Computer Modeling in
Engineering & Sciences

Tech Science Press

DOI: 10.32604/cmes.2024.048932

REVIEW

# A Comprehensive Survey on Federated Learning in the Healthcare Area: Concept and Applications

**Deepak Upreti[1], Eunmok Yang[2], Hyunil Kim[3,*] and Changho Seo[1,4]**

[1]Department of Convergence Science, Kongju National University, Chungcheongnam-do, Gongju-si, 32588, South Korea

[2]Department of Information Security, Cryptology, and Mathematics, Kookmin University, Seoul, 02707, South Korea

[3]Department of Information and Communication Engineering, Chosun University, Gwangju, 61452, South Korea

[4]Basic Science Research Institution, Kongju National University, Chungcheongnam-do, Gongju-si, 32588, South Korea

*Corresponding Author: Hyunil Kim. Email: hyunil@chosun.ac.kr

Received: 22 December 2023 Accepted: 09 May 2024 Published: 08 July 2024

**ABSTRACT**

Federated learning is an innovative machine learning technique that deals with centralized data storage issues while maintaining privacy and security. It involves constructing machine learning models using datasets spread across several data centers, including medical facilities, clinical research facilities, Internet of Things devices, and even mobile devices. The main goal of federated learning is to improve robust models that benefit from the collective knowledge of these disparate datasets without centralizing sensitive information, reducing the risk of data loss, privacy breaches, or data exposure. The application of federated learning in the healthcare industry holds significant promise due to the wealth of data generated from various sources, such as patient records, medical imaging, wearable devices, and clinical research surveys. This research conducts a systematic evaluation and highlights essential issues for the selection and implementation of federated learning approaches in healthcare. It evaluates the effectiveness of federated learning strategies in the field of healthcare. It offers a systematic analysis of federated learning in the healthcare domain, encompassing the evaluation metrics employed. In addition, this study highlights the increasing interest in federated learning applications in healthcare among scholars and provides foundations for further studies.

**KEYWORDS**

Federated learning; artificial intelligence; machine learning; privacy; healthcare

| **Acronyms** | **Definitions** |
|---|---|
| FL | Federated Learning |
| CFL | Clustered Federated Learning |
| IoMT | Internet of Medical Things |
| AI | Artificial Intelligence |
| DL | Deep Learning |
| IoT | Internet of Things |
| EHRs | Electronic Health Records |

| | |
|---|---|
| ECGs | Electro Cardiograms |
| CNN | Convolutional Neural Network |
| AUC | Area Under the Curve |
| ML | Machine Learning |
| NbAFL | Noising Before Model Aggregation Federated Learning |
| CXR | Chest X-Ray |
| SVM | Support Vector Machine |
| KNN | K-Nearest Neighbor |
| RF | Random Forest |
| NB | Naive Bayes |
| SSAE | Stacked Sparse Autoencoder |
| BRATS | Brain Tumor Segmentation |
| fMRI | Functional Magnetic Resonance Imaging |
| ADRs | Adverse Drug Reactions |
| HIPAA | Health Insurance Portability and Accountability Act |
| APPI | Act on the Protection of Personal Information |
| HITECH | Health Information Technology for Economic and Clinical Health Act |
| MLP | Multilayer Perceptron |

## 1 Introduction

Healthcare and related services are crucial in preventing illnesses, treating medical conditions, and promoting overall physical well-being. Healthcare providers have recently increasingly adopted technology to streamline various processes, such as data monitoring, patient registration, self-care applications, and lab testing. This technological integration empowers individuals to manage their physical or mental health conditions during planning.

The volume of digital health data has witnessed a significant surge in recent years, mainly due to the advent of the Internet of Medical Things (IoMT) revolution [1,2]. This technology enables sensing and transmitting an individual's health updates, which is invaluable in collecting comprehensive healthcare data. Hence, Artificial Intelligence (AI) is employed to analyze this vast amount of data, giving rise to various healthcare applications, such as remote patient monitoring and disease prognosis. Combining IoMT and AI has revolutionized the healthcare landscape, providing efficient and personalized healthcare solutions to improve patient outcomes [3–5].

The healthcare industry is characterized by generating and storing massive amounts of sensitive data related to patient's medical records, diagnoses, treatments, and other personal information [6]. Safeguarding these data and ensuring its management and security present considerable challenges for healthcare organizations. With the increasing awareness of data breaches and privacy concerns, users are becoming more cautious about how their personal information is handled and protected [7–9].

On the other hand, deep learning [10–12] is a subset of artificial intelligence [13,14] that has shown remarkable capabilities in analyzing vast datasets and extracting meaningful patterns and insights. DL algorithms can process raw data or historical patient records to create models that simulate real-time patient conditions. This ability makes DL highly valuable in various applications within the healthcare industry, including disease diagnosis, treatment planning, and personalized medicine. In addition, DL's potential extends beyond healthcare, as it is crucial in strengthening cybersecurity efforts to detect and prevent cyber threats, data breaches, and fraudulent activities in healthcare systems [15–18].

Over the last five years, federated learning [19] has emerged as a powerful and innovative approach in the field of machine learning, particularly in the context of privacy-preserving data analysis. FL operates on the principle of training machine learning models on decentralized data sources, such as individual hospitals, clinics, or patient devices, without directly sharing the raw data [20]. Instead, only the local models are updated and exchanged between the clients and a parameter server. This decentralized nature ensures that sensitive data remains localized and secure, mitigating the risks of data breaches or unauthorized access [20–23]. Federated learning consists of three main components: client devices, FL server, and communication protocol, as in Fig. 1.

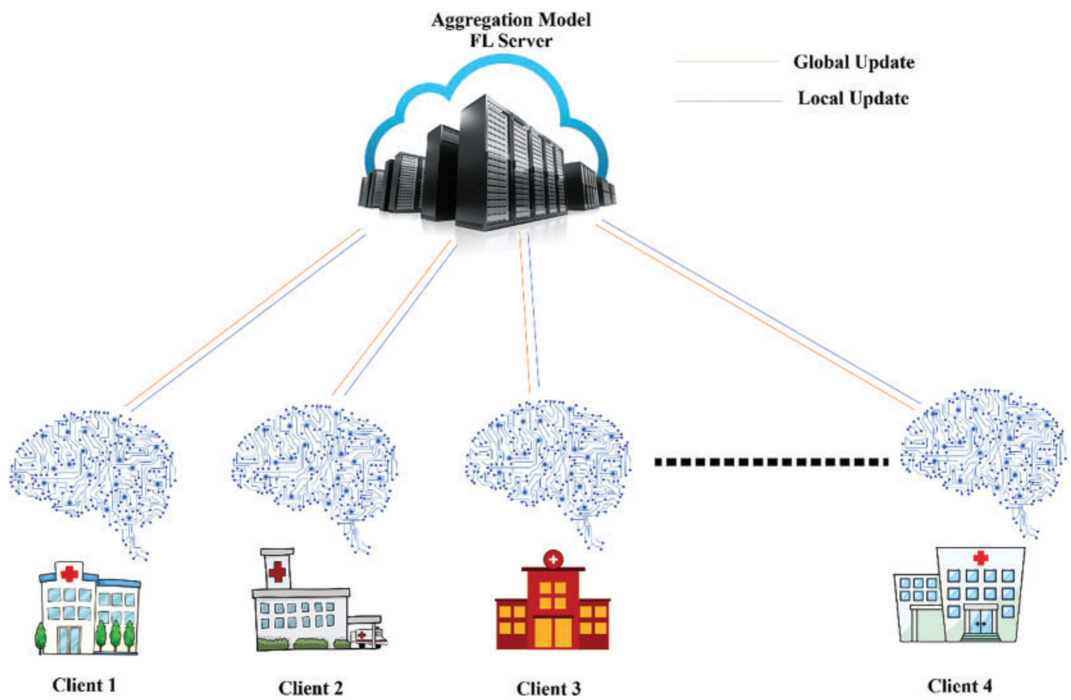


**Figure 1:** An overview of federated learning systems

1) Client devices: These are the edge devices that hold the data used for training the model [24]. Clients consider part of the learning process using their data to train local models and sharing updates with a central server.

2) Central server: The server is responsible for coordinating client learning. It compiles the model updates and gives them access to the new global model [25].

3) Communication protocol: The protocol controls how data is transferred between clients and servers. It defines how model updates are shared and aggregated and how the global model is allocated to clients.

In contrast, traditional smart healthcare systems have often relied on centralized AI architectures [26], where data from different sources are pooled into a central repository for training, as shown in Fig. 2. Although this approach can be effective, it has inherent drawbacks. Communication delays and data transmission inefficiencies can arise when dealing with large and distributed datasets, leading to slower model updates and potential privacy concerns [20–23,27]. In addition, centralization can hinder the system's scalability, as it can struggle to handle the ever-increasing volume of healthcare data generated [28].

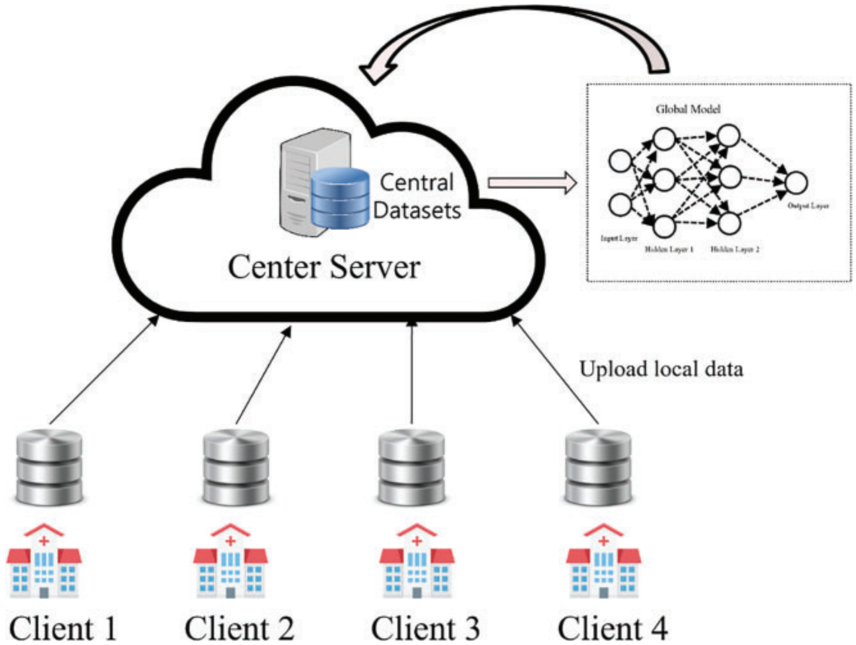


**Figure 2:** Centralized ML

Healthcare organizations can collaborate using federated learning to make better tools for helping patients and predicting diseases [29]. The big deal in this study is that they can do it without sharing private information. This is important as healthcare improves with data and computers, but they must keep information safe. A malicious node could upload a faulty model, disturbing the federated learning system. This is where blockchain technology arises, as it provides a ledger for secure and scalable solutions with blockchain; models can be decentralized without relying on a central server [30–33]. Blockchain also permits the secure collection of data models from several sources. This technology provides traceability, transparency, and immutability that can be combined with federated learning to improve privacy protection [7,8]. Blockchain-based federated learning has the potential to be a game-changer in healthcare. Integrating blockchain technology with federated learning has further strengthened the security and trustworthiness of FL systems [34]. Blockchain provides a tamper-resistant and decentralized ledger where the model updates can be recorded and verified, ensuring the integrity of the training process and preventing any malicious tampering with the data or model [35].

The adoption of blockchain-based federated learning in the healthcare industry has seen significant progress, especially in the Internet of Things (IoT) and smart healthcare applications [9,34,36]. Internet of Things devices collect real-time patient data, and federated learning enables them to build robust healthcare models while keeping patient data private within their respective systems.

There are some limitations, such as sensitive patient information, non-standardized data formats, low-quality data, incompatible systems, legal and regulatory constraints, data ownership issues, technological infrastructure, lack of standardization, patient consent and engagement, and Security concerns [37]. Healthcare systems can overcome these limitations and create a more efficient and secure infrastructure for data analysis and model development by adopting blockchain-based federated learning [38]. This shift in mindset allows healthcare institutions to take advantage of the collective knowledge of decentralized data sets while protecting patient information confidentially. As a result, it improves patient outcomes, diagnostic capabilities, and privacy in the healthcare sector [39].

The network edge needs to transition towards distributed AI technologies to establish intelligent healthcare systems that are both scalable and privacy-protecting [9,40]. In response to increasingly stringent data regulations worldwide, a machine learning approach known as federated learning has emerged, which operates on distributed data sets.

As a result, Federated learning offers numerous advantages compared to centralized learning [41]. Firstly, it enables training a global model using data distributed across multiple sources. Secondly, its primary focus is safeguarding data privacy by sharing only metadata and mathematical parameters, ensuring the underlying data remains as private as possible to prevent attacks and traceability. The long-term impacts of federated learning on healthcare research, treatment outcomes, and patient care are as follows: 1) Improved research collaboration can accelerate the discovery of novel treatments, biomarkers, and insights into various health conditions. 2) Patients can experience better treatment outcomes and improved quality of care. 3) Timely identification of health risks allows healthcare providers to implement preventive measures and early interventions, reducing the severity and progression of diseases. 4) Healthcare providers can benefit from more accurate diagnostic tools, leading to quicker and more reliable detection of cancer, cardiovascular disease, neurological illnesses, and other diseases. 5) Quicker and more cooperative drug discovery might lead to the development of novel medications and treatments for various diseases, expanding patients' healthcare alternatives.

In recent years, several research papers have been published on federated learning. However, there are still unresolved technical aspects related to its application in medicine and healthcare. Several recent surveys have highlighted the relevance of federated learning in these fields, with some focusing on areas such as the IoMT and electronic health records (EHRs) [39,42,43].

Nguyen et al. [20] investigated significant federated learning (FL) applications in smart healthcare, exploring sophisticated FL designs that benefit federated smart healthcare. They extensively addressed crucial applications of FL in intelligent healthcare, including the management of EHRs, remote health monitoring, medical imaging, and the detection of COVID-19 through federated approaches [44]. In addition, they investigated advanced FL designs that can potentially enhance federated smart healthcare. However, there is still a need for a comprehensive examination of FL's impact on medical imaging, IoT, and the management of COVID-19 outbreaks.

