A Look at Federated Learning Applications in Healthcare

Qiuxian Chen, Yizheng Tao
Institute of Computer Application, China Academy of Engineering Physics
Mianyang, China
chen_qiuxian@outlook.com, taolilan@126.com

***Abstract*—This paper provides an in-depth overview of the federated learning (FL) applications within the healthcare domain. Firstly, we discuss the background of FL, including its relationship with other machine learning (ML) technologies, the drawbacks of conventional ML methods in the field of healthcare, and how FL can alleviate these drawbacks. Second, we present three different FL frameworks based on data partitioning and describe possible medical scenarios. Thirdly, we classify current research on FL applications into two categories based on the types of data employed. Finally, we summarize the investigated work and propose a number of potential research directions for FL applications in healthcare.**

I. Introduction

Artificial intelligence (AI) is rapidly evolving and expanding these days, with machine learning (ML) being a specific approach and driving force behind many AI applications. Google introduced the concept of federated learning (FL) in 2016 [1], [2], which is an ML paradigm that develops a common model via multiple independent participants. The relationships between FL, distributed machine learning (DML), and centralized machine learning (CML) are depicted in Fig. 1. In fact, the main difference between FL and other ML approaches is how data are organized during model training. Unlike CML, which uses a centralized dataset for model training, FL enables data to remain distributed across multiple sites during training, which is similar to other DML approaches. But unlike typical DML based on data parallelism or model parallelism, where each participant holds a portion of the entire dataset or focuses on a particular part of the entire model, FL does not require data exchange, and each participant trains a model on its local dataset.

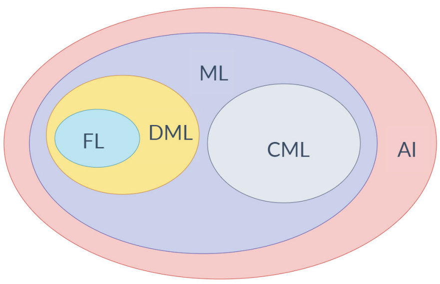

Fig. 1. Venn Diagram for FL, ML and AI

There have been numerous ML applications in healthcare, serving a variety of purposes, such as disease diagnosis and prediction [3], drug discovery [4], and personalized medicine [5]. While ML has the potential to revolutionize healthcare, there are challenges and potential problems that must be addressed to ensure successful and responsible implementation. For one thing, healthcare data are highly sensitive and subject to strict privacy regulations [6], [7], so using ML on patient data may raise concerns about data privacy and security breaches. For another, having access to a broad set of data is critical when developing ML models, particularly deep learning (DL) models. When data access is limited due to privacy rules and data owners' unwillingness to share their data, the available datasets may be smaller, more homogeneous, or lack adequate diversity. As a result, ML models trained on limited data may fail to reflect the entire complexity of healthcare scenarios, resulting in poor performance and incorrect predictions [8], [9].

Our study investigates the application of FL in the field of healthcare. Firstly, we introduce three types of FL frameworks and provide a potential healthcare scenario for each framework. Secondly, we outline the various kinds of data used in the investigated papers. Thirdly, based on whether the data used by the authors consists of image and text data or sensor data, we classify the current FL applications in healthcare. We conclude our investigation with a few ideas for future research directions.

II. Federated Learning

FL flameworks can be divided into horizontal federated learning, vertical federated learning, and federated transfer learning based on the way data are partitioned across different participants. Some of them inevitably require data alignment, so encryption or other privacy-preserving techniques must be employed in such cases.

A. Horizontal Federated Learning (HFL)

Datasets used in HFL have significant overlap in features but less overlap in sample IDs [10], as shown in Fig.2. The typical architecture and operation of HFL are shown in Fig. 3. A central server initializes and distributes a global model to participating clients. Each client trains a local model with the full feature space on its own dataset and sends the updated model parameters to the central server. After that, the server

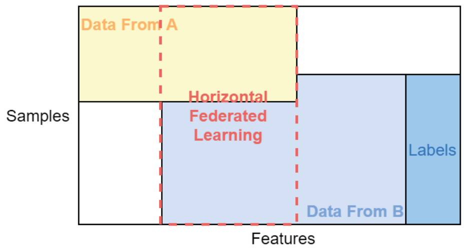

Fig. 2. Data Partition in HFL

aggregates these model updates based on a specific algorithm and updates the global model. These steps continue until certain requirements are met (e.g., the global model converges or time runs out), allowing the global model to benefit from distinct datasets without data exchange.

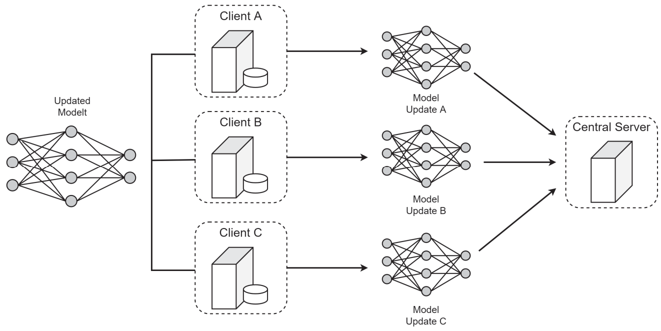

Fig. 3. Architecture and Operation of HFL

Generally speaking, the goal of HFL is to minimize the following objective function:

$$\min_{w} F(w), \text{where } F(w) := \sum_{k=1}^{m} p_k F_k(w)$$

Here:

m is the number of clients;

p_k is typically defined by the kth client, which specifies the client's relative impact; under the constrains of $p_k \geqslant 0$ and $\sum_k p_k = 1$, the natural settings are $p_k = \frac{1}{n}$ or $p_k = \frac{n_k}{n}$ [1], where $n = \sum_k n_k$ is the total number of data samples held by all participating clients;

F_k is the local objective function for the kth client, often defined as $F_k(w) = \frac{1}{n_k} \sum_{j_k=1}^{n_k} f_{j_k}(w; x_{j_k}, y_{j_k})$, where n_k is the number of data samples held by the kth client.

It is worth noting that the choice of local objective function needs to be determined by the specific algorithms (e.g., Logistic Regression (LR)) or ML models (e.g., Convolutional Neural Network (CNN)).

In a scenario of HFL for disease diagnosis, multiple hospitals possess diverse sets of patient data related to a specific disease and intend to develop a disease diagnosis model. Each hospital trains the local model on its own dataset and just needs to send model updates to the central server periodically.

B. Vertical Federated Learning (VFL)

Datasets used in VFL share a number of sample IDs but have little in common in the feature space, as shown in Fig. 4. VFL training differs from HFL training in that the features

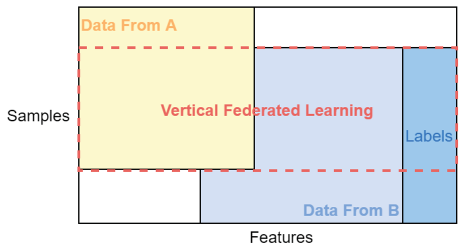

Fig. 4. Data Partition in VFL

of the same sample are distributed among various participants, and each participant cannot complete model training with its own dataset. To compute model updates, VFL requires two additional steps: aligning datasets with the same sample IDs and training models on these aligned datasets with privacy-preserving techniques, as demonstrated in Fig. 5.

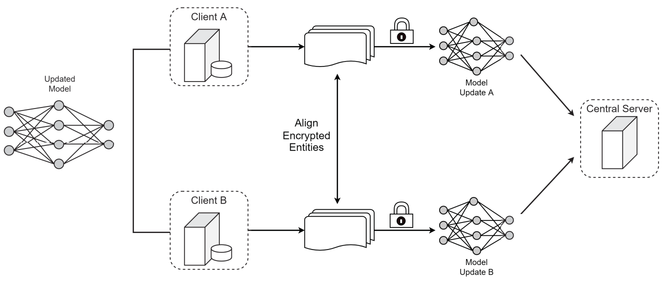

Fig. 5. Architecture and Operation of VFL

In a scenario of VFL regarding remote patient monitoring, wearable health devices and medical records are utilized to provide patients with customized medical care. To ensure that companies holding data from wearable devices and hospitals holding record data can align the same sample IDs, an encryption-based alignment technique [11], [12] is employed.

