

Federated Learning in Medical Image Analysis: A Systematic Survey

Fabiana Rodrigues da Silva ¹, Rui Camacho ²  and João Manuel R. S. Tavares ^{3,*} 

¹ Faculdade de Engenharia, Universidade do Porto, R. Dr. Roberto Frias, 4200-465 Porto, Portugal; up202100810@edu.fe.up.pt

² Departamento de Engenharia Informática, Faculdade de Engenharia, Universidade do Porto, R. Dr. Roberto Frias, 4200-465 Porto, Portugal; rcamacho@fe.up.pt

³ Instituto de Ciência e Inovação em Engenharia Mecânica e Engenharia Industrial, Departamento de Engenharia Mecânica, Faculdade de Engenharia, Universidade do Porto, R. Dr. Roberto Frias, 4200-465 Porto, Portugal

* Correspondence: tavares@fe.up.pt

Abstract: Medical image analysis is crucial for the efficient diagnosis of many diseases. Typically, hospitals maintain vast repositories of images, which can be leveraged for various purposes, including research. However, access to such image collections is largely restricted to safeguard the privacy of the individuals whose images are being stored, as data protection concerns come into play. Recently, the development of solutions for Automated Medical Image Analysis has gained significant attention, with Deep Learning being one solution that has achieved remarkable results in this area. One promising approach for medical image analysis is Federated Learning (FL), which enables the use of a set of physically distributed data repositories, usually known as nodes, satisfying the restriction that the data do not leave the repository. Under these conditions, FL can build high-quality, accurate deep-learning models using a lot of available data wherever it is. Therefore, FL can help researchers and clinicians diagnose diseases and support medical decisions more efficiently and robustly. This article provides a systematic survey of FL in medical image analysis, specifically based on Magnetic Resonance Imaging, Computed Tomography, X-radiography, and histology images. Hence, it discusses applications, contributions, limitations, and challenges and is, therefore, suitable for those who want to understand how FL can contribute to the medical imaging domain.



Citation: da Silva, F.R.; Camacho, R.; Tavares, J.M.R.S. Federated Learning in Medical Image Analysis:

A Systematic Survey. *Electronics* **2024**, *13*, 47. <https://doi.org/10.3390/electronics13010047>

Academic Editor: Jian Sun

Received: 29 July 2023

Revised: 17 December 2023

Accepted: 19 December 2023

Published: 21 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: image processing; artificial intelligence; machine learning; federated learning; medical image analysis

1. Introduction

Presently, produced data, both in industry and public research centers, is huge and growing daily. Some data may be of a sensitive or personal nature and must be, therefore, protected from public access. This scenario is no different in the medical environment, where Health Information Systems and Electronic Health Records can collect and store patient data on a large scale.

Analyzing medical images is a daunting task for healthcare professionals due to the heavy workload, the need to analyze many images, the intricate nature of cases, and the resemblance to other diseases, which often require additional studies within a constrained time frame. Despite the extensive training that healthcare experts undergo, they are prone to commit errors given the absence of information or insufficient time to analyze the images under study, which may impact the identification of the right treatment to prevent the growth and spread of the disease under analysis.

Different types of image modalities, such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), X-ray, or histology images are required to establish a correct

diagnosis. Therefore, a comprehensive study of the use of those modalities across several image repositories is crucial to developing accurate image analysis software.

Artificial Intelligence (AI) is in a development stage that can use medical images to successfully detect and diagnose pathological conditions successfully [1]. In this way, hospitals use technology by building models to discover new patterns of diseases and treatments from a given image dataset, mainly through data mining. However, institutions should be careful about the data privacy of their patients since health data are sensitive and legally protected; therefore, personal data cannot be exposed outside them [2].

Building very high-quality AI-based models, mainly deep machine-based models, for image analysis requires diversity and many images, i.e., examples used for training them, especially for complex problems. One way to access such a large number of medical images is to directly access or obtain the images in a set of hospital image repositories. However, transferring large amounts of data between the client and the central server can bring high communication costs, and the data can leak and be an attractive target for cyberattacks. In addition, centralized models may exhibit biases and generalization issues when the training dataset lacks diversity. Nevertheless, several problems may then occur. For example, acquiring the necessary permissions to access the images may entail lengthy bureaucratic procedures. Moreover, the images must be duly anonymized, and restrictions may prohibit removal from the original hospital repository.

Centralized data storage can pose difficulties in meeting legal and regulatory standards, particularly in regions with stringent privacy laws. Privacy concerns become pronounced in a centralized repository, as it becomes susceptible to vulnerabilities and unauthorized access, such as the lack of user control over personal information where individuals may use data without owners' consent or even expose them without permission. Consequently, the integrity of data security may be compromised.

To prioritize the privacy of patients' data and legal issues related to ethics, an approach called Federated learning (FL), which enables building better machine-learning models focusing on privacy-preserving [3–5], has emerged.

This innovative approach harnesses the potential of artificial intelligence and machine learning for analyzing medical images, aiming to amplify healthcare systems' diagnostic and predictive capabilities, all while safeguarding patient privacy and ensuring data security. Besides privacy preservation, FL also addresses collaboration in environments where data sensitivity and compliance are critical concerns, as demonstrated in the contributions outlined in [6] where the presented FL approach achieved privacy-protected tumor classification with high accuracy. In FL, the data remains at its source and does not need to be transferred to a central server, which mitigates the likelihood of data breaches. Additionally, FL facilitates real-time model updates, allowing models to constantly learn and adjust to changing data, rendering it well-suited for dynamic environments. Training models that can tackle data heterogeneity are essential to obtain a model in an FL environment that exhibits high performance across all devices, as is outlined in [7], where a method to overcome performance drop due to data heterogeneity and achieve high accuracy in federated learning is proposed.

Therefore, FL has been helping healthcare professionals extract meaningful information from images by applying machine-learning models, such as detecting a disease based on medical imaging analysis, resulting in a quicker time for disease identification and maintaining a knowledge base to be applied in other new cases. Moreover, healthcare delivery can be improved by automating imaging-based procedures in hospitals where access to related experts is limited.

The main purpose of this article is to present the current state-of-the-art related to medical image analysis with a special focus on Federated Learning as an adequate approach for building privacy-preserving competent machine-based models. This article is divided into the following sections: Section 2 presents an overview of FL; the adopted articles searching method is described in Section 3; frameworks that use FL in collaborative research

are discussed in Section 4; and to finalize, Section 5 presents the discussion and the main conclusions of this survey.

2. Federated Learning

Federated learning is a designation commonly used for collaborative machine learning without centralized training data, introduced in a study published by Google in 2017 [8]. Thus, data are distributed in different sites, locations, and devices. Therefore, in an FL framework, the process of training data happens locally, where the data are located, which is called decentralized learning. Hence, a device, i.e., an edge device, trains its model and stores its data locally. Then, the server aggregates the result of updated models from each device and updates the centralized, i.e., global, model. Multiple trained models enrich the global model once each local device, based on its individually trained model, provides feedback to the server, which maintains a global shared model and disseminates it to all institutions.

