Markers Initiative. 2018. Available online: https://www.ppmi-info.org (accessed on 26 December 2020).
101. Biobank UK: Enabling Scientific Discoveries that Improve Human Health. 2018. Available online: https://www.ukbiobank.ac.uk (accessed on 26 December 2020).
102. McMahan, H.B.; Ramage, D.; Talwar, K.; Zhang, L. Learning differentially private recurrent language models. *arXiv* **2017**, arXiv:1710.06963.
103. coMind Collaborative Machine Learning. 2018. Available online: https://medium.com/@Acuratio/collaborative-machine-learning-ba7961a75f3a (accessed on 18 December 2020).
104. Behrman, R.E.; Butler, A.S. *Preterm Birth: Causes, Consequences, and Prevention*; National Academies Press: Washington, DC, USA, 2007.
105. Bluemke, E. Privacy-Preserving AI in Medical Imaging: Federated Learning, Differential Privacy, and Encrypted Computation. 2019. Available online: https://blog.openmined.org/federated-learning-differential-privacy-and-encrypted-computation-for-medical-imaging/ (accessed on 3 November 2020).
106. Despotović, I.; Goossens, B.; Philips, W. MRI segmentation of the human brain: challenges, methods, and applications. *Comput. Math. Methods Med.* **2015**, *2015*, 450341. [CrossRef]
107. de Brebisson, A.; Montana, G. Deep neural networks for anatomical brain segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Boston, MA, USA, 7–12 June 2015; pp. 20–28.
108. Roy, A.G.; Conjeti, S.; Navab, N.; Wachinger, C.; Alzheimer's Disease Neuroimaging Initiative. QuickNAT: A fully convolutional network for quick and accurate segmentation of neuroanatomy. *NeuroImage* **2019**, *186*, 713–727.

109. Sheller, M.J.; Reina, G.A.; Edwards, B.; Martin, J.; Bakas, S. Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. In Proceedings of the International MICCAI Brainlesion Workshop, Granada, Spain, 16 September 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 92–104.
110. Rumbold, J.M.M.; Pierscionek, B. The effect of the general data protection regulation on medical research. *J. Med. Internet Res.* **2017**, *19*, e47. [CrossRef]
111. Portability, I.; Act, A. *Guidance Regarding Methods for de-Identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule*; Human Health Services: Washington, DC, USA, 2012.
112. Chaudhuri, K.; Imola, J.; Machanavajjhala, A. Capacity bounded differential privacy. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; pp. 3474–3483.
113. WHO. WHO Coronavirus Disease (COVID-19). 2021. Available online: https://covid19.who.int (accessed on 18 January 2021).
114. Vaishya, R.; Javaid, M.; Khan, I.H.; Haleem, A. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab. Syndr. Clin. Res. Rev.* **2020**, *14*, 337–339. [CrossRef] [PubMed]
115. Shahid, O.; Nasajpour, M.; Pouriyeh, S.; Parizi, R.M.; Han, M.; Valero, M.; Li, F.; Aledhari, M.; Sheng, Q.Z. Machine Learning Research Towards Combating COVID-19: Virus Detection, Spread Prevention, and Medical Assistance. *arXiv* **2020**, arXiv:2010.07036. [CrossRef] [PubMed]
116. Bragazzi, N.L.; Dai, H.; Damiani, G.; Behzadifar, M.; Martini, M.; Wu, J. How Big Data and Artificial Intelligence Can Help Better Manage the COVID-19 Pandemic. *Int. J. Environ. Res. Public Health* **2020**, *17*, 3176. [CrossRef]
117. Nasajpour, M.; Pouriyeh, S.; Parizi, R.M.; Dorodchi, M.; Valero, M.; Arabnia, H.R. Internet of Things for current COVID-19 and future pandemics: An exploratory study. *J. Healthc. Inform. Res.* **2020**, *4*, 325–364. [CrossRef]
118. Sun, J.; Wang, F.; Hu, J.; Edabollahi, S. Supervised patient similarity measure of heterogeneous patient records. *Acm Sigkdd Explor. Newsl.* **2012**, *14*, 16–24. [CrossRef]
119. Carroll, R.J.; Eyler, A.E.; Denny, J.C. Naïve electronic health record phenotype identification for rheumatoid arthritis. In Proceedings of the AMIA Annual Symposium Proceedings. American Medical Informatics Association, Washington DC, USA, 22–26 October 2011; Volume 2011, p. 189.
