Federated Learning for Multi-Institutional Medical Image Segmentation

Elvis Theyo
Indian Institute of Information Technology, Gwalior
Gwalior, India
elvistheyo@gmail.com

Dr. W Wilfred Godfrey
Indian Institute of Information Technology, Gwalior
Gwalior, India
godfrey@iiitm.ac.in

***Abstract*—Data-driven medical applications have emerged as a promising tool for designing reliable and scalable diagnostic and prognostic models from medical data due to the rise of Deep Learning (DL). Recently, these applications have attracted a great deal of attention from academia to industry which has undoubtedly enhanced the quality of healthcare delivery. However, data-driven medical applications still have poor adoption due to the large amounts of data required for training DL models. Furthermore, it is difficult to collect medical data from various sites to a centralized database due to growing data privacy concerns. Recent developments in Federated Learning (FL) have made it possible to train complex DL models in a distributed manner and has received unprecedented attention as a research domain, particularly for processing medical data at the edge of a network in a decentralized manner to preserve privacy and address security concerns. In this paper, we investigate the feasibility of applying Federated Learning for multiple medical image segmentation tasks to enable multi-institutional collaboration and preserve patient privacy.**

***Index Terms*—Deep Learning, Federated Learning, Medical Image Segmentation**

I. INTRODUCTION

Medical Image Segmentation is considered one of the most important areas in medical image analysis and is essential for monitoring, diagnosis and treatment. The aim of medical image segmentation is to identify a region of interest (RoI) such as tumor and lesion. Segmentation consists of two related tasks: object recognition and object delineation. Object recognition determines the target object's whereabouts on the image while object delineation draws the object's spatial extent and composition. The prevalent medical image segmentation tasks include brain and brain-tumor segmentation, liver and liver-tumor segmentation, cell segmentation, optic disc segmentation, lung segmentation, cardiac image segmentation, pulmonary nodule segmentation, etc. [1]

Deep Learning is being widely used for medical image segmentation and numerous papers have been presented documenting the success of DL in this field [2]. However, the performance of DL models strongly relies on the amount and diversity of data used for training. In the field of Medical Imaging, acquiring large-scale annotated medical imaging datasets for training is a significant challenge. Unlike photography images, labeling medical images require expert knowledge.

Usually, state-of-the-art algorithms are assessed on meticulously curated datasets which often originate from a small number of sources. This can result in biases where demographics or technical imbalances distort predictions and negatively affect the accuracy for certain groups or sites. Thus, in order to capture subtle relationships between disease patterns, socioeconomic and genetic factors, as well as complex and rare cases, it is crucial to expose a model to diverse cases.

Ideally, a collaboration between research institutions could address these challenges but sharing medical data to a centralized location faces various regulatory, legal, ethical, privacy, technical, and data-ownership challenges. This is a significant barrier to pursuing scientific collaboration across transnational medical research institutions.

Traditional Artificial Intelligence (AI) techniques require data collection to a centralized location and processing which is impractical in realistic healthcare scenarios due to the aforementioned challenges [2]. In recent years, Federated Learning [3] has emerged as a distributed collaborative AI paradigm that enables the collaborative training of DL models by coordinating with multiple clients (e.g., medical institutions) without the need of sharing raw data to a centralized location.

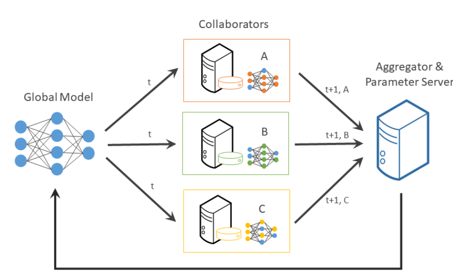

Fig. 1. System Architecture of Federated Learning.

In FL, each client trains its own model using local data, and only the model updates are sent to the central server. The server accumulates and aggregates the individual updates to yield a global model and then sends the new shared parameters to the clients for further training. In this way, the training data is kept private for each client and never leaves the client's site during the learning process. The central server only receives the model's updates, thus the patient data remains private and multi-institutional collaboration is enabled.

In this work, various variants of the U-Net model architectures [7] will be trained to evaluate the performance of each variant in the federated setting. Furthermore, three medical image segmentation tasks namely: Brain Tumor Segmentation, Liver Tumor Segmentation, Lung Segmentation will be explored using Federated Learning to evaluate the performance of federated learning models with standalone (centralized) training models on segmentation datasets from multiple institutions. To the best of my knowledge, this work is the first to utilise the various variants of the U-Net architecture in a federated setting for multiple multi-institutional medical image segmentation datasets.