Artificial intelligence and deep learning (DL) have a significant advantage in that they have the potential for early disease detection and diagnosis [45]. AI and DL algorithms can analyze vast amounts of medical data, including images, genetic information, and patient records, to identify patterns and anomalies that indicate various diseases [2,36]. This early detection allows for timely intervention and personalized treatment plans, ultimately increasing the chances of successful outcomes and reducing healthcare costs associated with late-stage interventions [46].

In addition, AI and DL empower healthcare professionals by automating routine tasks, enabling them to focus on more complex and critical aspects of patient care [2,3]. These technologies, from predictive analytics for patient deterioration to personalized treatment recommendations based on individual patient profiles, enhance clinical decision-making, as in Fig. 3. In addition, AI applications facilitate the efficient management of healthcare data, leading to improved interoperability between different healthcare systems and the creation of comprehensive patient records. This enhances the continuity of care and supports medical research by providing large datasets for epidemiological studies and drug discovery. Integrating AI and DL in healthcare promises more accurate diagnostics, personalized treatments, and more efficient healthcare delivery [2,47].

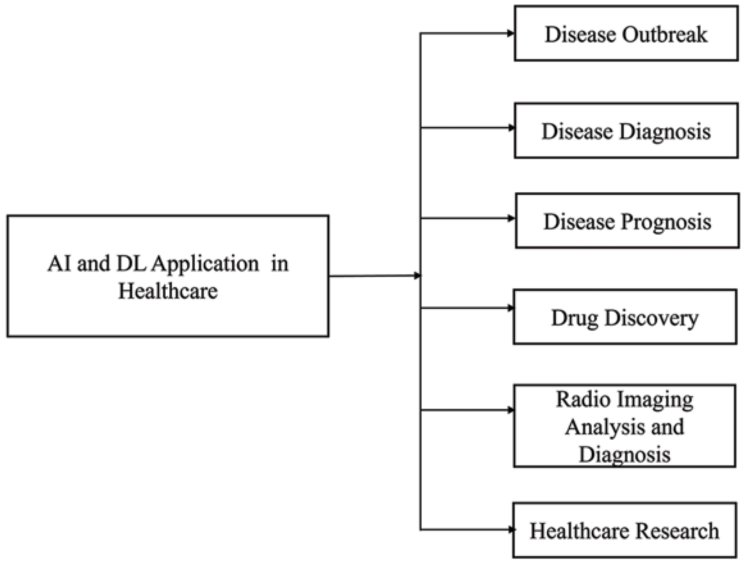


**Figure 3:** AI and DL applications in healthcare

This research provides an innovative overview of FL applications in critical healthcare domains, encompassing EHR maintenance, COVID outbreak management, IoT utilization, disease prediction, disease outbreak, drug discovery, healthcare research, and understanding of diseases, as shown in Fig. 2. The primary contributions of the study can be summarized in Fig. 4.

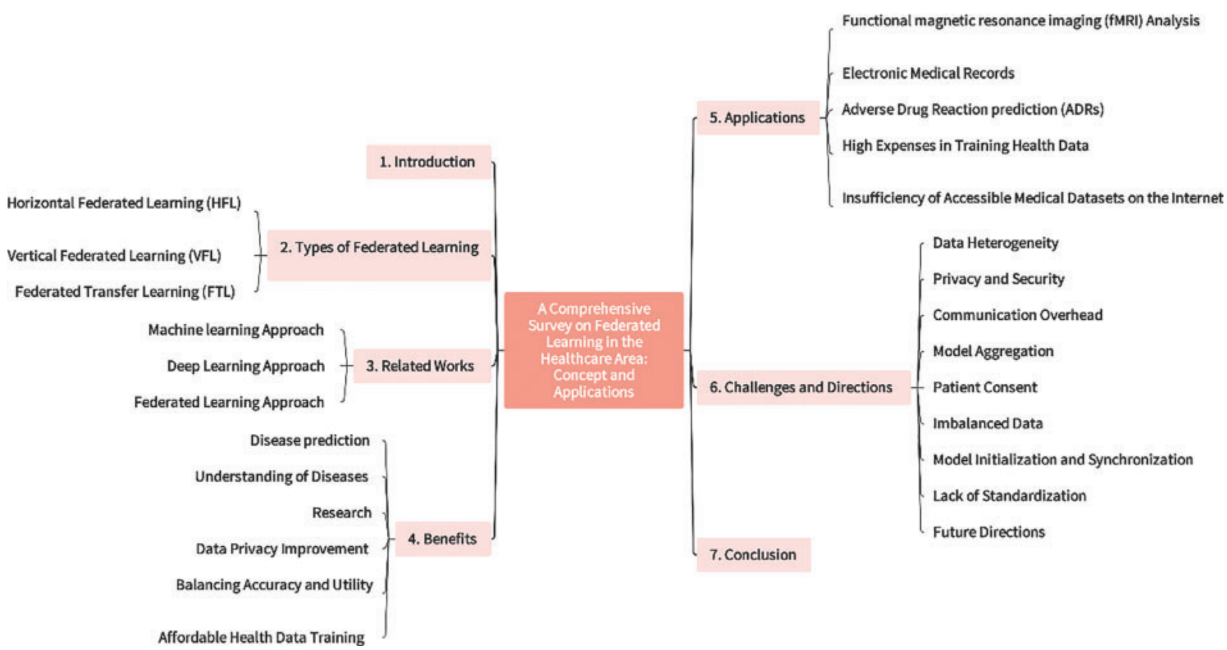


**Figure 4:** Primary structure of the article

1) This study begins with an introduction to the concept of federated learning and an exploration of its goals and the technological requirements for its use in smart healthcare.

2) Types of healthcare in federated learning are extended in the section.

3) In-depth exploration of related work in federated learning, machine learning, and deep learning are presented.

4) The detailed advantages of federated learning in the context of smart healthcare are initiated in this section. These include disease prediction, understanding of diseases, research, data privacy improvement, balancing accuracy and utility, and affordable health data training in federated learning.

5) Applications of federated learning in healthcare are explored.

6) The constraints of existing smart healthcare systems.

7) Lastly, the conclusion and address prospects in the scope of federated learning for smart healthcare are emphasized.

## 2 FL for Healthcare: Types

Each type of federated learning system offers unique advantages to different use cases, highlighting the versatility and potential of this decentralized machine learning approach. Federated learning systems come in various types, each designed to address specific challenges and requirements in decentralized machine learning scenarios. Types of FL are Horizontal Federated Learning, Vertical Federated Learning, and Federated Transfer Learning [48].

### *2.1 Horizontal Federated Learning (HFL)*

In health care, clients can collaboratively train a shared global model utilizing their respective datasets, as depicted in Fig. 5. Horizontal federated learning is characterized by sharing data horizontally, where multiple datasets have the same feature space but differ in samples [48–50].

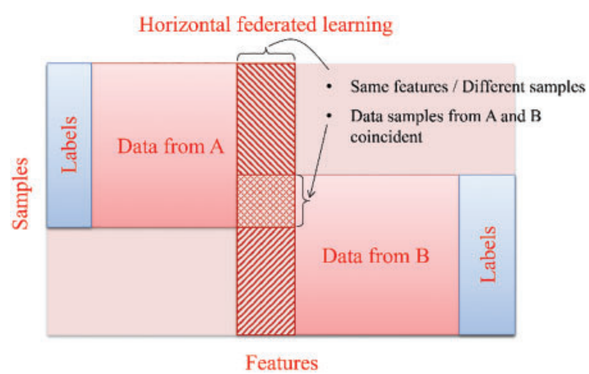


**Figure 5:** Horizontal federated learning

Every horizontal FL group possesses complete access to the entire set of features and labels, enabling them to train their local model using their dataset. The algorithm presented in [34] is an excellent illustration of a standard horizontal federated learning configuration.

### *2.2 Vertical Federated Learning (VFL)*

Vertical federated learning, also known as feature-based federated learning [20,48,50]. It comes into the picture when collaborating parties possess datasets with distinct features but share similar samples, as illustrated in Fig. 6. An instance of vertical federated learning within the Internet of Medical Things applications can be observed in a smart healthcare environment, where entities share a learning model for collaborative training. For example, historical medical records stored at hospitals and healthcare cost data at the insurance company can be utilized for intelligent healthcare decision-making.

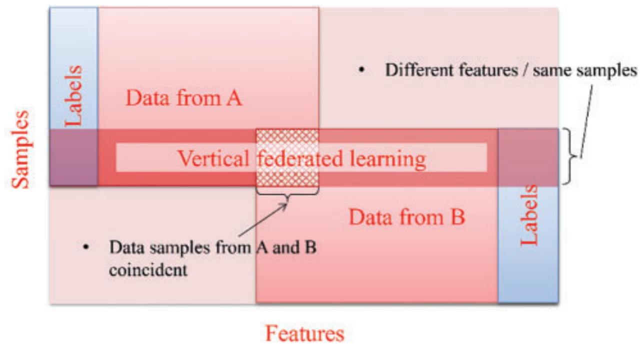


**Figure 6:** Vertical federated learning

### *2.3 Federated Transfer Learning (FTL)*

The introduction of federated transfer learning aims to address the difficulty of consolidating dispersed data and enhance statistical modeling when conducting data federation. Federated transfer learning operates independently of specific prerequisites such as a shared feature or common sample space. It facilitates transfer learning by offering solutions across the entire sample and feature space during the process of data federation, as shown in Fig. 7. FedHealth algorithm represents that notion of federated transfer learning within smart wearable healthcare devices [20,48,50]. The algorithm utilizes federated learning to aggregate data while employing transfer learning to construct personalized models, ensuring both model and data privacy and security are preserved [8,9].

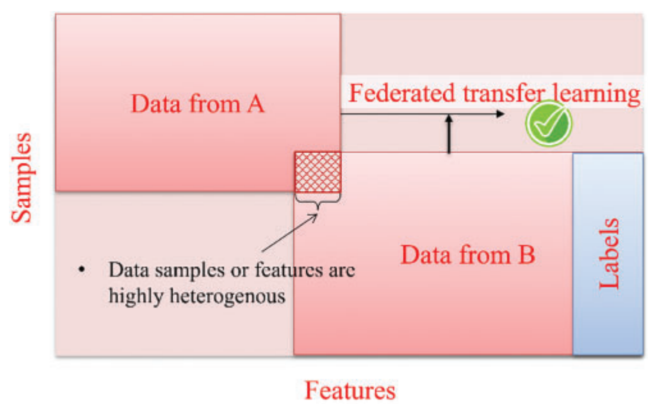


**Figure 7:** Federated transfer learning

## 3 Related Work

In this section, the literature review will examine research related to medical data utilizing machine learning, deep learning, and federated learning approaches. The review will specifically explore how federated learning is compared to other techniques, such as mechanical, deep, and distributed learning.

The primary focus is investigating various approaches involving machine, deep, and federated learning methodologies.

### *3.1 Machine Learning Approach*

"Machine learning" is a branch of artificial intelligence that focuses on creating statistical models and algorithms that enable computers to learn and make decisions without explicit programming. Machine learning techniques utilize algorithms trained on datasets to generate models capable of tasks such as image categorization, data analysis, trend prediction, and language translation, as depicted in Fig. 8. In addition, technologies based on machine learning have been employed to boost effectiveness. For example, classify Electroencephalogram data and detect seizures [51,52], blood glucose pattern classification and anomalies detection [53], prediction of distant metastasis from soft-tissue sarcomas [54], and robot-assisted surgeries [55,56]. Machine learning is an effective method for spotting patterns in medical images, but it must be utilized carefully because it can be abused if its advantages and disadvantages are not recognized [57]. Table 1 summarizes Machine learning in the survey with some recent methods.

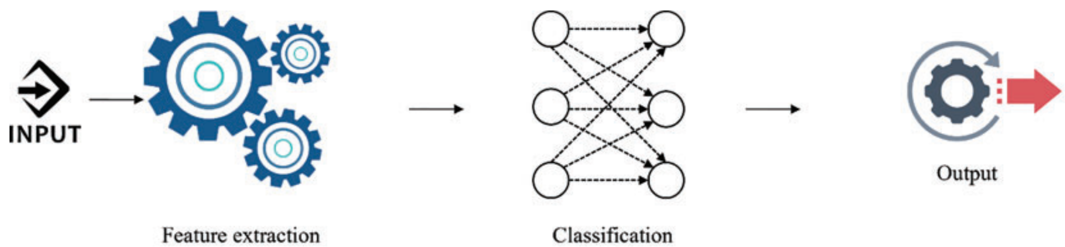


**Figure 8:** Machine learning approach

**Table 1:** Summary of related work (machine learning)

| Related works | Authors | Methods |
|---|---|---|
| [58] | Mattfeldt et al. | Artificial neural networks in machine learning |
| [59] | Kundu et al. | Machine learning, weka package's random forest, and gaussian process regression techniques |
| [60] | Bahrami et al. | An in-depth examination of ML and DL algorithms for single-lead Electrocardiogram sleep disorder screening |
| [61] | Bhatt et al. | Machine learning, k-mode clustering |
| [62] | Hu et al. | Convolutional neural networks |
| [63] | Khan et al. | An automated method for identifying brain tumors using three publicly available, unrestricted datasets |
| [64] | Ogundepo et al. | Machine learning technique |
| [65] | Jagadeesha et al. | Neural network, linear regression and SVM |

Mattfeldt et al. [58] compared prostatic cancer and postoperative tumor progression. In this survey, two applications of learning vector quantization and linear discriminant analysis were used in multi-layer feedforward networks with backpropagation. The learning vector quantization networks, linear discriminant analysis, and multi-layer feedforward networks with backpropagation networks produced the best results.

Kundu et al. [59] attempted to apply Machine Learning (ML) algorithm models on the protein-ligand binding affinity data that was already available. They applied the Weka package's Random Forest and Gaussian process regression techniques to analyze the protein and ligand binding information in the Protein data bank (Pdb) Bind database. Correlation coefficient R2 and root mean square error were employed to evaluate the models.

Bahrami et al. [60] focused on deep convolutional neural networks (CNN) such as visual geometry group (VGG16), Zeiler and Fergus networks (ZF-Net) and AlexNet, recurrent network, and hybrid deep neural networks (DNN). This study provides valuable information for sleep research on designing and selecting appropriate machine learning and deep learning algorithms for detecting sleep apnea.

