Article

Federated Learning for the Internet-of-Medical-Things: A Survey

Vivek Kumar Prasad [1], Pronaya Bhattacharya [1], Darshil Maru [1], Sudeep Tanwar [1,*], Ashwin Verma [1], Arunendra Singh [2], Amod Kumar Tiwari [3], Ravi Sharma [4], Ahmed Alkhayyat [5,6], Florin-Emilian Țurcanu [7,*] and Maria Simona Raboaca [8,9,10]

[1] Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad 382481, Gujarat, India
[2] Department of Information Technology, Pranveer Singh Institute of Technology, Kanpur 209305, Uttar Pradesh, India
[3] Department of Computer Science and Engineering, Rajikiya Engineering College, Sonbhadra 231206, Uttar Pradesh, India
[4] Centre for Inter-Disciplinary Research and Innovation, University of Petroleum and Energy Studies, Dehradun 248001, Uttarakhand, India
[5] College of Technical Engineering, The Islamic University, Najaf 54001, Iraq
[6] Department of Medical instruments Engineering Techniques, Al-Turath University College, Baghdad 10021, Iraq
[7] Department of Building Services, Faculty of Civil Engineering and Building Services, Technical University of Gheorghe Asachi, 700050 Iași, Romania
[8] National Research and Development Institute for Cryogenic and Isotopic Technologies—ICSI Râmicu Vâlcea, 240050 Râmicu Vâlcea, Romania
[9] Doctoral School, University Politehnica of Bucharest, 060042 Bucharest, Romania
[10] Faculty of Electrical Engineering and Computer Science, Ștefan cel Mare University, 720229 Suceava, Romania
* Correspondence: sudeep.tanwar@nirmauni.ac.in (S.T.); florin-emilian.turcanu@academic.tuiasi.ro (F.-E.T.)

Abstract: Recently, in healthcare organizations, real-time data have been collected from connected or implantable sensors, layered protocol stacks, lightweight communication frameworks, and end devices, named the Internet-of-Medical-Things (IoMT) ecosystems. IoMT is vital in driving healthcare analytics (HA) toward extracting meaningful data-driven insights. Recently, concerns have been raised over data sharing over IoMT, and stored electronic health records (EHRs) forms due to privacy regulations. Thus, with less data, the analytics model is deemed inaccurate. Thus, a transformative shift has started in HA from centralized learning paradigms towards distributed or edge-learning paradigms. In distributed learning, federated learning (FL) allows for training on local data without explicit data-sharing requirements. However, FL suffers from a high degree of statistical heterogeneity of learning models, level of data partitions, and fragmentation, which jeopardizes its accuracy during the learning and updating process. Recent surveys of FL in healthcare have yet to discuss the challenges of massive distributed datasets, sparsification, and scalability concerns. Because of this gap, the survey highlights the potential integration of FL in IoMT, the FL aggregation policies, reference architecture, and the use of distributed learning models to support FL in IoMT ecosystems. A case study of a trusted cross-cluster-based FL, named *Cross-FL*, is presented, highlighting the gradient aggregation policy over remotely connected and networked hospitals. Performance analysis is conducted regarding system latency, model accuracy, and the trust of consensus mechanism. The distributed FL outperforms the centralized FL approaches by a potential margin, which makes it viable for real-IoMT prototypes. As potential outcomes, the proposed survey addresses key solutions and the potential of FL in IoMT to support distributed networked healthcare organizations.

Keywords: federated Learning; healthcare; cloud computing; security; privacy; blockchain; machine learning

MSC: 92C50

1. Introduction

The recent decade has seen exponential and unbounded growth in the Internet-of-Medical-Things (IoMT) landscape. IoMT encompasses the collection of wearable sensors, which are mounted on the patient body to measure the patient's vital indicators, such as blood pressure, heartbeat, and oxygen levels. The sensors continuously monitor the data transmitted over wireless communication networks to gateway nodes that perform data analytics [1]. According to a recent report by *Research and Markets*, which was released in January 2022, the IoMT market is expected to reach $258 billion by 2026, which shows a significant increase of $203 billion in four years [2]. By 2025, the healthcare industry would constitute ≈30% of total Internet-of-Things (IoT) data volume, which is expected to rise a further 10% by 2030 [3]. The expected forecast in the IoMT domain promises new opportunities in healthcare analytics (HA). Thus, there are strong requirements to revisit the centralized machine learning (ML) algorithms owing to the issues of privacy, data leakage, and computational requirements [4]. Due to these challenges, the data collection process is distributed and often heterogeneous due to different autonomous stakeholders' involvement.

In centralized approaches, the analytics is mostly performed on cloud servers (Amazon Web Services, Google Cloud, Microsoft Azure, and others), where critical indicators are analyzed for future medical predictions using the cloud resources [5]. Thus, IoMT leverages remote patient monitoring and is controlled through connected networks. Recently, the growing privacy concerns with the health insurance portability and accountability act (HIPPA) and general data protection regulations (GDPR) act have restricted stakeholders from sharing electronic health records (EHRs) with medical practitioners (doctors, hospitals, and medical labs), and researchers with patient consent only [6]. IoMT enables users to access healthcare services via smart devices such as phones, tablets, and personal digital assistance [7]. Thus, with growing mobile applications (m-health services), data collection to form EHRs is frequently synchronized with centralized mobile cloud computing (CC) servers for planning and decision-making. At any stage, misconfiguration at the centralized server causes data leakage that impacts the privacy of the patient's sensitive data [8].

Data in different geographical locations have their sharing policies and rights. Thus, the sensitive attributes of the data are not shared without the prior consent of data owners [9]. The sharing policies are tightly scrutinized and regulated by law enforcement agencies or service providers due to legal policies. Moreover, in CC servers, an enemy can target these servers due to central and single-point vulnerabilities, which might result in privacy leakage, alterations, and authorization-based attacks. Thus, the sharing of data among remote channels should be restricted. Still, such a step hinders the analytics carried out by ML algorithms, as models are highly data-driven [10]. Thus, a fine balance between data sharing and data are required for the learning models. To address the aforementioned issues of ease of accessibility and storage of collected data, a decentralized approach seems feasible.

Thus, researchers have provided decentralized solutions in IoMT ecosystems [11,12]. The data are kept secure at local sites for analysis, which adds personalization and privacy; such data are also used to generate full-meaning information by applying dynamic association among the different attributes [13]. The distributed learning approach allows parallel learning models to train the data at local sites. The parallelism is carried out by two mechanisms—enforcing the same data to parallel nodes (data-parallel) or enforcing the same data to all nodes, where they train distinct portions of an ML model (model-parallel). However, with large data, the model division becomes complex, and challenges of computation and system heterogeneity arise owing to varying degrees of CPU, bandwidth, and storage capacity [14]. Another challenge is the non-independent and identical distribution (non-IID) data, which require unified feature learning and optimization [15]. To preserve the data homogeneity and privacy considerations, federated learning (FL) assures local learning in IoMT via two approaches, namely, the centralized FL (CFL) and distributed FL (DFL).

CFL employs cloud-based aggregation and is suitable for general models, whereas

DFL assumes the distribution of users as a prime condition and focuses on edge-based aggregation. Sometimes, a unified mix of CFL and DFL approach (hybrid FL) is possible and is termed edge-cloud-based aggregation. In CFL, every healthcare facility trains its model with local data aggregated and sent to improve the results of the global model. CFL improves the system's latency as sending the data to a global server is unnecessary to obtain the result. Instead, healthcare facilities update their local model parameters from global ones. However, in CFL, the healthcare nodes (or remote sites) have to trust the global server (a CC server) to correctly send the global parameters under which the local training would take place [16]. Thus, a malicious adversary might poison the global server or might poison the local updates, which would lead to incorrect training of the CFL model [17].

Thus, DFL models are created to address issues in CFL, where there is no requirement for a central model; instead, local learning is a peer learning process. The connected neighbors (single-hop) learn about local updates, which are sent to other nodes to improve the global update of the entire network. Thus, DFL has inherent potential in a variety of applications, ranging from vehicular networks [18], drones [19], IoT [20], financial systems [21], and others. Clinical data sets from remote hospitals and medical facilities are difficult to obtain owing to patients' sensitive medical information, which is unavailable and has restricted access. Hence, DFL allows numerous agents to learn a common prediction model together while preserving the confidentiality of the training data [22]. Most real-world applications for EHR are decentralized, and the HIPPA act prohibits sharing patient-level data with other parties, such as insurance providers and healthcare facilities.

Hospitals maintain many patient-level records at local facilities that contain a general description of each patient's characteristics with missing detailed information about an individual. It is impractical to train a complex model with limited and imbalanced data. CFL and DFL address issues surrounding the confidentiality and accuracy of medical data. In CFL, the computation takes place on a locally trained ML model, and only model updates or weights are transferred. In notation, we refer to ΔW_k as the slope m_k of the linear regression model for the kth healthcare entity. These updates are globally synchronized among the different entities via a centralized approach that consumes extra bandwidth and creates an additional overhead at CC servers. However, in practical setups, the bandwidth required at the local nodes is not uniform, owing to link constraints, which impose variable delays in the aggregation.

In contrast, DFL uses an edge aggregation strategy with low overheads in terms of resource consumption and variability in link delays. In such cases, the end devices seamlessly communicate with the edge aggregation server, owing to fewer devices in the coverage range. When nodes in a particular range increase, the edge aggregation server might communicate with the cloud server for offloading and resource requirements. Edge-cloud-based aggregation in DFL assures high scalability and improves the robustness of the IoMT ecosystem. Another DFL approach is collaborative learning [23], where IoMT devices share resources during the training process. In this case, devices send their local models to nearby devices rather than the edge server. However, the security of the local model is achieved through effective schemes such as homomorphic encryption, which allows users to perform computations on the encrypted model with the requirement of decryption [24]. Then, local updates are shared to edge servers for updates. Another approach is the gossip learning approach which deals with proper bandwidth utilization and shares the ΔW_k with its directly connected neighbor healthcare entity [25]. The weights are exchanged by the segmented gossip aggregation method. These DFL approaches maintain data integrity and, at the same time, assure privacy-preserving challenges.

As network hospitals are geographically dispersed in DFL, they work together to train ML models for data-driven applications. The raw data acquired at every IoMT sensor are not shared with other nodes, which allows FL to safeguard the privacy of EHR in the IoMT realm. Figure 1a,b, shows the CFL and DFL learning approaches. In both scenarios, $\{\Delta W_1, \Delta W_2, \Delta W_3, \ldots \Delta W_n\}$ are the gradient weights of the local model of n hospitals, presented as $\{H_1, H_2, H_3, \ldots, H_n\}$, respectively. The only difference between

CFL and DFL approaches is the aggregation policy, where the gradient ΔW_n is shared with the global model or peer nodes. Table 1 presents a comparative analysis of the CFL and DFL approaches in terms of diverse parameters. The table signifies that both CFL and DFL have significant advantages under the privacy-preservation view, where techniques such as differential privacy, anonymization, and diversity are frequently used. Another critical aspect is FL communication, where techniques are devised to improve the overall data rate, minimize latency, and manage resources effectively.
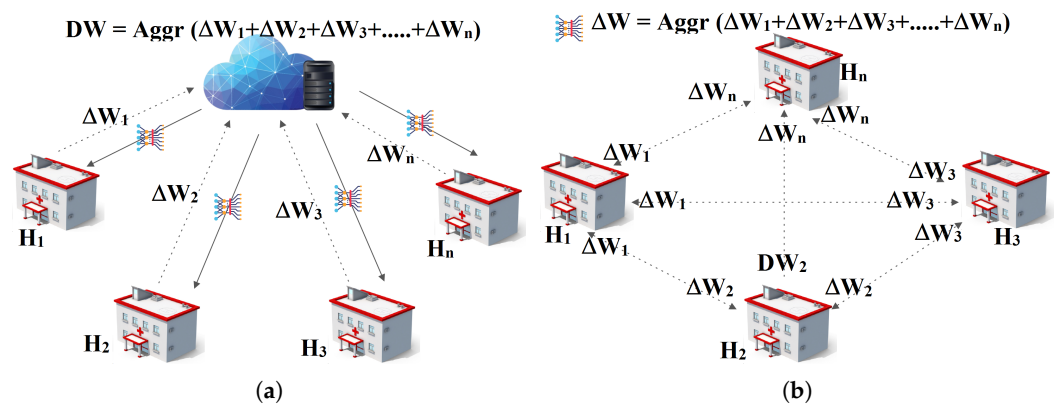

Figure 1. The two views of FL aggregation. (**a**) CFL Architecture [26]; (**b**) DFL Architecture [27].

Table 1. A comparative analysis of CFL and DFL schemes.

Parameters	Centralized FL	Decentralized FL
Model Selection	Global server (cloud) selects the model and initial hyperparameters	Peer-nodes form a consensus protocol to finalize the model
Gradient Update	Successive iterations between local (edge) nodes and global (cloud) nodes in reducing the loss function	Gradient-sharing is based on a ledger mechanism to record transactions
Algorithms	Federated Averaging (FedAvg), federated client selection (FedCS), FedProx, q-FedAvg	Gossip-based FL, incentive FL, resource-FL, BlockFL, ChainFL, BAFFLE
Communication Round	Synchronous with timestamping	Asynchronous with chunk-based processing
Supporting AI models	CNN-Unet, Deep NN, Sparse SVM, RNN, Deep-Q-networks	Partitioned ML, RL, Collaborative ML, Sparse Autoencoders, Hidden Markov models
Communication Mechanism	Client-Server, where local nodes download the initial model (downlink), and upload gradients (uplink)	Distributed with cross-verification followed by block propagation phase
Resource Management	Depends on the dataset size, available nodes energy, and CPU/cycles	Depends on peer (joint) optimization of peer models, link costs, communication delays, game theory-based optimization, relaxation-enabled resource allocation
Attacks	Data Poisoning, Label-flipping, Convergence-based attacks, Pseudorandom attacks, Secret-key sharing	Knowledge-based attacks, contracts and injection attacks, route poisoning, causative attacks, Sybil attacks, distributed denial-of-service attacks
Security mechanism	Homomorphic encryption, secure multi-party communication	Zero-knowledge proof, blockchain-based miner verification, turbo coding
Datasets	Xender trace, MNIST, Fashion-MNIST, CIFAR-10, Pascal VOC 2012	Fashion-MNIST, CIFAR-10

1.1. FL Research Trends

The above discussions indicate that the requirements of FL to support daily healthcare use cases are significant. As per the reports by NVIDIA, medical informatics would require FL to tackle the volumetric data. A recent study by Dayan et al. [28], which is published in Nature Magazine and led by general mass Birgham and NVIDIA, forms collaborative efforts of 20 diverse hospitals that train deep learning (DL)-based neural network (NN) to monitor the oxygen levels of COVID-19, and predicts the future requirements, which might require medical attention. A large pool of diverse data are required from different organizations to support accurate decision-making. Recent studies by IBM Corporation [29] indicate a merger with Healthcare, Inc., where a USD 1$ billion deal is finalized to set

up central servers for FL aggregation (the Watson supercomputer). In medical imaging, projects such as the cancer imaging archive (TCIA) support FL for image analysis and segmentation purposes. Thus, FL becomes a viable choice for distributed analytics, with the dual benefits of effective management of healthcare data-sharing rules. A study by Darzidehkalani et al. [30] suggested an increase in healthcare projects on medical imaging data, after the COVID-19 era. In magnetic resonance imaging (MRI), computed tomography (CT) scan, and real-time reverse transcription polymerase chain (RTPCR) tests for the SARS-COV2 virus, deep learning (DL) models (CNN, deep neural networks) are trained with high data volumes. In FL setups, a dataset size is typically 150 megabytes (MB). In local setups, FL harmonizes local training with parameter optimizations. Table 2 represents the abbreviations with their intended meanings used in the survey.

Table 2. Abbreviations table.

Abbreviation	Description	Abbreviation	Description
2PCC	Two-phase cross-chain consensus	GANs	Generative Adversarial Networks
5G	Fifth-generation	GDPR	General Data Protection Regulations
6G	Sixth-generation	HFL	Horizontal Federated Learning
AE	Autoencoders	HIDS	Host-based Intrusion Detection
AI	Artificial Intelligence	HIPPA	Health Insurance Portability and Accountability
APS	Artificial pancreas system	HstCon	Hasty Consensus
BFL	blockchain-based FL	IoMT	Internet of Medical Things
BLE	Bluetooth Low Energy	IoT	Internet of Things
BSNs	Body Sensor Networks	KNN	K-nearest Neighbour
CC	Cloud Computing	M2M	Machine to Machine
CCFL	cross cluster FL	MEC	mobile edge computing
CCGA	Cross-Cluster Gradients Aggregation	ML	Machine Learning
CFL	Centralized Federated Learning	MSQE	minimum square quantization error
CIA	Confidentiality, Integrity, and, Availability	NFC	Near-Field Communication
CNN	Convolution Neural Network	NN	Neural Network
DBN	Deep Belief Networks	PBFT	practical Byzantine fault tolerance
DefCon	Deferred consensus	PHR	Personal Health Records
DFL	Decentralized Federated Learning	PLS	physical layer security
DL	Deep Learning	RFID	Radio-Frequency Identification
DRL	deep reinforcement learning	RL	Reinforcement learning
EHR	Electronic Health Record	RNN	Recurrent Neural Networks
FDN	Feed Forward Deep Networks	SOM	Self-Organizing Map
FL	Federated Learning	SVM	Support Vector Machine
FL-IoMT	Federated Learning-Internet of Medical Things	VFL	Vertical Federated Learning
FRFs	Fuzzy Random Forest		

1.2. Survey Contributions

Listed below are the survey contributions:

- A reference model of FL-IoMT is presented, which discusses the FL aggregator, and the model update scenario. Insights on the communication and security perspectives of the reference architecture are also presented;
- Based on the research questions, we outline the key role of AI and trusted security solutions to support FL-IoMT. Key applications are presented to support modern healthcare scenarios- group learning, EHR data analysis, monitoring, and use of FL in medical imaging;
- A case-study, *Cross-FL* supports blockchain-based FL (BFL) mechanisms among different departments of the same hospital and inter-hospital data management. The peer learning process and the gradient aggregation strategy are discussed. Two algorithms to share updates—the HstCon algorithm and the DefCon algorithm—are presented.

Performance evaluation for these algorithms in terms of latency and accuracy of model parameters is discussed.

1.3. Survey Organization

The survey is organized as follows: Section 2 presents a functional classification (in terms of FL learning, FL networking, and FL security) of existing surveys and the state-of-the-art schemes in the IoMT domain. Section 3 presents the review methodology of the article, which poses the research questions the survey addresses. Section 4 presents a revisit of basic IoMT functions, where we discuss the different sensor types, IoMT environment structure, and computing requirements (in terms of processing at cloud, edge, fog, and dew nodes). This section allows the user the understand its basic nuts and bolts of it, which is useful to understand the FL-IoMT reference architecture in the next section. Section 5 discusses a reference architecture of FL-IoMT, with specific highlights on the learning process, aggregation policy, and security and communication paradigms. Based on existing surveys, we highlight the survey gaps. Section 6 presents ML and DL algorithms' role in assisting the FL training. Section 7 presents the inclusion of trust as a security parameter to FL algorithms and discusses emerging trusted FL concepts. Section 8 presents the different FL-based applications in a practical context. Section 9 discusses the open issues and challenges in deploying FL-IoMT and prospective solutions. Section 10 presents a case study, *Cross-FL* that forms a blockchain-based trusted solution for FL communication among intra and inter-communication setups in modern hospitals and associated stakeholders. The underlying architecture, along with consensus mechanisms, is presented in the case study. Finally, Section 11 presents the concluding remarks and future directions.

2. Related Work

In this section, we have functionally classified the surveys and state-of-the-art (SOTA) schemes in healthcare and IoMT ecosystems. We have classified the related papers in terms of FL learning models (both ML and DL schemes in FL environments), networking and resource requirements of FL nodes, and security and privacy requirements in FL designs. Figure 2 denotes the classification taxonomy. We have subdivided our related work into two subsections: survey articles and the SOTA approaches. The details are presented as follows.
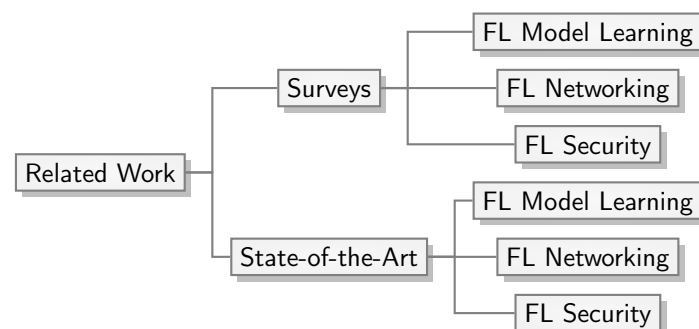
Figure 2. A functional taxonomy of related surveys and state-of-the-art approaches.

2.1. Survey Articles

In this subsection, we present the related surveys of FL in the healthcare domain. The details are presented as follows.

2.1.1. FL Model Learning

Nguyen et al. [26] discussed the potential of FL designs with the federated healthcare setups. The authors discussed FL integration with novel coronavirus disease-2019 (COVID-19) medical imaging. Antumes et al. [31] discussed the potential use cases with FL for model training, DL algorithms such as convolution neural networks (CNN), Auto Encoders (AE), and recurrent Boltzmann machine (RBM) for EHR disease diagnosis and

outcome prediction. KhoKhar et al. [32] presented a systematic survey on the FL-related schemes on healthcare imaging with secured and protected applications. The presented potential applications are cyber attack detection and recommender systems. Authors in [33] identified the challenges of the FL-based agnostic IoMT Healthcare system and EHR of the images (medical) system. A distributed FL and decentralized FL scheme are presented.

2.1.2. FL Networking

Authors in [34] presented an overview of the FL chain system with multi-access edge computing (MEC) integration. The design analysis was tested via metrics such as communication, security, protection, and incentive learning-based costs. Turjman et al. [35] surveyed the interdisciplinary characteristics of the FL-based smart healthcare applications for the IoMT domain, emphasizing communication and sensing mechanisms. ML algorithms such as K-means and decision trees are found beneficial for the study's signal enhancement of the IoMT communications. Aouedi et al. [36] presents different services in the FL-IoMT applicative verticals, such as medical imaging, healthcare diagnosis, optimization of drug compositions, and drug discovery processes. During the training process, the emphasis is laid on security and communication scenarios. The paper discusses communication cost reduction with sufficient privacy preservation.