II. Literature Review

Sheller *et al*. [6] (2018) proposed a FL approach for brain tumor segmentation to enable multi-institutional collaboration. The dataset from the Brain Tumor Segmentation (BraTS) Challenge 2018 [8]–[10] is used which contains data collected from 19 institutions, using various MRI scanners. The popular Convolutional Neural Network (CNN) architecture: U-Net [11] is used to train the data at each individual client.

Wenqi *et al*. [12] (2019) investigated the feasibility of applying differential-privacy techniques to protect the patient data in a FL setup for brain tumor segmentation using the BraTS dataset. The experimental results demonstrated that there is a trade off between model performance and privacy protection costs

Wang *et al*. [13] (2020) collaboratively generated and evaluated a FL model for Pancreas Segmentation. The data consisted of healthy and unhealthy pancreas collected at two institutions from different countries namely Taiwan and Japan. The model utilized was from coarse-to-fine network architecture search (C2FNAS) [14] with an additional variational auto-encoder (VAE) [15] branch to the encoder endpoint. It was demonstrated that FL dramatically improved the generalizability of models on the server-side and client-side for both datasets.

Yang *et al*. [16] (2020) proposed a novel system based on federated and semi-supervised learning for COVID Region Segmentation in Chest CT using multi-national Data from China, Italy, and Japan. FL gives sufficient flexibility to different institutions to collaboratively train DL models without data sharing while semi-supervised learning ensures effective training even when some sites have only limited amount of annotated data but large amount of unannotated data. A 3D U-shape Fully Convolution Network (FCN) [17] was used to segment the ground glass-like opaque (GGO) regions (COVID-19 affected regions) of the lung from the 3D chest CT.

Sarma *et al*. [18] (2021) demonstrated that FL is able to produce a model for whole prostate segmentation which exhibited superior performance with improved generalizability when applied to an external validation dataset. The data was collected from 3 US academic institutions namely: University of California, Los Angeles(UCLA), the State University of New York (SUNY) Upstate Medical University, and National Cancer Institute (NCI). Each institute collected 1 prostate MRI from each of a cohort of 100 patients enrolled in an IRB-approved protocol studying the use of MRI for prostate cancer diagnosis. The 3D Anisotropic Hybrid Network (3D AH-Net) [19] was used as the DL model for this study.

Ziller *et al*. [22] (2021) performed federated medical image segmentation under image-level differential privacy on the MSD Liver segmentation task. The experiments were carried out using U-Net-like architectures with MobileNet V2 [23] and ResNet-18 [24] as backbones to the encoder portion of the U-Net. It was observed that FL trained models are able to achieve performance on-par with locally trained models.

Liu *et al*. [20] (2021) presented a novel approach, Episodic Learning in Continuous Frequency Space (ELCFS) to solve federated domain generalization, which aims to learn a federated model from multiple distributed source domains such that it can directly generalize to unseen target domains. The method was extensively evaluated on two medical image segmentation tasks: optic disc and cup segmentation on retinal fundus images and the prostate segmentation on T2-weighted MRI.

Xu and Yan [21] (2022) proposed a novel federated multi-encoding U-Net (Fed-MENU) method for multi-organ segmentation. Each sub-network in the model is considered an expert of a specific organ and trained for that client. To evaluate the performance of their method, extensive experiments were conducted on four public abdominal CT datasets, each of which is annotated with a different set of abdominal organs. The results showed that without sharing the raw data, the proposed method can effectively utilize the isolated datasets with different partial labels to train a global model for multi-organ segmentation.

Xu *et al*. [25] (2022) conducted a theoretical analysis on the FL to reveal the problem of model aggregation during training on non-iid data and proposed a novel federated cross learning (FedCross) to resolve this problem. The method was evaluated using four public prostate MRI datasets: MSD, NCI-ISBI, PROMISE12, and PROSTATEx. The experimental results showed that FedCross outperforms mainstream FL methods on non-iid data.

Misonne and Jodogne [29] (2022) applied FL with U-Net to the task of Heart Segmentation in CT-scan images. The experimental results indicated that FL by a coalition of hospitals is a viable alternative to traditional centralized learning.