Bhatt et al. [61] conducted a survey that classifies cardiovascular disease occurrence using machine learning, which can assist doctors in reducing misdiagnosis. They employed a k-mode clustering algorithm with Huang as the starting point to increase classification accuracy. Models were trained on data divided 80:20 and achieved accuracy even with and without cross-validation.

Hu et al. [62] proposed model achieves an accuracy rate of 96% on the training set, 93% on the validation set, and 82% on the test set. For Natural language processing (NLP), the chatbot can manage basic conversations such as simple greetings, requests for input images, and basic suggestions based on prediction results.

Khan et al. [63] used a pre-trained CNN model, EfficientNetB0, which is adjusted and trained in two ways: on improved and tumor localization images. The most beneficial features were selected in the last stage with an updated dragonfly optimization method. Finally, the most significant characteristics were identified using extreme machine learning. The study was performed on three publicly available datasets and resulted in enhanced accuracy.

Ogundepo et al. [64] employed the Cleveland dataset, which was analyzed using the Chi-square test of independence, and then ten conventional classification models were trained for class predictions.

Jagadeesha et al. [65] trained new skin type classification models using reconstructed Hyperspectral (HS) images of skin and tested them on a clinical dataset. The proposed models outperform RGB based on image models.

### *3.2 Deep Learning Approach*

In the deep learning methodology, computers are taught to acquire knowledge from inputs and establish sophisticated connections among data utilizing neural networks. CNNs, Recurrent neural networks (RNNs), and transformer architectures illustrate complex artificial neural networks utilized in deep learning frameworks. In deep learning, the term "deep" denotes the abundance of concealed layers within the neural network, with some models comprising hundreds or even thousands of such layers. Training these algorithms and extracting features and patterns directly from data necessitates extensive amounts of labeled data during the training phase, as shown in Fig. 9. Table 2 lists deep learning in the survey with some recent methods.

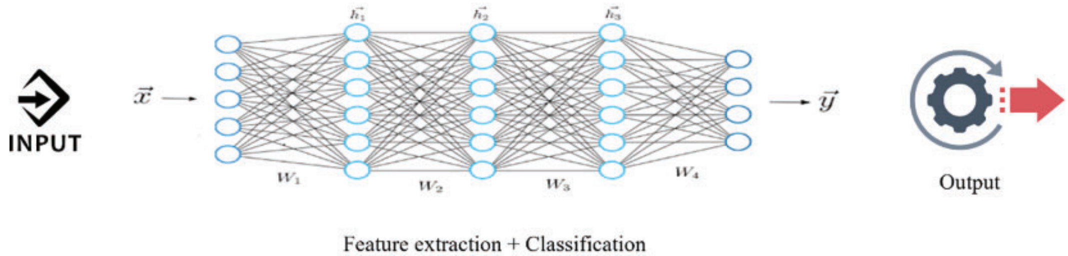


**Figure 9:** Deep learning approach

**Table 2:** Summary of related work (deep Learning)

| Related works | Authors | Methods |
|---|---|---|
| [66] | Sharma et al. | VGG16 |
| [67] | Jain et al. | CNN and deep learning techniques |
| [68] | Amin et al. | Stacked sparse autoencoder (SSAE) |
| [69] | Rashid et al. | MobileNetV2 |
| [70] | Balaha et al. | VGG16, VGG19, MobileNet, MobileNetV2, MobileNetV3Large, MobileNetV3Small, NASNetMobile, and NASNetLarge |

Sharma et al. [66] conducted a survey where they employed VGG16, a CNN model commonly used for image recognition, to detect and classify pneumonia using two chest x-ray (CXR) image datasets, as shown in Fig. 10. In the first dataset, they achieved an accuracy of 92.15%, a recall of 0.9308, a precision of 0.9428, and an F1 score of 0.937. In the second dataset, the accuracy, recall, precision, and F1 score were reported as 95.4%, 0.954, 0.954, and 0.954, respectively. The survey results indicated that VGG16 with neural network (NN) outperforms VGG16 with Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Random Forest (RF), and Naive Bayes (NB) in terms of performance for both datasets.

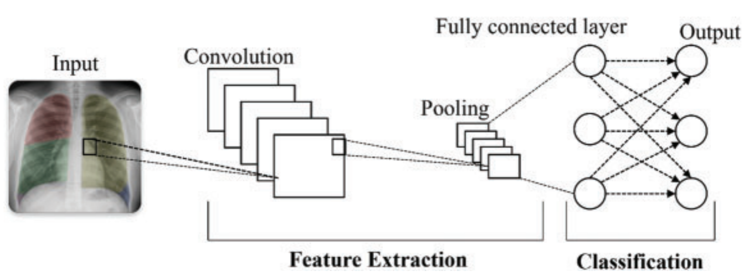


**Figure 10:** Convolutional neural network architectures

Jain et al. [67] conducted a survey using CNN and deep learning techniques to identify COVID-19-induced pneumonia in chest X-rays. They utilized transfer learning with Inception V3, Xception, and ResNeXt models and examined their accuracy. The experimental results demonstrated promising outcomes, and the Xception model achieved an accuracy of 98% in detecting COVID-19-induced pneumonia. The evaluation metrics also showed positive performance, including precision, recall, F1 score, specificity, erroneous omission rate, false negative rate, false positive rate, and false discovery rate. The research accurately identified COVID-19 exposure using experimental data and differentiated between COVID-19-induced pneumonia and regular pneumonia. Although both are common conditions, COVID-19 presents a more significant risk in terms of its severity and fatality.

Amin et al. [68] utilized a high-pass filter image to emphasize homogeneities in Magnetic resonance imaging (MRI) slices and then merge it with the input slices. The fused slices underwent a median filter to smooth and enhance the edges, improving slice quality. Hence, a 4-connected seed growth technique was applied based on the slice intensities, and a suitable threshold was employed to cluster related pixels from the input segments. Enhanced segments were fed into a refined two-layer stacked sparse autoencoder (SSAE) model for segmentation. The model's hyperparameters were thoroughly chosen through complete experimentation, featuring 200 hidden units in the initial layer and 400 in the subsequent layer. The SoftMax layer was employed for image prediction, distinguishing between images with and without tumors. It was trained and tested on various datasets, including Brain Tumor Segmentation (BRATS), to assess the model's performance. Multiple performance metrics were employed to evaluate the proposed model, and the results demonstrated superior performance compared to other approaches.

Rashid et al. [69] found that the likelihood of a patient surviving can be improved with an early and accurate diagnosis. They classified melanoma using MobileNetV2. Using the ISIC 2020 dataset, the proposed deep learning model's effectiveness was assessed. The experimental findings showed that the proposed deep learning technique performs better than cutting-edge deep learning algorithms in terms of accuracy and computing cost.

Balaha et al. [70] proposed a threshold-based automatic approach for skin cancer detection, classification, and segmentation utilizing a meta-heuristic optimizer named the sparrow search algorithm (SpaSA). Five U-Net models with different configurations and SpaSA optimizers were employed to optimize the hyperparameters using eight pre-trained CNN models. For the "skin diseases image" dataset, the best overall accuracy from the applied CNN experiments was 85.87% by the MobileNetV2 pre-trained model [70].

### *3.3 Federated Learning Approach*

The federated learning approach trains AI models without sharing raw data, in which different IoT devices contribute their data to train a centralized model. Each IoT device downloads the model, trains it on its data, and encrypts the updates, which are then sent back, decrypted, averaged, and integrated into the central model [71], as shown in Figs. 1 and 11. Federated learning has been gaining popularity due to its potential to address privacy concerns while enabling efficient and collaborative model training across decentralized data sources. Its appeal lies in its ability to train machine learning models without centralizing sensitive data, thereby preserving user privacy. This approach has gained interest across various sectors, including healthcare. Based on the graphs above, the number of articles published with the dimensions indexed federated learning in healthcare is rising, as shown in Fig. 12.

Pneumonia detection has traditionally been performed by trained individuals, including doctors and other healthcare experts [72]. The research conducted by Kareem et al. [73] aimed to utilize real-time datasets in a privacy-preserving manner. This was achieved by the federated learning framework. They conducted experiments using various CNN models, including Alexnet, DenseNet, Residual Neural Network-50 (ResNet50), Inception, Visual Geometry Group-19 (VGG19), and other advanced ML models for medical image classification. The evaluation metrics, including the Area Under the Curve (AUC), were employed to compare the results. Preliminary findings indicated that ResNet-50 exhibits notably high performance on the testing dataset, achieving an accuracy of 93%.

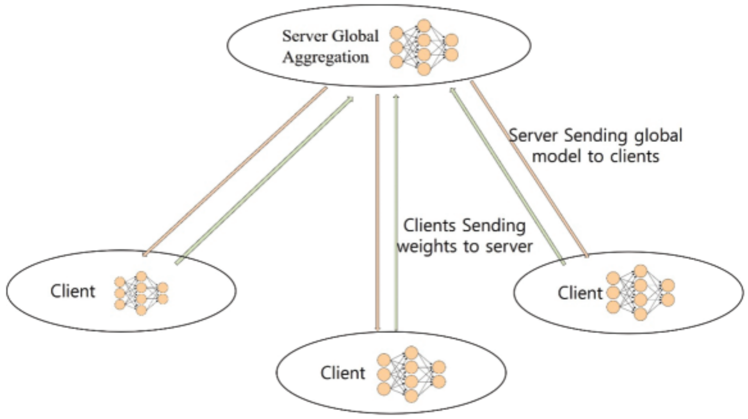


**Figure 11:** Federated learning approach

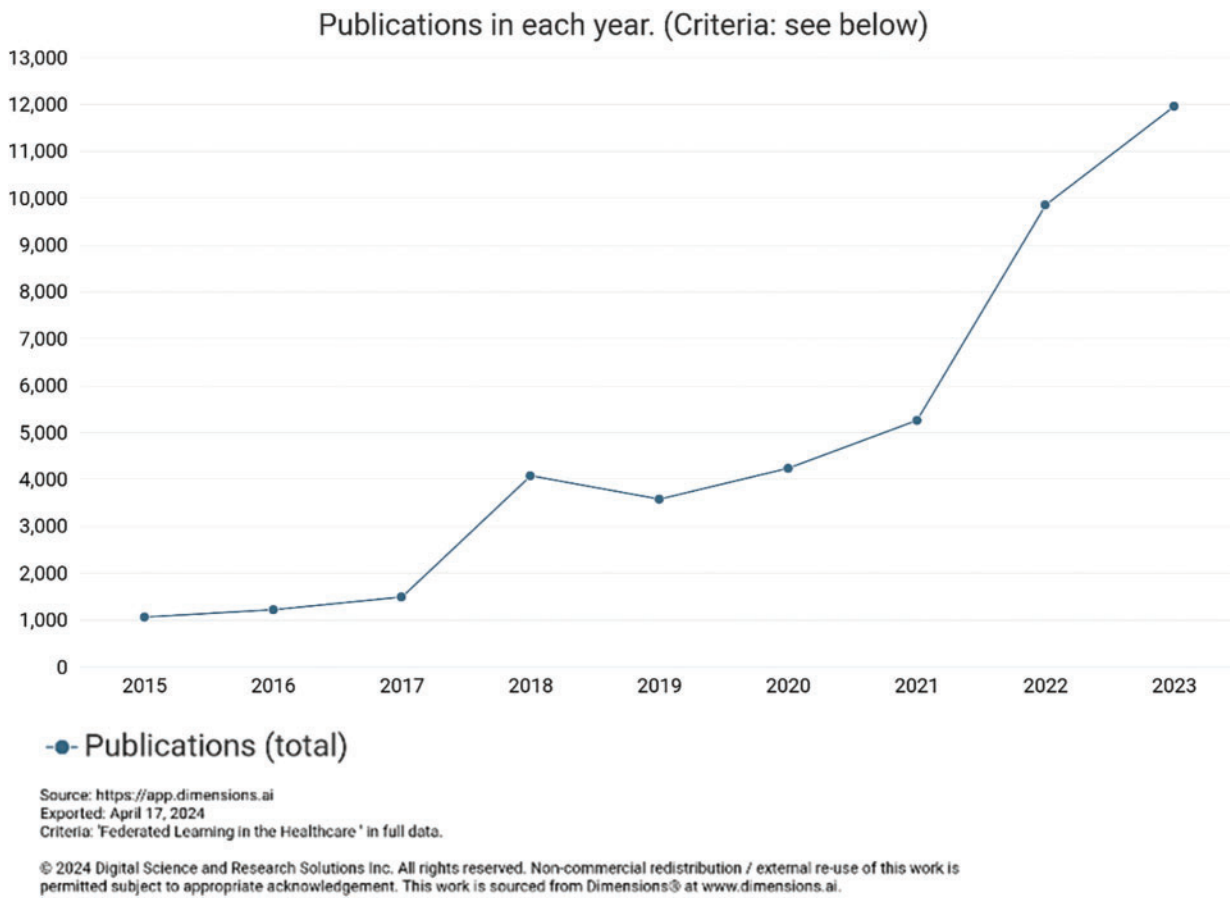


**Figure 12:** Trends of publication

Khan et al. [74] proposed a diagnostic system for pneumonia using a combination of a deep learning model and federated learning. The system achieved an accuracy of approximately 90% in pneumonia diagnosis. The approach involved training local models with patient data from each hospital and then aggregating the knowledge to create a central model while preserving individual patient data's privacy [75].

Qayyum et al. [76] proposed a collaborative learning approach for COVID-19 diagnosis, utilizing the concept of clustered federated learning. They aimed to create an automated diagnostic system for COVID-19, which can help to improve the burden on healthcare systems worldwide that have been under immense strain since the emergence of the COVID-19 pandemic in late 2019. They conducted experiments on two benchmark datasets, chest X-ray [77] and chest ultrasound images [78], to evaluate the effectiveness of their proposed framework. In their survey, they trained specialized models for different types of COVID-19 imagery using centralized data. The results obtained from this clustered federated learning approach show promising outcomes on both datasets, with performance comparable to the traditional centralized baseline. Specifically, the proposed method achieved 16% and 11% improvements in overall F1 scores over the traditional federated learning setup for the X-ray and ultrasound datasets.