The data exchanged is encrypted to ensure that no other devices access private information. With that, personal information is never sent to a central server, and just the weights, biases, and other parameters are learned by the local model trained in each device. When new hospitals enter this distributed learning environment, they bring more data and computational resources to the consortium. Nonetheless, some hospitals may generate more images than others and, in some cases, with minimal image heterogeneity. Consequently, the consortium may possess a substantial dataset in terms of capacity; still, the models will not acquire fresh features because of deficient image variety, which causes models to produce inaccurate data. On the other hand, hospitals that possess a substantial quantity and diversity of images will contribute positively to developing better learning models and consequently producing more dependable results.

Figures 1 and 2 depict how federated learning works.

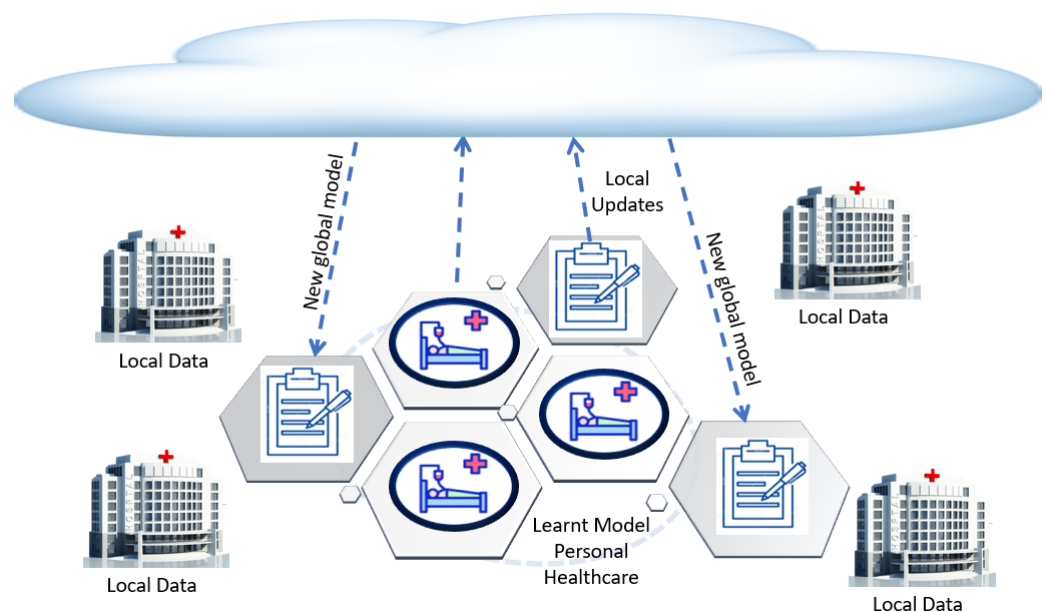


Figure 1. Example of a Federated learning approach.

Federated Machine Learning Process

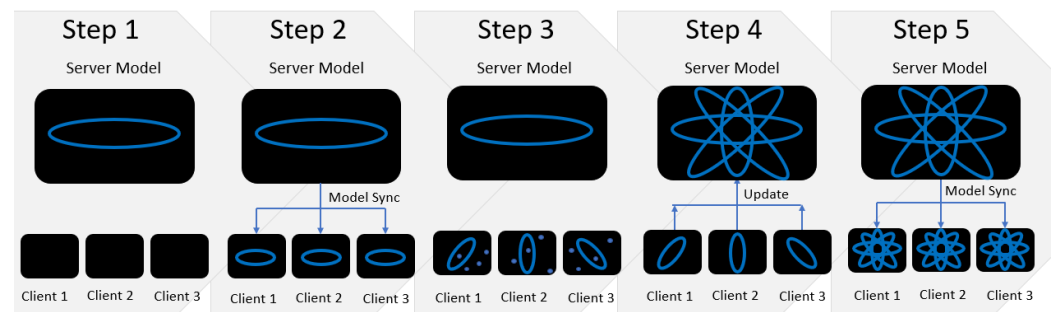


Figure 2. Steps of a common Federated learning approach.

Figure 2 shows that, in the first step, the server model selects the data sources and the machine-learning model to be used. The diverse data are stored locally on each client and are collected from those different sources. The second step involves sending an initial version of the machine-learning model to each data source, i.e., to the edge devices or clients, so the server and edge device models are synchronized. In the third step of the approach, the edge devices use their respective datasets to train the model locally. The locally trained models have similar parameters (represented by ellipses in Figure 2) but differ in their weights (represented by dots in Figure 2) due to their locally collected dataset. In step four, each edge device sends the updated weights of its model back to the central server. The central server aggregates the received information by averaging the updated model weights from each data source to build a new version of its model. In the fifth step, the updated server model is synchronized with the edge devices without accessing any data; then, an individual model is updated and ready to be evaluated on new and unseen data.

In FL, hospitals collaborate to train a shared machine-learning model, which is distributed to each hospital, where it learns from the local data while preserving data security. As a result, hospitals can achieve significant savings in resources and costs associated with data transfer and centralized storage, ultimately enhancing the efficiency of healthcare operations.

As aforementioned, federated learning is valuable when the data are sensitive or difficult to share due to regulatory or privacy concerns. Therefore, FL is being applied in diverse industries, especially in the health field, where the collaborative analysis of sensitive data are highly demanded.

3. Searching Method

This systematic literature review was carried out in the SCOPUS database. To find articles, the following queries were combined using the AND logical operator: (a) “federated learning” OR “federated machine learning”; (b) “medical image” OR “medical imaging”. Table 1 presents the used query and the total number of gathered articles.

Table 1. Total number of articles retrieved from the used electronic repository.

Repository	Query Performed	No. of Gathered Articles
Scopus	TITLE-ABS-KEY ((federated learning) OR (federated machine learning) AND (medical image) OR (medical imaging)) AND (LIMIT-TO (DOCTYPE, “ar”))	84

The current research topic is relatively new, and even though the conducted search was relative to the last 10 years, no articles published before 2019 were found. However,

from the gathered articles, the growing number of studies being conducted on the topic in recent years is noticeable; see Figure 3.

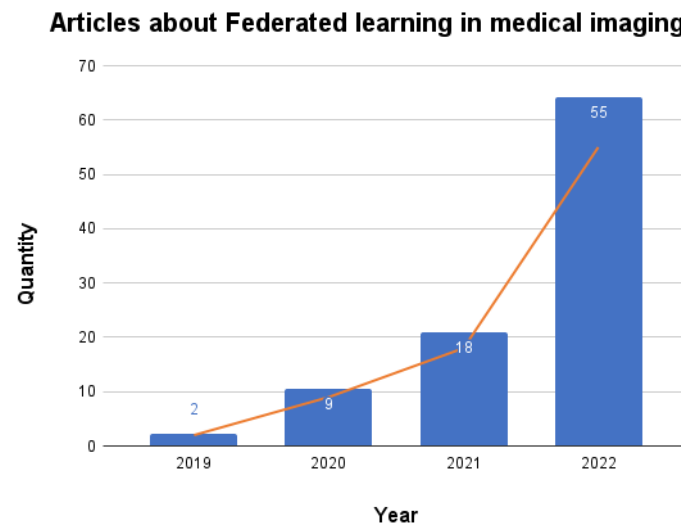


Figure 3. Articles published in the last ten years that were retrieved from the SCOPUS database using the keywords “federated learning”, “federated machine learning”, “medical image”, and “medical imaging”.

