

TOPICAL REVIEW

Decentralized Learning in Healthcare: A Review of Emerging Techniques

CHAMANI SHIRANTHIKA¹, (Graduate Student Member, IEEE),
PARVANEH SAEEDI¹, (Member, IEEE), AND IVAN V. BAJIĆ¹, (Senior Member, IEEE)

School of Engineering Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

Corresponding author: Chamani Shiranthika (csj5@sfu.ca)

This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada under Grant RGPIN-2021-02485 and Grant RGPAS-2021-00038.

ABSTRACT Recent developments in deep learning have contributed to numerous success stories in healthcare. The performance of a deep learning model generally improves with the size of the training data. However, there are privacy, ownership, and regulatory issues that prevent combining medical data into traditional centralized storage. Decentralized learning approaches enable collaborative model training by distributing the learning process among several nodes or devices. Conceptually, decentralized learning builds on earlier work in distributed optimization, but the focus of this paper is on recent and emerging techniques such as Federated Learning (FL), Split Learning (SL), and hybrid Split-Federated Learning (SFL). With common, universal deep learning models and centralized aggregator servers, FL overcomes the difficulties of centralized training. Additionally, patient data remains at the local party, upholding the security and anonymity of the data. SL enables machine learning without directly accessing data on clients or end devices. It further enhances privacy in a decentralized setting and mitigates clients' storage issues. In this survey, we first provide a contemporary survey of FL, SL, and SFL approaches. Second, we discuss their state-of-the-art applications in healthcare, particularly in medical image analysis. Third, we review these emerging decentralized learning approaches under challenging conditions such as statistical and system heterogeneity, privacy preservation, communication efficiency, fairness, etc. Then, we address existing approaches to tackle these challenges. We detail unique complications related to healthcare applications including data, privacy and security, and communication challenges. Finally, we outline potential areas for further research on emerging decentralized learning techniques in healthcare, including developing personalized models, reducing bias, incorporating hybrid non-IID features, hyperparameter tuning, developing sufficient incentive mechanisms, and incorporating domain expertise knowledge.

INDEX TERMS Data privacy, federated learning, hybrid split-federated learning, split learning, healthcare.

I. INTRODUCTION

Deep learning has demonstrated promising outcomes in healthcare applications that support clinical data-based medical diagnosis and treatment decisions. It helps with text detection in medical laboratory reports, brain tumor segmentation and classification in Magnetic Resonance Imaging (MRI) scans, cancer diagnosis and prognosis, and many other tasks [1]. When developing deep learning models, a large amount of training data is needed for reliable performance.

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamad Forouzanfar¹.

In the healthcare domain, such training data are obtained from various clinical resources such as biological sensors, clustered patients, hospitals, medical research institutions, pharmaceutical companies, etc. In the clinical setting, accessing healthcare data for training a specific deep learning model is challenging. Normally, the amount of data related to a specific disease/condition in a single institution is limited, while obtaining data from other institutions is complicated due to privacy and data protection-related regulations. In Canada, medical data are subject to the Personal Information Protection and Electronic Documents Act (PIPEDA) and also Provincial Privacy Laws such as British Columbia's

E-Health (Personal Health Information Access and Protection of Privacy) Act, Alberta's Healthcare Information Act, and Manitoba's Personal Health Information Act.¹ Data privacy has been termed as the "most important issue in the next decade" [2]. It has gained prominence as a result of laws like the California Consumer Privacy Act (CCPA) and the General Data Protection Regulation (GDPR) of the European Union.² Thus, considerable effort is required to enable model training on private data, such as patients' medical records.

In the machine-learning literature, several solutions for data privacy-related issues have been established. Homomorphic encryption [3], differential privacy [4], Private Aggregation of Teacher Ensembles (PATE) [5], and secure multiparty computation [6] are traditional techniques for data privacy in Artificial Intelligence (AI). Federated Learning (FL) [7] is a more recent technique, involving multiple clients/institutions, each with their own data. By retaining the data locally at each client, FL enables model training using data from many institutions while avoiding data sharing and the challenges associated with establishing and maintaining large central databases. Although FL addresses the issues with data sharing, it is not without its own challenges. For example, FL assumes that each institution has sufficient computational resources to train its own version of the full deep model (called a *local model*), which is not always realistic.

Split Learning (SL) [8] was developed to tackle some of the challenges associated with FL, specifically those involving computational resources. SL allows medical institutions to implement only a portion of the model on the hardware resources available to them, while the majority of the computation is carried out at a remote server. At the same time, original data can be kept at each institution to preserve privacy. The amalgamation of FL algorithms with the SL framework facilitates a combination of these approaches, where multiple institutions/clients with limited resources can participate in training a large model. Our objective in this paper is to survey the FL, SL, and emerging hybrid Split-Federated Learning (SFL) approaches in healthcare. There are numerous challenges to such learning due to data properties (e.g., imbalanced or non-identically distributed data at different clients), privacy (embodied in various attacks), hardware heterogeneity, etc. We will review these issues throughout this survey article.

There are several related surveys on FL and SL in the literature, but relatively few focus on healthcare, particularly medical image analysis. Moreover, SFL, as a more recent concept, has not been reviewed in detail. Yang et al. [9] summarized the early FL literature for general privacy-preserving techniques. Kulkarni et al. [10] highlighted the need to personalize global models to work better for individual clients. Jin et al. [11] provided a brief overview of the most common semi-supervised FL algorithms, detailing potential approaches, contexts, and difficulties. Lim et al. [12]

presented applications of FL for mobile edge network optimization. A survey of the research addressing communication constraints in FL has been presented in [13]. A survey of FL threats has been presented in [14], highlighting the intuition, various approaches, and fundamental premises of various attacks. The authors in [15] provided a detailed analysis of recent advances and open problems in FL.

In addition, [16] discussed the current state of FL in healthcare. The reference [17] presented a comprehensive review of FL in healthcare from the perspective of data properties and applications. Chowdhury et al. [18] reviewed FL in the field of oncology with its associated future clinical directions. Majorly focusing on the biomedical space, [19] presented their survey showing the challenges and potential solutions. Rieke et al. [20] explored how FL may benefit the future of digital health. Recently, [21] provided a detailed review of the current and future research trends of FL on medical applications, highlighting FL's statistical problems such as device challenges, security, privacy concerns, etc.

As seen above, the concept of FL has been reviewed extensively in the recent literature, even in the area of healthcare. However, SL, and especially SFL, have yet to be thoroughly reviewed. This is where our survey paper differs from existing literature on the topic, as we survey all three emerging decentralized learning paradigms – FL, SL, and SFL – in the context of healthcare. This allows us to draw parallels and explore the benefits and weaknesses of all three learning methodologies from the point of view of healthcare. We focus particular attention on medical image analysis, one of the main areas of healthcare research. We review the existing publicly available datasets that can be used in medical image analysis research. To the best of our knowledge, this is the first survey to present a unified treatment of existing FL, SL, and SFL approaches with a view toward medical image analysis, along with related datasets.

The structure of the paper is as follows. Section II presents the preliminaries on centralized and decentralized learning needed for the remainder of the paper. Section III reviews the applications in healthcare from existing literature on FL, SL, and SFL, with a special view toward medical image analysis. The general challenges faced by the three emerging decentralized learning paradigms and proposed solutions are discussed in Section IV. Section V discusses challenges specific to healthcare, current trends, and possible future research directions. Finally, Section VI concludes the paper with a summary of the main conclusions.

II. PRELIMINARIES

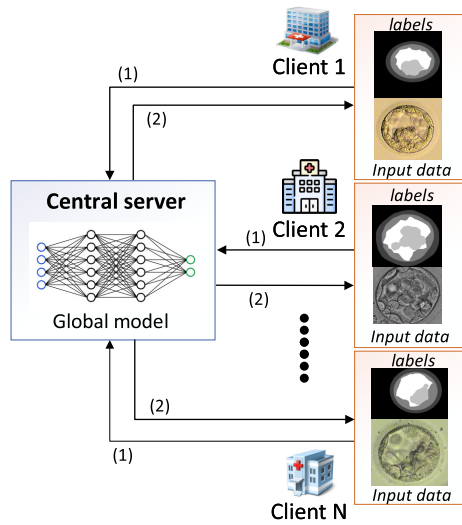
This section contains descriptions of the survey's primary topics. We describe centralized machine learning and decentralized machine learning, which are federated learning, split learning, and hybrid split-federated learning, in-depth.

A. CENTRALIZED MACHINE LEARNING

Centralized machine learning is carried out at one location, usually a powerful server or the cloud. The basic premise

¹<https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada>

²<https://www.osano.com/articles/data-privacy-laws>



(1) Transmission of input data and labels from clients
 (2) Transmission predictions from central server

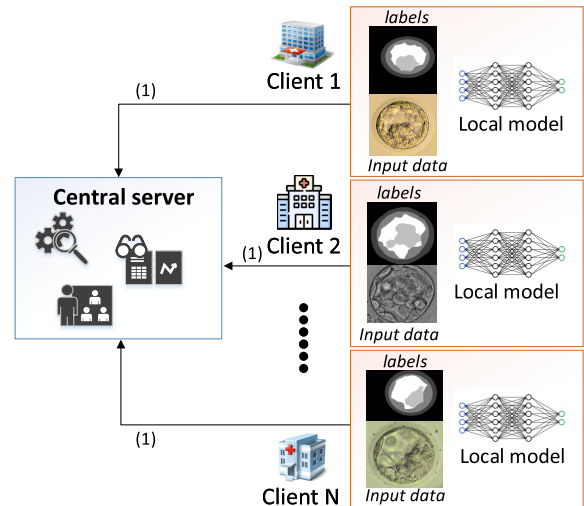
FIGURE 1. Centralized machine learning and inference.

is that all data is available at that location. In a healthcare setting, this means that all clients (clinics, hospitals) need to transmit their labeled data to the server where the global model is being trained. Upon training, the inference may also be carried out at the central server, especially if the clients lack the computational infrastructure to run the model. In this case, clients may upload their data to the server and get inference/prediction results back from the server. This framework is illustrated in Fig. 1. Health service providers have used cloud-centralized frameworks, such as Microsoft Azure Healthcare, to promote healthcare solutions.

While centralized solutions address the lack of powerful computational infrastructure for the clients, the main issue with their use is the protection of health data. Both training and inference require uploading data to another location, which may be in a different jurisdiction and subject to different (if any) health-related privacy protection regulations. For this reason, centralized machine learning (and inference) in healthcare may not be feasible except in very special cases.

B. DECENTRALIZED MACHINE LEARNING

Decentralized learning can alleviate issues with data protection by keeping data at the clients. In the most basic setting, each client would train its own local model. However, this type of solution has many drawbacks. First, clients might not have the infrastructure to train large models, so they would be forced to compromise on the ultimate accuracy of their model, compared to a larger centrally-trained model. Second, since local models are trained only on the client’s local data, they will likely not generalize as well as a centrally-trained model with access to more data. Finally, having different models for different clients may cause inequity in health outcomes, as the clients with more data (e.g., those in urban



(1) Transmission predictions to central server

FIGURE 2. Ensembling as a form of distributed inference.

centers with larger populations and better infrastructure) are likely to end up with better, more accurate models.

Although different local models can be pooled together through ensembling techniques [22], this would still require sharing data and/or predictions from different clients. A simple version of ensembling is illustrated in Fig. 2, where predictions made by the different local models are aggregated at the central server, but for this to work, local models would have to operate on the same input. Federated Learning (FL) [7] was developed as a decentralized learning strategy that avoids such data sharing, yet is able to create a model that benefits from all clients’ data. In the remainder of this section, we present the preliminaries of FL, as well as related decentralized learning strategies of Split Learning (SL) and hybrid Split-Federated Learning (SFL).

1) FEDERATED LEARNING

Federated Learning [7] is a decentralized machine learning approach that enables multiple clients (data centers, organizations, remote devices, ...) to train a global model without sharing their data. It was first introduced by Google to improve next-word prediction models on Android mobile phones. In FL, a model is trained using data from different clients, and the process is controlled by a central server, as illustrated in Fig. 3. The basic FL procedure has the following steps. Steps 2-5 comprise one *federated round* or *global epoch* and are repeated until stopping criteria are met:

- Step 1:** The server transmits the initial model to the clients.
- Step 2:** Clients train their *local models* with their own data for a certain number of *local epochs*.
- Step 3:** Each client sends the parameters of its local model to the server.

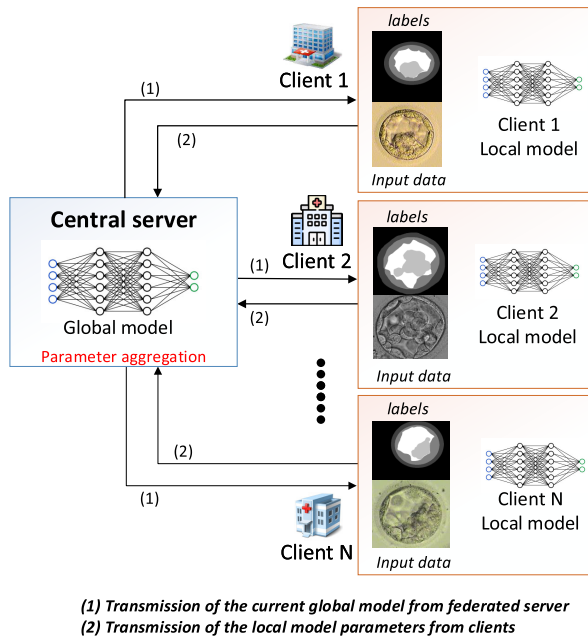


FIGURE 3. Illustration of federated learning.

Step 4: The server aggregates the model parameters from various clients using an averaging scheme and sends the aggregated parameters back to the clients.

Step 5: Clients update their local models with the aggregated model.

According to [15], FL methods can be divided into *cross-silo*, where FL takes place among data centers or organizations, and *cross-device*, where FL takes place among IoT/mobile devices, which have lower computational resources. Another classification exists considering clients' features and labels [11], [15]: *horizontal FL*, *vertical FL*, and the *federated transfer learning*. In horizontal FL, datasets at different clients share the same feature space, but different label spaces. An example would be different units in the same hospital, which use the same input data (patients' health records) but are interested in different diagnostics and/or outcomes. In vertical FL, clients share the same label space, but different feature spaces. An example would be oncology units in different hospitals or health authorities, which are interested in the same outcomes but may differ in the type of patient data they have access to. Finally, in federated transfer learning, neither feature space nor label space is shared across clients. An example would be the generic problem of "cancer detection" across different hospitals or health authorities. While available inputs and expected outputs may differ across different clients, knowledge of common characteristics of cancer may be transferable and useful for each client's model.