Following the observation from the review of prior related works above, although several works have dealt with the topic of Federated Learning for medical image segmentation, they either use a single variant of the U-Net architecture or another single DL model to train the local data in a federated setting. The aforementioned works also evaluated the performance of federated learning models on only a single multi-institutional medical image segmentation dataset and hence it limits the possibility of generalisation to other medical image segmentation tasks.

III. METHODOLOGY

The effectiveness of Federated Learning for medical image segmentation tasks will be demonstrated by training various variants of the U-Net architecture: Base U-Net, U-Net++, and Attention U-Net on three multi-institutional medical image datasets. The performance of each variant will be evaluated to identify the best model for each medical image segmentation task. The models trained in a federated setting aim to achieve better or comparable segmentation performance to centrally trained models while ensuring preservation of sensitive patient information and demonstrating multi-institutional collaboration. As demonstrated in Sheller *et al*. [6], along with Brain Tumor Segmentation, this work will extend towards two additional multi-institutional medical image segmentation tasks namely: Liver Tumor Segmentation and Lung Segmentation. Through this, the generalizability of models on server-side and client-side will be evaluated for various medical image segmentation tasks.

A. Federated Learning

In traditional machine learning or centralized learning, all collaborating institutions (collaborators) upload their data to a central server for training. On the contrary, In FL, the collaborators do not share their data but instead train a shared model locally and only send model updates to the central server. The server accumulates and aggregates the individual updates to yield a global model and then forwards the new shared parameters to each client for further training. Once the model updates have been applied, they are discarded by the central server as they are only required for enhancing the current global model.

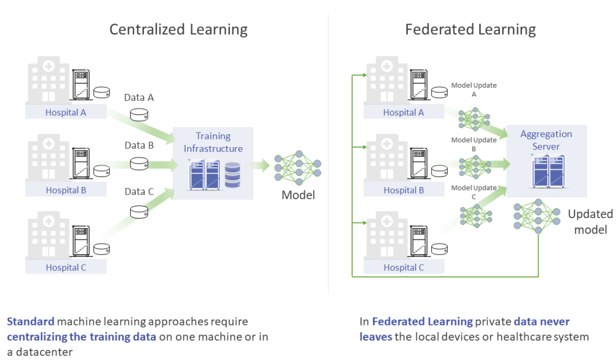

Fig. 2. Centralized Learning vs Federated Learning.

Federated Learning (FL) [3], [4] is a learning paradigm that seeks to address the problem of data governance and privacy by training algorithms collaboratively without exchanging the underlying datasets. This allows multiple collaborators in various sites to build a standard, robust ML model without storing data in a centralized location. Thus, critical issues such as data privacy, data security, data access rights, and access to heterogeneous data are addressed. Although FL was initially designed for mobile edge devices, it has attracted increasing attention in the healthcare domain because of its privacy-preserving nature of the patient information.

It is clear that some trust of the central server coordinating the training of the clients is still required. However, for applications where the training objective can be specified on the basis of data available at each client, FL can significantly reduce privacy and security risks by limiting the attack surface to only the device, rather than the device and the cloud [3].

Federated Learning brings in a few more components to the traditional centralized training pipeline:

- **Collaborator**
  It is a client in the federation that has access to the local train, validation, and test sets. By design, the collaborator is the only component of the federation with access to the local data. The local dataset should never leave the collaborator.
- **Parameter Server**
  The parameter/central server coordinates the training process in a FL system. parameter. It is responsible for sending copies of the global model to the collaborators. Parameter servers are often combined with aggregators on the same compute node.
- **Aggregator**
  The aggregator receives locally tuned models from collaborators and combines them into a new global model. Typically, federated averaging, (a weighted average) is the algorithm used to combine the locally tuned models.
- **Round**
  A federation round is defined as the interval (typically defined in terms of training steps) where an aggregation is performed. Collaborators may perform local training on the model for multiple epochs (or even partial epochs) within a single training round.

The training process of the Federated Learning system usually involves the following steps:

1) The collaborator receives the global model updates from the server and locally trains on their local data and sends the local model updates to the central server.
2) The central server receives the local model updates and performs secure aggregation without learning information about any collaborator to yield a global model.
3) The central server forwards the new shared parameters to the collaborators for further training.
4) Go back to 1 for another federated round.