Based on an analysis of the titles and abstracts of the gathered articles, a total of 84 articles were initially selected. Then, 42 articles were discarded for not meeting the criteria: medical images acquired by MRI, CT imaging, X-ray imaging, or histology imaging, bringing the total to 42 articles. Then, 20 articles were excluded once they were focused on topics unrelated to this study. In the end, 22 original articles were selected for the review. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram [9] shown in Figure 4 represents the performed systematic search. Using the same criteria, a second search was performed in the PubMed database, but no additional articles were retrieved.

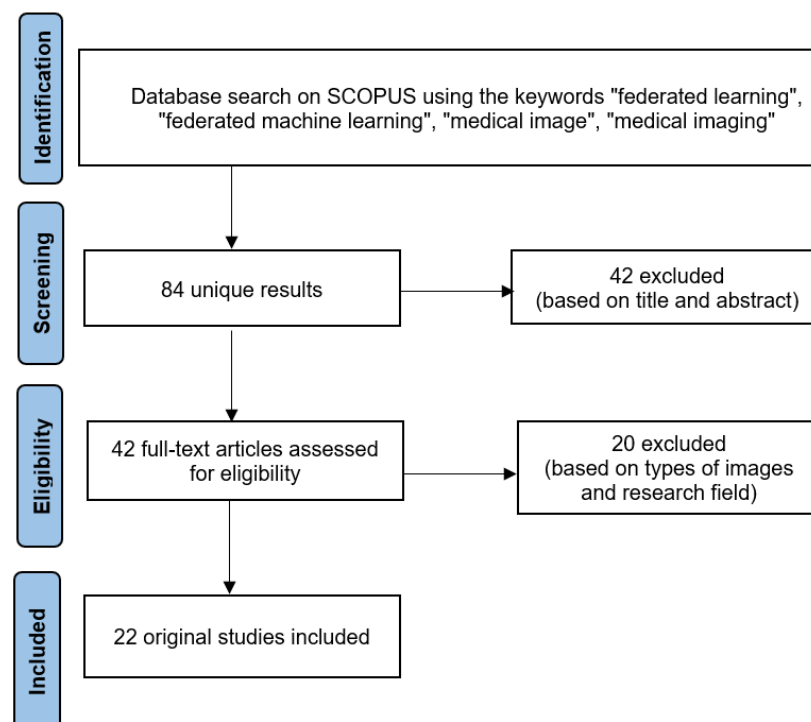


Figure 4. PRISMA diagram of the performed systematic literature review.

Figure 5 indicates the types of medical images studied in the selected articles per year, which allows the identification of the imaging modalities that FL frameworks have addressed. In 2022, for example, MRI led the number of works in this field, with nine articles, followed by X-ray, with four articles, one article involving histology images, and, lastly, CT, with one article.

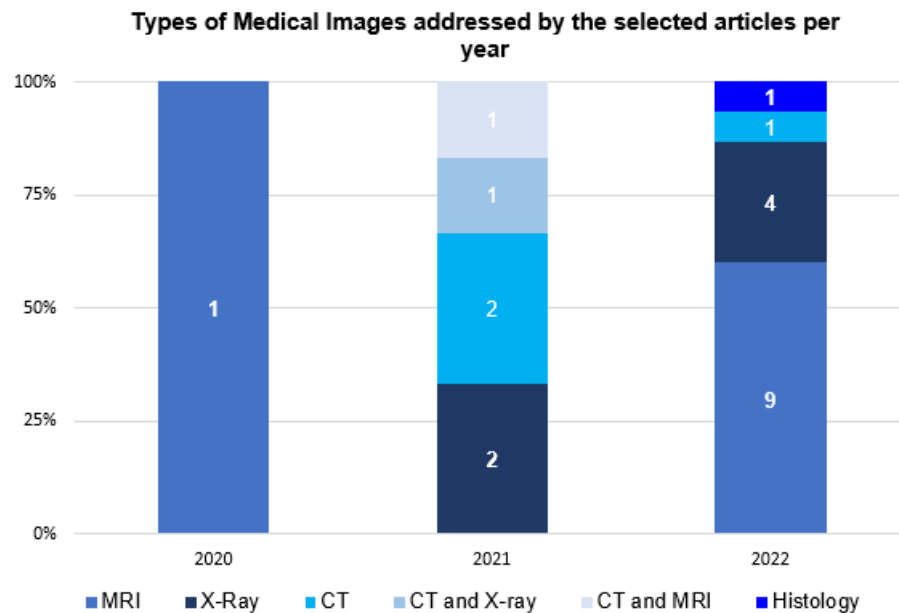


Figure 5. Types of imaging modality addressed by the selected articles per year.

A summary of the selected and studied articles, as to the used models, datasets, architectures, input dimension and imaging modality, research goal, main purpose, contributions and benefits, limitations, and reported accuracy, is presented in Tables 2–13 according to the used imaging modality(ies): Tables 2–5 are related to Magnetic Resonance (MR) images and include ten articles. Tables 6 and 7 detail three articles related to CT images. Table 8 is related to CT and MR images with one article. Table 9 includes one article addressing CT and X-ray images. Tables 10–12 present six articles related to X-ray images. Lastly, Table 13 presents one article related to histology images.

Table 2. Summary of the studied articles related to MR images (1 of 4). Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s), and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[10]	BT-small-2c and BT-large-3c	CNN End-to-end	2D, MRI	Binary classification to diagnose a brain tumor in a federated environment.	Designing the CNN model for each client and exchanging the training parameters between clients and server.	FL framework to detect brain cancers using 3 (2 public) datasets.	Proposed model is limited to 4 different strategies.	Higher accuracy on the larger dataset (BT-large-3c), 96%, than to the smallest one (BT-small-2c), 82%.
[11]	M&M and ACDC 2018	3D-CNN	3D, MRI	Diagnosis of hypertrophic cardiomyopathy, which is a cardiovascular disease.	Test CNN models on partitions of the centers seen during the training and also on unseen centers.	First simulated federated learning study on the modality of cardiovascular MRI.	Small dataset (M&M), several data augmentation techniques were needed to increase the size of the training set artificially.	-
[12]	SpineSagT2-Wdataset3, which comes from the Chinese Society of Biomedical Engineering	DAGs-U-Net	MRI	Segmentation of Vertebral Body images.	Create a Federated Learning-based Vertebral Body Segment Framework (FLVBSF) with a novel local Dual Attention Gates (DAGs)-based attention mechanism.	FLVBSF can strongly classify each vertebral body pixel from the background.	Several approaches applied on the U-Net. When the number of iterations or institutions is set to smaller or larger, the performance of the federated model varies.	The U-Net with DAGs achieved a Pixel-level Accuracy (PA) of 98.29%.

