A Review of Medical Federated Learning: Applications in Oncology and Cancer Research

Alexander Chowdhury[1], Hasan Kassem[3], Nicolas Padoy[3,4], Renato Umeton[1,2,5,6], and Alexandros Karargyris[3,4(✉)]

[1] Dana-Farber Cancer Institute, Boston, MA, USA
renato_umeton@dfci.harvard.edu
[2] Massachusetts Institute of Technology, Cambridge, MA, USA
[3] ICube, University of Strasbourg, CNRS, Strasbourg, France
[4] IHU Strasbourg, Strasbourg, France
alexandros.karargyris@ihu-strasbourg.eu
[5] Weill Cornell Medicine, New York, NY, USA
[6] Harvard T.H. Chan School of Public Health, Boston, MA, USA

Abstract. Machine learning has revolutionized every facet of human life, while also becoming more accessible and ubiquitous. Its prevalence has had a powerful impact in healthcare, with numerous applications and intelligent systems achieving clinical level expertise. However, building robust and generalizable systems relies on training algorithms in a centralized fashion using large, heterogeneous datasets. In medicine, these datasets are time consuming to annotate and difficult to collect centrally due to privacy concerns. Recently, Federated Learning has been proposed as a distributed learning technique to alleviate many of these privacy concerns by providing a decentralized training paradigm for models using large, distributed data. This new approach has become the defacto way of building machine learning models in multiple industries (e.g. edge computing, smartphones). Due to its strong potential, Federated Learning is also becoming a popular training method in healthcare, where patient privacy is of paramount concern. In this paper we performed an extensive literature review to identify state-of-the-art Federated Learning applications for cancer research and clinical oncology analysis. Our objective is to provide readers with an overview of the evolving Federated Learning landscape, with a focus on applications and algorithms in oncology space. Moreover, we hope that this review will help readers to identify potential needs and future directions for research and development.

Keywords: Federated learning · Cancer research · Clinical oncology · Privacy-preserving computation · Healthcare informatics · Distributed learning

A. Chowdhury and H. Kassem - Joint first authors,
R. Umeton and A. Karargyris - Joint senior authors.

A. Crimi and S. Bakas (Eds.): BrainLes 2021, LNCS 12962, pp. 3–24, 2022.
https://doi.org/10.1007/978-3-031-08999-2_1

Highlights

- Federated learning (FL) has the potential to become the primary learning paradigm for distributed cancer research, but specific hurdles have slowed its adoption in the clinical setting.
- Labeled medical data is still extremely scarce; this problem also affects federated learning. A plethora of cancer datasets exist (e.g. TCIA, TCGA, Gene Expression Omnibus, etc.), but few of them are labeled for supervised learning. The ones that are labeled (i.e., the Wisconsin Breast Cancer dataset - for classification, the BraTS dataset - for image segmentation, the Kaggle datasets for skin cancer) are the ones most commonly seen being used in FL.
- The largest majority of papers we found use cancer datasets for benchmarking purposes: very few federated learning works solve an actual clinically relevant question. Many of the papers we reviewed propose new software frameworks, and virtually none follow-up with a clinical trial. This leaves FL absent from the field of clinical oncology, based on our literature review.
- The compliance and security aspect of healthcare still poses the largest hurdle. Commercial entities such as EHR vendors (e.g., Epic Systems, Cerner, Meditech, Allscripts, etc.), PACS vendors (e.g., GE, Philips, Hitachi, Siemens, Canon, etc.), and other hardware manufacturers (e.g., Nvidia, Intel, etc.) seem to be the best positioned to start pulling together resources, data, and models that use FL to improve patient outcomes.

1 Introduction

Over the past decade, machine learning has witnessed rapid growth due to the proliferation of deep learning. Fueled by large-scale training databases [1], these data driven methods have gained significant popularity. Thanks to rapidly evolving architectures, (e.g., AlexNet [2], GoogLeNet [3], ResNet [4]) convolutional neural networks (CNNs) have demonstrated consistent improvement on difficult computer vision tasks including classification, object detection, and segmentation. Other areas of machine learning, such as natural language understanding, recommendation systems and speech recognition, have also seen outstanding results in their respective applications through the introduction of novel approaches such as transformers [5,6], DLRM [7] and RNN-T [8].

Such advancements in artificial intelligence and machine learning have already disrupted and transformed healthcare through applications ranging from medical image analysis to protein sequencing [9–12]. And yet, while there are over 150 AI-based interventions that are approved by the FDA (an updated list with a focus on radiology can be reviewed at https://aicentral.acrdsi.org), many open questions persist about how to best deploy existing AI solutions in healthcare environments [13]. In addition to getting existing solutions deployed, there are many challenges that must be overcome during the training process. A consistent bottleneck has been the need for large amounts of heterogeneous data to train accurate, robust and generalizable machine learning models. However, most healthcare organizations rarely carry data in such large quantities,

especially in the case of homogeneous populations or rare diseases with scarce amounts of cases.

A common way data scientists attempt to overcome this issue is by first pre-training a model on large, generic datasets (e.g., ImageNet [1]), and then fine-tuning them on specific medical tasks of interest. However, even with this approach, underperformance or generalizability issues [14] may persist. This is often the case for medical tasks where there exists a large domain shift between medical data (e.g., brain MRI, abdomen CT, genomics) and general purpose public datasets such as ImageNet [1], MIMIC-CXR [15], ChexPert [16], etc. More recently, Self Supervised Learning (SSL) approaches have demonstrated promising results in performance using large unlabelled datasets, thus alleviating the need for annotations; however, even with such SSL approaches, the need for access to large amounts of heterogeneous medical data is still necessary to train robust medical ML algorithms [17,18].

In addition to large, heterogeneous datasets, the other most common bottleneck for ML algorithm training is computational power. The need for access to considerably efficient computing resources (e.g., processing power, memory, storage space) led to the field of distributed systems [19]. Within this area, distributed machine learning has evolved as a setting where algorithms are implemented and run on multiple nodes, leveraging larger amounts of data and computational resources, thus improving performance and efficiency. The core concept of distributed learning lies in the parallelization of algorithms across computational nodes [19], but these processes are run without considering any constraints that might need to be imposed by these nodes (e.g., considering that data used across these nodes comes from different distributions). Because of this, the majority of practical applications in collaborative learning fail to keep the assumption of Independent-and-Identically-Distributed (IID) data across nodes, such as user data from mobile devices or healthcare data from different geographic and demographic properties. Federated Learning emerged as a distributed learning paradigm that takes into account several practical challenges, and differentiates itself from traditional distributed learning settings, as noted by Google [20], by addressing four main themes: statistical heterogeneity of data across nodes, data imbalance across nodes, limited communication in the distributed network (e.g., loss of synchronization, variability of communication capabilities), and the possibility of a large number of nodes relative to the amounts of data.

In the Federated Learning setting, a "federation" of client sites with their own datasets train models locally and then send their updates to a server. The weights are the only information passed over lines of communication aiming at preserving privacy. The model weights are then aggregated in the server from the client updates, and the resulting aggregated model weights are sent back to the clients for the next round of training. Because of its strong potential to preserve privacy with client sites, such as hospitals, by keeping their data in-house, Federated Learning has seen a rise in popularity over the last several years, especially in the medical domain.