2.1.3. FL Security

Ali et al. [37] discussed the privacy challenges in the area of FL-based IoMT architectures. The privacy threats in these architectures were identified using deep reinforcement learning (DRL), generative adversarial networks (GANs), and digital twins (DTs). In some cases, the efficiency and performance of the learning technique become affected by a malicious jammer robot in the FL-based IoMT ecosystem. Shen et al. [38] presents a systematic survey that addresses the security issues in FL when local nodes (IoT sensors) send encrypted updates to the global model. The survey addresses an in-depth analysis of differential privacy, secure multi-party computation, and trusted aggregation based on attack models. For the same, cryptanalysis of identified models is considered to identify the penetration depth of IoT devices. Based on the adversarial attacks, the open issues and challenges are discussed. Authors in [39] proposed a survey on distributed ML techniques where the local model is secured based on a multi-party cooperative mechanism. A sixth generation (6G) network is considered for FL communication, and analysis is carried out on the communication of different iterations of FL model weight updating. To optimize the results, adaptive gradient descent FL strategies are proposed for multi-clients where local DFL training is carried out collaboratively. The survey highlights the potential mechanisms to address the adaptive learning rate and avoid excessive fluctuations and model overfitting issues. Through experimental evaluations in a secured multi-party computation, communication costs are analyzed with high quantifiable robustness.

2.2. State-of-the-Art Approaches

In this subsection, we discuss the SOTA schemes of recent FL designs similarly as presented in Section 2.1. The details are presented as follows. Table 3 represents the comparative analysis of the proposed study with the existing state-of-the-art survey.

2.2.1. FL Model Learning

Sandi et al. [40] suggested a successful privacy-preserving scheme for the FL-IoMT secured misbehavior detection for the artificial pancreas system (APS). This has been achieved through blockchain technology and bidirectional long-short-term memory (Bi-LSTM). Alamleh et al. [41] proposed a framework for the FL-based IoMT to detect and mitigate the conditions that might arise due to intrusion attack vectors. The authors explored different ML/DL approaches to tackle intrusion detection systems (IDS) attack scenarios. Gupta et al. [42] focused on the anomaly detection pattern in the FL-IoMT scenario. Hence, implementing the same Hierarchical FL (HFL) that carries out aggregation

at different positions and for the various level of users for the remote patient monitoring use cases has been deployed. In the FL-based IoMT scenario, the resource utility patterns are also important, and this should not be wasted. Zhao et al. [43] proposed FL concepts in a reinforcement learning (RL) environment to train the system to evaluate and identify the gradients achieved by the clients of the large business models. Ahmed et al. [44] presented feature selections of various heterogeneous FL-based medical services, and their accuracy is analyzed. Xu et al. [45] proposed a scheme FL with minimum square quantization error (FED-MSQE). The FED-MSQE proposes a minimal square quantization in the federated ecosystem. The results show significant improvements for multi-FL silos.

Zhang et al. [46] presented an optimization scheme for fine-tuning hyper-parameters of complex DL models such as deep belief networks, Boltzmann machines, and CNNs. The data become hierarchical with more nodes, increasing the model's complexity. The authors propose a multi-layer extreme learning machine (ELM), which selects the important attributes in such cases. The scheme accelerates the DL model's convergence time. A non-iterative version is presented over stacked autoencoders, residual models, where middle connections are made scalable. It is also established that randomized algorithms are more optimal for the generation of hidden layer parameter selection in the case of CNN's. Randomization depends on task and data distribution. Authors in [47] discussed ELM for hidden feed-forward networks. ELMs reduce the modeling errors with the usage of the squared loss function. The work suggests a Gaussian error distribution strategy to induce robustness in the scheme, and the effect is compared for data with Gaussian and non-Gaussian noise. An objective function is formed for a mixture of Gaussian distribution to approximate incoming data.

2.2.2. FL Networking

Sanyal et al. [48] proposed a federated filtering framework for IoMT. The scheme takes advantage of the FL scenario, which enhances the system's energy efficiency and also provides privacy in the resource-controlled IoMT ecosystem at minimal latency costs. Furthermore, AI models may be trained on dispersed data due to FL. Consequently, a substantial amount of IoMT data can be utilized without data sharing. FL-based schemes with privacy preservation are proposed in IoMT ecosystems, where authors discussed security mechanisms for the physical layer. Thus, physical layer security (PLS) mechanisms are proposed; for example, authors in [49] explored this limitation and presented a scheme for power control uplink FL-IoMT. The scheme enhances the performance through the detection of the malicious jammer robot.

2.2.3. FL Security

Ferag et al. [50] did a comparative study on various FL-oriented DL schemes for the IoT application associated with cybersecurity. Examples include unmanned aerial vehicles (UAVs) for healthcare supplies, surveillance, and management. Authors in [51] developed deep learning algorithms in FL-based setups to trace heart activity. To preserve the sensitive attributes, anonymity and encryption schemes are considered. Dai et al. [52] presents the survey and case study of the multi-layer blockchain-based IoMT system. The system helps the healthcare organization manage COVID-19 patients and contributes to telemedicine and remote healthcare assistance. Cheng et al. [53] designed an FL-assisted expectation maximization Gaussian mixture model, and the scheme is validated through the IoMT data with minimal data leakage.

Table 3. Comparative analysis of the proposed survey with existing state-of-the-art surveys.

Author	Year	1	Parameters 2	3	4	Objective	Adavantages	Limitations
Proposed	2022	Y	Y	Y	Y	A comprehensive discussion of FL techniques to ensure security and privacy is presented	Applications of FL in IoMT domains are discussed, which is validated through a case study of BFL IoMT	A solution taxonomy of FL in the IoMT domain is not discussed
Nguyen et al. [26]	2022	Y	N	Y	Y	The authors surveyed recent healthcare advances in FL and presented the security and communication perspective	Recent FL designs are discussed based on computing resources, privacy, and incentives	Provable guarantees of FL security, data heterogeneity, and non-iidness are key limitations
Antunes et al. [31]	2022	Y	Y	Y	N	Authors have discussed FL adoption to EHR	FL with generative adversarial networks (GAN) is exploited for data analysis.	The secrecy of training data are still an open question that needs more research and suggestions.
KhoKhar et al. [32]	2022	Y	N	Y	Y	The secured FL training model is surveyed in IoMT	The authors considered an image-processing application and presented methods for enhancing FL security.	Privacy issues of FL and data aggregation issues require further discussions.
Khan and Alkaabi et al. [33]	2021	Y	Y	Y	Y	Surveyed AI potential to integrate with emerging cyber-physical technologies for smart communities	Emerging communication networks such as sixth-generation (6G) are combined with edge computing in healthcare and FL for decentralized analytics.	The two major concerns that need to be solved are communication costs and clients with heterogeneous characteristics.
Nguyen et al. [34]	2021	N	Y	Y	Y	Communication, incentive, security, and resource distribution are the important metrics addressed through an FLchain design.	Authors surveyed applicative domains to support edge data sharing and crowdsensing and explored its potential with FL learning for healthcare	Convincing mobile users to join the FLchain process is a fundamental challenge in real setups, as a mobile user serves as both a training and miner node.
Turjman et al. [35]	2020	Y	N	Y	N	The survey presents insights about new FL perspectives in the IoMT ecosystems.	Smart healthcare applications in classification, sensing technologies through wearables, and inter-connectedness in FL network communication are explicitly discussed.	Fewer insights are presented for the design of new protocols and open communication standards, which allows interoperability between local FL nodes with the global trained model
Sheng et al. [38]	2020	N	Y	Y	Y	Discusses backdoor and inference attacks as potential threats to FL systems.	Privacy and security countermeasures such as differential privacy, secure aggregation, and secure multiparty computation are discussed.	FL is still susceptible to data membership inference and backdoor attacks.
Wu et al. [39]	2020	N	N	N	Y	The survey presents insights about new FL perspectives in the IoMT ecosystems.	The model's privacy, security, and sensitivity to hyperparameters are to be improved.	The accessibility of massive data contexts and the convergence efficiency of the training model are still restricted.
Sanyal et al. [48]	2019	N	Y	N	N	FL-based filtering mechanisms are discussed for IoMT setups for prediction at central fog servers.	For IoMT aggregation, a lightweight, fully unsupervised local subroutine and filtering algorithm are introduced	Research challenges in procuring effective and unbiased data collection at IoMT testbeds at local setups are not discussed, which can help formulate a general relationship between decision accuracy and perturbation error.
Ferag et al. [50]	2021	Y	N	Y	N	Authors reviewed FL-based security and privacy solutions with blockchain integration	Blockchain-assisted FL schemes are selected to operate in conjunction with DL for cybersecurity applications	Designing a specialized IoT framework based on FL remains an essential research topic that must consider the underlying IoT architecture.
Can and Ersoy et al. [51]	2021	N	Y	Y	Y	Federated deep learning (FDL) algorithms are surveyed to address data heterogeneity and collection process for heart activity data.	IoT-based wearable monitoring schemes with FL are examined, and dataset challenges are presented.	The absence of control groups in the experimental setups lowers the validation functionality of wearables. The discussion on corrective measures is not carried out in the study.

Author	Year	Parameters 1	2	3	4	Objective	Adavantages	Limitations
Dai et al. [52]	2020	Y	N	Y	Y	To address the security and privacy issues with IoMT systems, the authors have presented a blockchain-enabled IoMT.	Blockchain technology can strengthen security and safeguard user privacy in IoMT systems.	It is not easy to improve performance by implementing more scalable consensus algorithms.
Cheng et al. [53]	2020	N	Y	N	Y	Blockchain-assisted FL verification architectures are surveyed, and for aggregation, homomorphic encryption is studied	Designing homomorphic encryption schemes during the FL aggregation process lowers the privacy risk and allows personalization in healthcare setups	The authors did not address the importance of credential-based authorization, which simplifies the complexity of the overall process.

1: Review Method, 2: Case Study, 3: FL architecture, 4: FL Applications, Y—shows that the parameter is present, N—shows that the parameter is absent.

Wang et al. [54] discussed the consideration of edge computing for the FL-IoMT and designed an architecture consisting of the lightweight privacy protocol with minimal overheads. Wu et al. [55] presented an incentive mechanism for the FL-IoMT, where the incentive proposals are proportional to the effectiveness of the learning mechanism, and it trivially depends on the data size of local training, privacy budget, and information asymmetry. Hence, to deal with such a problem, Lakhan et al. [56] proposed a security and privacy scheme that works with minimal resources. For validation, the authors have used metrics such as delay and cost. To manage the security and privacy in the COVID-19 situation, Samuel et al. [57] have introduced the blockchain concept in the FL-based IoMT. The scheme ensures trust, immutability, data availability, and information security. Choudhury et al. [58] offered a plan to cope with the challenges associated with centralized learning. The concept of differential privacy was also imposed on the Fl-IoMT-based case study. The performance was high and found to be usable with the IoMT framework.

2.3. The Necessity of the Survey

The market size of FL in HA healthcare analytics from 2020 to 2030 will rise to ≈12% [59]. Researchers have shifted towards incorporating FL solutions with modern learning capabilities that address data sharing and privacy issues. Most surveys and SOTA on FL discuss optimization of model parameters, gradient sharing with low power, and security issues. For the same, articles discuss CFL and DFL schemes where cloud, edge, or hybrid aggregation is presented. The other direction is towards the fairness and incentive mechanisms of DFL schemes. In edge FL schemes, numerous surveys discussed issues of caching at network devices to improve local learning. The articles have discussed these issues from model training and security viewpoint but have not addressed the specific challenges in the IoMT domain. Thus, the proposed survey underlines the integration of FL with IoMT with a high-level overview of the proposed architecture. Furthermore, recent advancement in IoMT has propelled the design of optimal FL algorithms that caters the communication and trust in such systems. Thus, blockchain integration with FL is a viable choice in distributed setups, and the proposed survey discusses a useful case study.

3. Review Method

This section represents the review methodology and highlights the research questions to assist the study.

3.1. Research Questions and Its Study

IoMT devices generate a vast amount of data per second, consisting of permutation of different sensors with time. The variety of this data, combined with IoMT devices, is limited to power and resources, particularly implanted medical devices. This places a large computational load on typical ML algorithms and limits their usefulness in IoMT devices. As a result, new tactics are necessary to utilize ML approaches effectively. It is critical to comprehend the IoMT system's security and privacy challenges and their related remedies based on ML approaches. Furthermore, it is critical to analyze the efficacy of currently deployed ML approaches, the strength of solutions provided to the IoMT system, and its difficulties. We discovered that there had been little attention devoted to these concerns in the literature. Consequently, this study includes a systematic literature review, which aims to expose the strengths and limitations of previous research in this area, followed by the creation of solid improved techniques. Table 4 highlights the research question identified along with the objective that assists the survey.

Table 4. Research question of the proposed survey.

Q.N.	Research Question	Objective
RQ1:	What is the need for FL in IoMT?	To understand the important features of FL that improve the privacy and security of the shared EHR among different stakeholders in the IoMT setup.
RQ2:	What is the drawback of the current centralized IoMT ecosystem?	To explore privacy attacks in the centralized IoMT system and how a decentralized FL-based system ensures the privacy of sensitive information.
RQ3:	What are the challenges of FL in the IoMT ecosystem?	To formulate the open issues and challenges to seek future directions for decentralized healthcare.
RQ4:	How does FL-IoMT benefits healthcare?	To conceptualize the use of FL with the integration of IoMT in the healthcare setups.

3.2. Review Methodology

The literature survey is employed on the most recent and relevant computer science healthcare databases such as IEEEXplore, PubMed, WHO database, Science Direct, and ACM Digital Library. This study also includes websites, blogs, reports, books, and patents in the healthcare and IoMT domains. An exhaustive literature survey is carried out on technologies related to AI, ML, DL, FL, and IoMT with their integration in the healthcare domain. The search is narrowed down by initial compilation and filtering. Figure 3 presents the search criteria including some common synonyms. With the defined search keywords, a number of articles is identified and based on the date and year of the publication. The research repositories are searched based on the combination of words "Federated Learning", "Internet-of-Medical-Things", "Healthcare", "Federated Learning for healthcare", "Security in Internet-of-Medical-Things", and "Artificial Intelligence in Healthcare." We also searched the academic databases for classification algorithms, applications, security parameters for FL, and their applicability in the healthcare domain.
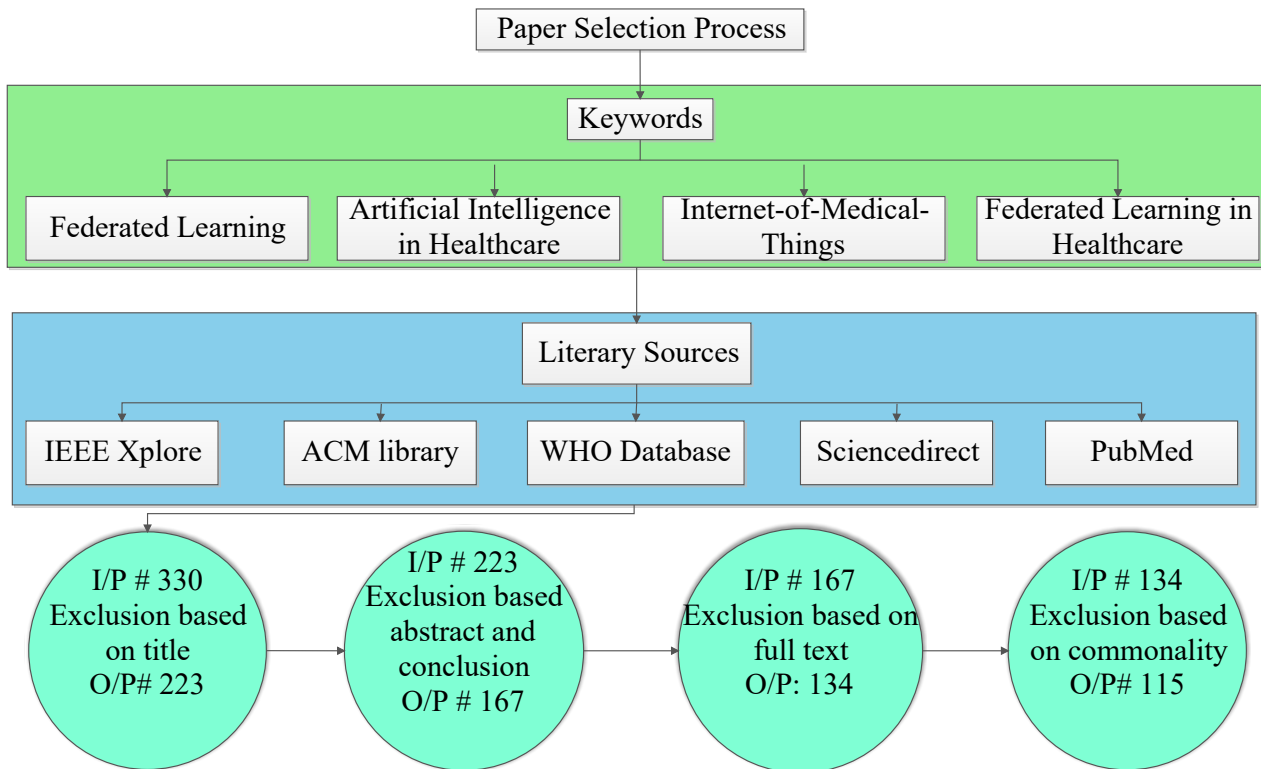
Figure 3. Survey methodology and Inclusion–Exclusion.

4. Nuts and Bolts of IoMT: Sensors, Node Setup, and Aggregation Policy

In this section, we present a background of the type of sensor nodes used in IoMT ecosystems. The nodes generate data required to undergo computational analysis at different computing capacities. The collected data are trained at cloud, fog, multi-access edge, or dew layer. Depending on application-specific requirements, data are trained via CFL and DFL models at different nodes. We begin our discussion by discussing different IoMT sensors and general IoMT setups and then shift towards the computational (processing) of collected data at different aggregations in a layerwise manner. The details are presented as follows.

4.1. IoMT Sensors

We have categorized IoMT sensors into two categories, normal medical sensors and biosensors. The details are presented as follows.

4.1.1. Normal Sensors

IoMT allows different sensor nodes to communicate over a wireless channel to remotely and continuously monitor the patient's health and subsequently form clinical indicators for HA. With constant monitoring, the patient's medical expenses are reduced as his body indicators are measured daily. IoMT has two categories of sensor nodes: normal medical sensors and biosensors. Normal sensors are mainly used in patient diagnosis and therapeutic support, such as blood pressure and glucose sensors (strip-based). Another category is patch sensors, which a patient wears like a torso, and critical indicators such as heart, pulse activity, sleep activity, and body fall movements are measured. These sensors are useful for elderly persons as they can measure triggers or spikes during fall detection analysis. Another category is connected sensors such as wearables and invasive sensors. Wearable sensors form direct contact with human skin and monitor individual activities (fitness trackers, watches, smart clothing, and others). Invasive sensors (such as electrocardiograms and electroencephalograms) generate signals that measure brain signals and eye

gaze motion to check patient alertness. A use-case application is when a person feels sleepy while driving. Then, it can monitor the person's eye pupils to raise alarms or notifications.

4.1.2. Biosensors

Biosensors are short-form for biological sensors and are normally formed by using any transducer with a biological element, which can be enzymes, nucleic acid, lectins, plant proteins, tissue slices, or antibodies. These elements, popularly known as bioelements, interacted with analytes and converted the biological reaction into electrical signals, which are measured as inputs [60]. Different categories of biosensors are used, such as immunosensors, canaries, biochips, glucometers, and others. Any biosensor mainly consists of two elements—the natural elements and the electronic element. The generated signal is mainly electrical, but in many cases might be thermal or optical. Mainly, biosensors are small, so they can be implanted inside human bodies (normally the size of rice grain) or taken as pills and are soluble in human body fluids [61]. Once inside the body, the biosensor electrical component reacts with gastric body fluids to trigger the electrochemical reaction. Once triggered, a digital voltage or spike is generated, which is measured by devices.

Recently, new biosensors have been designed to detect ribonucleic acid (RNA) viruses such as CRISPR-Cas9. Other categories include nucleic acid-based, antigen-Au/Ag nanoparticles-based electrochemical biosensors, optical biosensors, and surface plasmon resonance. Biosensors can identify dangerous metal concentrations in the water thanks to their metal-specific electrodes. In addition, it can recognize harmful pathogens and find elements of bio-recognition such as enzymes, antibodies, or biomolecules [62]. Biosensors also offer the chance to improve the post-operative care process. Table 5 summarizes the common types of sensors used to empower the Internet of Medical Things (IoMT). It presents the different categories of normal sensors and biosensors with descriptions, different types, and their potential use-case.

Table 5. Types of sensor devices in IoMT ecosystems.

Sensor Type	Name	Description	Sensor Name/Types	Use-Case
Biosensors	Electrochemical biosensor	This sensor uses an electrode to immobilize the electron and bio-molecules which causes a transducing biochemical event.	potentiometric, amperometric, and conductometric biosensors,	It is used to directly monitor the activity of the living cell by converting biological events to electrical signals.
	Physical biosensors	It senses the biological change such as a change in mass, fluorescence, and resonance of the patient.	piezoelectric biosensors and thermometric biosensors	Patients temperature, room humidity, and light intensity
	Optical biosensors	Based on the optical measurement principle, and use fiber optics as well as optoelectronic transducers.	Surface plasmon resonance (SPR), SPR imaging (SPRi), and localized SPR (LSPR)	Detection of deoxyribonucleic acid (DNA) hybridization, clinical diagnostics, bio-defense, food testing, clinical biomarkers, and toxin screening
	Wearable biosensors	These are portable integrated sensors and are easily implanted on the patient's body with the help of tattoos, body implants and clothing.	saliva, tear, and body sweat-based wearable biosensors	bio-sensing to detect vitamins and minerals, obtaining information regarding biological fluid within a body cavity.