B. U-Net Architectures

To train the data, the following U-net architectures were implemented:

- **Basic U-Net**
  Ronneberger *et al*. (2015) [11] proposed the U-Net architecture which was primarily designed for medical image segmentation. It consists of two paths: The first path is the contracting path (encoder or analysis path) which is similar to a regular convolution network and provides classification information. The second path is

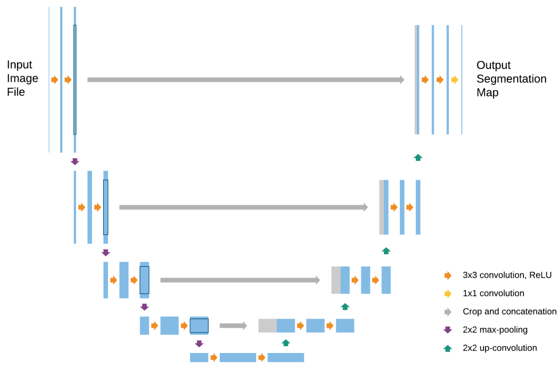

Fig. 3. Basic U-Net Architecture.

the expansion path (decoder or synthesis path) consisting of sequences of up-convolutions and concatenations with features from the corresponding contracting path. This expansion allows the network to learn localized classification information and also increases the resolution of the output which is then passed onto a final convolutional layer to create a fully segmented image. The resulting network is almost symmetrical which results in a U-like shape.

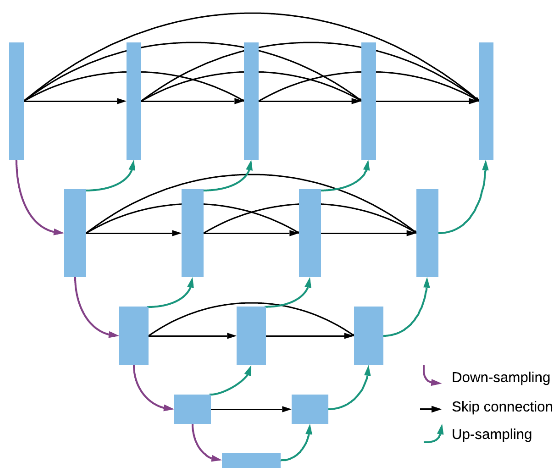

Fig. 4. U-Net++ Architecture.

- **U-Net++**
  Zhou *et al*. proposed U-Net++ (2018) [30] which was inspired from DenseNet. A dense network of skip connections is used as an intermediary grid between the contracting and the expansive path. This helps the network by propagating more semantic information between the two paths and thereby enabling it to segment images more accurately. The feature maps of the contracting path in a traditional U-net are directly concatenated onto the corresponding layers in the expansive path. However, U-net++ has a number of skip connection nodes between each corresponding layer.
- **Attention U-Net**
  Oktay *et al*. proposed Attention U-Net [31] which has the ability to focus on specific objects that are of importance while ignoring unnecessary areas. This is achieved by making use of the attention gate, a unit which trims features that are not relevant to the ongoing task.

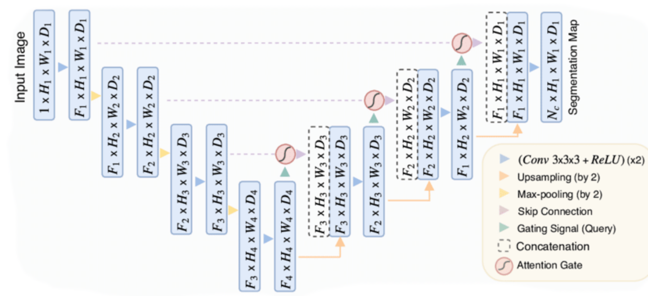

Fig. 5. Attention U-Net Architecture.

  Each layer in the expansive path has an attention gate through which the corresponding features from the contracting path must pass through before the features are concatenated with the upsampled features in the expansive path.