Secure multiparty computation with secret sharing is one implementation method for VFL [13]. Each participant shares their own features with the other participants using the secret sharing method at the beginning of the learning process. Hence, each participant possesses encrypted features in all dimensions. Then they use the encrypted features to train local models, which are later decrypted to obtain plaintext parameters.

C. Federated Transfer Learning (FTL)

FTL combines the methodology of transfer learning. Transfer learning typically involves two domains: the source domain, where the model is pretrained, and the target domain, where the model needs to be adapted or refined. A model is pretrained on a large dataset from the source domain, which functions as a knowledge base containing the learned general features and patterns from the source domain. Once the pretraining is complete, the model parameters are typically fine-tuned with data from the source domain to make the model appropriate for the target domain.

The basic idea of FTL is that the features learned during the source task can be useful for the target task, despite differences in data distribution, data scale, or context between the tasks. FTL is applicable in situations where datasets have limited overlap in both samples and features [14], as depicted in Fig. 6. But FTL does not have a fixed framework: FL can be used to train a model in the source domain before adapting to the target domain [15]; it can also be used to aggregate the models after adapting to the target domain [16].

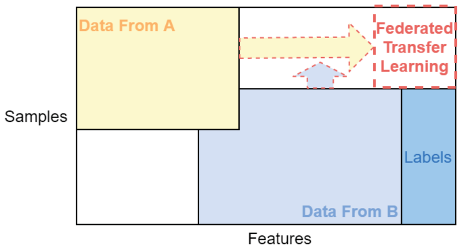

Fig. 6. Data Partition in FTL

In a scenario of FTL where two medical institutions in different cities hold different types of medical images, they can leverage each other's data to improve the accuracy of the model for disease detection and diagnosis as follows: each institution pretrains a local model on its own image dataset; then, they share a portion of their model's weights as knowledge transfer prior to building the global model.

D. Benefits and Challenges

FL has emerged as a promising approach for various healthcare applications. For one thing, it enables model training on decentralized data sources without sharing raw data, which preserves patient privacy. For another, since the size of model updates is typically much smaller than that of an actual dataset, FL helps to reduce communication costs associated with raw data transmission (e.g., latency and transmit power [17]–[20].

However, FL still faces challenges [17], [19], [21], including:

1) Statistical Heterogeneity: Data distributions at each client are likely to differ, and data quality is not guaranteed, resulting in poor global model performance.
2) Security Issues: Assuming that all of the clients are trustworthy is impossible, so additional privacy-preserving techniques are necessary to protect medical data from untrustworthy clients or third-party attackers.
3) Real-time Data Stream: real-time data produced by edge devices such as wearables is typically fast, massive, and dynamic, and must be handled in real time.
4) Client Selection: FL requires selecting eligible clients to participate in the training process based on some criteria. This can be challenging when the clients have limited resources or when the network is large. It also requires selecting clients in a fair and unbiased manner to ensure that all clients have an equal opportunity to participate.

III. Literature Analysis

A. Research Methodology

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) has become a widely acknowledged standard for presenting evidence from systematic reviews in health-related organizations and journals. Although our study is more of an extensive investigation than a systematic review, its research methodology is still guided by PRISMA.

B. Research Sources

The literature search is based on a variety of databases, including IEEE Xplore, ACM Digital Library, ScienceDirect, Springer, Web of Science, PubMed, and JMIR.

C. Inclusion and Exclusion Criteria

1) Publication Date: Our study includes articles published from 2018 to July 2023. Although FL was proposed in 2016, we observed that research on the application of federated learning in the healthcare field has only begun to grow significantly since 2018.
2) Research Type: Only journal articles and conference proceedings were included, while conference abstracts only, books, editorials, and commentaries were not.
3) Search Keywords: The search was carried out using keyword combinations such as "federated learning" AND ("healthcare" OR "medical*").
4) Initial Analysis: The title and abstract help in making an initial screening to exclude those that are not directly relevant to the intersection of FL and healthcare. Studies on COVID-19 were excluded because the common datasets used in these studies remain rooted in the past compared to the rapidly mutating virus and the growing variety of symptoms, which may make these studies, particularly the early ones, lack a high degree of adaptability.
5) Duplicate Removal: Papers from various sources were integrated, and duplicates were removed.

Based on PRISMA and the inclusion and exclusion criteria, 35 articles were included in our investigation.

IV. FL Applications in Healthcare

Disease diagnosis and prognosis are essential aspects of healthcare, which involve analyzing medical data, such as clinical records, imaging scans, laboratory results, and other relevant data sources, to determine the presence of a particular disease or predict the likelihood of its occurrence in the future. In addition, remote monitoring plays a significant role in disease diagnosis and prediction by enabling continuous and real-time monitoring of patients outside of traditional healthcare settings.

ML empowers healthcare professionals with advanced tools to accurately diagnose diseases and predict their occurrence.

Manual analysis of medical data is time-consuming and error-prone in traditional procedures. ML provides an engine for identifying and utilizing complex patterns, correlations, and features in data. However, while conventional ML techniques have made significant progress in healthcare, FL offers unique advantages that can amplify the accuracy, robustness, and ethical considerations of disease diagnosis and prognosis.

Our research categorizes FL healthcare applications into two primary categories based on the nature of the training data they use: image and text data, and sensor data.

1) **Image and text data** can be derived from the sources listed in Table I, which appear frequently in the papers we examined.

TABLE I. EXAMPLES FOR IMAGE AND TEXT DATA

Name	Description
Whole Slide Image (WSI) [22], [23]	a high-resolution digital scan of entire glass slides comprising biopsied tissue samples
Cancer Genome Atlas (TCGA) [24], [25]	a large-scale, collaborative research endeavor launched in the US by the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) to better understand the molecular basis of cancer
Magnetic Resonance Imaging (MRI) [26]	a common and non-invasive medical imaging technique that employs a combination of strong magnetic fields and radio waves to facilitate the in-depth visualization of internal structures within the human body
Electrocardiogram (ECG or EKG) [27], [28]	a medical test that measures the heart's electrical activity and can be used to evaluate the heart's rhythm, rate, and overall condition
Electroencephalogram (EEG) [29]	a medical test that measures the brain's electrical activity and can be used to diagnose various neurological conditions and monitor brain activity during various states
Electronic Health Record (EHR) [30], [31]	a digitalized compilation of a patient's medications,laboratory results, imaging reports, and clinical documentation

2) **Sensor data** are derived from various sensors embedded in medical devices, wearables, or other equipment that monitor specific physiological or environmental parameters and collect real-time data from patients, individuals, or the surrounding environment. In this case, sensor data are typically numerical measurements and readings and do not include image or text data.

Within each of these two categories, many studies have been conducted to examine the application of FL in different use cases.

A. Applications using Image and Text Data

*1) **Cancer**:* Cancer is a complex group of diseases with diverse manifestations and outcomes. Cancer and tumors are closely related. Tumor is a broad medical term that refers to any abnormal lump or mass of tissue. Tumors can be benign (non-cancerous) or malignant (cancerous) and cancer is a specific type of malignant tumor. The term "cancer" typically refers to a group of diseases characterized by the presence of malignant tumors. These malignant tumors can occur in various parts of the body. Tumor segmentation involves outlining the exact boundaries or contours of a tumor within an image (e.g., MRI). This allows medical professionals to visualize the size, shape, and location of the tumor.

