An In-Depth Exploration of Federated Learning's Revolutionary Influence on Medical Imaging

Monirul Haque
Computer Science and Engineering
Brac University
Dhaka, Bangladesh

Sheikh Araf Noshin
Computer Science and Engineering
Brac University
Dhaka, Bangladesh

Tanjina Akter Nipu
Computer Science and Engineering
Brac University
Dhaka, Bangladesh

Shihab Uddin Sikder
Computer Science and Engineering
Brac University
Dhaka, Bangladesh

Md Abu Ibrahim
Computer Science and Engineering
Brac University
Dhaka, Bangladesh

A.M.Tayeful Islam
Computer Science and Engineering
Brac University
Dhaka, Bangladesh

Sadiul Arefin Rafi
Computer Science and Engineering
Brac University
Dhaka, Bangladesh

Adib Muhammad Amit
Computer Science and Engineering
Brac University
Dhaka, Bangladesh

Annajiat Alim Rasel
Computer Science and Engineering
Brac University
Dhaka, Bangladesh

***Abstract*—Medical data, including structured clinical records, unstructured narratives, and multi-omics data, presents both opportunities and challenges. Advanced data storage, management, and governance are crucial for ensuring privacy, security, and compliance. Machine learning and artificial intelligence can extract insights, predict disease outcomes, and personalize treatment plans. However, ethical dilemmas, data ownership, patient consent, and cybersecurity measures remain. Healthcare systems must invest in scalable infrastructure and data science expertise to fully harness medical data. Federated learning, which enables cooperative training of machine learning models based on data from several sites without cross-site data exchange, has lately drawn a lot of interest as a possible option. In this article, we perform a thorough examination of the most recent advancements in federated learning techniques for medical picture processing. To deal with privacy protection and collaborative learning difficulties in medical imaging, we first explain the background information of federated learning. We give an analysis of the benchmark software platforms and medical imaging datasets used in the most recent federated learning research.**

***Index Terms*—Machine Learning, NLP, Binary Classification, federated learning, Medical Imaging, Ensemble.**

I. INTRODUCTION

In the modern era of automation and the advancement of computer technologies. Machine learning is quickly becoming the most interesting yet intimidating part of the world. The recent progression of machine learning in image classification, natural language analysis and other developments is creating scope for future research and implementation on a level that is unheard of. Chatgpt[1] has made a revolution in general day-to-day work as it can easily make formal writing, transcript and any other redundant work that a human may need to repeatedly do seem just one click away. To add more, voice-assisted AIs are also running the computer world which makes it so much easier to do any task by simply saying the instructions to your mobile/computer and it automatically interprets the instructions and does the task for you. Google's version of this voice assistant is called "OK Google". Although it all sounds like these kinds of innovative technology are making life easier, sometimes the situation can be too good to be true. In an artice[2] author perfectly represents this issue. In order to make these large-scale AI that can identify almost every word that's been spoken into the mic, it must have a massive amount of information gathered at first and trained in a massive computational system to be this versatile. The paper says, that a large model to achieve this type of accuracy the data which were used in training may or may not come from a certified source who gave access to the data. One common example can be shown in a famous social media platform called Facebook. In recent times, lots of people were concerned about the situation where if they mention the name of the product/groceries they wanted to buy, they immediately Facebook started to show advertisements of that exact product. This advertising tool was becoming so apparent that people started to find patterns and became very concerned about the privacy of the data they are sharing in social media[3]. Although it's claimed to only use for the purpose of selling advertisement or focused audience advertisement. This also begs the question of what if all the data was leaked and anybody can access it. They can easily exploit the sensitive information users

share on any platform and thus the rise of security issues of these massive corporate world machine learning models came to light. Along with that, machine-learning models getting increasingly bigger with more parameters meaning the need of higher computational power needed to train them is getting to a point that an average researcher willing to run those may not have access to it. To add more, if the nature of the model architecture is sequential then the time it takes to run becomes an issue. As a result distributed system to achieve these processing is being tested for quite a time. furthermore, one of the most important application parts of these models is medical imaging. Meaning that by examining medical images the model can accurately predict the issues within. Some common applications of medical imaging in machine learning can be found in this research [4] where the author describes multiple applications for medical imaging. For example, an x-ray image of the lungs, and a CT scan of the brain to classify whether it's healthy or not. The results are extremely reliable and helped the medical community by automating the preliminary process of medical treatment. Also, it helped to accommodate the less transportable area with a short-staffed committee. But with this importance, the sensitivity of the data and how crucial it is to preserve the information has been an issue also. Health data being leaked can be hazardous for the affected people as these data may be used to cause harm or taken advantage of. All of these reasons along with the urge to use the machine learning paradigm in distributed computing systems while also protecting the data's privacy gave birth to the concept of federated Learning. The federated learning approach ensures that no data is leaked from the local computer environment while training and simply after training is done the updated parameters/weight matrices are accumulated in all computers in that particular distributed system to achieve what a traditional machine learning algorithm would've done but parallelly and in a closed loop where each local environment doesn't have to share their data with any central system thus preventing data leak[5].

II. Federated Learning

The currently distributed networks that produce a plethora of data each day include mobile phones, wearable technology, and autonomous vehicles, to name just a few. Due to worries over sending private information and the increasing computational capability of these devices, it is more and more appealing to store data locally and move network computation to the device.