C. Datasets

The following multi-institutional medical image segmentation datasets were used:

- **BraTS Dataset**
  The BraTS 2020 [8]–[10] utilizes multi-institutional preoperative MRI scans. All scans are available as NIfTI files (.nii.gz) and describe a) native (T1) and b) post-contrast T1-weighted (T1Gd), c) T2-weighted (T2), and d) T2 Fluid Attenuated Inversion Recovery (T2-FLAIR) volumes. Annotations comprise the “GD-enhancing tumor”, the “peritumoral edema”, and the “necrotic and non-enhancing tumor” core. The MRI scans were acquired in accordance with different clinical protocols and various scanners from 19 institutions.
- **LiTS Dataset**
  This dataset was extracted from LiTS [33] – Liver Tumor Segmentation Challenge (LiTS17) organised in conjunction with ISBI 2017 and MICCAI 2017. The dataset is created in collaboration with seven hospitals and research institutions and manually blind reviewed by independent three radiologists.
- **Montgomery & Shenzen Chest X-Ray Dataset**
  This dataset contains over 500 x-rays scans which was labelled by radiologists. The data was acquired from the Department of Health and Human Services, Montgomery County, Maryland, USA and Shenzhen No. 3 People’s Hospital, China [34], [35]. Both the datasets contain normal and abnormal chest X-rays with manifestations of TB and include associated radiologist readings.

D. Mechanism

Prior to running the Federated Learning system, the appropriate medical image dataset required for the segmentation task is split into train, validation and test sets. Following this, n number of clients are created to represent the collaborating institutions.

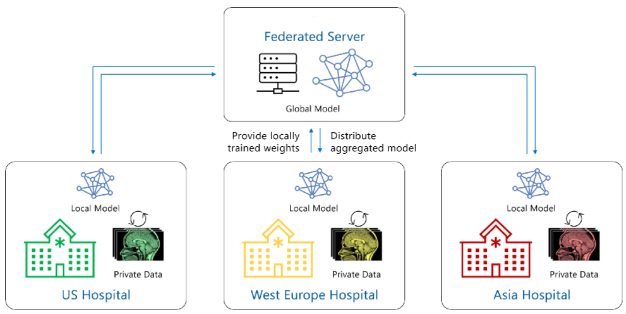

Fig. 6. System Architecture of Federated Learning for Multi-Institutional Medical Image Segmentation.

Each client's local data is represented by further partitioning the training set into n data shards to represent the number of clients. This ensures that each client will have its own local data to train on. In order to train the local data at each client, various variants of the U-Net architecture as mentioned above are used. The U-net architectures are implemented using the library: keras unet collection [32].

In the training process, initially, the global model is initialized and each local model receives the global model's current weights from the central server. The local models share the same model structure and hyperparameters as the global model. The local models trains on each of their respective local data which generates the updated parameters and are then sent to the central server.

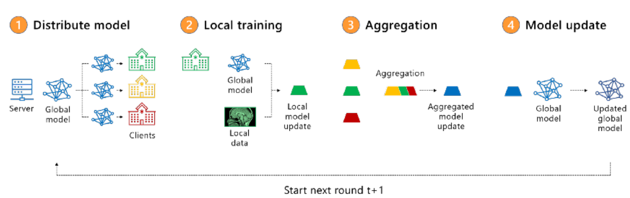

Fig. 7. Federated Learning Workflow.

The server collects the updated parameters from each client's local model and performs secure aggregation using the Federated Averaging (FedAvg) Algorithm [3]. The central server then forwards the new shared parameters to each client for further training. This marks the end of one federated round and this cycle persists until a certain number of federated rounds. In this manner, each client will receive the parameters learned from other clients and hence multi-institutional collaboration is achieved. Since the client has full authority over their data and local training data never leaves the client's site, sensitive patient information is preserved.

Algorithm 1 Federated Learning

```
initialize global model
Set n federated rounds
for n federated rounds do
    get global model's weights
    create m clients
    for m clients do
        initialize local model
        set global model's weights as local model model's weights
        train local model with client's data
        store local model's weights
        clear session
    end for
    aggregate the weights of the local models
    update global model with the aggregated weights
end for=0
```

E. Benchmarking Metric

The quantitative performance evaluation metric for all the medical image segmentation tasks is the Dice Coefficient (DC). The Dice Similarity Coefficient or simply the Dice Coefficient is a similarity measure ranging from 0 which indicates no overlap and 1 which indicates a perfect overlap.

IV. Experiments and Results

In order to simulate a Federated Learning setup, the building blocks of Federated Learning were implemented using Python, Keras and TensorFlow. All datasets used in the experiments are multi-institutional i.e it contains data from multiple institutions and thus provide heterogeneity. Each dataset is split into train, validation and test sets as (80, 10, 10) splits respectively. For all the experiments, 3 clients (C1, C2, C3) were created to represent the collaborating institutions. To represent the local data of each client, the training data is further split into 3 data shards and assigned to each client respectively. Before splitting, in order to make sure that each local dataset is representative of the overall distribution of the data, the training data is randomly shuffled.