1) **Brain Cancer**
Brain tumor segmentation is a critical component of brain cancer detection, diagnosis, treatment, and research.
Sheller et al. [32] first introduced FL to multi-institutional collaboration. For comparison, the authors use models based on institutional incremental learning (IIL), where each institution trains the model and then passes it to the next institution for training until all institutions have trained once, and CIIL (cyclic IIL), where IIL is performed in rounds with prescribed numbers of epochs. The experimental results show that FL models reach a performance comparable to models based on CML and outperform models based on IIL and CIIL approaches.
Differential privacy is a mathematical framework that enables the analysis of data without disclosing sensitive information by adding noise to the data in a way that preserves the overall statistical properties of the dataset while obscuring the data contribution. Li et al. [33] implement the first FL system for medical image analysis that protects patient data with differential privacy. The experimental results show a tradeoff between model performance and protection costs. The authors finally conclude that, even with a robust differential privacy assurance, the allocation of privacy costs is conservative.
Bercea et al. [34] propose the Federated Disentanglement (FedDis) for the collaborative training of an unsupervised deep convolutional autoencoder on MRI scans from four distinct institutions. These institutions train the parameters of shape to model the anatomical structures of a healthy brain. The experimental results show that FedDis improves anomaly segmentation results for tumors over locally trained models.

2) **Lung Cancer**
Rajendran et al. [35] focus on predicting the risk of tobacco- and radon-related diseases, both of which are closely associated with lung cancer and chronic obstructive pulmonary disease. The experimental results show that applying the FL process did not improve the LR model's performance as much as that of artificial neural network models, due to the LR algorithm's lower complexity and lack of iterative training.
Adnan et al. [36] combine FL with differential privacy to develop models for medical image analysis. Specifically, the training process for each local model consists of two steps: bag preparation and multiple-instance learning. Lung cancer images from the TCGA dataset are used to build a simulated environment to validate the approach. According to the experimental results, the suggested method achieves about the same level of performance as CML with extra privacy protection.

3) **Breast Cancer**
Along with gigapixel WSIs from several institutions, Lu et al. [37] use FL, multiple instance learning, and differential privacy to help with breast cancer histological subtyping classification tasks. The experimental results show that models trained with the proposed FL technique perform as well as or better than models based on CML.

Classifying breast density helps determine how much fibroglandular tissue increases breast cancer risk [38]. Roth et al. [39] employ FL to build a breast density classification model based on medical imaging and the Breast Imaging Reporting & Data System proposed by the American College of Radiology in 1986. Seven clinical institutions from around the globe participated in the experiments, and the results show that FL models perform better on average than models trained using only local data from one institute.

4) **Liver Cancer**
Hepatocellular carcinoma (HCC) detection is a typical component of liver pathology image analysis. However, HCC detection might fail when patches cover a small tissue area without enough information about the surrounding cell structure, since tumor and benign liver tissue exhibit similar apoptosis, necrosis, and steatosis. To address this issue, Yang et al. [40] propose the Feature Aligned Multi-Scale Convolutional Network architecture, which is based on WSIs and integrates elements from different magnification levels to reference additional surrounding information.

5) **Prostate Cancer**
Cross-client variation in medical image data poses a significant challenge for practical applications. To mitigate this issue, Yan et al. [41] propose a variation-aware federated learning framework where the variations among clients are minimized by transforming images of all clients onto a common image space with a privacy-preserving generative adversarial network. The experimental results indicate that models trained with the proposed framework perform comparably to those trained with CML and better than models trained on single datasets or even with typical HFL.

6) **Skin Cancer**
Cai et al. [42] show a way to find skin cancer using FL and a deep generation model called DualGANs (which is used to deal with the issue of incomplete data). The authors test and compare how well the proposed model works in a number of different situations, including IID and non-IID data, as well as fully connected and sparse convolutional neural networks. The experimental results show that the proposed method attains a high skin cancer detection rate with high accuracy.

*2) **Neurological Disorders**:* Neurological disorders are diseases of the central and peripheral nervous systems, which include stroke, Parkinson's disease, and Alzheimer's disease.

1) **Stroke**
A stroke occurs when the blood supply to a portion of the brain is interrupted or when a blood vessel in the brain ruptures. Several studies on the improvement of stroke diagnosis using ML have been undertaken during the last few decades [43].
WeBank and Tencen's tech experts created and implemented a federated version of a stroke prediction model, based on Tencen's previous work on the model, WeBank's self-developed privacy-preserving framework, and FATE (an industrial-level open-source secure computing framework) [44]. This is the second implementation of WeBank's new privacy-protecting technology, which could help millions of people prevent strokes.
Victor et al. [45] present FL-PSO, a FL-based system for brain stroke prediction that employs particle swarm optimization for model optimization. The experimental results show that selecting the best hyperparameters for global model training can increase accuracy.

2) **Parkinson's disease**
Dipro et al. [46] use an open-access dataset and biosample library of Parkinson's disease, such as single-photon emission computed tomography and MRI, to train three types of CNN models (VGG19, VGG16, and InceptionV3) for the detection of Parkinson's disease. The experimental results show that FL with VGG19 has the highest accuracy.
There have been data-driven computational methods relying on a large number of high-quality clinical assessments. Reyes et al. [47] investigate the data imputation and reconstruction of clinical scores from the Parkinson Progression Marker Initiative. They also compare the performance of two aggregation algorithms: FedAvg and precision-weighted FL. The experimental results show that the former provides more precise reconstruction, whereas the latter is better suited to handle data heterogeneity.

3) **Alzheimer's Disease**
Huang et al. [48] propose Federated Conditional Mutual Learning (FedCM), a framework for federated mutual distillation. Their work is the first to apply FL to the classification of Alzheimer's disease. FL with knowledge distillation [49] uses all available data without disclosing local private data. This method was previously employed by Federated Learning via Model Distillation (FedMD) [50]. FedCM's authors argue that, despite FedMD's current success, particularly on synthesized datasets, there are challenges in applying the framework to actual medical applications. FedCM considers clients' local performance and similarity, enabling client-aware mutual learning. The experimental results show that FedCM performs better than FedMD.

*3) **Dermatological Diseases**:* Dermatological diseases, also referred to as skin diseases, are a broad range of conditions affecting the skin.

Dermatology medical images are susceptible to attacks during transmission, which will result in malicious tampering or privacy data disclosure. To address the security issue, Han et al. [51] propose an FL-based robust zero-watermarking scheme. The experimental results reveal that the proposed scheme is more resistant to conventional and geometric attacks than six other zero-watermarking schemes. The proposed scheme is suitable for specific requirements of medical images, as it neither modifies important information contained in medical images nor discloses personal information.

Multiple FL studies presume that all data are labeled, which is impractical in many situations in the real world. Furthermore, the lack of labels renders supervised FL implausible. Wu et al. [52] propose Federated Contrastive Learning (FCL), an on-device framework for diagnosing dermatological diseases that takes into account inadequate data labeling. Specifically, FCL initializes the model using distributed unlabeled data and then conducts disease diagnosis using a limited number of labeled data. The experimental results show that FCL outperforms the other four pre-training baseline models (Random init, Local CL, Rotation, and SimCLR).

Elayan et al. [53] propose a framework to train DL models for skin disease detection, which employs FL and transfer learning to address the problem of limited healthcare data availability. They also propose an algorithm for the automated acquisition of training data. One of the Keras application DL models based on CML, ResNet50, was used to initialize the FL global model and as a baseline for comparison. The experimental results show that the proposed FL method achieves generally better performance and is more privacy-preserving than the basic ResNet50 model, but the increased model conversion time may compromise the quality of service to the user.

*4) **Cardiovascular Diseases**:* Cardiovascular diseases are a group of disorders of the heart and blood vessels. Arrhythmia is one type of cardiovascular disease that refers to an abnormal heart rhythm. ECGs provide extensive information about the cardiac rhythm and are vital to clinical treatment. Zhang et al. [54] propose a FL-based arrhythmia detection algorithm for auxiliary diagnosis and therapy. Since ECG data collected from different medical institutions are typically non-IID, which may lead to the non-convergence of FL-based algorithms, the authors optimize their algorithm by combining partial ECG data from each medical institution. Compared to baseline algorithms such as FedAvg [1] and FedCurv [55], their algorithm obtains significant improvements on non-IID ECG.

Raza et al. [56] design a federated healthcare framework with ECG, explainable artificial intelligence (XAI) and CNN. Specifically, CNN-based autoencoders and classifiers sort arrhythmias into different groups; an XAI-based module analyzes classification results and helps clinical practitioners make quick and reliable decisions. The experimental results show that the proposed classifier outperforms existing arrhythmia detection methods using either noisy or pristine data.