Specifically, large-scale projects have been developed for facilitating collaboration of medical institutions around the globe with the aid of Federated Learning, in both academic and industrial areas [21]. Trustworthy Federated Data Analytics [22], German Cancer Consortium's Joint Imaging Platform [23], and the Melloddy project [24] were developed to improve academic research in various healthcare applications by combining multiple institutions' efforts. In industry, the HealthChain project [25] aims to develop and deploy a Federated Learning framework across four hospitals in France to help determine effective treatments for melanoma and breast cancer patients. Additionally, the Federated Tumour Segmentation initiative (FeTS) [26,27] is an international collaboration between 30 healthcare institutions aimed at enhancing tumor boundary detection, for example, in breast and liver tumors. In another international effort [28], researchers trained ML models for mammogram assessment across a federation of US and Brazilian healthcare providers.

In light of all these efforts, and given the growing adoption of Federated Learning in healthcare, we believe that the cancer research community is lacking a much needed review of the current state-of-the-art. Therefore, with this review we aim at providing an comprehensive list of Federated Learning algorithms, applications and frameworks proposed for cancer analysis. We envision that this review can function as a quick reference for Federated Learning's applications in cancer and oncology, and provide a motivation for research in specific directions.

The review is structured as follows. In Sect. 2 we give an overview of Federated Learning to introduce the reader to related concepts. The main body of this review is found in Sect. 3, which we begin by providing the search query along with the inclusion/exclusion criteria for papers. After this, we provide a summary of the current literature for: 1) Federated Learning algorithms in cancer analysis, 2) Federated Learning frameworks developed for cancer research, and 3) Algorithms developed to preserve privacy under Federated Learning settings. Finally, we conclude this review by offering our thoughts on the needs and potential future directions for Federated Learning in the cancer research and clinical oncology space.

2 Federated Learning Overview

Federated Learning was first introduced as a decentralized distributed machine learning paradigm by Google [20]. The standard Federated Learning paradigm that is outlined in this paper is as follows: i) Multiple client sites, each containing a local dataset that remains at the client site during the entirety of training, connect to a global server; ii) A global model is initialized in the global server, and the weights from this global model are passed to each of the local client sites; iii) Each client site trains a local version of the global model on their respective dataset, and then sends the updated model weights to the global server; iv) The global server updates the global model by aggregating the weights it receives from the local clients, and then passes a copy of the updated global model to each of the clients. The process that occurs between steps i–iv is called a round, and during federated training, steps i–iv are repeated for multiple rounds until the global model converges to a local minima. The most important aspect of

this process is step iii. During this step, all data used for training is kept strictly on the local clients' machines. The only information that is passed between the clients and the server are weight updates. This enables multiple sites to pool their data for training of a global model while still maintaining data privacy. During step iv, the authors use an algorithm that they coin federated averaging to aggregate the weights. In this algorithm, each weight updated is weighted by the size of the client dataset from which it comes, relative to the size of the other client datasets. The aforementioned clients-server topology is known as Centralized Federated Learning. One other topology has been found in research [29], Decentralized Federated Learning, in which clients communicate peer-to-peer without a central server.

Federated Learning can be broken down into three main subtypes [30]: Horizontal Federated Learning, Vertical Federated Learning, and Federated Transfer Learning. All three of these subtypes follow the core Federated Learning paradigm, which is decentralized data pooling through the use of weight sharing and aggregation between multiple clients and a global server. They are distinguished by the way in which their data sources differ. In Horizontal Federated Learning, every client site has different users in their data, but all of these users share similar features that are extracted by the networks. In Vertical Federated Learning, users are the same across all client sites, but each client sites' data consists of different features, so the same user will be analyzed through different modalities depending on the client site. In Transfer Federated Learning, the client sites don't have users or features in common, but the tasks in their datasets are still marginally related, so pooling them together typically leads to more robust network training. For a more general review of Federated Learning, readers are referred to [29,31,32]. Here we also list common Federated Learning platforms: OpenFL [33], PySyft[1], Tensorflow-Federated[2], FedML [34], Flower[3], NVIDIA Clara[4], Personal Health Train (PHT[5]).

3 Review

3.1 Search Design

The literature review was conducted in October 2021 by searching Google Scholar for papers published between 2019 and 2021 that matched the query: *federated AND (cancer OR cancers OR tumor OR tumors OR oncology)*.

We chose this time period for our search query due to the fact that Google didn't publish their seminal Federated Learning paper [35] until 2017, so we didn't see a large amount of medical applications until than. A visual representation of the split of the material reviewed is presented in Fig. 1 and our review process is shown in Fig. 2.

[1] https://github.com/OpenMined/PySyft.
[2] https://medium.com/tensorflow/introducing-tensorflow-federated-a4147aa20041.
[3] https://flower.dev/.
[4] https://developer.nvidia.com/clara.
[5] https://pht.health-ri.nl/.

Through our review process we identified two main categories of Federated Learning applications related to cancer and oncology: whether the study was designed exclusively with cancer as its intended use-case, or whether cancer datasets were used for benchmarking a general method (Fig. 1-Category). Every

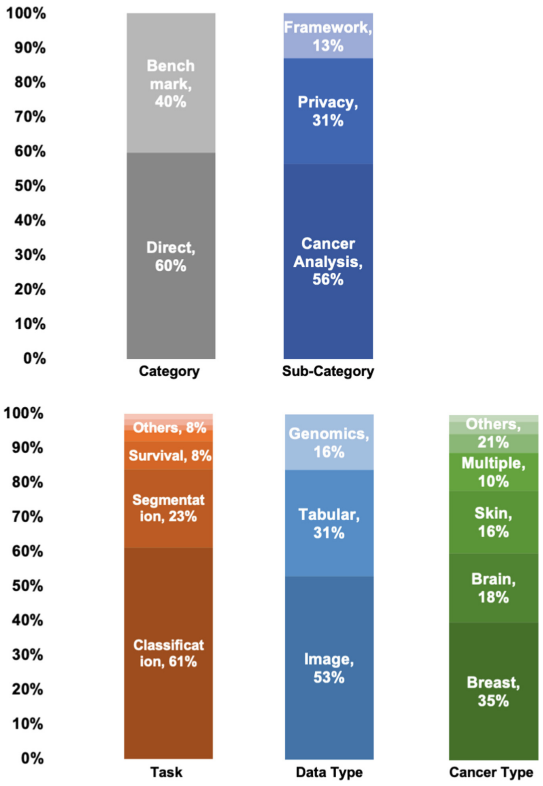

Fig. 1. Split of the papers reviewed: **Category** and **Sub-Category** represent the paper scope. **Task** refers to the machine learning task, while **Data Type** and **Cancer Type** relate to the FL input data.

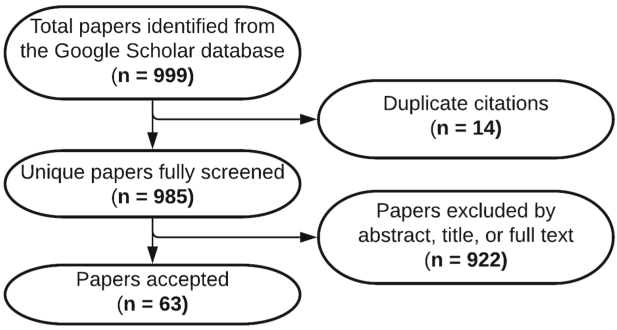

Fig. 2. A visual representation of our process for including papers for this review.

category is also further divided into three sub-categories: the first one contains the Federated Learning feasibility studies and methods that have been applied to the analysis of cancer datasets (i.e., 'Framework' in Fig. 1-Sub-Category). The second contains Federated Learning frameworks proposed or developed for 'Cancer Analysis', although almost all fail to secure relevant and novel cancer datasets and hence resort to open-access data. Finally, the third sub-category contains Federated Learning studies that address and analyze 'Privacy' of cancer data and computation.