Federated learning is a machine learning technique that trains an algorithm via multiple independent sessions, each using its own dataset. This method is vastly accepted and cherished in modern days because of the reasons above. Federated learning aims at training heterogeneous datasets on local storage. It's a client-based structure that will train the data from as many clients as it currently has as a local machine. That also results in unreliable

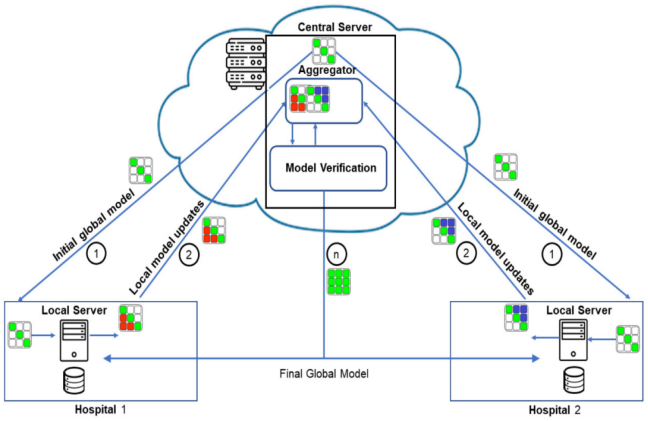

Figure 1: The overarching procedure of federated learning involves multiple training rounds and the active involvement of numerous participating hospitals, typically exceeding a mere two. Within each round, meticulous model verification is conducted both centrally and at the local participating client level.

situations where a client can be disconnected. Different solutions to each problem are being researched and discussed.

A. Centralized federated learning

In a centralized federated learning environment, a central server is utilized to coordinate all of the participating nodes and orchestrate the various steps of the algorithms. The server is in charge of selecting the nodes at the start of the training process and gathering the received model changes. The server could become the system's bottleneck because every single one of the chosen nodes must transmit updates to it.

B. Decentralized federated learning

The global model may be acquired by nodes working together in a decentralized federated learning environment. This configuration does not require a central server, hence lowering the possibility of single points of failure. Only nodes that are connected together may share model updates independently.

C. Heterogeneous federated learning

In various application domains, a wide range of clients, such as mobile devices and Internet of Things (IoT) gadgets, are being utilized. Most of the existing federated learning systems typically assume the presence of a shared global model architecture used in conjunction with local models. To address the requirements of diverse clients possessing varying computing and communication capabilities, a novel federated learning framework called HeteroFL was recently introduced. The HeteroFL approach enables the training of diverse local models, which can adapt to changing computational demands and handle

non-identically distributed (non-IID) data complexities. Ultimately, this results in the creation of a single accurate global inference model, as outlined in [6].

D. Federated Learning Process

There are a total of 5 processes of federated learning. They are Client selection, local training, model upload, model aggregation, and broadcast. Client selection is a process which the server selects a group of clients that are compatible with the requirements. Secondly, local training is a process where selected clients locally train a machine learning model through optimization methods such as stochastic gradient descent given that they have access to only the local data. The model weights can be controlled and initialized on both sides (server/client). The next step is called model upload which refers to the selection of clients and uploading the model to their local environment. Then, model aggregation is used where all the model trains their parameters and update a global model by aggregating all client models. Finally, the broadcast is done when the server sends/broadcasts the current shared global model to the selected clients. The accumulated model will be shared with the client and the client will retrieve that result and can update/fine-tune it to locally use their private data in their local environment.

III. Federated Learning in Medical Image Analysis

To initiate the survey. At first, we prioritized the research fields to a certain scope where both medical imaging and federated learning are incorporated. These modern machine learning models are referred to as data-driven ML models. The survey focuses on all the research papers that used different methods of federated learning to achieve classification with higher accuracy and precision. The research is then categorized with the method of federated learning used along with the particular medical field and the classification problem it solved. The problems are divided into three categories: client-end methods, server-end methods, and communication methods.

A. Methods Overview: A System Perspective

Federated learning (FL) offers a versatile framework for distributed learning while safeguarding privacy. It readily accommodates various machine learning techniques. FL encompasses data management, model learning, privacy safeguards, and communication architecture. In a system view, we classify current FL approaches for medical image analysis into three categories: 1) methods operating on client devices, 2) methods based on central servers, and 3) methods focusing on communication protocols. Each category groups different approaches based on their specific research objectives.

B. Client-End Learning

In the real world, the data distributions in medical imaging from various sites often exhibit notable dissimilarities, a phenomenon termed "data heterogeneity." This issue is commonly referred to as the "domain shift" problem [7], and within the context of federated learning (FL), it's also known as "client shift." When domain shifts occur within an FL system, they can pose significant challenges, complicating the convergence of the global model and leading to reduced performance for certain clients. In the following sections, we provide an overview of relevant research efforts aimed at mitigating domain shifts among clients within the field of FL.