*5) **Autism Spectrum Disorder (ASD)**:* Autism spectrum disorder is a term used to describe individuals with early-appearing social communication deficits, repetitive sensory-motor behaviors, and/or highly restricted interests that are associated with a strong genetic component and other causes [57].

Functional Magnetic Resonance Imaging (fMRI) measures the small changes in blood flow that occur with brain activities [58]. Li et al. [59] focus on identifying ASD based on resting-state fMRI data from the Autism Brain Imaging Data Exchange dataset. They employ a privacy-preserving FL technique that uses a randomization mechanism to alter shared local model weights. Besides, they propose two domain adaptation methods, considering systemic differences in fMRI distributions at various sites. The experimental results show that it is promising to use multi-site data without data sharing to improve the performance of neuroimage analysis and identify reliable disease-related biomarkers.

*6) **Psychiatric Disorders**:* Psychiatric disorders refer to a broad range of problems that disturb a person's thoughts, feelings, behavior, or mood.

1) **Depression**
   Depression is a common disease in the real world. Currently, the diagnosis of depression relies almost exclusively on the opinions of the physician and is determined through communication with the patient and relevant questionnaires.
   To train a sophisticated DL model, large and diverse patient data are required. However, DL models trained on restricted datasets have poor clinical performance in a new location with different data. Ahmed et al. [60] propose a method for extracting depression symptoms from text based on FL, Natural Language Processing (NLP), and attention-based learning. The experimental results indicate that FL has practical advantages over traditional supervised learning methods.
   Chhikara et al. [61] propose a FL-based schema that derives features from images and audio for automatic emotion recognition, which can detect depression for an individual at an earlier stage and recommend that the individual consult with a therapist. The classifier for facial expression recognition is a CNN-SVM model. The experimental results show that applying typical HFL to the CNN-SVM classifier can enhance the classifier's performance by achieving greater accuracy.

2) **Inpatient Violence**
   Inpatient violence is a serious issue in psychiatry that concerns hospital staff and patients. Knowing who is likely to become violent can affect personnel levels and reduce severity. Borger et al. [62] examine the application of FL and NLP to violence risk assessment. The experimental results show that the FL model outperforms local models and is comparable to the model based on CML, implying that FL can be applied successfully in a cross-institutional scenario and support novel clinical note-based applications.

*7) **Adverse Drug Reactions**:* An adverse drug reaction is a significantly harmful or unpleasant reaction resulting from an intervention related to the use of a medicinal product. It can be used to predict the targeted treatment, dosage regimen adjustment, or product withdrawal. [63].

Choudhury et al. [64] propose a FL-based method to build a global ADR prediction model using decentralized health data from multiple local sites. To demonstrate the effectiveness of their strategy, they predict the chronic use among opioid drug users and extrapyramidal symptoms among antipsychotic drug users. In addition, they present two new local model aggregation techniques to improve the global model's prediction. For two types of adverse drug reactions, the experimental results show that the proposed FL approach obtains comparable performance to models based on CML and outperforms models based on local learning.

*8) **Predict Hospitalization**:* FL with clinical data, including EHRs, holds promise for enhancing mortality and hospital stay predictions. EHRs provide a wealth of patient data, including medical history and demographics, enabling FL to collaboratively build accurate prediction models while respecting data privacy.

Brisimi et al. [65] develop the iterative cluster Primal Dual Splitting (cPDS) algorithm to solve the sparse Support Vector Machine (SVM) problem. They used cPDS on heart disease patients' de-identified electronic heart records from the Boston Medical Center. Demographics, diagnoses, admissions, and other medical histories are combined to define each patient. The experimental results show that cPDS can predict hospitalization within a specified year, and that cPDS improves convergence rates and lowers communication costs compared to traditional CML and DML solutions.

Huang et al. [66] propose the Community-based Federated Machine Learning (CBFL) algorithm to address the problem of non-IID ICU patient data. CBFL clusters EMR data into multiple communities and simultaneously trains one model per community, resulting in a more efficient learning process. In both mortality and stay time prediction tasks, the experimental results show that CBFL converges to higher predictive accuracy in fewer communication rounds than the typical HFL model.

B. Applications using Sensor Data

Internet of Medical Things (IoMT) refers to the network of interconnected medical devices, sensors, wearable devices, and healthcare systems that collect, transmit, and exchange medical data and information over the Internet. FL plays a crucial role in IoMT, especially in monitoring and data analysis. This segment focuses on leveraging FL's capabilities to support medical diagnoses and addressing some of the challenges encountered by FL-based applications within the IoMT domain.

*1) **Support Medical Diagnoses**:*

1) **Arterial Blood Pressure (ABP)**

Doctors employ ECGs and blood pressure measures to understand heart health. Given that continuous ABP monitoring is invasive and expensive, Brophy et al. [67] propose a federated framework that infers ABP from one optical photoplethysmogram (PPG) sensor (a simple non-invasive optical technique used to detect volumetric changes of blood in peripheral circulation). The study compares the proposed framework with a non-federated framework, both of which employ the same type of ML model. The experimental results show that the federated framework's models have slightly degraded performance compared to models trained using the non-federated framework. Since the federated framework has the advantage of preserving data privacy by not requiring data centralization, it still shows promising results in modeling ABP.

2) **Depression**

Xu et al. [68] implement FL to analyze and diagnose depression using data from a study based on BiAffect, a free mobile application. The authors develop a general multiview FL framework that employs multi-source data. They also fix conflicting multiview data time series with later fusion techniques. The experimental results show that FL with enough participants results in greater accuracy for depression prediction than local training.

3) **Daily Activities**

Daily activities have a significant impact on people's health. Recent advancements in wearable technologies assist individuals in understanding their health status through activity monitoring with devices such as smartphones, wristbands, and smart glasses.

Chen et al. [15] propose FedHealth, the first FTL framework for wearable healthcare. FedHealth aggregates data with FL and then creates relatively personalized models via transfer learning. Transfer learning allows for the use of pre-trained models and knowledge from other domains to improve the accuracy and efficiency of the models. This is especially useful in scenarios where user data is limited or difficult to obtain. Wearable activity recognition experiments and real Parkinson's disease auxiliary diagnosis applications have evaluated that FedHealth performs better than traditional methods (K-Nearest Neighbor, Random Forest, and SVM) trained on single datasets.

Fan et al. [69] propose the Federated Learning Driven IoMT (FLDIoMT) framework that facilitates flexible deployment of IoMT services while addressing privacy and security concerns. FLDIoMT implements iSmile, a system providing sleep monitoring and emotion-aware services to support psychological analysis. The experimental results show that the model accuracy of the

FL mechanism is close to that of the CML mechanism and higher than that of local training. The study also discusses the potential of the FL mechanism to achieve a more significant increment in accuracy as the number of local clients increases.

Liu et al. [70] apply ML to the publicly available Wearable Stress and Affect Detection dataset to predict individual stress levels and emotional states. The study examines the effectiveness of various ML models, including personalized and nonpersonalized FL models and server models (trained on a centralized dataset maintained on the server). A personalized model is an ML model tailored to the distinctive features of each user. Specifically, personalized models are constructed by adding user embeddings to the neural network, which encode user-specific variables and enable the model to learn from each user's input separately and produce personalized predictions. The experimental results show that the personalized server model and the personalized federated model have the highest accuracy, whereas the nonpersonalized server model and the nonpersonalized federated model have lower accuracy.

*2) **Address Technical Challenges**:* The research fields of the following studies are at the intersection of medical health and FL, but their focus is on solving certain technical challenges, particularly those likely to be encountered when FL is combined with IoT.

1) **Imbalanced and Non-IID Data**

Due to the diversity of sensor devices and their limited storage capacity, the size of a single local dataset will be small while the data types can be broad.

Wu et al. [71] propose FedHome, a cloud-edge FL framework for personalized in-home health monitoring. FL enables participants to train anomalous health detection models in collaboration while keeping health data private. FedHome further uses a generative convolutional autoencoder to address the issue of uneven and non-IID data. The study evaluates the performance of FedHome against several other models, including those based on traditional CML and typical FL. The experimental results show that FedHome outperforms other models in terms of test accuracy and communication overhead.