FL can be applied to any machine learning model, although the focus in the current literature is on Deep Neural Networks (DNNs). Examples include Multi-Layer Perceptron (MLP) [23], which is common for making predictions using tabular medical data; Convolutional Neural Networks (CNNs), which are common for medical image

analysis [24], and other models such as auto-encoders [25], Generative Adversarial Networks (GANs) [26], Long Short-Term Memory (LSTM) networks [27], Support Vector Machines (SVMs) [28], etc.

Although FL addresses data protection and privacy concerns, it requires all clients to train their local models, which are of the same size as the global model. This is challenging when considering modern DNNs with hundreds of millions of parameters since the clients might not have the required computational infrastructure and associated Information Technology (IT) support. Another distributed learning strategy – Split Learning – was developed to address this computational imbalance between the clients and the server.

2) SPLIT LEARNING

In Split Learning (SL), a DNN is split into several parts, which can be located on various devices and/or servers. In its most basic form, a front-end of a DNN (usually the initial few layers) is located on a client device, and a more computationally-demanding back-end is located on a server. SL is the learning counterpart of split inference, also known as *collaborative inference* or *collaborative intelligence* [29], and was introduced as SplitNN [8], where the authors have described the split network training approaches with and without label propagation.

Various SL configurations are possible depending on where the data and the labels are located: simple *vanilla SL*, where the client keeps its data but shares labels with the server (Fig. 4(a)); *U-shaped SL*, where the client keeps both its data and labels, without sharing with the server (Fig. 4(b)); and *SL with vertically partitioned data*, which involves multiple clients, each keeping its own data but sharing labels with the server (Fig. 4(c)). Another classification can be made depending on how the DNN is split between the client(s) and the server. In *horizontal SL*, the split is introduced between layers (as shown in Fig. 4), so that each layer is fully executed on a single device/client. In *vertical SL* [30], layers themselves can be split across multiple devices/clients.

The basic steps of vanilla SL (Fig. 4(a)) are outlined below. Steps 2-5 are repeated until the stopping criteria are met. The more sophisticated versions of SL are extensions of this procedure, involving additional transmission of data and gradients between the server and the client(s).

Step 1: A DNN model is split between the client and the server, taking into account computational resources available at the client, communication link between the client and the server, etc.

Step 2: A batch of data is loaded and passed through the client-side model front-end, and the features (also known as *smashed data*) are computed and sent to the server.

Step 3: The server receives the features, processes them through the back-end, and computes the outputs.

Step 4: The server compares the outputs with the ground-truth labels, computes error gradients, back-propagates them through the model back-end

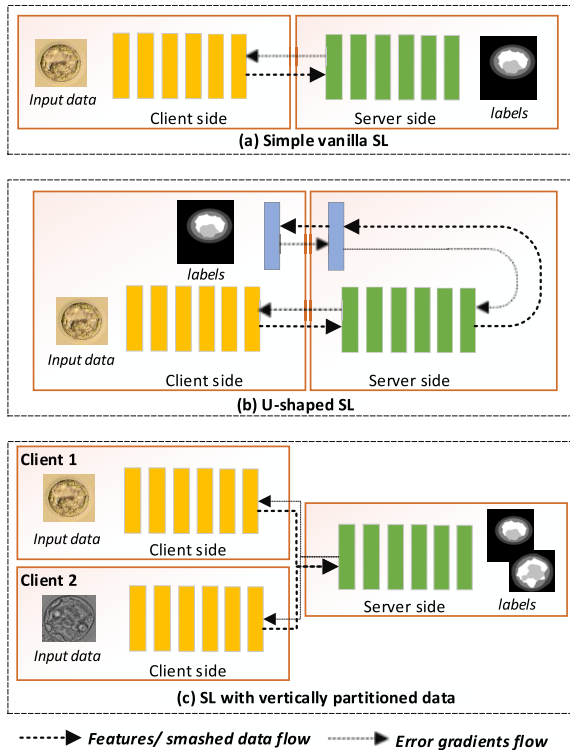


FIGURE 4. Split learning configurations.

(updating model parameters along the way), and sends the gradients from the split layer back to the client.

Step 5: The client receives the error gradients and back-propagates them through the front-end, updating the model parameters along the way.

In the literature, SL has been demonstrated on a number of popular datasets such as the Modified National Institute of Standards and Technology (MNIST) database [31], ImageNet Large Scale Visual Recognition Challenge (ILSVRC) dataset [32], and Canadian Institute For Advanced Research (CIFAR) datasets [33], and it was shown [8] that – assuming perfect communication links – its performance is similar to conventional centralized learning. In practice, SL has been used in real-world IoT settings [34], where the overhead, training time, power consumption, and memory usage were studied. Reference [35] described an efficient model training in IoT systems. Their approach, termed ARES (Adaptive Resource-Aware Split Learning for the Internet of Things), accounted for time-varying communication link throughput and computing resources. SplitNet [36] is an SL approach that splits a deep network into a tree of sub-networks, which allowed for a simple model parallelization while simultaneously reducing the number of parameters and computations.

SL solves the issue of computational imbalance between the client(s) and the server that exists in FL, and in certain cases has lower communication requirements than FL. For example, [37] found that SL is more communication-efficient³ than FL when there are many clients or when the

³Needs to communicate fewer data.

model is large. However, unlike FL, the basic SL schemes do not offer a way for clients to collaborate among themselves to increase the pool of data and potentially train better models. For this reason, hybrid split-federated learning (SFL) approaches were developed to combine the best of both worlds.

3) HYBRID SPLIT-FEDERATED LEARNING

As seen above, FL and SL each have their own advantages and disadvantages. Research has been done to combine these two approaches to yield the best of both worlds. We will refer to such combined approaches as hybrid Split-Federated Learning (SFL). The overall goal of SFL is to appropriately use the computational resources available at the clients and the server (like SL), while allowing clients to collaborate in model training and keep their data private (like FL).

The first reported SFL approach is SplitFed [38]. It eliminates FL’s and SL’s drawbacks with their modified architectural design. SFL enhances overall data privacy and model robustness compared to that of FL and SL [34], [38], [39], [40], [41]. An SFL model allows training better models with enhanced performance since it combines the strengths of SL and FL architectures.

Steps 2-10 comprise one *communication round* or *global round* or *global epoch* and are repeated until stopping criteria are met:

- Step 1:** Federated Server (FS) chooses a statistical model; which is supposed to be the client-side model to be trained (client-side global model).
- Step 2:** FS transmits this initial client-side global model to several units (clients).
- Step 3:** Clients train their client-side models locally with their own data (Forward propagation (FP)).
- Step 4:** Clients send their model features to the main server.
- Step 5:** The main server processes the forward propagation on its server-side model separately and in parallel.
- Step 6:** The main server does the backpropagation (BP) on the server-side model.
- Step 7:** The main server sends the gradients to the respective clients.
- Step 8:** The main server updates its server side global model.
- Step 9:** Each client receives the gradient update and performs the back-propagation on its client-side local model, and computes its gradients.
- Step 10:** Fed server updates its model.

SplitFed appears in two main variants: *SplitFedv1 (SFLV1)* and the *SplitFedv2 (SFLV2)* [38]. In SFLV1, the number of server-side sub-networks equals the number of clients. SFLV2 performs sequential server-side sub-network training over the smashed data of the clients and shares only one copy of the sub-network on the server-side. The authors in [34] also discussed a generalized version of SFL (SFLG) that merges SFLV1 and SFLV2 enabling a varying number of

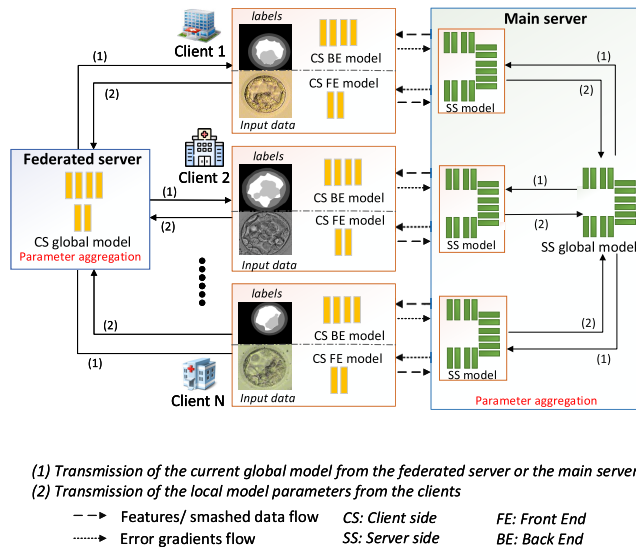


FIGURE 5. Illustration of SplitFedv1 (SFLV1) approach without label sharing.

server-side sub-networks. *SplitFedv3 (SFLV3)* [40] is another new version that involves unique client-side sub-networks and averaged server-side sub-network. The problem of catastrophic forgetting is avoided in SFLV3 due to averaging of the server-side network. Fig. 5 illustrates the SFLV1 approach without label sharing. Similar designs can be occupied for SFLV2, SFLG, and SFLV3.

Research has been particularly done to study the effect of client-side and server-side model synchronization. The Multi-Head Split Learning (MHSL) approach [41], [42] studied SplitFed without client-side model synchronization. The performance of ML models that did not use client-side parameter aggregation was comparable to that of those that did. The similar research Parallel SL with Split-Layer Gradient Averaging and Learning Rate Splitting (SGLR) [43] discussed a scalable SL framework. In contrast to SplitFed, the server in SGLR broadcasted a common gradient averaged at the split layer, simulating FL without any further client interactions.

The number of clients and servers was equal in the Federated Split Learning framework suggested in [44]. It trained the server-side model corresponding to each client on a separate server. Each client-server pair was simultaneously trained. After each global round, the server-side models were aggregated by an FS and updated on each server. This method leveraged the PySyft and PyGrid libraries. The authors confirmed better accuracy and the privacy guarantee, provided that each client has a reasonable amount of training data. The authors in [45] showed that the vanilla SL could have the possibility of overfitting due to the sequential nature of training and proposed the Parallel Split Learning Concept (PSL) that prevents overfitting.

Federated Deep Learning with Private Passport (FDL-PP) [46] considered the layers prior to the split as private layers and the following layers as public layers. They

suggested adopting FL in public layers to prevent any attacks at the split. The server only aggregated the model parameters of the public layers. PyVertical [47] is a framework that supports vertical federated learning using a split neural network. It enabled a data scientist to maintain raw data on an owner's device while training neural networks on vertically partitioned data features among several owners. Adasplit [48] is another hybrid approach of SL and FL, that enabled efficient scaling to SL to low resource scenarios in reducing bandwidth consumption and improving performance across heterogeneous clients. The authors in [49] suggested a hybrid approach to updating client- and server-side models simultaneously through local-loss-based training. Losses were calculated for each local split.

FedFly [50] is an approach to migrate a partitioned neural network when devices move (*device mobility challenge*) between the edge servers during FL training. FedLite [51] showed high communication costs associated with model splitting. The authors solved this problem by using a new clustering scheme and a gradient correction method to compress the extra communication. LocFedMix-SL [52] showed that existing parallel SL algorithms achieve neither scalability nor fast convergence since there is an imbalance between the FP and the BP updates. To fix this problem, the authors added local parallelism, FL, and mixup data augmentation techniques to parallel SL to keep the FP and BP updates balanced.

In accordance with the principle of “first-parallel-then-sequential”, federated parallel training in SL was made possible by the Cluster-based Parallel SL (CPSL) method [53]. Each training round in CPSL was divided into two phases: *sequential inter-cluster training* followed by *simultaneous intra-cluster training*. Client devices were organized into several clusters. All clients in the same cluster cooperated with the server to carry out parallel training, similar to SplitFedv2. After the completion of a single round of intra-cluster training, each client's client-side model was sent to the server for aggregation and updating. The client model was then initialized in the following cluster to begin intra-cluster training using the updated model.

Hybrid Split and Federated Learning (HSFL) scheme [54] was proposed to deal with challenges in collaborative learning in highly diverse IoT devices with heterogeneous resources and data distributions. This framework organized clients into two groups: one group performed FL, and the other group performed SL.

Figure 6 shows a summary of the described emerging decentralized learning approaches along with their evolutionary timeline. Table 1 outlines a comparative evaluation of FL, SL, and SFL under different criteria.

III. DECENTRALIZED LEARNING APPLICATIONS IN HEALTHCARE

This section reviews State-of-the-art (SOTA) FL, SL, and SFL applications in healthcare. Readers are requested to

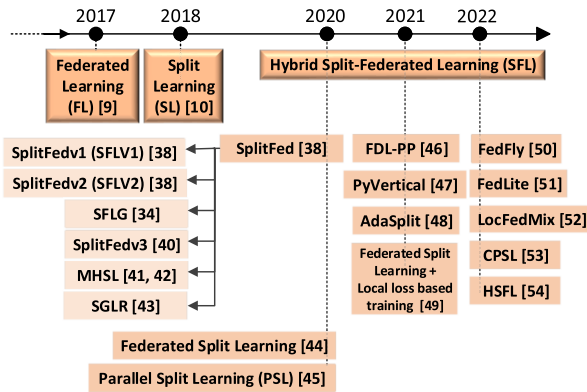


FIGURE 6. Evolutionary timeline of federated, split, and hybrid split-federated learning.

TABLE 1. Characteristics of FL, SL, and SFL.

Characteristics	FL	SL	SFL
Clients access to other clients' data	No access	No access	No access
Client responsible for training	Full model	Partial model	Partial model
Training process	Parallel training	Sequential training	Parallel training
Computation requirement of clients	High	Less and resource constrained clients perform well	Less and resource constrained clients perform well
Server computation	Lower server computation compared to SL	Higher server computation compared to FL	Higher server computation compared to FL
Bandwidth requirement	High	Low	Low
Model convergence	Faster than SL	Slower than FL	Closer to FL
Level of privacy	Better privacy preservation than CML	Better privacy preservation than FL	Better privacy preservation than FL
Server computation	Lower server computation compared to SL	Higher server computation compared to FL	Higher server computation compared to FL
Model aggregation	Conducts model aggregation	No model aggregation	Conducts model aggregation

follow Table 5 of the Appendix for descriptions of the publicly available medical imaging datasets mentioned in this section.