A. Brain Tumor Segmentation

Since the training set contains 249 subjects, after splitting, each client has 249/3 = 82 subjects respectively.

TABLE I
Hyperparameters for Brain Tumor Segmentation

Batch Size	32
Learning Rate	1e-4
Optimizer	Adam

As seen from the quantitative results in TABLE II, for brain tumor segmentation in a federated setting, the U-Net Architecture, U-Net++ gives the best performance.

TABLE II
COMPARING U-NET VARIANTS PERFORMANCE AFTER 5 FEDERATED ROUNDS FOR BRAIN TUMOR SEGMENTATION.

Metric	U-Net	U-Net++	Attention U-Net
Accuracy	96.71%	97.15%	96.32%
Loss	0.015	0.017	0.038
Dice Coefficient	0.57	0.78	0.65

B. Liver Tumor Segmentation

Since the training set contains 90 subjects, after splitting, each client has 90/3 = 30 subjects respectively.

TABLE III
HYPERPARAMETERS FOR LIVER TUMOR SEGMENTATION

Batch Size	16
Learning Rate	1e-3
Optimizer	Adam

TABLE IV
COMPARING U-NET VARIANTS PERFORMANCE AFTER 5 FEDERATED ROUNDS FOR LIVER TUMOR SEGMENTATION.

Metric	U-Net	U-Net++	Attention U-Net
Accuracy	94.88%	95.75%	96.31%
Loss	0.06	0.031	0.047
Dice Coefficient	0.82	0.76	0.87

As seen from the quantitative results in TABLE IV, for liver tumor segmentation in a federated setting, the U-Net Architecture, Attention U-Net gives the best performance.

C. Lung Segmentation

Since the training dataset contains 450 subjects, after splitting, each client has 450/3 = subjects respectively.

TABLE V
HYPERPARAMETERS FOR LUNG SEGMENTATION

Batch Size	16
Learning Rate	1e-5
Optimizer	Adam

TABLE VI
COMPARING U-NET VARIANTS PERFORMANCE AFTER 5 FEDERATED ROUNDS FOR LUNG SEGMENTATION.

Metric	U-Net	U-Net++	Attention U-Net
Accuracy	95.37%	96.12%	95.32%
Loss	0.93	0.67	1.24
Dice Coefficient	0.74	0.81	0.69

As seen from the quantitative results in TABLE VI, for lung segmentation in a federated setting, the U-Net Architecture, U-Net++ gives the best performance.

V. CONCLUSION

The experiments demonstrate that clinical institutions around the globe can collaborate to train a model without sharing their data and thereby ensuring patient data privacy by using Federated Learning. Furthermore, the FL models learn from a wide variety of heterogeneous data spread across various sites which allows for better performance on unseen data. Despite the assurances of FL regarding the protection of patient information, it is still prone to risks and so additional measures may be required to prevent information leakage through model updates. Further work can be done to improve the security concerns regarding the misuse of model updates to reconstruct training data. One approach to this problem is Differential Privacy (DP) [37] which provides a quantifiable measure of data anonymization. Using DP, noise can be added to the model updates at the client-side before forwarding it to the central server.

REFERENCES

[1] R. Wang, Tao Lei, R. Cui, B. Zhang, H. Meng, A K. Nandi, "Medical Image Segmentation Using Deep Learning: A Survey", 2020.

[2] Dinh C. Nguyen, Quoc-Viet Pham, Pubudu N. Pathirana, Ming Ding, Aruna Seneviratne, Zihuai Lin, Octavia A. Dobre, Won-Joo Hwang, "Federated Learning for Smart Healthcare: A Survey", 2021.

[3] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, Blaise Agüera y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data", 2016.

[4] Q. Yang, Y. Liu, T. Chen, Y. Tong, "Federated Machine Learning: Concept and Applications", 2019.

[5] Zhihua Liu, Lei Tong, Zheheng Jiang, Long Chen, Feixiang Zhou, Qianni Zhang, Xiangrong Zhang, Yaochu Jin, Huiyu Zhou, "Deep Learning Based Brain Tumor Segmentation: A Survey", 2020.

[6] Micah J Sheller, G Anthony Reina, Brandon Edwards, Jason Martin, and Spyridon Bakas, "Multi-Institutional Deep Learning Modeling Without Sharing Patient Data: A Feasibility Study on Brain Tumor Segmentation", 2018.