The success of this cooperative learning model for COVID-19 diagnosis indicated its potential to contribute to more efficient and accurate automated diagnostic systems. Healthcare institutions can collaborate and collectively train specialized models on their distributed data without sharing raw data using the power of clustered federated learning. This enhances privacy and security and has the potential to relieve the strain on healthcare systems by providing automated and reliable diagnostic support in the ongoing battle against COVID-19.

The Wei et al.'s [79] survey introduced a new framework called noising before model aggregation federated learning based on differential privacy. This approach involved adding artificial noise to the clients' side parameters before aggregating them, effectively preventing information leakage. The survey demonstrated that NbAFL can achieve differential privacy at different protection levels by appropriately adjusting the artificial noise variations. In addition, they established a theoretical convergence bound for the loss function of the FL model trained with NbAFL. They proposed a K-client random scheduling technique to enhance the aggregation process, randomly selecting K (where $1 \leq K < N$) clients from a total of N clients for each aggregation step. They identified an optimal value of K that provides the best convergence performance for a given privacy level.

The survey conducted by Yi et al. [80] employed a federated learning model called "SU-Net," based on the U-Net architecture, to perform brain tumor segmentation. The "SU-Net" model incorporated the advantage inception module and a dense block to extract features at multiple scales and efficiently reuse data from earlier layers, enhancing gradient flows and information transfer. The proposed "SU-Net" model achieved an impressive accuracy of 99.7%, surpassing both the traditional U-Net model and the DeepLabv3+ model in the semantic segmentation of medical images.

Kumar et al. [34] proposed a survey detecting COVID-19 patients. They proposed a blockchain-based federated learning method to train a global deep learning model utilizing a small quantity of data collected from diverse sources (hospitals). As the data were collected from various hospitals with various types of Computed Tomography (CT) scanners, they first proposed a data normalization technique that deals with the heterogeneity of the data. Secondly, they utilized Capsule Network-based segmentation and classification to find COVID-19 patients. Thirdly, they developed a technique using blockchain technology and federated learning to jointly train a global model while main-training anonymity.

The primary objective of the Heidari et al.'s [81] survey is to classify lung cancer based on severity and identify malignant lung nodules in a CT scan of the lungs. Modern deep learning techniques were employed to locate the malignant nodules. They trained a global DL model using a little data from many hospitals and block-chain-based federated learning. Blockchain technology was employed to authenticate the data, and FL trained the model globally while protecting the privacy of the enterprise. Additionally, lung cancer patients were categorized in local mode using the CapsNets approach. The Kaggle Data Science Bowl (KDSB), Lung nodule analysis 2016 (LUNA 16), the local dataset, and the Cancer Imaging Archive (CIA) datasets were utilized to test and train the learned algorithm. The results demonstrated that the methodology accurately identifies lung cancer patients with 99.69% accuracy and minimal potential category error.

Lincy et al. [82] introduced a strategy for early Type-2 Diabetes prediction, incorporating feature selection techniques alongside a federated MLP method. They compared centralized machine learning and decentralized, federated learning models. The experiments conducted on the widely used Pima Indian Diabetes dataset reveal that the decentralized, federated architecture prioritizes privacy by keeping data on individual client devices accessible only to them. Although this ensures data privacy, the downside is a reduction in accuracy due to the prolonged time required to gather separately stored data on the server. In summary, organizations emphasizing privacy can select a federated model, whereas those prioritizing accuracy can lean towards a centralized model.

Astillo et al. [83] provided a lightweight DL-based anomaly detection model designed for the Diabetes Management Control System. The classification model was designed to identify abnormal data points to compare CNN with MLP algorithms. They employed independent learning (IL) and FL techniques to preserve user data privacy. In addition, to get around DL's computationally taxing operations, the post-quantization compression technique was employed to make models into lightweight versions. The FL approach had a greater recall rate ($\geq$98.69%) than the IL method ($\leq$97.87%), based on the experiment's findings. In addition, the FL-supported CNN-based approach to anomaly detection outperforms the MLP-based method in terms of performance. The recall rate for the former was an average of 99.24%, whereas the recall rate for the latter was only 98.69%. When the original models were transformed into their lightweight version, the inference latency of the models drastically decreased from more than 300 ms to less than two milliseconds without sacrificing recall value.

The research by Hossen et al. [84] on Skin Disease Detection utilizing CNN for classification and a federated learning strategy for data privacy preservation exhibited notable results in medical imaging. In this survey, a custom image dataset with four categories of skin diseases was created, a CNN model was proposed and evaluated against three benchmark CNN algorithms, and a federated learning experiment was run to test how well data privacy can be maintained. An image augmentation approach was applied for acne, eczema, and psoriasis to increase the dataset and broaden the model. The proposed model attained precision of 86%, 43%, and 60%, and recall of 67%, 60%, and 60%. The model in the federated learning strategy had an average accuracy of 81.21%, 86.57%, 91.15%, and 94.15% when the dataset was distributed among 1000, 1500, 2000, and 2500 clients. The CNN-based skin disease classification combined with the federated learning approach is a breathtaking idea for categorizing human skin illnesses while maintaining data security.

Jiménez-Sánchez et al. [85] conducted a study utilizing a collaborative, decentralized model incorporating three clinical datasets from various vendors. They demonstrated that they can strategically schedule training samples by overseeing local and global classification predictions. This scheduling effectively enhanced the alignment between domain pairs, improving classification performance.

Giuseppi et al. [86] introduced a federated learning approach named Federated learning consensus (FedLCon), developed based on discrete-time average consensus theory findings. FedLCon is employed to decentralize two existing FL algorithms: Federated averaging (FedAvg) and Adaptive federated learning (AdaFed). The decentralization of FL algorithms through FedLCon enhances their implementation across federations with sparse communication graphs, improving privacy-related capabilities. The study detailed the outcomes achieved by the evaluated algorithms in various scenarios, specifically in the context of a COVID-19 detection task.

Balaji et al. [87] demonstrated that the superiority of aggregation methods such as Coordinate-wise median (COMED) and Geometric median (GEOMED) over FedAvg is apparent, with a 30% increase in accuracy due to their resilience to corrupted models. The study investigated the tradeoff between achieving the optimal model and the communication overhead associated with the frequent transmission of model weights from individual nodes to the global server. Employing higher local epochs with a single aggregation method resulted in a marginal decline in model performance. Examining various shuffling mechanisms, considering factors such as latency and the number of combinations, revealed that layer shuffling is the most cost-effective option. This indicates that federated learning and robust cryptographic measures make it feasible to enhance model performance without compromising privacy.

To enhance the updating process of the federated model [88], Mixed supervised federated learning (FedMix) employs an adaptive aggregation function that adjusts client weights based on both the quantity and quality of their data. Experimental findings on two segmentation tasks illustrated the efficiency of FedMix in learning from diverse supervisions, offering valuable insights into improving the annotation burden on medical experts. In the semi-supervised federated scenario, FedMix surpasses the state-of-the-art approach of federated style transfer learning (FedST) [89]. Compared to FedAvg [19], the suggested adaptive aggregation function consistently enhances performance on the two tasks in the fully supervised setting. Wicaksana et al. [88] suggested that the techniques introduced in FedMix hold broad applicability in federated learning for medical image analysis, particularly in scenarios involving mixed supervisions.

Hansen et al. [90] created a stratified Cox regression model using data from hospitals in three countries, ensuring that patient-specific information remained within the hospital premises to avoid any data leakage risks. The key factors influencing the survival model are tumor volume and performance status.

Cetinkaya et al. [91] addressed the classification of diverse chest diseases using a proposed convolutional neural network architecture applied to chest X-ray images. In addition, a novel dataset comprising 28,833 CXR images, including cases of COVID-19, non-Covid viral or bacterial pneumonia, lung opacity, and normal cases, is introduced by combining publicly available datasets for federated learning in the proposed models. The initial training of the proposed network is centralized using the integrated dataset, followed by federated training involving 20 institutions. Compression techniques are subsequently employed to reduce the model size, improving communication efficiency in federated training. Experimental comparisons are made among central training, federated training, and communication-efficient federated training. The results demonstrated that federated training for chest disease classification achieves an accuracy of 92.96%, presenting a viable alternative to central training, which attains 93.34% accuracy in this context. Implementing model compression on the client side reduces the communication cost of FL by ten times, and the model trained with communication efficiency methods attains an accuracy of 92.44%. The study's findings can motivate medical organizations to initiate or adopt FL methodologies in their research. Table 3 summarizes federated learning in the survey with these recent methods.

**Table 3:** Summary of related work (federated learning)

| Related works | Authors | Methods | Dataset |
|---|---|---|---|
| [73] | Kareem et al. | Federated learning | Real-time dataset from hospitals and medical institutions |
| [74] | Khan et al. | Deep learning model and a federated learning technique | X-ray images |
| [76] | Qayyum et al. | Clustered federated learning (CFL) | X-ray and ultrasound datasets |
| [79] | Wei et al. | Differential privacy (DP) | Real-world federated database |
| [80] | Yi et al. | SU-Net for brain tumor segmentation | LGG (Low-Grade Glioma) Segmentation dataset "Brain Magnetic resonance imaging (MRI) Segmentation" in Kaggle |
| [34] | Kumar et al. | Blockchain-based federated learning | CC-19, related to the latest family of coronavirus |
| [81] | Heidari et al. | Modern Deep Learning techniques and blockchain-based Federated Learning | Cancer Imaging Archive (CIA) dataset, Kaggle Data Science Bowl (KDSB), LUNA 16, and the local dataset |
| [82] | Lincy et al. | Centralized machine learning model and decentralized federated learning model | Type-2 diabetes dataset |
| [83] | Astillo et al. | Convolutional neural network and multilayer perceptron (MLP) | Locally stored datasets |
| [84] | Hossen et al. | CNN algorithms and a federated learning | Skin disease datasets |
| [85] | Jiménez Sánchez et al. | Federated learning implementation | 3 datasets of Full Field Digital Mammography (FFDM), coming from three different private and public vendors |
| [86] | Giuseppi et al. | FedLCon has been employed | COVID-19, Pneumonia and Normal Chest X-ray Posteroanterior (PA) Dataset |
| [87] | Balaji et al. | Aggregation techniques FedAvg, COMED and GEOMED have been implemented for predicting pneumonia | Pneumonia dataset |

(Continued)

**Table 3 (continued)**

| Related works | Authors | Methods | Dataset |
|---|---|---|---|
| [88] | Wicaksana et al. | FedMix have been implemented | Three public breast ultrasound datasets, namely Breast ultrasound (BUS) [37], Breast ultrasound image segmentation (BUSIS) [38], and UDIAT-Centre diagnostic [39], are used, Skin tumor segmentation. Human against machine with 10000 training images (HAM10K) |
| [90] | Hansen et al. | Federated learning system | Data were obtained from 1821 larynx cancer patient |
| [91] | Cetinkaya | Adopted federated learning approach | Covid chest X-ray, COVID-19 Radiography, Covid chest X-ray repository, Covid chest X-ray dataset |
| [51] | Sakib et al. | Investigated two federated learning | Four different publicly available datasets, Massachusetts institute of technology-beth israel hospital (MIT-BIH) Supraventricular Arrhythmia database MIT-BIH Arrhythmia database, the Institute of cardiological technics (INCART) 12-lead arrhythmia database, and the Sudden Cardiac death holter database |

Sakib et al. [51] suggested an asynchronously updating Federated Learning (Async-FL) architecture designed for mobile and deployable Ultra Edge Nodes (UENs). The goal is to establish decentralized and collaborative arrhythmia detection without the necessity of directly exchanging Electrocardiogram (ECG) data with the cloud. In addition, their proposal is expected to reduce network overhead, including lower bandwidth consumption, time, and memory requirements, as the number of UENs increases. Given the rising demand for remote patient monitoring, particularly in events like the novel coronavirus pandemic. Applying Async-FL in this ECG monitoring scenario can pave the way for realizing a future-generation smart and widespread remote health monitoring system.

## 4 Benefits of Federated Learning in Smart Healthcare

A critical component of medical research and healthcare that uses cutting-edge technologies and data analysis is the prediction of people's susceptibility to certain illnesses. Medical professionals can identify patterns and risk factors associated with various diseases, as shown in Fig. 13, by combining genetic information, lifestyle data, clinical records, and machine learning algorithms. This proactive approach improves patient outcomes and lessens the burden on healthcare systems by facilitating early intervention, individualized treatment plans, and preventative measures. Developing illness prediction models integrating big data and artificial intelligence strengthens the ability to anticipate and mitigate health risks. This leads to a more efficient and targeted approach to healthcare management.

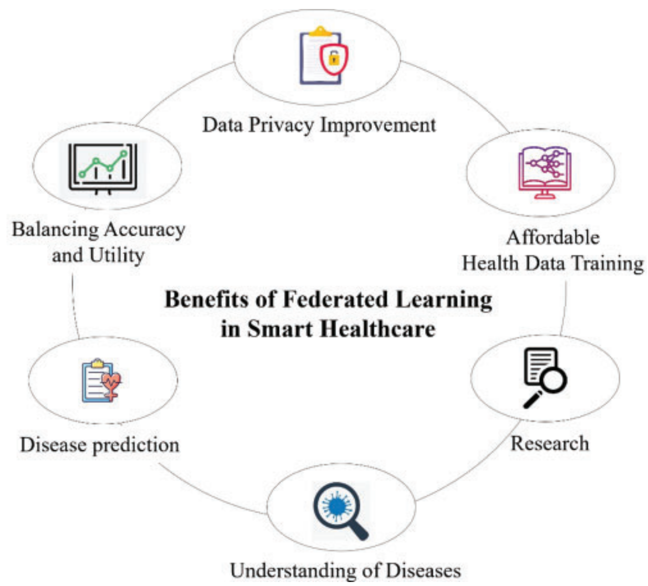


**Figure 13:** Benefits of federated learning in smart healthcare

### *4.1 Disease Prediction*

Federated learning is a revolutionary technology in Smart Healthcare, mainly when predicting a person's susceptibility to certain diseases. FL is a novel medical research and healthcare technique where data analysis and state-of-the-art technology are essential. Hospitals can utilize federated learning models to assess the probability of a particular patient developing an illness or contracting an infection upon admission. A survey by Moshawrab et al. [43] thoroughly examined federated learning, providing a theoretical explanation and comparing it to other technologies. Disease prediction also highlights the application of federated learning in diagnosing cancer [92], diabetes [82], and cardiovascular diseases [93].