2) **Class Imbalance**

Zhang et al. [72] propose FedSens, a FL framework that can address the challenge of class imbalance. FedSens employs both a new local update scheme inspired by the curiosity-driven reinforcement learning model and an adaptive global update scheme using online regret minimization. These two parts allow each edge device to choose the best local and global update strategy. This can make the AHD model more accurate when there is a class imbalance. The experiments on the stress detection application show that FedSens achieves a higher level of performance than the best-performing baseline, Astraea, which employs a weighted loss function to resolve the class imbalance problem.

3) **Willingness to Participate (WTP)**

Aside from privacy concerns, varying levels of WTP are a source of concern. In an FL system, some users may be more intrinsically motivated to contribute if the development of an application directly benefits them, whereas others may intend to take advantage of others' efforts [73], [74]. Moreover, many IoT-based healthcare applications rely on longitudinal data collected over an extended period of time [75], but the WTP of a user might change over time. Therefore, an incentive scheme should be developed to guarantee the long-term participation of users with dynamic WTP values in the presence of information asymmetry.

In light of this, Lim et al. [76] propose a dynamic contract design for the FL network that addresses the challenges of WTP. A self-revealing mechanism of contract design ensures that each user only chooses the contract that is designed for its type. By providing appropriate incentives that match their WTP, users can be incentivized to participate in the training process. The network also uses a profit function, a mathematical representation of the model owner's profit derived from the FL-based collaborative model training process, and takes into account model accuracy, data quantity, and contract rewards. The experimental results show that the dynamic contract scheme can lead to higher profit than the uniform pricing scheme (a fixed data quantity-contract reward bundle offered to all users).

4) **Constrained Resources**

Given the growing size and complexity of current neural network models, training models on wearable devices with limited resources becomes inefficient, if not impossible.

Guo et al. [77] propose the Federated Edge Learning (FEEL) system that boosts training efficiency with an edge-based training task offloading strategy and strengthens privacy protection through differential privacy. Models based on CML and local training are used as baselines in the experiments to evaluate the performance of FEEL. The experimental results show that the CML model has optimal and stable performance; the performance of FL is comparable to that of CML while satisfying privacy requirements; the average performance of models trained on local datasets is the worst and fluctuates greatly with different data distributions.

Table II shows the a summary of the papers examined in this study.

TABLE II. LITERATURE SUMMARY

No.	Ref.	Research Area	Training Data	FL Framework	Contributions
1	[40]	HCC; liver tumor	WSI	HFL	Mitigate the challenge of HCC detection failure caused by similarities between tumor and benign liver tissue, especially in cases where patches cover a limited tissue region without adequate surrounding cell structure information.
2	[37]	breast cancer; renal cell cancer	WSI	HFL	Combining FL and weakly supervised multiple instance learning; Consider the challenges of the lack of detailed annotations.
3	[41]	prostate cancer	ADC image	HFL	Mitigate the challenge posed by cross-client variation in medical image data and outperform HFL frameworks for the automated classification of prostate cancer.
4	[35]	lung cancer; COPD	EHR	HFL	Investigate and evaluate several FL implementations that predict the risks of tobacco and radon related diseases.
5	[36]	cancer	TCGA	HFL	Combining FL and DP; Investigate effects from IID and non-IID data, the number of healthcare providers, and dataset sizes; Achieve comparable performance to conventional training and offer a high privacy guarantee.
6	[42]	skin cancer	skin image	HFL	Evaluate the performance of the proposed model under several conditions: IID/non-IID data and sparse/fully connected CNNs; Attain a high skin cancer detection rate.
7	[39]	breast density	BI-RADS	HFL	Show that the FL model performs better on average than models trained with only local data from an institute.
8	[32]	brain tumor	MRI	HFL	Present FL for multi-institutional collaboration for the first time; Achieve comparable performance to models trained based on sharing data.
9	[33]	brain tumor	MRI	HFL	Combining FL and DP; Implement the first privacy-preserving FL system for medical image analysis.
10	[34]	brain tumor	MRI	HFL	Improve anomaly segmentation results of locally trained models; particularly helpful for institutes with both healthy and anomaly data.
11	[48]	Alzheimer's disease	MRI	HFL	Attain the highest recognition rates compared to FedM and some other frameworks.
12	[47]	PD	PPMI	HFL	Evaluate two types of aggregation algorithms in distributed learning environments for data imputation and reconstruction of clinical assessment.
13	[44]	stroke	EHR; prescription; inspection	HFL	The second implementation of the new privacy-preserving technology of WeBank.
14	[45]	stroke	text	HFL	Indicate that selecting the best hyperparamete for global model training can increase accuracy.
15	[51]	dermatological diseases	dermatology medical image	HFL	Provide an FL-based robust zero-watermarking scheme; Address privacy and security issues; being more resistant to conventional and geometric attacks than previous zero-watermarking schemes.
16	[52]	dermatological diseases	dermatology medical image	HFL	Consider the problem of insufficient data labeling; have a greater diagnostic accuracy than current methods in experiments on dermatological disease datasets with diverse skin colors .
17	[53]	dermatological diseases	dermatology medical image	FTL	Use transfer learning to overcome the challenge of restricted healthcare data availability; have higher AUC, accuracy, precision, recall, and F1 score.
18	[54]	arrhythmia	ECG	HFL	Combining FL, partial ECG and elastic weight consolidation; Obtain significant improvements in recall and precision for non-IID ECG.
19	[56]	arrhythmia	ECG	HFL	Outperform existing detection methods based on either noisy or pristine data.
20	[59]	ASD	fMRI	HFL	Provide privacy-preserving FL with a randomization mechanism to alter shared model updates; provide two domain adaptation methods considering systemic differences use multi-site data without sharing to improve the performance of neuro image analysis and identify reliable disease-related biomarkers.
21	[60]	depression	text	HFL	Combining FL, NLP and attention-based learning; propose a method that extracts depression symptoms from text; show that FL has practical advantages over traditional supervised learning methods.
22	[61]	depression	image; audio	HFL	Propose a schema extracting features from images and audio to solve the problem of automatic emotion recognition for a specific individual.
23	[62]	inpatient violence	text	HFL	The the federated model outperforms local models and is comparable to data-centralized models; imply that FL can supports novel clinical note-based applications.
24	[65]	hospitalization	EHR	HFL	Predict hospitalization within a specified year, solve the sparse SVM problem and improves convergence rates and communication costs.
25	[66]	hospitalization	EHR	HFL	Address the issue of non-IID ICU patient data and outperform basic FL in predicting mortality and ICU stay time.
26	[64]	adverse drug reaction	EMR	HFL	Introduce an FL-based strategy to build a global ADR prediction model from decentralized data; the aggregation methods outperform state-of-the-art techniques.
27	[67]	ABP	sensor data	HFL	The first example of a GAN with continuous ABP generation from an input of PPG signal from a sensor; prevent invasive and expensive monitoring.
28	[68]	depression	sensor data	HFL	Develop a general multiview FL framework; fix the conflicting multiview data time series with later fusion techniques; show that FL with enough participants predicts better depression scores than local training.
29	[15]	PD	sensor data	FTL	The first FTL framework for wearable healthcare.
30	[69]	sleep and emotion	sensor data	HFL	Provide sleep monitoring and emotion-aware services to support the psychological analysis; facilitate flexible deployment of IoMT services while addressing privacy and security concerns.
31	[70]	mental state	sensor data	HFL	Show that the FL model performs better than models based on data-centralized training and individualized training in state prediction for experiment subjects.
32	[71]	statistical heterogeneity	sensor data	HFL	Provide personalized in-home health monitoring; Solve the imbalanced and non-IID data problem.
33	[72]	statistical heterogeneity	sensor data	HFL	FedSens enhances the accuracy of AHD models in the presence of a severe class imbalance while reducing the energy consumption of edge devices.
34	[76]	incentive mechanism	sensor data	HFL	Provide a system with a contract-theoretic incentive mechanism that encourages users to participate in FL-based collaborative model training.
35	[77]	constrained resource	sensor data	HFL	Have higher training efficiency, better inference performance and stronger privacy protection for training tasks on resource-constrained wearable devices.