3.2 Federated Learning Algorithms

Algorithms Designed for Cancer: Based on our literature search we identified that Federated Learning has been explored in many cancer studies, where the aim is either comparing Federated Learning to conventional centralized data analysis approaches in terms of performance, or developing novel methods to solve various challenges faced when using Federated Learning (e.g., domain shift, label deficiency, ...). In the most common training scenario, researchers simulate a Federated Learning environment by taking an existing dataset and dividing it into subsets using a partitioning scheme, where each subset represents a client in a Federated Learning group.

Federated Learning has been applied on detecting brain tumors in several studies [36–39]. In [36], the authors used the 'Brain MRI Segmentation' dataset from Kaggle for low-grade glioma segmentation [40], dividing the dataset into 5 "client" sites. The authors designed a network that achieves state-of-the-art results on the task of glioma segmentation, and those results remained consistent when they applied it to a Federated Learning setting. In [37], two separate Federated Learning environments for brain tumor segmentation were simulated using the BraTS dataset [41]. In both environments, the Federated Learning model was compared against two other collaborative learning techniques, and outperformed both. It also achieved nearly 99% of the DICE score obtained by a model trained on the entire dataset with no decentralization. Similarly, [38] demonstrated comparable performance between federated averaging and data sharing for brain tumor segmentation on the BraTS dataset [41]. Sheller et al. also showed how Federated Learning improves the learning of each participating institution both in terms of performance on local data and performance on data from unseen domains. In [39], the authors presented a comparison between a Federated Learning approach and individual training of a 3D-Unet model to segment glioblastoma in 165 multi-parametric structural MRI (mpMRI) scans. The Federated Learning approach is shown to yield superior quantitative results.

Additional studies have explored Federated Learning on a variety of other cancers, including less common types. Some of the types covered in the uses cases we reviewed included: skin cancer [42,43], breast cancer [44,45], prostate cancer [46], lung cancer [47], pancreatic cancer, anal cancer, and thyroid cancer. [42] used the ISIC 2018 dataset [48] to simulate a Federated Learning environment for classifying skin lesions. They first partitioned the dataset among multiple mock client sites, then used a Dual-GAN [49] to augment each clients' dataset. A classifier was then trained in a federated environment on the augmented datasets.

In [43], the authors use the ISIC 2019 Dermoscopy dataset [48] to demonstrate proof-of-concept for a skin lesion detection device trained using federated learning. In Roth et al. [44], a real-world experiment of federated breast density classification was performed using NVIDIA's Clara framework. The authors developed a breast density classification model with mammography data from 7 different institutions. The global federated model showed significant improvements over the locally trained models when validated against their own data as well as external site validation. In [50] and [45], the authors demonstrate the ability to successfully apply vertical federated learning (VFL) to cancer analysis, using VFL to create a survival prediction model for breast cancer. [46] performed prostate image segmentation in a federated setting. They showed how Federated Learning improves model performance on local datasets. [47] described a large experiment on 20K lung cancer patients across 8 institutes and 5 countries. They trained a logistic regressor on these distributed data. To train the LR coefficients in a distributed manner they used the Alternating Direction Method of Multipliers (ADMM). The data included tumor staging and post-treatment survival information.

In [51], the authors tackle the task of pancreas segmentation for patients with pancreatic cancer. Advanced tools to correctly identify pancreatic cancer are extremely important since pancreatic cancer is normally only detectable once it is late-stage, leading to extremely low survival rates [52]. They used two datasets obtained from hospitals in Japan and Taiwan to simulate a Federated Learning environment. The resulting model was able to better identify pancreas from both datasets than models trained only on one site and validated on the other. Concluding with similar results, [53] tested several deep learning architectures for federated thyroid images classification, and Choudhury et al. [54] used data from 3 different sites to create a prediction model for patients with anal cancer, an extremely rare form of cancer, who received radical chemoradiotherapy. The large and diverse group of examples given here demonstrates the robustness and versatility of the Federated Learning paradigm, as well as its ability to improve automated analysis on more rare cancer cases [51,53,54].

In addition to having many use cases with specific cancer types, Federated Learning's applications in genomics have also been a popular focal point for research [55,56]. [55] performed federated gene expression analysis on breast cancer and skin cancer data. [56] adapted the Cox proportional hazards (PH) model [57] in a Federated Learning setting for survival analysis. Noting that adapting this method in a distributed manner is non-trivial due to its non-separable loss function, they implemented a discrete time extension of this model with a separable loss function, and validated their method on the Genome Atlas Data (TCGA)[6], showing comparable performance to the centralized approach.

While the bulk of the papers we've reviewed so far focus purely on designing federated algorithms that can predict different aspects of cancer with high degrees of accuracy, a large sub-group of the papers in our review also aim at addressing challenges federated learning currently faces. For many papers, that challenge is either data heterogeneity [58–65], a common barrier in the medi-

[6] https://www.cancer.gov/tcga.

cal field where patients can be subject to different geographic and demographic conditions, or label deficiency [66,67], where it is not always guaranteed that clients' sites will have access to labeled data.

Addressing label deficiency, [66] introduced a new Federated Semi-Supervised Learning (FSSL) approach for skin lesion classification. Their method is inspired by knowledge distillation [68], where they model disease relationships in each client by a relation matrix calculated from the local model output, then aggregate the relation matrices from all clients to form a global one that is used locally in each round to ensure that clients will have similar disease relationships. In [67], the authors proposed a semi-supervised Federated Learning method, FedPerl. The method was inspired by peer learning from educational psychology and ensemble averaging from committee machines and aims to gain extra knowledge by learning from similar clients i.e. peers. This encouraged the self-confidence of the clients by sharing their knowledge in a way that did not expose their identities. Experimental setup consisted of 71,000 skin lesion images collected from 5 publicly available datasets. With little annotated data, FedPerl outperformed state-of-the-art FSSL methods and the baselines by 1.8% and 15.8%, respectively. It also generalized better to an unseen client while being less sensitive to noisy ones.

Another challenge that frequently occurs in Federated Learning is domain shift, which is caused by heterogeneity in datasets due to different scanners and image acquisition protocols at different sites. Many papers modify the original FL algorithm to account for this. Jimenez et al. [58] designed a novel weight aggregation algorithm designed to address the problem of domain shift between data from different institutions. This study utilized one public and two private datasets, and the final global model outperformed previous Federated Learning approaches. Similarly, [59] introduced a new weight aggregation strategy and showed its efficiency on pancreas CT image segmentation. [60] built on the work of [51] by developing a Federated Learning algorithm that can learn multiple tasks from heterogeneous datasets, making use of a training paradigm the authors call dynamic weight averaging (DWA). Specifically, they trained a model on the binary-classification problem of segmenting the pancreas from background as well the multi-label classification problem of segmenting healthy and tumorous pancreatic tissue and background. During the global aggregation step, the weight value for each client update was adjusted based on the variation of loss values from the previous rounds. DWA outperforms federated averaging (FedAvg) and FedProx [69], another federated weight aggregation scheme designed to handle heterogeneous networks.