1) Tailored Learning for Specific Domains: Federated learning seeks to create a global model that performs well across all clients. However, due to the diversity of data across different sites, achieving optimal performance for all clients can be challenging. To address this, a strategy involves fine-tuning the global model using domain-specific (local) data, often referred to as personalized federated learning [8]. Another research [9] proposes an encoder-decoder architecture within the federated learning framework for magnetic resonance (MR) image reconstruction. They maintain a globally shared encoder on the server to learn representations that are invariant across domains. Simultaneously, a client-specific decoder is trained using local data to harness the domain-specific characteristics of each client. Moreover, [10] introduces a federated learning framework that combines Convolutional Neural Networks (CNN) and Graph Neural Networks (GNN) to tackle domain shift issues in chest X-ray image classification. They share model weights of the CNN across clients to capture site-independent features, while they employ a local GNN fine-tuned with local data on each client to address site-specific data variations, thereby enabling the learning of both site-independent and site-specific features. Also, [11] present an ensemble-based framework to handle client shifts in medical image segmentation. Their approach includes a global model, personalized models, and a model selector. Instead of relying solely on the global model to fit all client data, they utilize personalized models generated for different clients to adapt to their data distributions through a model selector.Jiang et al. (2023) propose training a locally adapted model that combines global gradients and local gradients to optimize model performance on each client. This resulted in an effective improvement in the biased performance of the global model incorporated with different clients caused by client domain shifts.[12] construct a federated learning framework based on Generative Adversarial Networks (GANs) for harmonizing histopathological images. Each client trains a local discriminator to capture client-specific image styles, while a global generator model on the server generates domain-invariant images, achieving harmonization of histopathological images. Similarly, [13] proposes

a GAN-based approach for histopathological image harmonization, assuming access to a reference dataset shared among all clients to assist in training local GANs.

2) Enhancing Generalizability through Domain Adaptation: Domain adaptation, a well-established machine learning technique[14], aims to mitigate domain shift between diverse datasets and improve a model's ability to generalize across them. Many federated learning (FL) studies employ domain adaptation techniques to enhance learning performance.[15]introduce the use of domain adaptation methods to align domain distribution differences among clients. In their approach, privacy protection is achieved by introducing noise into the data within each client. A domain discriminator/classifier is then trained on this noisy data to reduce domain shift. present a domain adaptation-based federated learning framework designed to eliminate domain shifts arising from different scanning devices. They assume that image features follow Gaussian distributions, allowing the sharing of mean and standard deviation values among clients. During the training of each client's model, a label classifier and a domain discriminator are jointly trained to produce domain-invariant features, effectively mitigating domain shift. [16] incorporate batch normalization (BN) into deep neural networks to address client shift. Inspired by BN-based domain adaptation, [17] propose a federated training approach that shares only BN parameters containing domain-invariant information while keeping BN statistics local, as these statistics are presumed to contain domain-specific information.Guo et al. (2021) [18] put forth a federated learning method for MRI reconstruction. In their approach, the intermediate latent features learned across different clients are aligned with the distribution of latent features from a reference site.

3) Addressing Data Heterogeneity through Image Harmonization: [19] introduce a generative replay strategy as a solution to manage data heterogeneity among clients. They start by training an auxiliary variational autoencoder (VAE) to generate medical images that closely resemble the input images. Each client then optimizes its local classifier using both its actual local data and synthesized data designed to resemble data from other clients. This approach effectively reduces domain shift. Another research[20] leverage cycleGAN to minimize variations among clients. They select one client (site) with relatively simple data patterns as a reference. cycleGAN is applied to harmonize images from other clients with the reference site, ensuring consistency across the dataset. [21] propose a frequency-based harmonization technique to mitigate domain differences among clients. In this method, images are transformed into the frequency domain, and phase components are retained locally, while the average amplitudes from each client are shared and normalized to harmonize all client images effectively.

C. Cleave Learning

1) Aggregation of Weight: At the server end, [22] suggest using a Progressive Fourier Aggregation technique. Only these low-frequency components are aggregated to share information gathered from many clients, while the high-frequency portions are ignored, based on prior research that shows low-frequency components of parameters constitute the backbone of deep network capabilities [23]. According to [11], the effect factor of weight aggregation is the training loss of each customer. A reduced weight will be assigned to the client whose performance was comparatively poor due to unbalanced data in the overall weight aggregation.

2) Clients' domains change: In Federated Learning, the term "domain shift among clients" refers to a frequent issue where the data distributions across several client devices or sources vary considerably [24]. The lack of data uniformity that results from this mismatch can be attributed to various geographic locations, device kinds, or data-gathering techniques [25]. The federated learning process might be hampered by this non-iid (non-independent and identically distributed) data distribution because models developed on the data of one client may not generalize effectively to the data of other clients[26]. In Federated Learning, addressing domain shift is essential to ensuring that the global model can adapt to the diverse data characteristics across clients, thereby boosting model performance and sustaining the integrity and efficacy of the federated learning system[27].

D. Client-server connection

The predominant imperative underpinning federated learning systems centers on the preservation of data privacy, specifically, the unequivocal guarantee that the data pertaining to each individual client remains impervious to unauthorized viewing or access by other clients or servers. Extensive prior research has unequivocally established that a machine learning model's gradient leakage can facilitate the reconstruction of pixel-level images, even in scenarios where inter-site data sharing is entirely absent[28][29]. Therefore, the investigation of state-of-the-art techniques for the proactive prevention of data leakage during the communication interface between the server and multiple clients assumes paramount importance. Notably, this subject matter has garnered significant attention in contemporary scholarly inquiries.