[7] Vinay Ummadi, "U-Net and its variants for Medical Image Segmentation: A short review", 2022.

[8] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging 34(10), 1993-2024 (2015) DOI: 10.1109/TMI.2014.2377694.

[9] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, et al., "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features", Nature Scientific Data, 4:170117 (2017) DOI: 10.1038/sdata.2017.117.

[10] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, et al., "Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge", arXiv preprint arXiv:1811.02629, 2018.

[11] Olaf Ronneberger, Philipp Fischer, Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", 2015.

[12] Wenqi Li, Fausto Milletarì, Daguang Xu, Nicola Rieke, Jonny Hancox, Wentao Zhu, Maximilian Baust, Yan Cheng, Sébastien Ourselin, M. Jorge Cardoso, Andrew Feng, "Privacy-preserving Federated Brain Tumour Segmentation", 2019.

[13] Pochuan Wang, Chen Shen, Holger R. Roth, Dong Yang, Daguang Xu, Masahiro Oda, Kazunari Misawa, Po-Ting Chen, Kao-Lang Liu, Wei-Chih Liao, Weichung Wang, Kensaku Mori, "Automated Pancreas Segmentation Using Multi-institutional Collaborative Deep Learning", 2020.

[14] Yu, Q., Yang, D., Roth, H., Bai, Y., Zhang, Y., Yuille, A.L., Xu, D, “C2FNAS: Coarse-to-Fine neural architecture search for 3D medical image segmentation”, 2019.
[15] Myronenko, “A 3D MRI brain tumor segmentation using autoencoder regularization”, In Brain lesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. pp. 311–320. Springer International Publishing, 2019.
[16] Dong Yang, Ziyue Xu, Wenqi Li, Andriy Myronenko, Holger R. Roth, Stephanie Harmon, Sheng Xu, Baris Turkbey, Evrim Turkbey, Xiaosong Wang, Wentao Zhu, Gianpaolo Carrafiello, Francesca Patella, Maurizio Cariati, Hirofumi Obinata, Hitoshi Mori, Kaku Tamura, Peng An, Bradford J. Wood, Daguang Xu, “Federated Semi-Supervised Learning for COVID Region Segmentation in Chest CT using Multi-National Data from China, Italy, Japan”, 2020.
[17] Liu, S., Xu, D., Zhou, S.K., Pauly, O., Grbic, S., Mertelmeier, T., Wicklein, J., Jerebko, A., Cai, W., Comaniciu, “3d anisotropic hybrid network:Transferring convolutional features from 2d images to 3d anisotropic volumes”, 2018.
[18] Karthik V. Sarma. Stephanie Harmon, Thomas Sanford, Holger R. Roth, Ziyue Xu, Jesse Tetreault,6 Daguang Xu, Mona G. Flores, Alex G. Raman, Rushikesh Kulkarni, Bradford J. Wood, Peter L. Choyke, Alan M. Priester, Leonard S. Marks, Steven S. Raman, Dieter Enzmann, Baris Turkbey, William Speier and Corey W. Arnold, “Federated learning improves site performance in multicenter deep learning without data sharing”, 2021.
[19] Liu S, Xu D, Zhou SK, et al., “3D anisotropic hybrid network: transferring convolutional features from 2D Images to 3D anisotropic volumes”, 2017.
[20] Quande Liu, Cheng Chen, Jing Qin, Qi Dou, Pheng-Ann Heng, “FedDG: Federated Domain Generalization on Medical Image Segmentation via Episodic Learning in Continuous Frequency Space”, 2021.
[21] X. Xu and P. Yan, “Federated Multi-organ Segmentation with Partially Labeled Data”, 2022.
[22] Alexander Ziller, Dmitrii Usynin, Nicolas Remerscheid, Moritz Knolle, Marcus Makowski, Rickmer Braren, Daniel Rueckert, Georgios Kaissis, “Differentially private federated deep learning for multi-site medical image segmentation”, 2021.
[23] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks”, 2018.