In Guo et al. [61], the authors addressed the problem of domain shift while applying their algorithm to the task of MRI reconstruction, using 4 different MRI datasets; FastMRI, BraTS, IXI, and HPKs. Their algorithm, Federated Learning-based Magnetic Resonance Imaging Reconstruction with Cross-site Modeling (FL-MRCM), uses an adversarial domain identifier to align latent features taken from the encoders of 2 different sites, avoiding sharing of data while taking advantage of multiple sites' data. In all experiments, FL-MRCM came closest to reaching the upper-bound score of training a network on the entire

dataset. In the same space, to alleviate domain shift performance impact, [62] proposed a new method to train deep learning algorithms in Federated Learning settings based on the disentanglement of the latent space into shape and appearance information. Their method only shared the shape parameters to mitigate domain shifts between individual clients. They presented promising results on multiple brain MRI datasets.

Researchers in [63] proposed a method to address domain shift issues in terms of performance and stability based on sharing the parameters of batch normalization across clients but keeping the batch norm statistics local. Given that these statistics are not shared with the central server they argued that there is better protection from privacy attacks. They demonstrated their algorithm on breast histopathology image analysis (Camelyon 2016[7] and Camelyon 2017[8] datasets). In [64] a key-problem of digital pathology is addressed via federated learning: stain normalization across multiple laboratories and sites. They apply GANs in a Federated Learning environment to solve the problem of color normalization that arises due to different staining techniques used at different sites. Here, a central discriminator is trained to be extremely robust by making use of several decentralized generators.

Domain shift in Federated Learning has been also studied in Neural Architecture Search (NAS). [65] applied AutoML, a NAS approach, in a federated setting for prostate image segmentation. To address domain shift, they trained a 'supernet' consisting of several deep learning modules in a federated setting, then personalize this supernet in each client by searching for the best path along the supernet components according to each client.

General Algorithms Benchmarked on Cancer Datasets: Cancer datasets are also commonly used as benchmarks for evaluating general Federated Learning approaches. BraTS [41], HAM10000 [70], Wisconsin Breast Cancer dataset [71], and TCGA[9] were the most common datasets used in the papers we sourced for this review.

The BraTS dataset is an imaging dataset used to train computer vision models for brain tumor segmentation. It is frequently used as a benchmark for state-of-the-art image analysis algorithms. Chang et al. [72] performed a Federated Learning experiment on BraTS [41] using GANs in a similar setting to [64]. They use several decentralized discriminators, placed at mock client sites, to train a centralized discriminator at the client. Receiving synthetic images from a large amount of generators allowed the authors to augment the dataset in a decentralized fashion and train the discriminator to achieve very high accuracy. In some cases the classifier was able to outperform non-Federated Learning trained models, using Area Under the Curve (AUC) as a performance metric. In [73], the authors address the problem of domain shift while benchmarking on BraTS. They partition the network, and place a copy of each partition at each client site. They then place the rest of the network on a centralized server. Lower-level features taken from each client site are aggregated and passed as input to the

[7] https://camelyon16.grand-challenge.org/.
[8] https://camelyon17.grand-challenge.org/.
[9] https://www.cancer.gov/tcga.

central network, which learns to be robust against domain shift. This paradigm leads to extremely strong training results, especially as the domain shift becomes more pronounced.

The HAM10000 dataset is a multi-source dermatoscopic image dataset of pigmented lesion used for skin lesion detection and segmentation. Similar to BraTS, it frequently appears in many computer vision applications, such as [74], where the authors proposed a new server aggregation method addressing statistical heterogeneity that may be present between the participating datasets. The weights are calculated to be inversely proportional to the difference between the corresponding client model distribution and the global model distribution. They validated their new method on several benchmarks, including HAM10000 [70]. In [75] a new Federated Learning strategy was introduced for tackling non iid-ness in data. Training one epoch on each local dataset was done over several communication rounds. The approach was evaluated on various datasets, including HAM10000, and showed superior results to similar methods, such as FedAVG.

The Wisconsin Breast Cancer dataset [71] is another versatile dataset that is used for benchmarking many different classification algorithms. It is a simple dataset that is easy to integrate into most ML workflows, consisting of positive and negative breast cancer samples, and several numerical features describing those samples. Salmeron et al. [76] used this dataset to simulate a Federated Learning environment. The authors then used this environment to train a Fuzzy Cognitive Map (FCM) [77] classifier that outperformed clients that were trained individually as well as a model trained on the entire dataset. Researchers in [78] extended SQL-based training data debugging (RAIN method) for Federated Learning. They demonstrated this extension on multiple datasets, including the Wisconsin Breast Cancer dataset [71]. [79] introduced a new Federated Learning strategy that showed comparable performance to federated averaging while giving two benefits: communication efficiency and trustworthiness, via Stein Variational Gradient Descent (SVGD) which is a non-parametric Bayesian framework that approximates a target posterior distribution via non-random and interacting particles. They performed extensive experiments on various benchmarks, including binary classification of breast cancer data. [80] introduced a new federated setup that requires less communication costs and no centralized model sharing; clients learn collaboratively and simultaneously without the need of synchronization. They validated their setup, termed gradient assisted learning, on various datasets including breast cancer, and showed comparable performance with state-of-the-art methods but with less communications costs. [81] investigated how to mitigate the effects of model poisoning, a scenario where one or more clients upload intentionally false model parameters (or are forced to do so, e.g. by being hacked). They introduced new model-poisoning attacks, and showed that the methods of mitigating the effects of these attacks still need development. In [82], a method for building a global model under the Federated Learning setting was proposed by learning the data distribution of each client and building a global model based on these shared distributions.

The Cancer Genome Atlas (TCGA) is a public consortium of cancer data created for the purpose of benchmarking healthcare analysis algorithms. In [83] a method was proposed for matrix factorization under Federated Learning settings. Specifically, they extended the FedAvg method to allow for robust matrix factorization. They benchmarked this method on the Cancer Genome Atlas (TCGA). Benchmarking on the same data, [84] introduced two Federated Learning algorithms for matrix factorization and applied them to a data clustering task.

3.3 Federated Learning Frameworks

Frameworks Developed for Cancer Analysis: In [85], the authors designed a decentralized framework which they coined Braintorrent. This framework removes the global server from the traditional FL paradigm, and instead allows sites to communicate their weights with one another directly. The framework was tested on the task of whole-brain segmentation, and demonstrates impressive results, outperforming traditional Federated Learning with a global server and achieving performance close to that of a model trained using pooled data. [86] designed an open source framework to facilitate analysis of local data between institutions in order to create a model for oral cavity cancer survival rates using data from multinational institutions. [87] introduced a framework, GenoPPML, that is a combination of Federated Learning and multiparty computation. The framework utilizes differential privacy and homomorphic encryption for guaranteeing preserved privacy, and it was mainly built for regression for genomics data. In [88] the authors proposed a framework to train on skin lesion images using IoT devices (smartphones). They further utilized Transfer Learning in this Federated Learning framework to circumvent the need of large, labelled data. The German National Cancer Center, an initiative whose primary goal is to foster multiclinical trials for development of improved diagnosis and treatment tools for cancer, recently released the Joint Imaging Platform (JIP) [89], a platform designed to build a foundation for Federated Learning scenarios. JIP provides containerized tools for Federated Learning, and many institutions have committed to testing JIP for use cases in the coming years. [90] provides another framework with multiple objectives and use cases. Here, the authors proposed a "marketplace" approach to federated learning: it provides the infrastructure and other computational resources for 3rd party applications to run in a Secure Multiparty Computation system; there, for sake of example, multiple computational tasks related to cancer research (from data normalization to Kaplan-Mayer analysis and COX regression) are treated as "Apps" and deployed into a secure and distributed environment.