V. Conclusion

Our study analyzes current research on the implementation of FL in the healthcare domain and emphasizes comparative experiments and their outcomes to demonstrate the advances of FL applications in this domain.

As a result of the diverse study emphases shown in prior studies, several potential research areas can be identified, which are as follows:

1) Healthcare datasets are highly diverse, including images, EHRs, genomic data, and more. The diversity of data within the healthcare domain requires researchers to investigate various FL frameworks that correlate with the specific dataset types. In addition, the effectiveness of FL methods varies depending on sample size and the characteristics of the datasets; for example, HFL often performs well with large datasets, while FTL can adapt to fewer, more specialized datasets. Given that the majority of current studies are based on HFL, future research can utilize various FL frameworks to develop flexible FL healthcare applications to enhance diagnosis and prognosis support.
2) The number of investigated studies on some diseases is significantly higher than on others, though this is partly due to the availability of training data, with healthcare research heavily depending on the availability of high-quality medical data. For diseases like cancer, data are often more comprehensive and accessible due to the frequency of cases and extensive research history. In contrast, diseases like Alzheimer's may have limited datasets available for study. Ensuring sufficient and reliable training data is essential for conducting meaningful research on any disease. If conditions are sufficient, future research can investigate more types of diseases.
3) Many researchers have proposed solutions to FL's challenges outside of the healthcare domain, but these solutions are not always applicable to the healthcare domain. Consequently, it is necessary to develop solutions that address challenges specific to the healthcare sector, such as regulatory compliance, interoperability, and clinical relevance. While some general solutions from other domains may provide inspiration, the distinctive nature of healthcare requires research efforts to tackle the challenges associated with FL's application in healthcare.

References

[1] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *International Conference on Artificial Intelligence and Statistics*, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:14955348

[2] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. B. Charles, G. Cormode, R. Cummings, R. G. L. D'Oliveira, S. Y. E. Rouayheb, D. Evans, J. Gardner, Z. Garrett, A. Gascón, B. Ghazi, P. B. Gibbons, M. Gruteser, Z. Harchaoui, C. He, L. He, Z. Huo, B. Hutchinson, J. Hsu, M. Jaggi, T. Javidi, G. Joshi, M. Khodak, J. Konecný, A. Korolova, F. Koushanfar, O. Koyejo, T. Lepoint, Y. Liu, P. Mittal, M. Mohri, R. Nock, A. Özgür, R. Pagh, M. Raykova, H. Qi, D. Ramage, R. Raskar, D. X. Song, W. Song, S. U. Stich, Z. Sun, A. T. Suresh, F. Tramèr, P. Vepakomma, J. Wang, L. Xiong, Z. Xu, Q. Yang, F. X. Yu, H. Yu, and S. Zhao, "Advances and Open Problems in Federated Learning," *Found. Trends Mach. Learn.*, vol. 14, pp. 1–210, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:209202606

[3] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," *New England Journal of Medicine*, vol. 380, no. 14, pp. 1347–1358, 2019, publisher: Mass Medical Soc.

[4] J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer, and others, "Applications of machine learning in drug discovery and development," *Nature reviews Drug discovery*, vol. 18, no. 6, pp. 463–477, 2019, publisher: Nature Publishing Group UK London.

[5] S. Zhang, S. M. H. Bamakan, Q. Qu, and S. Li, "Learning for personalized medicine: a comprehensive review from a deep learning perspective," *IEEE reviews in biomedical engineering*, vol. 12, pp. 194–208, 2018, publisher: IEEE.

[6] P. Voigt and A. Von dem Bussche, "The eu general data protection regulation (gdpr)," *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, vol. 10, no. 3152676, pp. 10–5555, 2017, publisher: Springer.

[7] K. Theodos and S. M. Sittig, "Health Information Privacy Laws in the Digital Age: HIPAA Doesn't Apply." *Perspectives in health information management*, vol. 18 Winter, p. 1l, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:232058362

[8] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: review, opportunities and challenges," *Briefings in bioinformatics*, vol. 19 6, pp. 1236–1246, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:2740197

[9] A. Qayyum, J. Qadir, M. Bilal, and A. I. Al-Fuqaha, "Secure and Robust Machine Learning for Healthcare: A Survey," *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 156–180, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:210859156

[10] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 2, pp. 1–19, 2019, publisher: ACM New York, NY, USA.

[11] G. Liang and S. S. Chawathe, "Privacy-Preserving Inter-database Operations," in *Intelligence and Security Informatics*, 2004. [Online]. Available: https://api.semanticscholar.org/CorpusID:1986261

[12] M. Scannapieco, I. Figotin, E. Bertino, and A. K. Elmagarmid, "Privacy preserving schema and data matching," in *ACM SIGMOD Conference*, 2007. [Online]. Available: https://api.semanticscholar.org/CorpusID:12585496

[13] H. Shi, Y. Jiang, H. Yu, Y. Xu, and L. Cui, "MVFLS: Multi-participant Vertical Federated Learning based on Secret Sharing."

[14] S. Saha and T. Ahmad, "Federated Transfer Learning: concept and applications," *Intelligenza Artificiale*, vol. 15, pp. 35–44, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:225103353

[15] Y. Chen, X. Qin, J. Wang, C. Yu, and W. Gao, "FedHealth: A Federated Transfer Learning Framework for Wearable Healthcare," *IEEE Intelligent Systems*, vol. 35, no. 4, pp. 83–93, Jul. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9076082/

[16] C. Ju, D. Gao, R. Mane, B. Tan, Y. Liu, and C. Guan, "Federated Transfer Learning for EEG Signal Classification," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. Montreal, QC, Canada: IEEE, Jul. 2020, pp. 3040–3045. [Online]. Available: https://ieeexplore.ieee.org/document/9175344/

[17] M. Joshi, A. Pal, and M. Sankarasubbu, "Federated Learning for Healthcare Domain - Pipeline, Applications and Challenges," *ACM Transactions on Computing for Healthcare*, vol. 3, no. 4, pp. 1–36, Oct. 2022, arXiv:2211.07893 [cs]. [Online]. Available: http://arxiv.org/abs/2211.07893

[18] B. Pfitzner, N. Steckhan, and B. Arnrich, "Federated Learning in a Medical Context: A Systematic Literature Review," *ACM Transactions on Internet Technology*, vol. 21, no. 2, pp. 1–31, Jun. 2021. [Online]. Available: https://dl.acm.org/doi/10.1145/3412357

[19] H. Li, C. Li, J. Wang, A. Yang, Z. Ma, Z. Zhang, and D. Hua, "Review on security of federated learning and its application in healthcare," *Future Generation Computer Systems*, vol. 144, pp. 271–290, Jul. 2023. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167739X23000626

[20] A. Rahman, M. R. Hossain, G. Muhammad, D. Kundu, T. Debnath, M. Rahman, M. S. I. Khan, P. Tiwari, and S. S. Band, "Federated

learning-based AI approaches in smart healthcare: concepts, taxonomies, challenges and open issues," *Cluster Computing*, pp. 1 – 41, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:251659681

[21] T. Wang, Y. Du, Y. Gong, K.-K. R. Choo, and Y. Guo, "Applications of Federated Learning in Mobile Health: Scoping Review," *Journal of Medical Internet Research*, vol. 25, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:257449260

[22] L. X. Pantanowitz, P. Valenstein, A. J. Evans, K. J. Kaplan, J. D. Pfeifer, D. C. Wilbur, L. C. Collins, and T. J. Colgan, "Review of the current state of whole slide imaging in pathology," *Journal of Pathology Informatics*, vol. 2, 2011. [Online]. Available: https://api.semanticscholar.org/CorpusID:11787374

[23] N. Kumar, R. Gupta, and S. Gupta, "Whole Slide Imaging (WSI) in Pathology: Current Perspectives and Future Directions," *Journal of Digital Imaging*, vol. 33, pp. 1034 – 1040, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:218978672

[24] K. Tomczak, P. Czerwińska, and M. Wiznerowicz, "The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge," *Contemporary Oncology*, vol. 19, pp. A68 – A77, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:12829250

[25] Z. Wang, M. A. Jensen, and J. C. Zenklusen, "A practical guide to the cancer genome atlas (TCGA)," *Statistical Genomics: Methods and Protocols*, pp. 111–141, 2016, publisher: Springer.