120. Ho, J.C.; Ghosh, J.; Steinhubl, S.R.; Stewart, W.F.; Denny, J.C.; Malin, B.A.; Sun, J. Limestone: High-throughput candidate phenotype generation via tensor factorization. *J. Biomed. Inform.* **2014**, *52*, 199–211. [CrossRef]
121. Dwork, C. Differential privacy: A survey of results. In Proceedings of the International Conference on Theory and Applications of Models of Computation, Xi'an, China, 25–29 April 2008; Springer: Berlin/Heidelberg, Germany, 2008; pp. 1–19.
122. Boyd, S.; Parikh, N.; Chu, E. *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*; Now Publishers Inc.: Hanover, MA, USA, 2011.
123. Dligach, D.; Miller, T. Learning Patient Representations from Text. In Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics, New Orleans, LA, USA, 5–6 June 2018; pp. 119–123. [CrossRef]
124. Miotto, R.; Li, L.; Kidd, B.A.; Dudley, J.T. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Sci. Rep.* **2016**, *6*, 26094. [CrossRef]
125. Islam, S.R.; Kwak, D.; Kabir, M.H.; Hossain, M.; Kwak, K.S. The internet of things for health care: a comprehensive survey. *IEEE Access* **2015**, *3*, 678–708. [CrossRef]
126. Firouzi, F.; Rahmani, A.M.; Mankodiya, K.; Badaroglu, M.; Merrett, G.V.; Wong, P.; Farahani, B. Internet-of-Things and big data for smarter healthcare: From device to architecture, applications and analytics. *Future Gener. Comput. Syst.* **2018**, *78*, 583–586. [CrossRef]
127. Lipsey, M.W.; Wilson, D.B. *Practical Meta-Snalysis*; SAGE publications, Inc.: New York, NY, USA, 2001.
128. Wei, E.; Ozdaglar, A. Distributed alternating direction method of multipliers. In Proceedings of the 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), Maui, HI, USA, 10–13 December 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 5445–5450.
129. Konečnỳ, J.; McMahan, H.B.; Yu, F.X.; Richtárik, P.; Suresh, A.T.; Bacon, D. Federated learning: Strategies for improving communication efficiency. *arXiv* **2016**, arXiv:1610.05492.
130. Li, T.; Sahu, A.K.; Talwalkar, A.; Smith, V. Federated learning: Challenges, methods, and future directions. *IEEE Signal Process. Mag.* **2020**, *37*, 50–60. [CrossRef]
131. Kamp, M.; Adilova, L.; Sicking, J.; Hüger, F.; Schlicht, P.; Wirtz, T.; Wrobel, S. Efficient decentralized deep learning by dynamic model averaging. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Dublin, Ireland, 10–14 September 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 393–409.
132. Lin, Y.; Han, S.; Mao, H.; Wang, Y.; Dally, W.J. Deep Gradient Compression: Reducing the Communication Bandwidth for Distributed Training. In Proceedings of the International Conference on Learning Representations (ICLR), Vancouver, BC, Canada, 30 April–3 May 2018.