The privacy of individuals can potentially be compromised through the inadvertent inclusion of sensitive information within the gradient data of a deep neural network, rendering it susceptible to reconstruction by malicious entities. To mitigate this risk, the application of differential privacy [30] techniques can diminish the extent to which an individual's presence within the training dataset can be inferred.

Furthermore, recent research endeavors have advocated the incorporation of Gaussian random noise into the com-

puted gradients derived from each client's imaging data [31]. This measure is undertaken to safeguard patient privacy not only from the server but also from other clients partaking in the collaborative learning process.

IV. Analysis of Federated Learning Implementation in Medical Imaging

A. Methods, Architectures, and Frameworks

Federated learning endeavors to mitigate the prevailing concerns associated with data governance and privacy by facilitating the cooperative training of algorithms, all while avoiding the explicit exchange of raw data. The conventional approach of amalgamating data from diverse sources invariably engenders substantial apprehensions regarding patient privacy and data security. The capability to train machine learning models comprehensively across various medical institutions without necessitating the transfer of the underlying data represents an indispensable technological advancement in this context.

It is imperative to acknowledge that the scholarly work entitled "The Future of Digital Health with Federated Learning" [32] was disseminated through the esteemed academic platform of Nature Digital Medicine in September 2020. In this seminal piece, the authors meticulously examine the potential role of federated learning in shaping the trajectory of digital health in the future. Additionally, the authors underscore the formidable challenges and multifaceted considerations that demand thoughtful attention within this burgeoning field of research. In a scholarly contribution featured in the esteemed journal Nature Medicine, titled "Federated Learning for Predicting Clinical Outcomes in Patients with COVID-19" [33], the authors have prominently demonstrated the precision and dependability of a federated artificial intelligence model. This model was designed for the critical task of predicting oxygen requirements among patients afflicted with COVID-19 infections.

Additionally, within the confines of another scholarly work, specifically titled "A Systematic Review of Federated Learning in the Healthcare Domain: An Assessment of Data Characteristics and Applications," [6] the researcher endeavors to offer a comprehensive analysis of the challenges inherent to federated learning, concentrating particularly on the nuanced considerations originating from medical data properties and their applications.

B. Performance Improvements

In the context of intermittent client participation, characterized by fluctuating numbers of clients, an empirical assessment was undertaken to evaluate the efficacy and scalability of pneumonia classification using Chest X-ray (CXR) images [34]. Clients within this system possess the option to enroll for participation in the training cycle or to withdraw at their discretion. While the study did not explicitly factor in communication costs, the proposed systems designed to oversee the data management of such intermittently participating clients exhibited notable advantages in terms of classification accuracy when compared to a centralized approach, along with a discernible reduction in processing time [34]. The recommendation arose to implement a customized Federated Learning (CusFL) framework, wherein each client refines a personalized model by leveraging the collective knowledge of the federated model [35].

An extensive range of mental health disorders were diagnosed employing the Federated Multi-Task Learning Framework for Joint Diagnosis (FMTLJD), thereby showcasing the capacity of shared knowledge in enhancing the generalizability of diagnostic outcomes [36]. This comprehensive investigation involved the participation of eight distinct clients with varying sample sizes, systematically evaluating the incremental influence of incorporating clients into the learning process [36]. This proposed approach has been empirically validated as efficacious, particularly for institutions characterized by limited datasets, thereby facilitating more effective learning [36].

V. Medical Image Datasets

A. Overview of Medical Image data

In the field of medical image analysis, federated learning (FL) offers a lot of potential since it provides a solution that encourages cooperative model training across many data sources while upholding the crucial need for data privacy. Strict privacy regulations make it difficult to share medical photos from different websites in order to build big training datasets. Two main approaches for efficiently leveraging various imaging datasets in simulations and experiments emerge in the framework of FL research for medical image analysis.

The first tactic entails making direct use of databases gathered from various medical websites and facilities. These databases often result from cross-center research collaboration projects. As a result, they stand as the best option for creating a mimicked FL environment. Partitioning a large medical picture dataset [15] into smaller parts and treating each as a separate client dataset is another popular strategy for creating an FL experimental platform. Due to the fact that no data is shared between the sites, this clever strategy enables the use of a sizable dataset without compromising the privacy of individual users. These subgroups may be carefully chosen using a variety of factors, such as imaging modality, illness category, or patient demographics. Working with subsets [10] also makes it easier to investigate domain adaptation strategies that are intended to improve FL model performance across various subcategories.

Both of these methods provide efficient solutions to the problem of small sample sizes in medical picture analysis. They open the door for the creation of reliable FL models with broad applicability in this important field.

B. Chest Images

The publicly available COVID-19 Chest X-ray database, commonly referred to as COVID-19 CXR and cited as per [37], contains a range of chest X-ray images. This dataset encompasses 3,616 instances of individuals testing positive for COVID-19, 10,192 cases of healthy individuals, 6,012 cases featuring lung opacity unrelated to COVID, and 1,345 instances of viral pneumonia.

The COVIDx dataset, as referenced in [38], is a thorough and freely accessible database, encompassing 13,975 chest X-ray images originating from 13,870 patients. This dataset incorporates 358 chest X-ray images from 266 patients with COVID-19, 8,066 images portraying normal cases, and 5,538 images illustrating cases of pneumonia not related to COVID-19.