A. FL IN HEALTHCARE

FL has found its way into healthcare, from medical imaging use cases to Electronic Health Record (EHR) management as outlined in Fig. 7. SOTA applications in medical imaging

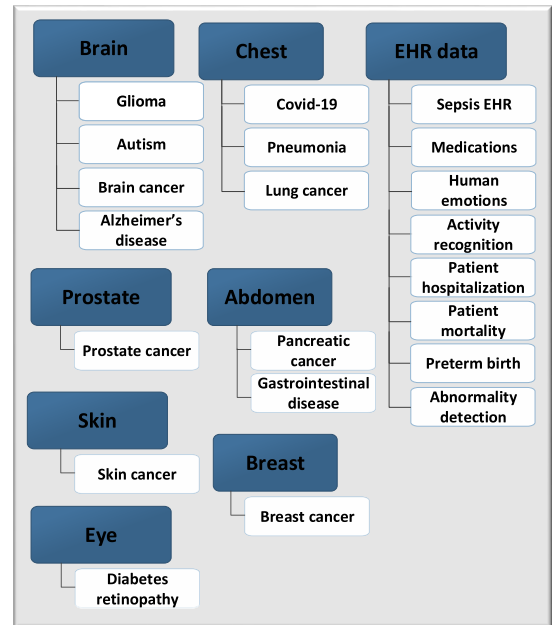


FIGURE 7. FL applications in healthcare.

cover a wide range, including the brain, chest, skin, breast, eye, prostate, skin, and abdomen which are discussed in this Section. Some other interesting advanced studies associated with FL in the healthcare domain include patient similarity learning [55], patient representation learning [56], phenotyping [57], and predictive modeling [28] etc.

1) FL APPLICATIONS ASSOCIATED WITH BRAIN

The first reported FL used for multi-institutional applications in medical imaging was done by Intel in collaboration with the Centre for Biomedical Image Computing and Analytics at the University of Pennsylvania for brain image segmentation [58]. The data was a collection of multi-institutional, multi-model brain MRI scans from glioma patients that are publicly available as part of the Brain Tumor Image Segmentation Benchmark (BraTS) challenge 2018 [59]. This system was based on a U-Net. The server aggregated the model parameters of chosen clients to form the global model. They proved that the semantic segmentation performance of the federated model is superior to that of the centralized model. FedDis [60] is a disentangled FL application for unsupervised brain pathology segmentation. The model is trained with two brain MRI scan datasets: Open Access Series of Imaging Studies-3 (OASIS-3) [61] and Alzheimer's Disease Neuroimaging Initiative-3 (ADNI-3) [62]. The model was tested with two public Multiple Sclerosis (MS) datasets (MSLUB [63], MSISBI [64]) and an in-house MS and glioblastoma database. Silva et al. [65] suggested an FL architecture, to securely access and meta-analyze any biomedical data without disclosing personal information. They investigated brain structural relationships across

diseases and clinical cohorts. Li et al. [24] proposed a privacy-preserving approach for multi-site Functional Magnetic Resonance Imaging (fMRI) classification using federated transfer learning and domain adaptation. Experiments were conducted with fMRI data from the Autism Brain Imaging Data Exchange (ABIDE) dataset [66]. Promising results were achieved while improving neural image analysis performance and identifying valid disease-related biomarkers. FedBioMed [67] proposed a robust and scalable open-source FL framework accommodating different models and optimization methods and tested it with brain imaging datasets from four institutions. The authors in [68] investigated the feasibility of applying differential privacy techniques by applying their FL algorithm on the BraTS 2018 dataset [59].

2) FL APPLICATIONS ASSOCIATED WITH CHEST

FL research related to the chest has mostly been done with pneumonia Computerized Tomography (CT) lung cancer scans, and coronavirus disease (COVID)-19 CT scans. An interesting FL application, C-DistriM (Chained Distributed Machine learning) delivered a superior performance with Non-Small Cell Lung Cancer (NSCLC) radiomics dataset [69] in predicting two-year lung cancer survival [70]. With the COVID-19 pandemic, a major focus of AI was shifted to institutional data collaboration through FL. The authors in [71] showed the viability of an FL approach for identifying CT anomalies associated with COVID-19. The authors investigated FL strategies to create a COVID-19 medical image diagnosis AI model with strong generalizability on seven international datasets. The public dataset developed by Radiology Ai One-Stop solution (RAIOSS) [72] was also used during the model validations. The FL research [73] performed collaborative training of multiple medical institutions' models using the two public chest X-ray screening datasets: Montgomery County chest X-ray dataset (MC) [74] and Shenzhen chest X-ray dataset [74]. The authors investigated several key specificities of FL settings, including non-(Independent and Identically Distributed) IID and unbalanced data distributions. The FL research in [75] developed an abnormal chest radiograph classification model on GoogLeNet-22 and ResNet-50 using the public Chest X-ray (CXR) dataset [76]. The authors discussed the challenges of sample size and label distribution variability in FL.

3) FL APPLICATIONS ASSOCIATED WITH SKIN

FedPerl [77] is a semi-supervised FL method developed to classify skin carcinoma data. It encouraged the community to use collaborative learning to generate more accurate pseudo-labels for unlabeled data. They used 71,000 skin lesion images from four publicly available datasets: International Skin Imaging Collaboration (ISIC)19 [78], Human Against Machine with 10,000 training images (HAM10K) [79], Derm7pt [80], PAD-UFES [81]. The FL model in [82] focused on multimodal melanoma detection using ISIC19 [78] data. The FL model's

performance was comparatively better than that of the centralized model. A gradient aggregation method that better extracts the shareable information from multiple local servers is presented in [83]. The authors applied it to skin lesion segmentation using seven datasets, including HAM10K [79] and Derm7pt [80] datasets. FedMix [84] is a federated skin lesion segmentation model developed with the HAM10K dataset [79].

4) FL APPLICATIONS ASSOCIATED WITH BREAST

FedMix [84] also tested their FL model using the breast tumor datasets: Breast Ultrasound (BUS) [85], Breast Ultrasound Image Segmentation (BUSIS) [86], and UDIAT [87]. Using a cooperative strategy with seven clinical institutes worldwide, Roth et al. [88] developed an FL model for breast density categorization. According to their experimental findings, the FL model beat the individually-trained models on the local data of each institute by an average of 6.3%. Additionally, when the FL model was assessed using external testing datasets from other participating sites, a relative improvement of 45.8% in the model's generalizability was established. The memory-aware curriculum learning method for FL is proposed in [89]. Their findings with three mammography imaging datasets demonstrated the benefits of federated adversarial learning for multi-site breast cancer categorization.

5) FL APPLICATIONS ASSOCIATED WITH EYE

The authors in [90] evaluated the performance of an FL framework for DNN-based retinal microvasculature segmentation and referable diabetic retinopathy classification using Optical Coherence Tomography (OCT) and OCT angiography (OCTA) images. FL is applied to diabetic retinopathy detection in [91]. The authors used five datasets: EyePACS [92], Methods to Evaluate Segmentation and Indexing Techniques in the field of Retinal Ophthalmology (MESSIDOR) [93], Indian Diabetic Retinopathy image Dataset (IDRiD) [94], Asia Pacific Tele-Ophthalmology Society (APTOS) [95], and University of Auckland (UoA) diabetic retinopathy database [96]. The model was investigated with three approaches: standard transfer learning, federated averaging, and federated proximal framework. The authors in [75] also evaluated their models with a diabetic retinopathy detection task using the Diabetic Retinopathy binary classification dataset [97].

6) FL APPLICATIONS ASSOCIATED WITH PROSTATE

The authors in [98] introduced a flexible FL framework for cross-site training, validation, and evaluation of deep prostate cancer detection. Their method used an abstract representation of the model architecture and data, enabling NVFlare FL framework to be used to train prototype deep learning models. Prostate biopsy data were collected from two University of California research hospitals. Their method showed improvements in prostate cancer detection and classification compared to that of SOTA. A personalized FL model was

developed for prostate segmentation in [99]. The authors suggested an adaptation strategy that enables unique model architecture for each client. The suggested strategy was assessed on MSD1 [100], PROSTATEx [101], PROMISE12 [102], and National Cancer Institute–International Symposium on Biomedical Imaging (NCI-ISBI) [103] datasets. It was demonstrated to enhance the local models' performance following adaptation. The authors in [104] applied a Federated Cross Learning (FedCross) algorithm on prostate cancer MRI segmentation using the same datasets as in [99].

7) FL APPLICATIONS ASSOCIATED WITH ABDOMEN

The authors in [105] investigated the automatic segmentation of pancreatic cancer CT scans with three datasets including the publicly available MSD1 [100] pancreatic dataset and Synapse dataset [106]. The authors applied heterogeneous optimization techniques to achieve significant performance. MoNet [107] is a highly optimized federated pancreatic CT scans segmentation algorithm, that used the MSD1 [100] pancreatic dataset.

8) FL APPLICATIONS ASSOCIATED WITH EHR DATA

EHRs are a vital source of real-world healthcare data. EHRs contain information such as a patient's medical history, diagnoses, medications, treatment plans, immunization dates, allergies, radiology images, and laboratory test results. Several FL research with EHR is available in the literature. They cover areas of medications [108], [109], [110], human emotions [108], activity recognition [111], [112], EEG classification [113], personalized wearable healthcare applications [111], patient hospitalization [114], patient mortality [114], and preterm birth data [115].

B. SL IN HEALTHCARE

The applicability of the SL in healthcare was first shown by the authors in [8], [116] with their SplitNN model pointing to its promising advantages. Then the first application of SL based approach in the medical field was reported in [117]. The authors applied their model to a binary classification of fundus images and a multi-label classification of chest X-rays. The authors in [118] first showed the clinical feasibility of SL. They discussed the inference performance aspects, convergence rates, computational efficiency, and communication requirements in relation to clinical feasibility. SL applications in healthcare still only cover a limited range, including the chest, eye, brain, and EHR.

1) SL APPLICATIONS ASSOCIATED WITH CHEST AND EYE

The binary classification task in [117] used a dataset of 9,000 fundus images, and the multi-label classification task used a dataset of 156,535 chest X-rays. As the first study in the SL and medical settings, their research paved the way for future developments in the collaborative training of DNNs. The authors in [118] developed a 121-layer denseNet

for multi-label classification of chest X-ray images from the Chexpert dataset [119].

2) SL APPLICATIONS ASSOCIATED WITH BRAIN

The authors in [118] also used their approach for brain tumor segmentation with the BraTS datasets [59]. One another brain tumor classification task is presented in [120]. The authors vertically distributed brain MRI scans and classified them as healthy or tumorous using a partitioned neural network.

3) SL APPLICATIONS ASSOCIATED WITH EHR DATA

An SL framework for Electrocardiogram (ECG) classification was proposed in [121]. The system was trained on the Physikalisch-Technische Bundesanstalt (PTB-XL) ECG dataset [122]. Results showed a significant reduction in the computation and communication overhead with minimal performance loss. SL approach in [123] was based on a 1D CNN model, where heart abnormalities are detected with ECG data. The authors studied the effects of privacy preservation on model performance.

C. SFL IN HEALTHCARE

The first-ever SFL approach in healthcare was initiated by The authors in SplitFed [38], by using their proposed model on a skin lesion segmentation task. Similar to SL, SFL applications in healthcare are comparatively less prevalent than FL. These cover areas in healthcare related to the skin, chest, bones, stomach, brain, eye, cervix, and EHR.

1) SFL APPLICATIONS ASSOCIATED WITH SKIN

SplitFed [38] authors evaluated their experiments using the dermatoscopic images of the HAM10K dataset [79]. MHSL [41], [42], which is described in the Section II-B3 was tested on ResNet-18 and 1D CNN architectures with the HAM10K [79], MNIST [31], and CIFAR [33] datasets.

2) SFL APPLICATIONS ASSOCIATED WITH BRAIN

Split-U-Net [124] applied SFL with vertical FL on four institutions for multi-model brain tumor segmentation. Experiments were done using the Medical Segmentation Decathlon (MSD1) brain tumor dataset [100]. It was the first application of SFL to a multi-modal image segmentation task. The authors quantified the amount of data leakage in biomedical image segmentation and presented defense strategies. SplitAVG and SplitAVG-V2 [125] are two interesting research that focused on heterogeneity-aware FL. They were tested on a brain tumor segmentation model with the BraTS datasets [59].

3) SFL APPLICATIONS ASSOCIATED WITH EYE

SplitAVG and SplitAVG-V2 were also evaluated in a diabetic retinopathy binary classification task. It achieved 96.25% of accuracy on the Diabetic Retinopathy binary classification dataset [97].

4) SFL APPLICATIONS ASSOCIATED WITH CHEST, BONES, STOMACH, AND EHR

“Spatio-temporal split learning” presented in [126] allowed collaboration among privacy-sensitive organizations. They spatially spread many clients to cover a variety of datasets from various participants. The algorithm was tested on MURA X-ray images [127], and COVID-19 CT scans from the COVID-CT dataset [128]. FL and SL have achieved 95.7% and 98.5% respectively, for the classification task. The same authors came up with “multi-site split learning” [129] that enabled secure medical data transfer between hospitals. They explored the optimal number of clients via experimental analysis. They empirically investigated the optimal data split ratio for the best-split learning performance. Federated Learning on Medical Datasets using Partial Networks (FLOP) [130] presented a framework that only shared a partial model between the server and the clients, keeping the rest of the layers private within the client’s own space. Experiments were carried out with the COVIDx dataset [131] and the Kvasir dataset [132]. FLOP achieved better performance while reducing privacy and security risks.

Apart from the above-mentioned, [133] applied Vertical SL to five medical datasets related to cervical cancer, diabetes, heart disease, stroke, and stroke rehabilitation. The authors evaluated the impact of different networks and feature distributions on predictive performance and compared the performance with the centralized architecture.

IV. DECENTRALIZED LEARNING CHALLENGES AND SOTA SOLUTIONS

This section gives a summary of the major challenges in FL, SL, and SFL and the SOTA solutions in tackling them.

A. CHALLENGES ASSOCIATED WITH FL

Existing research on FL is mostly focused on reducing the effects of corrupted or noisy clients, dealing with statistical and system heterogeneity issues, improving communication efficiency, addressing privacy issues, improving fairness, handling biased sources, and dealing with system challenges. Researchers have taken steps to evaluate their algorithms under different challenging circumstances. These techniques have helped in developing more robust FL algorithms.

1) RESEARCH ON HANDLING CORRUPTED OR NOISY CLIENTS IN FL

Data annotation is a complex task, even for experts, and depending on the nature of the data and the considerations for labeling, it is time-consuming. Since in an FL setting, local data are gathered by different clients, it is difficult to ensure that the data are annotated correctly. Furthermore, clients may have different standards according to their domain expert’s knowledge. Thus, clients’ annotated data may have distinct distributions compared to that of a centralized annotation. As a result, data corruption or data mislabelling could occur. Similarly, clients’ local data might include noise. Noise can

TABLE 2. Prior research on FL handling corrupted and/or noisy clients issues.