General Frameworks: Because decentralized analysis of medical data is one of the most natural use cases for federated learning, cancer datasets are frequently included when benchmarking general federated learning frameworks. [91] introduced a framework for federated meta learning; a library for fast and efficient algorithm selection. They evaluated a prototype on various datasets including breast cancer dataset, showing better efficiency of their framework in finding the best algorithm for a given dataset against the ordinary grid search approach. In

[92], the authors design a classification framework for breast cancer that incorporates differential privacy. Similarly, [50] uses the Wisconsin Breast Dataset as once of their use cases for a privacy-verification FL framework.

3.4 Privacy Protection in Federated Learning Settings

One important benefit of Federated Learning for healthcare is its potential to mitigate privacy concerns. Although Federated Learning allows for multiple sites to train ML models on their data safely, there are still ways that this paradigm can be exploited. One very common exploitation is that dataset labels can be reconstructed from the gradients used during model training [93,94].

In this section we discuss research that addresses privacy concerns of Federated Learning in cancer. We present papers that either benchmark their privacy-concerned investigations and methods on cancer data, or those which study Federated Learning privacy exclusively for cancer applications.

Privacy Methods for Cancer: In [95], the authors proposed a combination of meta-heuristic methods to operate the whole mechanism of aggregation, separation of models as well as evaluation. They analyzed the results in terms of the accuracy of the general model as well as for security against poisoning attacks. [96] implemented differentially privacy SGD training in a cyclic Federated Learning setting of two clients, and did an extensive study on the trade-off between privacy and accuracy. They achieved an acceptable trade-off between accuracy and privacy, and tested their experiments on classification of tumorous genes. In [97] the authors benchmarked various differential privacy methods against skin lesion classification in Federated Learning settings. [98] demonstrated an approach to prevent access to intermediate model weights by using a layer for privacy protection. The aggregation server prevented direct connections between hosts so that interim model weights cannot be viewed during training.

In [99], the authors studied the effect that two different techniques to preserve privacy had on a Federated Learning environment: injecting samples with noise or sharing only a fraction of the model's weights. Using the BraTS dataset [41] for brain tumor segmentation, they found that leaving out up to 40% of the model's weights only affected accuracy by a negligible amount. Using the BraTS dataset [41] the authors in [100] extended Private Aggregation of Teacher Ensembles (PATE) [101] which is used as an aggregation function using the teacher-student paradigm to enable privacy preserving training: teacher models are trained on private datasets and the student model (global) is trained on a public dataset using those teacher models. This extension applied a dimensionality reduction method to increase sensitivity for segmentation tasks. They validated their approach on three (2) common dimensionality reduction methods to assess differential privacy: PCA, Autoencoder and Wavelet transforms. [102] used noise injection as a successful privacy preservation technique for analyzing gigapixel whole slide images. [103] created a hybrid environment for encryption of medical data using blockchain technologies, Federated Learning, and homomorphic encryption. Homomorphic encryption is also used in [104], where it is leveraged to show

secure and accurate computation of essential biomedical analysis tasks, including Kaplan-Meier survival analysis in oncology and genome-wide association studies (GWAS) in medical genetics. The authors demonstrate this through the use of their framework, FAMHE. GWAS data was also at the center of the SAFETY framework [105], where a hybrid deployment of both homomorphic encryption and secure hardware (Intel SGX) provides a good trade-off in terms of efficiency and computational support for secure statistical analysis. Rrajotte et al. [106] developed a framework called FELICIA (Federated Learning with a Centralized Adversary), which uses the PrivGAN architecture [107] to make use of data from multiple institutions and create higher-quality synthetic training data without sharing data among sites. [108] used differential privacy and demonstrated how the performance was still comparable to the centralized experiments despite the privacy-performance trade-off. They also showed empirically how the model with differential privacy became immune against adversarial attacks, and evaluated all their approaches on liver image segmentation.

General Privacy-Preserving Methods Benchmarked on Cancer Data sets: [109] introduced Federboost, a Federated Learning method for gradient boosting decision trees (GDBT). Their method can be applied for vertical and horizontal Federated Learning, and is characterized by the ease of ensuring secure model sharing. They demonstrated security and comparable performance to centralized settings using various datasets including breast cancer gene data from TCGA. [110] introduced a new Federated Learning approach for mitigating possible privacy breaches when sharing model weights. Their method was evaluated on various benchmark datasets including breast cancer data, and showed comparable performance to the conventional Federated Learning approaches while being more robust to gradient leaks, i.e. more privacy-preserving. [111] developed a homomorphic encryption framework on FPGA, aiming to accelerate the training phase under Federated Learning with the most possible encryption. They demonstrated performance improvement in speed benchmarking on multiple datasets including the Wisconsin Breast Cancer dataset.

In [112], the authors proposed attacks for two machine learning algorithms, logistic regression and XGBoost, in a Federated Learning setting. In this study the adversary does not deviate from the defined learning protocol, but attempts to infer private training data from the legitimately received information. In [113], the authors proposed an approach, self-taught Federated Learning, to address the limitations of current methods when handling heterogeneous datasets (e.g. a slow training speed, impractical for real-world applications). It exploited unsupervised feature extraction techniques for Federated Learning with heterogeneous datasets while preserving data privacy. In [114] a method is proposed to identify malicious poisoning attacks by having the server itself bootstrap trust. Specifically, the server collects a small, clean training dataset (called the root dataset) for the learning task and maintains a model (called server model) based on this to bootstrap trust. In each iteration, the server first assigns a trust score to each local model update from the clients, where a local model update has

a lower trust score. They benchmarked their method against CH-MNIST; a medical image classification dataset consisting of 5,000 images of histology tiles collected from colorectal cancer patients. Where privacy is concerned, quantum cryptography is probably the next frontier of the security battleground, and some authors have started developing in this direction while using cancer datasets for benchmarking their secure federated learning frameworks [115]. Figure 1 presents an overall synopsis of all the studies reviewed in this paper based on AI tasks, cancer type, data type and category of work.

4 Conclusion and Discussion

Data decentralization is a crucial setting for developing data-driven models in healthcare due to the sensitive nature of medical data. Federated Learning, while still a new research field, has already demonstrated its potential use to support a distributed learning setup for healthcare. While the general field of Federated Learning research is very active with a focus on improving model aggregation and efficient communication between nodes, model and data privacy is a very challenging and open problem [32]. The data privacy aspect is very important especially in healthcare where legal, ethical and regulatory constraints impose tremendous restrictions and pressure to data providers (e.g., healthcare networks, research institutions)

While the Federated Learning research community is engaged in addressing the aforementioned open problems, in this paper we aimed at presenting the current status of Federated Learning in the domain of cancer and oncology because we believe that the machine learning community in this particular space can benefit from a quick review and perhaps direct research efforts in specific subareas. Our review highlighted that although a lot of works have been developed for Federated Learning only 56% of them have been exclusively proposed for cancer research or clinical oncology. This demonstrates the need for solutions designed specifically within this space. For example, privacy preserving methods may need to be researched and explored under the scope of the cancer field given that privacy requirements and guarantees can be significantly different from other areas (e.g., finance). In a similar fashion, while data heterogeneity is an open challenge in the general machine learning community, cancer and oncology datasets manifest unique properties which may require deeper clinical and medical device expertise involvement when developing methods that aim at overcoming model degredation in largely heterogeneous medical data.