Table 5. *Cont.*

Sensor Type	Name	Description	Sensor Name/Types	Use-Case
Normal sensors	Flow sensors	Flow sensor accurately measures the flow of liquid and gases in ventilators and anaesthesia workstations	contact and non-contact flow sensors	Used as a respiratory gas monitor
	Fingerprint sensors	Captures a digital image of the fingerprint pattern	optical scanners, capacitance scanners, ultrasonic scanners, and thermal scanners	Fingerprints can be used to identify an injured person during an emergency, which can be checked against public healthcare databases for vital health information of the person.
	Temperature Sensor	Temperature from a source and converts it into human and device understandable form.	Thermistors, resistance temperature detectors and thermocouples	Measuring blood or body fluid temperatures
	Pulse Heart Rate Sensor	Heartbeat sensor can sense a change in the volume of blood flow in the vessel with the help of red and green lights	photoelectric pulse wave and blood pressure measurement.	Sleep tracking, anxiety monitoring, health bands, remote patient monitoring, alarm systems

4.2. IoMT Node Setup

The section discusses IoMT sensor nodes setup and required communication protocols. IoMT network scales itself in three categories, namely, the IoMT body setup (on-body), hospital setup (local FL client systems), or community setup (a shared group of local FL clients). In the case of in-house IoMT (both on-body and hospital IoMT), the sensor nodes send clinical data to servers to assist decision-making capability. These setups are computationally constrained and thus have small datasets to execute the ML and DL models. These models can send patient data from residences and hospitals to clinics (doctors) or healthcare practitioners. The underlying communication stack employs low-powered communication protocols, and the energy consumption is ≈ 1 joule/sec. The devices communicate with a power range of 0.4 watts to 0.7 watts. The transfer network varies from personal area network (PAN) to master–slave piconets and scatters nets (in the case of Bluetooth low energy). The network range is mostly indoors and falls under the IEEE 802.15.4 protocol. The other protocols mostly used are 6lowPAN, a routing protocol for low-power and lossy networks (RPL) supported by an IPv6 address. The device identification is made through protocols such as uCode, and uniform resource identifiers, supported by transport/application layer services such as message queue telemetry transport (MQTT), constrained application protocol (CoAP), and lightweight machine-to-machine (LwM2M) to interface with web services. For faster connections, WebSocket is another preferred choice. Other suitable networks are Z-wave, Zigbee, low power wide area networking (LoRAWAN), LoRAPAN, radio frequency identification (RFID), wireless highway addressable remote transducer protocol (Wireless HART), and others.

In-house-IoMT prevents hospital readmissions by constantly monitoring and identifying critical health indicators, known as telehealth. Telehealth allows pristinely discharged patients to contact doctors in remote mode. Community IoMT refers to utilizing IoMT devices over a larger region or city. Mobility services, for instance, use equipment to track patients when traveling in a car. Paramedics and other first responders monitor patient statistics beyond the boundaries of the hospitals using emergency service intelligence systems. It sets up point-of-care devices, which are used to diagnose health problems. In long-range communication, cellular networks are a viable choice (low-powered long-term evolution, 5G and beyond), long-range radio (LoRA), SigFox, and others. LoRa supports a range higher than 10 km, whereas SigFox supports a range of 10–20 km in rural and urban conditions. Sigfox also supports the advanced encryption standard of 128-bit key size (AES-128) for communicating data to the receiver nodes, with a high interference immunity due to adaptive modulation.

4.3. Aggregation Policy

In this subsection, the aggregation policies are discussed in detail. We form a comparative analysis of cloud-based, edge-based, fog-based, and dew-based FL aggregation. Table 6 presents a comparative analysis of the varying degree of computing at these layers in different FL parameters. The section above discusses that a central server (global server) is called an aggregator in CFL setup [63]. By design, the local nodes retain data and perform local model training. The local updates are shared with the cloud aggregator, which fuses the gradients to form an optimal weight (gradient) strategy for the global server [64]. As cloud servers have high computing resources, the global model training is supported (with high backend GPU accelerators). For example, consider the scenario where local hospitals execute a CNN model for COVID-19 detection on MRI images. In such cases, a cross-device scenario is applicable where a large number of client nodes (>100) participate in the training process [65]. In such cases, FL models such as federated averaging (FedAvg) are a popular choice [66]. The global model is normally trained via a stochastic gradient descent (SGD) model to minimize the classification and prediction bias. Every party has a small dataset and constrained operating capability (mainly mobile devices). The communication is synchronous, and the global training updates are sent back to local nodes for model optimization. In edge-based FL, nodes communicate with the local edge server as aggregators. Edge is the closest to end devices, and it allows for computing to be carried out and a small amount of data to be stored directly on devices, applications, and edge gateways. Thus, edge aggregation supports the DFL scenarios in a better manner. In such cases, FedAvg does not cater well to requirements owing to simple averaging, as data might have diverse parameters affecting global convergence and learning rates. The local objectives also vary significantly in these models and might become bounded by local optima instead of global solutions. One solution is to increase the iterations to improve convergence, but it would significantly consume useful edge resources, making the network non-scalable to support the cross-device scenario.

In edge aggregation, FedProx is considered a better solution that adds a proximal term to the local function defined by FedAvg [67]. It also addresses the statistical heterogeneity and non-IID data distribution with minimal computational constraints. However, when the edge aggregator faces bottlenecks, a fog computing layer is introduced, which acts as the resource manager between the edge and the cloud node. This simplifies the logistic operation at the local nodes, as fog nodes analyze important communication parameters. Thus, the data transfer has low latency, even in low bandwidth networks. However, a joint optimization strategy is required for resource allocation and gradient selection at the fog layer. At the fog layer, it is necessary to virtualize resources, administer the edge aggregator policy, store sensitive data, and connect multiple edge nodes using the cloud.

Recently, dew computing [68] is considered an optimal fit to support the DFL scheme, owing to the property of independence and collaboration. In dew computing, the independence property supports the ability of the dew layer to conserve partial learning functionality of local models, where connection loss to the edge aggregator is normal. Dew supports a hierarchical aggregation to the edge node, where the DFL nodes learn the local models collaboratively. Dew layers support dew intelligence, where micromanagement of data selection is carried out asynchronously. This principle is denoted as the dew layer's collaborative property, and many local devices create a shared knowledge of data, where the updates are shared via shared links between DFL nodes. Thus, dew intelligence is suitable in IoMT applications, as it supports the highest degree of personalization during local training. In the dew layer, authors in [69] presented a scheme, FedHealth that forms a collaborative learning model for shared EHRs among multiple nodes. The Dew layer supports another learning paradigm, split learning [70], where a common model trains different data modalities. It solves the problems of vanilla FL, proposed by Google, where local gradients are ensembled at the central global node. The convergence time of vanilla FL (for example, GBoard) is high. Another FL approach, community-based FL (CBFL) [71],

is also supported by the dew layer, which has high predictive accuracy and a high degree of data privacy.

Table 6. A comparative analysis of IoMT devices at varying levels of computing nodes (Cloud, Fog, Edge, and Dew).

FL Parameters	Cloud Computing	Fog Computing	Edge Computing	Dew Computing
Latency	50–150 ms	15–20 ms	5–10 ms	1–2 ms
Architecture	Centralized	Distributed	Distributed	Hybrid
Hardware	Routers, layer 3 switches, bridges, and gateways	Bridges and gateways	Access points, base stations	base stations, dew servers
Nodes	Mostly Servers with high-end computing	Portable handled devices (Laptops, Mobiles)	Portable with limited capacity (Mobiles)	Embedded Controllers (Sensors with proportional integral derivative control)
Distance	Far from the edge (No. of Hops—1 to 30)	Network closest to the edge (1–10 Hops)	At the edge (1 or 2 Hops)	On Device (at most 1 hop)
Access Network	Both Wired and Wireless	Wireless	Wireless	Both wireless and wired
Data Analysis	permanent storage, less time-sensitive data processing	Real-time picks whether to send to the cloud or process locally	real-time, immediate decision-making	Real-time and offline (disconnected) computation
Scalability	High, Easy to Scale (No Limit)	Scalable within network	Hard to scale	No scaling
Computing Cost	$400 monthly to $15000 based on the type of model	$100 monthly to $999 based on the type of model	$199 monthly to $1299 based on the type of model	$50 monthly to $100 depending on model types
Usage	Analytics, processing, warehousing	Local data analytics, processing, storage	Real-time processing, visualization, micro storage	Real-time processing, micro storage, and miniature protocols required
Virtualization	Yes	Yes	Yes	No

5. FL-IoMT: Fundamentals and Key Technicalities

This section provides the fundamentals and key technicalities through an FL-IoMT architecture and presents the FL learning process, aggregation, and the security and communication process.

In medical setups, wearable sensors communicate with gateway edge nodes that leverage machine-to-machine (M2M) connections driven by IoT protocols. It enables various links to communicate with each other over the Internet [72]. Different AI models are designed over these indicators and are utilized to find statistical inferences to make informed decisions. Due to data inconsistencies, we use interpretable analytics that provides valid reasoning to AI models and allows explainability for AI models [73]. The local data are sent to the central servers for analysis, but it increases the computational requirements four-fold [74]. According to a recent study, by 2025, the network's edge will process $\approx$ 950 ZettaBytes of data in digital reports and files. On the other hand, the global data center traffic is expected to reach 21.6 ZettaBytes [75]. With massive data collection, security is an important parameter and requires privacy preservation among security channels and third-party applications. To cater to this, recent researchers have suggested the use of privacy preservation [76], crypto primitives [77], and trust building using blockchain [78,79].

5.1. A Reference Architecture for FL-IoMT

This section represents a typical FL model training procedure in IoMT ecosystems. Figure 4 presents the architecture which specifies the overall communication process of FL-IoMT. The architecture consists of two main components-IoMT sensor nodes that capture the data and the other is the aggregation server. We consider that *n* hospital entities participate collaboratively to train their local models. The IoMT devices (sensors) and the global servers are positioned at the service access point to aggregate the local updates. The different architectural components are discussed as follows.
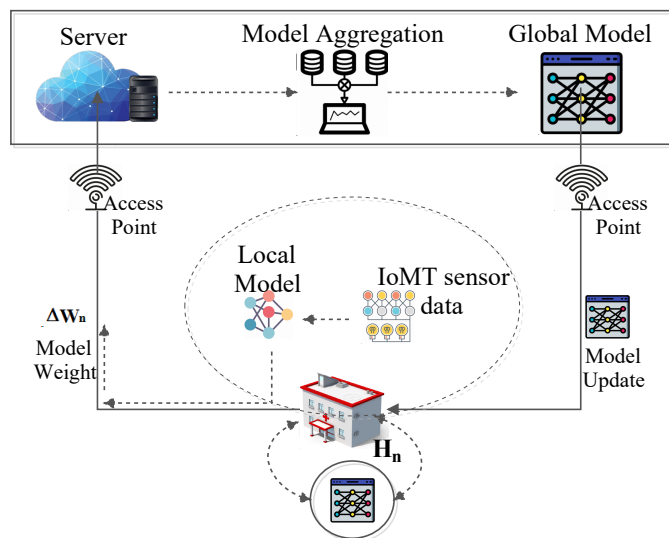
Figure 4. FL-IoMT architecture and communication process [80].

5.1.1. IoMT Client

Every H_n receives the initial model M_0. Each local model M_0 is trained with local IoMT data D_n, and model weights ΔW_n are computed by minimizing the loss function f^n. For example, consider a linear regression model where the weight optimization function is denoted as $f^n(W)$ and can be written as follows:

$$f^n(w) = 1/2(X_i - Y_i)^2 \tag{1}$$

where $\{X_i, Y_i\}$ are the input pairs. The objective function can be fed to a back propagation model, which can optimize the weights more efficiently. Once the ΔW_n is computed, it is sent to the aggregation server via the access point.

5.1.2. Aggregation Server

The main server (global) collects the model updates in the form of gradient ΔW_n. Based on the received gradient, the server computes the new version of the global model, which is depicted as follows:

$$M_g = \frac{1}{\sum |D_i|} \sum_{i=1}^{n} |D_i| |W_i| \tag{2}$$

where D_i and W_i are the local data and local model weights at the ith hospital, respectively. The constraint is specified for all the H_n, which needs to use and send to the same learning model after every training iteration. After training the new model, a new ΔW_n is sent to compute the next round of global model M_g.

5.1.3. The FL Learning Process

FL uses statistical data models to train while keeping the data safe at a local site. With the rise of data and the computing capability of devices, the practical FL learning process in IoMT has taken significant leaps. However, owing to privacy concerns, the data need to be processed locally, and the computations are supported at edge nodes. This improves the latency of FL communication rounds [81]. The important FL use cases normally include wearables which require quick decisions in critical life-saving emergencies. The challenge with FL remains to construct a single statistical model from the raw data saved on millions of devices over the Internet. The FL learning process's final aim is to minimize the global loss function, presented in Equation (3) as follows:

$$f = min(\Delta W)\frac{1}{K}\sum_{i=1}^{n}\sum_{j=1}^{k_n} f_k(\Delta W) \tag{3}$$

where ΔW represents the weights of the global model, and K is the total data points from IoMT devices. Each n IoMT device has its local data set d_n of k_n data points.

5.2. FL Categorization Based on the Data Partition View

In FL, the data are partitioned between multiple client nodes. The scheme in which the data are partitioned significantly affects the aggregation models. The partitioning view can be categorized into three categories, namely, vertical FL (VFL), horizontal FL (HFL), and federated Transfer Learning (FTL). The details are presented as follows. These are categorized according to different samples, and feature space lies in the distribution pattern. The same sample space refers to the set of the possible outcomes of a random experiment, while the same feature space refers to the set of features used to categorize the data. Figure 5 represents the overview of HFL, VFL, and FTL.
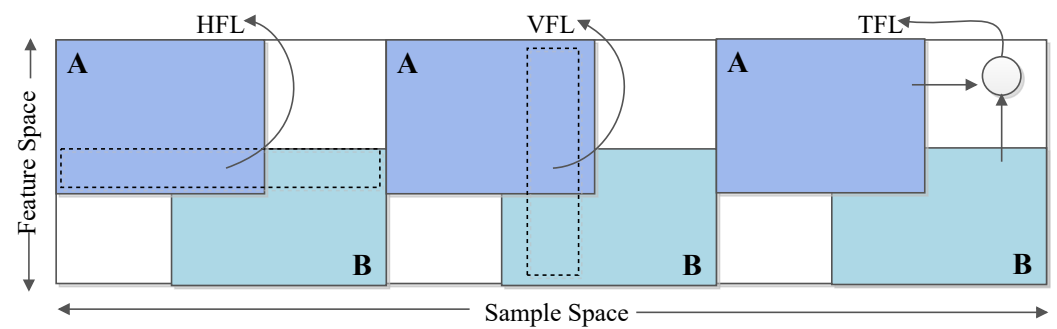

Figure 5. Categorization of FL based on Data Partition.

Horizontal FL

HFL is applicable when the feature space of two different datasets of users A and B overlaps more than the users' closeness. For example, consider two different IoMT devices from different regions that might have some overlapping data collected from their respective patient node. In HFL, healthcare stakeholders develop the distributed global model using datasets with the same feature space but different sample spaces. The same AI model might be used by the local FL players (artificial NN) to train their respective datasets. Without accessing the local data directly, the server might make a global update by including local changes received by local participants. The detection of speech disorders is considered an example of an HFL example. Many people say the same phrase (feature space) on their cellphones with varying voices (sample space). A parameter server averages the local speaking information to create a worldwide speech recognition model. Thus, an FL-enabled speech (voice) recognition system would become more personalized for a specific user (owing to local samples) but would also incorporate the averaging information of other connected users. In HFL, the clients compute their local gradients and upload the information, but the collection nodes might leak private information. Thus, adding a random noise with the local gradient is a viable approach. In the case of DFL approaches, the authors have proposed an HFL framework named BlockFL [82], which uses blockchain-based ledgers from where transactions from mobile clients are recorded to update the local models. Another scheme termed MOCHA [83] is designed to solve security and multitasking issues of edge aggregation, and the scheme assures fault-tolerance of distributed nodes during FL training.

5.3. Vertical FL

VFL specializes in federated training of medical datasets with similar sample spaces but distinct data feature spaces. VFL is applicable when the feature sets of user A and user

B in the database overlap less than the user similarity of A and B. For example, consider two IoMT devices at one hospital entity, which collect patient body movements, and postures, and predict fall detection, sleep cycle, and other details. Patients are common (user similarity), but each has different clinical indicators owing to the difference in body movements. During local training, it uses solutions based on entity alignment in combination with encryption measures to alleviate the problem of data sample overlap at distant clients. A shared learning model among entities, such as health centers and insurance companies, might be an example of VFL in IoMT applications in a smart healthcare system. In this frame of reference, a VFL approach is used by a hospital and an insurance company that serves patients (same sample space) to collaboratively train an AI model for smart treatment choices utilizing their datasets, such as healthcare costs at insurance companies and previous medical records at hospitals. ML models such as classification [84], gradient descent computation, and linear regression are applicable in the case of such vertical divisions. In DFL, peer training strategies such as SecureBoost are proposed, where all peer nodes summate the user features to train their models [85]. This strengthens the decision-model accuracy, and data losses are minimized. To assure data privacy during communication, Paillier and addictive homomorphic encryption are proposed by users [86].

Federated TL

FTL works with datasets with various sample spaces and feature spaces, unlike the VFL systems. It is applicable when two users, A and B, rarely overlap, and we do not find any common segment. In such situations, we use transfer learning to find similarities in data, such as two different IoMT devices of two different regions sensing different body movements (for example, one is monitoring the heart pulses, and the other is measuring the patient body temperature. A transfer learning approach estimates feature values by incorporating several feature areas into a single representation, which is trained using regional datasets. Random masks and other encryption techniques provide extra privacy and security between clients and servers during the gradient transfer. FTL might help in illness identification in smart healthcare by working with countries that have a variety of facilities with a variety of people (sample space) and varied therapeutic plans (feature space). FTL may improve the shared AI model output that enhances diagnostic accuracy.

5.4. Security Paradigms in FL

Both conventional and zero-day attacks that might compromise IoMT devices. This is mainly due to the lack of well-established security standards and safeguards in device manufacturing, device configuration, and underlying IoT-based communication protocols. IoT devices are small embedded chips with limited capacities, and thus they have limited computing and power capacity. Modern cryptographic algorithms employ multiple levels of security (in terms of rounds, key generations, and hash computations) and thus require high processing and computational power for robust security in large organizational perimeters. Moreover, with heterogeneous networks and distinct protocols at each tier, finding a unified security solution that addresses all requirements is daunting. According to Statista (The statistics portal), the number of medical IoT devices in the European Union (EU) might reach 25.8 million by 2025. Besides the ever-increasing range of smart medical equipment and low-cost wireless sensors, privacy and security issues have surfaced as major concerns. Furthermore, the volume of data generated rises in lockstep with the number of internet-connected devices. Statistic estimates that IoT devices will generate 79.4 Zettabytes by 2025 [87]. In reality, privacy and data disclosure are now the most important issues IoMT infrastructure faces, as IoMT devices and their data are vulnerable to cyber-attacks. The privacy and security requirements for IoMT are distinct from those for conventional networks, often alluded to as the confidentiality, integrity, and availability (CIA) triad.

To assure the security and privacy of FL data, encryption strategies such as homomor-

phic encryption (HE), differential privacy (DP), and secure multi-party communication (MPC) is preferred. The details are presented as follows:

1. *Homomorphic Encryption*—Normal encryption strategies require the FL users to decrypt the data based on the shared key and the encryption algorithm. In such cases, there are high risks of key disclosure; thus, in secure FL design, HE is a preferred choice. In HE, any user (FL client), denoted by C, encrypts the training data D using its private key Pr_C. The encrypted data are then sent to other users or global cloud servers. The other user forms an evaluation function, normally denoted as *Eval*, and the sending user queries a homomorphic function $f(.)$ to server *Eval*. The server S or other local users reply through a function on the same data D, denoted by $Enc(f(D))$. Once the sending user C receives $Enc(f(D))$, it operates decryption on the function, which is represented as $Dec(Enc(f(D)))$, which returns the data D as output. Thus, any user can compute the function and process the encryption without prior knowledge of the encryption key Pr_C. In the process, users not authorized with the server *Eval* do not know the homomorphic function $f(.)$, and thus data are not disclosed. Thus, there are low risks of leakage of gradient values during the exchange process, and the model parameters are shared securely between multiple clients (in the case of DFL) and a global server (in the case of CFL). Different variants of HE are proposed, such as fully HE, partial HE, and somewhat HE. Fully HE supports the construction of functions with the desired functionality, and the program never forms any decryption at any node. Fully HE has high overheads (as function operators are complex); thus, for edge-based aggregation, partial HE uses fewer function parameters for computation (mostly addition and multiplication operations only). Somewhat HE only works on a subset of total nodes in the network and is used for local node training.

2. *Differential Privacy*—DP is a security mechanism where information can be shared publicly depending on local data patterns, but explicit identifiable attributes are hidden. Thus, DP assures that, from a given released public healthcare dataset, the attributes are non-identifiable, which limits the disclosure level. DP is used when the IoMT ecosystem supports a high degree of confidentiality, as it mixes the information of other neighboring users in the shared dataset. Any record in a dataset is insensitive to overall statistical changes in the complete dataset. In DP, the process is to normally add a noise factor to the output during the computation of the gradient function. The preferable noise distributions used are Gaussian and Laplacian distributions. However, suppose a high amount of noise is added to the dataset. In that case, it undermines the statistical distribution of the dataset, and thus there is a trade-off between high privacy and dataset stability. Normally, in edge-based aggregation, owing to computational requirements, DP is quantized and compressed to improve the communication performance of the local network. A popular variant of DP, named ϵ-DP, is preferred. As inputs, we supply $\epsilon \, \epsilon \in (R)$, and a random algorithm $R(A)$ on the dataset D as input (the trusted party which works on the dataset and its release). An image of the random algorithm is computed, denoted as $I_{R(A)}$. By definition, we consider datasets D_1 and D_2, which differ in one row (a constant c), and any malicious adversary does not know c. We say that the algorithm $R(A)$ assures ϵ-DP iff $\forall$ subsets $S \in I_{R(A)}$, we have the condition as follows:

$$Pr[R(A)_{D_1} \in S] \leq e^{\epsilon c}.Pr[R(A)_{D_2} \in S] \tag{4}$$

   This means that the datasets are bounded by e^ϵ for c items, or more specifically, the datasets which differ by c, are protected by a noise factor of ϵc. The values are ϵ, and c are fine-tuned depending on the number of FL users and the aggregation operation to assure low-powered communication in IoMT ecosystems. In research, fine-tuning is achieved via minimax optimization and gradient-based convex optimization.

3. *Secure MPC*—In secure MPC, we assume that k clients $\{C_1, C_2, \ldots, C_k\}$ have their local gradients $\{G_1, G_2, \ldots, G_k\}$ to share with one another. Any k^{th} user C_k splits its

gradient into k partitions, keeps one partition secret to itself, and shares the remaining $(k-1)$ sub-gradients with other users. Similarly, all other users follow the same approach and keep one secret gradient to themselves. Each participant has the $(k-1)$ parts with them, which are shared by other $(k-1)$ users. The trick is that which part they have kept secret with them is not known to another user. Now an addictive function $A(.)$ is performed on the received $(k-1)$ gradient shared with them, and they add their secret gradient $S(G_k)$ to this to obtain the partial gradient computation at their end. In general, the overall sum and average computation at the aggregator node remains constant, which is shown as $\sum_{i=1}^{k} G_k$ and $\sum_{i=1}^{k} G_k/k$, respectively.