A large dataset called CheXpert [39] has 65,240 patients' 224,316 chest radiographs. These pictures were taken from the medical center of Stanford University.

The ChestX-ray [40] is a sizable and freely accessible medical image collection that includes 112,120 X-ray pictures of 30,805 individuals with 14 illness classifications in frontal view. It adds six more thorax disorders, including Edema, Emphysema, Fibrosis, Hernia, Pleural, and Thickening, to the ChestX-ray8 dataset.

The Automatic Cardiac Diagnosis Challenge (ACDC) [41] is a sizable dataset made exclusively for evaluating cardiac MRI. It is publically accessible. Based on well-defined physiological criteria for cardiac examination, it consists of data from 150 patients divided into five groups.

A publicly available dataset centered on cardiac MRI is called the Multi-Center, Multi-Vendor, and Multi-Disease Cardiac Segmentation Challenge (M&Ms) [42]. It contains information from 375 people who were affiliated with six hospitals in Spain, Canada, and Germany. Four distinct scanner manufacturers, including GE, Siemens, Philips, and Canon, were used to obtain the cardiac MRIs in this dataset.

C. Skin Images

The International Skin Imaging Collaboration (ISIC) challenge dataset [43] is a substantial database offering various challenges for processing skin lesion images. ISIC serves as a widely used benchmark dataset in the field of dermatoscopic image analysis.

The "Human Against Machine with 10000 Training Images" (HAM10000) [44] is a widely recognized and extensive dataset for identifying skin lesions which are pigmented. It comprises a total of 10,015 dermatoscopic images collected from various sources. This dataset covers cases from all major diagnostic categories of pigmented skin lesions.

D. Brain Images

The largest and most important benchmark for research on Alzheimer's disease (AD) is the Alzheimer's Disease Neuroimaging Initiative (ADNI) [45] which includes the ADNI-1, ADNI-2, ADNI-GO, and ADNI-3 studies. We provide for analysis and study structural brain MRI, functional brain MRI, and positron emission tomography (PET) data from 1,900+ participants and 59 centers.

A sizable dataset for brain imaging made up of 100,000 people in the UK Biobank [46] includes structural, functional, and diffusion MRI scans of the participants' brains.

A dataset known as the Multimodal Brain Tumor Image Segmentation Benchmark (BraTS) [47] is commonly used in the Brain Tumor Segmentation Challenge and is regularly updated. It contains brain MRIs collected from approximately 19 different independent organizations, with data acquired using various types of MRI scanners.

Another benchmark database for research on autism spectrum disorder is the Autism Brain Imaging Data Exchange (ABIDE) program [48]. Structured and functional brain pictures from over 24 imaging facilities/sites throughout the world are separately gathered and included in ABIDE.

Lastly, the Radiological Society of North America (RSNA) [49] dataset is a sizable collection of CT scans created for the purpose of identifying cerebral hemorrhages. The 874,035 photos that make up this dataset were acquired from three different institutions: Universidade Federal de So Paulo in So Paulo, Brazil; Thomas Jefferson University Hospital in Philadelphia, USA; and Stanford University in Palo Alto, USA.

E. Knee

The fastMRI dataset [50] is an extensive collection for applying machine learning in medical image reconstruction. It encompasses over 1,500 knee MRIs (with 1.5 and 3 Tesla) and includes DICOM images derived from 10,000 clinical knee MRIs (also featuring 1.5 and 3 Tesla).

F. Histology

According to the Cancer Genome Atlas Research Network in 2013, the Cancer Genome Atlas (TCGA) [51] is a notable cancer genomics resource. For histology and microscopy study, it provides a wealth of information,

including whole-slide photos for both healthy controls and cancer patients.

G. Eye

A sizable collection of color digital retinal fundus photos intended to identify diabetic retinopathy makes up the Kaggle Diabetic Retinopathy (Retina) dataset **diabetic-retinopathy-detection**. There are 17,563 pairings of these photos in all. Each picture in the collection has a label that, on a scale from 0 to 4, indicates if diabetic retinopathy is present and how severe it is. The severity of the condition is indicated by an intermediate value between 0 (no diabetic retinopathy) and 4 (proliferative diabetic retinopathy).

H. Abdomen

The PROMISE12 dataset [52] was produced for the evaluation of prostate MRI segmentation methods and is linked to the MICCAI 2012 Prostate MR Image Segmentation challenge. It includes 100 prostate MRIs that were acquired using different scanners at four separate medical facilities: Beth Israel Deaconess Medical Center in Boston, Massachusetts, Radboud University Nijmegen Medical Centre in the Netherlands, and University College London in the United Kingdom.

I. MedMNIST

MedMNIST [53], introduced by Yang ,serves as a dataset designed for medical image classification. It follows a format akin to the MNIST dataset, with all images being of size 28 x 28 pixels. MedMNIST consists of ten pre-processed subsets that encompass various primary modalities like MR, CT, X-ray, Ultrasound, and OCT. Due to its versatility and manageable size, MedMNIST proves valuable for swiftly prototyping machine learning algorithms.