Research	Approach	Results
FedMix [84]	Proposed an adaptive weight assignment procedure when clients have different levels of annotations (mislabeling) such as strong pixel-level labels, weak bounding box labels, and weakest image-level class labels.	Resulted in 80.6% and 85.3% segmentation accuracy with breast and skin data, respectively.
FedNCL [134]	Presented data quality measurement and noise-robust aggregation approach when clients have noisy data.	Boosted performance of different DNNs with noisy clients.
Tuor et al. [135]	Proposed a method where clients could select highly relevant data for the FL process, eliminating noisy and irrelevant data.	Showed up to 25% accuracy improvement compared to training with all the data.
Fang et al. [136]	Proposed a method for dealing with noisy and heterogeneous clients by utilizing a client confidence re-weighting scheme.	Proved efficiency in minimizing model homogeneity and heterogeneity.
Yang et al. [137]	Proposed a method of server cooperating with local models to maintain consistent decision boundaries when clients have noisy data.	Resulted in a robust FL algorithm to deal with noisy clients.
ARFL [138]	Presented a model that jointly learns the global model and the weights of local updates when clients have corrupted data.	Performed robustly against label flipping, label shuffles, and noisy features.
Yang et al. [139]	Proposed a client selection framework for FL when clients have label noise based on <i>copeland score</i> and <i>multi-arm bandits</i> .	Resulted in an accurate algorithm on the server side without extra client operations.

arise from clients using different data acquisition sources or devices, communication barriers, system errors, etc. In the collaborative FL environment, if the clients’ data are corrupted or noisy, the global aggregated model produces wrong predictions. Table 2 describes prior research on handling corrupted or noisy client issues in the FL network [84], [134], [135], [136], [137], [138], [139].

2) RESEARCH ON HANDLING HETEROGENEITY ISSUES IN FL Heterogeneity is defined as the quality or state of being diverse in character or content. Although data heterogeneity is not new in machine learning, it is significantly more prominent in FL than in CML. Heterogeneity is classified into two categories in FL: *Statistical heterogeneity*, which is the non-IID data across the network, and the *Systems heterogeneity*, which is the significant variability of system characteristics [7], [140], [141]. The systems’ characteristics include storage, computational, and communicational

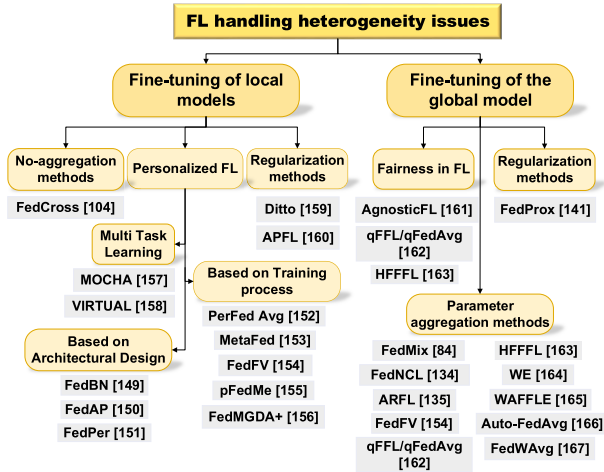


FIGURE 8. FL handling heterogeneity issues.

capabilities (memory and CPU/GPU), network connectivity (3G, 4G, 5G, or Wi-Fi), and power capabilities (battery level).

The bounds on convergence caused by heterogeneity were theoretically presented in [142], and the impact of adjusting heterogeneity on model performance was empirically demonstrated in [143]. The challenges of heterogeneity for FL were discussed in [140]. The authors in [144] demonstrated how the clients’ local model weights diverge as a result of the data heterogeneity. New insights for protecting against the non-uniformity introduced by data heterogeneity in FL as a solution for backdooring attacks were shown in [145]. FedKL [146] used federated reinforcement learning to tackle heterogeneity issues. Data problems and resource heterogeneity in FL were discussed in [147]. The authors in [148] came up with the concept of *Virtual Homogeneity Learning (VHL)*, where a virtual dataset addresses data heterogeneity issues. VHL improved the speed of convergence and performance generalization drastically.

Fig. 8 illustrates a categorical representation of FL in handling heterogeneity issues. Methods for handling heterogeneity can be categorized into two groups: fine-tuning local models and fine-tuning the global model. The local model fine-tuning group utilizes personalized FL methods, regularization methods, and non-aggregated methods of training. A variety of parameter aggregation variants (Table 3), fairness tackling methods, and regularization methods are used in global model fine-tuning.

SOTA on personalized FL can be categorized based on the model’s architectural design, training process, and multi-task learning. FedBN [149], FedAP [150], and FedPer [151] come under the personalizing models based on architectural design. PerFed Avg [152], MetaFed [153], FedFV [154], pFedMe [155], and FedMGDA+ [156] come under personalizing approaches based on the training process. MOCHA [157] and VIRTUAL [158] designed FL for multi-task learning. FedCross [104] presented an FL approach that doesn’t involve parameter aggregation, where it sequentially trained the global model across different clients in a

TABLE 3. Federated parameter aggregation in depth.

Parameter aggregation algorithm	Criteria for parameter aggregation	Addressed FL challenge
FedMix [84]	Data and pseudo-label quality.	Clients with mislabeled data.
FedNCL [134]	Data distribution, data quality, and clients’ training performance.	Clients with different noise distributions.
ARFL [138]	Minimized the weighted sum of empirical risks for clients with a regularization term. Weights were allocated by comparing the empirical risk of each client with the average empirical risk of the best P clients. Optimal weights were obtained using <i>Karush-Kuhn-Tucker (KKT)</i> conditions.	Non-IID data distribution, clients with data corruption.
FedFV [154]	Used <i>cosine similarity</i> to detect gradient conflicts and iteratively eliminated conflicts by modifying the direction and the magnitude of the gradients.	Heterogeneous and imbalanced data, Client selection, Client dropping out.
WE [164]	Euclidean distance between clients’ updates.	Non-IID data distribution.
WAFFLE [165]	Euclidean distance between clients’ updates.	Non-IID data distribution.
Auto-FedAvg [166]	Data distribution and clients’ training performance.	Non-IID data distribution.
qFedAvg/qFFL [162]	Clients’ training performance.	Non-IID data distribution.
HFFFL [163]	Data distribution.	Non-IID data distribution.
FedWAvg [167]	Number of forgetting examples in each round.	Label imbalanced in extremely Non-IID data distribution.

round-robin manner. Ditto [159] and Adaptive Personalized Federated Learning (APFL) [160] performed local model fine-tuning using regularization methods.

Fairness in FL is inspired by fair resource allocation and encourages equitable or uniform resource distribution throughout a federated network. Agnostic Federated Learning [161], q-Fair Federated Learning (qFFL)/qFedAvg [162], and Hierarchically Fair Federated Learning (HFFFL) [163] are the most distinguished research on fairness in FL literature. FedProx [141] is a popular regularization method that facilitates stable convergence in heterogeneous settings. It used a proximal term to the local sub-problem to minimize the impact of varying local updates on clients.

3) RESEARCH ON HANDLING COMMUNICATION EFFICIENCY ISSUES IN FL

In FL settings, communication efficiency is a popular topic with several SOTA studies. FedKD [143] showed a communication-efficient FL algorithm based on knowledge distillation and gradient compression techniques. It reduced communication costs (by 94.89%) and achieved competitive results on medical Named Entity Recognition (NER) datasets. FedPAQ [168] is another communication efficient scheme based on averaging models periodically at the server, using only a fraction of clients in training and quantized message passing at edge nodes. It achieved near-optimal theoretical guarantees for strongly convex and non-convex loss functions and demonstrated the communication-computation trade-off. FedAGM [169] improved server-side aggregation using an accelerated model to guide local gradient updates. It achieved communication efficiency even with low client participation.

B. CHALLENGES ASSOCIATED WITH SL

Although SL offers improved privacy compared to CML and FL settings [8], [37], [116], [117], SL still has several privacy loopholes. Smaller client-side models can lead to private data leakage.

1) RESEARCH ON HANDLING PRIVACY ISSUES IN SL

The authors in [123] examined SL's privacy-preserving training in 1D CNN models. The authors adopted two privacy mitigation techniques: adding more hidden layers to the client side and applying differential privacy. SplitHE [170] presented an empirical security evaluation using membership inference applying homomorphic encryption. The model was built on top of the SplitNN architecture. Marvell (optiMized perturbAtion to pReVEnt Label Leakage) [171] is a realistic threat model that proposed a privacy loss metric to quantify label leakage in SL. By minimizing the amount of label leakage, it found the structure of the noise disturbance in a smart way. TPSL (Transcript Private Split Learning) [172] is a similar gradient perturbation-based SL approach that provides a provable differential privacy guarantee. Their experiments demonstrated robustness and effectiveness against label leakage attacks. SL in the context of the private collaborative inference against reconstruction attacks are analyzed in [173]. Their approach modified model training to reduce data leakage while maintaining accuracy. SplitGuard [174], [175] detected and mitigated hijacking attacks in SL. UnSplit [176] studied data oblivious model inversion, model stealing, and label inference attacks against SL. These methods showed that vanilla SL is not very secure and showed how important it is to take additional steps to make secure protocols.

C. CHALLENGES ASSOCIATED WITH SFL

Similar to SL, SFL offers privacy preservation advantages, as previously discussed. However, SFL can still suffer from several privacy loopholes. Like SL, private data can leak

through the smashed data with smaller client-side models. Also, research has been done to tackle the heterogeneity issues associated with SFL.

1) RESEARCH ON HANDLING PRIVACY ISSUES IN SFL

Research has enabled SFL to be performed in a privacy-preserving manner. To protect sensitive data and limit the amount of data sent during training, encryption, and other security measures are used. SplitFed authors [38] came up with several privacy preservation mechanisms on the client-side, the FS, and the main server. Among these were adding a PixelDP noise layer and using differential privacy to train the client-side model. These methods allowed for model training without disclosing sensitive information. The recent work, [177] performed an empirical analysis of SplitFed's robustness to strong model poisoning attacks.

2) RESEARCH ON HANDLING HETEROGENEITY ISSUES IN SFL

The SFL framework of [178] proposed an energy- and loss-aware selective updating method for heterogeneous systems, updating client-side models based on clients' energy and loss changes. Experiments were conducted using CIFAR [33] datasets. SplitAVG [125], as previously stated, also dealt with heterogeneous clients.

V. HEALTHCARE VIEWPOINT ON CHALLENGES, CURRENT TRENDS, AND FUTURE DIRECTIONS

In this section, we discuss the problems, the current research trends, and some possible directions for the future from the point of view of healthcare.

A. CHALLENGES SPECIFIC TO HEALTHCARE

Despite the fact that these decentralized learning approaches have enormous advantages and work well in healthcare applications, this field is still in its early stages due to the challenges that are frequently rising [179], [180]. We categorized these healthcare-specific challenges into three: data challenges, privacy and security challenges, and communication challenges. Data challenges involve issues with data quality, data heterogeneity, or data biases of the healthcare participants. Privacy and security challenges include data poisoning attacks, adversarial attacks, membership inference attacks, or free-riding attacks in the collaborative network. Communication challenges might occur due to failure or dropouts of healthcare participants, energy consumption issues, or computational and communication overhead prevalent with systems. In contrast to the technical challenges, some other specific challenges exist, as shown under Section V-C.

B. CURRENT TRENDS SPECIFIC TO HEALTHCARE

It is evident that the research communities are truly active in mitigating associated challenges, as we detailed in Section IV. Prior to their local model training during the collaboration, all of the healthcare participants should have mutually agreed upon these mechanisms.

Decentralized learning paradigms require agreements to specify the scope, objective, and technology used, which might be challenging to determine because they are still relatively new. Research has revealed that massive projects being undertaken now are paving the way for future norms of inventive, safe, and fair collaboration in healthcare applications. Several university and industrial research institutions jointly lead Federated Learning (FL) research in the healthcare area. Associated institutions have successfully implemented several projects to guarantee efficient standards for confidential, equitable, and creative healthcare application collaboration. These collaborative projects and consortiums aim to unite researchers and practitioners interested in FL. Students, professors, and business leaders from around the world get together for these projects or consortiums to learn more about the topic, find technical problems, and discuss possible solutions. Table 4 highlights popular consortiums and collaborative healthcare projects associated with FL.

Currently, some open-source software has also been developed to facilitate proofs-of-concept and experiments. These software hold open-source federated datasets, including medical datasets, strict evaluation frameworks, and reference implementations that try to show the problems and complexities of real-world federated environments. Some of the most accessible medical datasets relate to COVID, Alzheimer’s disease, and predictive maintenance. Federated AI Technology Enabler (FATE),⁴ Substra,⁵ OpenFL,⁶ Tensorflow Federated,⁷ IBM Federated Learning,⁸ NVIDIA Clara,⁹ PySyft¹⁰+PyGrid platform,¹¹ and enterprise-grade FL platforms such as Apheris,¹² are some popular open-source software.

C. FUTURE DIRECTIONS SPECIFIC TO HEALTHCARE

Here are some possible directions for future research on how FL, SL, and SFL can be used in healthcare.

1) DEVELOPING PERSONALIZED MODELS

FL and SFL approaches produce a common global model for all participants. Healthcare decisions are generally scientifically designed in a personalized health management scheme. Care might vary institution-wise, physician-wise, and patient-wise. SOTA has some general research about FL personalization, as shown in Fig. 8 under *personalized FL*. However, how to combine the medical domain knowledge and produce personalized versions of the global models in collaborative networks is yet to be investigated.

⁴<https://fate.fedai.org/>
⁵<https://www.substra.ai/>
⁶<https://github.com/intel/openfl/>
⁷<https://www.tensorflow.org/federated/>
⁸<https://ibmfl.mybluemix.net/>
⁹<https://developer.nvidia.com/clara/>
¹⁰<https://blog.openmined.org/tag/pysyft/>
¹¹<https://github.com/OpenMined/pygrid-admin/>
¹²<https://www.apheris.com/>

TABLE 4. Popular consortia/ collaborative healthcare projects associated with FL.