Table 3. Summary of the studied articles related to MR images (2 of 4). Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s), and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[7]	ACDC MICCAI 2017	Momentum Contrast (MoCo)	3D, MRI	Segmentation and generalization performance on a cardiac MRI dataset.	Two federated self-supervised learning frameworks for volumetric medical image segmentation with limited annotations.	The first framework has high accuracy and fits high-performance servers with high-speed connections. The second framework addresses lower communication costs, applicable to mobile devices.	The proposed optimized method (FCLOpt) does not rely on negative samples, which reduces the communication cost of contrastive learning.	First framework, FCL, has high accuracy ($82.4 \pm 2.5\%$) with feature sharing; the second framework, FCLOpt, has an accuracy of $82.1 \pm 2\%$.
[13]	BraTS 2017, Retina and BoneAge	ResNet34 pre-trained on ImageNet as the base network for all methods on Retina and BoneAge datasets. Used pre-trained U-Net for the BraTS dataset.	MRI	Implement SplitAVG method to drop the performance from 3 heterogeneous datasets.	Method SplitAVG applied on a CNN to overcome the performance drop from data heterogeneity in federated learning.	Used heterogeneous data in real-world federated learning settings. The SpliAVG method does not need any complex hyper-parameter tuning, training heuristics, or additional training/fine-tuning.	Data privacy concerns reconstructing raw images from shared feature maps of the cut layer.	Achieved 96.2% of accuracy.
[14]	MSLUB, MISBI, MSKRI, GBKRI, BraTS	Convolutional auto-encoders	MRI	Segmentation of an unsupervised brain pathology.	Train an unsupervised CNN using Federated Disentangled representation learning, called FedDis.	Apply healthy reconstructions by increasing the anomaly scores, leveraging the global anatomical structure, and detaching the parameters while mitigating domain shifts.	Trains the model locally it is aggregated only the updated models of the local institutions.	The anomaly segmentation has the results improved by 99.74% for multiple sclerosis and 40.45% for tumors over locally trained models.

Table 4. Summary of the studied articles related to MR images (3 of 4). Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s), and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[6]	ABIDE, ADHD-200 and COBRE	STM, SBM, and MMoE	MRI	Identify multiple related mental disorders.	Compare different architectures to help medical professionals set up early detection and personalized treatment in FL.	The FL framework has effective learning for those participating institutions with relatively small datasets.	Regardless of the increasing available datasets, most client models fail to converge during the training process, which leads to significant performance degradation.	Accuracy of $69.48 \pm 1.6\%$ in autism spectrum disorder, $71.44 \pm 3.2\%$ in attention deficit/hyperactivity disorder, and $83.29 \pm 3.2\%$ in schizophrenia.
[15]	UK Data Service dataset	CNN Ensemble	Axial T2 and Coronal slices of MR images.	Brain tumor identification.	Used 6 pre-trained models, combined 3 best models based on the accuracy on an average CNN (InceptionV3, VGG16, DenseNet) for Voting Ensemble, to compare the performance to select the global model in FL.	FL approach achieved privacy-protected tumor classification with high accuracy.	Small dataset, which degrades the global performance if many clients have really small datasets and may heavily overfit. Within the clients, no measure was taken to address class distribution imbalance.	FL with an accuracy of 91.05% compared to 96.68% obtained by the base ensemble model.
[16]	ABIDE: Autism brain imaging data exchange ADNI: Alzheimer’s disease neuroimaging initiative	GCNs	MRI	Predict neural diseases such as Autism and Alzheimer’s.	Train a global Graph Convolutional Neural Networks (GCNs) node classifier in many institutions using a federated graph learning platform called FedNI.	Designed a federated network where a missing node generator grants each institution the ability to generate missing nodes and edges and federated network inpainting. FedAvg was used as an aggregation strategy.	Training an effective GCN model for node classification requires a bigger dataset.	Accuracy of 66.7% ($\pm 0.6\%$) on the ABIDE dataset, and 75.8% ($\pm 0.7\%$) on the ADNI dataset.

Table 5. Summary of the studied articles related to MR images (4 of 4). Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s) and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[17]	ABIDE	Multi-layer perceptron (MLP) 6105-16-2:6105 nodes for the input (first) layer, 16 nodes for the hidden layer, and 2 nodes for the output layer)	MRI	Identify autism spectrum disorders (ASD) or healthy control (HC).	Applied two methods: Mixture of Experts (MoE), adaptation near the output layer, and Adversarial domain alignment, adaptation on the data knowledge representation level.	Overcome the domain shift issue; the federated learning model has revealed possible brain biomarkers for identifying ASD. Fed-MoE outperformed Fed on NYU, UM, and UCLA sites, and Fed-Align outperformed Fed on NYU, UM, and UCLA sites in terms of mean accuracy. All Fed and Fed+Domain Adaptation strategies significantly improved compared to Cross, Single, and Ensemble strategies.	Implications into other disease areas, particularly rare diseases with few patients.	Results of using different training strategies, but the best accuracy was 78.9% ($\pm 15.3\%$) when the Fed framework was used on ASD.

Table 6. Summary of the studied articles related to CT images (1 of 2). Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s), and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[18]	Private large dataset with unsegmented 2230 axial chest CT images from various institutions, divided into (a) COVID-19 (1016 images), (b) lung cancer, and non-COVID-19 lung infections (610 images), and (c) normal lung aspect (604 images)	CNN	CT lung images.	Multiclass classification of COVID-19, cancer, and non-COVID-19 lung infections or normal lung.	Used VGG-16 pre-trained on the ImageNet dataset with transfer learning and the FedAvg method. Three individual clients were deployed independently instead of training the model on the entire dataset in a centralized way.	The proposed AI-assisted software can identify a healthy lung without COVID-19 to disregard lung cancer or a non-COVID-9 lung infection.	High computation cost, being possible to train and validate only 2 clients from three rounds of the total of ten chosen for training.	The proposed method—FL VGG-16 had an accuracy of 83.82% on images during training, and 79.32% during validation.
[19]	3 private datasets from local hospitals from Germany, China, and one publicly available dataset	2D CNN based on an improved version of RetinaNet18	CT	Detecting lung abnormalities in COVID-19 patients.	Used DeepLesion CT lesion dataset as a pre-trained model, applying three independent networks: (1) Learning with limited annotations, (2) Learning to segment COVID-19 CT scans from non-COVID-19 CT scans, and (3) Learning with both COVID-19 and non-COVID-19 CT scans. Then, a model ensemble of these 3 individual models will be used, and lastly, a baseline of training a single joint model with all data centralized.	To support clinical disease management, using diverse and large datasets led to a good impact of AI providing low-cost and scalable tools for lesion burden estimation.	Limitation of concept shift, where the manual annotations in the cohort were not directly compatible with training data.	-

Table 7. Summary of the studied articles related to CT images (2 of 2). Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s), and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[20]	Private datasets from hospitals in China, Italy, and Japan	CNN	CT	Detect COVID-19 infections.	Semi-supervised FL learning for segmentation of abnormal regions related to COVID-19 using 3D U-shape FCN as a baseline model.	Semi-supervision potentially reduces the annotation burden under a distributed setting. The proposed approach performed better in real-world datasets compared to the default setting of federated learning.	Conventional data sharing instead of model weight sharing. The model is not trained to discriminate against other types of abnormalities, e.g., pneumonia or cancer. Semi-supervised framework complexity is higher compared to regular FL.	-