133. Zhang, S.Q.; Lin, J.; Zhang, Q.; McDanel, B.; Kim, S.W. Field-Aware Adaptive Communication for Collaborative Learning. *IEEE Trans. Mob. Comput.* **2021**, *20*, 2981–2994. [CrossRef]

134. Kairouz, P.; McMahan, H.B.; Avent, B.; Bellet, A.; Bennis, M.; Bhagoji, A.N.; Bonawitz, K.; Charles, Z.; Cormode, G.; Cummings, R.; et al. Advances and open problems in federated learning. *arXiv* **2019**, arXiv:1912.04977.
135. Dwork, C.; Kenthapadi, K.; McSherry, F.; Mironov, I.; Naor, M. Our data, ourselves: Privacy via distributed noise generation. In Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques, St. Petersburg, Russia, 28 May 28–1 June 2006; Springer: Berlin/Heidelberg, Germany, 2006; pp. 486–503.
136. Goldreich, O. Secure multi-party computation. *Manuscript. Prelim. Version* **1998**, *78*, 1–108.
137. Ma, C.; Li, J.; Ding, M.; Yang, H.H.; Shu, F.; Quek, T.Q.; Poor, H.V. On Safeguarding Privacy and Security in the Framework of Federated Learning. *IEEE Netw.* **2020**, *34*, 242–248. [CrossRef]
138. Wu, B.; Zhao, S.; Sun, G.; Zhang, X.; Su, Z.; Zeng, C.; Liu, Z. P3sgd: Patient privacy preserving sgd for regularizing deep cnns in pathological image classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 2099–2108.
139. Zhu, L.; Liu, Z.; Han, S. Deep leakage from gradients. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; pp. 14774–14784.
140. Wang, Z.; Song, M.; Zhang, Z.; Song, Y.; Wang, Q.; Qi, H. Beyond inferring class representatives: User-level privacy leakage from federated learning. In Proceedings of the IEEE INFOCOM 2019-IEEE Conference on Computer Communications, Paris, France, 29 April–2 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 2512–2520.
141. Sindhusaranya, B.; Yamini, R.; Manimekalai, M.A.P.; Geetha, K. Homomorphic Encryption-Based Privacy-Preserving Federated Learning in IoT-Enabled Healthcare System. *IEEE Trans. Netw. Sci. Eng.* **2022**, *10*, 2864–2880. [CrossRef]
142. Wei, K.; Li, J.; Ding, M.; Ma, C.; Yang, H.H.; Farokhi, F.; Jin, S.; Quek, T.Q.S.; Poor, H.V. Personalized Federated Learning with Differential Privacy. *IEEE Internet Things J.* **2020**, *7*, 9650–9662. [CrossRef]
143. Verma, D.C.; White, G.; Julier, S.; Pasteris, S.; Chakraborty, S.; Cirincione, G. Approaches to address the data skew problem in federated learning. In Proceedings of the Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications. International Society for Optics and Photonics, Baltimore, MD, USA, 15–17 April 2019; Volume 11006, p. 110061I.
144. Zhao, Y.; Li, M.; Lai, L.; Suda, N.; Civin, D.; Chandra, V. Federated Learning with Non-IID Data. *arXiv* **2018**, arXiv:1806.00582. [CrossRef]
145. Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated Optimization in Heterogeneous Networks. In Proceedings of the Machine Learning and Systems (MLSys), Austin, TX, USA, 2–4 March 2020.
146. Li, Z.; Huang, L. Tracing requirements as a problem of machine learning. *Int. J. Softw. Eng. Appl. (IJSEA)* **2018**, *9*, 21–36. [CrossRef]
147. Rodriguez, R.; Svensson, G.; Wood, G. Sustainability trends in public hospitals: Efforts and priorities. *Eval. Program Plan.* **2020**, *78*, 101742. [CrossRef]
148. Trichelair, P.; Andreux, M.; Trower, A. Data Heterogeneity in Healthcare. Available online: https://owkin.com/federated-learning/heterogeneity-in-healthcare/ (accessed on 3 December 2020).