Consortia/ Collaborative project	Description
Trustworthy Federated Data Analytics (TFDA) project	Invented the “Federated radiation therapy study” to assure the reliability of the study and transparency of the analysis without infringing privacy restrictions. Available at: https://tfda.hmsp.center/
German Cancer Consortium’s Joint Imaging Platform	Facilitated collaborative, decentralized research amongst German institutions using CT scans of cancerous organs. Available at: https://jip.dktk.dkfz.de/jiphompage/
NVIDIA Clara FL Mammogram Assessment	Developed AI models for mammogram assessment in FL networks. Available at: https://blogs.nvidia.com
HealthChain project	Developed FL tools that improved patient privacy in collaboration with healthcare facilities that treat melanoma and breast cancer in France. Available at: https://www.healthchain.ca/
Federated Tumour Segmentation (FeTS) initiative	Developed an open-source FL framework with a graphical user interface to use by a global alliance of 30 devoted healthcare organizations. Aimed to enhance tumor boundary identification, including bone lesions from patients with multiple myeloma, breast tumors, liver tumors, and brain gliomas. Addressed issues of ownership, privacy, and the law within multi-institutional collaborations. Available at: https://www.med.upenn.edu/cbica/fets/
Melloddy project	Implemented multi-task FL for drug discovery with datasets of 10 pharmaceutical institutions. Available at: https://www.melloddy.eu/
Doc.ai consortium	Used FL in a zero trust environment to create strong, actionable insights while maintaining data confidentiality and privacy. Available at: https://doc.ai/consortium/
Federated Learning for Medicine (FL4M)	Developed a hospital-based FL for medicine consortium that collaborated with seven healthcare institutions around four continents for brain aneurysm detection. Available at: https://www.rhinohealth.com/

2) REDUCING BIAS

Although FL and SFL limit the effects of a biased dataset by utilizing multiple datasets, systems with biased data propagation may exist due to poor system design. In the collaborative network, individual datasets should be appropriately weighted to reduce the risk of biased or insecure data. Even though they’re weighted properly initially, bias might still emerge later in training. The reason is that client features such as data distribution may still vary over time (e.g., with the addition of new patients or the demise of existing patients). Thus, developing more robust models resistant to handling

TABLE 5. Publicly available medical imaging datasets used in the recent decentralized learning research.

Dataset	Focused healthcare area	Description
BraTS 2017 and 2018 datasets [59]	Brain cancer	Contains brain MRI scans of 359 subjects collected from 13 institutions. Data belong to high-grade glioma, glioblastoma, and low-grade glioma. Data are heterogeneous in appearance, shape, and histology. Available at: https://www.med.upenn.edu/sbia/brats2017/data.html/ , https://www.kaggle.com/datasets/sanglequang/brats2018/
Open Access Series of Imaging Studies (OASIS-3) [61]	Alzheimer's disease	Consist of brain MRI scans and Positron Emission Tomography (PET) images collected from 1,098 subjects. Available at: https://www.oasis-brains.org/
Alzheimer's Disease Neuroimaging Initiative (ADNI-3) dataset [62]	Alzheimer's disease	Contains brain MRI scans and PET images collected from 1,068 subjects. Available at: https://adni.loni.usc.edu/data-samples/access-data/
MSLUB dataset [63]	Alzheimer's disease	Contains T1-weighted (T1), T2-weighted (T2), and Fluid Attenuated Inversion Recovery (FLAIR) MRI scans of 30 subjects with multiple sclerosis. Available at: http://lit.fe.uni-lj.si/tools/
MSISBI dataset [64]	Alzheimer's disease	Contains brain MRI scans divided into the training set and test set. The training set and test set contain data from 5 and 14 subjects. Available at: https://smart-stats-tools.org/lesion-challenge-2015/
Autism Brain Imaging Data Exchange (ABIDE) dataset [66]	Autism	Contains 1,112 brain fMRI scans from 539 subjects with autism spectrum disorders. Available at: http://fcon_1000.projects.nitrc.org/indi/abide/
Non-Small Cell Lung Cancer (NSCLC)-Radiomics dataset [69]	Lung cancer	Contains chest CT scans of 422 NSCLC subjects, paired with Gross Tumor Volume (GTV) segmentations. Available at: https://wiki.cancerimagingarchive.net/display/Public/NSCLC-Radiomics/
Radiology Ai One-Stop solution (RAIOSS) dataset [72]	COVID-19	Contains 10 cases of chest CT scans of COVID-19 patients. Available at: https://coronacases.org/
Montgomery County chest X-ray (MC) dataset [74]	Tuberculosis	Contains 138 frontal chest X-ray images: 80 normal images and 58 images affected to tuberculosis. Available at: https://openi.nlm.nih.gov/faq#faq-tb-coll
Shenzhen chest X-ray dataset [74]	Tuberculosis	Contains 662 frontal chest X-ray images: 326 normal images and 336 images affected to tuberculosis. Available at: https://openi.nlm.nih.gov/faq#faq-tb-coll
Chest X-ray (CXR) dataset [76]	Thoracic diseases	Contains 108,948 frontal view chest X-ray images of 32,717 subjects with eight common thoracic diseases (atelectasis, cardiomegaly, effusion, infiltration, mass, nodule, pneumonia, pneumothorax). Available at: https://www.kaggle.com/datasets/nih-chest-xrays/data
International Skin Imaging Collaboration (ISIC)19 dataset [78]	Skin cancer	Consists of 25,000 dermatoscopic images with eight classes: Melanoma (MEL), melanocytic Nevus (NV), Basal Cell Carcinoma (BCC), Actinic Keratosis (AK), Benign Keratosis (BKL), Dermatofibroma (DF), Vascular lesion (VASC), and Squamous Cell Carcinoma (SCC). Available at: https://challenge.isic-archive.com/data/

TABLE 5. (Continued.) Publicly available medical imaging datasets used in the recent decentralized learning research.

HAM10K Skin lesion dataset [79]	Skin cancer	Contains 10,000 dermoscopic images with seven classes: Actinic Keratoses and Intraepithelial Carcinoma (AKIEC), MEL, NV, BCC, BKL, DF, and VASC. Available at: https://doi.org/10.7910/DVN/DBW86T
Derm7pt dataset [80]	Skin cancer	Contains 1,011 dermoscopic images classified into six classes. Available at: https://derm.cs.sfu.ca/Download.html
PAD-UFES dataset [81]	Skin cancer	Contains 2,298 dermoscopic images collected from 1,373 subjects classified into six classes: Seborrheic Keratosis (SEK), Bowen's disease (BOD), MEL, NV, BCC, SCC, and AK. Available at: https://data.mendeley.com/datasets/zr7vgbcyr2/1/
Breast Ultrasound (BUS) dataset [85]	Breast cancer	Consists of 780 breast Ultrasound (US) images: 487 benign images, 210 malignant images, and 133 normal images. Available at: https://scholar.cu.edu.eg/?q=afahmy/pages/dataset
Breast Ultrasound Image Segmentation (BUSIS) dataset [86]	Breast cancer	Consists of 562 breast US images. Available at: http://cvprip.cs.usu.edu/busbench/
UDIAT dataset [87]	Breast cancer	Consists of 163 breast US images: 53 cancerous images, and 110 benign images. Available at: http://www2.docm.mmu.ac.uk/STAFF/m.yap/dataset.php/
EyePACS dataset [92]	Diabetic retinopathy	Consists of 1,122 OCT and OCTA images collected from 53 clinicians. Available at: https://www.kaggle.com/datasets/mariaherrero/eyepacspreprocess/
Methods to Evaluate Segmentation and Indexing Techniques in the field of Retinal Ophthalmology (MESSIDOR) database [93]	Diabetic retinopathy	Contains 1,200 fundus images captured from Charged-Coupled Device (CCD) cameras. The retinopathy grade, scaled from 0–3, and the risk of macular edema, scaled from 0–2, are provided for each image. Available at: https://www.adcis.net/en/third-party/messidor/
Indian Diabetic Retinopathy image Dataset (IDRiD) [94]	Diabetic retinopathy	Contains 516 fundus images classified into those with and without diabetic retinopathy (DR). Available at: https://iee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid/
Asia Pacific Tele-Ophthalmology Society (APTOS) 2019 dataset [95]	Diabetic retinopathy	Contains 703 fundus images classified into five scales of DR: 0 - No DR, 1 - mild, 2 - moderate, 3 - severe, and 4 - proliferative DR. Available at: https://www.kaggle.com/c/aptos2019-blindness-detection/data/
University of Auckland (UoA) diabetic retinopathy database [96]	Diabetic retinopathy	Contains 200 fundus images captured from VISUCAM 500 fundus camera: 114 non-proliferative DR images, 30 proliferative DR images, and 56 healthy images. Available at: https://auckland.figshare.com/articles/journal_contribution/UoA-DR_Database_Info/5985208/5
Diabetic Retinopathy binary classification dataset [97]	Diabetic retinopathy	Consists of 88,702 left and right eye color digital fundus images, collected from 44,351 subjects. They're classified into five scales of DR mentioned in [95]. Available at: https://www.kaggle.com/competitions/diabetic-retinopathy-detection/data

TABLE 5. (Continued.) Publicly available medical imaging datasets used in the recent decentralized learning research.

Medical Segmentation Decathlon (MSD1) dataset [100]	Brain, liver, lung, colon, pancreas, hepatic vessels cancers	Contains CT scans of liver, lung, colon, pancreas, and hepatic vessels and brain MRI scans. Available at: http://medicaldecathlon.com/
SPIE-AAPM PROSTATEx dataset [101]	Prostate cancer	Contains 343 prostate MRI scans captured from two types of Siemens 3T MR scanners, the MAGNETOM Trio and Skyra. Available at: https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=23691656
PROMISE12 dataset [102]	Prostate cancer	Consists of 100 prostate MRI scans from four different institutions. Available at: https://promise12.grand-challenge.org/Download/
National Cancer Institute–International Symposium on Biomedical Imaging (NCI-ISBI) dataset [103]	Prostate cancer	Consists of prostate MRI scans collected from 80 subjects. Available at: https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=21267207
Synapse dataset [106]	Abdoman and cervix cancers	Includes CT scans of abdomen and cervix. Available at: https://www.synapse.org/#!/Synapse:syn3193805/wiki/217753/
Chexpert dataset [119]	Chest pathology	Consists of 224,316 chest radiograph images from 65,240 subjects. Available at: https://stanfordaimi.azurewebsites.net/datasets/8cbd9ed4-2eb9-4565-affe-111cf4f7ebe2/
MUsculoskeletal Radiographs (MURA) dataset [127]	Abnormality detection	Includes 40,561 X-ray images of seven body parts: wrist, humerus, forearm, hand, elbow, finger, and forearm collected from 14,863 cases. Each case is labeled as either normal or abnormal. Available at: https://stanfordmlgroup.github.io/competitions/mura/
COVID-CT dataset [128]	COVID-19	Contains chest CT scans collected from 216 subjects: 349 COVID-19 images, and 463 non-COVID-19 images. Available at: https://github.com/UCSD-AI4H/COVID-CT/
COVIDx dataset [131]	COVID-19	Contains 13,975 chest CT scans collected from 13,870 subjects: 7,966 normal images, 5,471 pneumonia images, and 517 COVID-19 images. Available at: https://github.com/lindawangg/COVID-Net/
Kvasir dataset [132]	Gastrointestinal disease	Contains 8,000 endoscopic images of the Gastrointestinal tract, classified into anatomical landmarks (Z-line, pylorus, cecum), pathological findings (esophagitis, polyps, ulcerative colitis), and polyp removal (dyed and lifted polyp, dyed resection margins). Available at: https://datasets.simula.no/kvasir/

bias with proper parameter aggregation schemes is an important research direction.

3) INCORPORATING HYBRID NON-IID PROPERTIES

Healthcare institutions have different sample sizes, label distributions, resolutions, data measurement frequencies, types of laboratory tests, laboratory tools, data acquisition mechanisms, and various demographic features of subjects such as age, gender, etc. Therefore, medical datasets often involve several non-IID features [181]. Most FL studies focus exclusively on one of the non-IID features, such as noisy or

corrupted labels. There hasn't been a thorough investigation of various non-IID features in medical datasets. Therefore, in the future, another focus for the decentralized learning paradigm could be addressing challenges related to multiple non-IID features in medical data.

4) HYPERPARAMETER TUNING

Hyperparameter tuning in machine learning is crucial but time-consuming. Optimizing hyperparameters is significantly challenging in collaborative learning [182], while it is extremely difficult in healthcare collaborations.

Future research will urgently require automated techniques or frameworks to choose the best hyperparameters in decentralized learning models in healthcare.

The following are identified as non-technical directions that need specific attention.

5) DEVELOPING SUFFICIENT INCENTIVE MECHANISMS

Healthcare organizations may have limited trust in the new technological frameworks. They might even be required to undergo considerable communication and computation overheads. Naturally, they might be unwilling to participate in collaborative learning tasks without well-designed incentives. Finding effective ways to encourage organizations with standard imaging or related data to participate in the learning process is a fundamental problem for the future.

6) INCORPORATING DOMAIN EXPERTISE KNOWLEDGE

The healthcare industry relies on professionalism and accuracy. Physicians or subject-matter experts delicately perform tasks like devising treatment plans, prescribing medications, or forecasting specific health conditions. Even if there is a large medical dataset, healthcare experts will not trust or accept the predictions of a collaborative model trained in a collaborative network in the first round. Such predictions require expert knowledge, supervision, and intervention. Experts in the medical field could oversee the entire process of collaborative learning, from the first step of collecting data to the last step of making a prediction based on the global model. Such oversight by experts will lead to more accurate and confident results.

VI. CONCLUSION

In this survey, we reviewed the emerging decentralized learning approaches of FL, SL, and SFL and their applications in healthcare, with a particular focus on medical imaging. In general, several studies have tried to address some of the challenges in FL, SL, and SFL settings. We detailed major challenges with SOTA solutions. From the healthcare perspective, these challenges still need to be fully addressed. Since FL, SL, and SFL will be active research areas for the next ten years, we detailed current trends and possible directions for future research in their healthcare applications. There is a lot of room and need for systems and algorithms that are more realistic, perform better, and ensure security and privacy.

APPENDIX

PUBLICLY AVAILABLE MEDICAL IMAGING DATASETS USED IN RECENT DECENTRALIZED LEARNING RESEARCH

Having open access to data is an essential component of research. In healthcare, publicly available data would benefit researchers and academia to a greater extent. Open data helps, without a doubt, in coming up with effective and efficient solutions to life-threatening diseases or long-lasting problems that plague humanity.

As part of the survey, we contributed to the design of a collection of publicly available medical imaging datasets that have been used in the recent decentralized learning research. It is identified that these datasets cover a large range, including data related to the brain, chest, skin, breast, eye, prostate, skin, abdomen, etc. Table 5 lists our findings.