VI. Limitations and Future Research

A. Future Directions

Exploring the future of federated learning applied to medical imaging will open up a world of ground-breaking possibilities. The fusion of cutting-edge technologies has the promise of improving diagnostic accuracy as well as enabling innovative applications in predictive medicine. These cutting-edge technologies include enhanced deep learning architectures, federated optimization algorithms, and secure multi-party computation. The pursuit of interoperability across various imaging modalities and healthcare systems will take center stage in order to facilitate smooth institutional collaboration and data harmonization. Furthermore, federated learning has promise for customized medicine by allowing models to adjust to unique patient profiles while protecting data privacy. As federated learning develops, the combination of interpretability and explainable AI approaches will further shed light on the mystery that is deep neural networks, fostering trust and openness in clinical decision support systems. The use of federated learning in global health projects and situations with limited resources also calls, for portending a time when everyone has access to cutting-edge medical image analysis regardless of location. The management of ethical data sharing and patient-centric privacy measures will continue to be crucial throughout this complicated path as federated learning emerges as a transformational force in the always-changing world of healthcare innovation.

B. Overcoming Security Threats

As participant identities and contributions are known, federated learning (FL) collaboration among massive healthcare institutions according to a managed, restricted membership framework boosts security against harmful assaults. For on-device collaborations, where any device might contribute data, possibly with malicious intent, enforcing such limits can be difficult FL frequently substitutes gradients for raw data when sending information to a central server to ensure privacy, it's crucial to keep in mind that sensitive images can still be reconstructed from gradients [54]. Protecting sensitive data requires the use of privacy-preserving techniques like differential privacy and homomorphic encryption [55]. [56] [57] provide comprehensive information on various privacy measures and an evaluation of their efficacy.

Security is essential in federated learning (FL), where clients communicate with a central server. Attacks using poison and attacks using gradient inversion are two frequent risks.

1) Attacks with Poison: False actors may tamper with model updates or client training datasets. The resilience of FL was increased by the proposal of a distance-based outlier suppression (DOS) algorithm to protect against different poisoning attacks.

2) Gradient Inversion Attacks: These approaches seek to compare the real data with the trainable input data, jeopardizing privacy. The privacy enhancement potential of Gaussian and Laplace randomization noise levels was investigated, and a trade-off between privacy and model precision was found. The risk of such attacks is quantified by the improved Rank Analysis Index (RA-I).

3) Model Inversion Attacks: Adding Gaussian noise to model updates can help defend against server-side model inversion attacks, which are especially common in CXR image classification.

4) Secure Framework: The FedAvg algorithm on datasets of chest X-rays produced a 15% improvement using the MediSecFed framework, even in the presence of malevolent clients.

VII. Conclusion

To summarize the survey, this paper compiles all information regarding federated learning that impacts the biomedical digitization process. Additionally, with the advantages of federated learning, inherent dangers and security issues arise. This paper shed light on that matter as well. Furthermore, each medical sector that was being researched and innovative technologies implemented through federated learning are collected for the reader to get proper guidance while navigating this topic to gather knowledge. This survey aims to have a perfect representation of organized data for relatively new researchers or experienced learners who want to get a general idea about where the trend of federated learning is going and which should be exploited and upgraded to help develop medical imaging with machine learning through federated learning. The survey hopes to be regarded as a guide map for young enthusiasts seeking to learn federated learning to contribute to the distributed computing system space.

References

[1] Y. Liu, T. Han, S. Ma, *et al.*, "Summary of chatgpt-related research and perspective towards the future of large language models," *Meta-Radiology*, p. 100 017, 2023.

[2] M. Vimalkumar, S. K. Sharma, J. B. Singh, and Y. K. Dwivedi, "'okay google, what about my privacy?': User's privacy perceptions and acceptance of voice based digital assistants," *Computers in Human Behavior*, vol. 120, p. 106 763, 2021.

[3] J. G. Cabañas, Á. Cuevas, A. Arrate, and R. Cuevas, "Does facebook use sensitive data for advertising purposes?" *Communications of the ACM*, vol. 64, no. 1, pp. 62–69, 2020.

[4] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual review of biomedical engineering*, vol. 19, pp. 221–248, 2017.

[5] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," *IEEE signal processing magazine*, vol. 37, no. 3, pp. 50–60, 2020.

[6] Prayitno, C.-R. Shyu, K. T. Putra, *et al.*, "A systematic review of federated learning in the healthcare area: From the perspective of data properties and applications," *Applied Sciences*, vol. 11, no. 23, p. 11 191, 2021.

[7] H. Guan and M. Liu, "Domain adaptation for medical image analysis: A survey," *IEEE Transactions on Biomedical Engineering*, vol. 69, pp. 1173–1185, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:231951465.

[8] C. T. Dinh, N. H. Tran, and T. D. Nguyen, "Personalized federated learning with moreau envelopes," *ArXiv*, vol. abs/2006.08848, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:219708331.

[9] C.-M. Feng, Y.-b. Yan, H. Fu, Y. Xu, and L. Shao, "Specificity-preserving federated learning for mr image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 42, pp. 2010–2021, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:245124470.

[10] A. Chakravarty, A. Kar, R. Sethuraman, and D. Sheet, "Federated learning for site aware chest radiograph screening," *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pp. 1077–1081, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:235206998.

[11] Z. Li, X. Xu, X. Cao, *et al.*, "Integrated cnn and federated learning for covid-19 detection on chest x-ray images.," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. PP, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:249886072.