149. Crelier, B.; Streit, S.; Donzé, J. Patient Complexity Characteristics in the Hospital Setting. *Am. J. Accountable Care* **2018**, *9*, 3–8.
150. Laguel, Y.; Pillutla, K.; Malick, J.; Harchaoui, Z. Device Heterogeneity in Federated Learning: A Superquantile Approach. *arXiv* **2020**, arXiv:2002.11223.
151. Hahn, S.J.; Lee, J. Privacy-preserving Federated Bayesian Learning of a Generative Model for Imbalanced Classification of Clinical Data. *arXiv* **2019**, arXiv:1910.08489.
152. Liu, L.; Cao, J.; Liu, M.; Liu, Y. Client-Edge-Cloud Hierarchical Federated Learning. In Proceedings of the IEEE International Conference on Communications (ICC), Dublin, Ireland, 7–11 June 2020; pp. 1–6. [CrossRef]
153. Xie, C.; Koyejo, S.; Ray, I. Asynchronous Federated Optimization. *arXiv* **2019**, arXiv:1903.03934.
154. Reina, G.A.; Gruzdev, A.; Foley, P.; Perepelkina, O.; Sharma, M.; Davidyuk, I.; Trushkin, I.; Radionov, M.; Mokrov, A.; Agapov, D.; et al. OpenFL: An Open-Source Framework for Federated Learning. *arXiv* **2021**, arXiv:2105.06413.
155. Shen, Y.; Yuan, K.; Yang, M.; Tang, B.; Li, Y. Ontology-Based Data Integration for Federated Learning in Healthcare. In Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Seoul, Republic of Korea, 16–19 December 2020; pp. 1483–1488. [CrossRef]
156. Hashash, O.; Sharafeddine, S.; Dawy, Z. MEC-based energy-aware distributed feature extraction for mHealth applications with strict latency requirements. In Proceedings of the 2021 IEEE International Conference on Communications Workshops (ICC Workshops), Montreal, QC, Canada, 14–23 June 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–6.
157. Patel, V.A.; Bhattacharya, P.; Tanwar, S.; Gupta, R.; Sharma, G.; Bokoro, P.N.; Sharma, R. Adoption of federated learning for healthcare informatics: Emerging applications and future directions. *IEEE Access* **2022**, *10*, 90792–90826. [CrossRef]
158. Fu, L.; Zhang, H.; Gao, G.; Zhang, M.; Liu, X. Client selection in federated learning: Principles, challenges, and opportunities. *IEEE Internet Things J.* **2023**, *10*, 21811–21819. [CrossRef]
159. Smith, V.; Chiang, C.K.; Sanjabi, M.; Talwalkar, A.S. Federated multi-task learning. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 4424–4434. [CrossRef]

160. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [CrossRef] [PubMed]
161. Zhang, H.; Jiang, S.; Xuan, S. Decentralized federated learning based on blockchain: Concepts, framework, and challenges. *Comput. Commun.* **2024**, *216*, 140–150. [CrossRef]
162. Abreha, H.G.; Hayajneh, M.; Serhani, M.A. Federated learning in edge computing: A systematic survey. *Sensors* **2022**, *22*, 450. [CrossRef]
163. Holly, S.; Hiessl, T.; Lakani, S.R.; Schall, D.; Heitzinger, C.; Kemnitz, J. Evaluation of hyperparameter-optimization approaches in an industrial federated learning system. In *Data Science—Analytics and Applications, Proceedings of the 4th International Data Science Conference–iDSC2021, Online, 21–22 May 2021* ; Springer: Berlin/Heidelberg, Germany, 2022; pp. 6–13.
164. Deng, Y.; Lyu, F.; Ren, J.; Wu, H.; Zhou, Y.; Zhang, Y.; Shen, X. AUCTION: Automated and quality-aware client selection framework for efficient federated learning. *IEEE Trans. Parallel Distrib. Syst.* **2021**, *33*, 1996–2009. [CrossRef]