[26] R. W. Brown, Y.-C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan, "Magnetic Resonance Imaging: Physical Principles and Sequence Design," 1999. [Online]. Available: https://api.semanticscholar.org/CorpusID:119991289

[27] S. Ansari, N. Farzaneh, M. Duda, K. Horan, H. B. Andersson, Z. D. Goldberger, B. K. Nallamothu, and K. Najarian, "A Review of Automated Methods for Detection of Myocardial Ischemia and Infarction Using Electrocardiogram and Electronic Health Records," *IEEE Reviews in Biomedical Engineering*, vol. 10, pp. 264–298, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:34546946

[28] l. Qiao, Y. Zhou, J. Shang, C. Xiao, and J. Sun, "Opportunities and Challenges in Deep Learning Methods on Electrocardiogram Data: A Systematic Review," *ArXiv*, vol. abs/2001.01550, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:209862062

[29] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: a review," *Journal of Neural Engineering*, vol. 16, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:73508614

[30] M. Graña and K. Jackowski, "Electronic Health Record: A review," *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1375–1382, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:1970789

[31] N. S. Abul-Husn and E. E. Kenny, "Personalized Medicine and the Power of Electronic Health Records," *Cell*, vol. 177, pp. 58–69, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:84843749

[32] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin, and S. Bakas, "Multi-Institutional Deep Learning Modeling Without Sharing Patient Data: A Feasibility Study on Brain Tumor Segmentation," *Brainlesion : glioma, multiple sclerosis, stroke and traumatic brain injuries : second International Workshop, BrainLes 2016, with the challenges on BRATS, ISLES and mTOP 2016, held in conjunction with MICCAI 2016, Athens, Greece, Octob...*, vol. 11383, pp. 92–104, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:52956484

[33] W. Li, F. Milletarì, D. Xu, N. Rieke, J. Hancox, W. Zhu, M. Baust, Y. Cheng, S. Ourselin, M. J. Cardoso, and A. Feng, "Privacy-Preserving Federated Brain Tumour Segmentation," in *Machine Learning in Medical Imaging*, H.-I. Suk, M. Liu, P. Yan, and C. Lian, Eds. Cham: Springer International Publishing, 2019, vol. 11861, pp. 133–141, series Title: Lecture Notes in Computer Science. [Online]. Available: http://link.springer.com/10.1007/978-3-030-32692-0_16

[34] C. I. Bercea, B. Wiestler, D. Rueckert, and S. Albarqouni, "Federated disentangled representation learning for unsupervised brain anomaly detection," *Nature Machine Intelligence*, vol. 4, pp. 685 – 695, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:238810898

[35] S. Rajendran, J. S. Obeid, H. Binol, R. B. D. Agostino, K. L. Foley, W. Zhang, P. F. Austin, J. Brakefield, M. N. Gurcan, and U. Topaloglu, "Cloud-Based Federated Learning Implementation Across Medical Centers," *JCO Clinical Cancer Informatics*, vol. 5, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:231193486

[36] M. Adnan, S. Kalra, J. C. Cresswell, G. W. Taylor, and H. R. Tizhoosh, "Federated learning and differential privacy for medical image analysis," *Scientific Reports*, vol. 12, no. 1, p. 1953, Feb. 2022. [Online]. Available: https://www.nature.com/articles/s41598-022-05539-7

[37] M. Y. Lu, D. Kong, J. Lipková, R. J. Chen, R. Singh, D. F. K. Williamson, T. Y. Chen, and F. Mahmood, "Federated learning for computational pathology on gigapixel whole slide images," *Medical image analysis*, vol. 76, pp. 102 298 – 102 298, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:221836394

[38] A. T. Wang, C. M. Vachon, K. R. Brandt, and K. Ghosh, "Breast density and breast cancer risk: a practical review." *Mayo Clinic proceedings*, vol. 89 4, pp. 548–57, 2014. [Online]. Available: https://api.semanticscholar.org/CorpusID:9472922

[39] H. R. Roth, K. Chang, P. Singh, N. Neumark, W. Li, V. Gupta, S. Gupta, L. Qu, A. Ihsani, B. C. Bizzo, Y. Wen, V. Buch, M. Shah, F. Kitamura, M. Mendonça, V. Lavor, A. Harouni, C. Compas, J. Tetreault, P. Dogra, Y. Cheng, S. Erdal, R. White, B. Hashemian, T. Schultz, M. Zhang, A. McCarthy, B. M. Yun, E. Sharaf, K. V. Hoebel, J. B. Patel, B. Chen, S. Ko, E. Leibovitz, E. D. Pisano, L. Coombs, D. Xu, K. J. Dreyer, I. Dayan, R. C. Naidu, M. Flores, D. Rubin, and J. Kalpathy-Cramer, "Federated Learning for Breast Density Classification: A Real-World Implementation," in *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*, S. Albarqouni, S. Bakas, K. Kamnitsas, M. J. Cardoso, B. Landman, W. Li, F. Milletari, N. Rieke, H. Roth, D. Xu, and Z. Xu, Eds. Cham: Springer International Publishing, 2020, vol. 12444, pp. 181–191, series Title: Lecture Notes in Computer Science. [Online]. Available: https://link.springer.com/10.1007/978-3-030-60548-3_18

[40] T.-L. Yang, H.-W. Tsai, W.-C. Huang, J.-C. Lin, J.-B. Liao, N.-H. Chow, and P.-C. Chung, "Pathologic liver tumor detection using feature aligned multi-scale convolutional network," *Artificial intelligence in medicine*, vol. 125, p. 102244, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:245875295

[41] Z. Yan, J. Wicaksana, Z. Wang, X. Yang, and K.-T. Cheng, "Variation-Aware Federated Learning With Multi-Source Decentralized Medical Image Data," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, pp. 2615–2628, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:227167201

[42] X. Cai, Y. Lan, Z. Zhang, J. Wen, Z. Cui, and W. Zhang, "A Many-Objective Optimization Based Federal Deep Generation Model for Enhancing Data Processing Capability in IoT," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 561–569, Jan. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/9472967/

[43] M. S. Sirsat, E. Fermé, and J. Camara, "Machine learning for brain stroke: a review," *Journal of Stroke and Cerebrovascular Diseases*, vol. 29, no. 10, p. 105162, 2020, publisher: Elsevier.

[44] C. Ju, R. Zhao, J. Sun, X. Wei, B. Zhao, Y. Liu, H. Li, T. Chen, X. Zhang, D. Gao, B. Tan, H. Yu, C. He, and Y. Jin, "Privacy-Preserving Technology to Help Millions of People: Federated Prediction Model for Stroke Prevention," Dec. 2020, arXiv:2006.10517 [cs]. [Online]. Available: http://arxiv.org/abs/2006.10517

[45] N. Victor, S. Bhattacharya, P. K. R. Maddikunta, F. M. Alotaibi, T. R. Gadekallu, and R. H. Jhaveri, "FL-PSO: A Federated Learning approach with Particle Swarm Optimization for Brain Stroke Prediction," in *2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing Workshops (CCGridW)*. Bangalore, India: IEEE, May 2023, pp. 33–38. [Online]. Available: https://ieeexplore.ieee.org/document/10181200/

[46] S. H. Dipro, M. Islam, A. Al Nahian, M. Sharmita Azad, A. Chakrabarty, and T. Reza, "A Federated Learning Based Privacy Preserving Approach for Detecting Parkinson's Disease Using Deep Learning," in *2022 25th International Conference on Computer and Information Technology (ICCIT)*. Cox's Bazar, Bangladesh: IEEE, Dec. 2022, pp. 139–144. [Online]. Available: https://ieeexplore.ieee.org/document/10055787/

[47] J. Reyes, Y. Xiao, and M. Kersten-Oertel, "Data Imputation and Reconstruction of Distributed Parkinson's Disease Clinical Assessments: A Comparative Evaluation of Two Aggregation Algorithms," in *CLIP/DCL/LL-COVID19/PPML@MICCAI*, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:244110278

[48] Y.-L. Huang, H.-C. Yang, and C.-C. Lee, "Federated Learning via Conditional Mutual Learning for Alzheimer's Disease Classification on T1w MRI," *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 2427–2432,

2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:243920645

[49] T. Furlanello, Z. C. Lipton, M. Tschannen, L. Itti, and A. Anandkumar, "Born-Again Neural Networks."