5.5. Communication Perspective in FL

In this subsection, we discuss the requirement of resource management in FL networks in terms of communication and computation costs. During the training process, the available computational resources are local or global. In the case of local training, the important factors to be considered are the dataset size $S(D)$, the node available energy E_N, and the processor computing power (in terms of CPU cycles/sec). The energy dissipation is proportional to the frequency of items in the dataset, denoted by f_d, $S(D)$, and the number of training iterations I_t, and the CPU processing parameters γ and τ. γ is the parameter for CPU cycles, and τ is the processing capacity parameter, which varies according to I_t [88]

$$D_e = I_t(\gamma \tau S(D) f_d{}^2) \tag{5}$$

Based on D_e, the model computational latency L_t is computed as follows:

$$L_t = I_t \frac{\tau S(D)}{f_d} \tag{6}$$

Equations (5) and (6) suggest that the computational time and energy have an inherent trade-off and depend on the hardware and system parameters and the convergence time of the training model. Thus, in the case of small-sized datasets, the computational requirements are low, which are further minimized in the case of a fog-edge layered scheme. The fog layer provides computational offloading services to FL clients, and the operating frequency is managed properly with the assistance of CPU computation. However, the local device parameters might be heterogeneous, affecting the model accuracy ω. To solve this, data portioning approaches play an important role in reducing the data items' statistical disparity. Once the local modeling constraints are satisfied, the gradients are communicated to the aggregator node. In the case of vanilla FL, the global updates are carried out at cloud/edge nodes, and the updates are sent back to FL clients, which corresponds to the CFL scheme. For distributed nodes, the model learning is achieved via low-powered consensus approaches, where the validator nodes form a meta-information block of the shared updates. It strongly depends on the data rate D_n of the underlying network, where we consider the transmission latency L_t of local parameter size p_s to be communicated to the edge server as follows [89]:

$$L_t = \sum_{n \in N} \frac{p_s}{D_n} \tag{7}$$

The required power P_n to communicate the updates is carried out as follows:

$$P_n = \sum_{n \in N} \frac{W_n}{L_t} \tag{8}$$

where W_n denotes the work carried out in total by n FL clients in L_t. Summing up the discussions from Equations (7) and (8), the model communication and computation costs depend on the number of iterations I_t and the minimization of the local loss function and through increased D_n. In such cases, the signal rate S_n also becomes an important

factor, and thus efficient channel encoding mechanisms are required. The models are fitted according to the designed local accuracy ω to update the global iterations G_t and global loss δ as follows:

$$G_t = \frac{\omega log(\frac{1}{\delta})}{1 - \omega} \tag{9}$$

Thus, there is a requirement for joint optimization of energy and latency in low-powered IoMT wireless networks.

6. Learning Models to Support FL Training in IoMT Ecosystems

FL is a framework for distributed ML and DL methods. FL allows different devices or clients to solve a common problem while keeping the privacy of the data at local devices in healthcare or IoMT infrastructure, where one global server takes control of the global aggregated model. This section mostly classifies FLs using DL and ML and their associated techniques. ML is an AI technique that learns without being explicitly programmed from data and experiences. For continuous data analysis and the production of useful information in the IoMT, ML might be crucial, particularly at nodes for computation such as cloud [90] and fog computing [91]. For FL, ensemble techniques, probabilistic fuzzy random forest, and DL mechanisms are found to be useful [92]. Table 7 presents related works of ML and DL to support FL-IoMT ecosystems.

Table 7. Contribution of ML and DL models to support FL-IoMT.

Author	Year	Model	Description	Contribution of Models to Support FL-IoMT	Dataset
Celestine et al. [93]	2020	ML and DL	ML and DL algorithm are used on healthcare dataset to identify which food need to be given to certain patients.	Data sets are collected from an internet source, and FL-IoMT models are used to enhance the accuracy of food selection.	health base medical dataset
Nighat et al. [94]	2020	ResNet-34 or DenseNet-121 models	DL model to detect Leukemia patients an IoMT-based model is proposed.	To detect four different subtypes, an IoMT-based model is used to enhance the accuracy up to 99.50% and 99.56% with GA SVM and ResNet 34, respectively.	ASH image bank and ALL-IDB global data-set
Muham-mad et al. [95]	2021	SVM, DT and K-nearest neighbor	A smart effective healthcare model that monitors the health of elderly people through IoMT devices.	The ioMT-based model interface provides higher validation and accuracy by 91.00% and 92.00%, respectively.	UCI Repository IoMT Dataset
Sekhar et al. [96]	2021	CNN, SVM, KNN	Implemented a TL-based approach to detect brain tumors from MRI images.	The model achieves accurate results for 3-class tumor classification.	Figshare dataset and Harvard Dataset
Atta-ur Rahman et al. [97]	2022	SVM and KNN.	Predict mitochondrial and multifactorial genetic inheritance disorders.	The SVM-based model gives better accuracy up to 87.00% than the KNN-based model.	Genome disorder dataset from Kaggle

Gathering information and integrating it into a model with sophisticated ML/DL models to construct an AI-based system is a critical and sensitive procedure concerning FL setups. Suppose an intruder alters information from origin to destination. In that case, the prediction will be erroneous, and the safety of the information will be jeopardized because this technique interacts with susceptible information. Thus, FL assures the privacy of data during the integration process.

Edge-based FL aggregation is advantageous as the quantity of knowledge obtained from nodes is fairly high when local procedures are used. Device heterogeneity can be maintained in a secure environment for data evaluation and training. An FL-associated AI framework can be incorporated to generate a secured and low-powered computation in healthcare setups. In the same direction, Table 8 provides a summary of the understanding of ML/DL support with FL-IoMT with a discussion on applicative areas, challenges, and potential solutions. Table 9 presents a summary of the various approaches of datasets that

combine ML/DL with XAI in FL-oriented setups, and Table 10 presents examples of the local learning models in IoMT applications.

Table 8. Applicative view of FL with ML/DL in IoMT setups with potential challenges and scope.

Author	Various Fields of Applications	Challenges and Its Addressing	Proposed Solution
Nguyen et al. [26]	IoT, edge computing, blockchain,FL	data volume	To improve the access and security of deploying FL, Blockchain is being used to provide decentralized learning via FL without the need for a central network connection.
Kim et al. [82]	Blockchain, FL, ML	The complexity of the architecture	Evaluated a latency prototype of FL based on blockchain to reflect the best block generation frequency while accounting for processing delays and communication.
Brisimi et al. [98]	learning mechanism for heart related disease	Issue related to the Sparse SVM	Proposed the design architecture for a cPDS that could distinguish between individuals who desired to be admitted to the hospital and individuals who did not.
Silva et al. [99]	Brain imaging data using the FL approach	There is not any production ready process based on FL.	A software central module and the base client to carry out the experiment learning experience.
Holbl et al. [100]	healthcare informative, distributed systems, consensus, Blockchain	Data confidentiality and Methods of encryption	Realizing the possibilities of blockchain systems and focusing on the challenges and contributions of blockchain-oriented research in the healthcare sector.
Long et al. [101]	Bio-informative, FL, healthcare	Data with various characteristics	Investigation of FL to allow for the development of an open health system using AI. FL's current challenges and prospective solutions are examined.
Esteva et al. [102]	AI, Healthcare, FL, NLP, Computer Vision	Ethical issue, data leakage, The NLP model is challenging to train.	A extensive examination of object recognition in biomedical image informatics, as well as an explanation of the usage of NLP in domains such as EHR datasets.

Table 9. Dataset summary of FL works with ML/DL and XAI in healthcare.

Authors	Year	Dataset Used and Its Description	XAI	FL	AI	Healthcare	Implementation Areas
Raza et al. [103]	2022	Beth Israel Hospital at the Massachusetts Institute of Technology maintains an arrhythmia database (MIT-BIH)	Y	Y	Y	Y	ECG-based prediction of arrhythmia using both noisy and clean data.
Shukla et al. [104]	2022	Datasets from 3DIRCAD	N	N	Y	Y	Detecting and Predicting Liver Cancer
Thomsen et al. [105]	2022	Statistics Denmark and the Danish Colorectal Cancer Screening Private Database	Y	N	Y	Y	Screening for colorectal cancer.
Flores et. al. [106]	2021	Image of a chest X-ray from Mass General Brigham	N	Y	Y	Y	To forecast COVID-19 cases based on chest X-ray analysis
Barbiero et al. [107]	2021	dataset of CUB	Y	N	Y	N	Method of logical explanation based on entropy

Table 10. IoT local learning models.

Local Learning Model	Authors	IoT Applications
SVM	[108]	Network slicing.
CNN	[109]	using the classification of images to diagnose disease.
LSTM	[110]	Used for the edge caching
K-means	[111]	To reduce packet error rates, sensor networks can be clustered. IoT networks can also use proactive caching.
RNN	[112]	forecasting using time series

ML and DL approaches simplify healthcare decision analytics, but understandability is still an open problem. The models act as 'black boxes'; thus, the output predictions are not validated. In healthcare, this is crucial to have confidence in the predictions, as inputs are vital health indicators of patients. Thus, explainable AI (XAI) has emerged as a groundbreaking approach to solve the explainability of these models. In XAI, rule-based and fuzzy mechanisms are researched worldwide. Rule-based systems allow model explainability depending on inputs and the processing steps. Fuzzy approaches produce forms in normal language to make the models easier for physicians and patients to understand. Authors in [113] discussed the CNN-supported model, which categorizes the wound. In the model, an XAI local interpretable model-agnostic explanation (LIME) replaces the model output with an interpretable form. Authors in [114] exploited the concept of XAI to diagnose glioblastoma based on textual characteristics, where the relationships between attributes are formed in a graph-based topology.

In FL communication, the underlying IoT protocols play a major impact on local learning. Based on the communication standards and application requirements, models are selected. For example, in the case of image-based classification, a fully-connected feed-forward NN (FNN) is preferred owing to low complexity. Different NNs necessitate the right application of activation functions, which control a NN response. Binary step functions, nonlinear activation, and linear functions are the three main categories for activation functions. These functions must be selected for local learning. Local methods should have fewer complications and a high degree of performance.

To determine the ideal number of local repetitions, authors in [115] took into consideration a variety of data patterns for study, namely, the data distribution, local datasets, and cognitive patterns. It is concluded that these factors influence the number of local learning iterations. In such cases, the concepts of CNN and SVM are used with edge-based aggregation. The architecture considers a network with nine layers and a convolution layer of $5 \times 5 \times 32$. A 2×2 MaxPool has been used to normalize the local response with $z \times 256$ fully connected system, with a Softmax activation function. Other approaches use squared SVM and double deep Q-network in the case of DFL schemes, which minimizes the transmission and energy cost of CFL networks [116]. Double deep Q-network is constructed through a single-layer fully-connected FNN with 200 neurons. The network integrates computational offloading and edges caching at the local sites. In [69], an FL-learning framework is presented with access to local repositories for authorized users. In work, convolutional and max-pooling layers are fixed at endpoints during the model transfer. These layers observe the miniature details of low-level user activity-related information. The performance is further optimized via TL. In addition, an incremental learning approach can be tailored to the architecture. The evaluation testbed was created using the two layers of the convolution and two pooling layers with the three fully connected layers. Here, a convolution size of 1×9 is used. Another case study is on an augmented reality-based application, which uses the CIFAR-10 dataset [117], and a CNN model is constructed. The analysis shows that, with a rise in local iterations, there is also a proportionate rise in effectiveness. A simple FedAvg scheme is carried out in the CNN network layer with a

structure size of $3 \times 6 \times 5$, with 2×2 max pool, and $6 \times 16 \times 5$ convolution layer. ReLU has been used as the activation function for the same. In [118], the proposed methodology revisits the performance issues of local iterations of FL and bias minimization. To address the same, an edge node collects the local models, and it is designed using an FNN model, which consists of two hidden layers of 30 and 50 neurons, respectively.

7. Trusted FL-IoMT

The section discusses trust as an important phenomenon in IoMT systems for a diverse set of activities such as data sharing, report monitoring, patient tracking, data collection and analysis, hygienic care, and preventive device maintenance. For the same, blockchain-based ledgers are considered that provide provenance and auditability in training systems [119].

7.1. Blockchain in IoMT

Blockchain is a distributed system that keeps track of transactions between networked nodes. The IoMT has expanded demand for distributed computing, and blockchain addresses many challenges regarding the security of healthcare system players. As IoMT devices generate data frequently, it is impossible to immediately store each sensing data in the blockchain, as it involves a high transaction cost [120]. To overcome these limitations, the environment of the IoMT device uses a lightweight, scalable blockchain that handles the frequent transactions from IoMT devices. Figure 6 represents the blockchain for the IoMT setup, where the cluster head (CH) is responsible for reducing the data flow from IoMT devices to the blockchain. CH also reduces the packet overhead by creating a pool of transaction information of IoMT devices after a certain epoch(δt). Permissioned blockchain stores the $< H_n, \Delta W_n >$ by executing the smart contract S_n. This crucial step ensures that the model updates are immutable and free from any modification by an adversary. Similarly, patient information (P_k) can be stored on the blockchain by executing the smart contract. As the size of P_k is generally large, a single transaction might span multiple blocks, increasing the storage overhead and proportionally affecting the transaction fees. To overcome this, we can use an interplanetary file system (IPFS) that stores P_k in local offline storage. Thus, a single unique 32-byte content address is returned irrespective of the content size. Thus, comparatively storing a single transaction in one block (on-chain), a drastic improvement of $\approx 32k$ entries is achieved with IPFS (off-chain) storage. In the FL-IoMT scenario, the blocks are linked in a network and use distributed time-based consensus algorithm for block verification after collecting data IoMT nodes and CH, and at δt time sends the block to mining nodes for verification.
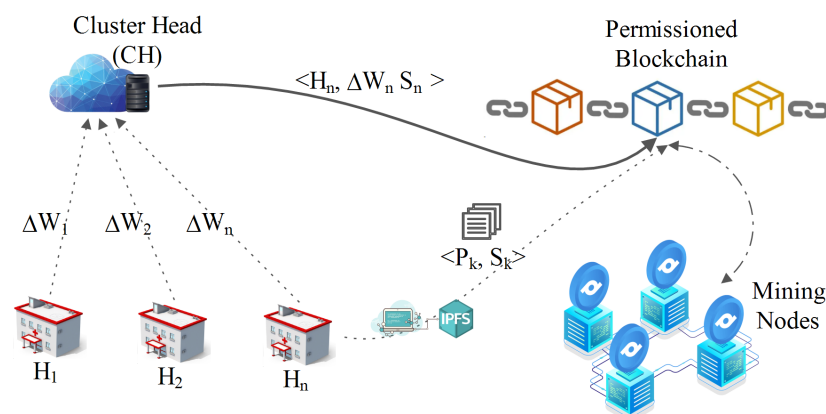

Figure 6. Block-Chain for IoMT.

The data exchanged between network nodes is saved and utilized for cross-references. This approach aids in pinpointing the specific source of miscreants in the network since these blocks carry information from preceding blocks. The unidentified blocks in the

network are therefore deleted, clearing the path for blockchain to be viewed as a trustworthy technique in FL–IoMT-based information exchange systems [121].

7.2. Trusted FL

The critical component of FL is aggregating the local gradients at the central server. It has been analyzed that model updates may contain some sensitive parameters that can be used to gain insights into users' details. Thus, user privacy may be jeopardized throughout the training, making FL vulnerable and discouraging medical sites from participating in cooperative learning. This aspect makes the development of a trustworthy server for the coordination of model aggregation a necessary condition for ensuring dependable FL operations in smart healthcare. According to the contract, entities such as global servers and local hospitals must guarantee a trusted model aggregation. This is especially important in the context of smart patient monitoring, where the processing of sensitive health information must be trusted to deliver trustworthy FL-based healthcare [26]. Researchers explored novel solutions, such as creating decentralized and blockchain-based trusted servers or offering safe aggregation techniques, to increase confidence in the server further. In this context, a novel approach is presented to prevent fake updates from unreliable local healthcare setups or devices. Such dependable device selection is crucial in reducing the impact of various security assaults.

As previously indicated, there are some difficulties in establishing trustworthy user selection in smart healthcare systems. There is no uniform selection methodology, for instance, and there are no real-time user monitoring techniques. Blockchain technology controls FL users' reputations to overcome these issues. In particular, the adoption of blockchain in federated environments decentralizes FL, enabling the removal of a single central server from the model aggregation process. The peer-to-peer block consensus function of the blockchain enables peer-to-peer coordination of the global model calculation. Blockchain technology is employed in decentralized FL-enabled healthcare systems to stop rogue servers and outside intrusions [122].

7.3. Incentive FL

All IoMT devices must communicate their local model updates with the aggregation server following conventional FL techniques, although this server is not always accessible in real-world scenarios [123]. IoMT devices typically have constrained computer power, radio bandwidth, confidentiality issues for user data, and server dependability, so they are unwilling to share their algorithms. Incentive-aware FL solutions motivate more FL users and enhance FL-enabled occupational health and safety effectiveness. A recent study discovered that FL incentive systems could be grouped by various criteria, along with the device's information. FL can be categorized according to various factors, such as based on the information contribution of the device, resource usage, and device reputation.

- *Device information contribution*: Information quantity and information quality are two crucial variables used to analyze it. Information quantity refers to the frequency of the local model updates and how often we train with sample data; it is commonly measured using the Shapely value.
- *Device reputation*: When creating FL incentive systems, device reputation is a crucial measure to consider. Reputation generally represents how accurate information is supplied for model training and trusted local updates.

7.4. Resource-Aware FL

Instead of aggregating the weights of each local model, a resource-aware FL is used to aggregate an ensemble of local information collected from edge models, which is subsequently distilled into a resilient global knowledge as the server model using learning extraction. Deep knowledge exchange combines the local model and the global information into a small-scale knowledge network. The deployment of a resource-aware model and multi-model knowledge fusion is made possible by such knowledge extraction, which

preserves communication effectiveness and modeling diversity. If the mutual training of the local model trains with the global information, the results present a small-size network to extract the local knowledge [124]. The tiny-size network is then transmitted to the service for multi-model synthesis and global knowledge distillation. As a result, better communication efficiency and multi-model FL for the resource-aware FL are produced.

7.5. Gossip-Based FL

A decentralized alternative to FL is called gossip learning and does not call for a central server or any other component. It relies on a simpler infrastructure—message transfer and no cloud resources. Gossip learning is likely to be strictly less efficient than FL. However, decentralized FL is a viable answer to the issue of unreliable parameter servers in centralized FL for secure federated competent healthcare. Decentralized FL implementations frequently use consensus, diffusion, and gossip methods. For instance, it is suggested that a decentralized FL scheme can be used in conjunction with a segmented gossip aggregation technique to improve training effectiveness. To maximize the use of all clients' available bandwidth, each client can function as a worker who randomly chooses a small number of nearby workers to send the model segment during each training iteration [125,126].

8. FL-IoMT for Modern Healthcare and Its Applications

In a conventional IoMT scenario, for various reasons, AI-based systems collect data from medical records, such as sickness diagnoses, medical images, clinical trials, medication discovery, and EHR, among others. In this circumstance, exchanging EHR with remote data centers or the cloud to generate medical data raises significant privacy concerns. This suggests that eliminating or missing metadata such as patient information may not be adequately protected in terms of privacy, notably in difficult healthcare situations. Because data must be sent to analyze, most AI systems depend on a central server. FL combines more data with increased privacy awareness to suggest alternate options. Table 11 presents a comparative analysis of emerging FL-assisted IoMT schemes regarding trust, aggregation, privacy, and computation costs.

8.1. Facilitates Group Learning

To address the problem of a small sample size for training secure collaborating information and shared ML model, the FL method gathers data from a variety of clients [127]. It is essential to choose a data partition in the HFL and DFL to handle the limited sample size and limited sample characteristics. This technique offers decentralized ML model training without using an organized central aggregate server. Medical institutions train their DL models locally before periodically sending them to the global server. The aggregate model is distributed to all local facilities.

During the training process, the data are always kept private to each node. Medical information is kept private by only transmitting the model's weight and characteristics. Because it keeps sensitive and private information while allowing many medical organizations to collaborate, FL eradicates numerous security worries. Hence, the FL approach collects data from several clients to address the issue of limited-size data for training ML-based shared models. To address the sample data's limited size and features, it is important to select a good data partition scheme in HFL and DFL. This method provides decentralized ML model training through a central global server.

Table 11. A comparative analysis of FL-assisted IoMT schemes.