Table 8. Summary of studied article related to CT and MR images. Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s), and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[21]	Two private datasets from Medical Segmentation Decathlon (MSD): pancreas and brain tumor segmentation.	CNN (MoNet, a shallow U-Netlike architecture)	CT and MRI	Pancreatic segmentation in CT and brain tumor segmentation in MRI.	Proposed the MoNet architecture for federated learning applications to segment pancreatic in CT and brain tumors in MRI.	MoNet has a small number of parameters and less computational cost compared to U-Net-16, and extracts more robust features that generalize better to out-of-sample data.	Not considered the use of larger, multi-institutional training and validation sets, nor considered the use of other larger models as well.	-

Table 9. Summary of studied article related to CT and X-ray images. Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s), and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[22]	LIDC dataset and 2 other datasets from the National Institutes of Health	CNNs (GhostNet, ResNet50, and ResNet101)	CT and X-ray	COVID region segmentation in chest CT images.	Proposed an architecture for a dynamic fusion-based learning approach for semi-supervised medical image analysis to detect COVID-19 infections to improve communication efficiency and model performance.	Conducted 18 groups of experiments using three different models: GhostNet, ResNet50, and ResNet101. Used aggregated global model. The fusion-based federated performs better in real-world data sets than the default setting of federated learning, which learning can ensure fault tolerance and robustness.	The study does not consider the communication efficiency and model accuracy issues of federated learning.	-

Table 10. Summary of studied articles related to X-ray images (1 of 3). Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s), and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[23]	COVIDX-8A, COVIDX-8B	CNNs (ResNet18 and ResNet34)	Chest X-ray	Classification of COVID on chest X-ray images to ensure additional security and privacy features.	Propose MediSecFed: a secure framework for federated learning in a difficult environment applied on chest X-ray datasets. Developed Knowledge Distillation (KD), a new technique for compacting and accelerating neural networks, to maintain comparable performance while keeping the constraints in mind.	To generate the synthetic data used for KD, a publicly available frozen DenseNet-121 model trained on CheXpert for chest X-ray image generation was used. The method does not allow the server to perform model inversion attacks since the model parameters are never shared.	The method outperforms FedAvg by 15% on both datasets in a difficult environment.	-

Table 10. Cont.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[24]	CheXpert and Mendeley	CNNs (Dense-Net121 and ResNet50)	Chest X-ray	Federated learning for chest X-ray pneumonia classification to prevent data privacy attacks.	Propose an FL as a solution for privacy-preserving distributed learning by integrating Rényi differential privacy with a Gaussian noise mechanism into the federated learning process.	Used 2 architectures pre-trained on ImageNet data, showing that the classification produced better results on Mendeley than on CheXpert data. Images reconstruction from shared model updates within the FL setting from networks using the DLG attack. Used FedAvg aggregation strategy to send local models back to the server.	The effectiveness of model aggregation is limited by data heterogeneity and imbalance.	-

Table 11. Summary of studied articles related to X-ray images (2 of 3). Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s), and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[25]	DarkCOVID, ChestCOVID	GAN, CNNs	COVID-19 X-ray	FL scheme (FedGAN) to generate realistic COVID-19 images, in order to facilitate privacy-enhanced COVID-19 detection with GANs in edge cloud computing.	COVID-19 multiclass classification with a blockchain-based FedGAN framework for security.	Enhances COVID-19 data privacy and detection performance with a blockchain-based FedGAN framework for secure COVID-19 data analytics. Compared to the state-of-the-art schemes, the proposal has a high detection accuracy rate and a low running time. Achieves significant accuracy by combining FL and GAN, composing the FedGAN.	GAN models have a high computational cost. Limitations on Edge nodes on its resources to train image datasets, which are needed for an FL-GAN process to have a better result.	Accuracy of 99.2% and 98.5% on the DarkCOVID and ChestCOVID datasets, respectively.
[26]	COVIDX8, composed of sub-datasets from open-source chest radiography datasets	CNN (ResNet18)	Chest X-ray	COVID-19 detection.	Dynamic fusion FL framework to detect COVID-19, called FedFocus.	Applied a blockchain-based solution to decentralize the aggregation process. Highly restores the imbalance of the real dataset by simulating the division of the training set based on the population and the infected cases of 3 real cities. Compared to FedAvg, FedFocus had significantly better stability.	Imbalance of samples in each training set of non-independent and identically distributed (Non-IDD) data, resulting in different training losses. The aggregation method and the optimal setting have the same effect on the optimal dynamic and resilience factors.	FedFocus has achieved similar accuracy to FedAvg, both higher than 92%.

Table 12. Summary of studied articles related to X-ray images (3 of 3). Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s), and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[27]	A public pediatric pneumonia dataset called MedNIST	CNN (ResNet18)	Chest X-ray	Classifies pediatric chest X-ray images into one of three categories: normal (no signs of infection), viral pneumonia or bacterial pneumonia, using large and multi-national datasets.	Proposed an FL-based solution, created to be an open-source software framework called Privacy-preserving Medical Image Analysis (PriMIA).	Compatible with a wide range of medical imaging data formats, user-configurable, which introduces functional improvements to FL training.	High computational cost on an imbalanced dataset. High data quality on nodes is needed to succeed in FL models.	The best accuracy was 91% on training and 92% on validation, which occurred in the centrally trained experiment.
[28]	A public chest X-ray dataset acquired from the Department of Health and Human Services, Montgomery County, Maryland, USA, and Shenzhen No. 3 People’s Hospital in China.	CNNs (VGG-16 and ResNet50)	Chest X-ray	Chest X-ray images classification to identify COVID-19 from non-COVID-19 cases.	Proposed 2 FL architectures based on non-independent and identically distributed (Non-IID) data, using unbalanced data.	The first study addressing the problem of federated learning on X-ray images for COVID-19 detection. CNNs applied pre-trained weights on ImageNet. The FL framework led to a comparable performance with a centralized learning process and remained robust.	Needed to apply data augmentation on the experiments once a small dataset was used.	Accuracy of 94.40% on FL-VGG16 with data augmentation, and 97.0% on FL-ResNet50 with data augmentation.

Table 13. Summary of studied article related to histology images. Ref means Reference, IDM means Input Dimension and Imaging Modalities, Arch means Model Architecture(s), and RG means Research Goals.

Ref	Datasets	Arch.	IDM	RG	Propose	Contributions	Limitations	Accuracy
[29]	Private datasets from multiple institutions	ResNet50	Histology whole-slide imaging (WSI) with only slide-level labels.	FL for digitized gigapixel whole-slide images for binary and muti-class classification on breast cancer and renal cell cancer.	Proposed an FL-based solution focusing on weakly supervised deep-learning models to demonstrate the feasibility and effectiveness of privacy-preserving using only slide-level labels for supervision on survival prediction.	Software package available for usage on the GitHub website, enabling multiple institutions to integrate their WSI datasets and train their models. The developed FL framework has the clear potential to be utilized in numerous crucial computational pathology assignments beyond those demonstrated in this study.	Limited to weakly supervised federated multiple-instance learning.	Best-balanced accuracy of 0.900 ± 0.020 on Renal Cell Carcinoma (RCC) on sub-typing test reported as a five-fold mean. The best-balanced accuracy of 0.756 ± 0.026 on Breast Invasive Carcinoma (BRCA) as to the sub-typing test was reported as a five-fold mean, both in a federated environment where $\alpha = 0.1$.