Although there are quite a few frameworks developed specifically for cancer analysis (i.e., 13% Fig. 1), there is the potential risk of a fragmented platform landscape. This is true when it comes to the general Federated Learning community in which a large number of frameworks are currently being developed and maintained. Indeed, such efforts can lead to improved solutions but it is usually collaborative efforts that can achieve better adoption. In the cancer domain data scientists can benefit from platforms that aim at developing tools for distributed annotation, distributed model training workflows, and moreover the adoption of

data standardization and thus better integration of Federated Learning into the clinical workflow.

When it comes to tasks (Fig. 1) we observed that the majority of algorithms are related to classification and segmentation, and use images (either from radiology or pathology) as input data type. This highlights the need for a broader exploration of other important tasks in cancer analysis such as survival prediction, genomics expression, precision medicine, patient treatment planning, and advanced patient diagnosis/prognosis through multi-modal data. Furthermore, within the context of cancer type we identified that almost 70% of the studies were addressing only a specific type of cancer: either brain tumor, or breast cancer, or skin lesions. This reaffirms our previous statement that Federated Learning should expand its application on multiple cancer types. Perhaps the reason for this increased focus on these three specific cancer types comes from the fact that these three areas have been well-established through the release of large public datasets. This emphasizes the overall need for large medical datasets being available to the research community. Ideally, federations that are currently being developed to support distributed learning (e.g., Federated Learning) will provide support in the future for secure remote machine learning development on geographically distributed data providers through robust privacy-preserving layers.

As with any new research field, Federated Learning for healthcare and in particular for cancer and oncology is still in its early days. However, whether the studies were simulating Federated Learning environments or conducting small experiments across hospitals with real private data, they constitute solid basis for future work. Federated Learning infrastructures are continuously being developed specifically for healthcare and cancer research to facilitate true collaboration between healthcare institutions across the world.

Acknowledgments. NP and AK were supported by French State Funds managed by the "Agence Nationale de la Recherche (ANR)" - "Investissements d'Avenir" (Investments for the Future), Grant ANR-10-IAHU-02 (IHU Strasbourg). NP and HK received support from the French ANR National AI Chair program (ANR-20-CHIA-0029-01, Chair AI4ORSafety).

References

1. Deng, J., et al.: ImageNet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE (2009)
2. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. Adv. Neural. Inf. Process. Syst. **25**, 1097–1105 (2012)
3. Szegedy, C., et al.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)
4. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)

5. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, pp. 5998–6008 (2017)
6. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)
7. Naumov, M., et al.: Deep learning recommendation model for personalization and recommendation systems. CoRR abs/1906.00091 (2019). https://arxiv.org/abs/1906.00091
8. Graves, A.: Sequence transduction with recurrent neural networks. arXiv preprint arXiv:1211.3711 (2012)
9. Suzuki, K.: Overview of deep learning in medical imaging. Radiol. Phys. Technol. **10**(3), 257–273 (2017). https://doi.org/10.1007/s12194-017-0406-5
10. Biswas, M., et al.: State-of-the-art review on deep learning in medical imaging. Front. Biosci. (Landmark Ed) **24**, 392–426 (2019)
11. Korfiatis, P., et al.: Residual deep convolutional neural network predicts MGMT methylation status. J. Digit. Imaging **30**, 622–628 (2017)
12. Jumper, J., et al.: Highly accurate protein structure prediction with AlphaFold. Nature **596**, 583–589 (2021)
13. Karargyris, A., et al.: MedPerf: open benchmarking platform for medical artificial intelligence using federated evaluation. arXiv preprint arXiv:2110.01406 (2021)
14. Chang, K., et al.: Distributed deep learning networks among institutions for medical imaging. J. Am. Med. Inform. Assoc. **25**, 945–954 (2018)
15. Johnson, A.E., et al.: MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. Sci. Data **6**, 1–8 (2019)
16. Irvin, J., et al.: CheXpert: a large chest radiograph dataset with uncertainty labels and expert comparison. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 590–597 (2019)
17. Chowdhury, A., Rosenthal, J., Waring, J., Umeton, R.: Applying self-supervised learning to medicine: review of the state of the art and medical implementations. Informatics **8**, 59 (2021)
18. Doersch, C., Zisserman, A.: Multi-task self-supervised visual learning. In: 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2070–2079 (2017)
19. Van Steen, M., Tanenbaum, A.: Distributed systems principles and paradigms. Network **2**, 28 (2002)
20. McMahan, B., Moore, E., Ramage, D., Hampson, S., Agüera y Arcasy, B.: Communication-efficient learning of deep networks from decentralized data. In: Artificial Intelligence and Statistics, pp. 1273–1282. PMLR (2017)
21. Rieke, N., et al.: The future of digital health with federated learning. NPJ Digit. Med. **3**, 1–7 (2020)
22. Trustworthy federated data analytics (TFDA) (2020). https://tfda.hmsp.center/
23. Joint Imaging Platform (JIP) (2020). https://jip.dktk.dkfz.de/jiphomepage/
24. Machine learning ledger orchestration for drug discovery (2020). https://cordis.europa.eu/project/id/831472
25. Healthchain consortium (2020). https://www.substra.ai/en/healthchain-project
26. The federated tumor segmentation (FeTS) initiative (2020). https://www.fets.ai
27. Pati, S., et al.: The federated tumor segmentation (FeTS) challenge. arXiv preprint arXiv:2105.05874 (2021)
28. Medical institutions collaborate to improve mammogram assessment AI (2020). https://blogs.nvidia.com/blog/2020/04/15/federated-learning-mammogram-assessment/

29. Li, T., Sahu, A.K., Talwalkar, A., Smith, V.: Federated Learning: challenges, methods, and future directions. IEEE Sig. Process. Mag. **37**, 50–60 (2020)
30. Yang, Q., et al.: Federated learning. In: Synthesis Lectures on Artificial Intelligence and Machine Learning, vol. 13, pp. 1–207 (2019)
31. Li, L., Fan, Y., Tse, M., Lin, K.-Y.: A review of applications in federated learning. Comput. Ind. Eng. **149**, 106854 (2020)
32. Kairouz, P., et al.: Advances and open problems in federated learning. arXiv preprint arXiv:1912.04977 (2019)
33. Reina, G. A., et al.: OpenFL: an open-source framework for federated learning. arXiv preprint arXiv:2105.06413 (2021)
34. He, C., et al.: Fedml: a research library and benchmark for federated machine learning. arXiv preprint arXiv:2007.13518 (2020)
35. Konečný, J., et al.: Federated learning: strategies for improving communication efficiency. In: NIPS Workshop on Private Multi-Party Machine Learning (2016). https://arxiv.org/abs/1610.05492
36. Yi, L., Zhang, J., Zhang, R., Shi, J., Wang, G., Liu, X.: SU-Net: an efficient encoder-decoder model of federated learning for brain tumor segmentation. In: Farkaš, I., Masulli, P., Wermter, S. (eds.) ICANN 2020. LNCS, vol. 12396, pp. 761–773. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-61609-0_60
37. Sheller, M.J., Reina, G.A., Edwards, B., Martin, J., Bakas, S.: Multi-institutional deep learning modeling without sharing patient data: a feasibility study on brain tumor segmentation. In: Crimi, A., Bakas, S., Kuijf, H., Keyvan, F., Reyes, M., van Walsum, T. (eds.) BrainLes 2018. LNCS, vol. 11383, pp. 92–104. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11723-8_9
38. Sheller, M.J., et al.: Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. Sci. Rep. **10**, 1–12 (2020)
39. Sheller, M., Edwards, B., Reina, G. A., Martin, J., Bakas, S.: NIMG-68. Federated learning in neuro-oncology for multi-institutional collaborations without sharing patient data. Neuro-oncology **21**, vi176 (2019)
40. Mazurowski, M.A., et al.: Radiogenomics of lower-grade glioma: algorithmically-assessed tumor shape is associated with tumor genomic subtypes and patient outcomes in a multi-institutional study with the cancer genome atlas data. J. Neuro-oncol. **133**, 27–35 (2017)
41. Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (BraTS). IEEE Trans. Med. Imaging **34**, 1993–2024 (2014)
42. Cai, X., et al.: A many-objective optimization based federal deep generation model for enhancing data processing capability in IOT. IEEE Trans. Ind. Inform. **99**, 1–1 (2021)
43. Hashmani, M.A., Jameel, S.M., Rizvi, S.S.H., Shukla, S.: An adaptive federated machine learning-based intelligent system for skin disease detection: a step toward an intelligent dermoscopy device. Appl. Sci. **11**, 2145 (2021)
44. Roth, H.R., et al.: Federated learning for breast density classification: a real-world implementation. In: Albarqouni, S., et al. (eds.) DART/DCL -2020. LNCS, vol. 12444, pp. 181–191. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60548-3_18
45. Rooijakkers, T.: Convinced-enabling privacy-preserving survival analyses using multi-party computation (2020)
46. Sarma, K.V., et al.: Federated learning improves site performance in multicenter deep learning without data sharing. J. Am. Med. Inf. Assoc. **28**, 1259–1264 (2021)
47. Deist, T.M., et al.: Distributed learning on 20 000+ lung cancer patients-the personal health train. Radiother. Oncol. **144**, 189–200 (2020)