REFERENCES

- [1] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2018.
- [2] M. Meehan. (Nov. 2019). *Data Privacy Will Be the Most Important Issue in the Next Decade*. [Online]. Available: <https://www.forbes.com/sites/marymeehan/2019/11/26/data-privacy-will-be-the-most-important-issue-in-the-next-decade/?sh=1e3459111882>
- [3] C. Fontaine and F. Galand, "A survey of homomorphic encryption for nonspecialists," *EURASIP J. Inf. Secur.*, vol. 2007, pp. 1–10, Dec. 2007.
- [4] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, "Deep learning with differential privacy," in *Proc. CCS*, 2016, pp. 308–318.
- [5] N. Papernot, S. Song, I. Mironov, A. Raghunathan, K. Talwar, and U. Erlingsson, "Scalable private learning with PATE," in *Proc. ICLR*, Vancouver, BC, Canada, 2018, p. 34.
- [6] P. Bogetoft, "Secure multiparty computation Goes live," in *Proc. Int. Conf. Financial Cryptogr. Data Secur.* Berlin, Germany: Springer, 2009, pp. 325–343.
- [7] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. AISTATS*, 2017, pp. 1273–1282.
- [8] O. Gupta and R. Raskar, "Distributed learning of deep neural network over multiple agents," *J. Netw. Comput. Appl.*, vol. 116, pp. 1–8, Aug. 2018.
- [9] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, p. 19, Feb. 2019.
- [10] V. Kulkarni, M. Kulkarni, and A. Pant, "Survey of personalization techniques for federated learning," in *Proc. 4th World Conf. Smart Trends Syst., Secur. Sustainability (WorldS)*, Jul. 2020, pp. 794–797.
- [11] Y. Jin, X. Wei, Y. Liu, and Q. Yang, "Towards utilizing unlabeled data in federated learning: A survey and prospective," 2020, *arXiv:2002.11545*.
- [12] W. Y. B. Lim, N. C. Luong, D. T. Hoang, Y. Jiao, Y. Liang, Q. Yang, D. Niyato, and C. Miao, "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031–2063, 3rd Quart., 2020.
- [13] O. Shahid, S. Pouriyeh, R. M. Parizi, Q. Z. Sheng, G. Srivastava, and L. Zhao, "Communication efficiency in federated learning: Achievements and challenges," 2021, *arXiv:2107.10996*.
- [14] L. Lyu, H. Yu, and Q. Yang, "Threats to federated learning: A survey," 2020, *arXiv:2003.02133*.
- [15] P. Kairouz, "Advances and open problems in federated learning," *Found. Trends Mach. Learn.*, vol. 14, nos. 1–2, pp. 1–210, Jun. 2021.
- [16] N. Mouhni, A. Elkalay, M. Chakraoui, A. Abdali, A. Ammoumou, and I. Amalou, "Federated learning for medical imaging: An updated state of the art," *Ingénierie des Systèmes d Inf.*, vol. 27, no. 1, pp. 143–150, Feb. 2022.
- [17] Prayitno, C.-R. Shyu, K. T. Putra, H.-C. Chen, Y.-Y. Tsai, K. S. M. T. Hossain, W. Jiang, and Z.-Y. Shae, "A systematic review of federated learning in the healthcare area: From the perspective of data properties and applications," *Appl. Sci.*, vol. 11, no. 23, Nov. 2021, Art. no. 11191.
- [18] A. Chowdhury, H. Kassem, N. Padoy, R. Umeton, and A. Karargyris, "A review of medical federated learning: Applications in oncology and cancer research," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2022, pp. 3–24.
- [19] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, "Federated learning for healthcare informatics," *J. Healthcare Informat. Res.*, vol. 5, no. 1, pp. 1–19, Mar. 2021.
- [20] N. Rieke, J. Hancox, W. Li, F. Milletari, H. Roth, S. Albarqouni, S. Bakas, M. N. Galtier, B. Landman, K. Maier-Hein, S. Ourselin, M. Sheller, R. M. Summers, A. Trask, D. Xu, M. Baust, and M. J. Cardoso, "The future of digital health with federated learning," *NPJ Digit. Med.*, vol. 3, no. 1, p. 119, Dec. 2020.

- [21] A. Rauniar, D. H. Hagos, D. Jha, J. E. Håkegård, U. Bagci, D. B. Rawat, and V. Vlassov, "Federated learning for medical applications: A taxonomy, current trends, challenges, and future research directions," 2022, *arXiv:2208.03392*.
- [22] Z.-H. Zhou, *Ensemble Methods: Foundations and Algorithms*. London, U.K.: Chapman & Hall, 2012.
- [23] Z. Xue, P. Zhou, Z. Xu, X. Wang, Y. Xie, X. Ding, and S. Wen, "A resource-constrained and privacy-preserving edge-computing-enabled clinical decision system: A federated reinforcement learning approach," *IEEE Internet Things J.*, vol. 8, no. 11, pp. 9122–9138, Jun. 2021.
- [24] 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," *Med. Image Anal.*, vol. 65, Oct. 2020, Art. no. 101765.
- [25] 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 Things J.*, vol. 8, no. 8, pp. 6949–6962, Apr. 2021.
- [26] L. Zhang, B. Shen, A. Barnawi, S. Xi, N. Kumar, and Y. Wu, "FedDPGAN: Federated differentially private generative adversarial networks framework for the detection of COVID-19 pneumonia," *Inf. Syst. Frontiers*, vol. 23, no. 6, pp. 1403–1415, Dec. 2021.
- [27] Y. Chen, X. Sun, and Y. Jin, "Communication-efficient federated deep learning with Layerwise asynchronous model update and temporally weighted aggregation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 10, pp. 4229–4238, Oct. 2020.
- [28] 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," *Int. J. Med. Informat.*, vol. 112, pp. 59–67, Apr. 2018.
- [29] I. V. Bajic, W. Lin, and Y. Tian, "Collaborative intelligence: Challenges and opportunities," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 8493–8497.
- [30] I. Ceballos, V. Sharma, E. Mugica, A. Singh, A. Roman, P. Vepakomma, and R. Raskar, "SplitNN-driven vertical partitioning," 2020, *arXiv:2008.04137*.
- [31] L. Deng, "The MNIST database of handwritten digit images for machine learning research [best of the web]," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 141–142, Nov. 2012.
- [32] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
- [33] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," Univ. Toronto, Toronto, ON, Canada, Tech. Rep., 2009.
- [34] Y. Gao, M. Kim, C. Thapa, A. Abuadba, Z. Zhang, S. Camtepe, H. Kim, and S. Nepal, "Evaluation and optimization of distributed machine learning techniques for Internet of Things," *IEEE Trans. Comput.*, vol. 71, no. 10, pp. 2538–2552, Oct. 2022.
- [35] E. Samikwa, A. D. Maio, and T. Braun, "ARES: Adaptive resource-aware split learning for Internet of Things," *Comput. Netw.*, vol. 218, Dec. 2022, Art. no. 109380.
- [36] J. Kim, Y. Park, G. Kim, and S. J. Hwang, "SplitNet: Learning to semantically split deep networks for parameter reduction and model parallelization," in *Proc. 34th Int. Conf. Mach. Learn. (ICML)*, vol. 70, Aug. 2017, pp. 1866–1874.
- [37] A. Singh, P. Vepakomma, O. Gupta, and R. Raskar, "Detailed comparison of communication efficiency of split learning and federated learning," 2019, *arXiv:1909.09145*.
- [38] C. Thapa, M. A. P. Chamikara, S. Camtepe, and L. Sun, "SplitFed: When federated learning meets split learning," 2020, *arXiv:2004.12088*.
- [39] C. Thapa, P. C. M. Arachchige, S. Camtepe, and L. Sun, "SplitFed: When federated learning meets split learning," in *Proc. AAAI*, vol. 36, 2022, pp. 8485–8493.
- [40] M. Gawali, C. S. Arvind, S. Suryavanshi, H. Madaan, A. Gaikwad, K. N. B. Prakash, V. Kulkarni, and A. Pant, "Comparison of privacy-preserving distributed deep learning methods in healthcare," 2020, *arXiv:2012.12591*.
- [41] P. Joshi, C. Thapa, S. Camtepe, M. Hasanuzzaman, T. Scully, and H. Afli, "Performance and information leakage in split learning and multi-head split learning in healthcare data and beyond," *Methods Protocols*, vol. 5, no. 4, p. 60, Jul. 2022.
- [42] P. Joshi, C. Thapa, S. Camtepe, M. Hasanuzzaman, T. Scully, and H. Afli, "SplitFed learning without client-side synchronization: Analyzing client-side split network portion size to overall performance," 2021, *arXiv:2109.09246*.
- [43] S. Pal, M. Uniyal, J. Park, P. Vepakomma, R. Raskar, M. Bennis, M. Jeon, and J. Choi, "Server-side local gradient averaging and learning rate acceleration for scalable split learning," 2021, *arXiv:2112.05929*.
- [44] V. Turina, Z. Zhang, F. Esposito, and I. Matta, "Combining split and federated architectures for efficiency and privacy in deep learning," in *Proc. 16th Int. Conf. Emerg. Netw. Experiments Technol.*, Nov. 2020, pp. 562–563.
- [45] J. Jeon and J. Kim, "Privacy-sensitive parallel split learning," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Jan. 2020, pp. 7–9.
- [46] H. Gu, L. Fan, B. Li, Y. Kang, Y. Yao, Q. Yang, and A. Group, "A novel approach to simultaneously improve privacy, efficiency and reliability of federated DNN learning," in *Proc. Int. Workshop Federated Transf. Learn. Data Sparsity Confidentiality Conjoint With (IJCAI)*, 2021, pp. 7–9.
- [47] D. Romanini, P. Papadopoulos, A. J. Hall, T. Titcombe, A. Ismail, T. Ceberer, R. Sandmann, R. Roehm, and M. A. Hoeh, "Pyvertical: A vertical federated learning framework for multi-headed SplitNN," in *Proc. ICLR Workshop DPML*, 2021, p. 9.
- [48] A. Chopra, S. K. Sahu, A. Singh, A. Java, P. Vepakomma, V. Sharma, and R. Raskar, "AdaSplit: Adaptive trade-offs for resource-constrained distributed deep learning," 2021, *arXiv:2112.01637*.
- [49] D. J. Han, H. I. Bhatti, J. Lee, and J. Moon, "Accelerating federated learning with split learning on locally generated losses," in *Proc. Int. Workshop Federated Learn. User Privacy Data Confidentiality Conjoint With (ICML)*, 2021, p. 12, 2021.
- [50] R. Ullah, D. Wu, P. Harvey, P. Kilpatrick, I. Spence, and B. Varghese, "FedFly: Towards migration in edge-based distributed federated learning," *IEEE Commun. Mag.*, vol. 60, no. 11, pp. 42–48, Nov. 2022.
- [51] J. Wang, H. Qi, A. S. Rawat, S. Reddi, S. Waghmare, F. X. Yu, and G. Joshi, "FedLite: A scalable approach for federated learning on resource-constrained clients," 2022, *arXiv:2201.11865*.
- [52] S. Oh, J. Park, P. Vepakomma, S. Baek, R. Raskar, M. Bennis, and S.-L. Kim, "LocFedMix-SL: Localize, federate, and mix for improved scalability, convergence, and latency in split learning," in *Proc. ACM Web Conf.*, Apr. 2022, pp. 3347–3357.
- [53] W. Wu, M. Li, K. Qu, C. Zhou, S. Xuemin, W. Zhuang, X. Li, and W. Shi, "Split learning over wireless networks: Parallel design and resource management," 2022, *arXiv:2204.08119*.
- [54] X. Liu, Y. Deng, and T. Mahmoodi, "Energy efficient user scheduling for hybrid split and federated learning in wireless UAV networks," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 1–6.
- [55] J. Xu, Z. Xu, P. Walker, and F. Wang, "Federated patient hashing," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 4, Apr. 2020, pp. 6486–6493.
- [56] D. Liu, D. Dligach, and T. Miller, "Two-stage federated phenotyping and patient representation learning," in *Proc. 18th BioNLP Workshop Shared Task*, 2019, p. 283.
- [57] Y. Kim, J. Sun, H. Yu, and X. Jiang, "Federated tensor factorization for computational phenotyping," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2017, pp. 887–895.
- [58] M. J. Sheader, 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," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 11383, A. Crimi, S. Bakas, H. Kuijff, F. Keyvan, M. Reyes, and T. van Walsum, Eds. Cham, Switzerland: Springer, 2019, pp. 92–104.
- [59] B. H. Menze, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.
- [60] C. I. Bercea, B. Wiestler, D. Rueckert, and S. Albarqouni, "FedDis: Disentangled federated learning for unsupervised brain pathology segmentation," 2021, *arXiv:2103.03705*.
- [61] P. J. LaMontagne, T. L. S. Benzinger, J. C. Morris, S. Keefe, R. Hornbeck, C. Xiong, E. Grant, J. Hassenstab, K. Moulder, A. G. Vlassenko, M. E. Raichle, C. Cruchaga, and D. Marcus, "OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease," medRxiv, Cold Spring Harbor Lab. Press, Dec. 2019, doi: 10.1101/2019.12.13.19014902.