Author	Year	Objectives	1	2	3	4	5	6	Contribution	Limitations
Aouedi et al. [36]	2022	A transfer learning approach in FL communication is considered to simplify and save time in the data collection process	Y	Y	Y	Y	Y	Y	With FL, blockchain is integrated and encryption methods are used during aggregation to preserve the privacy of transfer learning	There are still unresolved privacy and security concerns with client synchronization and the use of non-IID (independent and identically distributed) datasets.
Ali et al. [37]	2022	A triage of DRL, DT, and GANs with FL setups is established to detect privacy risks.	Y	N	Y	Y	Y	Y	The efficacy of FL algorithms utilizing DRL, DNN, and GANs in FL architectures was demonstrated through experimental simulations.	The elimination of bias elements in data preparation for ML is a critical issue. The challenge becomes more complex in FL as data are stored in a scattered manner.
Sandi et al. [40]	2022	To deploy a smart contract ecosystem to implement the privacy-preserving bidirectional long-short-term memory and enhance security for the FL-based IoMT.	N	Y	Y	Y	Y	Y	The efficient privacy-preserving framework for secure misbehavior	It is essential to check the suggested strategy in a real setting using actual EHR.
Alamleh et al. [41]	2022	Standardization of multiple-criteria decision-making (MCDM) is considered for evaluation and benchmarking of FL-based intrusion detection system datasets	N	Y	Y	Y	Y	Y	Via FL, the intrusion detection datasets are distributed at remote locations, which improved the model accuracy.	IDS classifiers were only used for the binary classification problem
Zhao et al. [43]	2022	Federated Reinforce Client Contribution Evaluation (F-RCCE) is designed as an evaluation method for the client.	N	Y	Y	Y	Y	Y	Privacy preservation and accuracy are gained using differential privacy and deep neural networks (DNN)	More robust and lightweight encryption algorithm is needed.
Ahmed et al. [44]	2022	In FL-based IoMT, the idea of physical layer security (PLS) is introduced and is utilized to effectively preserve data privacy.	Y	Y	Y	Y	Y	Y	The improvement of the secrecy rate of PLS in FL-based IoMT is reported, and it is based on clustering.	For the development of an energy-efficient and secure FL in IoMT, PLS and cryptographic mechanisms are required to be merged and tested for efficiency.
Xu et al. [45]	2022	An MSQE quantizer mechanism is integrated with FL to address the quantization issues of IoMT sensors and wearables.	N	N	Y	Y	Y	Y	With quantization, security concerns are also addressed to assure privacy and improved accuracy of the training model.	The difficulty lies in deciding how many parameters to use for quantization, which involves bncing performance gain and overhead.
Ruby et al. [49]	2022	The work addressed jamming-based attacks in FL communication networks by a malicious adversary.	N	Y	Y	Y	Y	Y	The interaction between the FL network and a jammer is modeled using a hierarchical game-based technique to identify the optimal strategy for each player.	The suggested anti-jamming technique has excellent energy and learning performance when a single jammer is present, but it is insufficiently protective when there are several jammers present.
Wang et al. [54]	2022	To reduce serious privacy and security issues while transferring the patient's data to the central server.	N	Y	Y	Y	Y	Y	explains the FL training framework and creates a simple protocol for privacy protection based on secret sharing and weight masks.	Techniques such as a single point of failure and model poisoning attacks are not implemented.
Fan et al. [126]	2022	A COVID-19 ecosystem is considered, where authors used FL to solve privacy and security concerns of locally trained nodes	N	Y	Y	N	Y	Y	Diverse and heterogeneous data sets are considered to assure personalization in model learning, and the data are aggregated for global update through the privacy-preserving mechanism.	The transmission of model updates to a global server might be attacked by an adversarial model, and thus central privacy protection needs to be explored.
Gupta et al. [42]	2021	Security and privacy risks in the anomaly detection (AD) models are considered while sharing the patient's confidential data	Y	Y	Y	Y	Y	Y	A lightweight IoMT ecosystem is devised, where devices are identified based on anomalous classes.	FL communication metrics such as communication time and local training time are not considered during evaluation
Wu et al.[55]	2021	An incentive method for rewarding data owners to participate in FL in the IoMT is considered with privacy risk into account.	Y	Y	Y	Y	Y	Y	Multi-dimensional cost optimization using incentive mechanism for FL privacy.	Numerical methods are needed for performance evaluation.
Lakhan et al. [56]	2021	A multi-level local fog-cloud ecosystem is designed that guarantees privacy preservation and fraud detection.	Y	Y	Y	Y	Y	Y	Energy consideration is minimized and communication delay for healthcare devices is considered for IoMT applications.	Dynamic and run-time unknown threats in edge-fog design are not considered.
Samuel et al. [57]	2021	Addressed the issue of ineffective model detection.	Y	Y	Y	Y	Y	Y	A COVID-19 infrastructure design is presented to overcome the communication constraints.	A flow model of the implementation is not discussed. Furthermore, there is a concern about the scalability of the real-time design.
Choudhury et al. [58]	2020	To apply an FL strategy to decrease the problems caused by centralized learning.	N	Y	Y	Y	N	Y	A comparative analysis between accuracy and privacy.	FL algorithm has convergence issues and the same is not addressed.

1: Threat Model, 2: FL Client Nodes, 3: Aggregator Server Type, 4: Computation cost, 5: Communication cost, 6: Privacy, Y—shows that the parameter is present, N—shows that the parameter is absent.

8.2. EHR Data Analysis

In EHR analysis, obtaining reliable results across populations is challenging since healthcare data are frequently fragmented and private. This poses a challenge to creating generalizable, effective analytical approaches that require a variety of "big data" based learning. FL holds significant promise for collaborating with different healthcare data sources while ensuring patients' privacy. FL uses a centralized approach where a global server builds a model by receiving the inputs from different local healthcare setups while keeping the sensitive data in local establishments where it belongs.

As an example, patient pathology data are often acquired digitally and in the form of an EHR; it is beneficial for medical inspections, sickness diagnosis, and locating essential information. CFL and DFL are two types of FL for managing the data centrally and in a distributed manner. EHRs include systematic and stochastic biases that restrict the generality of the conclusions, even though they give a vast quantity of EHR for research. FL

allows medical organizations to integrate EHR data practically, allowing people to maintain their privacy. Because of the iterative perceived advantages of gaining knowledge from big and varied volumes of medical data, the FL model will perform well.

8.3. Healthcare Monitoring

As a result of COVID-19, physical and mental health problems are becoming more prevalent among people. Due to this, maintaining our bodily and mental health requires personal daily healthcare management and monitoring. The IoMT has developed due to the convergence of the IoT with healthcare services to offer intelligent medical services. Its widespread use has been hampered by privacy and security issues. Healthcare systems using the IoT have been continuously improving, which handles the growing need to carry patient data. This expansion has made it possible to use various health gadgets, such as smart technology-based sensors, which may track and analyze various personal health characteristics and, in some cases, act as a trigger for potential health incidents. Most linked healthcare systems currently use ML and DL algorithms to generate decisions automatically for accurate predictions. While this has led to more accurate disease diagnosis and quicker disease detection, some drawbacks, such as a lack of labelled data for training models, make systems unreliable and inefficient.

Additionally, improper and malicious coordination of ML models could result in possible attacks and information leaking. In some circumstances, this could result in the privacy of patient data being violated, putting data security at risk in similar circumstances and sowing distrust amongst various parties. Therefore, it is necessary to incorporate experts' domain knowledge to provide a heuristic-based knowledge system that maintains confidentiality to improve model accuracy. To protect data privacy, a model is trained locally using patient data. In this situation, local adaptation or contextualization of the ML model is still possible, which is more efficient than a model that trains on fewer data. Additionally, utilizing their data, all nodes collaborate to train the model. All ML models distribute their knowledge among all connected nodes. This platform has a powerful anonymity feature and can repel adversarial attacks throughout incremental learning. The data leakage problem is solved when local data are kept at the site and shared with the trained model parameter. Hence, an active FL-based health monitoring platform is beneficial. This platform has a powerful anonymity feature and can repel adversarial attacks throughout incremental learning.

8.4. Imaging in the Medical Field and COVID-19

Another way FL might be employed in an IoMT environment is in medical imaging. In collaboration with some medical institutes, it is utilized for activities such as brain tumor segmentation and diagnosis using Magnetic Resonance Imaging (MRI), CT scans, and X-ray images of the chest. To detect COVID-19, AI and computer vision help [128] to determine infection level. The hospital does not allow sharing of medical data without authorization due to patient privacy concerns and protection. Such training data took a lot of work to gather. This will result in insufficient data samples while using DL techniques to find the COVID-19 infection spread. Such types of problems or challenges can be solved using the FL mechanism. Without obtaining local data, it can resolve the problem of data silos and produce a common model.

To collaboratively create a shared model, FL relies on data sets spread over some local healthcare setups across the globe. This perfectly safeguards patient information. FL experiments for COVID-19 and medical imaging are required, which is still spread across the globe in different states of the counties. In [129], the FL concepts have experimented on COVID-19 identification with the help of chest X-ray images. Four alternative models—MobileNet [130], ResNet18 [131], MobileNet, and COVID-Net—were used in the experiment to train CXR pictures [132]. The comparative experiment of training with and without FL was identified. The experimental findings demonstrate that ResNet18 performs optimally in both FL- and non-FL-training. In images containing COVID-19

labels, ResNeXt [133] performs best. The most minimal set of settings is in MobileNet. Research shows that ResNet18 is a better option among the four well-known COVID-19 detection models.

With the assistance of numerous hospitals and a data center, FL can be useful to develop cooperative education procedures for EHRs analytic management systems that are resource-conscious and respect privacy [134]. Despite the rigorous privacy regulations imposed by distant medical centers [135], the use of FL creates new possibilities for federated EHR analytics because it only permits the transmission of model parameters; raw data are preserved at local locations due to the FL learning nature [58,136]. Due to distributed data resources being used and the computing power of several silos, FL is an extremely effective learning method to boost the success rates of training AI models, as demonstrated in [137,138]. By preparing a worldwide model, FL is useful for easing at-home health monitoring from scattered houses under a data server's supervision by keeping user data locally while avoiding data leaks [139]. For instance, FL can help healthcare applications such as assisted living and fall detection by enabling mobile activity monitoring [69,140]. It is also found that the FL method outperforms other traditional methods, such as centralized education and regional training without federation, apps for smart health monitoring that have excellent accuracy rates, and respect user privacy [141,142].

It is recognized that FL, by combining several medical institutions into a single, unified entity through the federated data training process, significantly supports medical imaging applications [143]. In particular, FL can be used to connect medical devices such as MRI scanners to cloud servers because these devices have sufficient local datasets and adequate CPU power to compute AI updates [144,145]. The potential for aiding the identification of acute neurological signs such as headache or unconsciousness is greatly increased by a recent study of the functions of FL, particularly in X-rays [146,147]. FL is very helpful in promoting COVID-19 diagnosis and detection with escalating privacy issues, with coordination of enormous hospitals to construct a shared AI model [148–150]. For instance, it was discovered that FL could be used by several medical institutions in collaboration for COVID-19 screening from X-rays [151], while working with hospitals to detect COVID-19 and categorization and feature extraction of X-ray is carried out [152,153].

9. Open Issues and Future Research Directions

The section discusses the open issues and potential research directions of FL-IoMT and its key integration in terms of security and networking concepts. The details are provided below.

- *Size of Blockchain*: The majority of public blockchains are significantly smaller than personal healthcare data. Storing EHRs on the blockchain is a significant challenge due to high computational mining overheads. The blockchain size might increase significantly, and thus handling large volumes of data is a prime challenge. To cope with the storage issue, the information should be maintained in a separate off-chain storage-based system. Rather than the whole data management, the blockchain comprises the hash-based references; this is a case of the typical database system with the clinical data held off-chain and off-chain clinical patient information may now be accessed through immutable hashes of healthcare data kept on-chain [154].
- *Communication in FL-IoMT*: Communication is crucial to the success of FL-enabled healthcare services between the aggregate server and FL users [155]. Effective communication resource allocation strategies can greatly boost learning outcomes. When numerous devices using IoMT must establish a connection with the aggregation server for downlink model broadcasting and uplink model updating, it carries huge significance [156]. According to numerous current research studies, in such a scenario, the aggregation server can use efficient scheduling rules to choose an appropriate group of IoMT devices [157].
- *The dynamicity of the wireless channels*: Owing to variable delays in different connected links and the dynamicity of the connected topology in wireless channels, a propor-

tional dependence in the present in the training of the local models, and communication of the learning updates from aggregation server to FL clients, and vice versa [158]. Possible solutions include optimizing objectives where the link dynamicity and temporal behavior are considered in the approach. Secondly, trustworthy design objectives such as outage probability and device reliability should be considered.

- *Blockchain framework for FL-based IoMT*: Although the literature has identified several positive effects of FL-enabled healthcare services, for the identified issue, there is no accepted terminology for managing the usefulness of the various explanations or solutions. For the various blockchain frameworks suggested in the literature, it indicates that the requirement of the central server with trustworthiness is to be achieved with the limited changes of the prescribed IoMT policies. They are available for various (healthcare) scenarios, and various network configurations and data sets are used to assess their effectiveness; it is challenging to compare such systems [159].

- *Standards associated with the protocols*: The ubiquity and standardization of communication protocols, device technology, deployment situations, and aggregation techniques provide significant obstacles. Recently, IEEE Std 3652.1-2020 [160] provided instructions for FL architecture and design aspects. Additionally, this guideline highlights important FL-related issues such as privacy, security, performance effectiveness, and economic feasibility, as well as assessment techniques and success factors for FL platforms.

- *AI Techniques and FL-based IoMT*: Recently, concerns have been raised about slow learning rates and high bias in the FL learning process. Moreover, with a high number of local nodes (silos), the communication overhead with the central server also increases. Thus, it is highly imperative to design flexible and on-the-fly training hyperparameters so that learning rates can be optimized [161]. For example, the design of efficient incentive systems in DRL is used for optimizing network parameters in FL communication, with lower dimensionality [43,162]. Moreover, to implement trusted FL, effective incentive design mechanisms should be present, which synergize with the model cost [163,164]. Another approach is to design adaptable FL techniques where different ML models are introduced on a plug-and-play basis. The underlying design remains intact (at the local silos and the central server) [165].

- *Security in FL based IoMT*: Distinct customers might have various datasets, including time series, text, images, and audio, as well as many data elements, including skin temperature, heart rate, face photos, and blood type, in realistic healthcare settings[43]. Most FL architectures generally employ single datasets, for example, diabetic retinopathy [58]. To train these datasets, heterogeneous FL architecture must be designed, where ensemble learning mechanisms are included [166].

- *6G and its impact on FL-IoMT*: The rise of 6G communication networks has opened interesting applicative use cases in the IoMT domain. We are transitioning towards micro-level chip implants, which require low networking power to communicate [167]. New technologies such as compressive sensing, blockchain, THz and visible light communications, quantum transmission, 3D networking, and huge intelligence surfaces coincide with the considerably tighter 6G criteria [168]. Future research should focus on employing sixth-generation (6G) service on smart devices, including smart wearables and implants, for FL-based healthcare on a large scale. AI and FL capabilities, for example, should enhance future e-health services, improving patient quality of life and reducing hospitalization rates [43,169].

- *Data privacy and FL-IoMT*: Despite FL's significant potential to safeguard user data privacy, many privacy concerns need to be appropriately handled, particularly in situations related to smart healthcare given the high sensitivities of the wellness data [58]. Attacks using association extrapolation, unintended disclosures, and generative adversarial networks are three categories under which FL privacy problems might be grouped [170]. For instance, the attacker might improperly use the global FL model to determine if a sample of data is included in the FL health data collection. Furthermore,

the moment the patient's apparatus transmits local model updates to the main server located in healthcare centers and hospitals, the patient's information can be deduced. Building privacy-preserving FL healthcare solutions using differential privacy, AI, and cutting-edge encryption methodology seems challenging. Numerous studies consider using various other types of privacy to improve the privacy of FL systems [171]. Further study of this area of research is necessary for smart healthcare systems, including how artificial noise is introduced to model upgrades from individual gadgets. In such a way, the model based on the centralized computer will also be affected [172].

- *Attacks and its defense mechanism*: Several client-side participants in the FL-based oriented healthcare approach could pretend to be attackers and try to provide false or poisoned model updates to damage the model aggregation. When local clients and the central server are communicating the model during local data training or model transmission, an adversary may also contaminate information about data features. Attacks on the server side may be used by an outside attacker that potentially steals information from the combined global model, posing serious privacy concerns such as information leakage. A fundamental obstacle for FL-based smart healthcare systems is finding a solution to these security issues. Consider other methods, such as differential privacy, to protect training datasets against leaks. Additionally, creating aggregation with the safety techniques is a viable way to offer a framework with double masking for local updates encrypting, facilitating a key exchange between clients and the main server, and protecting clients from data theft and attack [173].

- *Non-iidness of healthcare dataset*: A critical issue that has to be resolved is the non-iidness of the database related to the medical, which might lead the FL training to diverge in training to obtain an outstanding training result in FL-based healthcare services. A hospital might, for instance, possess a greater prevalence of a certain local ailment compared to hospitals in other cities. The label distributions, in this instance, vary between medical institutions, making it difficult for them to participate in the federated data training. Without mentioning the non-iidness challenge, the data instruction would substantially decline in quality. To assure effective information support in FL-based intelligent healthcare, remedies to the non-iid issue must be developed, such as creating a new subset of databases to distribute equitably among clients. Solutions to the non-iid problem must be developed, such as creating a supplementary subset of databases to distribute evenly among the end-users, ensuring efficient data training for intelligent healthcare. Within the smart healthcare industry powered by FL, non-iid data must be evaluated using quantitative standards such as standard deviation, accuracy with precision, and correctness regarding the label/feature distribution skew division and the homogeneous mechanism [174].

10. Case Study: *Cross-FL-*Trusted Cross-Cluster Federated Learning in IoMT

This section outlines a case study by Jin et al. [175] and addresses the issue of sparse clustering in heterogeneous IoMT ecosystems. The authors integrated blockchain with FL (BFL) to address the trust issues among remote hospital nodes. They proposed the generation of large BFL clusters, which internally can be subdivided into smaller clusters. To address the interoperability of communication between intra-clusters, a cross-cluster FL (CCFL) approach is presented, and consensus protocols are discussed to send immutable updates through the FL chain. To obtain a consensus, the blockchain ecosystems implemented in the cluster require regular network connectivity and communications. The consensus effectiveness and efficiency might be quite low due to the high-latency interactions and regular communications. Furthermore, BFL demands that model changes be distributed across the FL-based cluster. Figure 7 depicts cross-cluster gradient aggregation across two hospitals. The gradients which are produced in hospital A would be transferred to hospital B. Then, the gradients are aggregated into the final step. The local model updates are sent out by the devices/nodes installed in hospital A. After FL training, the FL consensus FL is formed on the blockchain ledger based on the intrachain mechanism, where the gradients

are recorded from the updated models within the premises of hospital A. These gradients, which are identified from the various devices, are then consolidated in the process of the consensus. The resultant updates of the aggregated gradients are then transferred to hospital B, and the step is termed the gradient exchange. Once this is received by hospital B, the validations will be performed. The consensus is accepted if the intrachain consensus is confirmed by hospital B. When the confirmation receipt is received, hospital A is acknowledged the same. The process of validation from different devices supported by the intrachain consensus is called receipt consensus. The resultant (receipt consensus) is added to the transactional ledger, which is stored at hospital A. To communicate in the FL, every IoMT device in a hospital forms a cluster, and each cluster has a permissioned blockchain system installed onto it. The major outcomes and the contributions of the case study are outlined as follows:

1. The issues with the current BFL were identified, particularly the problem of data sparsity, privacy leakage, and low efficiency were discussed;
2. Cross-cluster federated learning schemes were implemented to provide secure communication across the cluster;
3. To show the viability and effectiveness of CFL, prototypes are used, and comprehensive trials are run.

As indicated in Figure 7, the transaction updates are based on the gradient computation, the confirmation of local update status, the application response, and the miner incentive in the CFL protocol. The blockchain follows the Merkle tree data structure, where apart from the gradient computations at local nodes, the aggregation strategy (for both CFL and DFL) is based on Merkle addition, and the transactions are proposed on a block packager. The packaged data (in the block) are added with a time-generated nonce and sent for validation.

The blockchain system uses the consensus mechanism to enable distributed training task orchestration and model update aggregation [176]. The BFL setup runs in local hospitals, where clusters are created and are called healthcare centers. Interchangeably, the term node means any general IoMT device, and the cluster represents healthcare centers. We consider that the training data d is distributed among D nodes in any cluster, where any node n_k works on k samples, denoted by s_k, with the trivial constraint $1 \leq k \leq D$. The objective of the training node n_k is to minimize the function $f_k(w)$:

$$f_k(w) = 1/n_k \sum_{p=1}^{n_k} l(x_{k_l}, y_{k_l}, w) \tag{10}$$

In Equation (10), the (x_{k_l}, y_{k_l}) denotes the sample data with an indexing l which is in node n_k, and w denotes the weights. The $l(x_{k_l}, y_{k_l}, w)$ is the loss function used to make the sample prediction. Once model learning is complete, the gradient descent algorithm is executed to minimize the local model loss as presented in Equation (11) as follows:

$$w_{t+1} \leftarrow w_t - \gamma \bigtriangledown f_k(w_t) \tag{11}$$

Here, γ denotes the unit step function, and t denotes the model learning rate as inputs to the gradient descent algorithm. We consider that $\bigtriangledown f_k(w)$ computes gradients of small unit-step functions, represented as $f_k(w)$. For learning, the BFL gathers all updates from nodes (IoMT devices) and sends them to healthcare centers (clusters), which operate as aggregators. The training objective $O(w)$ is presented in Equation (12) as follows:

$$O_w = \sum_{d=1}^{D} (n_k/n) f_d(w) \tag{12}$$

where n represents the total number of samples in the training data, n_k is the k_{th} value of the aggregation, and O_w is the loss functionality. To identify the resultant value of the loss function at unit steps, denoted as $S(O_w)$, Equation (13) is presented:

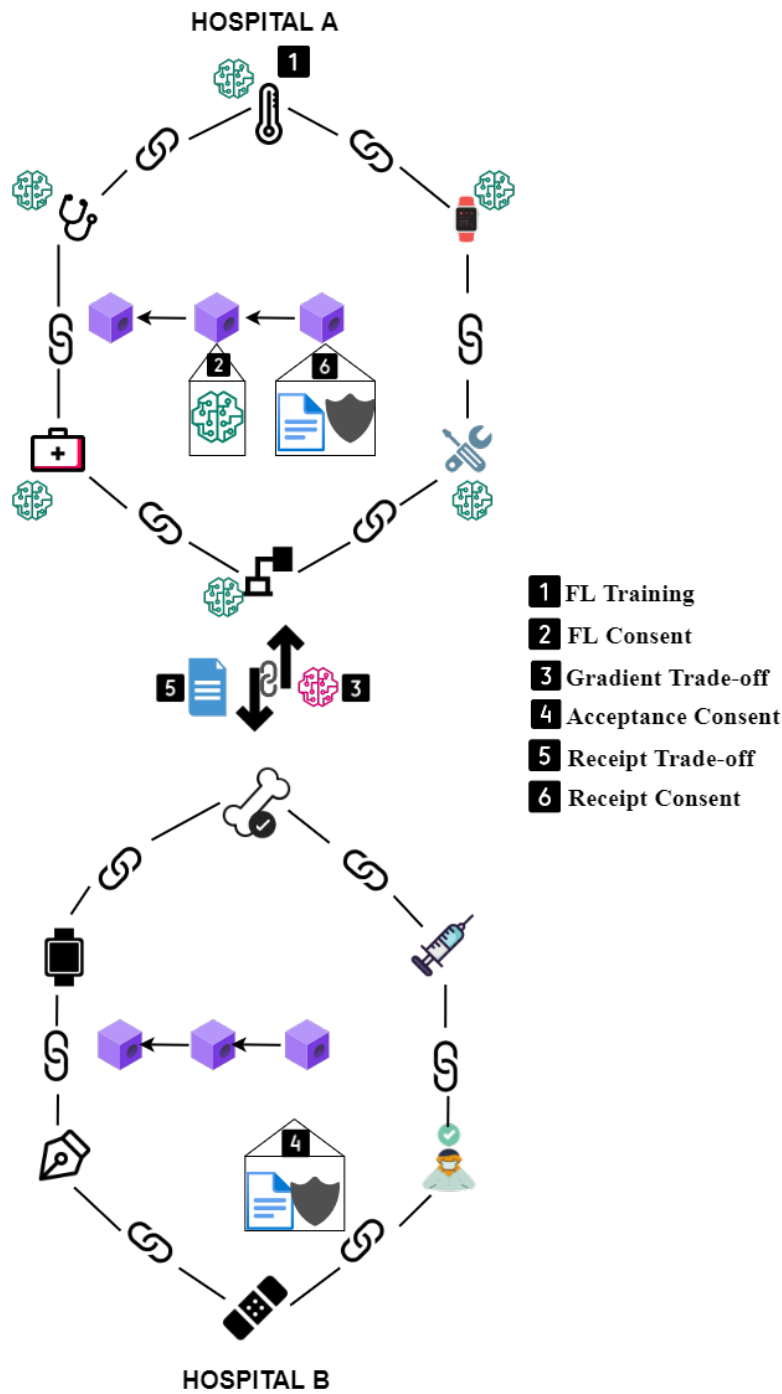$$w_{t+1} \leftarrow w_t - \gamma \sum_{d=1}^{D} (n_k/n) \triangledown f_k(w_t) \tag{13}$$

Figure 7. Gradient aggregation for the cross-cluster architecture [175].