48. Codella, N. C., et al.: Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC). In: 2018 IEEE 15th International Symposium Biomedical imaging (ISBI 2018), pp. 168–172. IEEE (2018)
49. Yi, Z., Zhang, H., Tan, P., Gong, M.: DualGAN: unsupervised dual learning for image-to-image translation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2849–2857 (2017)
50. Zhang, C., Zhang, J., Chai, D., Chen, K.: Aegis: a trusted, automatic and accurate verification framework for vertical federated learning. arXiv preprint arXiv:2108.06958 (2021)
51. Wang, P., et al.: Automated pancreas segmentation using multi-institutional collaborative deep learning. In: Albarqouni, S., Bakas, S., Kamnitsas, K., Cardoso, M.J., Landman, B., Li, W., Milletari, F., Rieke, N., Roth, H., Xu, D., Xu, Z. (eds.) DART/DCL -2020. LNCS, vol. 12444, pp. 192–200. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60548-3_19
52. McGuigan, A., et al.: Pancreatic cancer: a review of clinical diagnosis, epidemiology, treatment and outcomes. World J. Gastroenterol. **24**, 4846 (2018)
53. Lee, H., et al.: Federated learning for thyroid ultrasound image analysis to protect personal information: validation study in a real health care environment. JMIR Med. Inform. **9**, e25869 (2021)
54. Choudhury, A., et al.: Predicting outcomes in anal cancer patients using multicentre data and distributed learning-a proof-of-concept study. Radiother. Oncol. **159**, 183–189 (2021)
55. Zolotareva, O., et al.: Flimma: a federated and privacy-preserving tool for differential gene expression analysis. arXiv preprint arXiv:2010.16403 (2020)
56. Andreux, M., Manoel, A., Menuet, R., Saillard, C., Simpson, C.: Federated survival analysis with discrete-time Cox models. arXiv preprint arXiv:2006.08997 (2020)
57. Cox, D.R.: Regression models and life-tables. J. R. Stat. Soc. Ser. B (Methodological) **34**, 187–220 (1972). http://www.jstor.org/stable/2985181
58. Jiménez-Sánchez, A., Tardy, M., Ballester, M.A.G., Mateus, D., Piella, G.: Memory-aware curriculum federated learning for breast cancer classification. arXiv preprint arXiv:2107.02504 (2021)
59. Xia, Y., et al.: Auto-FedAvg: learnable federated averaging for multi-institutional medical image segmentation. arXiv preprint arXiv:2104.10195 (2021)
60. Shen, C., et al.: Multi-task federated learning for heterogeneous pancreas segmentation. arXiv preprint arXiv:2108.08537 (2021)
61. Guo, P., Wang, P., Zhou, J., Jiang, S., Patel, V.M.: Multi-institutional collaborations for improving deep learning-based magnetic resonance image reconstruction using federated learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2423–2432 (2021)
62. Bercea, C. I., Wiestler, B., Rueckert, D., Albarqouni, S.: FedDis: disentangled federated learning for unsupervised brain pathology segmentation. arXiv preprint arXiv:2103.03705 (2021)
63. Andreux, M., du Terrail, J.O., Beguier, C., Tramel, E.W.: Siloed federated learning for multi-centric histopathology datasets. In: Albarqouni, S., et al. (eds.) DART/DCL -2020. LNCS, vol. 12444, pp. 129–139. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60548-3_13
64. Ke, J., Shen, Y., Lu, Y.: Style normalization in histology with federated learning. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), pp. 953–956. IEEE (2021)

65. Roth, H.R., et al.: Federated whole prostate segmentation in MRI with personalized neural architectures. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12903, pp. 357–366. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87199-4_34
66. Liu, Q., Yang, H., Dou, Q., Heng, P.-A.: Federated semi-supervised medical image classification via inter-client relation matching. arXiv preprint arXiv:2106.08600 (2021)
67. Bdair, T., Navab, N., Albarqouni, S.: Peer learning for skin lesion classification. arXiv preprint arXiv:2103.03703 (2021)
68. Seo, H., Park, J., Oh, S., Bennis, M., Kim, S.-L.: Federated knowledge distillation. arXiv preprint arXiv:2011.02367 (2020)
69. Li, T., et al.: Federated optimization in heterogeneous networks. arXiv preprint arXiv:1812.06127 (2018)
70. Tschandl, P., Rosendahl, C., Kittler, H.: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci. Data **5**, 1–9 (2018)
71. Blake, C.: UCI repository of machine learning databases (1998). http://www.ics.uci.edu/~mlearn/MLRepository.html
72. Chang, Q., et al.: Synthetic learning: learn from distributed asynchronized discriminator GAN without sharing medical image data. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 13856–13866 (2020)
73. Zhang, M., Qu, L., Singh, P., Kalpathy-Cramer, J., Rubin, D. L.: SplitAVG: a heterogeneity-aware federated deep learning method for medical imaging. arXiv preprint arXiv:2107.02375 (2021)
74. Yeganeh, Y., Farshad, A., Navab, N., Albarqouni, S.: Inverse distance aggregation for federated learning with Non-IID data. In: Albarqouni, S., et al. (eds.) DART/DCL -2020. LNCS, vol. 12444, pp. 150–159. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60548-3_15
75. Nasirigerdeh, R., et al.: Federated multi-mini-batch: an efficient training approach to federated learning in Non-IID environments. arXiv preprint arXiv:2011.07006 (2020)
76. Salmeron, J.L., Arévalo, I.: A privacy-preserving, distributed and cooperative FCM-based learning approach for cancer research. In: Bello, R., Miao, D., Falcon, R., Nakata, M., Rosete, A., Ciucci, D. (eds.) IJCRS 2020. LNCS (LNAI), vol. 12179, pp. 477–487. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-52705-1_35
77. Kosko, B.: Fuzzy cognitive maps. Int. J. Man Mach. Stud. **24**, 65–75 (1986)
78. Liu, Y., Wu, W., Flokas, L., Wang, J., Wu, E.: Enabling SQL-based training data debugging for federated learning. arXiv preprint arXiv:2108.11884 (2021)
79. Kassab, R., Simeone, O.: Federated generalized Bayesian learning via distributed stein variational gradient descent. arXiv preprint arXiv:2009.06419 (2020)
80. Diao, E., Ding, J., Tarokh, V.: Gradient assisted learning. arXiv preprint arXiv:2106.01425 (2021)
81. Fang, M., Cao, X., Jia, J., Gong, N.: Local model poisoning attacks to byzantine-robust federated learning. In: 29th Security Symposium Security 2020, pp. 1605–1622 (2020)
82. Kasturi, A., Ellore, A.R., Hota, C.: Fusion learning: a one shot federated learning. In: Krzhizhanovskaya, V.V., Závodszky, G., Lees, M.H., Dongarra, J.J., Sloot, P.M.A., Brissos, S., Teixeira, J. (eds.) ICCS 2020. LNCS, vol. 12139, pp. 424–436. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-50420-5_31