### *4.2 Understanding of Diseases*

The application of federated learning in healthcare is a promising avenue for accelerating medical research and improving patient outcomes, particularly in the context of understanding the genetic complexities underlying diseases such as cancer and Alzheimer's disease [94]. Federated learning offers a transformative approach to unlocking valuable insights and propelling the future of precision medicine by harnessing the power of distributed data while prioritizing data privacy.

Federated learning enables healthcare providers to collaborate and exchange knowledge and insights while ensuring the privacy and security of their data. FL is utilized by both healthcare providers, who can use it to develop prediction models for chronic illness risks using electronic health records (EHR) [95], and consumers (patients), who can benefit from FL in tasks such as cardiac medical examinations using wearable devices with electro cardiograms (ECGs) [74,76,96].

### *4.3 Research*

Recently, an increasing number of researchers have directed their attention toward addressing the challenges associated with federated learning and exploring more effective research approaches. The future development of FL will primarily revolve around finding solutions to overcome the current bottleneck issues [97–99]. Subsequent research in the field of FL will focus on devising mechanisms to ensure privacy and security, exploring client cooperative training methods, and developing fair, robust, and personalized federated learning techniques. These advancements aim to enable the widespread deployment and application of FL technology in various domains, facilitating in-depth investigations and research opportunities [75].

### *4.4 Data Privacy Improvement*

Due to the requirement to develop more precise and privacy-preserving ML/DL models, federated learning is receiving much attention in the healthcare industry. Various privacy-preserving methods must be applied to further enhance privacy preservation in FL. Data-filtering [100], sanitization [101], adversarial training [102], robust aggregation [103], homomorphic encryption [104], safe multiparty computation [105], and differential privacy [106] are the techniques that are most frequently utilized in FL to maintain privacy [107–109]. Differential privacy is frequently employed in real-time applications since it is scalable and has less overhead [79,110]. A formal framework called differential privacy allows one to measure how much privacy a protocol offers [111]. The fundamental theory of DP is that privacy should be seen as a resource, one that is depleted as data is pulled from a dataset. Private data analysis aims to obtain the most valuable information while consuming the least amount of privacy.

Differentially private federated learning is proposed as a potential technique for learning from distributed medical data, such as histopathology scans [112–114]. Federated learning enables the training of models without openly disclosing patient information, reducing privacy and confidentiality concerns related to healthcare data [115]. They suggested that this is supplemented by differential privacy's quantitative limits on the level of privacy offered. They used independent and identically distributed data (IID) and non-IID data distributions to show the effectiveness of federated learning (FedAvg) with simulated real-world data. Private federated learning is an option for distributed training on medical data since it produces results comparable to traditional centralized training [90]. A great alternative for building smart and safe IoT devices is FL's security features, especially considering the General Data Protection Regulation (GDPR), Health Insurance Portability and Accountability Act (HIPAA) [116], Health Information Technology for Economic and Clinical Health (HITECH) [117], Act on the Protection of Personal Information (APPI) [118] increasingly strict data privacy protection laws [119,120]. GDPR, HIPAA, HITECH, and APPI laws must be followed by healthcare organizations and entities handling Protected Health Information (PHI) to guarantee patient security and confidentiality. Considering these rules is essential to preserving confidence and safeguarding the private integrity of medical records. For federated learning to be used in healthcare in a way that is both compliant and successful, security protocols and any biases must be addressed. Flexibility, adaptability, and a commitment to patient-centric and ethical principles will be essential in shaping regulations that support the responsible advancement of federated learning in healthcare. Research effective and efficient privacy-preserving techniques must be investigated to incorporate them into FL's upcoming real-world projects.

### 4.5 Balancing Accuracy and Utility

Learning provides a balanced compromise between accuracy and utility while encouraging privacy [121]. In addition, FL training preserves model generalizability with only minimal accuracy reduction. As a result, FL empowers the smart healthcare system with enhanced scalability, owing to its distributed learning capabilities.

### 4.6 Affordable Health Data Training

FL offers substantial benefits regarding communication cost reduction, such as minimizing latency and lowering the transmit power required for raw data transmission [122]. This advantage stems from model gradients typically being much smaller than the original datasets. Hence, FL significantly conserves network bandwidth and helps mitigate network congestion risk, particularly in large healthcare networks.

## 5 Applications of FL

The rapid development of artificial intelligence is advancing intelligent digital transformation in numerous businesses. However, this progress gives rise to data fragmentation and islands issues, impeding the seamless sharing of data across various domains and hindering its full potential value realization. Federated learning can impact patients in the healthcare data-sharing process in several ways: privacy protection, informed consent, data ownership and control, data portability, enhanced clinical decision support systems, and ethical considerations. Federated learning holds immense promise in revolutionizing healthcare across various domains [95], particularly in prognosis and diagnosis, as shown in Fig. 14. In prognosis, federated learning can be instrumental in predicting and managing chronic conditions such as heart disease, COVID-19, and diabetes [123]. Using data from diverse sources such as hospitals, clinics, and wearable devices, a federated learning model can provide more accurate and personalized prognostic insights, facilitating early intervention and improved patient outcomes. In addition, in the diagnosis domain, federated learning can contribute significantly to assessing mortality risks in COVID-19 patients [124], predicting outcomes in conditions such as Parkinson's disease [125], and discerning various types of cancer [126]. By collaboratively training models on decentralized healthcare data while preserving privacy, federated learning enables the development of robust diagnostic tools that enhance disease identification and treatment planning, ultimately advancing the quality of healthcare delivery. Fig. 14 shows the application of FL in the healthcare domain. This survey focuses on disease prediction and prevention and medical imaging data, including X-rays, MRIs, and CT scans, for tasks such as tumor detection, drug discovery, and development. Federated learning works especially effectively in healthcare applications where data confidentiality and privacy are critical. Federated learning has shown potential in the following machine learning model types and healthcare-related task examples: Clinical Prediction Models, Disease Diagnosis Models, Epidemiological Surveillance, Natural Language Processing for EHRs, Personalized Treatment Recommendations, Drug Discovery and Development, and Health Monitoring and Wearables.

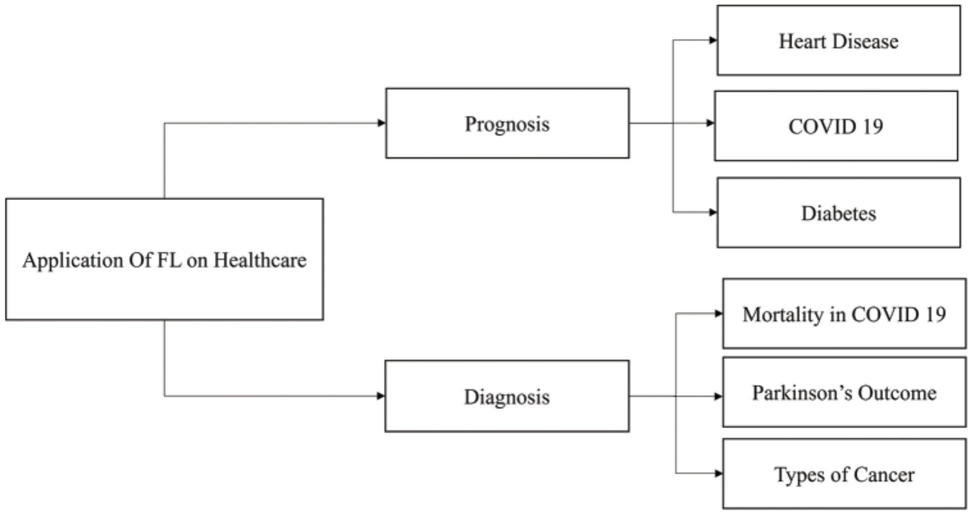


**Figure 14:** Applications of FL in healthcare

### *5.1 Functional Magnetic Resonance Imaging (fMRI) Analysis*

Healthcare professionals are concerned about potentially exploiting their data if it becomes publicly accessible. Various neurological conditions and disorders were linked to fMRI data [127–129]. The study successfully develops the proposed framework, which consists of two crucial components: the global server and data sharing. This proposed framework highlights the benefits of federated learning and proposes potential applications, such as detecting rarer tumors with limited patient data [129,130].

### *5.2 Electronic Medical Records*

Electronic Medical Records (EMRs) support the development of machine learning algorithms for predicting disease incidence, patient response to treatment, and other healthcare events [131]. Integrating Federated Learning with Electronic Medical Records (EMRs) addresses critical concerns surrounding the privacy and security of sensitive health information. Federated Learning eliminates the need to centralize or share this private data by allowing machine learning models to be trained locally on devices where the data resides. Only aggregated model updates are exchanged, ensuring that individual patient records remain confidential. This approach safeguards patient privacy, promotes collaborative research, and models improvement in healthcare while adhering to stringent data security regulations. The extensive utilization of electronic health records has significantly enhanced the feasibility of collaboration across multiple institutions [132]. Various neurological conditions and disorders were linked to fMRI data. The study effectively develops the proposed framework, which consists of a global server and data sharing as its core components. This proposed framework highlights the benefits of federated learning and presents potential applications, such as identifying rarer tumors with a limited number of patients [132–134].

### *5.3 Adverse Drug Reaction Prediction (ADRs)*

The pharmaceutical company, healthcare system, and medical experts are all extremely concerned about Adverse Drug Reactions (ADRs). Machine learning techniques play a crucial role in pharmaceutical surveillance, helping to detect, predict, and understand adverse drug reactions. The structure allows the training of a global model utilizing data from each site without uploading the raw data from those sites. For the initial instance, federated machine learning algorithms utilized distributed electronic health data to predict adverse drug reactions (ADRs) [135–137].

### 5.4 High Expenses in Training Health Data

High costs in health data training arise from the substantial financial investments required throughout preparing and utilizing healthcare data to train machine learning models and artificial intelligence algorithms. Collecting data from diverse sources, such as hospitals, clinics, wearable devices, and electronic health records, involves significant expenses in negotiating data-sharing agreements and ensuring compliance with privacy regulations [138].

Setting up and maintaining the necessary storage and computing infrastructure to handle these large-scale datasets also contributes to the overall costs. In addition, protecting patient data through robust security measures and ensuring compliance with strict privacy regulations further adds to the financial burden [139]. Employing skilled data scientists, machine learning engineers, and healthcare domain experts, along with the time-consuming process of obtaining accurate data annotations, compounds the expenses.

Training complex machine learning models also requires substantial computational power, increasing electricity consumption and operational costs [140]. In order to manage these high costs, data collaboration, cloud-based services, privacy-preserving techniques, and financial support from governments and private organizations are potential solutions to foster responsible and sustainable AI development in the healthcare industry.

### 5.5 Insufficiency of Accessible Medical Datasets on the Internet

The shortage of available medical datasets online presents a significant challenge for researchers and developers seeking to advance healthcare AI and machine learning applications [28]. Although the demand for data-driven healthcare solutions is rapidly growing, the availability of high-quality, publicly accessible medical datasets remains limited [141]. Several factors contribute to this shortage. First, medical data is inherently sensitive and subject to strict privacy regulations, such as HIPAA [116], HITECH [117], APPI [118], and GDPR [111,119], making it challenging to share patient information without compromising privacy [89]. Second, healthcare institutions often lack incentives to release their valuable and comprehensive datasets due to concerns about potential data misuse or competitive advantage [142]. In addition, the diversity and complexity of medical data, including electronic health records, medical images, and genomic information, impose large-scale, diverse datasets to build robust AI models [143]. However, assembling and maintaining such datasets requires significant resources and collaborations across multiple institutions. As a result, researchers and developers may encounter barriers in accessing sufficient data to adequately train and validate AI models, hampering progress in developing innovative healthcare solutions that can revolutionize diagnostics, personalized treatments, and healthcare management. In order to address this challenge, it is crucial for governments, regulatory bodies, and healthcare organizations to foster data-sharing collaborations, incentivize data contribution through appropriate data governance frameworks, and promote responsible data anonymization techniques to protect patient privacy while facilitating the growth and accessibility of medical datasets for the greater benefit of healthcare innovation [144].

## 6 Challenges and Directions

Implementing federated learning for healthcare applications comes with challenges. Federated learning presents several difficulties in the healthcare industry, such as data heterogeneity, privacy and security, communication overhead, model aggregation, patient consent, imbalanced data, model initialization and synchronization, and lack of standardization, as illustrated in Fig. 15.

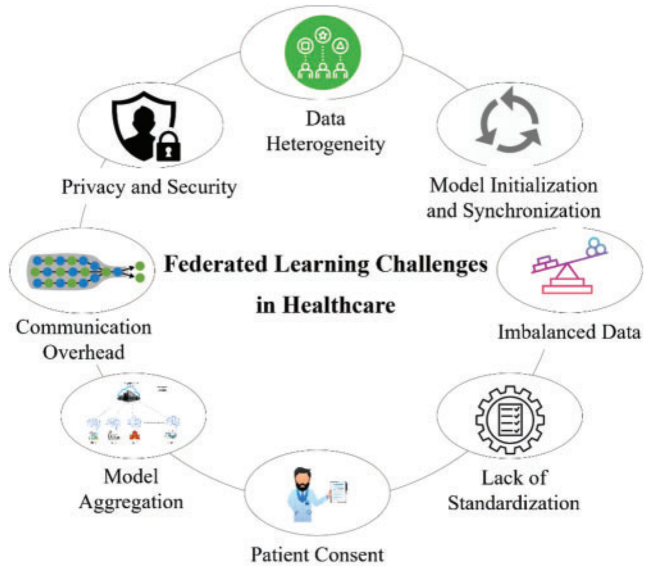


**Figure 15:** Federated learning challenges in healthcare

### *6.1 Data Heterogeneity*

Data were gathered from several users or devices in a federated learning healthcare environment. These users or devices can represent diverse demographics, behaviors, and preferences [145,146]. As a result, there can be variations in the patterns, characteristics, and distributions of the data gathered from healthcare. Different devices can generate data in different types or representations, including unstructured (text, pictures, audio), semi-structured, and structured data. Federated learning can have difficulties integrating and processing such a wide range of data types while maintaining security and privacy in healthcare [146]. Data transmission consistency across healthcare within a federated learning system can be affected [147]. An imbalance in data distribution can result in some devices having more data than others. Federated learning algorithms' performance may be impacted by this, mainly if it is not properly addressed. Federated learning must address data heterogeneity to guarantee that the models trained on decentralized data properly generalize across healthcare clients while preserving privacy and security. In federated learning systems, methods including data preparation, adaptive algorithms, model aggregation methods, and differential privacy protections are frequently employed to mitigate the effects of data heterogeneity.