3.1. Magnetic Resonance Imaging

Among the ten selected articles related to MRI, half (5) proposed Convolutional Neural Networks (CNNs) containing several deep layers to extract features from the input images. Of those five articles, two used pre-trained models: one used ResNet34 and U-Net [7], and the other article used six pre-trained models [6] and applied a CNN ensemble combining the three best models based on the average accuracy (InceptionV3, VGG16, DenseNet) for the Voting Ensemble model. From the group of CNNs, only 1 (one) article used convolutional auto-encoders [13], and two articles developed end-to-end CNNs [13,30]. The remaining articles did not make use of CNNs and applied different architectures, including one article that used Single Task Model (STM), Shared-bottom Model (SBM), and Multi-gate Mixture of Experts (MMoE) [14], another Graph Convolutional Neural Networks (GCNs) [15], other Multi-layer perceptron (MLP) [31], another DAG's [11] and a last one MoCo [12].

In terms of used datasets, in [14,15,31], the experiments were conducted on the same dataset, called Autism Brain Imaging Data Exchange (ABIDE). With small datasets, common techniques were applied in the pre-processing step, mainly of data augmentation [10,14,30].

3.2. Computed Tomography

Among the five reviewed articles related to CT imaging, in all of them, private large datasets were used, and the research goal was related to COVID-19. They also used pre-trained models to train their CNN-based models, such as VGG-16 [16], an improved version of RetinaNet18 [17]. GhostNet, ResNet50, and ResNet101 were used as pre-trained models using 3D U-shape fully convolution networks (FCN) as the baseline model [18]. Those studies deal with a common topic, and the applied FL framework suffers from high computational cost and communication efficiency issues.

The authors of [20] proposed an architecture for a Dynamic-Fusion-Based Federated by conducting experiments in two types of images: CT and X-ray images. They used three public datasets and applied CNN-based models: pre-trained GhostNet, ResNet50, and ResNet101 models.

Knolle et al. [19] used CT images for pancreatic segmentation and MR images for brain tumor segmentation, both on private datasets, and applied a CNN in a pre-trained model called MoNet, a shallow U-Netlike architecture, giving the possibility to extract more robust features, which generalize better to out-of-sample data.

3.3. X-ray

Among the seven selected articles about X-ray images, all of them used CNNs, six used pre-trained models, and almost half of them applied more than one pre-trained model to compare results [21,22,26]. They used ResNet18 and ResNet34, DenseNet121 and ResNet50, VGG-16, and ResNet50, respectively, and share some common models.

Nguyen et al. [23] used a generative adversarial network (GAN) that generates realistic COVID-19 images to facilitate privacy-enhanced COVID-19 detection with GANs in edge cloud computing and applied them on an FL scheme called FedGAN.

The authors in [24,25] used chest X-ray images for COVID-19 and Pneumonia detection, respectively, and both applied ResNet-18 as a pre-trained model. In [24], it is proposed an architecture for a Dynamic Fusion FL Framework and two strategies were compared: FedAvg and FedFocus, where both achieved similar results. On the other hand, in [25], the authors proposed an open-source software framework called Privacy-preserving Medical Image Analysis (PriMIA), but the proposed solution had a high computational cost due to the imbalanced dataset used.

As shown in Tables 10–12, all the selected articles indicate high accuracy (if reported) achieved using the proposed FL approach. Feki et al. [26], for example, achieved a high accuracy after applying data augmentation to the training data.

3.4. Histology

The single article found in the performed literature search that is related to histology images used different private datasets to predict and solve two different diagnostic problems: classification of breast and renal cell cancers. It used thousands of histology whole-slide images with only slide-level labels and a pre-trained ResNet50 CNN encoder.

There is a lack of annotations in most real-world whole-slide histopathology datasets, and as explained in [27], the proposed FL solution can be linked to weakly supervised multiple-instance learning to solve binary and multiclass classification problems. In addition, it presented accurate results without direct data exchange and its related intricacies while likewise maintaining differential privacy through randomized noise generation.

4. Frameworks

FL strategies help AI solutions enrich data on the training dataset, generating more accurate results by allowing multiple collaborators to build a robust machine-learning model using a large dataset. This is possible because there is no direct data sharing, as federated learning prioritizes the privacy of patients' data. In the health field, for example, when new hospitals participate in this collaborative environment, they bring more data and more computational resources. Although individual hospitals that do not have a large dataset can benefit from the rich datasets without providing much data, it can be, at the same time, a concern to big hospitals. There can be another challenge of equitable allocation, where a hospital may produce considerably more images than another, but the diversity of its images may be low. Differences in images in terms of acquisition protocols and labeling methodologies, coming from different devices, may contribute poorly or even negatively to the central model as well.

Some open-source software solutions are available for secure bioscience collaboration based on FL, as the ones presented in the following sections.

4.1. OBiBa

According to the OBiBa website, OBiBa is an international project that offers open-source software for epidemiological studies [28] and is suitable for applications built on a federated database infrastructure. The software includes five main tools, which are:

- Collect with Onyx: web application to collect data in clinics or assessment centers and manage interviews;
- Store and Document with Opal: central data repository for epidemiological studies;
- Analyze with R and DataSHIELD: uses Rock server application (Rest API) to provide statistical analysis services without accessing individual-level data;
- Publish with Mica: a web application that can be used to create web data portals for epidemiological studies [28]. This tool provides accessibility to query datasets stored in Opal databases, search variable dictionaries, and create study catalogs;
- Manage Users with Agate: central authentication server, which can be used for email notification services. This might help notify researchers of new data available for studies.

OBiBa infrastructure enables researchers to access patients' medical information related to epidemiology but does not allow access to personal data, such as the address or phone number of the patients. This infrastructure helps researchers to create studies or collect more data by respecting data privacy, and the process of enriching data improves the quality of the study once there is data diversity from which models learn.

4.2. DataSHIELD

DataSHIELD is free and open-source software developed for biomedicine, social science, and public health, which uses the OBiBa infrastructure. This analytical tool has been used to develop projects in collaboration with health by impacting social effects, lifestyle, and healthcare [29].

A large dataset contains images suitable for analyzing “remotely”, but transferring them has a high computational cost and a risk of leakage once shared throughout the network. In addition, as already mentioned, privacy data are crucial in the health sector because institutions deal with ethical issues when dealing with patients’ data, where security and data confidentiality are important factors to consider.

DataSHIELD is helpful for co-analysis of individual-level data from multiple studies or sources, but not physically sharing them.