83. Wang, S., Suwandi, R. C., Chang, T.-H.: Demystifying model averaging for communication-efficient federated matrix factorization. In: ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3680–3684. IEEE (2021)
84. Wang, S., Chang, T.-H.: Federated matrix factorization: algorithm design and application to data clustering. arXiv preprint arXiv:2002.04930 (2020)
85. Roy, A. G., Siddiqui, S., Pölsterl, S., Navab, N., Wachinger, C.: BrainTorrent: a peer-to-peer environment for decentralized federated learning. arXiv preprint arXiv:1905.06731 (2019)
86. Geleijnse, G., et al.: Prognostic factors analysis for oral cavity cancer survival in the Netherlands and Taiwan using a privacy-preserving federated infrastructure. Sci. Rep. **10**, 1–9 (2020)
87. Carpov, S., Gama, N., Georgieva, M., Jetchev, D.: GenoPPML-a framework for genomic privacy-preserving machine learning. Cryptol. ePrint Arch. (2021)
88. Elayan, H., Aloqaily, M., Guizani, M.: Deep federated learning for IOT-based decentralized healthcare systems. In: 2021 International Wireless Communications and Mobile Computing (IWCMC), pp. 105–109. IEEE (2021)
89. Scherer, J., et al.: Joint imaging platform for federated clinical data analytics. JCO Clin. Cancer Inform. **4**, 1027–1038 (2020)
90. Matschinske, J., et al.: The featurecloud AI store for federated learning in biomedicine and beyond. arXiv preprint arXiv:2105.05734 (2021)
91. Arambakam, M., Beel, J.: Federated meta-learning: democratizing algorithm selection across disciplines and software libraries. In: 7th ICML Workshop on Automated Machine Learning (AutoML) (2020)
92. Chen, X., Wang, X., Yang, K.: Asynchronous blockchain-based privacy-preserving training framework for disease diagnosis. In: 2019 IEEE International Conference on Big Data (Big Data), pp. 5469–5473. IEEE (2019)
93. Zhu, L., Han, S.: Deep leakage from gradients. In: Yang, Q., Fan, L., Yu, H. (eds.) Federated Learning. LNCS (LNAI), vol. 12500, pp. 17–31. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-63076-8_2
94. Wei, W., et al.: A framework for evaluating gradient leakage attacks in federated learning. arXiv preprint arXiv:2004.10397 (2020)
95. Połap, D., Woźniak, M.: Meta-heuristic as manager in federated learning approaches for image processing purposes. Appl. Soft Comput. **113**, 107872 (2021)
96. Beguier, C., Terrail, J. O. d., Meah, I., Andreux, M., Tramel, E. W.: Differentially private federated learning for cancer prediction. arXiv preprint arXiv:2101.02997 (2021)
97. Shah, M.M., et al.: Distributed machine learning on differentially private skin cancer data. Solid State Technol. **63**, 1777–1786 (2020)
98. Zhang, C., et al.: Feasibility of privacy-preserving federated deep learning on medical images. Int. J. Radiat. Oncol. Biol. Phys. **108**, e778 (2020)
99. Li, W., et al.: Privacy-preserving federated brain tumour segmentation. In: Suk, H.-I., Liu, M., Yan, P., Lian, C. (eds.) MLMI 2019. LNCS, vol. 11861, pp. 133–141. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32692-0_16
100. Fay, D., Sjölund, J., Oechtering, T. J.: Decentralized differentially private segmentation with pate. arXiv preprint arXiv:2004.06567 (2020)
101. Papernot, N., Abadi, M., Erlingsson, U., Goodfellow, I., Talwar, K.: Semi-supervised knowledge transfer for deep learning from private training data. arXiv preprint arXiv:1610.05755 (2016)
102. Lu, M. Y., et al.: Federated learning for computational pathology on gigapixel whole slide images. arXiv preprint arXiv:2009.10190 (2020)

103. Cheng, W., et al.: A privacy-protection model for patients. Secur. Commun. Networks **2020**, 12 (2020)
104. Froelicher, D., et al.: Truly privacy-preserving federated analytics for precision medicine with multiparty homomorphic encryption. bioRxiv (2021)
105. Sadat, M.N., et al.: SAFETY: secure gwAs in federated environment through a hYbrid solution. IEEE/ACM Trans. Comput. Biol. Bioinform. **16**, 93–102 (2018)
106. Rajotte, J.-F., et al.: Reducing bias and increasing utility by federated generative modeling of medical images using a centralized adversary. arXiv preprint arXiv:2101.07235 (2021)
107. Mukherjee, S., Xu, Y., Trivedi, A., Patowary, N., Ferres, J.L.: PrivGAN: protecting GANs from membership inference attacks at low cost to utility. Proc. Priv. Enhancing Technol. **2021**, 142–163 (2021)
108. Ziller, A., et al.: Differentially private federated deep learning for multi-site medical image segmentation. arXiv preprint arXiv:2107.02586 (2021)
109. Tian, Z., Zhang, R., Hou, X., Liu, J., Ren, K.: FederBoost: private federated learning for GBDT. arXiv preprint arXiv:2011.02796 (2020)
110. Wei, W., Liu, L., Wut, Y., Su, G., Iyengar, A.: Gradient-leakage resilient federated learning. In: 2021 IEEE 41st International Conference on Distributed Computing Systems (ICDCS), Washington DC, pp. 797–807 (2021)
111. Yang, Z., Hu, S., Chen, K.: FPGA-based hardware accelerator of homomorphic encryption for efficient federated learning. arXiv preprint arXiv:2007.10560 (2020)
112. Weng, H., et al.: Privacy leakage of real-world vertical federated learning. arXiv preprint arXiv:2011.09290 (2020)
113. Chu, K.-F., Zhang, L.: Privacy-preserving self-taught federated learning for heterogeneous data. arXiv preprint arXiv:2102.05883 (2021)
114. Cao, X., Fang, M., Liu, J., Gong, N. Z.: FLTrust: byzantine-robust federated learning via trust bootstrapping. arXiv preprint arXiv:2012.13995 (2020)
115. Li, W., Lu, S., Deng, D.-L.: Quantum federated learning through blind quantum computing. Sci. China Phys. Mech. Astron. **64**(10), 1–8 (2021). https://doi.org/10.1007/s11433-021-1753-3