### *6.2 Privacy and Security*

In federated learning, models are trained using data spread among several clients or devices, sometimes without direct access to the raw data [8]. However, during model updates or aggregation, there is a chance that adversaries can infer private information regarding the training dataset from the model's gradients or parameters. When aggregating updates from several clients, methods including differential privacy, secure aggregation, and encrypted model updates are employed to keep private information safe. To guarantee the confidentiality, integrity, and privacy of sensitive data throughout the federated learning process, various cryptographic techniques, privacy-preserving algorithms, reliable authorization procedures, and adversarial defense strategies must be used [9].

### 6.3 Communication Overhead

Although federated learning requires several healthcare data to communicate with a central server or one another to coordinate the training process, communication overhead is a significant obstacle. High bandwidth requirements, network latency, communication bottlenecks, heterogeneous networks, privacy concerns, and asynchronous updates are some issues with communication overhead in federated learning [148]. The design and development of effective communication protocols, optimization strategies, and distributed algorithms adapted to the features of federated learning environments are necessary to address communication overhead in federated learning. By minimizing the quantity of data transferred between devices and the central server, techniques including differential privacy, federated averaging, and model compression can assist in minimizing communication overhead while maintaining model performance and privacy.

### 6.4 Model Aggregation

In federated learning, particularly in healthcare settings, model aggregation combines locally trained models from several clients or devices to generate a global model while maintaining data confidentiality and privacy [9,75,149,150]. Developing effective communication protocols, privacy-preserving algorithms, strong aggregation strategies, and methods for managing heterogeneity and reliability concerns among distributed devices or clients in federated learning systems are all necessary to address these model aggregation challenges.

### 6.5 Patient Consent

Patient consent in federated learning healthcare presents significant challenges, particularly in healthcare settings involving sensitive medical data. To overcome these obstacles, a multidisciplinary group comprising medical professionals, legal experts, ethicists, and technologists must create consent management frameworks that balance patient autonomy, privacy protection, and the federated learning-based advancement of medical research and healthcare innovation [151,152]. In addition, utilizing technologies such as differential privacy and safe multiparty computation can provide a cooperative study of sensitive healthcare data while improving patient privacy.

### 6.6 Imbalanced Data

Imbalanced data in federated learning healthcare settings poses several challenges that must be addressed to ensure the effectiveness and fairness of machine learning models [153]. Addressing these challenges requires a combination of algorithmic approaches, data preprocessing techniques, privacy-preserving mechanisms, and evaluation strategies designed specifically for the imbalanced nature of healthcare data in federated learning settings [154,155]. Collaboration between healthcare providers, data scientists, and privacy experts is essential to develop robust and equitable machine learning solutions for healthcare applications.

### 6.7 Model Initialization and Synchronization

Model initialization and synchronization present difficulties in federated learning in the healthcare industry because of the sensitive nature of medical data, the diversity of data sources, and legal restrictions [156,157]. Addressing these challenges in model initialization and synchronization in federated learning healthcare requires a multidisciplinary approach involving expertise in machine learning, healthcare informatics, privacy-preserving techniques, regulatory compliance, and domain-specific knowledge of healthcare systems and practices. Cooperation between researchers, healthcare providers, legislators, and technology developers is necessary to overcome these obstacles and realize the potential advantages of federated learning in enhancing healthcare outcomes.

### *6.8 Lack of Standardization*

The development of standardized protocols, compatible systems, and ethical guidelines for the secure and privacy-preserving sharing and analysis of healthcare data is necessary to address the lack of standardization in federated learning healthcare [158]. This collaboration involves healthcare providers, researchers, policymakers, and regulatory bodies. Standardization initiatives should focus on standardizing data formats, protocols, privacy-preserving strategies, regulatory frameworks, and governance procedures to promote federated learning in healthcare while guaranteeing patient privacy, data security, and regulatory compliance [159].

Federated learning in healthcare should concentrate on resolving current issues, boosting productivity, improving privacy-preserving techniques, maintaining standardization, addressing non-IID data, optimizing communication overhead, real-time federated learning, ethical and regulatory frameworks, and broadening its range of uses in the future [150] as shown in Fig. 16.

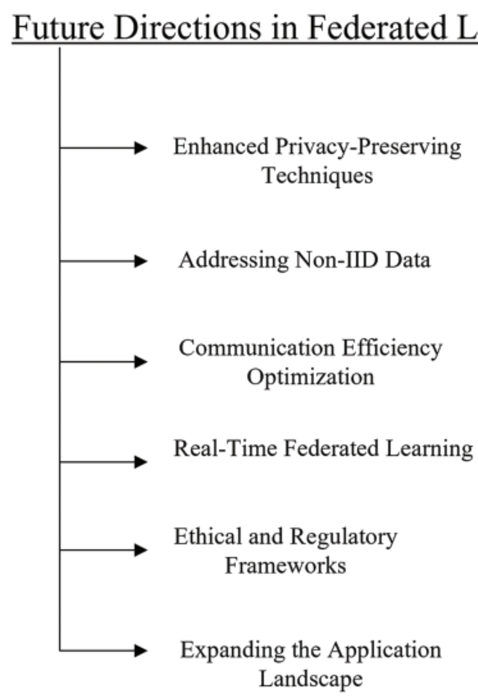


**Figure 16:** Future directions in federated learning

Federated learning can enhance the possible benefits of precision medicine and enhance healthcare in the context of global health [160]. Collaborative efforts between researchers, healthcare providers, and technology developers are essential to overcome these challenges and realize the potential of federated learning in improving healthcare outcomes [161].

## 7 Conclusion

Federated Learning, a promising decentralized AI strategy, has generated significant enthusiasm for achieving privacy-enhanced and scalable healthcare services and applications. This study examines the potential that Federated Learning brings to revolution in healthcare. It involves a comprehensive survey and in-depth conversations centered around the latest research in this domain. In this survey, an extensive examination is conducted on the fundamental concept of federated learning and its associated frameworks, technologies, and recent research on various FL-related subjects. The survey establishes a solid basis for understanding different FL components, exploring their advantages and disadvantages, and exploring their implementation strategies.

**Acknowledgement:** We thank the anonymous reviewers for their constructive feedback.

**Funding Statement:** This work was supported by a research fund from Chosun University, 2023.

**Author Contributions:** The authors confirm their contribution to the paper as follows: study conception and design: H.-I. K., D.U.; data collection: E.-M. Y. and C.-H.S.; analysis and interpretation of results: D.U., and E.-M.Y.; draft manuscript preparation: D.U., H.-I.K., and C.-H.S. All authors reviewed the results and approved the final version of the manuscript.

**Availability of Data and Materials:** This review paper primarily consolidates existing methods and findings from the literature. All data used in this study originates from publicly available datasets, which readers can access through the sources listed in the references section. Since the data comes from publicly accessible repositories, there are no restrictions on its availability. Hence, all materials and data utilized in this review are easily accessible to interested readers.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

1. Muhammad G, Alshehri F, Karray F, El Saddik A, Alsulaiman M, Falk TH. A comprehensive survey on multimodal medical signals fusion for smart healthcare systems. Inf Fusion. 2021;76(8):355–75. doi:10.1016/j.inffus.2021.06.007.
2. Amiri Z, Heidari A, Navimipour NJ, Esmaeilpour M, Yazdani Y. The deep learning applications in IoTbased bio-and medical informatics: a systematic literature review. Neural Comput Appl. 2024;36: 1–41.
3. Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ. 2023;23(1):689. doi:10.1186/s12909-023-04698-z.
4. Monteiro ACB, França RP, Arthur R, Iano Y. An overview of the Internet of medical things (IoMT): applications, benefits, and challenges. In: Security and Privacy Issues in Internet of Medical Things. 2023. p. 83–98.
5. Nguyen DC, Cheng P, Ding M, Lopez-Perez D, Pathirana PN, Li J, et al. Enabling AI in future wireless networks: a data life cycle perspective. IEEE Commun Surv Tut. 2020;23(1):553–95. doi:10.1109/COMST.2020.3024783.
6. Aledhari M, Razzak R, Parizi RM, Saeed F. Federated learning: a survey on enabling technologies, protocols, and applications. IEEE Access. 2020;8:140699–725. doi:10.1109/ACCESS.2020.3013541.

7. Sangaiah AK, Javadpour A, Ja'fari F, Pinto P, Chuang HM. Privacy-aware and AI techniques for healthcare based on k-anonymity model in internet of things. IEEE Trans Eng Manag. 2023, 1–15.
8. Wu G, Chen X, Gao Z, Zhang H, Yu S, Shen S. Privacy-preserving offloading scheme in multi-access mobile edge computing based on MADRL. J Parallel Distr Comput. 2024;183:104775. doi:10.1016/j.jpdc.2023.104775.
9. Shen Y, Shen S, Li Q, Zhou H, Wu Z, Qu Y. Evolutionary privacy-preserving learning strategies for edge-based IoT data sharing schemes. Digit Commun Netw. 2023;9(4):906–19. doi:10.1016/j.dcan.2022.05.004.
10. Beam AL, Kohane IS. Big data and machine learning in health care. JAMA. 2018;319(13):1317–8. doi:10.1001/jama.2017.18391.
11. LeCun Y, Bengio Y, Hinton G. Deep learning. Nat. 2015;521(7553):436–44. doi:10.1038/nature14539.
12. Rusk N. Deep learning. Nat Methods. 2016;13(1):35–35. doi:10.1038/nmeth.3707.
13. Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism. 2017;69(3):S36–S40. doi:10.1016/j.metabol.2017.01.011.
14. Noorbakhsh-Sabet N, Zand R, Zhang Y, Abedi V. Artificial intelligence transforms the future of health care. Am J Med. 2019;132(7):795–801. doi:10.1016/j.amjmed.2019.01.017.
15. Alonazi WB. Fraud and Abuse in the Saudi healthcare system: a triangulation analysis. Inquiry: J Health Care Org, Prov, Fin. 2020;57. doi:10.1177/0046958020954624.
16. Bordoloi D, Singh V, Sanober S, Buhari SM, Ujjan JA, Boddu R. Deep learning in healthcare system for quality of service. J Healthc Eng. 2022;2022(4):1–11. doi:10.1155/2022/8169203.
17. Thornton D, Mueller RM, Schoutsen P, van Hillegersberg J. Predicting healthcare fraud in medicaid: a multidimensional data model and analysis techniques for fraud detection. Proc Technol. 2013;9(1):1252–64. doi:10.1016/j.protcy.2013.12.140.
18. Li H, Li C, Wang J, Yang A, Ma Z, Zhang Z, et al. Review on security of federated learning and its application in healthcare. Future Gener Comput Syst. 2023;144:271–90. doi:10.1016/j.future.2023.02.021.
19. McMahan B, Moore E, Ramage D, Hampson S, Arcas BA. Communication-efficient learning of deep networks from decentralized data. Artif Intell Stat. 2017;54:1273–82. PMLR.
20. Nguyen DC, Pham QV, Pathirana PN, Ding M, Seneviratne A, Lin Z, et al. Federated learning for smart healthcare: a survey. ACM Comput Surveys. 2022;55(3):1–37.
21. Agrawal S, Sarkar S, Aouedi O, Yenduri G, Piamrat K, Alazab M, et al. Federated learning for intrusion detection system: concepts, challenges and future directions. Comput Commun. 2022;195(5):346–61. doi:10.1016/j.comcom.2022.09.012.
22. Demertzi V, Demertzis S, Demertzis K. An overview of privacy dimensions on industrial internet of things (IIoT). arXiv preprint arXiv:2301.06172. 2023.
23. Gadekallu TR, Pham QV, Huynh-The T, Bhattacharya S, Maddikunta PKR, Liyanage M. Federated learning for big data: a survey on opportunities, applications, and future directions. arXiv preprint arXiv:2110.04160. 2021.
24. Chen J, Ran X. Deep learning with edge computing: a review. Proc IEEE. 2019;107(8):1655–74. doi:10.1109/JPROC.2019.2921977.
25. 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. 2022;34(16):e6002. doi:10.1002/cpe.6002.
26. Pronovost PJ, Mathews SC, Chute CG, Rosen A. Creating a purpose-driven learning and improving health system: the Johns Hopkins Medicine quality and safety experience. Learn Health Syst. 2017;1(1):e10018. doi:10.1002/lrh2.10018.
27. Lyu L, Yu H, Yang Q. Threats to federated learning: a survey. arXiv preprint arXiv:2003.02133. 2020.
28. Khan MJ, Tawose OT, Hu R, Zhao D. Exploring the efficacy of data-decoupled federated learning for image classification and medical imaging analysis. In: International Workshop on Federated Learning for Distributed; 2023; Long Beach, California, USA.