4.3. Virtual Pooling and Analysis of Research Data

Virtual Pooling and Analysis of Research Data (ViPAR) is a software platform that implements database federation techniques. Researchers can remotely analyze datasets located in different locations across the globe with a web-based environment. ViPAR is an open-source, simple, powerful framework for centralized data management and facilitates the sharing of research data, respecting ethical and privacy issues [32]. The tool facilitates the standard analyses without accessing individual-level data, which is pre-configured for R, SAS, and Stata programming languages and can be shared with other researchers.

Regarding comparison, Vipar pools site-specific raw data, while DataShield pools site-specific statistics. In this way, DataShield can be considered a better tool in terms of collaborations since no data leaves a particular site on any occasion, while ViPAR analysis data leave a data-contributing site momentarily [32].

4.4. Data Safe-Havens

According to Burton et al. [33], data safe-havens (DSH) make for a secure database for sensitive data related to biomedicine, health, and healthcare systems, accessible only by authorized users with an appropriate informatics system and governance, who are working on projects and investigations that enhance these fields.

In a secure and reliable environment for performing medical studies, DSH provides access to medical records, financial data, and other sensitive information. It follows compliance regulations, respects data protection policies, and helps reduce the risk of data breaches and cyberattacks, facilitating collaboration among researchers by accessing sensitive data in a secure environment.

However, maintaining DSH can be expensive since it often requires significant financial and technical resources. DSH can also restrict or limit access to specific data types, resulting in difficulties for researchers accessing it. Researchers may also face difficulties regarding inconsistent policies, where each environment may be under data protection policies that differ from others.

The decentralized nature of federated learning, facilitated by data safe-havens, champions privacy by design, allowing data to stay within its source and preventing the need for centralized data repositories. Nevertheless, the effectiveness of data safe-havens depends on the robust implementation of privacy-preserving techniques, and there remains a risk of privacy breaches if not executed carefully.

Although data safe-havens promote collaboration and collective model training, a critical challenge emerges in finding a balance between collaboration and upholding data privacy. The system must manage potential trade-offs between improving model quality through diverse datasets and the necessity to safeguard highly sensitive information. In the context of federated learning, it is essential to guarantee the sustained success of data safe-havens through continuous evaluation and adaptation to evolving security threats and regulatory changes.

The concept of DSH in AI is becoming increasingly important as AI systems become more complex and integrated into everyday life.

5. Discussion and Conclusions

There are several studies proposing different federated learning approaches for disease identification and classification, going from lower to higher complexity.

Most of the reviewed articles were published last year, which proves that researchers are becoming interested in this field, and FL has the potential to inspire and attract them to advance in the development of more competent AI-based medical imaging solutions. Moreover, most authors use CNNs to extract features from the input images and use pre-trained models, which enables neural network models to learn faster.

The authors have used different approaches, datasets, and architectures to build their FL frameworks, but overall, MR images have led to better accuracy than CT and X-ray images, and that might be the reason why MR images have gained more attention in the studied articles, once MR images are a better option than X-ray or CT images when specialists need to visualize soft tissues. Likewise, there are several publicly available datasets of MR and X-ray images, but just a few with CT images, and this research difficulty should be addressed.

Histology images are another important type of medical imaging modality that facilitates clinical decision-making, but there is a lack of FL solutions for this type of image, as demonstrated in the presented literature search. In this way, there are opportunities and demands for those images to be explored in an FL environment to solve different problems.

The direct quantitative comparisons between models applied on different datasets are unsuitable because each study has a different context and goal, e.g., detecting lung abnormalities in COVID-19 patients or binary classification to diagnose a brain tumor. However, the selected articles show that the FL-based frameworks have yielded interesting and promising results.

The use of multiple techniques in federated learning is due to the need to address the diverse and dynamic nature of decentralized data, as well as to optimize communication efficiency, enhance privacy, and fortify the system's robustness in the face of diverse challenges.

More than one dataset was considered in most of the reviewed articles because it contributes to increased model generalization by providing a more comprehensive understanding of data patterns. Additionally, incorporating diverse datasets enhances model robustness and resilience to variations, addresses biases, promotes fair models, and adapts to heterogeneous environments.

This article also summarized free open sources for bioscience collaboration, which aims to improve human health through science. OBiBa, DataSHIELD, and Vipar encourage collaborative learning using federation principles and facilitating research relationships. In addition, data safe havens can support researchers in storing, analyzing, and working on confidential individuals' data. These environments protect sensitive personal data, establish strict data protection policies, and are accessible only by authorized personnel.

From the current review, it can be realized that in the near future, researchers, clinicians, and hospitals might use more frequent applications that utilize FL to predict mortality, propose treatments to start earlier improvement of patient care and maintain the security of patient data.

Through its comprehensive review, this survey can be used as a reference for future explorations in the FL medical domain.

Author Contributions: Conceptualization and supervision by J.M.R.S.T.; investigation, data collection, formal analysis and writing—original draft preparation by F.R.d.S.; writing—review and editing by R.C. and J.M.R.S.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ng, D.; Lan, X.; Yao, M.M.-S.; Chan, W.P.; Feng, M. Federated learning: A collaborative effort to achieve better medical imaging models for individual sites with small labelled datasets. *Quant. Imaging Med. Surg.* **2021**, *11*, 852. [[CrossRef](#)]