- [62] M. W. Weiner, D. P. Veitch, P. S. Aisen, L. A. Beckett, N. J. Cairns, R. C. Green, D. Harvey, C. R. Jack, W. Jagust, J. C. Morris, R. C. Petersen, J. Salazar, A. J. Saykin, L. M. Shaw, A. W. Toga, and J. Q. Trojanowski, "The Alzheimer's disease neuroimaging initiative 3: Continued innovation for clinical trial improvement," *Alzheimer's Dementia*, vol. 13, no. 5, pp. 561–571, May 2017.
- [63] Ž. Lesjak, A. Galimzianova, A. Koren, M. Lukin, F. Pernuš, B. Likar, and Ž. Špiclin, "A novel public MR image dataset of multiple sclerosis patients with lesion segmentations based on multi-rater consensus," *Neuroinformatics*, vol. 16, no. 1, pp. 51–63, Jan. 2018.
- [64] A. Carass, S. Roy, A. Jog, J. L. Cuzzocreo, E. Magrath, A. Gherman, J. Button, J. Nguyen, P.-L. Bazin, P. A. Calabresi, C. M. Crainiceanu, L. M. Ellingsen, D. S. Reich, J. L. Prince, and D. L. Pham, "Longitudinal multiple sclerosis lesion segmentation data resource," *Data Brief*, vol. 12, pp. 346–350, Jun. 2017.
- [65] S. Silva, B. A. Gutman, E. Romero, P. M. Thompson, A. Altmann, and M. Lorenzi, "Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 270–274.
- [66] A. D. Martino, "The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism," *Mol. Psychiatry*, vol. 19, no. 6, pp. 659–667, Jun. 2014.
- [67] S. Silva, A. Altmann, B. Gutman, and M. Lorenzi, "Fed-BioMed: A general open-source frontend framework for federated learning in healthcare," in *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*, vol. 12444, 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, Switzerland: Springer, 2020, pp. 201–210.
- [68] W. Li, "Privacy-preserving federated brain tumour segmentation," in *Proc. MLMI*. Cham, Switzerland: Springer, 2019, pp. 133–141.
- [69] H. J. W. L. Aerts, L. Wee, E. R. Velazquez, R. T. H. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers, M. M. Rietbergen, C. R. Leemans, A. Dekker, J. Quackenbush, R. J. Gillies, and P. Lambin. (2019). *Data From NSCLC-Radiomics Version Number: 4 Type: Dataset*. [Online]. Available: <https://wiki.cancerimagingarchive.net/display/Public/NSCLC-Radiomics>
- [70] F. Zerka, V. Urovi, A. Vaidyanathan, S. Barakat, R. T. H. Leijenaar, S. Walsh, H. Gabrani-Juma, B. Miraglio, H. C. Woodruff, M. Dumontier, and P. Lambin, "Blockchain for privacy preserving and trustworthy distributed machine learning in multicentric medical imaging (C-DistriM)," *IEEE Access*, vol. 8, pp. 183939–183951, 2020.
- [71] Q. Dou, "Federated deep learning for detecting COVID-19 lung abnormalities in CT: A privacy-preserving multinational validation study," *NPJ Digit. Med.*, vol. 4, no. 1, p. 60, Mar. 2021.
- [72] *Coronacases.org*. Accessed: Mar. 18, 2023. [Online]. Available: <https://coronacases.org/>
- [73] I. Feki, S. Ammar, Y. Kessentini, and K. Muhammad, "Federated learning for COVID-19 screening from chest X-ray images," *Appl. Soft Comput.*, vol. 106, Jul. 2021, Art. no. 107330.
- [74] S. Jaeger, S. Candemir, S. Antani, Y. X. Wang, P. X. Lu, and G. Thoma, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," *Quant. Imag. Med. Surg.*, vol. 4, p. 475, Dec. 2014.
- [75] N. Balachandar, K. Chang, J. Kalpathy-Cramer, and D. L. Rubin, "Accounting for data variability in multi-institutional distributed deep learning for medical imaging," *J. Amer. Med. Inform. Assoc.*, vol. 27, no. 5, pp. 700–708, May 2020.
- [76] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3462–3471.
- [77] T. Bdair, N. Navab, and S. Albarqouni, "FedPerl: Semi-supervised peer learning for skin lesion classification," in *Proc. MICCAI*. Cham, Switzerland: Springer, 2021, pp. 336–346.
- [78] N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, D. Gutman, B. Helba, A. Kallou, K. Liopyris, M. Marchetti, H. Kittler, and A. Halpern, "Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC)," 2019, *arXiv:1902.03368*.
- [79] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci. Data*, vol. 5, no. 1, Aug. 2018, Art. no. 180161.
- [80] J. Kawahara, S. Daneshvar, G. Argenziano, and G. Hamarneh, "Seven-point checklist and skin lesion classification using multitask multimodal neural nets," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 2, pp. 538–546, Mar. 2019.
- [81] A. G. C. Pacheco, G. R. Lima, A. S. Salomão, B. Krohling, I. P. Biral, G. G. de Angelo, F. C. R. Alves Jr., J. G. M. Esgario, A. C. Simora, P. B. C. Castro, F. B. Rodrigues, P. H. L. Frasson, R. A. Krohling, H. Knidel, M. C. S. Santos, R. B. do Espírito Santo, T. L. S. G. Macedo, T. R. P. Canuto, and L. F. S. de Barros, "PAD-UFES-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones," *Data Brief*, vol. 32, Oct. 2020, Art. no. 106221.
- [82] B. L. Y. Agleby, J. Li, A. U. Haq, E. K. Bankas, S. Ahmad, I. O. Agye-mang, D. Kulevome, W. D. Ndiaye, B. Cobbinah, and S. Latipova, "Multimodal melanoma detection with federated learning," in *Proc. 18th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process. (ICCWAMTIP)*, Dec. 2021, pp. 238–244.
- [83] C. Xing Tian, H. Li, Y. Wang, and S. Wang, "Privacy-preserving constrained domain generalization for medical image classification," 2021, *arXiv:2105.08511*.
- [84] J. Wicaksana, Z. Yan, D. Zhang, X. Huang, H. Wu, X. Yang, and K.-T. Cheng, "FedMix: Mixed supervised federated learning for medical image segmentation," 2022, *arXiv:2205.01840*.
- [85] W. Al-Dhabyani, M. Goma, H. Khaled, and A. Fahmy, "Dataset of breast ultrasound images," *Data Brief*, vol. 28, Feb. 2020, Art. no. 104863.
- [86] Y. Zhang, M. Xian, H.-D. Cheng, B. Shareef, J. Ding, F. Xu, K. Huang, B. Zhang, C. Ning, and Y. Wang, "BUSIS: A benchmark for breast ultrasound image segmentation," *Healthcare*, vol. 10, no. 4, p. 729, Apr. 2022.
- [87] M. H. Yap, G. Pons, J. Martí, S. Ganau, M. Sentís, R. Zwiggelaar, A. K. Davison, and R. Martí, "Automated breast ultrasound lesions detection using convolutional neural networks," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 4, pp. 1218–1226, Jul. 2018.
- [88] H. R. Roth, "Federated learning for breast density classification: A real-world implementation," in *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*. Cham, Switzerland: Springer, 2020, pp. 181–191.
- [89] A. Jiménez-Sánchez, M. Tardy, M. A. González Ballester, D. Mateus, and G. Piella, "Memory-aware curriculum federated learning for breast cancer classification," *Comput. Methods Programs Biomed.*, vol. 229, Feb. 2023, Art. no. 107318.
- [90] J. Lo, T. T. Yu, D. Ma, P. Zang, J. P. Owen, Q. Zhang, R. K. Wang, M. F. Beg, A. Y. Lee, Y. Jia, and M. V. Sarunic, "Federated learning for microvasculature segmentation and diabetic retinopathy classification of OCT data," *Ophthalmol. Sci.*, vol. 1, no. 4, Dec. 2021, Art. no. 100069.
- [91] M. Nasajpour, M. Karakaya, S. Pouriyeh, and R. M. Parizi, "Federated transfer learning for diabetic retinopathy detection using CNN architectures," in *Proc. SoutheastCon*, Mar. 2022, pp. 655–660.
- [92] J. Cuadros and I. Sim, "EyePACS: An open source clinical communication system for eye care," in *Proc. MEDINFO*. Amsterdam, The Netherlands: IOS Press, 2004, pp. 207–211.
- [93] E. Decencière, "Feedback on a publicly distributed image database: The Messidor database," *Image Anal. Stereol.*, vol. 33, no. 3, pp. 231–234, 2014.
- [94] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabudhe, and F. Meriaudeau, "Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research," *Data*, vol. 3, no. 3, p. 25, Jul. 2018.
- [95] *APTOS 2019 Blindness Detection*. Accessed: Mar. 18, 2023. [Online]. Available: <https://www.kaggle.com/c/aptos2019-blindness-detection>
- [96] R. J. Chalakkal, W. H. Abdulla, and S. Sinumol, "Comparative analysis of University of Auckland diabetic retinopathy database," in *Proc. 9th Int. Conf. Signal Process. Syst.*, Nov. 2017, pp. 235–239.
- [97] *Diabetic Retinopathy Detection*. Accessed: Mar. 18, 2023. [Online]. Available: <https://www.kaggle.com/c/diabetic-retinopathy-detection>
- [98] A. Rajagopal, E. Redekop, A. Kemiseti, R. Kulkarni, S. Raman, K. Magudia, C. W. Arnold, and P. E. Z. Larson, "Federated learning with research prototypes for multi-center MRI-based detection of prostate cancer with diverse histopathology," 2022, *arXiv:2206.05617*.
- [99] H. R. Roth, D. Yang, W. Li, A. Myronenko, W. Zhu, Z. Xu, X. Wang, and D. Xu, "Federated whole prostate segmentation in MRI with personalized neural architectures," in *Proc. Int. Conf. MICCAI*. Strasbourg, France: Springer, 2021, pp. 357–366.
- [100] M. Antonelli, "The medical segmentation decathlon," *Nature Commun.*, vol. 13, no. 1, p. 4128, Jul. 2022.

- [101] G. Litjens, O. Debats, J. Barentsz, N. Karssemeijer, and H. Huisman. (2017). *SPIE-AAPM PROSTATEx Challenge Data, Version Number: 2 Type: Dataset*. [Online]. Available: <https://wiki.cancerimagingarchive.net/x/iIFpAQ>
- [102] G. Litjens, "Evaluation of prostate segmentation algorithms for MRI: The PROMISE12 challenge," *Med. Image Anal.*, vol. 18, no. 2, pp. 359–373, Feb. 2014.
- [103] *NCI-ISBI 2013 Challenge: Automated Segmentation of Prostate Structures (ISBI-MR-Prostate-2013)*. Accessed: Mar. 18, 2023. [Online]. Available: <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=21267207>
- [104] X. Xu, H. H. Deng, T. Chen, T. Kuang, J. C. Barber, D. Kim, J. Gateno, J. J. Xia, and P. Yan, "Federated cross learning for medical image segmentation," 2022, *arXiv:2204.02450*.
- [105] C. Shen, "Multi-task federated learning for heterogeneous pancreas segmentation," in *Clinical Image-Based Procedures, Distributed and Collaborative Learning, Artificial Intelligence for Combating COVID-19 and Secure and Privacy-Preserving Machine Learning*. Cham, Switzerland: Springer, 2021, pp. 101–110.
- [106] I. J. Goodfellow, "Challenges in representation learning: A report on three machine learning contests," in *Proc. ICONIP*. Cham, Switzerland: Springer, 2013, pp. 117–124.
- [107] M. Knolle, G. Kaissis, F. Jungmann, S. Ziegelmayr, D. Sasse, M. Makowski, D. Rueckert, and R. Braren, "Efficient, high-performance semantic segmentation using multi-scale feature extraction," *PLoS ONE*, vol. 16, no. 8, Aug. 2021, Art. no. e0255397.
- [108] R. Duan, M. R. Boland, Z. Liu, Y. Liu, H. H. Chang, H. Xu, H. Chu, C. H. Schmid, C. B. Forrest, J. H. Holmes, M. J. Schuemie, J. A. Berlin, J. H. Moore, and Y. Chen, "Learning from electronic health records across multiple sites: A communication-efficient and privacy-preserving distributed algorithm," *J. Amer. Med. Inform. Assoc.*, vol. 27, no. 3, pp. 376–385, Mar. 2020.
- [109] Z. Li, K. Roberts, X. Jiang, and Q. Long, "Distributed learning from multiple EHR databases: Contextual embedding models for medical events," *J. Biomed. Informat.*, vol. 92, Apr. 2019, Art. no. 103138.
- [110] S. Lu, Y. Zhang, Y. Wang, and C. Mack, "Learn electronic health records by fully decentralized federated learning," 2019, *arXiv:1912.01792*.
- [111] Y. Chen, X. Qin, J. Wang, C. Yu, and W. Gao, "FedHealth: A federated transfer learning framework for wearable healthcare," *IEEE Intell. Syst.*, vol. 35, no. 4, pp. 83–93, Jul. 2020.
- [112] 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 uHealth*, vol. 9, no. 3, Mar. 2021, Art. no. e23728.
- [113] C. Ju, D. Gao, R. Mane, B. Tan, Y. Liu, and C. Guan, "Federated transfer learning for EEG signal classification," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 3040–3045.
- [114] 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," *J. Biomed. Informat.*, vol. 99, Nov. 2019, Art. no. 103291.
- [115] S. Boughorbel, F. Jarray, N. Venugopal, S. Moosa, H. Elhadi, and M. Makhoul, "Federated uncertainty-aware learning for distributed hospital EHR data," 2019, *arXiv:1910.12191*.
- [116] P. Vepakomma, O. Gupta, T. Swedish, and R. Raskar, "Split learning for health: Distributed deep learning without sharing raw patient data," 2018, *arXiv:1812.00564*.
- [117] M. G. Poirot, P. Vepakomma, K. Chang, J. Kalpathy-Cramer, R. Gupta, and R. Raskar, "Split learning for collaborative deep learning in healthcare," 2019, *arXiv:1912.12115*.
- [118] M. G. Poirot, "Split learning in health care," M.S. thesis, Sci. Technol. Fac. (TNW), Med. Imag. Interventions Specialization Tech. Med. Program, Univ. Twente, Enschede, The Netherlands, 2020. [Online]. Available: <https://essay.utwente.nl/80601>
- [119] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, C. Chute, R. Ball, J. Seekins, S. S. Halabi, R. Jones, D. B. Larson, C. P. Langlotz, B. N. Patel, and M. P. Lungren, "CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, Jul. 2019, pp. 590–597.
- [120] O. S. Ads, M. M. Alfares, and M. A.-M. Salem, "Multi-limb split learning for tumor classification on vertically distributed data," in *Proc. 10th Int. Conf. Intell. Comput. Inf. Syst. (ICICIS)*, Dec. 2021, pp. 88–92.
- [121] A. Ayad, M. Frei, and A. Schmeink, "Efficient and private ECG classification on the edge using a modified split learning mechanism," in *Proc. IEEE 10th Int. Conf. Healthcare Informat. (ICHI)*, Jun. 2022, pp. 1–6.
- [122] P. Wagner, N. Strodthoff, R.-D. Boussejot, D. Kreiseler, F. I. Lunze, W. Samek, and T. Schaeffter, "PTB-XL, a large publicly available electrocardiography dataset," *Sci. Data*, vol. 7, no. 1, p. 154, May 2020.
- [123] S. Abuadba, K. Kim, M. Kim, C. Thapa, S. A. Camtepe, Y. Gao, H. Kim, and S. Nepal, "Can we use split learning on 1D CNN models for privacy preserving training?" in *Proc. 15th ACM Asia Conf. Comput. Commun. Secur.*, Oct. 2020, pp. 305–318.
- [124] H. R. Roth, A. Hatamizadeh, Z. Xu, C. Zhao, W. Li, A. Myronenko, and D. Xu, "Split-U-Net: Preventing data leakage in split learning for collaborative multi-modal brain tumor segmentation," in *Proc. Int. Workshop Distrib. Collaborative, Federated Learn., Workshop Affordable Healthcare AI Resource Diverse Global Health*. Cham, Switzerland: Springer, 2022, pp. 47–57.
- [125] M. Zhang, L. Qu, P. Singh, J. Kalpathy-Cramer, and D. L. Rubin, "SplitAVG: A heterogeneity-aware federated deep learning method for medical imaging," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 9, pp. 4635–4644, Sep. 2022.
- [126] Y. J. Ha, M. Yoo, G. Lee, S. W. Choi, J. Kim, and S. Yoo, "Spatio-temporal split learning for privacy-preserving medical platforms: Case studies with COVID-19 CT, X-ray, and cholesterol data," *IEEE Access*, vol. 9, pp. 121046–121059, 2021.
- [127] P. Rajpurkar, J. Irvin, A. Bagul, D. Ding, T. Duan, H. Mehta, B. Yang, K. Zhu, D. Laird, R. L. Ball, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng, "MURA: Large dataset for abnormality detection in musculoskeletal radiographs," 2017, *arXiv:1712.06957*.
- [128] X. Yang, X. He, J. Zhao, Y. Zhang, S. Zhang, and P. Xie, "COVID-CT-dataset: A CT scan dataset about COVID-19," 2020, *arXiv:2003.13865*.
- [129] Y. J. Ha, G. Lee, M. Yoo, S. Jung, S. Yoo, and J. Kim, "Feasibility study of multi-site split learning for privacy-preserving medical systems under data imbalance constraints in COVID-19, X-ray, and cholesterol dataset," *Sci. Rep.*, vol. 12, no. 1, p. 1534, Jan. 2022.
- [130] Q. Yang, J. Zhang, W. Hao, G. P. Spell, and L. Carin, "FLOP: Federated learning on medical datasets using partial networks," in *Proc. 27th ACM SIGKDD Conf. Knowl. Discovery Data Mining*, Aug. 2021, pp. 3845–3853.
- [131] 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," *Sci. Rep.*, vol. 10, no. 1, Nov. 2020, Art. no. 19549.
- [132] K. Pogorelov, K. R. Randel, C. Griwodz, S. L. Eskeland, T. de Lange, D. Johansen, C. Spampinato, D. T. Dang-Nguyen, M. Lux, P. T. Schmidt, and M. Riegler, "KVASIR: A multi-class image dataset for computer aided gastrointestinal disease detection," in *Proc. 8th ACM Multimedia Syst. Conf.*, Jun. 2017, pp. 164–169.
- [133] C. G. Allaart, B. Keyser, H. Bal, and A. van Halteren, "Vertical split learning—An exploration of predictive performance in medical and other use cases," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2022, pp. 1–8.
- [134] K. Tam, L. Li, B. Han, C. Xu, and H. Fu, "Federated noisy client learning," 2021, *arXiv:2106.13239*.
- [135] T. Tuor, S. Wang, B. J. Ko, C. Liu, and K. K. Leung, "Overcoming noisy and irrelevant data in federated learning," in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 5020–5027.
- [136] X. Fang and M. Ye, "Robust federated learning with noisy and heterogeneous clients," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 10062–10071.
- [137] S. Yang, H. Park, J. Byun, and C. Kim, "Robust federated learning with noisy labels," *IEEE Intell. Syst.*, vol. 37, no. 2, pp. 35–43, Mar. 2022.
- [138] S. Li, E. Ngai, F. Ye, and T. Voigt, "Auto-weighted robust federated learning with corrupted data sources," *ACM Trans. Intell. Syst. Technol.*, vol. 13, no. 5, pp. 1–20, Oct. 2022.
- [139] M. Yang, H. Qian, X. Wang, Y. Zhou, and H. Zhu, "Client selection for federated learning with label noise," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 2193–2197, Feb. 2022.
- [140] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 50–60, May 2020.
- [141] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," in *Proc. Mach. Learn. Syst.*, vol. 2, 2020, pp. 429–450.
- [142] X. Li, W. Yang, K. Huang, S. Wang, and Z. Zhang, "On the convergence of FedAvg on non-IID," in *Proc. ICLR*, 2020, p. 26.