29. Rieke N, Hancox J, Li W, Milletari F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. npj Digit Med. 2020;3(1):1–7. doi:10.1038/s41746-020-00323-1.
30. Biswas S, Sharif K, Li F, Nour B, Wang Y. A scalable blockchain framework for secure transactions in IoT. IEEE Internet Things J. 2018;6(3):4650–9. doi:10.1109/JIOT.2018.2874095.
31. Hasavari S, Song YT. A secure and scalable data source for emergency medical care using blockchain technology. In: 2019 IEEE 17th International Conference on Software Engineering Research, Management and Applications (SERA); 2019; Honolulu, Hawaii, USA, IEEE. p. 71–5.
32. Kaur G, Gandhi C. Scalability in blockchain: challenges and solutions. In: Handbook of research on blockchain technology. Amsterdam, Elsevier; 2020. p. 373–406.
33. Ogiela MR, Majcher M. Security of distributed ledger solutions based on blockchain technologies. In: 2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA); Krakow, Poland, IEEE. p. 1089–95.
34. Kumar R, Khan AA, Kumar J, Golilarz NA, Zhang S, Ting Y, et al. Blockchain-federated-learning and deep learning models for COVID-19 detection using CT imaging. IEEE Sens J. 2021;21(14):16301–14. doi:10.1109/JSEN.2021.3076767.
35. Himeur Y, Sayed A, Alsalemi A, Bensaali F, Amira A, Varlamis I, et al. Blockchain-based recommender systems: applications, challenges and future opportunities. Comput Sci Rev. 2022;43(5):100439. doi:10.1016/j.cosrev.2021.100439.
36. Aminizadeh S, Heidari A, Toumaj S, Darbandi M, Navimipour NJ, Rezaei M, et al. The applications of machine learning techniques in medical data processing based on distributed computing and the Internet of Things. Comput Methods Prog Biomed. 2023;241(October):107745. doi:10.1016/j.cmpb.2023.107745.
37. Ehrenstein V, Kharrazi H, Lehmann H, Taylor CO. Obtaining data from electronic health records. in: Tools and technologies for registry interoperability, registries for evaluating patient outcomes: a user's guide, 3rd Edition, Addendum 2. Rockville (MD), USA, Agency for Healthcare Research and Quality; 2019.
38. Chang Y, Fang C, Sun W. A blockchain-based federated learning method for smart healthcare. Comput Intell Neurosci. 2021;2021(3):1–12. doi:10.1155/2021/4376418.
39. 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;27(2):790–803. doi:10.1109/JBHI.2022.3185673.
40. Mwase C, Jin Y, Westerlund T, Tenhunen H, Zou Z. Communication-efficient distributed AI strategies for the IoT edge. Future Gener Comput Syst. 2022;131(3):292–308. doi:10.1016/j.future.2022.01.013.
41. Joshi M, Pal A, Sankarasubbu M. Federated learning for healthcare domain-pipeline, applications and challenges. ACM Trans Comput Healthc. 2022;3(4):1–36. doi:10.1145/3533708.
42. Antunes RS, André da Costa C, Küderle A, Yari IA, Eskofier B. Federated learning for healthcare: systematic review and architecture proposal. ACM Trans Intell Syst Technol. 2022;13(4):1–23. doi:10.1145/3501813.
43. Moshawrab M, Adda M, Bouzouane A, Ibrahim H, Raad A. Reviewing federated machine learning and its use in diseases prediction. Sensors. 2023;23(4):2112, 1–39. doi:10.3390/s23042112.
44. Sadilek A, Liu L, Nguyen D, Kamruzzaman M, Serghiou S, Rader B, et al. Privacy-first health research with federated learning. npj Digit Med. 2021;4(1):132, 1–8. doi:10.1038/s41746-021-00489-2.
45. Balyen L, Peto T. Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology. Asia-Pacific J Ophthalmol. 2019;8(3):264–72.
46. Panayides AS, Amini A, Filipovic ND, Sharma A, Tsaftaris SA, Young A, et al. AI in medical imaging informatics: current challenges and future directions. IEEE J Biomed Health Inform. 2020;24(7):1837–57. doi:10.1109/JBHI.2020.2991043.

47. Ahmed Z, Mohamed K, Zeeshan S, Dong X. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database. 2020;2020(104):1–35. doi:10.1093/database/baaa010.
48. Zhang T, Mao S. An introduction to the federated learning standard. GetMobile: Mob Comput Commun. 2022;25(3):18–22. doi:10.1145/3511285.3511291.
49. Fan Z, Fang H, Zhou Z, Pei J, Friedlander MP, Zhang Y. Fair and efficient contribution valuation for vertical federated learning. arXiv preprint arXiv:2201.02658. 2022.
50. Rahman KJ, Ahmed F, Akhter N, Hasan M, Amin R, Aziz KE, et al. Challenges, applications and design aspects of federated learning: a survey. IEEE Access. 2021;9:124682–700. doi:10.1109/ACCESS.2021.3111118.
51. Sakib S, Fouda MM, Fadlullah ZM, Abualsaud K, Yaacoub E, Guizani M. Asynchronous federated learning-based ECG analysis for arrhythmia detection. In: 2021 IEEE International Mediterranean Conference on Communications and Networking (MeditCom); 2021; Athens, Greece, IEEE. p. 277–82.
52. Siddiqui MK, Morales-Menendez R, Huang X, Hussain N. A review of epileptic seizure detection using machine learning classifiers. Brain Inform. 2020;7(1):1–18. doi:10.1186/s40708-020-00105-1.
53. Woldaregay AZ, Årsand E, Botsis T, Albers D, Mamykina L, Hartvigsen G. Data-driven blood glucose pattern classification and anomalies detection: machine-learning applications in type 1 diabetes. J Med Internet Res. 2019;21(5):e11030.
54. Tian L, Zhang D, Bao S, Nie P, Hao D, Liu Y, et al. Radiomics-based machine-learning method for prediction of distant metastasis from soft-tissue sarcomas. Clin Radiol. 2021;76(2):158–e19.
55. Kaouk JH, Garisto J, Eltemamy M, Bertolo R. Robot-assisted surgery for benign distal ureteral strictures: step-by-step technique using the SP® surgical system. BJU Int. 2019;123(4):733–9.
56. Lanfranco AR, Castellanos AE, Desai JP, Meyers WC. Robotic surgery: a current perspective. Ann Surg. 2004;239(1):14–21.
57. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. Radiographics. 2017;37(2):505–15.
58. Mattfeldt T, Kestler HA, Hautmann R, Gottfried HW. Prediction of prostatic cancer progression after radical prostatectomy using artificial neural networks: a feasibility study. BJU Int. 1999;84(3):316–23. doi:10.1046/j.1464-410x.1999.00209.x.
59. Kundu I, Paul G, Banerjee R. A machine learning approach towards the prediction of protein-ligand binding affinity based on fundamental molecular properties. RSC Adv. 2018;8(22):12127–37. doi:10.1039/C8RA00003D.
60. Bahrami M, Forouzanfar M. Sleep apnea detection from single-lead ECG: a comprehensive analysis of machine learning and deep learning algorithms. IEEE Trans Instrum Meas. 2022;71(1):1–11. doi:10.1109/TIM.2022.3151947.
61. Bhatt CM, Patel P, Ghetia T, Mazzeo PL. Effective heart disease prediction using machine learning techniques. Algorithms. 2023;16(2):1–14. doi:10.3390/a16020088.
62. Hu Q, Xia H, Zhang T. Chatbot combined with deep convolutional neural network for skin cancer detection. In: 2022 3rd International Conference on Artificial Intelligence and Education (IC-ICAIE 2022); 2022; Chengdu, China, Atlantis Press. p. 35–41.
63. Khan MA, Khan A, Alhaisoni M, Alqahtani A, Alsubai S, Alharbi M, et al. Multimodal brain tumor detection and classification using deep saliency map and improved dragonfly optimization algorithm. Int J Imaging Syst Technol. 2023;33(2):572–87. doi:10.1002/ima.22831.
64. Ogundepo EA, Yahya WB. Performance analysis of supervised classification models on heart disease prediction. Innov Syst Softw Eng. 2023;19(1):129–44. doi:10.1007/s11334-022-00524-9.

65. Jagadeesha N, Trisal A, Tiwar VN. Skin tone assessment using hyperspectral reconstruction from RGB image. In: 2023 15th International Conference on Communication Systems & Networks (COMSNETS); 2023; Bengaluru, India, IEEE. p. 125–8.
66. Sharma S, Guleria K. A deep learning based model for the detection of pneumonia from chest X-ray images using VGG-16 and neural networks. Proc Comput Sci. 2023;218(3-4):357–66. doi:10.1016/j.procs.2023.01.018.
67. Jain R, Gupta M, Taneja S, Hemanth DJ. Deep learning based detection and analysis of COVID-19 on chest X-ray images. Appl Intell. 2021;51(3):1690–700. doi:10.1007/s10489-020-01902-1.
68. Amin J, Sharif M, Gul N, Raza M, Anjum MA, Nisar MW, et al. Brain tumor detection by using stacked autoencoders in deep learning. J Med Syst. 2020;44(2):1–12. doi:10.1007/s10916-019-1483-2.
69. Rashid J, Ishfaq M, Ali G, Saeed MR, Hussain M, Alkhalifah T, et al. Skin cancer disease detection using transfer learning technique. Appl Sci. 2022;12(11):1–16. doi:10.3390/app12115714.
70. Balaha HM, Hassan AES. Skin cancer diagnosis based on deep transfer learning and sparrow search algorithm. Neural Comput Appl. 2023;35(1):815–53. doi:10.1007/s00521-022-07762-9.
71. Mishra A, Saha S, Mishra S, Bagade P. A federated learning approach for smart healthcare systems. CSI Trans ICT. 2023;11(1):39–44. doi:10.1007/s40012-023-00382-1.
72. Chouhan V, Singh SK, Khamparia A, Gupta D, Tiwari P, Moreira C, et al. A novel transfer learning based approach for pneumonia detection in chest X-ray images. Appl Sci. 2020;10(2):1–17. doi:10.3390/app10020559.
73. Kareem A, Liu H, Velisavljevic V. A federated learning framework for pneumonia image detection using distributed data. Healthc Anal. 2023;4(9773):1–14. doi:10.1016/j.health.2023.100204.
74. Khan SH, Alam MGR. A federated learning approach to pneumonia detection. In: 2021 International Conference on Emerging Engineering Technology (ICEET); 2021; Istanbul, Turkey, IEEE. p. 1–6.
75. Wen J, Zhang Z, Lan Y, Cui Z, Cai J, Zhang W. A survey on federated learning: challenges and applications. Int J Mach Learn Cybern. 2023;14(2):513–35. doi:10.1007/s13042-022-01647-y.
76. Qayyum A, Ahmad K, Ahsan MA, Al-Fuqaha A, Qadir J. Collaborative federated learning for healthcare: multi-modal COVID-19 diagnosis at the edge. IEEE Open J Comput Soc. 2022;3(12):172–84. doi:10.1109/OJCS.2022.3206407.
77. Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M. COVID-19 image data collection: prospective predictions are the future. arXiv preprint arXiv:2006.11988. 2020.
78. Born J, Brändle G, Cossio M, Disdier M, Goulet J, Roulin J, et al. POCOVID-Net: automatic detection of COVID-19 from a new lung ultrasound imaging dataset (POCUS). arXiv preprint arXiv:2004.12084. 2020.
79. Wei K, Li J, Ding M, Ma C, Yang HH, Farokhi F, et al. Federated learning with differential privacy: algorithms and performance analysis. IEEE Trans Inf Forensics Secur. 2020;15:3454–69. doi:10.1109/TIFS.2020.2988575.
80. Yi L, Zhang J, Zhang R, Shi J, Wang G, Liu X. SU-Net: an efficient encoder-decoder model of federated learning for brain tumor segmentation. In: International Conference on Artificial Neural Networks; 2020; Slovakia, Springer International Publishing. p. 761–73.
81. Heidari A, Javaheri D, Toumaj S, Navimipour NJ, Rezaei M, Unal M. A new lung cancer detection method based on the chest CT images using Federated Learning and blockchain systems. Artif Intell Med. 2023;141(2):102572. doi:10.1016/j.artmed.2023.102572.
82. Lincy M, Kowshalya AM. Early detection of type-2 diabetes using federated learning. Int J Sci Res Sci Eng Technol. 2020;12(3):257–67. doi:10.32628/IJSRSET.
83. Astillo PV, Duguma DG, Park H, Kim J, Kim B, You I. Federated intelligence of anomaly detection agent in IoTMD-enabled diabetes management control system. Future Gener Comput Syst. 2022;128(8):395–405. doi:10.1016/j.future.2021.10.023.

84. Hossen MN, Panneerselvam V, Koundal D, Ahmed K, Bui FM, Ibrahim SM. Federated machine learning for detection of skin diseases and enhancement of internet of medical things (IoMT) security. IEEE J Biomed Health Inform.. 2022;27(2):835–41. doi:10.1109/JBHI.2022.3149288.
85. Jiménez-Sánchez A, Tardy M, Ballester MAG, Mateus D, Piella G. Memory-aware curriculum federated learning for breast cancer classification. Comput Methods Programs Biomed. 2023;229(3–4):107318. doi:10.1016/j.cmpb.2022.107318.
86. Giuseppi A, Manfredi S, Menegatti D, Poli C, Pietrabissa A. Decentralised federated learning for hospital networks with application to COVID-19 detection. IEEE Access. 2022;10(1):92681–91. doi:10.1109/ACCESS.2022.3202922.
87. Balaji NNA, Narayanan N, Thelly KJ, Babu C. Investigating federated learning strategies for pneumonia image classification. In: 2021 IEEE MIT Undergraduate Research Technology Conference (URTC); 2021; Massachusetts, USA, IEEE. p. 1–5.
88. Wicaksana J, Yan Z, Zhang D, Huang X, Wu H, Yang X, et al. FedMix: mixed supervised federated learning for medical image segmentation. IEEE Trans Med Imaging. 2022;42(7):1955–68. doi:10.1109/TMI.2022.3233405.
89. Liang Z, Wang H. FedST: federated shapelet transformation for interpretable time series classification. arXiv preprint arXiv:2302.10631. 2023.
90. Hansen CR, Price G, Field M, Sarup N, Zukauskaite R, Johansen J, et al. Larynx cancer survival model developed through open-source federated learning. Radiother Oncol. 2022;176:179–86. doi:10.1016/j.radonc.2022.09.023.
91. Cetinkaya AE, Akin M, Sagiroglu S. A communication efficient federated learning approach to multi chest diseases classification. In: 2021 6th International Conference on Computer Science and Engineering (UBMK); 2021; Ankara, Turkey, IEEE. p. 429–34.
92. Ma Z, Zhang M, Liu J, Yang A, Li H, Wang J, et al. An assisted diagnosis model for cancer patients based on federated learning. Front Oncol. 2022;12:860532. doi:10.3389/fonc.2022.860532.
93. Yaqoob MM, Nazir M, Khan MA, Qureshi S, Al-Rasheed A. Hybrid classifier-based federated learning in health service providers for cardiovascular disease prediction. Appl Sci. 2023;13(3):1911.
94. Qian C, Xiong H, Li J. FeDeFo: a personalized federated deep forest framework for Alzheimer's disease diagnosis. In: International Conference on Software Engineering and Knowledge Engineering; 2023; USA, Virtual Conference Center. p. 1–6.
95. Rauniyar A, Hagos DH, Jha D, Håkegård JE, Bagci U, Rawat DB, et al. Federated learning for medical applications: a taxonomy, current trends, challenges, and future research directions. IEEE Internet Things J. 2023;11(5):7374–98. doi:10.1109/JIOT.2023.3329061.
96. Goto S, Solanki D, John JE, Yagi R, Homilius M, Ichihara G, et al. Multinational federated learning approach to train ECG and echocardiogram models for hypertrophic cardiomyopathy detection. Circulation. 2022;146(10):755–69. doi:10.1161/CIRCULATIONAHA.121.058696.
97. Konečný J, McMahan HB, Ramage D, Richtárik P. Federated optimization: distributed machine learning for on-device intelligence. arXiv preprint arXiv:1610.02527. 2016.
98. Suresh AT, Felix XY, Kumar S, McMahan HB. Distributed mean estimation with limited communication. In: International Conference on Machine Learning; 2017; Stockholm, Sweden, PMLR. p. 3329–37.
99. Uddin MP, Xiang Y, Lu X, Yearwood J, Gao L. Federated learning via disentangled information bottleneck. IEEE Trans Serv Comput. 2022;16(3):1874–89. doi:10.1109/TSC.2022.3187962.
100. Xu X, Li H, Li Z, Zhou X. Safe: synergic data filtering for federated learning in cloud-edge computing. IEEE Trans Ind Inform. 2022;19(2):1655–65. doi:10.1109/TII.2022.3195896.
101. Chen C, Gao Y, Huang S, Yan X. Avoid attacks: a federated data sanitization defense in IoMT systems. In: IEEE INFOCOM 2023—IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS); 2023; USA, IEEE. p. 1–6.