The proposed approach adds the updated models of the BFL cluster to the CFL. In total, if we consider F BFL healthcare setups, where each cluster contains T_i training data with m_i samples, with the condition $1 \leq j \leq F$. We build a CFL, where the model requires

the connection of M BFL clusters and aggregates each cluster's information with each update. Equations (14) and (15) show the condition that, whenever gradients are updated, we present two parameters and k and k_i, which indicates the c_i. Here, k_i denotes the total number of devices in c_i, and k is the total number of nodes. If we consider n as the total number of samples in all clusters, then $n = \sum_{j=1}^{F} n_i$, and $k = \sum_{j=1}^{F} D_j$ is considered:

$$w_{t+1} \leftarrow w_t - \gamma \sum_{j=1}^{D} (n_j/n) \tag{14}$$

$$w_t - \gamma \sum_{d=1}^{D} (n_d/n) \bigtriangledown f_d(w_t) \tag{15}$$

Here, the D_i and D denote the number of nodes in the cluster and the number of nodes in the cluster. n represents the number of cluster samples. Now, this can be validated that CFL generates a comparable model to the BFL if the same sample data sets are the foundation for both. The blockchain setup is permission-based, and proper authorization is required to access the local gradients of another system.

Working of the Cross-Cluster Gradients Aggregation (CCGA): We compute the produced gradients for two hospitals, say A and B. At the first hospital A, at the local level, every device performs local training at the device level, and then it disseminates the updated gradients to hospital B. Before the dissemination, the model updates are stored via an intra-consensus mechanism to ensure trust in the shared updates. In this manner, the updated gradients are aggregated from various on-devices and sent for further processing. The gradient exchange (AB) refers to the data shared from A to B, denoted as $A \rightarrow B$, and vice versa.

As soon as hospital B receives the updates from hospital A, it runs a verification process where it matches the sent data with the stored data on the blockchain ledger. Normally, due to the restricted memory of the blockchain, the data are stored off-chain, and the content key hash is only stored on the ledger record. Hospital B uses the content key of off-chain and the public key of A to decrypt the stored off-chain data. This dual verification process is called consensus undertaking, and once the details are verified, it is logged as a received transaction. An acknowledgement from hospital B is issued to A as a receipt exchange, and the different local gadgets of A and B are synchronized to operate in unison. The synchronization process is timestamped and recorded as a transactional entry in the blockchain.

Hasty Consensus (HstCon): The intrachain consensus and interchain consensus are two sub-protocols that make up the CFL consensus process. The latter is performed using a two-phase cross-chain consensus (2PCC) process. The classic single-chain blockchain consensus may be used to implement the former as practical Byzantine fault tolerance (PBFT). The two-phase commit (2PC) protocol used in the dataset of industry served as the model for the 2PCC protocol, which provides secure data sharing between two clusters. Regarding the union of 2PCC and traditional single-chain consensus, as soon as they obtain the model changes, nodes in this protocol operate quickly. The workflow of the algorithm's protocol is shown in Figure 8. In the figure, the protocol considers two phase of operation, namely, the prepare and merging and discarding phases. In the merging phase, a single chain consensus is preferred, and the updates made to the local models are validated. The updates sent to the validators in the merging phase are usually received from various other clusters. While receiving the updates, the local data would be sent to the merge/discard remote cluster decisions. The received updates and local updates are combined based on the criteria that, for subsequent rounds of learning, the local results and the results obtained from the remote model (normally a global server) should match. It is then considered an acceptable state in the process. If this does not match, then the local updates would be discarded by the cluster, which is considered as the discarding

phase. Two binary protocols, named the interchain and intrachain consensus, are the part of the CFL consensus mechanism. These protocols operate on the practical Byzantine fault tolerance and the 2PCC protocol in the cross-cluster gradient mechanism. The latter secures information communication between any two clusters. As mentioned before, the aggregated updates are collected from the various nodes available in the cluster, and these are referred to as a consensus. Due to the extensive presence of nodes, this takes a long time for the operational consensus in the blockchain. This results in the lower efficiency of the HstCon.

To address this, another proposed consensus, termed DefCon, provides a representation for every cluster and creates a proportional reward/punishment system. A closer examination of the 2PCC mechanism reveals that it comprises two phases, similar to HstCon: (1) preparation and (2) merging/discarding. The model updates that each cluster obtained from another cluster are verified using single-chain consensus during the preparation phase.
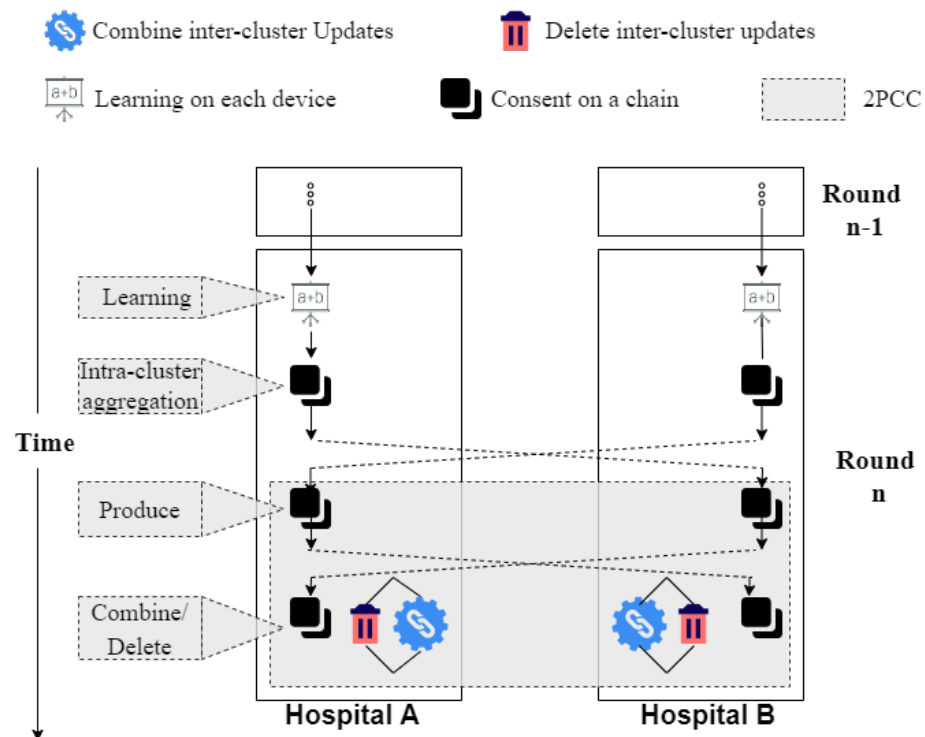
Figure 8. The HstCon algorithm's process flow [175].

Deferred consensus (DefCon): Although CFL and *HstCon* appear to operate, the system efficiency issue poses a significant obstacle. In DefCon, every cluster's representative is introduced and creates an appropriate positive reward and punishment system. DefCon often selects a similar characteristics-based representative to reduce the detrimental effects of frequent consensuses. DefCon frequently chooses a representative in each cluster to orchestrate and coordinate inter-cluster and intra-cluster learning. The configuration for the same is shown in Figure 9. Thus, DefCon improves the system's competence. The workflow is classified as successive sequences and is related to the representative of the corresponding tenure. A single cycle contains modified 2PCC* and the cross-cluster tentative learning task with the k rounds. Each supervisor conducts two confirmations, to be exact. Just before the cycle begins, one evaluates the model, and the other does the same at its end. The peer will agree to approve all previous actions if the latter validating findings are greater than the earlier ones. The peer will choose to disapprove of it if it does not. Additionally, the new representative's election is held during the preparation step. To stop the representative from

doing anything bad when establishing the reward/penalty system for the representative, DefCon asks them to mortgage certain assets.

Figure 9 systematically presents the DefCon procedure with the representative to increase system effectiveness. The process is broken down into several periodic rotations according to each representative's period in the service. A cycle in Defcon comprises *k* rounds of the modified 2*PCC* technique and tentative cross-cluster learning tasks. In this cycle, the latter is used to validate tentative learning assignments' outcomes. The intrachain and interchain consensus is not involved in a round in DefCon, unlike a cross-cluster FL round in HstCon. Instead, the representative's job is to compile the updates from the cluster's various nodes. Additionally, it chooses whether to include the updates from the far-off cluster. No ineffective consensus may be reached during a round. Therefore, it is anticipated that CFL with DefCon will operate more efficiently.
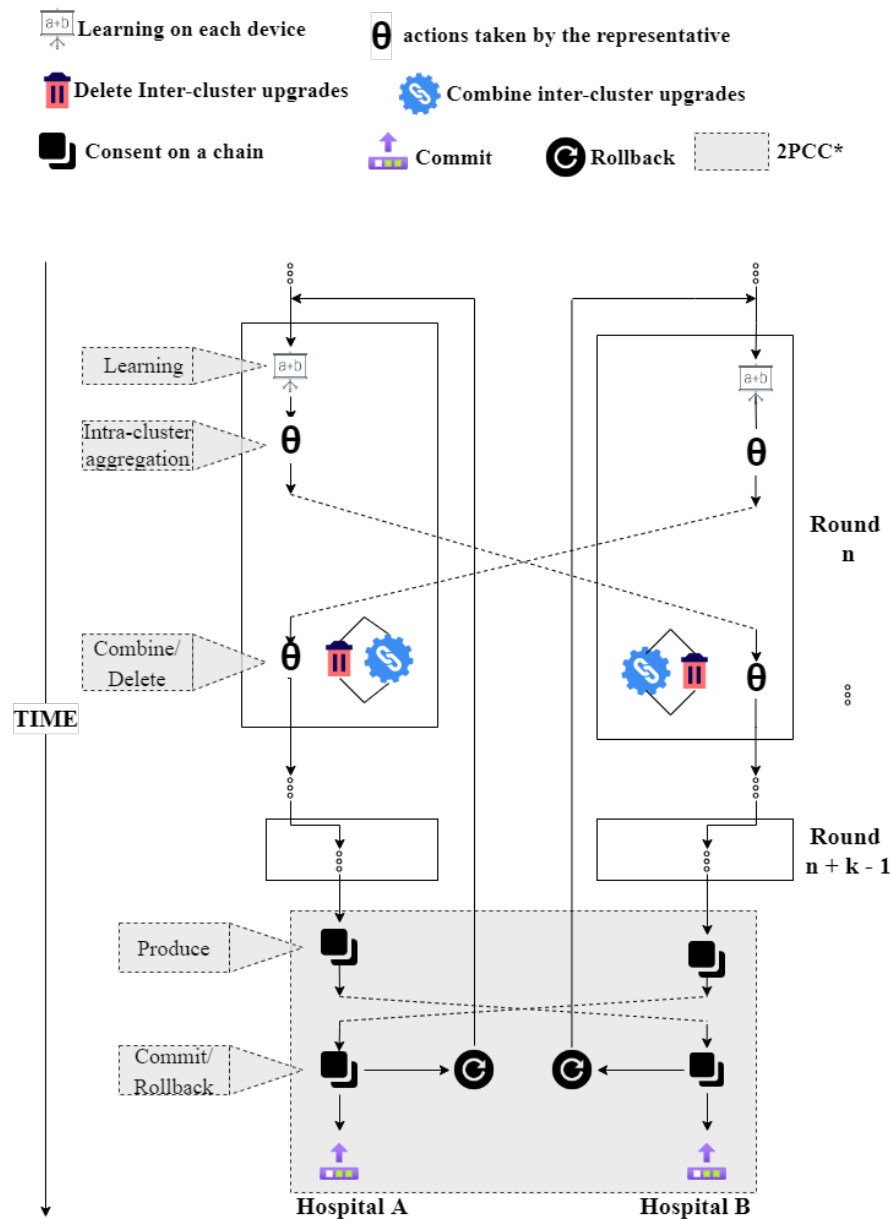
Figure 9. The DefCon algorithm's process flow [175].

Details of Consensus Round in DefCon—After *k* local rounds, the consensus protocol has an election for representative selection. Any collaborating peer node can apply for the election, in exchange for some assets to be mortgaged in the process. This is similar to

investment-based consensus approaches, where the node gets motivated to perform in a fair manner owing to the applied stake. In general, any candidate has a mortgage value, and all such mortgage values are stored in a set data structure. A sorting mechanism is applied, where the number of candidates is fixed. For example, q top-order candidates are allowed in the nomination process and are added to the representative pool. A randomized algorithm (designed on the ash process of the last added block at time t) is taken and is rounded off to map to a candidate serial number in the nomination pool. This makes the election process fair, random, verifiable, and traceable in the entire chain.

At the end of any working DefCon cycle, the efficiency of the 2PCC is examined, and in case of any inconsistency, the last verified state is rolled back. All the operations in the current cycle are deemed invalid. To validate the honesty of any miner node, a time-based difference Δ is computed, which is based on the current mining difficulty, and the previous nonce value. The block is added, and the miner is rewarded iff the time to add the new block $T_b \leq \Delta$. This step assures the elimination of foul play, and scenarios of collusion attacks are minimized. In case later it is found out that any false block is added, a factor $\mu(t)$ is computed, which is multiplied by the mortgaged amount $\mu(t) \times A$, and the value is decremented from the associated node as a penalty. Thus, in the next cycle, the node would have less stake and might be eliminated during the sorting cycle of top q miners. This reward/penalty scheme in DefCon allows the nodes to behave properly, and thus the gradient computation remains fair and valid.

10.1. Performance Evaluation

The installation of the CFL prototypes has been set up in the server machines to test the efficacy of our framework. Two prototypes—CFL-HstCon and CFL-DefCon—are implemented under the various consensus mechanisms intended for CFL. The authors have performed results based on the observations taken from [175], and have designed the same setup for the case-study. Table 12 shows the simulation parameters for the node and cluster setup. The experiments primarily focus on CFL- DefCon because it is more effective than CFL-HstCon. The CFL-DefCon outcome analysis has been discussed throughout the remainder of this section. To conduct the same, three different approaches, $Blockchain - FL_{26}$, which creates a network of 26 nodes, with 13 nodes in every shard, and HstCon and DefCon have been compared. In addition, the authors have identified trust probability as an important aspect in the event of a collusion attack. They have performed the effectiveness of the HstCon and DefCon consensus in the scenario when malicious nodes increase in the system, which is not conducted in the earlier study by [175]. The result signifies that consensus mechanisms must be robust to deal with fake transactions generated from the device to corrupt the learning process.

Table 12. Simulation parameter table.

Parameters	Value
Processor	2 eight-core Intel Xeon E5-2670
RAM	64 GB
Harddisk	8 TB
OS	Cent OS 7.2
No. of nodes used	13
The network capacity of the nodes across the racks	100 Mb
Latency added	200 ms

Figure 10a represents the learning latency of all three approaches. It is evident that, as we increase the number of rounds for collaborative aggregation, the system's latency also grows, consisting of training time and consensus time. The consensus time comprises both mainnet and crossnet, which verifies the block with certain rules. The $Blockchain - FL_{26}$

has higher latency than the cross-cluster (Hasty-Consensus). As we approach the 80 rounds of training, the algorithm takes $\approx 35\%$ less time as compared to $Blockchain - FL_{26}$; cross-cluster also outperforms the HstCon version in training the model.

Figure 10b represents the model comparison of different approaches where the dataset shards are divided into 13 random nodes, as $Blockchain - FL_{26}$ consists of two shards of 26 nodes. To reduce experimental error, every experiment is performed five times. Hence, this represents the mean value of five experiments and training rounds. With these settings, a significant improvement in accuracy is registered from 38% to 74%.

Figure 10c represents the trust model percentage of Proof-of-Work (PoW), Proof-of-Stake (PoS), and distributed lightweight consensus protocol (DLCP). The figure shows that PoW and PoS cannot identify the difference between genuine and fake transactions. As the number of fake transactions increases, the trust drops drastically in PoW and PoS approaches. Thus, these consensus algorithms cannot detect genuine transactions and add them to the main chain, which reduces trust in the blockchain network. The simulation is conducted on a PC with a virtual machine installed with certain assumptions. We assume that a single node can contain 1000 transactions, where we allowed 100 to 900 fake transactions to attack the consensus mechanism. The trust percentage is computed based on each device's trust rating, which is updated periodically if the transaction is successfully added to the block. Any new transaction generated by any device is forwarded to the neighboring node. In this case, the peer node checks the trust rating of initiating node and classifies the trustworthiness of the initiator.
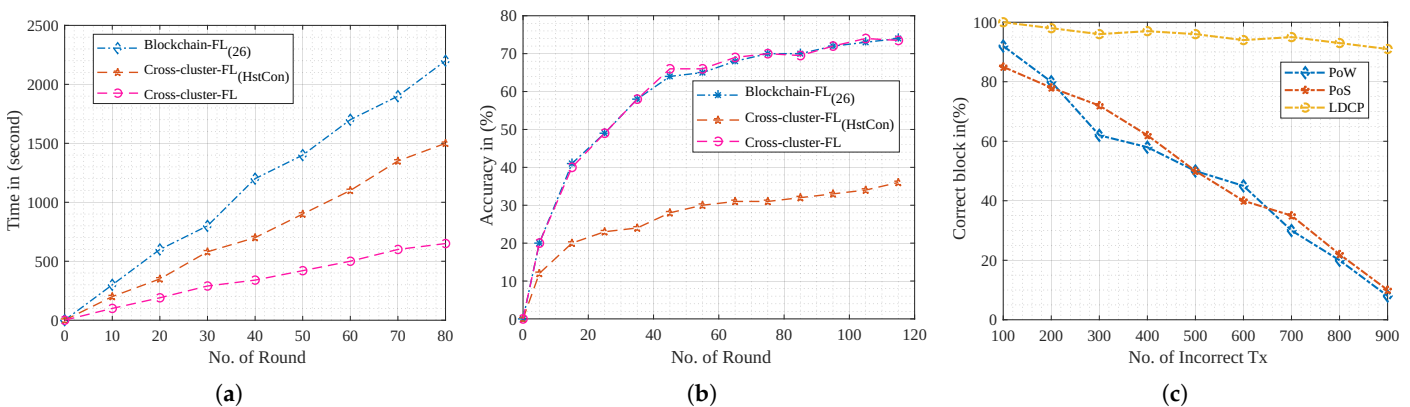

Figure 10. Performance Evaluation of the case-study. (**a**) System latency of different approaches. (**b**) Accuracy of model performance. (**c**) Trust percentage of consensus mechanism against fake transaction.

10.2. Learning Outcome of the Case Study

Integrating FL-IoMT in distributed setups opens up many new opportunities, but the data sharing and gradient exchange need to be secured and trusted. To alleviate the problem, BFL allows auditability and provenance in network setup. In BFL, however, resources are scarce, and thus communication-efficient consensus schemes are designed that address the issue of low efficiency and data sparsity. For an effective training outcome, it is crucial to identify and populate nodes geographically dispersed across a wide area into several small BFL clusters and use CFL to link these clusters. The aggregated improvements are transferred between the connected clusters to enhance the quality of the information samples for every cluster. As a result, aggregated updates' size becomes minimal, which lowers communication costs and significantly boosts the system's performance.

Despite the benefits, there are some unresolved challenges with BFL clusters trained via CFL setups. Even though CFL and BFL can enhance training results by communicating model updates, they might pose a significant communication and computational burden on these connected nodes, owing to the computational mining requirements, which eat up power and I/O resources in the network to assign common transactional ledger

snapshot state at all nodes. On the contrary, most IoMT devices typically have limited resources and thus cannot support these demands. To solve the issue, a fog-edge based computational might perform mining-as-a-service to nearby BFL nodes, simplifying the communication costs.

11. Conclusions and Future Scope

FL has been a game-changer in distributed AI learning paradigms. Owing to the large data conflux, the healthcare industry has shifted its core towards resource-effective FL algorithms. In IoMT ecosystems, however, with ease of learning comes the added burden of security and challenges of statistical disparity due to diverse and heterogeneous data at different nodes. FL addresses the dual issues of minimizing the learning convergence of models and keeping data in local silos, assuring data privacy and security. Coupled with blockchain, FL addresses the impending issue of trust, and recent research promises interesting blockchain-based FL architectures. The proposed survey orients its readers towards the key FL principles where discussion on resource requirements and analysis of cloud, edge, fog, and dew computing is discussed. We presented the FL-IoMT architecture, with an added emphasis on security and networking requirements, supported with practical use-case setups. We highlighted a useful case study of cross-FL in IoMT and presented the recent challenges in computation, data divergence, security, and networking issues.

In the future, we would explore the potential of secured multi-party communication with blockchain-based FL in the medical imaging domain. A differential-privacy preserving FL scheme would be proposed, with optimization conditions over the networking channels to address the network bandwidth and latency usage issues.