[12] J. Ke, Y. Shen, and Y. Lu, "Style normalization in histology with federated learning," *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pp. 953–956, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:235206349.

[13] N. Wagner, M. Fuchs, Y. Tolkach, and A. Mukhopadhyay, "Federated stain normalization for computational pathology," *ArXiv*, vol. abs/2209.14849, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:252369484.

[14] W. M. Kouw and M. Loog, "A review of domain adaptation without target labels," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, pp. 766–785, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:198898096.

[15] 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. 101 765, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:210702063.

[16] M. Andreux, J. O. du Terrail, C. Béguier, and E. W. Tramel, "Siloed federated learning for multi-centric histopathology datasets," in *DART/DCL@MICCAI*, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:221139661.

[17] Y. Li, N. Wang, J. Shi, X. Hou, and J. Liu, "Adaptive batch normalization for practical domain adaptation," *Pattern Recognit.*, vol. 80, pp. 109–117, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:13695858.

[18] P. Guo, P. Wang, J. Zhou, S. Jiang, and V. M. Patel, "Multi-institutional collaborations for improving deep learning-based magnetic resonance image reconstruction using federated learning," *2021 IEEE/CVF Conference on Computer Vision and*

Pattern Recognition (CVPR), pp. 2423–2432, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:232104977.

[19] L. Qu, N. Balachandar, M. Zhang, and D. Rubin, "Handling data heterogeneity with generative replay in collaborative learning for medical imaging," *Medical image analysis*, vol. 78, p. 102 424, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:235624146.

[20] 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.

[21] M. Jiang, Z. Wang, and Q. Dou, "Harmofl: Harmonizing local and global drifts in federated learning on heterogeneous medical images," in *AAAI Conference on Artificial Intelligence*, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:245353364.

[22] Y. Chen, N. J. Conroy, and V. L. Rubin, "Misleading online content: Recognizing clickbait as" false news"," in *Proceedings of the 2015 ACM on workshop on multimodal deception detection*, 2015, pp. 15–19.

[23] P. J. Liu, M. Saleh, E. Pot, *et al.*, "Generating wikipedia by summarizing long sequences," *arXiv preprint arXiv:1801.10198*, 2018.

[24] S. M. Hosseini, M. Sikaroudi, M. Babaie, and H. R. Tizhoosh, "Proportionally fair hospital collaborations in federated learning of histopathology images," *IEEE Transactions on Medical Imaging*, vol. 42, pp. 1982–1995, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:255701416.

[25] Z. Fan, J. Su, K. Gao, D. Hu, and L. Zeng, "A federated deep learning framework for 3d brain mri images," *2021 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–6, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:237619433.

[26] J. Luo and S. Wu, "Fedsld: Federated learning with shared label distribution for medical image classification," in *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2022, pp. 1–5.

[27] N. Alkhunaizi, D. Kamzolov, M. Tak'avc, and K. Nandakumar, "Suppressing poisoning attacks on federated learning for medical imaging," *ArXiv*, vol. abs/2207.10804, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:251018678.

[28] L. Zhu, Z. Liu, and S. Han, "Deep leakage from gradients," *CoRR*, vol. abs/1906.08935, 2019. arXiv: 1906.08935. [Online]. Available: http://arxiv.org/abs/1906.08935.

[29] H. Yin, A. Mallya, A. Vahdat, J. M. Álvarez, J. Kautz, and P. Molchanov, "See through gradients: Image batch recovery via gradinversion," *CoRR*, vol. abs/2104.07586, 2021. arXiv: 2104.07586. [Online]. Available: https://arxiv.org/abs/2104.07586.

[30] C. Dwork, A. Roth, *et al.*, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[31] X. Li, Y. Gu, N. Dvornek, L. H. Staib, P. Ventola, and J. S. Duncan, "Multi-site fmri analysis using privacy-preserving federated learning and domain adaptation: Abide results," *Medical Image Analysis*, vol. 65, p. 101 765, 2020.

[32] N. Rieke, J. Hancox, W. Li, *et al.*, "The future of digital health with federated learning," *NPJ digital medicine*, vol. 3, no. 1, p. 119, 2020.

[33] I. Dayan, H. R. Roth, A. Zhong, *et al.*, "Federated learning for predicting clinical outcomes in patients with covid-19," *Nature medicine*, vol. 27, no. 10, pp. 1735–1743, 2021.

[34] J. S.-P. Díaz and Á. L. García, "Study of the performance and scalability of federated learning for medical imaging with intermittent clients," *Neurocomputing*, vol. 518, pp. 142–154, 2023.

[35] J. Wicaksana, Z. Yan, X. Yang, Y. Liu, L. Fan, and K.-T. Cheng, "Customized federated learning for multi-source decentralized medical image classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 11, pp. 5596–5607, 2022.

[36] Z.-A. Huang, Y. Hu, R. Liu, *et al.*, "Federated multi-task learning for joint diagnosis of multiple mental disorders on mri scans," *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 4, pp. 1137–1149, 2022.

[37] M. E. H. Chowdhury, T. Rahman, A. Khandakar, *et al.*, "Can ai help in screening viral and covid-19 pneumonia?" *IEEE Access*, vol. 8, pp. 132 665–132 676, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:214713518.