102. Zizzo G, Rawat A, Sinn M, Buesser B. Fat: federated adversarial training. arXiv preprint arXiv:2012.01791. 2020.
103. Pillutla K, Kakade SM, Harchaoui Z. Robust aggregation for federated learning. IEEE Trans Signal Process. 2022;70:1142–54. doi:10.1109/TSP.2022.3153135.
104. Fang H, Qian Q. Privacy preserving machine learning with homomorphic encryption and federated learning. Future Internet. 2021;13(4):94. doi:10.3390/fi13040094.
105. 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, USA; 2020. p. 1–9.
106. Batool H, Anjum A, Khan A, Izzo S, Mazzocca C, Jeon G. A secure and privacy preserved infrastructure for VANETs based on federated learning with local differential privacy. Inf Sci. 2024;652(2):119717. doi:10.1016/j.ins.2023.119717.
107. Aziz R, Banerjee S, Bouzefrane S, Le Vinh T. Exploring homomorphic encryption and differential privacy techniques towards secure federated learning paradigm. Future Internet. 2023;15(9):310. doi:10.3390/fi15090310.
108. Li Y, Zhou Y, Jolfaei A, Yu D, Xu G, Zheng X. Privacy-preserving federated learning framework based on chained secure multiparty computing. IEEE Internet Things J. 2020;8(8):6178–86. doi:10.1109/JIOT.2020.3022911.
109. Zhang Q, Xin C, Wu H. Privacy-preserving deep learning based on multiparty secure computation: a survey. IEEE Internet Things J. 2021;8(13):10412–29. doi:10.1109/JIOT.2021.3058638.
110. Zhao Y, Zhao J, Yang M, Wang T, Wang N, Lyu L, et al. Local differential privacy-based federated learning for internet of things. IEEE Internet Things J. 2020;8(11):8836–53. doi:10.1109/JIOT.2020.3037194.
111. Truong N, Sun K, Wang S, Guitton F, Guo Y. Privacy preservation in federated learning: an insightful survey from the GDPR perspective. Comput Secur. 2021;110(4):102402. doi:10.1016/j.cose.2021.102402.
112. Adnan M, Kalra S, Cresswell JC, Taylor GW, Tizhoosh HR. Federated learning and differential privacy for medical image analysis. Sci Rep. 2022;12(1):1953.
113. Andreux M, du Terrail JO, Beguier C, Tramel EW. Siloed federated learning for multi-centric histopathology datasets. In: Domain adaptation and representation transfer, and distributed and collaborative learning. Lima, Peru, Springer; 2020. p. 129–39.
114. Gunesli GN, Bilal M, Raza SEA, Rajpoot NM. Feddropoutavg: generalizable federated learning for histopathology image classification. arXiv preprint arXiv:2111.13230. 2021.
115. Mothukuri V, Parizi RM, Pouriyeh S, Huang Y, Dehghantanha A, Srivastava G. A survey on security and privacy of federated learning. Future Gener Comput Syst. 2021;115(4):619–40. doi:10.1016/j.future.2020.10.007.
116. Gostin LO, Levit LA, Nass SJ. Beyond the HIPAA privacy rule: enhancing privacy, improving health through research. Washington DC, USA, Institute of Medicine of National Academies. 2009. p. 1–335.
117. Burde H. The HITECH act: an overview. AMA J Ethics. 2011;13(3):172–5. doi:10.1001/virtualmentor.2011.13.3.hlaw1-1103.
118. Kirchhoff U, Schiebe T. The reform of the Japanese act on protection of personal information. From the practitioner's perspective. Z Japanisches Recht. 2017;22(44):199–212.
119. Regulation GDP. General data protection regulation (GDPR). Intersoft Consult. 2018;24(1):1–8.
120. Khan MSI, Rahman A, Islam S, Nasir MK, Band SS, Mosavi A. IoT and WSN based effluent treatment plant monitoring system. EasyChair Preprint. 2021;5023:1–19.
121. Yuan X, Ni W, Ding M, Wei K, Li J, Poor HV. Amplitude-varying perturbation for balancing privacy and utility in federated learning. IEEE Trans Inf Forensics Secur. 2023;18:1884–97. doi:10.1109/TIFS.2023.3258255.

122. Xu J, Glicksberg BS, Su C, Walker P, Bian J, Wang F. Federated learning for healthcare informatics. J Healthc Inform Res. 2021;5(1):1–19. doi:10.1007/s41666-020-00082-4.
123. Badidi E. Edge AI for early detection of chronic diseases and the spread of infectious diseases: opportunities, challenges, and future directions. Future Internet. 2023;15(11):370, 1–34. doi:10.3390/fi15110370.
124. Flores M, Dayan I, Roth H, Zhong A, Harouni A, Gentili A, et al. Federated learning used for predicting outcomes in SARS-COV-2 patients. Res Sq. 2021;1–30. doi:10.21203/rs.3.rs-126892/v1.
125. Dipro SH, Islam M, Nahian MAA, Azad MS. A federated learning approach for detecting Parkinson's disease through privacy preserving by blockchain (Thesis & Report). 2022. Brac University: Bangladesh; 2022.
126. Chowdhury A, Kassem H, Padoy N, Umeton R, Karargyris A. A review of medical federated learning: applications in oncology and cancer research. In: International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) Brainlesion Workshop; 2021; Springer International Publishing. p. 3–24.
127. Bey R, Goussault R, Grolleau F, Benchoufi M, Porcher R. Fold-stratified cross-validation for unbiased and privacy-preserving federated learning. J Am Med Inform Assoc. 2020;27(8):1244–51. doi:10.1093/jamia/ocaa096.
128. KhoKhar FA, Shah JH, Khan MA, Sharif M, Tariq U, Kadry S. A review on federated learning towards image processing. Comput Electr Eng. 2022;99:107818.
129. Li X, Gu Y, Dvornek N, Staib LH, Ventola P, Duncan JS. Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: aBIDE results. Med Image Anal. 2020;65:101765.
130. Mahlool DH, Abed MH. A comprehensive survey on federated learning: concept and applications. Mobile Computing and Sustainable Informatics. 2022. vol. 126, p. 539–53.
131. Huang L, Shea AL, 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.
132. Vaid A, Jaladanki SK, Xu J, Teng S, Kumar A, Lee S, et al. Federated learning of electronic health records to improve mortality prediction in hospitalized patients with COVID-19: machine learning approach. JMIR Med Inform. 2021;9(1):e24207.
133. Dang TK, Lan X, Weng J, Feng M. Federated learning for electronic health records. ACM Trans Intell Syst Technol. 2022;13(5):1–17.
134. Lu S, Zhang Y, Wang Y. Decentralized federated learning for electronic health records. In: 2020 54th Annual Conference on Information Sciences and Systems (CISS); 2020; Princeton, NJ, USA, IEEE. p. 1–5.
135. Choudhury O, Park Y, Salonidis T, Gkoulalas-Divanis A, Sylla I. Predicting adverse drug reactions on distributed health data using federated learning. AMIA Annu Symp Proc. 2019;2019:313–22.
136. Islam TU, Ghasemi R, Mohammed N. Privacy-preserving federated learning model for healthcare data. In: 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC); 2022; Las Vegas, NV, USA, IEEE. p. 281–7.
137. Yoo JH, Jeong H, Lee J, Chung TM. Federated learning: issues in medical application. In: Future data and security engineering. Springer; 2021. p. 3–22.
138. Bonawitz K, Kairouz P, McMahan B, Ramage D. Federated learning and privacy: building privacy-preserving systems for machine learning and data science on decentralized data. Queue. 2021;19(5): 87–114.
139. Yu B, Mao W, Lv Y, Zhang C, Xie Y. A survey on federated learning in data mining. Wiley Interdisciplinary Rev Data Mining Knowl Discov. 2022;12(1):e1443. doi:10.1002/widm.1443.
140. Sater RA, Hamza AB. Afederated learning approach to anomaly detection in smart buildings.ACM. Trans Internet Things. 2021;2(4):1–23. doi:10.1145/3467981.

141. Crowson MG, Moukheiber D, Arévalo AR, Lam BD, Mantena S, Rana A, et al. A systematic review of federated learning applications for biomedical data. PLoS Digit Health. 2022;1(5):e0000033. doi:10.1371/journal.pdig.0000033.
142. Lim WY, Luong NC, Hoang DT, Jiao Y, Liang YC, Yang Q, et al. Federated learning in mobile edge networks: a comprehensive survey. IEEE Commun Surveys Tut. 2020;22(3):2031–63. doi:10.1109/COMST.2020.2986024.
143. Pati S, Baid U, Edwards B, Sheller M, Wang SH, Reina GA, et al. Federated learning enables big data for rare cancer boundary detection. Nat Commun. 2022;13(1):7346. doi:10.1038/s41467-022-33407-5.
144. Li Q, Wen Z, Wu Z, Hu S, Wang N, Li Y, et al. A survey on federated learning systems: vision, hype and reality for data privacy and protection. IEEE Trans Knowl Data Eng. 2021;35(4):3347–66. doi:10.1109/TKDE.2021.3124599.
145. Myrzashova R, Alsamhi SH, Shvetsov AV, Hawbani A, Wei X. Blockchain meets federated learning in healthcare: a systematic review with challenges and opportunities. IEEE Internet Things J.. 2023;10(16):14418–37. doi:10.1109/JIOT.2023.3263598.
146. Guan H, Yap PT, Bozoki A, Liu M. Federated learning for medical image analysis: a survey. Pattern Recog. 2024;151(3):1–17. doi:10.1016/j.patcog.2024.110424.
147. Sugianto N, Tjondronegoro D, Sorwar G. Collaborative federated learning framework to minimize data transmission for AI-enabled video surveillance. Inf Technol People. 2024;1–25. doi:10.1108/ITP-08-2021-0598.
148. Yang J, Zhou Y, Wen W, Zhou J, Zhang Q. Asynchronous hierarchical federated learning based on bandwidth allocation and client scheduling. Appl Sci. 2023;13(20):11134. doi:10.3390/app132011134.
149. Shi Y, Fan P, Zhu Z, Peng C, Wang F, Letaief KB. SAM: an efficient approach with selective aggregation of models in federated learning. IEEE Internet Things J. 2024;1–15. doi:10.1109/JIOT.2024.3373822.
150. Shen S, Tople S, Saxena P. Auror: defending against poisoning attacks in collaborative deep learning systems. In: Proc 32nd Annual Conference Computer Security Applications; New York, USA; 2016. p. 508–19.
151. Al Salman H, Al-Rakhami MS, Alfakih T, Hassan MM. Federated learning approach for breast cancer detection based on DCNN. IEEE Access. 2024;12:40114–38. doi:10.1109/ACCESS.2024.3374650.
152. Handa T, Singhal I, Chakraborty P, Kaur G. Recent trends of federated learning for smart healthcare systems. In: Federated learning and AI for Healthcare 5.0. IGI Global. 2024. p. 78–103.
153. Mitrovska A, Safari P, Ritter K, Shariati B, Fischer JK. Secure federated learning for Alzheimer's disease detection. Front Aging Neurosci. 2024;16:1324032. doi:10.3389/fnagi.2024.1324032.
154. Rafi TH, Noor FA, Hussain T, Chae DK. Fairness and privacy preserving in federated learning: a survey. Inf Fusion. 2024;105(4):102198. doi:10.1016/j.inffus.2023.102198.
155. Roszel M. Towards trustworthy artificial intelligence in privacy-preserving collaborativemachine learning (Doctoral Thesis). Université Du Luxembourg, The Faculty of Science, Technology and Medicine; 2024.
156. Llasag Rosero R, Silva C, Ribeiro B, Santos BF. Label synchronization for Hybrid federated learning in manufacturing and predictive maintenance. J Intell Manuf. 2024;1–20. doi:10.1007/s10845-023-02298-8.
157. Hard A, Girgis AM, Amid E, Augenstein S, McConnaughey L, Mathews R, et al. Learning from straggler clients in federated learning. arXiv preprint arXiv:2403.09086. 2024.
158. Thummisetti BSP, Atluri H. Advancing healthcare informatics for empowering privacy and security through federated learning paradigms. Int J Sustain Dev Comput Sci. 2024;1(1):1–16.
159. Dihan MS, Akash AI, Tasneem Z, Das P, Das SK, Islam MR, et al. Digital twin: data exploration, architecture, implementation and future. Heliyon. 2024;10(5):e26503. doi:10.1016/j.heliyon.2024.e26503.

160. de Falco I, Della Cioppa A, Koutny T, Scafuri U, Tarantino E. Model-free-communication federated learning: framework and application to precision medicine. Biomed Signal Process Control. 2024;87(9):105416. doi:10.1016/j.bspc.2023.105416.

161. Issac BM, Kumar S. Collaborative federated learning in healthcare systems. In: Handbook on federated learning. 1st Edition. CRC Press; 2024. p. 226–44.