Author Contributions: Conceptualization: P.B., V.K.P., D.M., A.V. and S.T; writing—original draft preparation: V.K.P., P.B., A.S., A.K.T. and D.M.; methodology: P.B., R.S., A.A. and S.T.; writing—review and editing: S.T, F.-E.Ț., M.S.R., A.S. and A.K.T.; Software: A.S., A.A., R.S., P.B. and A.V; Visualization: M.S.R., F.-E.Ț., A.A., A.K.T. and S.T.; Investigation: R.S., A.S., V.K.P., S.T. and A.K.T. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Razdan, S.; Sharma, S. Internet of Medical Things (IoMT): Overview, Emerging Technologies, and Case Studies. *IETE Tech. Rev.* **2021**, *39*, 1–14. [CrossRef]
2. $55 Billion Global Internet of Medical Things Market Expected to Grow at a CAGR of 24.55% Between 2021 and 2026. Available online: https://www.globenewswire.com/news-release/2022/01/28/2374901/28124/en/55-Billion-Global-Internet-of-Medical-Things-Market-Expected-to-Grow-at-a-CAGR-of-24-55-Between-2021-and-2026 (accessed on 30 October 2022).
3. Internet of Medical Things (IoMT) Advances and Brings New Challenges. Available online: https://network-king.net/internet-of-medical-things-iomt-advances-and-brings-new-challenges/ (accessed on 30 October 2022).
4. Khan, L.U.; Saad, W.; Han, Z.; Hossain, E.; Hong, C.S. Federated Learning for Internet of Things: Recent Advances, Taxonomy, and Open Challenges. *IEEE Commun. Surv. Tutorials* **2021**, *23*, 1759–1799. [CrossRef]

5. Ghayvat, H.; Pandya, S.; Bhattacharya, P.; Zuhair, M.; Rashid, M.; Hakak, S.; Dev, K. *CP-BDHCA*: Blockchain-Based Confidentiality-Privacy Preserving Big Data Scheme for Healthcare Clouds and Applications. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 1937–1948. [CrossRef] [PubMed]

6. Shuaib, M.; Alam, S.; Shabbir Alam, M.; Shahnawaz Nasir, M. Compliance with HIPAA and GDPR in blockchain-based electronic health record. *Mater. Today Proc.* 2021, *in press*. [CrossRef]

7. Tortorella, G.L.; Fogliatto, F.S.; Mac Cawley Vergara, A.; Vassolo, R.; Sawhney, R. Healthcare 4.0: Trends, challenges and research directions. *Prod. Plan. Control.* **2020**, *31*, 1245–1260. [CrossRef]

8. Prasad, V.K.; Bhavsar, M.D. SLAMMP framework for cloud resource management and its impact on healthcare computational techniques. *Int. J. e-Health Med. Commun. IJEHMC* **2021**, *12*, 1–31. [CrossRef]

9. Saraswat, D.; Ladhiya, K.; Bhattacharya, P.; Zuhair, M. PHBio: A Pallier Homomorphic Biometric Encryption Scheme in Healthcare 4.0 Ecosystems. In Proceedings of the 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), London, UK, 27–29 April 2022; pp. 306–312. [CrossRef]

10. Khatri, S.; Vachhani, H.; Shah, S.; Bhatia, J.; Chaturvedi, M.; Tanwar, S.; Kumar, N. Machine Learning Models and Techniques for VANET Based Traffic Management: Implementation Issues and Challenges. *Peer-to-Peer Netw. Appl.* **2021**, *14*, 1778–1805. [CrossRef]

11. Giannakis, G.B.; Ling, Q.; Mateos, G.; Schizas, I.D.; Zhu, H. Decentralized learning for wireless communications and networking. In *Splitting Methods in Communication, Imaging, Science, and Engineering*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 461–497.

12. Bhattacharya, P.; Mehta, P.; Tanwar, S.; Obaidat, M.S.; Hsiao, K.F. HeaL: A blockchain-envisioned signcryption scheme for healthcare IoT ecosystems. In Proceedings of the 2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI), Sharjah, United Arab Emirates, 3–5 November 2020; pp. 1–6. [CrossRef]

13. Verma, A.; Bhattacharya, P.; Bodkhe, U.; Ladha, A.; Tanwar, S. DAMS: Dynamic Association for View Materialization Based on Rule Mining Scheme. *Lect. Notes Electr. Eng.* **2021**, *701*, 529–544. [CrossRef]

14. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the Artificial Intelligence and Statistics, PMLR, Ft. Lauderdale, FL, USA, 20–27 April 2017; pp. 1273–1282.

15. Chai, Z.; Fayyaz, H.; Fayyaz, Z.; Anwar, A.; Zhou, Y.; Baracaldo, N.; Ludwig, H.; Cheng, Y. Towards Taming the Resource and Data Heterogeneity in Federated Learning. In Proceedings of the 2019 USENIX Conference on Operational Machine Learning (OpML 19), Santa Clara, CA, USA, 20 May 2019; pp. 19–21.

16. Prasad, V.K.; Bhavsar, M. Efficient resource monitoring and prediction techniques in an IaaS level of cloud computing: Survey. In Proceedings of the International Conference on Future Internet Technologies and Trends, Surat, India, 31 August–2 September 2017; pp. 47–55.

17. Ghimire, B.; Rawat, D.B. Recent Advances on Federated Learning for Cybersecurity and Cybersecurity for Federated Learning for Internet of Things. *IEEE Internet Things J.* **2022**, *9*, 8229–8249. [CrossRef]

18. Patel, V.A.; Bhattacharya, P.; Tanwar, S.; Jadav, N.K.; Gupta, R. BFLEdge: Blockchain based federated edge learning scheme in V2X underlying 6G communications. In Proceedings of the 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 27–28 January 2022; pp. 146–152. [CrossRef]

19. Saraswat, D.; Verma, A.; Bhattacharya, P.; Tanwar, S.; Sharma, G.; Bokoro, P.N.; Sharma, R. Blockchain-Based Federated Learning in UAVs Beyond 5G Networks: A Solution Taxonomy and Future Directions. *IEEE Access* **2022**, *10*, 33154–33182. [CrossRef]

20. Alam, T.; Gupta, R. Federated Learning and Its Role in the Privacy Preservation of IoT Devices. *Future Internet* **2022**, *14*, 246. [CrossRef]

21. Byrd, D.; Polychroniadou, A. Differentially Private Secure Multi-Party Computation for Federated Learning in Financial Applications. In Proceedings of the First ACM International Conference on AI in Finance, New York, NY, USA, 15–16 October 2020. [CrossRef]

22. Verma, A.; Bhattacharya, P.; Zuhair, M.; Tanwar, S.; Kumar, N. VaCoChain: Blockchain-Based 5G-Assisted UAV Vaccine Distribution Scheme for Future Pandemics. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 1997–2007. [CrossRef] [PubMed]

23. Chen, M.; Poor, H.V.; Saad, W.; Cui, S. Wireless Communications for Collaborative Federated Learning. *IEEE Commun. Mag.* **2020**, *58*, 48–54. [CrossRef]

24. Falcetta, A.; Roveri, M. Privacy-Preserving Deep Learning With Homomorphic Encryption: An Introduction. *IEEE Comput. Intell. Mag.* **2022**, *17*, 14–25. [CrossRef]

25. Pappas, C.; Chatzopoulos, D.; Lalis, S.; Vavalis, M. Ipls: A framework for decentralized federated learning. In Proceedings of the 2021 IFIP Networking Conference (IFIP Networking), Espoo and Helsinki, Finland, 21–24 June 2021; pp. 1–6.

26. Nguyen, D.C.; Pham, Q.V.; Pathirana, P.N.; Ding, M.; Seneviratne, A.; Lin, Z.; Dobre, O.; Hwang, W.J. Federated Learning for Smart Healthcare: A Survey. *ACM Comput. Surv.* **2022**, *55*, 60. [CrossRef]

27. Camajori Tedeschini, B.; Savazzi, S.; Stoklasa, R.; Barbieri, L.; Stathopoulos, I.; Nicoli, M.; Serio, L. Decentralized Federated Learning for Healthcare Networks: A Case Study on Tumor Segmentation. *IEEE Access* **2022**, *10*, 8693–8708. [CrossRef]

28. Dayan, I.; Roth, H.R.; Zhong, A.; Harouni, A.; Gentili, A.; Abidin, A.Z.; Liu, A.; Costa, A.B.; Wood, B.J.; Tsai, C.S.; et al. Federated learning for predicting clinical outcomes in patients with COVID-19. *Nat. Med.* **2021**, *27*, 1735–1743. [CrossRef]

29. IBM to Buy Merge Healthcare in 1 Billion Deal. Available online: https://www.reuters.com/article/us-merge-healthcare-m-a-ibm/ibm-to-buy-merge-healthcare-in-1-billion-deal-idUSKCN0QB1ML20150806 (accessed on 9 October 2022).

30. Darzidehkalani, E.; Ghasemi-rad, M.; van Ooijen, P. Federated Learning in Medical Imaging: Part I: Toward Multicentral Health Care Ecosystems. *J. Am. Coll. Radiol.* **2022**, *19*, 969–974. [CrossRef]

31. Antunes, R.S.; André da Costa, C.; Küderle, A.; Yari, I.A.; Eskofier, B. Federated Learning for Healthcare: Systematic Review and Architecture Proposal. *ACM Trans. Intell. Syst. Technol. TIST* **2022**, *13*, 1–23. [CrossRef]

32. KhoKhar, F.A.; Shah, J.H.; Khan, M.A.; Sharif, M.; Tariq, U.; Kadry, S. A review on federated learning towards image processing. *Comput. Electr. Eng.* **2022**, *99*, 107818. . [CrossRef]

33. Khan, M.A.; Alkaabi, N. Rebirth of Distributed AI—A Review of eHealth Research. *Sensors* **2021**, *21*, 4999. [CrossRef] [PubMed]

34. Nguyen, D.C.; Ding, M.; Pham, Q.V.; Pathirana, P.N.; Le, L.B.; Seneviratne, A.; Li, J.; Niyato, D.; Poor, H.V. Federated learning meets blockchain in edge computing: Opportunities and challenges. *IEEE Internet Things J.* **2021**, *8*, 12806–12825. [CrossRef]

35. Al-Turjman, F.; Nawaz, M.H.; Ulusar, U.D. Intelligence in the Internet of Medical Things era: A systematic review of current and future trends. *Comput. Commun.* **2020**, *150*, 644–660. [CrossRef]

36. Aouedi, O.; Sacco, A.; Piamrat, K.; Marchetto, G. Handling Privacy-Sensitive Medical Data With Federated Learning: Challenges and Future Directions. *IEEE J. Biomed. Health Inform.* **2022**, *in press*. [CrossRef] [PubMed]

37. Ali, M.; Naeem, F.; Tariq, M.; Kaddoum, G. Federated Learning for Privacy Preservation in Smart Healthcare Systems: A Comprehensive Survey. *arXiv* **2022**, arXiv:2203.09702.

38. Shen, S.; Zhu, T.; Wu, D.; Wang, W.; Zhou, W. From distributed machine learning to federated learning: In the view of data privacy and security. *Concurr. Comput. Pract. Exp.* **2022**, *34*, e6002. [CrossRef]

39. Wu, X.; Zhang, Y.; Shi, M.; Li, P.; Li, R.; Xiong, N.N. An adaptive federated learning scheme with differential privacy preserving. *Future Gener. Comput. Syst.* **2022**, *127*, 362–372. [CrossRef]

40. Rahmadika, S.; Astillo, P.V.; Choudhary, G.; Duguma, D.G.; Sharma, V.; You, I. Blockchain-based Privacy Preservation Scheme for Misbehavior Detection in Lightweight IoMT Devices. *IEEE J. Biomed. Health Inform.* **2022**, *in press*. [CrossRef]

41. Alamleh, A.; Albahri, O.S.; Zaidan, A.; Albahri, A.; Alamoodi, A.; Zaidan, B.; Qahtan, S.; Alsatar, H.; Al-Samarraay, M.S.; Jasim, A.N. Federated Learning for IoMT Applications: A Standardisation and Benchmarking Framework of Intrusion Detection Systems. *IEEE J. Biomed. Health Inform.* **2022**, *in press*. [CrossRef]

42. Gupta, D.; Kayode, O.; Bhatt, S.; Gupta, M.; Tosun, A.S. Hierarchical federated learning based anomaly detection using digital twins for smart healthcare. In Proceedings of the 2021 IEEE 7th International Conference on Collaboration and Internet Computing (CIC), Atlanta, GA, USA, 13–15 December 2021; pp. 16–25.

43. Zhao, J.; Zhu, X.; Wang, J.; Xiao, J. Efficient Client Contribution Evaluation for Horizontal Federated Learning. In Proceedings of the ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021; pp. 3060–3064. [CrossRef]

44. Ahmed, J.; Nguyen, T.N.; Ali, B.; Javed, A.; Mirza, J. On the Physical Layer Security of Federated Learning based IoMT Networks. *IEEE J. Biomed. Health Inform.* **2022**, *in press*. [CrossRef]

45. Xu, Z.; Guo, Y.; Chakraborty, C.; Hua, Q.; Chen, S.; Yu, K. A Simple Federated Learning-based Scheme for Security Enhancement over Internet of Medical Things. *IEEE J. Biomed. Health Inform.* **2022**, *in press*. [CrossRef] [PubMed]

46. Zhang, J.; Li, Y.; Xiao, W.; Zhang, Z. Non-iterative and Fast Deep Learning: Multilayer Extreme Learning Machines. *J. Frankl. Inst.* **2020**, *357*, 8925–8955. [CrossRef]

47. Zhang, J.; Li, Y.; Xiao, W.; Zhang, Z. Robust extreme learning machine for modeling with unknown noise. *J. Frankl. Inst.* **2020**, *357*, 9885–9908. [CrossRef]

48. Sanyal, S.; Wu, D.; Nour, B. A federated filtering framework for Internet of medical things. In Proceedings of the ICC 2019—2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; pp. 1–6.

49. Ruby, R.; Yang, H.; Wu, K. Anti-Jamming Strategies for Federated Learning Internet of Medical Things: A Game Approach. *IEEE J. Biomed. Health Inform.* **2022**, *in press*. [CrossRef] [PubMed]

50. Ferrag, M.A.; Friha, O.; Maglaras, L.; Janicke, H.; Shu, L. Federated Deep Learning for Cyber Security in the Internet of Things: Concepts, Applications, and Experimental Analysis. *IEEE Access* **2021**, *9*, 138509–138542. [CrossRef]

51. Can, Y.S.; Ersoy, C. Privacy-preserving federated deep learning for wearable IoT-based biomedical monitoring. *ACM Trans. Internet Technol. TOIT* **2021**, *21*, 1–17. [CrossRef]

52. Dai, H.N.; Imran, M.; Haider, N. Blockchain-enabled internet of medical things to combat COVID-19. *IEEE Internet Things Mag.* **2020**, *3*, 52–57. [CrossRef]

53. Cheng, W.; Ou, W.; Yin, X.; Yan, W.; Liu, D.; Liu, C. A privacy-protection model for patients. *Secur. Commun. Networks* **2020**, *2020*. [CrossRef]

54. Wang, R.; Lai, J.; Zhang, Z.; Li, X.; Vijayakumar, P.; Karuppiah, M. Privacy-Preserving Federated Learning for Internet of Medical Things under Edge Computing. *IEEE J. Biomed. Health Inform.* **2022**, *in press*. [CrossRef]

55. Wu, M.; Ye, D.; Ding, J.; Guo, Y.; Yu, R.; Pan, M. Incentivizing Differentially Private Federated Learning: A Multidimensional Contract Approach. *IEEE Internet Things J.* **2021**, *8*, 10639–10651. [CrossRef]

56. Lakhan, A.; Mohammed, M.A.; Nedoma, J.; Martinek, R.; Tiwari, P.; Vidyarthi, A.; Alkhayyat, A.; Wang, W. Federated-Learning Based Privacy Preservation and Fraud-Enabled Blockchain IoMT System for Healthcare. *IEEE J. Biomed. Health Inform.* **2022**, *in press*. [CrossRef] [PubMed]

57. Samuel, O.; Omojo, A.B.; Onuja, A.M.; Sunday, Y.; Tiwari, P.; Gupta, D.; Hafeez, G.; Yahaya, A.S.; Fatoba, O.J.; Shamshirband, S. IoMT: A COVID-19 Healthcare System driven by Federated Learning and Blockchain. *IEEE J. Biomed. Health Inform.* 2022, *in press*. [CrossRef] [PubMed]

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

59. Federated Learning Global Forecast to 2028. Available online: https://www.marketsandmarkets.com/ (accessed on 27 September 2022).

60. Omidfar, K.; Ahmadi, A.; Syedmoradi, L.; Khoshfetrat, S.M.; Larijani, B. Point-of-care biosensors in medicine: A brief overview of our achievements in this field based on the conducted research in EMRI (endocrinology and metabolism research Institute of Tehran University of medical sciences) over the past fourteen years. *J. Diabetes Metab. Disord.* **2020**, *in press*. [CrossRef] [PubMed]

61. Bahl, S.; Bagha, A.K.; Rab, S.; Javaid, M.; Haleem, A.; Singh, R.P. Advancements in biosensor technologies for medical field and COVID-19 pandemic. *J. Ind. Integr. Manag.* **2021**, *6*, 175–191. [CrossRef]

62. Li, J.; Stachowski, M.; Zhang, Z. 11 - Application of responsive polymers in implantable medical devices and biosensors. In *Switchable and Responsive Surfaces and Materials for Biomedical Applications*; Zhang, Z., Ed.; Woodhead Publishing: Oxford, UK, 2015; pp. 259–298. . [CrossRef]

63. Srinivasan, K.; Mahendran, N.; Vincent, D.R.; Chang, C.Y.; Syed-Abdul, S. Realizing an Integrated Multistage Support Vector Machine Model for Augmented Recognition of Unipolar Depression. *Electronics* **2020**, *9*, 647. [CrossRef]

64. Guizani, K.; Guizani, S. IoT Healthcare Monitoring Systems Overview for Elderly Population. In Proceedings of the 2020 International Wireless Communications and Mobile Computing (IWCMC), Limassol, Cyprus, 15–19 June 2020; pp. 2005–2009. [CrossRef]

65. Kazemi, Z.; Papadimitriou, A.; Hely, D.; Fazcli, M.; Beroulle, V. Hardware Security Evaluation Platform for MCU-Based Connected Devices: Application to Healthcare IoT. In Proceedings of the 2018 IEEE 3rd International Verification and Security Workshop (IVSW), Costa Brava, Spain, 2–4 July 2018; pp. 87–92. [CrossRef]

66. Zhou, Y.; Ye, Q.; Lv, J. Communication-Efficient Federated Learning With Compensated Overlap-FedAvg. *IEEE Trans. Parallel Distrib. Syst.* **2022**, *33*, 192–205. [CrossRef]

67. Yuan, X.T.; Li, P. On Convergence of FedProx: Local Dissimilarity Invariant Bounds, Non-smoothness and Beyond. *arXiv* **2022**, arXiv:2206.05187.

68. Guberović, E.; Lipić, T.; Čavrak, I. Dew Intelligence: Federated learning perspective. In Proceedings of the 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), Madrid, Spain, 12–16 July 2021; pp. 1819–1824. [CrossRef]

69. 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]

70. Poirot, M.G.; Vepakomma, P.; Chang, K.; Kalpathy-Cramer, J.; Gupta, R.; Raskar, R. Split Learning for collaborative deep learning in healthcare. *arXiv* **2019**, arXiv:1912.12115.

71. 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] [PubMed]

72. Bodkhe, U.; Tanwar, S.; Bhattacharya, P.; Verma, A. Blockchain Adoption for Trusted Medical Records in Healthcare 4.0 Applications: A Survey. *Lect. Notes Netw. Syst.* **2021**, *203 LNNS*, 759–774. [CrossRef]

73. Saraswat, D.; Bhattacharya, P.; Verma, A.; Prasad, V.K.; Tanwar, S.; Sharma, G.; Bokoro, P.N.; Sharma, R. Explainable AI for Healthcare 5.0: Opportunities and Challenges. *IEEE Access* **2022**, *10*, 84486–84517. [CrossRef]

74. Verma, A.; Bhattacharya, P.; Bodkhe, U.; Zuhair, M.; Dewangan, R.K. Blockchain-Based Federated Cloud Environment: Issues and Challenges. *Blockchain for Information Security and Privacy*; CRC Press: Boca Raton, Fl, USA, 2021; pp. 155–176.

75. Cloud Computing: Statistics and Facts. Available online: https://www.statista.com/topics/1695/cloud-computing/ (accessed on 27 September 2022).

76. Sahi, M.A.; Abbas, H.; Saleem, K.; Yang, X.; Derhab, A.; Orgun, M.A.; Iqbal, W.; Rashid, I.; Yaseen, A. Privacy Preservation in e-Healthcare Environments: State of the Art and Future Directions. *IEEE Access* **2018**, *6*, 464–478. [CrossRef]

77. Chen, D.; Wang, H.; Zhang, N.; Nie, X.; Dai, H.N.; Zhang, K.; Raymond Choo, K.K. Privacy-Preserving Encrypted Traffic Inspection With Symmetric Cryptographic Techniques in IoT. *IEEE Internet Things J.* **2022**, *9*, 17265–17279. [CrossRef]

78. Verma, A.; Bhattacharya, P.; Saraswat, D.; Tanwar, S. NyaYa: Blockchain-based electronic law record management scheme for judicial investigations. *J. Inf. Secur. Appl.* **2021**, *63*, 103025. [CrossRef]

79. Yuan, B.; Ge, S.; Xing, W. A federated learning framework for healthcare iot devices. *arXiv* **2020**, arXiv:2005.05083.

80. Thilakarathne, N.N.; Muneeswari, G.; Parthasarathy, V.; Alassery, F.; Hamam, H.; Mahendran, R.K.; Shafiq, M. Federated Learning for Privacy-Preserved Medical Internet of Things. *Intell. Autom. Soft Comput.* **2022**, *33*, 157–172. [CrossRef]

81. Verma, A.; Bhattacharya, P.; Patel, Y.; Shah, K.; Tanwar, S.; Khan, B. Data Localization and Privacy-Preserving Healthcare for Big Data Applications: Architecture and Future Directions. *Lect. Notes Electr. Eng.* **2022**, *875*, 233—244._18. [CrossRef]

82. Kim, H.; Park, J.; Bennis, M.; Kim, S.L. Blockchained On-Device Federated Learning. *arXiv* **2018**, arxiv:1808.03949.

83. Smith, V.; Chiang, C.K.; Sanjabi, M.; Talwalkar, A.S. Federated Multi-Task Learning. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30.

84. Tanwar, S.; Vora, J.; Kaneriya, S.; Tyagi, S.; Kumar, N.; Sharma, V.; You, I. Human Arthritis Analysis in Fog Computing Environment Using Bayesian Network Classifier and Thread Protocol. *IEEE Consum. Electron. Mag.* **2020**, *9*, 88–94. [CrossRef]

85. Cheng, K.; Fan, T.; Jin, Y.; Liu, Y.; Chen, T.; Papadopoulos, D.; Yang, Q. SecureBoost: A Lossless Federated Learning Framework. *IEEE Intell. Syst.* **2021**, *36*, 87–98. [CrossRef]

86. Fang, H.; Qian, Q. Privacy Preserving Machine Learning with Homomorphic Encryption and Federated Learning. *Future Internet* **2021**, *13*, 94. [CrossRef]

87. IoT: Medical Devices, Statistics and Facts. Available online: https://www.statista.com/outlook/hmo/medical-technology/medical-devices/india (accessed on 27 September 2022).