[50] D. Li and J. Wang, "FedMD: Heterogenous Federated Learning via Model Distillation," *ArXiv*, vol. abs/1910.03581, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:203951869

[51] B. Han, R. H. Jhaveri, H. Wang, D. Qiao, and J. Du, "Application of Robust Zero-Watermarking Scheme Based on Federated Learning for Securing the Healthcare Data," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, pp. 804–813, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:240228935

[52] Y. Wu, D. Zeng, Z. Wang, Y. Sheng, L. Yang, A. J. James, Y. Shi, and J. Hu, "Federated Contrastive Learning for Dermatological Disease Diagnosis via On-device Learning (Invited Paper)," in *2021 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*. Munich, Germany: IEEE, Nov. 2021, pp. 1–7. [Online]. Available: https://ieeexplore.ieee.org/document/9643454/

[53] H. Elayan, M. Aloqaily, and M. Guizani, "Deep Federated Learning for IoT-based Decentralized Healthcare Systems," in *2021 International Wireless Communications and Mobile Computing (IWCMC)*. Harbin City, China: IEEE, Jun. 2021, pp. 105–109. [Online]. Available: https://ieeexplore.ieee.org/document/9498820/

[54] M. Zhang, Y. Wang, and T. Luo, "Federated Learning for Arrhythmia Detection of Non-IID ECG," *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, pp. 1176–1180, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:231919596

[55] N. Shoham, T. Avidor, A. Keren, N. Israel, D. Benditkis, L. Mor-Yosef, and I. Zeitak, "Overcoming Forgetting in Federated Learning on Non-IID Data," *ArXiv*, vol. abs/1910.07796, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:204743727

[56] A. Raza, K. P. Tran, L. Koehl, and S. Li, "Designing ECG Monitoring Healthcare System with Federated Transfer Learning and Explainable AI," *Knowl. Based Syst.*, vol. 236, p. 107763, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:235195935

[57] C. Lord, E. H. Cook, B. L. Leventhal, and D. G. Amaral, "Autism spectrum disorders," *Neuron*, vol. 28, no. 2, pp. 355–363, 2000, publisher: Elsevier.

[58] E. A. DeYoe, P. A. Bandettini, J. Neitz, D. Miller, and P. Winans, "Functional magnetic resonance imaging (FMRI) of the human brain," *Journal of Neuroscience Methods*, vol. 54, pp. 171–187, 1994. [Online]. Available: https://api.semanticscholar.org/CorpusID:3718293

[59] X. Li, Y. Gu, N. C. Dvornek, L. H. Staib, P. Ventola, and J. S. Duncan, "Multi-site fMRI Analysis Using Privacy-preserving Federated Learning and Domain Adaptation: ABIDE Results," *Medical image analysis*, vol. 65, p. 101765, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:210702063

[60] U. Ahmed, J. C.-W. Lin, and G. S. Srivastava, "Hyper-Graph Attention Based Federated Learning Methods for Use in Mental Health Detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 2, pp. 768–777, Feb. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/9767700/

[61] P. Chhikara, P. Singh, R. Tekchandani, N. Kumar, and M. Guizani, "Federated Learning Meets Human Emotions: A Decentralized Framework for Human–Computer Interaction for IoT Applications," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6949–6962, Apr. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9253631/

[62] T. Borger, P. Mosteiro, H. Kaya, E. Rijcken, A. A. Salah, F. E. Scheepers, and M. Spruit, "Federated learning for violence incident prediction in a simulated cross-institutional psychiatric setting," *Expert Syst. Appl.*, vol. 199, p. 116720, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:247405036

[63] I. R. Edwards and J. K. Aronson, "Adverse drug reactions: definitions, diagnosis, and management," *The Lancet*, vol. 356, no. 9237, pp. 1255–1259, Oct. 2000. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0140673600027999

[64] O. Choudhury, Y. Park, T. Salonidis, A. Gkoulalas-Divanis, I. Sylla, and A. K. Das, "Predicting Adverse Drug Reactions on Distributed Health Data using Federated Learning," *AMIA ... Annual Symposium proceedings. AMIA Symposium*, vol. 2019, pp. 313–322, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:216030941

[65] T. S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I. C. Paschalidis, and W. Shi, "Federated learning of predictive models from federated Electronic Health Records," *International Journal of Medical Informatics*, vol. 112, pp. 59–67, Apr. 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S138650561830008X

[66] L. Huang, A. L. Shea, H. Qian, A. Masurkar, H. Deng, and D. Liu, "Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records," *Journal of Biomedical Informatics*, vol. 99, p. 103291, Nov. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1532046419302102

[67] E. Brophy, M. D. Vos, G. B. Boylan, and T. E. Ward, "Estimation of Continuous Blood Pressure from PPG via a Federated Learning Approach," *Sensors (Basel, Switzerland)*, vol. 21, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:232035696

[68] X. Xu, H. Peng, M. Z. A. Bhuiyan, Z. Hao, L. Liu, L. Sun, and L. He, "Privacy-Preserving Federated Depression Detection From Multisource Mobile Health Data," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 7, pp. 4788–4797, Jul. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9540999/

[69] J. Fan, X. Wang, Y. Guo, X. Hu, and B. Hu, "Federated Learning Driven Secure Internet of Medical Things," *IEEE Wireless Communications*, vol. 29, no. 2, pp. 68–75, Apr. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9801732/

[70] J. C. Liu, J. Goetz, S. Sen, and A. Tewari, "Learning from others without sacrificing privacy: Simulation comparing centralized and federated machine learning on mobile health data," *JMIR mHealth and uHealth*, vol. 9, no. 3, p. e23728, 2021, publisher: JMIR Publications Toronto, Canada.

[71] Q. Wu, X. Chen, Z. Zhou, and J. Zhang, "FedHome: Cloud-Edge Based Personalized Federated Learning for In-Home Health Monitoring," *IEEE Transactions on Mobile Computing*, vol. 21, no. 8, pp. 2818–2832, Aug. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9296274/

[72] D. Y. Zhang, Z. Kou, and D. Wang, "FedSens: A Federated Learning Approach for Smart Health Sensing with Class Imbalance in Resource Constrained Edge Computing," *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, pp. 1–10, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:236480980

[73] T. Luo, S. S. Kanhere, J. Huang, S. K. Das, and F. Wu, "Sustainable Incentives for Mobile Crowdsensing: Auctions, Lotteries, and Trust and Reputation Systems," *IEEE Communications Magazine*, vol. 55, pp. 68–74, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:2930744

[74] X. Zhu, J. An, M. Yang, L. Xiang, Q. Yang, and X. Gui, "A Fair Incentive Mechanism for Crowdsourcing in Crowd Sensing," *IEEE Internet of Things Journal*, vol. 3, pp. 1364–1372, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:281275

[75] F. Lin, A. Wang, Y. Zhuang, M. R. Tomita, and W. Xu, "Smart Insole: A Wearable Sensor Device for Unobtrusive Gait Monitoring in Daily Life," *IEEE Transactions on Industrial Informatics*, vol. 12, pp. 2281–2291, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:10992630

[76] W. Y. B. Lim, S. Garg, Z. Xiong, D. Niyato, C. Leung, C. Miao, and M. Guizani, "Dynamic Contract Design for Federated Learning in Smart Healthcare Applications," *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 16 853–16 862, Dec. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9241078/

[77] Y. Guo, F. Liu, Z. Cai, L. Chen, and N. Xiao, "FEEL: A Federated Edge Learning System for Efficient and Privacy-Preserving Mobile Healthcare," *Proceedings of the 49th International Conference on Parallel Processing*, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:221070115