- [143] F. Sattler, S. Wiedemann, K.-R. Müller, and W. Samek, "Robust and communication-efficient federated learning from non-IID data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 9, pp. 3400–3413, Sep. 2020.
- [144] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, "Federated learning with non-IID data," 2018, *arXiv:1806.00582*.
- [145] S. Zawad, A. Ali, P.-Y. Chen, A. Anwar, Y. Zhou, N. Baracaldo, Y. Tian, and F. Yan, "Curse or redemption? How data heterogeneity affects the robustness of federated learning," in *Proc. AAAI*, vol. 35, no. 12, May 2021, pp. 10807–10814.
- [146] Z. Xie and S. H. Song, "FedKL: Tackling data heterogeneity in federated reinforcement learning by penalizing KL divergence," 2022, *arXiv:2204.08125*.
- [147] Z. Chai, Z. Fayyaz, and H. Ludwig, "Towards taming the resource and data heterogeneity in federated learning," in *Proc. OpML*, 2019, p. 4.
- [148] Z. Tang, Y. Zhang, S. Shi, X. He, B. Han, and X. Chu, "Virtual homogeneity learning: Defending against data heterogeneity in federated learning," in *Proc. PMLR*, 2022, p. 22.
- [149] X. Li, M. Jiang, X. Zhang, M. Kamp, and Q. Dou, "FedBN: Federated learning on non-IID features via local batch normalization," 2021, *arXiv:2102.07623*.
- [150] W. Lu, J. Wang, Y. Chen, X. Qin, R. Xu, D. Dimitriadis, and T. Qin, "Personalized federated learning with adaptive batchnorm for healthcare," *IEEE Trans. Big Data*, early access, May 23, 2022, doi: 10.1109/TBDDATA.2022.3177197.
- [151] M. G. Arivazhagan, V. Aggarwal, A. K. Singh, and S. Choudhary, "Federated learning with personalization layers," 2019, *arXiv:1912.00818*.
- [152] A. Fallah, A. Mokhtari, and A. Ozdaglar, "Personalized federated learning with theoretical guarantees: A model-agnostic meta-learning approach," in *Proc. Adv. Neural Inf. Process Syst.*, 2020, p. 12.
- [153] Y. Chen, W. Lu, X. Qin, J. Wang, and X. Xie, "MetaFed: Federated learning among federations with cyclic knowledge distillation for personalized healthcare," 2022, *arXiv:2206.08516*.
- [154] Z. Wang, X. Fan, J. Qi, C. Wen, C. Wang, and R. Yu, "Federated learning with fair averaging," in *Proc. 13th Int. Joint Conf. Artif. Intell.*, Aug. 2021, pp. 1615–1623.
- [155] C. T. Dinh, N. H. Tran, and T. D. Nguyen, "Personalized federated learning with Moreau envelopes," in *Proc. Adv. Neural Inf. Process Syst.*, 2020, p. 12.
- [156] Z. Hu, K. Shaloudegi, G. Zhang, and Y. Yu, "Federated learning meets multi-objective optimization," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 4, pp. 2039–2051, Jul. 2022.
- [157] V. Smith, K. Chiang, M. Sanjabi, and A. S. Talwalkar, "Federated multi-task learning," in *Proc. Adv. Neural Inf. Process Syst.*, vol. 30, 2017, pp. 1–19.
- [158] L. Corinzia, A. Beuret, and J. M. Buhmann, "Variational federated multi-task learning," 2019, *arXiv:1906.06268*.
- [159] T. Li, S. Hu, A. Beirami, and V. Smith, "Ditto: Fair and robust federated learning through personalization," in *Proc. PMLR*, 2021, pp. 6357–6368.
- [160] Y. Deng, M. M. Kamani, and M. Mahdavi, "Adaptive personalized federated learning," 2020, *arXiv:2003.13461*.
- [161] M. Mohri, G. Sivek, and A. T. Suresh, "Agnostic federated learning," in *Proc. PMLR*, 2019, pp. 4615–4625.
- [162] T. Li, M. Sanjabi, A. Beirami, and V. Smith, "Fair resource allocation in federated learning," in *Proc. ICLR*, New Orleans, LA, USA, 2019, p. 27.
- [163] J. Zhang, C. Li, A. Robles-Kelly, and M. Kankanhalli, "Hierarchically fair federated learning," 2020, *arXiv:2004.10386*.
- [164] F. Grimberg, A. Hartley, Mary, M. Jaggi, and S. P. Karimireddy, "Weight erosion: An update aggregation scheme for personalized collaborative machine learning," in *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*, vol. 12444, 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, Switzerland: Springer, 2020, pp. 160–169.
- [165] M. Beaussart, M. A. Hartley, F. Grimberg, and M. Jaggi, "WAF-FLE: Weighted averaging for personalized federated learning," in *Proc. NeurIPS Workshop New Frontiers Federated Learn. (NFFL)*, 2021, p. 9.
- [166] Y. Xia, D. Yang, W. Li, A. Myronenko, D. Xu, H. Obinata, H. Mori, P. An, S. Harmon, E. Turkbey, B. Turkbey, B. Wood, F. Patella, E. Stellato, G. Carrafiello, A. Ierardi, A. Yuille, and H. Roth, "Auto-FedAvg: Learnable federated averaging for multi-institutional medical image segmentation," 2021, *arXiv:2104.10195*.
- [167] M. Hong, S.-K. Kang, and J.-H. Lee, "Weighted averaging federated learning based on example forgetting events in label imbalanced non-IID," *Appl. Sci.*, vol. 12, no. 12, p. 5806, Jun. 2022.
- [168] L. Qu, S. Song, and C.-Y. Tsui, "FedDQ: Communication-efficient federated learning with descending quantization," in *Proc. GLOBECOM/IEEE Global Commun. Conf.*, Dec. 2022, pp. 2021–2031.
- [169] G. Kim, J. Kim, and B. Han, "Communication-efficient federated learning with acceleration of global momentum," 2022, *arXiv:2201.03172*.
- [170] G.-L. Pereteau, A. Alansary, and J. Passerat-Palmbach, "Split HE: Fast secure inference combining split learning and homomorphic encryption," 2022, *arXiv:2202.13351*.
- [171] O. Li, J. Sun, X. Yang, W. Gao, H. Zhang, J. Xie, V. Smith, and C. Wang, "Label leakage and protection in two-party split learning," in *Proc. ICLR*, 2021, p. 27.
- [172] X. Yang, J. Sun, Y. Yao, J. Xie, and C. Wang, "Differentially private label protection in split learning," 2022, *arXiv:2203.02073*.
- [173] E. T. Zhang, "Computational privacy with split learning: Benchmarking of algorithmic defenses against reconstruction attacks," M.S. thesis, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2021. [Online]. Available: <https://hdl.handle.net/1721.1/139497>
- [174] E. Erdogan, A. Küpçü, and A. E. Cicek, "SplitGuard: Detecting and mitigating training-hijacking attacks in split learning," in *Proc. 21st Workshop Privacy Electron. Soc.*, Nov. 2022, pp. 125–137.
- [175] D. Pasquini, G. Ateniese, and M. Bernaschi, "Unleashing the tiger: Inference attacks on split learning," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Nov. 2021, pp. 2113–2129.
- [176] E. Erdoğan, A. Küpçü, and A. E. Çiçek, "UnSplit: Data-oblivious model inversion, model stealing, and label inference attacks against split learning," in *Proc. 21st Workshop Privacy Electron. Soc.*, Nov. 2022, pp. 115–124.
- [177] M. A. Khan, V. Shejwalkar, A. Houmansadr, and F. M. Anwar, "Security analysis of SplitFed learning," 2022, *arXiv:2212.01716*.
- [178] X. Chen, J. Li, and C. Chakrabarti, "Energy and loss-aware selective updating for SplitFed learning with energy harvesting-powered devices," *J. Signal Process. Syst.*, vol. 94, pp. 961–975, Jul. 2022.
- [179] M. Joshi, A. Pal, and M. Sankarasubbu, "Federated learning for healthcare domain-pipeline, applications and challenges," *ACM Trans. Comput. Healthcare*, vol. 3, no. 4, pp. 1–36, Oct. 2022.
- [180] J. H. Yoo, H. Jeong, J. Lee, and T.-M. Chung, "Open problems in medical federated learning," *Int. J. Web Inf. Syst.*, vol. 18, nos. 2–3, pp. 77–99, Oct. 2022.
- [181] Z. Shae, K. Chen, C. Chang, Y. Tsai, C. Chou, W. I. Baskett, C. Shyu, and J. J. P. Tsai, "Thoughts on non-IID data impact in healthcare with federated learning medical blockchain," in *Proc. IEEE 4th Int. Conf. Cognit. Mach. Intell. (CogMI)*, Dec. 2022, pp. 20–26.
- [182] M. Khodak, R. Tu, T. Li, L. Li, M.-F. F. Balcan, V. Smith, and A. Talwalkar, "Federated hyperparameter tuning: Challenges, baselines, and connections to weight-sharing," in *Proc. NeurIPS*, vol. 34, 2021, pp. 19184–19197.



CHAMANI SHIRANTHIKA (Graduate Student Member, IEEE) received the B.A.Sc. degree (Hons.) in information technology from the University of Moratuwa, Sri Lanka, in 2018, and the M.A.Sc. degree in electrical engineering from National Taipei University, Taiwan, in 2021. She is currently pursuing the Ph.D. degree with the School of Engineering Science, Simon Fraser University, Canada, under the supervision of Prof. Parvaneh Saedi and Prof. Ivan V. Bajić. She is also a Research Assistant with the School of Engineering Science, Simon Fraser University, under the supervision of Prof. Parvaneh Saedi and Prof. Ivan V. Bajić. Her research interests include medical imaging, collaborative intelligence, and reinforcement learning. She received the 2021 Falcon Reward for Character and Academic Excellence from the National Taipei University and the 2021 Best Master Thesis Award from the IEEE Taipei Section.



PARVANEH SAEEDI (Member, IEEE) received the M.A.Sc. and Ph.D. degrees from the University of British Columbia, in 1998 and 2004, respectively. From 2004 to 2006, she was with MacDonald, Dettwiler, and Associates Ltd., (MDA), Richmond, BC, Canada, where she was a Research Scientist for various projects in collaboration with Canadian Space Agency, European Space Agency, and Defence Research and Development Canada. Since 2007, she has been with the School of Engineering Science, Simon Fraser University, where she is currently a Professor. Her research interests include medical image processing and analysis, geospatial data processing and analysis (optical and radar data), 3D reconstruction, computer vision, and machine learning.



IVAN V. BAJIĆ (Senior Member, IEEE) is currently a Professor in engineering science and the Co-Director of the Multimedia Laboratory, Simon Fraser University, Burnaby, BC, Canada. His research interests include signal processing and machine learning, with applications in multimedia signal compression, processing, and collaborative intelligence. His groups work has received the 2023 IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY Best Paper Award, conference paper awards at ICME 2012, ICIP 2019, and MMSP 2022, and other recognitions (e.g., paper award finalist, top n%) at Asilomar, ICIP, ICME, ISBI, and CVPR. He is currently the Chair of the IEEE Multimedia Signal Processing Technical Committee. He has served on the organizing and/or program committees of the main conferences in his field. He has received several awards in these roles, including the Outstanding Reviewer Award (six times), the Outstanding Area Chair Award, and the Outstanding Service Award. He was on the editorial boards of the IEEE TRANSACTIONS ON MULTIMEDIA, *IEEE Signal Processing Magazine*, and *Signal Processing: Image Communication*. He is currently a Senior Area Editor of the IEEE SIGNAL PROCESSING LETTERS.

• • •