[24] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, “Deep residual learning for image recognition”, 2016.
[25] Xuanang Xu, Tianyi Chen, Han Deng, Tianshu Kuang, Joshua C. Barber, Daeseung Kim, Jaime Gateno, Pingkun Yan, James J. Xia, “Federated Cross Learning for Medical Image Segmentation”, 2022.
[26] Nicholas Bloch, A.M.: NCI-ISBI 2013 challenge: Automated segmentation of prostate structures. https://doi.org/10.7937/K9/TCIA.2015.ZF0VLOPV, https://wiki.cancerimagingarchive.net/x/B4NEAQ, publisher: The Cancer Imaging Archive
[27] Geert Litjens, Robert Toth, Wendy van de Ven, Caroline Hoeks, Sjoerd Kerkstra, Bram van Ginneken 3, Graham Vincent, Gwenael Guillard, Neil Birbeck, Jindang Zhang, Robin Strand, Filip Malmberg 6, Yangming Ou 7, Christos Davatzikos, Matthias Kirschner, Florian Jung, Jing Yuan, Wu Qiu, Qinquan Gao 10, Philip Eddie Edwards 10, Bianca Maan 11, Ferdinand van der Heijden 11, Soumya Ghose , Jhimli Mitra , Jason Dowling , Dean Barratt , Henkjan Huisman, Anant Madabhushi, “Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge”, 2014.
[28] Samuel G Armato 3rd 1, Henkjan Huisman 2, Karen Drukker 1, Lubomir Hadjiiski 3, Justin S Kirby 4, Nicholas Petrick 5, George Redmond 6, Maryellen L Giger 1, Kenny Cha 3 5, Artem Mamonov 7, Jayashree Kalpathy-Cramer 7, Keyvan Farahani, “PROSTATEx Challenges for computerized classification of prostate lesions from multiparametric magnetic resonance images”, 2018.
[29] T. Misonne & S. Jodogne, “Federated Learning for Heart Segmentation”, 2022.
[30] Z. Zhou, M. Mahfuzur R. Siddiquee, N. Tajbakhsh, J. Liang, “UNet++: A Nested U-Net Architecture for Medical Image Segmentation”, 2018.
[31] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y Hammerla, B. Kainz, B. Glocker, D. Rueckert, “Attention U-Net: Learning Where to Look for the Pancreas:, 2018.
[32] Sha, Y., Keras-unet-collection, 2021.
[33] P. Bilic, P. Ferdinand Christ, E. Vorontsov, G. Chlebus, H. Chen, Q. Dou, C. Fu, X. Han, P. Heng, J. Hesser, S. Kadoury, T. Konopczynski, M. Le, C. Li, X. Li, J. Lipkovà, J. Lowengrub, H. Meine, J. Hendrik Moltz, C. Pal, M. Piraud, X. Qi, J. Qi, M. Rempfler, K. Roth, A. Schenk, A. Sekuboyina, E. Vorontsov, Ping Zhou, C. Hülsemeyer, M. Beetz, F. Ettlinger, F. Gruen, G. Kaissis, F. Lohöfer, R. Braren, J. Holch, F. Hofmann, W. Sommer, V. Heinemann, C. Jacobs, G. Efrain Humpire Mamani, B. van Ginneken, G. Chartrand, A. Tang, M. Drozdzal, A. Ben-Cohen, E. Klang, M. M. Amitai, E. Konen, H. Greenspan, J. Moreau, A. Hostettler, L. Soler, R. Vivanti, A. Szeskin, N. Lev-Cohain, J. Sosna, L. Joskowicz,B. H. Menze, “The Liver Tumor Segmentation Benchmark (LiTS)”, 2019.
[34] Jaeger S, Karargyris A, Candemir S, Folio L, Siegelman J, Callaghan F, Xue Z, Palaniappan K, Singh RK, Antani S, Thoma G, Wang YX, Lu PX, McDonald CJ, “Automatic tuberculosis screening using chest radiographs”, IEEE Trans Med Imaging. 2014 Feb;33(2):233-45. doi: 10.1109/TMI.2013.2284099. PMID: 24108713
[35] Candemir S, Jaeger S, Palaniappan K, Musco JP, Singh RK, Xue Z, Karargyris A, Antani S, Thoma G, McDonald CJ, “Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration”, IEEE Trans Med Imaging. 2014 Feb;33(2):577-90. doi: 10.1109/TMI.2013.2290491. PMID: 24239990
[36] Intel, Open Federated Learning (OpenFL) Documentation.
[37] Kang Wei, Jun Li, Ming Ding, Chuan Ma, Howard H. Yang, Farokhi Farhad, Shi Jin, Tony Q. S. Quek, H. Vincent Poor, “Federated Learning with Differential Privacy: Algorithms and Performance Analysis”, 2019.