2. Mouhni, N.; Elkalay, A.; Chakraoui, M.; Abdali, A.; Ammoumou, A.; Amalou, I. Federated learning for medical imaging: An updated state of the art. *Ing. Syst. D'Inf.* **2022**, *27*, 143–150. [\[CrossRef\]](#)
3. Gomathisankaran, M.; Yuan, X.; Kamongi, P. Ensure privacy and security in the process of medical image analysis. In Proceedings of the 2013 IEEE International Conference on Granular Computing (GrC), Beijing, China, 13–15 December 2013; pp. 120–125. [\[CrossRef\]](#)
4. Li, W.; Milletari, F.; Xu, D.; Rieke, N.; Hancox, J.; Zhu, W.; Baust, M.; Cheng, Y.; Ourselin, S.; Cardoso, M.J.; et al. Privacy-preserving federated brain tumour segmentation. In Proceedings of the Machine Learning in Medical Imaging: 10th International Workshop, MLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, 13 October 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 133–141. [\[CrossRef\]](#)
5. Gkoulalas-Divanis, A.; Loukides, G. Introduction to medical data privacy. In *Medical Data Privacy Handbook*; Springer: Switzerland, 2015; pp. 1–14. [\[CrossRef\]](#)
6. Islam, M.; Reza, M.T.; Kaosar, M.; Parvez, M.Z. Effectiveness of federated learning and CNN ensemble architectures for identifying brain tumors using MRI images. *Neural Process. Lett.* **2022**, *55*, 3779–3809. [\[CrossRef\]](#)
7. Zhang, M.; Qu, L.; Singh, P.; Kalpathy-Cramer, J.; Rubin, D.L. SplitAVG: A heterogeneity-aware federated deep learning method for medical imaging. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 4635–4644. [\[CrossRef\]](#)
8. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the Artificial Intelligence and Statistics, PMLR, Fort Lauderdale, FL, USA, 20–22 April 2017; pp. 1273–1282. [\[CrossRef\]](#)
9. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Int. J. Surg.* **2021**, *88*, 105906. [\[CrossRef\]](#)
10. Linardos, A.; Kushibar, K.; Walsh, S.; Gkontra, P.; Lekadir, K. Federated learning for multi-center imaging diagnostics: A simulation study in cardiovascular disease. *Sci. Rep.* **2022**, *12*, 3551. [\[CrossRef\]](#)
11. Liu, J.; Liang, X.; Yang, R.; Luo, Y.; Lu, H.; Li, L.; Zhang, S.; Yang, S. Federated learning-based vertebral body segmentation. *Eng. Artif. Intell.* **2022**, *116*, 105451. [\[CrossRef\]](#)
12. Wu, Y.; Zeng, D.; Wang, Z.; Shi, Y.; Hu, J. Distributed contrastive learning for medical image segmentation. *Med. Image Anal.* **2022**, *81*, 102564. [\[CrossRef\]](#)
13. Bercea, C.I.; Wiestler, B.; Rueckert, D.; Albarqouni, S. Federated disentangled representation learning for unsupervised brain anomaly detection. *Nat. Mach. Intell.* **2022**, *4*, 685–695. [\[CrossRef\]](#)
14. Huang, Z.-A.; Hu, Y.; Liu, R.; Xue, X.; Zhu, Z.; Song, L.; Tan, K.C. Fed-erated multi-task learning for joint diagnosis of multiple mental disorders on MRI scans. *IEEE Trans. Biomed. Eng.* **2022**, *70*, 1137–1149. [\[CrossRef\]](#)
15. Peng, L.; Wang, N.; Dvornek, N.; Zhu, X.; Li, X. Fedni: Federated graph learning with network inpainting for population-based disease prediction. *IEEE Trans. Med. Imaging* **2022**, *42*, 2032–2043. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Florescu, L.M.; Streba, C.T.; Șerbănescu, M.S.; Mămuleanu, M.; Florescu, D.N.; Teică, R.V.; Nica, R.E.; Gheonea, I.A. Federated learning approach with pre-trained deep learning models for COVID-19 detection from unsegmented CT images. *Life* **2022**, *12*, 958. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Dou, Q.; So, T.Y.; Jiang, M.; Liu, Q.; Vardhanabhuti, V.; Kaissis, G.; Li, Z.; Si, W.; Lee, H.H.; Yu, K.; et al. Federated deep learning for detect- ing COVID-19 lung abnormalities in CT: A privacy-preserving multinational validation study. *NPJ Digit. Med.* **2021**, *4*, 60. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Zhang, W.; Zhou, T.; Lu, Q.; Wang, X.; Zhu, C.; Sun, H.; Wang, Z.; Lo, S.K.; Wang, F.-Y. Dynamic-fusion-based federated learning for COVID-19 detection. *IEEE Internet Things J.* **2021**, *8*, 15884–15891. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Knolle, M.; Kaissis, G.; Jungmann, F.; Ziegelmayer, S.; Sasse, D.; Makowski, M.; Rueckert, D.; Braren, R. Efficient, high-performance semantic segmentation using multi-scale feature extraction. *PLoS ONE* **2021**, *16*, 0255397. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Yang, D.; Xu, Z.; Li, W.; Myronenko, A.; Roth, H.R.; Harmon, S.; Xu, S.; Turkbey, B.; Turkbey, E.; Wang, X.; et al. Federated semi-supervised learning for covid region segmentation in chest CT using multi-national data from China, Italy, Japan. *Med. Image Anal.* **2021**, *70*, 101992. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Kumar, A.; Purohit, V.; Bharti, V.; Singh, R.; Singh, S.K. MediSecFed: Private and secure medical image classification in the presence of malicious clients. *IEEE Trans. Ind. Inform.* **2021**, *18*, 5648–5657. [\[CrossRef\]](#)
22. Ziegler, J.; Pfitzner, B.; Schulz, H.; Saalbach, A.; Arnrich, B. Defend- ing against reconstruction attacks through differentially private federated learning for classification of heterogeneous Chest X-ray data. *Sensors* **2022**, *22*, 5195. [\[CrossRef\]](#)
23. Nguyen, D.C.; Ding, M.; Pathirana, P.N.; Seneviratne, A.; Zomaya, A.Y. Federated learning for COVID-19 detection with generative adversarial net- works in edge cloud computing. *IEEE Internet Things J.* **2021**, *9*, 10257–10271. [\[CrossRef\]](#)
24. Li, Z.; Xu, X.; Cao, X.; Liu, W.; Zhang, Y.; Chen, D.; Dai, H. Integrated CNN and Federated Learning for COVID-19 Detection on Chest X-ray images. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2022**, in press. [\[CrossRef\]](#)
25. Kaissis, G.; Ziller, A.; Passerat-Palmbach, J.; Ryffel, T.; Usynin, D.; Trask, A.; Lima, I., Jr.; Mancuso, J.; Jungmann, F.; Steinborn, M.-M.; et al. End-to-end privacy preserving deep learning on multi-institutional medical imaging. *Nat. Mach. Intell.* **2021**, *3*, 473–484. [\[CrossRef\]](#)
26. Feki, I.; Ammar, S.; Kessentini, Y.; Muhammad, K. Federated learning for COVID-19 screening from Chest X-ray images. *Appl. Soft Comput.* **2021**, *106*, 107330. [\[CrossRef\]](#) [\[PubMed\]](#)

27. Lu, M.Y.; Chen, R.J.; Kong, D.; Lipkova, J.; Singh, R.; Williamson, D.F.; Chen, T.Y.; Mahmood, F. Federated learning for computational pathology on gigapixel whole slide images. *Med. Image Anal.* **2022**, *76*, 102298. [CrossRef]
28. OBiBa: Open Source Software for Epidemiology. Available online: <http://www.obiba.org/> (accessed on 31 January 2023).
29. DataSHIELD: A Software Solution for Secure Bioscience Collaboration. Available online: <https://www.datashield.org/> (accessed on 31 January 2023).
30. Mahloul, D.H.; Abed, M.H. Distributed brain tumor diagnosis using a federated learning environment. *Bull. Electr. Eng. Inform.* **2022**, *11*, 3313–3321. [CrossRef]
31. Li, X.; Gu, Y.; Dvornek, N.; Staib, L.H.; Ventola, P.; Duncan, J.S. Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: ABIDE results. *Med. Image Anal.* **2020**, *65*, 101765. [CrossRef]
32. Carter, K.W.; Francis, R.W.; Carter, K.; Francis, R.; Bresnahan, M.; Gissler, M.; Grønberg, T.; Gross, R.; Gunnes, N.; Hammond, G.; et al. Viper: A software platform for the virtual pooling and analysis of research data. *Int. J. Epidemiol.* **2016**, *45*, 408–416. [CrossRef]
33. Burton, P.R.; Murtagh, M.J.; Boyd, A.; Williams, J.B.; Dove, E.S.; Wallace, S.E.; Tassé, A.M.; Little, J.; Chisholm, R.L.; Gaye, A.; et al. Data Safe Havens in health research and healthcare. *Bioinformatics* **2015**, *31*, 3241–3248. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.