[38] L. Wang, Z. Q. Lin, and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images," *Scientific Reports*, vol. 10, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:215768886.

[39] J. A. Irvin, P. Rajpurkar, M. Ko, *et al.*, "Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison," in *AAAI Conference on Artificial Intelligence*, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:58981871.

[40] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray: Hospital-scale chest x-ray database and benchmarks on weakly supervised classification and localization of common tho-

rax diseases," in *Deep Learning and Convolutional Neural Networks for Medical Imaging and Clinical Informatics*, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:8945673.

[41] O. Bernard, A. Lalande, C. Zotti, *et al.*, "Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: Is the problem solved?" *IEEE Transactions on Medical Imaging*, vol. 37, pp. 2514–2525, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:51610194.

[42] V. M. Campello, P. Gkontra, C. Izquierdo, *et al.*, "Multi-centre, multi-vendor and multi-disease cardiac segmentation: The m&ms challenge," *IEEE Transactions on Medical Imaging*, vol. 40, pp. 3543–3554, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:235471536.

[43] B. Cassidy, C. Kendrick, A. Brodzicki, J. Jaworek-Korjakowska, and M. H. Yap, "Analysis of the isic image datasets: Usage, benchmarks and recommendations," *Medical Image Analysis*, Nov. 2021. DOI: 10.1016/j.media.2021.102305.

[44] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, vol. 5, no. 1, Aug. 2018. DOI: 10.1038/sdata.2018.161. [Online]. Available: https://doi.org/10.1038%2Fsdata.2018.161.

[45] C. R. Jack, M. A. Bernstein, N. C. Fox, *et al.*, "The alzheimer's disease neuroimaging initiative (adni): Mri methods," *Journal of Magnetic Resonance Imaging*, vol. 27, 2008. [Online]. Available: https://api.semanticscholar.org/CorpusID:3272607.

[46] K. L. Miller, F. Alfaro-Almagro, N. K. Bangerter, *et al.*, "Multimodal population brain imaging in the uk biobank prospective epidemiological study," *Nature neuroscience*, vol. 19, pp. 1523–1536, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:1018393.

[47] B. H. Menze, A. Jakab, S. Bauer, *et al.*, "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE Transactions on Medical Imaging*, vol. 34, pp. 1993–2024, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:1739295.

[48] A. di Martino, C. Yan, Q. Li, *et al.*, "The autism brain imaging data exchange: Towards large-scale evaluation of the intrinsic brain architecture in autism," *Molecular psychiatry*, vol. 19, pp. 659–667, 2013. [Online]. Available: https://api.semanticscholar.org/CorpusID:13785515.

[49] A. E. Flanders, L. M. Prevedello, G. Shih, *et al.*, "Construction of a machine learning dataset through collaboration: The rsna 2019 brain ct hemorrhage challenge.," *Radiology. Artificial intelligence*, vol. 2 3, e190211, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:219068068.

[50] F. Knoll, J. Zbontar, A. Sriram, *et al.*, "Fastmri: A publicly available raw k-space and dicom dataset of knee images for accelerated mr image reconstruction using machine learning.," *Radiology. Artificial intelligence*, vol. 2 1, e190007, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:211214833.

[51] *National cancer institute*. [Online]. Available: https://www.cancer.gov/ccg/research/genome-sequencing/tcga.

[52] G. J. S. Litjens, R. Toth, W. J. M. van de Ven, *et al.*, "Evaluation of prostate segmentation algorithms for mri: The promise12 challenge," *Medical image analysis*, vol. 18 2, pp. 359–73, 2014. [Online]. Available: https://api.semanticscholar.org/CorpusID:10331142.

[53] Q. Liu, H. Yang, Q. Dou, and P.-A. Heng, "Federated semi-supervised medical image classification via inter-client relation matching," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:235446484.

[54] C. Zhang, S. Ekanut, L. Zhen, and Z. Li, "Augmented multi-party computation against gradient leakage in federated learning," *IEEE Transactions on Big Data*, pp. 1–10, 2022. DOI: 10.1109/TBDATA.2022.3208736.

[55] A. Peyvandi, B. Majidi, S. Peyvandi, and J. C. Patra, "Privacy-preserving federated learning for scalable and high data quality computational-intelligence-as-a-service in society 5.0," *Multimedia Tools and Applications*, vol. 81, no. 18, pp. 25 029–25 050, Mar. 2022. DOI: 10.1007/s11042-022-12900-5. [Online]. Available: https://doi.org/10.1007/s11042-022-12900-5.

[56] T.-T. Ho, K.-D. Tran, and Y. Huang, "Fedsgdcovid: Federated sgd covid-19 detection under local differential privacy using chest x-ray images and symptom information," *Sensors*, vol. 22, no. 10, 2022, ISSN: 1424-8220. DOI: 10.3390/s22103728. [Online]. Available: https://www.mdpi.com/1424-8220/22/10/3728.

[57] L. Zhang, J. Xu, P. Vijayakumar, P. K. Sharma, and U. Ghosh, "Homomorphic encryption-based privacy-preserving federated learning in IoT-enabled healthcare system," *IEEE Transactions on Network Science and Engineering*, pp. 1–17, 2022. DOI: 10.1109/tnse.2022.3185327. [Online]. Available: https://doi.org/10.1109%2Ftnse.2022.3185327.