88. Wadu, M.M.; Samarakoon, S.; Bennis, M. Federated Learning under Channel Uncertainty: Joint Client Scheduling and Resource Allocation. *arXiv* **2020**, arXiv:2002.00802.

89. Konečný, J.; McMahan, H.B.; Ramage, D.; Richtárik, P. Federated Optimization: Distributed Machine Learning for On-Device Intelligence. *arXiv* **2016**, arXiv:1610.02527.

90. Prasad, V.K.; Bhavsar, M.D. Monitoring IaaS cloud for healthcare systems: Healthcare information management and cloud resources utilization. *Int. J. e-Health Med Commun. IJEHMC* **2020**, *11*, 54–70. [CrossRef]

91. Prasad, V.K.; Bhavsar, M.D.; Tanwar, S. Influence of montoring: Fog and edge computing. *Scalable Comput. Pract. Exp.* **2019**, *20*, 365–376. [CrossRef]

92. Divya, K.; Sirohi, A.; Pande, S.; Malik, R. An IoMT assisted heart disease diagnostic system using machine learning techniques. In *Cognitive Internet of Medical Things for Smart Healthcare*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 145–161.

93. Iwendi, C.; Khan, S.; Anajemba, J.H.; Bashir, A.K.; Noor, F. Realizing an Efficient IoMT-Assisted Patient Diet Recommendation System Through Machine Learning Model. *IEEE Access* **2020**, *8*, 28462–28474. [CrossRef]

94. Bibi, N.; Sikandar, M.; Ud Din, I.; Almogren, A.; Ali, S. IoMT-based automated detection and classification of leukemia using deep learning. *J. Healthc. Eng.* **2020**, *2020*, 6648574. [CrossRef]

95. Khan, M.F.; Ghazal, T.M.; Said, R.A.; Fatima, A.; Abbas, S.; Khan, M.; Issa, G.F.; Ahmad, M.; Khan, M.A. An IoMT-Enabled Smart Healthcare Model to Monitor Elderly People Using Machine Learning Technique. *Comput. Intell. Neurosci.* **2021**. [CrossRef]

96. Sekhar, A.; Biswas, S.; Hazra, R.; Sunaniya, A.K.; Mukherjee, A.; Yang, L. Brain Tumor Classification Using Fine-Tuned GoogLeNet Features and Machine Learning Algorithms: IoMT Enabled CAD System. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 983–991. [CrossRef] [PubMed]

97. Rahman, A.u.; Nasir, M.U.; Gollapalli, M.; Alsaif, S.A.; Almadhor, A.S.; Mehmood, S.; Khan, M.A.; Mosavi, A. IoMT-Based Mitochondrial and Multifactorial Genetic Inheritance Disorder Prediction Using Machine Learning. *Comput. Intell. Neurosci.* **2022**, *2022*, 2650742. [CrossRef] [PubMed]

98. 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] [PubMed]

99. Silva, S.; Altmann, A.; Gutman, B.; Lorenzi, M. Fed-biomed: A general open-source frontend framework for federated learning in healthcare. In *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 201–210.

100. Hölbl, M.; Kompara, M.; Kamišalić, A.; Nemec Zlatolas, L. A systematic review of the use of blockchain in healthcare. *Symmetry* **2018**, *10*, 470. [CrossRef]

101. Long, G.; Shen, T.; Tan, Y.; Gerrard, L.; Clarke, A.; Jiang, J. Federated learning for privacy-preserving open innovation future on digital health. In *Humanity Driven AI*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 113–133.

102. Esteva, A.; Robicquet, A.; Ramsundar, B.; Kuleshov, V.; DePristo, M.; Chou, K.; Cui, C.; Corrado, G.; Thrun, S.; Dean, J. A guide to deep learning in healthcare. *Nat. Med.* **2019**, *25*, 24–29. [CrossRef] [PubMed]

103. Raza, A.; Tran, K.P.; Koehl, L.; Li, S. Designing ecg monitoring healthcare system with federated transfer learning and explainable ai. *Knowl.-Based Syst.* **2022**, *236*, 107763. [CrossRef]

104. Shukla, P.K.; Zakariah, M.; Hatamleh, W.A.; Tarazi, H.; Tiwari, B. AI-DRIVEN novel approach for liver cancer screening and prediction using cascaded fully convolutional neural network. *J. Healthc. Eng.* **2022**, *2022*, 4277436. [CrossRef]

105. Thomsen, M.K.; Pedersen, L.; Erichsen, R.; Lash, T.L.; Sørensen, H.T.; Mikkelsen, E.M. Risk-stratified selection to colonoscopy in FIT colorectal cancer screening: Development and temporal validation of a prediction model. *Br. J. Cancer* **2022**, *126*, 1229–1235. [CrossRef]

106. Flores, M.; Dayan, I.; Roth, H.; Zhong, A.; Harouni, A.; Gentili, A.; Abidin, A.; Liu, A.; Costa, A.; Wood, B.; et al. Federated Learning used for predicting outcomes in SARS-COV-2 patients. *Res. Sq.* **2021**, *in press*.

107. Barbiero, P.; Ciravegna, G.; Giannini, F.; Lió, P.; Gori, M.; Melacci, S. Entropy-based logic explanations of neural networks. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual Event, 22 February–1 March 2022; Volume 36, pp. 6046–6054.

108. Singh, S.K.; Salim, M.M.; Cha, J.; Pan, Y.; Park, J.H. Machine learning-based network sub-slicing framework in a sustainable 5g environment. *Sustainability* **2020**, *12*, 6250. [CrossRef]

109. Yadav, S.S.; Jadhav, S.M. Deep convolutional neural network based medical image classification for disease diagnosis. *J. Big Data* **2019**, *6*, 1–18. [CrossRef]

110. Du, X.; Zhang, H.; Van Nguyen, H.; Han, Z. Stacked LSTM deep learning model for traffic prediction in vehicle-to-vehicle communication. In Proceedings of the 2017 IEEE 86th Vehicular Technology Conference (VTC-Fall), Toronto, ON, Canada, 24–27 September 2017; pp. 1–5.

111. Li, L.; Xu, Y.; Zhang, Z.; Yin, J.; Chen, W.; Han, Z. A prediction-based charging policy and interference mitigation approach in the wireless powered Internet of Things. *IEEE J. Sel. Areas Commun.* **2018**, *37*, 439–451. [CrossRef]

112. Hewamalage, H.; Bergmeir, C.; Bandara, K. Recurrent neural networks for time series forecasting: Current status and future directions. *Int. J. Forecast.* **2021**, *37*, 388–427. [CrossRef]

113. Sarp, S.; Kuzlu, M.; Wilson, E.; Cali, U.; Guler, O. The enlightening role of explainable artificial intelligence in chronic wound classification. *Electronics* **2021**, *10*, 1406. [CrossRef]

114. Rucco, M.; Viticchi, G.; Falsetti, L. Towards personalized diagnosis of glioblastoma in fluid-attenuated inversion recovery (FLAIR) by topological interpretable machine learning. *Mathematics* **2020**, *8*, 770. [CrossRef]

115. Wang, S.; Tuor, T.; Salonidis, T.; Leung, K.K.; Makaya, C.; He, T.; Chan, K. Adaptive federated learning in resource constrained edge computing systems. *IEEE J. Sel. Areas Commun.* **2019**, *37*, 1205–1221. [CrossRef]

116. Wang, X.; Han, Y.; Wang, C.; Zhao, Q.; Chen, X.; Chen, M. In-edge ai: Intelligentizing mobile edge computing, caching and communication by federated learning. *IEEE Netw.* **2019**, *33*, 156–165. [CrossRef]

117. Chen, D.; Xie, L.J.; Kim, B.; Wang, L.; Hong, C.S.; Wang, L.C.; Han, Z. Federated learning based mobile edge computing for augmented reality applications. In Proceedings of the 2020 International Conference on Computing, Networking and Communications (ICNC), Big Island, HI, USA, 17–20 February, 2020; pp. 767–773.

118. Feraudo, A.; Yadav, P.; Safronov, V.; Popescu, D.A.; Mortier, R.; Wang, S.; Bellavista, P.; Crowcroft, J. CoLearn: Enabling federated learning in MUD-compliant IoT edge networks. In Proceedings of the Third ACM International Workshop on Edge Systems, Analytics and Networking, Heraklion Greece, 27 April 2020; pp. 25–30.

119. Połap, D.; Srivastava, G.; Yu, K. Agent architecture of an intelligent medical system based on federated learning and blockchain technology. *J. Inf. Secur. Appl.* **2021**, *58*, 102748. [CrossRef]

120. Kumar, Y.; Singla, R. Federated learning systems for healthcare: Perspective and recent progress. *Fed. Learn. Syst.* **2021**, 141–156. [CrossRef]

121. Rehman, A.; Abbas, S.; Khan, M.; Ghazal, T.M.; Adnan, K.M.; Mosavi, A. A secure healthcare 5.0 system based on blockchain technology entangled with federated learning technique. *Comput. Biol. Med.* **2022**, 106019. [CrossRef]

122. Zuhair, M.; Patel, F.; Navapara, D.; Bhattacharya, P.; Saraswat, D. BloCoV6: A blockchain-based 6G-assisted UAV contact tracing scheme for COVID-19 pandemic. In Proceedings of the 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), London, UK, 28–30 April 2021; pp. 271–276. [CrossRef]

123. Arikumar, K.S.; Prathiba, S.B.; Alazab, M.; Gadekallu, T.R.; Pandya, S.; Khan, J.M.; Moorthy, R.S. FL-PMI: Federated Learning-Based Person Movement Identification through Wearable Devices in Smart Healthcare Systems. *Sensors* **2022**, *22*, 1377. [CrossRef] [PubMed]

124. Yu, S.; Qian, W.; Jannesari, A. Resource-aware Federated Learning using Knowledge Extraction and Multi-model Fusion. *arXiv* **2022**, arXiv:2208.07978.

125. Wang, Z.; Hu, Y.; Yan, S.; Wang, Z.; Hou, R.; Wu, C. Efficient Ring-Topology Decentralized Federated Learning with Deep Generative Models for Medical Data in eHealthcare Systems. *Electronics* **2022**, *11*, 1548. [CrossRef]

126. Fan, J.; Wang, X.; Guo, Y.; Hu, X.; Hu, B. Federated Learning Driven Secure Internet of Medical Things. *IEEE Wirel. Commun.* **2022**, *29*, 68–75. [CrossRef]

127. Jiang, J.C.; Kantarci, B.; Oktug, S.; Soyata, T. Federated learning in smart city sensing: Challenges and opportunities. *Sensors* **2020**, *20*, 6230. [CrossRef]

128. Prasad, V.K.; Bhattacharya, P.; Bhavsar, M.; Verma, A.; Tanwar, S.; Sharma, G.; Bokoro, P.N.; Sharma, R. ABV-CoViD: An Ensemble Forecasting Model to Predict Availability of Beds and Ventilators for COVID-19 Like Pandemics. *IEEE Access* **2022**, *10*, 74131–74151. [CrossRef]

129. Yan, B.; Wang, J.; Cheng, J.; Zhou, Y.; Zhang, Y.; Yang, Y.; Liu, L.; Zhao, H.; Wang, C.; Liu, B. Experiments of federated learning for COVID-19 chest X-ray images. In Proceedings of the International Conference on Artificial Intelligence and Security, Dublin, Ireland, 19–23 July 2021; pp. 41–53.

130. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.

131. Ayyachamy, S.; Alex, V.; Khened, M.; Krishnamurthi, G. Medical image retrieval using Resnet-18. In Proceedings of the Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications, San Diego, CA, USA, 17–18 February 2019; Volume 10954, pp. 233–241.

132. Wang, L.; Lin, Z.Q.; Wong, A. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest X-ray images. *Sci. Rep.* **2020**, *10*, 1–12. [CrossRef]

133. Sharma, A.; Muttoo, S.K. Spatial image steganalysis based on resnext. In Proceedings of the 2018 IEEE 18th International Conference on Communication Technology (ICCT), Chongqing, China, 8–11 October 2018; pp. 1213–1216.

134. Zhao, Y.; Li, M.; Lai, L.; Suda, N.; Civin, D.; Chandra, V. Federated learning with non-iid data. *arXiv* **2018**, arXiv:1806.00582.

135. Kholod, I.; Yanaki, E.; Fomichev, D.; Shalugin, E.; Novikova, E.; Filippov, E.; Nordlund, M. Open-source federated learning frameworks for IoT: A comparative review and analysis. *Sensors* **2020**, *21*, 167. [CrossRef]

136. Hao, M.; Li, H.; Xu, G.; Liu, Z.; Chen, Z. Privacy-aware and resource-saving collaborative learning for healthcare in cloud computing. In Proceedings of the ICC 2020—2020 IEEE International Conference on Communications (ICC), Online, 7–11 June 2020; pp. 1–6.

137. Vepakomma, P.; Gupta, O.; Swedish, T.; Raskar, R. Split learning for health: Distributed deep learning without sharing raw patient data. *arXiv* **2018**, arXiv:1812.00564.

138. 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.

139. Lu, Z.; Cao, G.; La Porta, T. Teamphone: Networking smartphones for disaster recovery. *IEEE Trans. Mob. Comput.* **2017**, *16*, 3554–3567. [CrossRef]

140. Liu, D.; Dligach, D.; Miller, T. Two-stage federated phenotyping and patient representation learning. In Proceedings of the Association for Computational Linguistics, Conference, Florence, Italy, July 2019, Volume 2019, p. 283.

141. Liu, D.; Fox, K.; Weber, G.; Miller, T. Confederated machine learning on horizontally and vertically separated medical data for large-scale health system intelligence. *arXiv* **2019**, arXiv:1910.02109.

142. Xu, X.; Peng, H.; Sun, L.; Bhuiyan, M.Z.A.; Liu, L.; He, L. Fedmood: Federated learning on mobile health data for mood detection. *arXiv* **2021**, arXiv:2102.09342.

143. Yan, Z.; Wicaksana, J.; Wang, Z.; Yang, X.; Cheng, K.T. Variation-aware federated learning with multi-source decentralized medical image data. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 2615–2628. [CrossRef]

144. 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 tumor segmentation. In Proceedings of the International Workshop on Machine Learning in Medical Imaging, Shenzhen, China, 13 October 2019; pp. 133–141.

145. 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; pp. 270–274.

146. Srivastava, U.C.; Upadhyay, D.; Sharma, V. Intracranial hemorrhage detection using neural network based methods with federated learning. *arXiv* **2020**, arXiv:2005.08644.

147. Malekzadeh, M.; Hasircioglu, B.; Mital, N.; Katarya, K.; Ozfatura, M.E.; Gündüz, D. Dopamine: Differentially private federated learning on medical data. *arXiv* **2021**, arXiv:2101.11693.

148. Kumar, R.; Khan, A.A.; Kumar, J.; Golilarz, N.A.; Zhang, S.; Ting, Y.; Zheng, C.; Wang, W.; et al. Blockchain-federated-learning and deep learning models for covid-19 detection using ct imaging. *IEEE Sens. J.* **2021**, *21*, 16301–16314. [CrossRef]

149. 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.

150. Zhang, W.; Zhou, T.; Lu, Q.; Wang, X.; Zhu, C.; Sun, H.; Wang, Z.; Lo, S.K.; Wang, F.Y. Dynamic-fusion-based federated learning for COVID-19 detection. *IEEE Internet Things J.* **2021**, *8*, 15884–15891. [CrossRef]

151. Daroudi, S.; Kazemipoor, H.; Najafi, E.; Fallah, M. The minimum latency in location routing fuzzy inventory problem for perishable multi-product materials. *Appl. Soft Comput.* **2021**, *110*, 107543. [CrossRef]

152. Dou, Q.; Thus, T.Y.; Jiang, M.; Liu, Q.; Vardhanabhuti, V.; Kaissis, G.; Li, Z.; Si, W.; Lee, H.H.; Yu, K.; et al. Federated deep learning for detecting COVID-19 lung abnormalities in CT: A privacy-preserving multinational validation study. *NPJ Digit. Med.* **2021**, *4*, 1–11. [CrossRef] [PubMed]

153. Yang, D.; Xu, Z.; Li, W.; Myronenko, A.; Roth, H.R.; Harmon, S.; Xu, S.; Turkbey, B.; Turkbey, E.; Wang, X.; et al. Federated semi-supervised learning for COVID region segmentation in chest CT using multi-national data from China, Italy, Japan. *Med Image Anal.* **2021**, *70*, 101992. [CrossRef] [PubMed]

154. Nguyen, D.C.; Pathirana, P.N.; Ding, M.; Seneviratne, A. BEdgeHealth: A decentralized architecture for edge-based IoMT networks using blockchain. *IEEE Internet Things J.* **2021**, *8*, 11743–11757. [CrossRef]

155. Xu, B.; Xia, W.; Zhang, J.; Quek, T.Q.; Zhu, H. Online client scheduling for fast federated learning. *IEEE Wirel. Commun. Lett.* **2021**, *10*, 1434–1438. [CrossRef]

156. Luo, S.; Chen, X.; Wu, Q.; Zhou, Z.; Yu, S. HFEL: Joint edge association and resource allocation for cost-efficient hierarchical federated edge learning. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 6535–6548. [CrossRef]

157. Yang, H.H.; Liu, Z.; Quek, T.Q.; Poor, H.V. Scheduling policies for federated learning in wireless networks. *IEEE Trans. Commun.* **2019**, *68*, 317–333. [CrossRef]

158. Bonawitz, K.; Ivanov, V.; Kreuter, B.; Marcedone, A.; McMahan, H.B.; Patel, S.; Ramage, D.; Segal, A.; Seth, K. Practical secure aggregation for federated learning on user-held data. *arXiv* **2016**, arXiv:1611.04482.

159. Thus, J.; Güler, B.; Avestimehr, A.S. Turbo-aggregate: Breaking the quadratic aggregation barrier in secure federated learning. *IEEE J. Sel. Areas Inf. Theory* **2021**, *2*, 479–489.

160. Yang, Q.; Fan, L.; Tong, R.; Lv, A. IEEE Federated Machine Learning. *IEEE Fed. Mach. Learn.* **2021**, 1–18. Available online: https://ieeexplore.ieee.org/document/9456823 (accessed on 30 October 2022).

161. Lu, S.; Zhang, Y.; Wang, Y. Decentralized federated learning for electronic health records. In Proceedings of the 2020 54th Annual Conference on Information Sciences and Systems (CISS), Princeton, NJ, USA, 18–20 March 2020; pp. 1–5.

162. Sarikaya, Y.; Ercetin, O. Motivating workers in federated learning: A stackelberg game perspective. *IEEE Netw. Lett.* **2019**, *2*, 23–27. [CrossRef]

163. Khan, L.U.; Pandey, S.R.; Tran, N.H.; Saad, W.; Han, Z.; Nguyen, M.N.; Hong, C.S. Federated learning for edge networks: Resource optimization and incentive mechanism. *IEEE Commun. Mag.* **2020**, *58*, 88–93. [CrossRef]

164. Xu, J.; Wang, H.; Chen, L. Bandwidth allocation for multiple federated learning services in wireless edge networks. *IEEE Trans. Wirel. Commun.* **2021**, *21*, 2534–2546. [CrossRef]

165. Kang, J.; Xiong, Z.; Niyato, D.; Xie, S.; Zhang, J. Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory. *IEEE Internet Things J.* **2019**, *6*, 10700–10714. [CrossRef]

166. Choquette-Choo, C.A.; Dullerud, N.; Dziedzic, A.; Zhang, Y.; Jha, S.; Papernot, N.; Wang, X. Capc learning: Confidential and private collaborative learning. *arXiv* **2021**, arXiv:2102.05188.

167. De Alwis, C.; Kalla, A.; Pham, Q.V.; Kumar, P.; Dev, K.; Hwang, W.J.; Liyanage, M. Survey on 6G frontiers: Trends, applications, requirements, technologies and future research. *IEEE Open J. Commun. Soc.* **2021**, *2*, 836–886. [CrossRef]

168. Mucchi, L.; Jayousi, S.; Caputo, S.; Paoletti, E.; Zoppi, P.; Geli, S.; Dioniso, P. How 6G technology can change the future wireless healthcare. In Proceedings of the 2020 2nd 6G Wireless Summit (6G SUMMIT), Online, 17 March 2020; pp. 1–6.

169. Wang, L.; An, H.; Zhu, H.; Liu, W. MobiKey: Mobility-based secret key generation in smart home. *IEEE Internet Things J.* **2020**, *7*, 7590–7600. [CrossRef]

170. Kerkouche, R.; Acs, G.; Castelluccia, C.; Genevès, P. Privacy-preserving and bandwidth-efficient federated learning: An application to in-hospital mortality prediction. In Proceedings of the Conference on Health, Inference, and Learning, Virtual Event, 8–10 April 2021; pp. 25–35.

171. Hu, R.; Guo, Y.; Li, H.; Pei, Q.; Gong, Y. Personalized federated learning with differential privacy. *IEEE Internet Things J.* **2020**, *7*, 9530–9539. [CrossRef]

172. Fereidooni, H.; Marchal, S.; Miettinen, M.; Mirhoseini, A.; Möllering, H.; Nguyen, T.D.; Rieger, P.; Sadeghi, A.R.; Schneider, T.; Yalame, H.; et al. SAFELearn: Secure aggregation for private federated learning. In Proceedings of the 2021 IEEE Security and Privacy Workshops (SPW), San Francisco, CA, USA, 27 May 2021; pp. 56–62.

173. Li, X.; Jiang, M.; Zhang, X.; Kamp, M.; Dou, Q. Fedbn: Federated learning on non-iid features via local batch normalization. *arXiv* **2021**, arXiv:2102.07623.

174. Li, Q.; Diao, Y.; Chen, Q.; He, B. Federated learning on non-iid data silos: An experimental study. *arXiv* **2021**, arXiv:2102.02079.

175. Jin, H.; Dai, X.; Xiao, J.; Li, B.; Li, H.; Zhang, Y. Cross-Cluster Federated Learning and Blockchain for Internet of Medical Things. *IEEE Internet Things J.* **2021**, *8*, 15776–15784. [CrossRef]

176. Gupta, R.; Shukla, A.; Mehta, P.; Bhattacharya, P.; Tanwar, S.; Tyagi, S.; Kumar, N. VAHAK: A Blockchain-based Outdoor Delivery Scheme using UAV for Healthcare 4.0 Services. In Proceedings of the IEEE INFOCOM 2020—IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Toronto, ON, Canada, 6–9 July 2020, pp. 255–260. [CrossRef]